

**ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ УЧРЕЖДЕНИЕ
«ФЕДЕРАЛЬНЫЙ НАУЧНО-КЛИНИЧЕСКИЙ ЦЕНТР ФИЗИКО-
ХИМИЧЕСКОЙ МЕДИЦИНЫ ФЕДЕРАЛЬНОГО МЕДИКО-
БИОЛОГИЧЕСКОГО АГЕНТСТВА»**

на правах рукописи

БЕСПЯТЫХ ЮЛИЯ АНДРЕЕВНА

**Геномная и протеомная характеристика
штаммов *Mycobacterium tuberculosis* кластера
Beijing B0/W148**

03.01.04 – биохимия

Диссертация на соискание ученой степени
кандидата биологических наук

Научный руководитель
д.б.н., профессор РАН Ильина Е.Н.

Москва 2016

Оглавление

ВВЕДЕНИЕ	4
1 ОБЗОР ЛИТЕРАТУРЫ	10
1.1 Общая характеристика <i>Mycobacterium tuberculosis</i>	10
1.2 Основные подходы к изучению <i>Mycobacterium tuberculosis</i>	14
1.2.1 Молекулярно-генетический анализ	15
1.2.2 Протеомный анализ	24
1.2.3 Транскриптомный анализ	34
1.2.4 Модельные эксперименты	36
1.3 Популяционная структура <i>Mycobacterium tuberculosis</i>	38
1.4 Генетическое семейство Beijing	41
1.5 Характеристика кластера Beijing B0/W148	47
1.6 Заключение	51
2. МАТЕРИАЛЫ И МЕТОДЫ	52
2.1 Бактериальные штаммы	52
2.2 Культивирование микобактерий	52
2.3 Генетический анализ <i>Mycobacterium tuberculosis</i>	52
2.3.1 Геномные последовательности	52
2.3.2 Амплификация фрагментов генома	53
2.3.3 Секвенирование фрагментов генома модифицированным методом Сенгера	54
2.3.4 Рестрикционный анализ	55
2.3.5 Полногеномное секвенирование	55
2.3.6 Анализ данных полногеномного секвенирования	56
2.4 Протеомный анализ микобактерий	57
2.4.1 Экстракция белка из клеток микобактерий	57
2.4.2 Денатурирующий электрофорез в полиакриламидном геле	58
2.4.3 Триптическое расщепление белков	58
2.4.4 Масс-спектрометрический анализ	60
2.4.5 Валидация данных методом мониторинга множественных реакций ...	61
2.4.6 Биоинформационный анализ данных	62

3. РЕЗУЛЬТАТЫ	65
3.1 Формирование экспериментальных групп	65
3.2 Полногеномное секвенирование	65
3.3 Геномные особенности кластера BeijingB0/W148	66
3.3.1 Кластер-специфические SNPs	68
3.3.2 Структурная организация генома BeijingB0/W148	69
3.4 Отработка протокола экстракции белка из клеток микобактерий.....	73
3.5 Протеомный анализ <i>Mycobacterium tuberculosis</i>	78
3.5.1. Валидация данных методом MRM анализа	96
3.6 Сравнительный анализ геномных и протеомных данных	99
4. ОБСУЖДЕНИЕ	101
4.1. Формирование экспериментальных групп штаммов и геномных последовательностей	101
4.2. Генетические особенности штаммов <i>Mycobacterium tuberculosis</i> кластера Beijing B0/W148	102
4.3. Разработка метода протеомного анализа микобактерий	106
4.4. Протеомные особенности изолятов <i>Mycobacterium tuberculosis</i> кластера Beijing B0/W148	110
4.5 Сравнительный анализ геномных и протеомных данных	113
ЗАКЛЮЧЕНИЕ	119
ВЫВОДЫ	122
Список используемых сокращений	123
Список литературы	125
ПРИЛОЖЕНИЯ	161

ВВЕДЕНИЕ

Актуальность темы исследования и степень ее разработанности

Системная биология прокариот стремится к пониманию того, как физико-химические свойства молекул и характер их взаимодействий приводят к формированию фенотипических свойств бактерий. В настоящее время последовательность прокариотического генома может быть расшифрована в течение нескольких часов. Тем не менее, несмотря на то, что геном прямо или косвенно предопределяет структуру всех биомолекул клетки, таких как РНК, белки и липиды, предсказать их изменения только по геномной последовательности невозможно. Для исследования структур, функций и механизмов регулирования таких молекулярных систем необходимы точные и воспроизводимые методы количественных измерений всех компонентов в различных состояниях. Сегодня такие измерения стали обычными на уровне мРНК (Talaat *et al.*, 2004; Talaat *et al.*, 2007; Tailleux *et al.*, 2008; Azhikina *et al.*, 2010), однако из-за технических ограничений, такие измерения на уровне белка пока отстают в чувствительности, репрезентативности.

В настоящее время наиболее актуальными являются системные исследования клинически значимых микроорганизмов, таких как микобактерии. Самые известные виды микобактерий являются патогенами человека, это *Mycobacterium tuberculosis*, *Mycobacterium leprae* и *Mycobacterium ulcerans*. В России туберкулез до сих пор остается одной из наиболее значимых проблем здравоохранения и социального развития страны. В настоящее время можно говорить о наметившейся тенденции к снижению заболеваемости, но, тем не менее, она продолжает оставаться на достаточно высоком уровне. Только за 2014 год, по данным Всемирной организации здравоохранения (ВОЗ), в стране было выявлено около 100 тысяч вновь заболевших (ВНО, 2015).

До недавнего времени основные усилия по исследованию возбудителя туберкулеза были сосредоточены в области изучения особенностей геномной организации патогена. Сегодня доступны данные геномного секвенирования более чем для 5000 штаммов *M. tuberculosis* с различными фенотипами и

генотипами. Однако, опираясь на опыт применения технологии полногеномного секвенирования с последующим сравнительным анализом, стали очевидны ограничения такого подхода для полноценного описания причин развития лекарственной устойчивости и проявления патогенности (Warner and Mizrahi, 2013). Так, большинство мутаций обнаруживается в белках с гипотетической функцией, роль которых в физиологии микобактерий не ясна. В этой связи актуальным становится функциональный анализ информации, реализуемой геномом патогена, что возможно с привлечением методов протеомного тестирования, в том числе количественной протеомики.

Стоит отметить, что популяция возбудителя туберкулеза в России имеет свои региональные особенности. Показано, что в структуре популяции преобладают штаммы генетического семейства Beijing (до 60 % случаев), для которых доказаны строгая ассоциация с формированием лекарственной устойчивости (Bifani *et al.*, 2002; Toungousova *et al.*, 2002; Lipin *et al.*, 2007) и большая вирулентность (Dormans *et al.*, 2004) по сравнению с другими генетическими семействами. В свою очередь среди представителей семейства Beijing можно выделить наиболее распространенные кластеры. Одним из таких кластеров является Beijing B0/W148, доля которого среди российских штаммов неуклонно возрастает. В англоязычной литературе он носит название «успешный» клон *M. tuberculosis*. Представители этого кластера выявляются в 25 % случаев от общего количества изолятов семейства Beijing и обладают рядом характерных особенностей по сравнению с другими генотипами: повышенной вирулентностью, более строгой ассоциацией с лекарственной устойчивостью, увеличенной трансмиссивностью и фитнес-успешностью (Mokrousov, 2013). Представленные обстоятельства делают эндемичные для России штаммы *M. tuberculosis* кластера Beijing B0/W148 наиболее актуальным объектом для исследования их уникальных особенностей.

Цель исследования

Описание характерных протеомных и геномных особенностей штаммов *Mycobacterium tuberculosis* кластера Beijing B0/W148 для выяснения потенциальных факторов патогенности.

Задачи исследования

1. Формирование экспериментальной группы клинических изолятов *M. tuberculosis* кластера Beijing B0/W148
2. Описание геномной организации и мутационного профиля изолятов *M. tuberculosis* кластера Beijing B0/W148
3. Разработка протокола инвентаризации белков микобактерий с привлечением средств количественного протеомного анализа
4. Изучение характерных протеомных особенностей кластера Beijing B0/W148 по сравнению с лабораторным штаммом H37Rv
5. Анализ совокупных протеомных и геномных данных для выяснения потенциальных факторов патогенности кластера Beijing B0/W148

Научная новизна и практическая значимость работы

Впервые описаны характерные геномные и протеомные особенности клональной группы штаммов возбудителя туберкулеза - *M. tuberculosis* кластера Beijing B0/W148, успешно распространяющегося на территории Российской Федерации.

Впервые разработан оригинальный протокол инвентаризации белков микобактерий средствами LC-MS/MS анализа.

Впервые для штаммов *M. tuberculosis* кластера Beijing B0/W148 проведено соответствие между данными геномного и протеомного анализа.

Впервые на основании протеомных данных предсказаны биохимические особенности метаболизма представителей *M. tuberculosis* кластера Beijing B0/W148, которые можно рассматривать как потенциальные факторы патогенности, объясняющие эпидемиологическую успешность этих штаммов.

Данная работа вносит существенный вклад в фундаментальные исследования особенностей молекулярной организации и функционирования микобактерий. Полученные результаты могут быть использованы в проектах по разработке новых противотуберкулезных вакцин и лекарственных препаратов.

Внедрение результатов работы в практику

На основе результатов работы подготовлены следующие документы:

1. Патент на изобретение №2551764 (Ильина Е. Н., Шитиков Е. А., Беспярых Ю. А., Способ обнаружения микобактерий туберкулеза генетического кластера Beijing B0/W148). Зарегистрировано в Государственном реестре изобретений РФ 23.04.2015. Приоритет: 04.07.2013.

2. Патент на изобретение №2548797 (Ильина Е. Н., Шитиков Е. А., Беспярых Ю. А., Способ одновременной амплификации и флуоресцентного маркирования нескольких сегментов генома микобактерий туберкулезного комплекса). Зарегистрировано в Государственном реестре изобретений РФ 24.03.2015. Приоритет: 04.07.2013.

Методология и методы исследования

При выполнении диссертационной работы использованы различные биохимические подходы к обработке бактериальной культуры для экстракции природных биополимеров – белков и ДНК. Для последующего анализа биомолекул использованы современные высокопроизводительные методы: полногеномное секвенирование, высокопроизводительная жидкостная хроматография в сочетании с масс-спектрометрическим анализом. Обработка данных проводилась биоинформатическими методами.

Основные положения диссертации, выносимые на защиту

1. Показано, что *M. tuberculosis* кластер Beijing B0/W148 представляет собой обособленную генетическую группу микобактерий, геном которых содержит 59 уникальных кластер-специфических полиморфизмов и два инвертированных участка.

2. Разработанный протокол экстракции и последующей масс-спектрометрической идентификации белков микобактерий позволяет достоверно оценить совокупность синтезируемых бактериальных белков и провести сопоставление между отдельными группами штаммов.

3. Показано, что профиль белков, синтезируемых микобактериями кластера Beijing B0/W148, существенно отличается от лабораторного штамма H37Rv, часть этих отличий объясняется отличиями в структуре геномов.

4. Гиперпродукцию ферментов, ответственных за биосинтез длинноцепочечных жирных кислот, можно рассматривать как один из основных факторов вирулентности микобактерий кластера Beijing B0/W148, который наряду с гиперпродукцией белка HsaA способствует повышенной выживаемости этих штаммов в макрофагах.

5. Низкий уровень продукции белка SseA в штаммах микобактерий кластера Beijing B0/W148 возможно приводит к накоплению активных форм кислорода в клетке и, как следствие, повреждению ДНК. Это в свою очередь может вести к образованию широкого спектра генетических вариантов, способствующих выживанию бактериальной клетки под действием селективного отбора, в частности во время лекарственной терапии.

6. Характерные для штаммов микобактерий кластера Beijing B0/W148 изменения в представленности белков DosR регулона свидетельствуют в пользу реализации ими отличного от H37Rv ответа на гипоксию, что также может способствовать лучшей выживаемости внутри макрофагов.

Степень достоверности и апробация работы

Достоверность результатов проведенных исследований подтверждается использованием современных измерительных технологий и методик, реализуемых с использованием сертифицированного оборудования. Материалы диссертации представлены на IV международной научно-практической конференции «Постгеномные методы анализа в биологии, лабораторной и клинической медицине» (Казань, 2014); Российско-китайской научно-

практической конференции по медицинской микробиологии и клинической микологии (XVIII Кашкинские чтения) (Санкт-Петербург, 2015); 6th European Conference on Prokaryotic and Fungal Genomics (Германия, 2015); итоговой научно-практической конференции ФНКЦ ФХМ ФМБА России (Москва, 2015); 13th International Conference on Molecular Epidemiology and Evolutionary Genetics of Infectious Diseases (Бельгия, 2016); Всероссийской научно-практической конференции по медицинской микробиологии и клинической микологии (XIX Кашкинские чтения) (Санкт-Петербург, 2016).

Публикации

Основное содержание работы отражено в 13 публикациях, в том числе в пяти статьях, опубликованных в рецензируемых журналах, рекомендованных ВАК, и в восьми публикациях трудов научных конференций.

1 ОБЗОР ЛИТЕРАТУРЫ

1.1 Общая характеристика *Mycobacterium tuberculosis*

Mycobacterium tuberculosis – возбудитель туберкулеза (ТБ), заболевания известного еще с древности и являющегося одной из главных причин смертности во всем мире. По данным Всемирной организации здравоохранения за 2014 год выявлено 9.6 миллионов новых и 1.5 миллиона смертельных случаев ТБ (WHO, 2015). При этом Россия является 1 из 22 стран с наибольшим бременем туберкулеза, на долю которых приходится 80 % всех случаев ТБ.

Согласно современной таксономии *M. tuberculosis* относится к царству Bacteria, типу Actinobacteria, классу Actinobacteridae, подклассу Actinomycetales, отряду Firmicutes, подотряду Corynebacterineae, семейству Mycobacteriaceae, роду *Mycobacterium* (Криг, 2007). Большинство микобактерий относятся к сапрофитам, но есть небольшая группа патогенных бактерий. Род включает в себя свыше 90 видов бактерий, которые разделяют на две основные группы согласно их скорости роста: медленно и быстро растущие виды (Godreuil *et al.*, 2007). Быстро растущие бактерии образуют видимые колонии на питательной среде при оптимальных условиях роста через семь дней. Данная группа включает в себя такие виды как *Mycobacterium smegmatis*, *Mycobacterium fortuitum* и *Mycobacterium abscessus*, представители которых обладают ограниченным патогенным эффектом и вызывают нетипичные клинические признаки у людей и животных. В противоположность этому, медленно растущая группа, включающая микобактерии туберкулезного комплекса (*M. tuberculosis* complex), *Mycobacterium tuberculosis*, *Mycobacterium avium*, *Mycobacterium leprae* и *Mycobacterium kansasii*, характеризуется большей патогенностью для человека и животных и вызывает хронические заболевания (Godreuil *et al.*, 2007).

Бактерии *M. tuberculosis* имеют вид прямой или слегка изогнутой палочки с шириной от 0.2 до 0.6 мкм и длиной варьирующей от 1 до 10 мкм (Godreuil *et al.*, 2007). Это неподвижные, не спорообразующие, грамположительные аэробные или факультативно анаэробные бактерии. Микобактерии туберкулеза демонстрируют тропность к тканям легких, вероятно, из-за наличия большого

количества кислорода. Они так же, обычно, являются факультативными внутриклеточными патогенами, инфицирующими макрофаги, где они размножаются и затем разносятся по различным частям тела/организма (Van Soolingen, 2001).

Бактерии имеют уникальную структуру клеточной стенки, которая играет решающую роль для их выживания (Godreuil *et al.*, 2007). Клеточная стенка содержит большое количество жирных микотиоловых кислот, ковалентно связанных с пептидогликаном и полисахаридом арабиногалактаном, обеспечивая тем самым сильный липидный барьер. Отличительной особенностью *M. tuberculosis* является кислото-спирто-щелочеустойчивость, что так же обусловлено высоким содержанием липидов и восков в клеточной стенке бактерии.

Выраженная внешняя капсула отсутствует, но имеется полисахаридная микрокапсула, отделенная от клеточной стенки осмиефобной зоной, обеспечивающая бактериям устойчивость к неблагоприятным воздействиям внешней среды.

Микобактериям необходим кислород для нормального развития, так как они являются аэрофилами (диапазон температур 37 — 42°C), при этом 5 – 10 % CO₂ способствует более быстрому росту. Щелочной оптимум среды 7.0 – 7.2. Имеются данные, что бактерии могут трансформироваться в микроаэрофилы или становиться анаэробами за счет изменения метаболизма в неблагоприятных условиях. Размножаются микобактерии простым поперечным делением, процесс происходит очень медленно – одно деление, в среднем, за 14 – 18 ч.

Микроорганизм достаточно требователен к составу питательных сред: для роста необходим белковый субстрат и глицерин, а так же набор микроэлементов и факторов роста (биотин, рибофлавин и другие).

В связи с этим, плотная яичная среда Левенштейна — Йенсена рекомендована ВОЗ в качестве стандартной среды для первичного выращивания микобактерий туберкулеза и определения их устойчивости к химиопрепаратам. В

настоящее время в качестве неселективных широко используют жидкие и плотные среды Middlebrook (Воробьев и Быков, 2003).

Первичную идентификацию *M. tuberculosis* всегда проводят на плотных питательных средах, на которых бактерии растут в виде R-колоний (rough – грубый, шероховатый). В большинстве случаев колонии морщинистые и сухие (редко могут встречаться влажные), имеют желтоватый или кремовый оттенок. При этом видимый рост обнаруживается только на 14 – 20 сутки. Среда Левенштейна-Йенсена достаточно трудная для приготовления, в связи с этим в лабораторной практике чаще всего используют полностью готовые коммерческие среды с добавками. В настоящее время наиболее распространенной является агар Middlebrook 7H11.

На жидкой питательной среде микобактерии туберкулеза образуют морщинистую пленку на поверхности (среда остается прозрачной), заметную на 7 – 10 сутки (Поздеев, 2001). Одним из стандартизованных методов культивирования *M. tuberculosis* является посев клеток в жидкую обогащенную среду Школьниковой (Драбкина, 1963). Видимый рост появляется на 14 день, однако эффективность роста посева невелика и зависит от особенностей выращиваемой культуры.

Полный геном *M. tuberculosis* H37Rv был опубликован в 1998 (Cole *et al.*, 1998). Кольцевая последовательность генома содержит около 4.4 миллионов пар оснований, соответствующих примерно 4000 генов (Рисунок 1). Геном *M. tuberculosis* является GC богатым (65.6 %), что характерно для аэробных прокариот в целом (Naya *et al.*, 2002). Первая аннотация содержала 3924 открытые рамки считывания (ОРС), но уже через несколько лет авторы скорректировали число ОРС до 3995, а так же предсказали функции 2058 генов (52 %) (Samus *et al.*, 2002). В настоящее время аннотация *M. tuberculosis* H37Rv (27 версия согласно базе данных TubercuList (<http://tuberculist.epfl.ch/>)) содержит 4018 белок-кодирующих генов, из которых 26 % относятся к классу белков с гипотетической функцией.

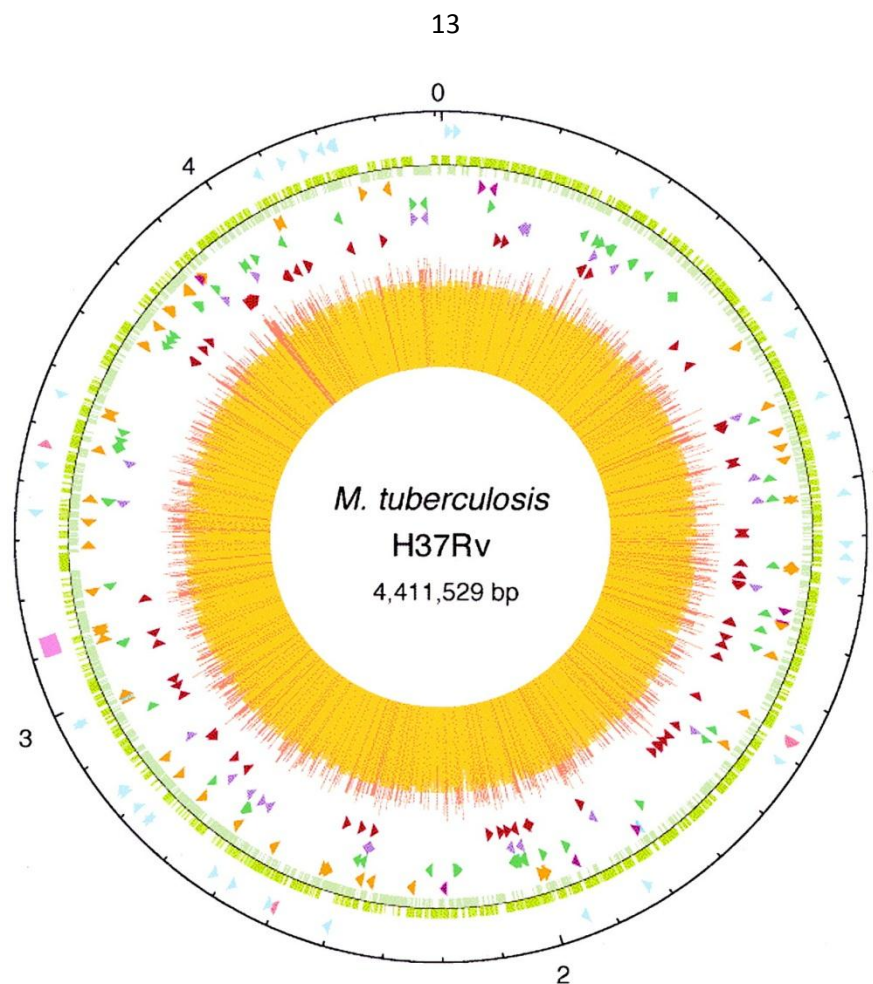


Рисунок 1. Круговая карта генома *M. tuberculosis* штамма H37Rv.

По направлению изнутри наружу: 1) гистограмма G+C состава (красным > 65 %, желтым < 65 %); 2) первое (темно-красное), второе (лиловое) и третье (зеленое) кольца отражают положение генов семейств PPE, PE и PE-PGRS, соответственно; 3) четвертое кольцо показывает повторяющиеся ДНК (IS-элементы (оранжевым цветом), семейство белков 13E12REP (темно-розовым), профаги (синим)); 4) пятое кольцо показывает кодирующие участки на плюс-цепи (темно-зеленые) и на минус-цепи (светло-зеленые); 5) шестое кольцо отражает положение генов РНК (тРНК-голубые, остальные - розовые) и регион DR (розовый прямоугольник). Внешнее кольцо отражает шкалу в т.п.о. (Cole *et al.*, 1998).

Микобактерии обладают всеми генами, необходимыми для синтеза незаменимых аминокислот, витаминов, ферментов и кофакторов (Godreuil *et al.*, 2007). Было отмечено, что имеется высокая доля генов, кодирующих ферменты, участвующие в липогенезе и липолизе. Кроме того, *M. tuberculosis* обладает генами, необходимыми для синтеза гликолитических ферментов и ферментов анаболического пентозофосфатного пути, который генерирует NADPH и пентозы, ферменты цикла Кребса и глиоксилатного цикла, который синтезирует углеводы из жиров. Туберкулезная палочка также обладает ферментами, используемыми в

аэробном, микроаэрофильном и бескислородном переносе электронов. Показано, что микобактерии способны выживать в различных условиях среды, в том числе богатых кислородом легких, макрофагах и в центре казеозных гранул (Cole *et al.*, 1998).

Геном микобактерий богат генами метаболизма жирных кислот, в том числе миколовых кислот, содержащих кислые аспарагин- и глицин-богатые полипептиды. Также большую часть генома составляют гены белков семейств PE (proline-glutamate – n = 99) и PPE (proline- proline-glutamate – n = 68) (Cole *et al.*, 1998). Вариабельность которых, предположительно, обеспечивает различия антигенов и способность ингибировать иммунный ответ (Poulet and Cole, 1995; Poulet and Cole, 1995; Karboul *et al.*, 2006).

Одной из особенностей генома *M. tuberculosis* является большое число повторяющихся последовательностей ДНК. Например, инсерционные элементы (от англ. insertion sequences (IS)), способствующие ДНК-полиморфизму микобактерий и вариабельное количество tandemных повторов (от англ. Variable Number of Tandem Repeats (VNTR)). Наряду с IS-элементами содержатся прямые повторы (от англ. direct repeat (DR) region), разделенные вариабельными последовательностями – спейсерами, а так же основные полиморфные tandemные повторы (MPTR) и полиморфная GC-богатая повторяющаяся последовательность (PGRS) (Poulet and Cole, 1995). Дополнительно, в геноме H37Rv были обнаружены последовательности профагов *phiRv1* и *phiRv2*. В связи с тем, что они не были обнаружены в геноме авирулентных H37Ra и *M. bovis* BCG, их ассоциируют с факторами патогенности.

1.2 Основные подходы к изучению *Mycobacterium tuberculosis*

Специфичность строения мощной гидрофобной стенки, переключение метаболизма из латентного состояния в активное и высокая степень адаптации к иммунной системе макроорганизма делает *M. tuberculosis* невероятно успешным патогеном. Дополнительно ситуацию осложняет глобальное распространение штаммов с множественной лекарственной устойчивостью (МЛУ) и широкой

лекарственной устойчивостью (ШЛУ), против которых практически неэффективна стандартная схема лечения. В связи с этим в последние годы все более актуальными становятся исследования, направленные на изучение взаимодействия микро- и макроорганизма, описание механизмов резистентности и поиск новых мишеней для лекарственных средств и вакцин.

Наиболее интересным в этой связи является системный подход к изучению описанных выше проблем. Триггером к этому послужило значительное совершенствование технологий в последние десятилетия, что привело к возникновению новых разделов молекулярной биологии, таких как геномика, транскриптомика и протеомика, позволяющих изучать микроорганизмы на различных биологических уровнях.

1.2.1 Молекулярно-генетический анализ

Генетические анализы, базирующиеся на методах типирования штаммов *Mycobacterium tuberculosis*, являются важным инструментом в дифференциации штаммов. Данные методы позволили заменить биохимические, использовавшиеся ранее (Wellman, 2012), которые, как правило, трудоемки и недостаточны для выявления филогенетических взаимоотношений. Первым стандартным методом генетического типирования штаммов был предложен анализ полиморфизма длины рестрикционных фрагментов IS6110 (от англ. Restriction Fragment Length Polymorphisms IS6110; IS6110 RFLP). Впоследствии были разработаны другие дифференцирующие методы, такие как сполиготипирование, анализ числа tandemных повторов в различных локусах генома и платформы для полногеномного секвенирования (Kamerbeek *et al.*, 1997).

Принцип RFLP типирования заключается в определении расположения и количества копий IS элементов в геноме бактерии (Eisenach *et al.*, 1986; Eisenach *et al.*, 1988; Ross *et al.*, 1991; Palittapongarnpim *et al.*, 1993). В настоящее время для анализа *M. tuberculosis* наиболее широко используется IS6110 (van Embden *et al.*, 1993). У различных штаммов *M. tuberculosis* количество копий данного элемента может варьировать от 1 до 26.

Стандартизованный вариант данного метода впервые был предложен van Embden с соавт. еще в 1993 г. и заключался в гидролизе ДНК эндонуклеазой рестрикции *PvuII*, с целью получения фрагментов, длина и количество которых зависят от количества и локализации *IS6110* (van Embden *et al.*, 1993). Затем полученные фрагменты разделяют электрофоретически, переносят на мембрану и гибридизуют с мечеными зондами, комплементарными 3'-концу последовательности *IS6110*. Паттерны полос гибридизации визуализируют на рентгенограмме. Каждая полоса соответствует единичной копии *IS6110*, окруженной фланкирующей ДНК различной длины, в зависимости от расстояния между сайтами *PvuII* (van Embden *et al.*, 1993). Метод *IS6110*-RFLP имеет высокую дискриминирующую способность, легко воспроизводится в лабораторных условиях и, следовательно, уже давно считается «золотым стандартом» среди методов типирования, применяемых для *M. tuberculosis*. Данный метод широко и успешно использовался для решения различных прикладных задач молекулярной эпидемиологии (выявление случаев кросс-контаминации, установление путей передачи инфекции), а так же изучения фундаментальных вопросов касающихся эволюции и филогении (Poynтен *et al.*, 2002; Hawkey *et al.*, 2003; Reddy *et al.*, 2009).

Тем не менее, метод имеет ряд ограничений. Во-первых, он довольно трудоемкий и занимает много времени. Во-вторых, для рестрикции необходимо большое количество (приблизительно от 2 до 10 мкг) высококачественной ДНК. Кроме того, метод является технически сложным и требует как сложное и дорогостоящее компьютерное программное обеспечение, так и высококвалифицированный персонал. Серьезным недостатком является отсутствие дискриминации в отношении штаммов *M. tuberculosis* содержащих менее 6 копий *IS6110* (Kremer *et al.*, 1999; Cowan *et al.*, 2002; Thong-On *et al.*, 2010; Varma-Basil *et al.*, 2011), а также *M. bovis* (Barlow *et al.*, 2001; Viana-Niero *et al.*, 2006; Otal *et al.*, 2008). Еще одно ограничение заключается в том, что данные, полученные в независимых исследованиях трудно сравнивать между различными лабораториями из-за отсутствия воспроизводимости и переносимости результатов

(Mazars *et al.*, 2001), несмотря на предпринимаемые попытки стандартизировать процедуру (van Embden *et al.*, 1993). В связи со всем выше сказанным стало очевидно, что необходимы альтернативные методы типирования *M. tuberculosis*.

Широкое распространение получили методы молекулярно-генетического типирования, основанные на амплификационных технологиях. К основным преимуществам данных методов относится возможность использования малого количества ДНК (несколько нанограмм) не высокого качества очистки, что позволяет использовать их для анализа клинического материала (Kanduma *et al.*, 2003). Данные методы так же менее трудоемки и менее затратные.

Одним из таких методов является анализ числа tandemных повторов (минисателлитный анализ, VNTR-типирование или Mycobacterial Interspersed Repetitive Unit, MIRU-типирование). VNTR, tandemно-повторяющиеся регионы, разбросанные по всему геному *M. tuberculosis* и напоминают полиморфные минисателлитные последовательности в геномах эукариот.

Суть MIRU-VNTR типирования заключается в выявлении различий между штаммами путем оценки количества и длины tandemных повторов в каждом локусе для каждого изолята. Изменчивость конкретного локуса часто зависит от географического источника образца и естественной генетической изменчивости штамма. Протокол MIRU-VNTR типирования включает ПЦР амплификацию для каждого локуса с использованием специфических праймеров, комплементарных фланкирующим участкам, и анализ ПЦР продуктов, которые разделяют гелем электрофорезом. Размер продуктов амплификации соответствует количеству tandemно-повторяющихся элементов, который определяется по отношению к известному размеру повторяющегося элемента в пределах каждого определяемого локуса. В хорошо оборудованных лабораториях, MIRU-VNTR типирование может быть выполнено с помощью капиллярного электрофореза и флуоресцентно меченных праймеров для ПЦР, с дополнительным преимуществом мультиплексирования (Frothingham and Meeker-O'Connell, 1998; Supply *et al.*, 2000; Roring *et al.*, 2002).

Типирование с использованием первой комбинации из 12 MIRU-VNTR локусов показало немного меньшую дискриминирующую способность по сравнению с IS6110-RFLP анализом для изолятов *M. tuberculosis* с большим количеством копий IS6110 (Hawkey *et al.*, 2003; Blackwood *et al.*, 2004). В связи с этим было предложено использовать 12-локусный MIRU-VNTR анализ в сочетании с другими методами генотипирования (например, спוליготипирование) чтобы приблизить дискриминирующую способность к IS6110 профилированию (Blackwood *et al.*, 2004; Alonso-Rodriguez *et al.*, 2008). В тоже время 12-локусный MIRU-VNTR анализ может иметь большую дискриминирующую способность, чем IS6110-RFLP анализ, если рассматривать изоляты *M. tuberculosis*, содержащие небольшое количество копий IS6110 (Cowan *et al.*, 2002; Lee *et al.*, 2002; Thong-On *et al.*, 2010; Asgharzadeh *et al.*, 2011; Varma-Basil *et al.*, 2011). Тем не менее, 12-локусный MIRU-VNTR анализ не может быть использован как единственный метод типирования, так как он переоценивает количество истинных эпидемиологических связей, особенно в крупных популяционных исследованиях (Cowan *et al.*, 2005; Scott *et al.*, 2005). В 2006 году в качестве международного стандарта была предложена оптимизированная схема 24-локусного (включая используемые ранее 12 локусов) MIRU-VNTR анализа (Supply *et al.*, 2006; Thong-On *et al.*, 2010; Brossier *et al.*, 2014). Филогенетически 24-локусный формат оказался более информативным, так же он может быть использован для отслеживания путей передачи возбудителей ТБ (Wirth *et al.*, 2008; Vadwai *et al.*, 2012). Дискриминирующая способность 24-локусной системы типирования сравнима с IS6110-RFLP типированием (Oelemann *et al.*, 2007; Alonso-Rodriguez *et al.*, 2008; Vadwai *et al.*, 2012).

Таким образом, 24-локусное MIRU-VNTR типирование было предложено в качестве нового «золотого стандарта» для молекулярного типирования штаммов *M. tuberculosis* complex. В то время как увеличение числа анализируемых локусов увеличивает дискриминирующую силу этого метода, оно так же повышает стоимость анализа и усложняет интерпретацию данных. Интересно, только 15 из 24 локусов (включая 6 ранее изученных локусов) составляли 96 % всех

обнаруживаемых полиморфизмов в штаммах *M. tuberculosis* (de Beer *et al.*, 2014). Стоит отметить, что ограниченность метода была показана для штаммов семейства Beijing, демонстрирующих гомоплазию генома (Hanekom *et al.*, 2008; Mokrousov, 2008). Для дифференциации штаммов генотипа Beijing были предложены дополнительные VNTR локусы 3232, 3820 и 4120, не включенные в стандартную 24-локусную систему (Niemann *et al.*, 2010).

Другой метод, сполиготипирование, так же основан на использовании амплификационных технологий. Данный метод заключается в определении структуры локуса хромосомы микобактерий (DR-регион), состоящего из нескольких десятков повторяющихся последовательностей (35 п.о.) и уникальных по своему нуклеотидному составу промежутков (спейсеров) между ними (длиной от 36 до 41 п.о). Локусы амплифицируются в ходе ПЦР с использованием инвертированных праймеров, комплементарных последовательностям DR. Затем ПЦР-продукты гибридизуют с мембраной, содержащей 43 иммобилизованных, ковалентно связанных синтетических олигонуклеотидных зонда, каждый из которых представляет собой уникальную последовательность спейсера. После гибридизации мембрану инкубируют со стрептавидин-пероксидазой или стрептавидин-щелочной фосфатазой, и гибридизационные сигналы обнаруживаются путем хемилюминесценции. Штаммоспецифические паттерны (сполиготипы) визуализируют на рентгеновской пленке, а затем дифференцируют согласно наличию или отсутствию индивидуальных спейсеров в сравнении с 43-спейсерным набором (Kamerbeek *et al.*, 1997). Поскольку результаты сполиготипирования могут быть представлены в двоичном формате (присутствует/отсутствует), они легко интерпретируются, оцифровываются и могут сравниваться между различными лабораториями (Dale *et al.*, 2001). В настоящее время существуют две крупные online базы данных, содержащие сполигопаттерны, это SpolDB4 (<http://www.pasteur-guadeloupe.fr:8081/SITVITDemo/>) и SITVIT (http://www.pasteur-guadeloupe.fr:8081/SITVIT_ONLINE/).

С момента внедрения сполиготипирования в лабораторную практику в конце 1990-х годов, были предложены несколько его модификаций. В частности, получил применение метод, основанный на проведении реакции удлинения зонда с последующей MALDI-ToF (матрично-активированная лазерная десорбция/ионизация времяпролетная от англ. Matrix Assisted Laser Desorption Ionization Time-of-Flight) масс-спектрометрической (МС) детекцией (Zhang et al., 2010; Shitikov et al., 2012). Метод основан на реакции ферментативного достраивания олигонуклеотидного зонда с участием дидезоксирибонуклеозидтрифосфатов, так называемых терминаторов, которые, в свою очередь, ответственны за специфическую остановку ферментативной реакции. На завершающем этапе проводится MALDI-ToF МС детекция олигонуклеотидов, отличающихся на массу одного основания от исходного зонда. Данная технология обеспечивает более высокую воспроизводимость, пропускную способность, большую простоту в использовании, и более глубокий анализ данных, чем классический метод, но основным ограничением является необходимость дорогостоящего оборудования (Honisch et al., 2010).

Еще одной альтернативой классического метода является сполиготипирование с использованием биочипов. Стоит отметить, что подобные биочипы имеются и отечественного производства, это СПОЛИГО-БИОЧИП (Gryadunov et al., 2005; Vespyatykh et al., 2014). Методика проведения анализа с использованием набора СПОЛИГО-БИОЧИП включает одну стадию ПЦР с последующей гибридизацией полученных ПЦР-продуктов на биочипе. Анализ результатов гибридизации проводят на универсальном аппаратно-программном комплексе для анализа биочипов. В данном случае цена необходимого оборудования значительно ниже, нежели для классической постановки или с использованием MALDI-ToF.

Расшифровка полной последовательности генома *M. tuberculosis* H37Rv (Cole et al., 1998; Camus et al., 2002) и затем ряда других референсных штаммов *M. tuberculosis* способствовала развитию новых методов типирования микобактерий. Данные методы различаются по дискриминационной силе, однако предоставляют

возможность для построения строгой филогении микобактерий и могут помочь в высокоточных эпидемиологических исследованиях (Roetzer *et al.*, 2013).

В ходе сравнительного анализа полных геномных последовательностей были выявлены различия между штаммами на уровне одиночных нуклеотидов (однонуклеотидные полиморфизмы, от англ. Single-Nucleotide Polymorphisms; SNP) и более протяженных фрагментов (large-sequence polymorphisms; LSP) (Fleischmann *et al.*, 2002). Таким образом, было предложено использовать SNP и LSP анализ для семейство-специфического типирования и изучения мутаций, ведущих к формированию лекарственной устойчивости.

Позиция SNP может быть детектирована с помощью REA (Ahmad *et al.*, 2004) или различных вариантов ПЦР (Herrera-Leon *et al.*, 2005). Современные технологии позволяют идентифицировать как одиночные замены, так и выявлять несколько сайтов SNP одновременно (Piatek *et al.*, 2000; Chakravorty *et al.*, 2012). Так, например, SNaPshot анализ (Thermo Fisher Scientific) позволяет обнаруживать до 10 SNPs в одном эксперименте с использованием капиллярного электрофореза. Данный анализ состоит из мультиплексной ПЦР с использованием праймеров различной длины, в результате чего длина продуктов специфична для каждого SNP. Удлинение подобных праймеров осуществляется одним специфически меченым дидезоксинуклеотидом (Wang *et al.*, 2010). Схожий принцип реализуется и в iPLEX Gold технологии. Здесь, продукты амплифицируются с помощью масс-модифицированных терминаторов, детектируемых с помощью MALDI-TOF (Bouakaze *et al.*, 2011). Таким образом, технология микрочипов может быть использована для одновременного обнаружения ряда мутаций в различных генах (Linger *et al.*, 2014).

Выявление SNPs в генах, ассоциированных с лекарственной устойчивостью, положило основу для разработки коммерчески доступных тестов для определения лекарственной устойчивости *M. tuberculosis* complex, включая Inno-LiPA Rif.TB (Innogenetics, Бельгия), GenoType MTBDR and Geno-Type MTBDRplus (Hain Lifescience GmbH, Германия), TBbiochip (ООО Biochip-IMB, Россия) и Xpert

MTB/RIF (Cepheid). GenoType MTBDR и Xpert MTB/RIF были одобрены ВОЗ для определения устойчивости к лекарственным препаратам (Gazi *et al.*, 2015).

Секвенирование обеспечивает высочайший уровень точности для анализа последовательности генов, поскольку позволяет анализировать всю последовательность данного гена. Оно используется для видовой идентификации и обнаружения SNP, в том числе новых полиморфизмов, возможно, участвующих в формировании устойчивости к лекарственным препаратам (Yuksel and Tansel, 2009).

Развитие секвенирования второго поколения (second-generation sequencing (SGS)) и платформ нового поколения вывело исследования по эпидемиологии микобактерий на новый высокоинформативный уровень. Так первое полногеномное секвенирование микобактерий было проведено методом секвенирования по Сенгеру (Cole *et al.*, 1998) и послужило основой для дальнейших исследований в этой области. Сегодня же используются пять основных SGS платформ: 454 pyrosequencing (Roche), переставшее использоваться в 2014; HiSeq/MiSeq (Illumina); SOLiD (Thermo Fisher Scientific), объявлено о прекращении в 2017; и Ion Torrent/Ion Proton (Thermo Fisher Scientific), за которыми следуют технологии секвенирования третьего поколения, PacBio (Pacific Biosystems) (Liu *et al.*, 2012).

Все SGS реализуют схожие протоколы. Первоначально геном фрагментируют и амплифицируют на твердой подложке с целью увеличения эмиссии сигнала. SGS секвенаторы могут прочитывать короткие последовательности от 150 п.о. до 800 п.о., в зависимости от платформы. Эти последовательности дополнительно либо выравниваются на известный шаблон, либо собираются *de novo* на основании перекрывающихся областей.

Стоит отметить, что данные, полученные с помощью имеющихся в настоящее время технологий секвенирования с высокой пропускной способностью, могут содержать ошибки секвенирования, характерные для каждой отдельной платформы. Например, платформы Ion Torrent и 454 имеют трудности из-за наличия в геноме *M. tuberculosis* большого количества повторяющихся

регионов (гомополимеры и микросателлиты). Важно учитывать эти особенности в ходе выполнения эпидемиологических исследований, где незначительные различия последовательностей могут привести к ложной интерпретации результатов (Niemann *et al.*, 2009). Кроме того, ошибки, генерируемые в процессе секвенирования, могут повлиять на эффективность *de novo* сборки. Последнее также затруднено, так как SGS риды короткие. В то время как технологии секвенирования четвертого поколения, MinION (Oxford Nanopore Technologies), позволяют секвенировать одну молекулу из нескольких тысяч оснований без амплификации (Feng *et al.*, 2015), дальнейшее сокращение частоты ошибок секвенирования необходимы прежде, чем эти технологии будут доступны для эпидемиологических исследований.

WGS предоставляет информацию обо всем геноме и позволяет определить практически все маркеры, детектируемые описанными выше методами. Поэтому он является более точным и достоверным при обнаружении изменчивости среди штаммов и предоставляет большее количество информации на каждом возможном уровне, от глобального (популяционного) до локального (сообщества) и от индивидуального хозяина (один пациент) к патогену (штамм).

На глобальном уровне WGS может быть использовано для изучения генетических взаимосвязей между разными видами, которые ранее уже были описаны для *M. ulcerans* и *M. marinum* (Stinear *et al.*, 2007) и членами *M. abscessus* комплекса (Cho *et al.*, 2013). Оно так же может быть полезно в исследовании важных вопросов касающихся биологии микобактерий, включая эволюционные механизмы формирования их популяции (Choo *et al.*, 2014), устойчивость к антибиотикам (Coll *et al.*, 2015), вирулентность и иммуногенность (Fleischmann *et al.*, 2002; Comas *et al.*, 2010).

В настоящее время основной акцент сделан на сравнительную геномику *M. tuberculosis*, где можно выделить следующие направления: реконструкция эволюционных событий, определение популяционной структуры, поиск новых механизмов резистентности и новых мишеней действия лекарственных препаратов. Так, например, в 2006 году Gagneux с соавт., в ходе анализа

делетированных участков генома, определил основные филогенетические линии (Gagneux *et al.*, 2006). А в 2013 году было показано, что представители различных линий мутируют с разной частотой, что отчасти может объяснить успешность распространения некоторых из них (Ford *et al.*, 2013). В том же 2013 году были опубликованы результаты двух крупных исследований по поиску генетических детерминант устойчивости методами сравнительной геномики большого количества образцов. Так, в исследовании Farhat с соавт. был проведен анализ 123 геномов *M. tuberculosis*, представляющих основные филогенетические линии (Farhat *et al.*, 2013). Применяв новый разработанный филогенетический тест, авторы обнаружили как известные детерминанты устойчивости, так и неописанные ранее мутации, локализованные в 30 генах и межгенных участках. Что не маловажно, исследователи экспериментально показали влияние одного из генов, *ponA1*, на устойчивость к противотуберкулезным препаратам. В другом исследовании, Zhang с соавт. проанализировали выборку из Китая, включающую 161 штамм *M. tuberculosis* (Zhang *et al.*, 2013). Как и в первом исследовании, авторам удалось идентифицировать классические маркеры устойчивости, при этом ассоциация с устойчивостью к препаратам первого и второго ряда была дополнительно выявлена для 70 генов и 19 межгенных участков. В тоже время авторами было показано, что противотуберкулезные препараты оказывают слабый позитивный отбор на геном *M. tuberculosis*. В 2014 году Casali с соавт. опубликовали работу, в которой описаны результаты секвенирования 1000 геномов из Самарской области (Casali *et al.*, 2014). Авторами была определена популяционная структура патогена в регионе, выявлены основные маркеры резистентности, а также описаны компенсаторные мутации, способствующие распространению *M. tuberculosis*.

1.2.2 Протеомный анализ

Как видно из предыдущей главы в последние годы достаточно много усилий по исследованию возбудителя туберкулеза были сосредоточены в области

изучения особенностей геномной организации патогена. Благодаря этому сегодня доступны данные геномного секвенирования более чем для 5000 штаммов *M. tuberculosis* с различными фенотипами и генотипами. Однако, опираясь на опыт применения технологии полногеномного секвенирования с последующим сравнительным анализом, стали очевидны ограничения такого подхода для полноценного описания причин развития лекарственной устойчивости и проявления патогенности (Warner and Mizrahi, 2013). Так, большинство замен обнаруживается в промоторных областях генов и белках с гипотетической функцией, роль которых в физиологии микобактерий не ясна. В этой связи актуальным становится функциональный анализ информации, реализуемой геномом патогена, что возможно с привлечение различных методов протеомного тестирования. В данном разделе обзора мы более детально рассмотрим развитие протеомики и ее направлений в исследованиях *M. tuberculosis*, а так же современные методы масс-спектрометрии, которые позволяют быстро осуществлять чувствительное и точное измерение представленности белка.

1.2.2.1 Развитие методов протеомного анализа

Марк Уилкинс в 1986 впервые ввел термин «протеом», объединив два слова: «белок» (англ. PROTein) и «геном» (англ. genOME) (Wilkins *et al.*, 1996). Протеом представляет совокупность всех белков в клетке, в том числе их изменения, происходящие со временем или под действием каких либо факторов. В 1997 был придуман термин «протеомика», по аналогии с геномикой, определяющей изучение генов и их функций (James, 1997).

Протеомика исследует всю совокупность белков, синтезируемых организмом/клеткой в данной среде и на конкретном этапе клеточного цикла. Она описывает их качественный состав, относительную представленность, взаимодействие с другими макромолекулами, а так же посттрансляционные модификации (ПТМ) (Hakkinen *et al.*, 2000; Molloy and Witzmann, 2002; Monteoliva and Albar, 2004). Белки играют важную роль почти во всех биологических процессах, соответственно в клетках существуют тысячи белков,

каждый из которых подвергается взаимодействию, как с другими белками, так и с целыми клеточными компартментами.

Традиционно, исследования белковых молекул базировались на специфическом взаимодействии с антителами, которые позволяют обнаруживать белки с помощью иммуноблотинга и иммуноферментного анализа, или непосредственно в интактных клетках и тканях иммуноцитохимически и иммуногистохимически, соответственно (Rosenkrands *et al.*, 2000). Сегодня для получения тотального протеома наиболее мощными являются технологии, основанные на применении методов масс-спектрометрического анализа (Molloy and Witzmann, 2002; Monteoliva and Albar, 2004). За последние годы значительно увеличилось покрытие протеома за счет развития МС технологий. Протеомные стратегии можно разделить на два типа: (I) Top-down протеомика, в основе которой лежит изучение интактных белков из биологических образцов, с использованием классических методов разделения, такие как двумерный гель электрофорез (2D-GE) и (II) Bottom-up протеомика основанная на том, что белки сначала протеолитически расщепляют на пептиды, которые затем разделяют с помощью жидкостной хроматографии, ионизируют и вводят в масс-спектрометр (Рисунок 2). Ранние исследования, использовавшие 2D-GE, идентифицировали и квантифицировали менее 100 белков в образце микобактерий (Rosenkrands *et al.*, 2000). Развитие методов жидкостной хроматографии сопряженной с тандемной масс-спектрометрией (от англ. liquid chromatography-coupled tandem mass spectrometry, LC-MS/MS), позволило идентифицировать и квантифицировать от 100 до 1000 белков за один МС запуск (Mehaffy *et al.*, 2010). Всё разнообразие bottom-up методик на уровне масс-спектрометрического эксперимента можно разделить на (1) исследовательские или панорамные (shotgun-proteomics, HRM), созданные для идентификации и количественного анализа как можно большего числа белков, и (2) подтверждающие (таргетные, целевые), созданные для отслеживания в множестве образцов сравнительно небольшого, определённого до начала эксперимента, набора белков/пептидов с наивысшей доступной чувствительностью, точностью, воспроизводимостью и пропускной способностью

метода (мониторинг множественных/выбранных реакций, от англ. Multiple/Selected Reaction Monitoring, MRM/SRM). Основные bottom-up подходы приведены на Рисунке 3.

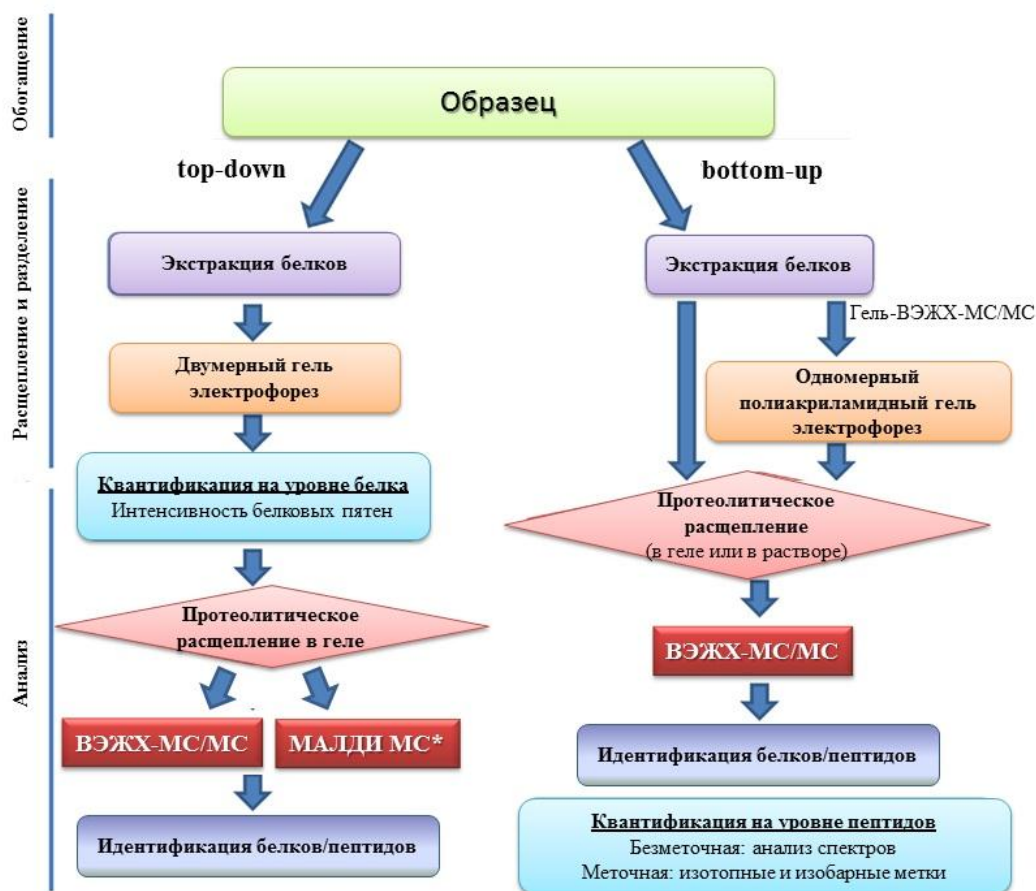


Рисунок 2. Основные протеомные подходы top-down и bottom-up протеомики. Адаптировано из (Yadeta *et al.*, 2013) с модификациями.

* МАЛДИ МС - матрично-активированная лазерная десорбция/ионизация, совмещенная с масс-спектрометрией

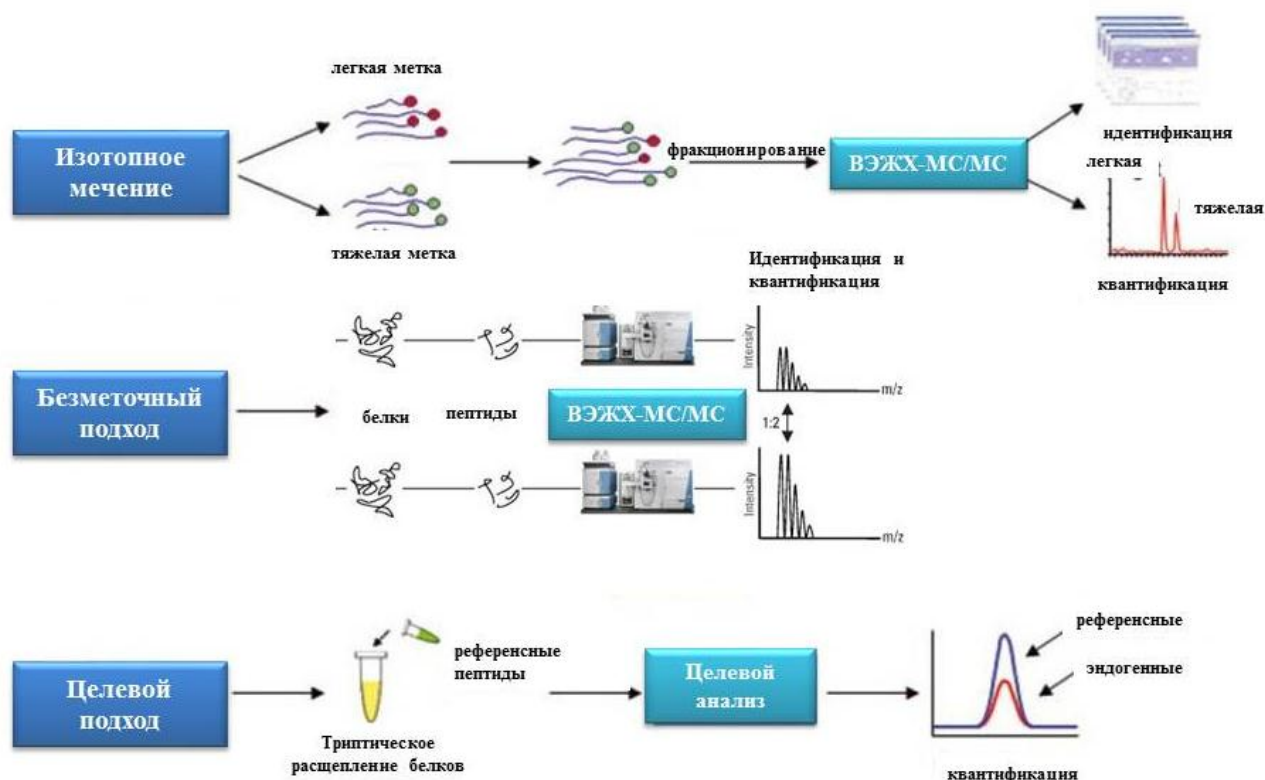


Рисунок 3. Методы протеомного анализа, реализующие bottom-up подходы

Одни из методов, реализующих shotgun подход, основаны на применении изотопных меток с последующим анализом белковой смеси тандемной масс-спектрометрией (Treumann and Thiede, 2010). Эти метки могут добавляться в среду и встраиваться в процессе роста клеток или путем химического мечения целого белка или триптических пептидов. Так, метод ICAT (isotope-coded affinity tags, сайт-специфическое мечение по цистеинам) основан на ковалентном мечении цистеинового аминокислотного остатка в полипептидной цепи химически идентичными, но изотопно разными реагентами (Gygi *et al.*, 1999; Schmidt *et al.*, 2004). При этом к белкам одного из образцов добавляют «легкую» метку, а к белкам другого «тяжелую». Другой метод, SILAC (Stable Isotope Labeling by Amino acids in cell Culture, изотопное мечение белков в культуре клеток), основан на встраивании метки в ходе метаболизма незаменимых аминокислот, содержащих стабильные изотопы (обычно ^{13}C / ^{15}N аргинин и/или лизин) (Ong *et al.*, 2002; Wang *et al.*, 2013). В данном случае экзогенные аминокислоты выступают в роли «тяжелой» метки. В обоих методах относительное содержание белков рассчитывается путем сравнения

интенсивности сигнала каждого из пептидов в легкой и тяжелой форме. Стоит отметить, что метод ICAT применим только к пептидам, содержащим остатки цистеина. В свою очередь метод SILAC подразумевает, что клетки бактерии не должны синтезировать лизин и аргинин самостоятельно и использовать исключительно находящиеся в среде меченые аминокислоты. Однако, *M. tuberculosis* обладает способностью синтезировать лизин эндогенно, что сразу ограничивает использование метода SILAC. Таким образом, в последнее время основной акцент делается на использовании безметочных методов MS квантификации белков, обладающих большей чувствительностью и производительностью (de Souza *et al.*, 2010; Gunawardena *et al.*, 2013). Последние в свою очередь основаны на анализе тотального пула белков каждого образца в отдельности. Таким образом, дальнейший сравнительный и количественный анализ можно проводить для большего числа образцов, в том числе полученных в разное время (Zhu *et al.*, 2010).

В свою очередь, целью таргетных технологий является анализ определенного набора белков в исследуемых образцах. Для этого в пептидомных и протеомных исследованиях необходимо определить так называемые протеотипические пептиды - фрагменты анализируемых белков, имеющие высокую степень ионизации и хорошую фрагментацию в условиях выбранных методов. В дальнейшем масс-спектрометр производит поиск и анализ только выбранных прекурсоров. Стоит отметить, что использование таргетных технологий позволяет существенно повысить чувствительность масс-спектрометрического определения необходимых белков, снижая тем самым временные затраты на работу прибора и обработку данных.

Новые протеомные подходы (например, SWATH-MS) совмещают панорамные, данный-независимые (independent data acquisition (IDA)) и таргетные данный-зависимые (dependent data acquisition (DDA)) методики с целью минимизировать ограничения каждой из методик (Gillet *et al.*, 2012). При этом для SWATH не требуется отбор единичных родительских ионов, а ионы прекурсоры пропускаются большими окнами (к примеру, в 25 Да). Таким образом, SWATH

позволяет так же идентифицировать и квантифицировать большой набор белков, подобно классическим shotgun подходам, но с точностью и воспроизводимостью SRM для большего количества образцов.

Развитие протеомики открыло новые пути для изучения туберкулеза, облегчив изучение многих сложных вопросов, в том числе касающихся взаимодействия бактерии с клеткой хозяина. Несмотря на то, что протеомика отстает от геномики и транскриптомики из-за инструментальных проблем и недостаточной чувствительности, все больше появляется работ, посвященных изучению инфекционных агентов с использованием протеомных подходов. Такие подходы позволяют выявлять потенциальные биомаркеры для диагностики (Liu *et al.*, 2010), факторы вирулентности и механизмы их действия (Malen *et al.*, 2011; Shui *et al.*, 2011), изучать реакцию хозяина и патогена на инфекционный процесс (Bhavsar *et al.*, 2010; Boshoff and Lun, 2010)). Стоит отметить, что наиболее весомый вклад в развитие протеомики микобактерий внесли работы группы Aebersold R. (Gygi *et al.*, 1999; Schubert *et al.*, 2013).

1.2.2.2 Протеомный анализ микобактерий

На сегодняшний день штамм H37Rv является наиболее хорошо изученным среди всех микобактерий. Как сказано выше, последовательность его генома была полностью расшифрована еще в 1998 году (Cole *et al.*, 1998). Не удивительно, что протеом этого штамма описан так же максимально полно. В частности, исследованы пулы белков мембраны (Xiong *et al.*, 2005; Malen *et al.*, 2011), культуральных фильтратов (Samanich *et al.*, 1998), клеточной стенки (Wolfe *et al.*, 2010; Feltcher *et al.*, 2015) и цитозольные (Covert *et al.*, 2001; Mawuenyega *et al.*, 2005; Mehaffy *et al.*, 2010). Исследования ведутся как на уровне единичных белков (Mascarello *et al.*, 2013; Parandhaman *et al.*, 2014; Schmidt-Schultz and Schultz, 2015), так и на уровне полного протеома (Schmidt *et al.*, 2004; Bell *et al.*, 2012). Так, на сегодняшний день, с использованием комплексных протеомных подходов, подтверждено существование 97 % белков от 4012 аннотированных для *M. tuberculosis* (Schubert *et al.*, 2013). В этой же работе авторы провели анализ белков

DosR регулона, ассоциированного с анаэробным выживанием *M. tuberculosis*, и зафиксировали изменения их представленности во время гипоксии. В частности, в условиях гипоксии представленность белка HspX увеличивалась в 340 раз. Стоит отметить, что до этого все исследования данного регулона базировались исключительно на уровне транскриптов (Park *et al.*, 2003; Rodriguez *et al.*, 2014).

Отдельно стоит упомянуть работы, авторы которых моделировали инфекционный процесс и оценивали белковый профиль *M. tuberculosis* в условиях, максимально приближенных к живому организму. В работе Cho с соавт. был проведен сравнительный анализ нерепликативной и персистирующий форм штамма H37Rv с использованием метода ICAT (Cho *et al.*, 2006). В данной работе было выявлено 38 высокопредставленных белков и 119 низкопредставленных белков нерепликативной формы по сравнению с контролем. Большинство этих белков относилось к системам деградации и энергетическому метаболизму, соответственно. В свою очередь метод SILAC использовался для описания изменений белкового профиля *M. tuberculosis*, при взаимодействии с макрофагами в ходе инфекционного процесса. Так был изучен предполагаемый фактор вирулентности, PPE38, относящийся к белковому семейству PE/PPE. Суть эксперимента заключалась в инфицировании макрофагов диким и мутантным по данному гену *Mycobacterium marinum*, модельным организмом и близким родственником *M. tuberculosis* (Wang *et al.*, 2013). Было выявлено, что данный белок участвует в ряде процессов, происходящих в макрофагах, в том числе в воспалительной реакции.

Большое количество работ посвящено поиску различий между вирулентным штаммом H37Rv и авирулентными микобактериями (*M. tuberculosis* H37Ra, *M. bovis* BCG). Основной целью данных работ является поиск факторов вирулентности *M. tuberculosis* и потенциальных кандидатов для разработки вакцин. Так, Malen с соавт. при сравнении мембранных белков штаммов H37Rv и H37Ra обнаружили, что представленность большинства белков одинакова в обоих штаммах, однако, 121 белок имеет различия в представленности более чем в пять раз (Malen *et al.*, 2011). Далее авторы сфокусировали свое внимание только на

мембранных липопротеинах, и, основываясь на наблюдаемом регулировании этих белков, предположили, что изменение метаболического состояния может играть роль в повышенной вирулентности. Другая группа авторов особое внимание уделяет белкам системы Esat-6 и ESAT-6-подобным белкам, которые обнаруживаются только в штаммах H37Rv, но не в H37Ra. Дополнительно было показано, гены пяти ESAT-6-подобных белков в штаммах H37Ra содержат мутации. Стоит так же отметить, что 6 кДа антиген (Esat-6), образует гетеродимерный комплекс с белком CFP-10 (Singhal *et al.*, 2012). Таким образом, предполагают, что система ESAT-6/CFP10 ассоциирована с вирулентностью *M. tuberculosis* и препятствует слиянию фагосомы и лизосомы внутри макрофагов хозяина, тем самым предотвращая разрушение клеток микобактерий (Tan *et al.*, 2006).

Не смотря на то, что H37Rv и *Mycobacterium bovis* BCG на геномном уровне имеют > 99.9 % гомологии, было выявлено 294 белка статистически отличающихся между двумя штаммами (Gunawardena *et al.*, 2013). Ранее сравнительный геномный анализ этих штаммов показал отсутствие определенных участков (RD) в геноме BCG и с потерей соответствующих генов было связано отсутствие патогенности (Brosch *et al.*, 2001). Среди описанных 294 различающихся белков, часть из них соответствовала описанным RD (Mattow *et al.*, 2001; Mattow *et al.*, 2003; Gunawardena *et al.*, 2013). Среди них особое внимание так же было уделено системе ESAT-6, белки которой ранее были предложены в качестве кандидатов для разработки новой вакцины (He *et al.*, 2003). Дополнительно в качестве кандидатных для разработки диагностических и вакцинных препаратов были описаны 22 дифференциальных белка, таких как ацетил-CoA C-ацетилтрансфераза (Rv0243) и некоторые Esat-6-подобные белки (Rv1198, Rv1793) (Mattow *et al.*, 2003).

В последние годы во многих странах мира наблюдается повсеместное распространение лекарственно-устойчивых штаммов *M. tuberculosis*. Феномен лекарственной устойчивости *M. tuberculosis* был обнаружен вскоре после начала применения стрептомицина, первого антибиотика, используемого в виде

монотерапии заболевания (Crofton and Mitchison, 1948). Устойчивость к противотуберкулезным препаратам у микобактерий наиболее часто связана с точечными мутациями, приводящими к модификации мишени действия лекарственного препарата или нарушению процесса его активации (Zhang *et al.*, 2005; Almeida Da Silva and Palomino, 2011). В последние два-три десятка лет шло интенсивное накопление информации о генетических механизмах устойчивости, как к препаратам первого, так и второго ряда. Однако, несмотря на разнообразие описанных мутаций до сих пор в ряде случаев не удается установить причины лекарственной устойчивости.

В связи с этим не удивительно, что проблему лекарственной устойчивости пытаются решить и на протеомном уровне. Ряд работ описывает особенности устойчивых к лекарственным препаратам штаммов *M. tuberculosis*, проводя их сравнительный анализ с чувствительными (Mattow *et al.*, 2006; Kumar *et al.*, 2013; Sharma *et al.*, 2015). Стоит отметить, что в основном были описаны белки, соответствующие известным маркерам лекарственной устойчивости. Однако, есть и ряд новых данных, например, описаны пять белков (Rv0491, Rv1446c, Rv2145, Rv2971 и Rv3028c) имеющие повышенную представленность в штаммах, устойчивых к изониазиду (Mattow *et al.*, 2006). Данные белки, как и классические маркеры, относились к мембранным белкам и потенциально могут являться мишенью для новых лекарственных препаратов. Анализ штаммов устойчивых к аминогликозидам показал в них повышенную представленность белков Rv0685, Rv1876 и Rv3841, ассоциированных с метаболизмом железа (Sharma *et al.*, 2015). Ассимиляция и утилизация железа для *M. tuberculosis* играют важную роль в росте, вирулентности и формировании латентной формы. Pandey и Rodriguez предположили, что ферритин (Rv3841) является обязательным для поддержания гомеостаза железа в клетках микобактерий и при его недостатке бактерии становятся более восприимчивы к воздействию антибиотиков (Pandey and Rodriguez, 2012). В схожем исследовании так же было отмечено увеличение представленности белков Rv1876 и Rv3224, участвующих в обмене железа, с указанием их возможной роли в формировании устойчивости к

противотуберкулезным препаратам второго ряда (Kumar *et al.*, 2013). В работе Singhal с соавт. при комплексном сравнении чувствительных и МЛУ штаммов были определены факторы вирулентности, такие как каталаза/пероксидаза (Rv1908c), активирующаяся в фагосомах, в резистентных штаммах (Singhal *et al.*, 2012). Ранее было показано, что каталаза/пероксидазная активность необходима для роста клеток и их персистенции в мышах и морских свинках (Li *et al.*, 1998), а так же в человеческих моноцитах периферической крови (Manca *et al.*, 1999). Дополнительно были отмечены белки Rv0036, Rv2032c, Rv0635, Rv1827 и Rv2896c, вовлеченные в клеточный метаболизм и способствующие внутриклеточному выживанию. В одной из работ последних лет в качестве биомаркеров для эффективного использования в серодиагностике резистентных штаммов туберкулеза были предложены белки Rv2031c, Rv3692 и Rv0444c (Zhang *et al.*, 2012).

Все описанные исследования показывают, что геномные вариабельности штаммов микобактерий действительно проявляются на уровне протеома и, таким образом, сравнительные протеомные исследования могут помочь в понимании фенотипических различий, таких как лекарственная устойчивость и вирулентность. Для выяснения, как генетическая изменчивость микобактерий отражается на их фенотипических различиях, необходимы более обширные сравнения, как на транскриптомном, так и на протеомном уровнях.

1.2.3 Транскриптомный анализ

Изучение транскриптома (комплекс транскриптов, продуцируемых бактерией) является так же дополнением к секвенированию геномов и реализуется с помощью различных подходов. ДНК-микрочипы являются одним из самых мощных инструментов для изучения дифференциальной экспрессии генов, исследования транскриптома. Самый распространенный формат представляет собой ПЦР продукты, специфичные для каждой ОРС в геноме, прикрепленные к микроскопическому предметному стеклу (Brown and Botstein, 1999). Профилирование экспрессии осуществляют двухцветной гибридизацией

(Shalon *et al.*, 1996). Меченную различными флюорохромами ДНК, получают из рНК, выделенной в двух экспериментальных условиях. Две популяции смешиваются в равных долях и гибридизуются на чипе. Сканирование чипа дает количественные интенсивности флуоресценции для каждого Fluor в каждой точке, из которой рассчитываются коэффициенты, чтобы определить относительные уровни экспрессии для каждого гена.

Технология микрочипов так же используется для видовой идентификации микобактерий на основе особой специфической области последовательности гена 16S рНК, и для обнаружения мутаций в генах, обуславливающих резистентность к противотуберкулезным препаратам.

В настоящее время транскриптомные технологии широко используются по всему миру с целью нахождения специфических молекулярных ответов по ключевым аспектам жизнедеятельности микроорганизма. При этом большая часть работ направлена на поиск новых мишеней для лекарственных препаратов и на анализ изменения экспрессии генов во время взаимодействия с клетками хозяина или меняющимися условиями среды. Первой работой, в которой дана характеристика экспрессии генов *M. tuberculosis* на полногеномном уровне *in vivo*, является работа Schnappinger с соавт. (Schnappinger *et al.*, 2003). В данной работе в качестве модельной системы использовались макрофаги мыши (активированные INF- γ и неактивированные). Вскоре после этого Cappelli с соавт. дали характеристику экспрессии генов *M. tuberculosis* уже из макрофагов человека на полногеномном уровне *in vivo* (Cappelli *et al.*, 2006). На настоящий момент исследование экспрессии генов микобактерий *in vivo*, проведенное Homolka с соавт. (Homolka *et al.*, 2010), является наиболее обширным. В нем был предпринят сравнительный анализ профилей экспрессии 3 клинических изолятов *M. africanum*, 12 клинических изолятов *M. tuberculosis*, а также 2 референсных лабораторных штаммов, *M. tuberculosis* H37Rv и CDC1551, в активированных и неактивированных макрофагах мыши. На основании полученных данных авторы выделили основной (коровый) транскриптом, включающий 280 генов (168 с универсально повышенной и 112 с универсально пониженной экспрессией),

экспрессия которых изменялась вне зависимости от штамма и статуса активации макрофага, тем самым позволяя сделать предположение об универсальной необходимости этих генов для обеспечения освоения внутриклеточной ниши микобактериями.

Другое направление транскриптомики связано с изучением процессов, происходящих в микробной клетке при непосредственном воздействии лекарственных препаратов. Первая работа такого рода была опубликована в 1999 году Behr с соавт. (Behr *et al.*, 1999). В ходе исследования был изучен транскрипционный профиль *M. tuberculosis* в ответ на воздействие противотуберкулезного препарата изониазида. Полученные результаты были подтверждены в работе Fu с соавт. в 2007 году (Fu and Shinnick, 2007). В ходе исследования был изучен ответ микробной клетки на воздействие этионамида, другого противотуберкулезного препарата со схожим механизмом действия. К настоящему времени существует большое количество работ, посвященных транскрипционному профилированию в ответ на воздействие, как отдельных противотуберкулезных препаратов, так и группы препаратов. Группой авторов предпринята попытка по созданию международной базы данных, объединяющей данные по аннотации генома патогена и экспрессии генов под воздействием различными внешних факторов (Raman *et al.*, 2008; Reddy *et al.*, 2009).

Несмотря на то, что в настоящее время достаточно много работ, посвященных изучению транскрипционной активности отдельных генов *M. tuberculosis* и их групп, как *in vitro*, так и *in vivo*, они не дают представления обо всех возможных изменениях и различиях между разными семействами и кластерами патогена.

1.2.4 Модельные эксперименты

Mycobacterium smegmatis является привлекательным модельным организмом для изучения особенностей патогенных микобактерий по нескольким причинам. Основным преимуществом является быстрый рост бактерии, примерно в 10 раз выше, чем у *M. tuberculosis* или *M. bovis* BCG.

Скорость деления у *M. smegmatis* составляет всего 2 – 3 часа по сравнению с 24 часами для *M. tuberculosis*. Таким образом, *M. smegmatis* образует видимые колонии на 2 – 4 день, в то время как туберкулезные микобактерии только на 3 – 4 неделе. В дополнении, *M. smegmatis* растет на простых питательных средах, содержащих глюкозу и соли. Так же, в отличие от *M. tuberculosis*, *M. smegmatis* не патогенный микроорганизм, поэтому с ним безопасно работать в стандартных лабораторных условиях.

Стоит отметить, что для 12 из 19 генов вирулентности *M. tuberculosis* были выявлены гомологи в *M. smegmatis* (Reyrat and Kahn, 2001). Первоначально некоторые из этих генов были описаны как уникальные для *M. tuberculosis* и ассоциированные с их вирулентностью. Затем их большая часть была отнесена к генам домашнего хозяйства микобактерий. Ввиду консервативности транскрипционного аппарата *M. tuberculosis* и *M. smegmatis* (Bashyam *et al.*, 1996), а также генов, кодирующих сигма факторы, включая SigF, систему DosR-DosS, считается, что *M. smegmatis* может служить полезной системой для изучения механизма адаптивного ответа *M. tuberculosis* на гипоксию.

M. smegmatis mc² 155 наиболее широко используемый лабораторный штамм. Он является мутантным вариантом родительского штамма mc⁶, который приобрел успешный фенотип после трансформации плазмидой (ept) (Snapper *et al.*, 1990). Следовательно, штамм mc² 155 легко трансформируется чужеродной ДНК в ходе гомологичной рекомбинации и может экспрессировать гены возбудителей лепры и туберкулеза что делает его уникальным инструментом для исследования микобактерий. Экспрессия чужеродных генов микобактерий позже стала возможна благодаря конструированию шаттл векторов, в которых использовали промотор *hsp60* *M. bovis* BCG для конститутивной экспрессии (например, pMV261 и pMV361) (Stover *et al.*, 1991). С другой стороны индуцибельные системы экспрессии дают возможность оценить экспрессию регулирующего гена и позволяют провести прямое сравнение индуцированных и неиндуцированных клеток. Стоит так же отметить, что в отношении исследования особенностей возбудителя туберкулеза *M. smegmatis* обладает значительными

преимуществами перед классическим модельным организмом *Escherichia coli* (Таблица 1). Таким образом, *M. smegmatis* играет важную роль в понимании биологии микобактерий. Учитывая широкий спектр плазмид, разрабатываемых научным сообществом (Triccas *et al.*, 1996), для экспрессии и очистки рекомбинантных белков из *M. smegmatis* клеток-хозяев были использованы различные индуцируемые системы экспрессий.

В настоящее время *M. smegmatis* широко используется для исследования физиологии возбудителя туберкулёзной инфекции, в том числе изучения ответа на окислительный стресс (Jain *et al.*, 2007; Singh *et al.*, 2015) и перехода в некультивируемое состояние (Cordone *et al.*, 2011).

Таблица 1.

Сравнительный анализ модельных организмов для изучения
M. tuberculosis

Характеристика	<i>E. coli</i>	<i>M. smegmatis</i>
Время необходимое для формирования колоний	8 часов	3 - 4 дня
Распознавание промоторов <i>M. tuberculosis</i>	не всегда	да
Гликозилирование белков <i>M. tuberculosis</i>	не всегда	да
Наличие миколовых кислот	нет	да
Репликация фагов <i>M. tuberculosis</i>	нет	да
Репликация рAL5000 эписомы	нет	да

1.3 Популяционная структура *Mycobacterium tuberculosis*

Первые результаты анализа геномных последовательностей *M. tuberculosis* показали крайне малое разнообразие патогена на генетическом уровне по сравнению с другими видами бактерий. Дополнительно, методами популяционной генетики, была установлена крайняя степень клональности популяционной структуры (Sreevatsan *et al.*, 1997; Musser *et al.*, 2000; Hirsh *et al.*,

2004). Это обстоятельство в свое время наложило ограничение на использование метода мультилокусного типирования, столь широко применяемого для других бактериальных патогенов. Однако с развитием сравнительной геномики и усовершенствованием технологий полногеномного секвенирования на геномном уровне было найдено значительное количество вариаций, которые могут быть использованы для построения филогенетически надежных классификаций (Gagneux and Small, 2007). Такими информативными маркерами для *M. tuberculosis* послужили LSPs и SNPs. Методы, используемые для их нахождения, описаны в главе 1.2.1 Молекулярно-генетический анализ. Здесь же важно подчеркнуть, что из-за достаточно малой вариации геномной последовательности патогена (штаммы в среднем отличаются друг от друга не более чем на 2000 SNPs) вероятность возникновения независимой рекуррентной мутации крайне мала. Отсутствие горизонтального переноса генов дополнительно уменьшает возможность обратной мутации, а в случае с LSPs предотвращает появление геномных последовательностей, потерянных ранее. Основные работы, посвященные изучению популяционной структуры патогена, представлены на Рисунке 4.

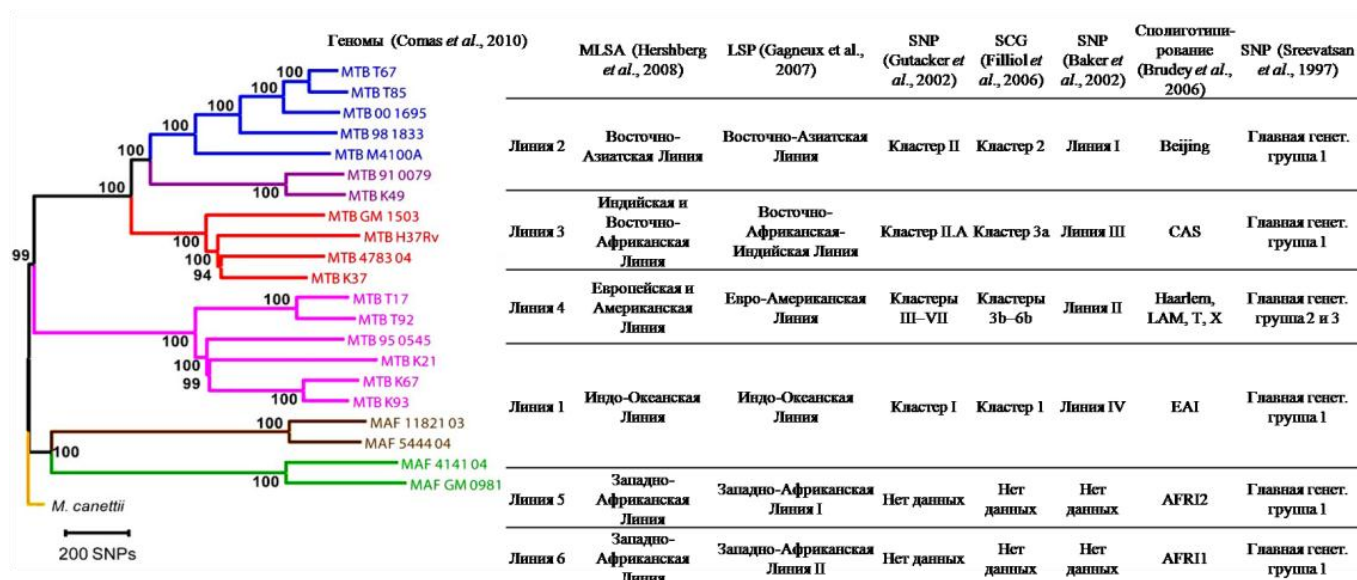


Рисунок 4. Молекулярные маркеры, используемые для определения основных линий внутри *M. tuberculosis* и *M. africanum*.

В 1997 году Sreevatsan с соавт. представили результаты анализа последовательностей 26 структурных генов на большой коллекции образцов (Sreevatsan *et al.*, 1997). В своей работе авторы разделили штаммы *M. tuberculosis* на три главные генетические группы на основании двух нейтральных несинонимичных замен (Leu463Arg гена *katG* и Thr95Ser гена *gyrA*). Отметим, что это было сделано за год до определения полной геномной последовательности штамма H37Rv и менее чем за 10 лет до построения строгой глобальной филогении патогена.

Так, уже в 2004 году Baker с соавт. провели типирование 225 клинических изолятов по 7 генам домашнего хозяйства и разделили *M. tuberculosis* на 4 основные линии (Baker *et al.*, 2004). В дальнейшем Gagneux с соавт. проанализировали филогенетически информативные LSPs на глобальной выборке штаммов *M. tuberculosis*, состоящей из 875 образцов из 80 стран (Gagneux *et al.*, 2006). Все изоляты принадлежали к одной из 6 основных филогенетических линий, включая две линии, которые традиционно относятся к *M. africanum*. При этом 4 остальные филогенетические линии были конгруэнтны с линиями, определенными в исследовании Baker с соавт. (Gagneux and Small, 2007). В свою очередь, Gutaker с соавт. проанализировали 5069 клинических изолятов по 36 синонимичным полиморфизмам, выявленных в ходе *in silico* анализа *M. tuberculosis* (Gutacker *et al.*, 2006). Авторы описали 9 основных филогенетических линий. В схожем исследовании Filliol с соавт. провели анализ 219 образцов *M. tuberculosis* и *M. bovis* по 159 sSNPs (Filliol *et al.*, 2006). В ходе работы было выявлено 10 основных групп, включая линию, специфическую для *M. bovis*.

Отдельно стоит отметить работы Comas с соавт., в которых авторы на уровне выравнивания целых геномов подтвердили основные филогенетические линии *M. tuberculosis*, дополнительно определив эволюционные взаимосвязи между ними (Comas *et al.*, 2010). Стоит подчеркнуть, что в ходе последней работы коллеги описали филогенетическую линию 7, ранее не определенную ни одной из научных групп (Comas *et al.*, 2013).

Помимо решения задач по классификации и типированию штаммов, методы сравнительной геномики также внесли большой вклад в изучение эволюции патогена. В ходе исследований было определено, что представители *M. tuberculosis* complex зародились в Африке, а *M. canettii* является ранней эволюционной ветвью микобактерий туберкулеза (Comas *et al.*, 2013; Supply *et al.*, 2013). *M. canettii*, также называемые «гладкие» микобактерии и/или *Mycobacterium prototuberculosis* имеют гладкую форму колоний при культивировании на питательной среде, в отличие от *M. tuberculosis* complex имеют большое количество однонуклеотидных полиморфизмов (около 60000 SNPs при сравнении с геномом *M. tuberculosis* H37Rv) и характеризуется частыми рекомбинационными событиями (Supply *et al.*, 2013). Gutierrez с соавт. было выдвинуто предположение, что микобактерии туберкулезного комплекса являются успешной клональной субпопуляцией *M. canettii*, а общий предок мог существовать около 3 млн. лет назад (Gutierrez *et al.*, 2005). Дополнительно, в ходе исследования геномов 259 штаммов *M. tuberculosis* complex методом коалесцентного анализа было показано, что представители комплекса образовались около 70000 лет назад и в дальнейшем расселились по миру в ходе миграционных потоков из Африки. При этом следует отметить выраженную взаимосвязь биогеографии человеческой популяции и популяции патогена, своего рода коэволюцию хозяин-патоген, приведшую к образованию линия-специфических патобиологических особенностей (Homolka *et al.*, 2010; Brites and Gagneux, 2012). Но при этом все же существуют филогенетические линии *M. tuberculosis*, преобладающие во всем мире и, тем самым, считающиеся успешными. Об одной из таких линий, филогенетической линии Beijing, более подробно будет изложено в следующей главе.

1.4 Генетическое семейство Beijing

Впервые представители генотипа Beijing были обнаружены в 90х годах XX века, в двух независимых исследованиях, проведенных группами исследователей из Голландии и Америки (van Soolingen *et al.*, 1995; Bifani *et al.*, 1996). В ходе

IS6110 RFLP анализа и сполиготипирования коллекции изолятов *M. tuberculosis*, полученных от больных ТБ в 1992 - 1994 годах в Китайской Народной Республике и Монголии, van Soolingen с соавт. выявили доминирующий генотип, характеризующийся уникальным сполигопротипом и схожими паттернами рестрикционного анализа (van Soolingen *et al.*, 1995). При этом наиболее часто представители генотипа встречались в окрестностях Пекина (англ. Beijing), отчего и получили свое название (van Soolingen *et al.*, 1995). Параллельно этому исследованию, Bifani с соавт. из Научно-исследовательского института общественного здравоохранения США (Public Health Research Institute, NY, США), в 1996 году методами молекулярно-генетической эпидемиологии описали вспышку лекарственно-устойчивого туберкулеза, произошедшую в Нью-Йорке в начале 1990х. Выявленные штаммы характеризовались крайне схожими паттернами IS6110 профилей и были названы «W» (Bifani *et al.*, 1996). В дальнейшем эти названия были объединены в W/Beijing или просто Beijing (Kurepina *et al.*, 1998; Van Soolingen, 2001). При этом название как нельзя лучше отражает реальное место зарождения генотипа. Его представители наиболее часто встречаются в Восточной Азии и, по мнению Мокроусова с соавт., генотип Beijing возник в Северном Китае более 1000 лет назад. Дальнейшее его распространение было связано с миграционными потоками: со средневековых времен в Россию, совсем недавно в ЮАР (с XVII века) и в Австралию (в XIX веке) (Mokrousov *et al.*, 2008). В свою очередь, согласно Merker с соавт., генотип в целом возник более 6000 лет назад в географической зоне, включающей в себя Северо-Восток Китая, Корею и Японию (Merker *et al.*, 2015).

Основные филогенетические маркеры представителей генотипа отображены на Рисунке 4. Стоит отметить, что в настоящее время наиболее простым и быстрым методом выявления штаммов Beijing считается сполиготипирование. Согласно интернациональной базе данных SITVIT (http://www.pasteur-guadeloupe.fr:8081/SITVIT_ONLINE/) (Demay *et al.*, 2012) штаммы генотипа характеризуются сполиготипом SIT 1 (от англ. Spoligotype International Type, международный сполиготип), который определяется наличием в DR локусе 9 из

43 спейсеров (с 35 по 43). Внутри генотипа исследователи выделяют от 2 до более десятка групп. Дифференциация зачастую носит весьма условный характер и зависит от выбранного метода типирования (Ebrahimi-Rad *et al.*, 2003; Tsolaki *et al.*, 2005; Filliol *et al.*, 2006; Gagneux and Small, 2007; Mokrousov *et al.*, 2008). При этом наиболее клинически значимым является деление представителей генотипа на атипичные/древние и типичные/современные штаммы (Tsolaki *et al.*, 2004; Mokrousov *et al.*, 2005; Tsolaki *et al.*, 2005). Особенности типичных/современных штаммов при этом будет являться делеция RD150 и/или RD142, а также наличие повторяющегося элемента IS6110 в области NTF. Наиболее распространенные классификации штаммов генотипа Beijing приведены на рисунке 4.

Согласно международной базе данных сполиготипирования SpolDB4, штаммы Beijing присутствуют в наибольшем количестве стран на глобальном уровне (13 % от мирового количества изолятов), являясь по этому показателю уникальным генотипом (Brudey *et al.*, 2006). Здесь также следует отметить ассоциацию представителей генотипа с многочисленными вспышками заболеваний во всем мире, многие из которых были лекарственно устойчивые (Frieden *et al.*, 1996; Agerton *et al.*, 1999; Caminero *et al.*, 2001; Affolabi *et al.*, 2009). В структуре популяции возбудителя туберкулеза в России доля штаммов Beijing составляет от 50 % до 80 % (Норкина с соавт., 2003; Mokrousov *et al.*, 2003), причем, крайне выражена ассоциация штаммов с лекарственной устойчивостью (Casali *et al.*, 2014). Исходя из сказанного выше, предполагается, что у штаммов данной эволюционной линии, возможно, развились уникальные свойства, которые позволили им распространиться по всему миру (клональная экспансия). По мнению многих авторов, этими свойствами являются: 1) способность «ускользнуть» от БЦЖ-вакцинирования 2) способность штаммов относительно быстро приобретать устойчивость к противотуберкулезным препаратам.

Гипотеза о сниженной эффективности БЦЖ-вакцинации против штаммов генотипа Beijing была высказана еще в 1995 году van Soolingen с соавт.(van Soolingen *et al.*, 1995). Авторы сделали это предположение на основании того, что

наиболее часто штаммы генотипа были выявлены среди БЦЖ-вакцинированных пациентов. В дальнейшем это предположение подтверждалось и опровергалось в различных популяционных исследованиях, и до сих пор мировое сообщество не пришло к единому мнению. Так, например, в исследованиях по циркуляции микобактерий в Китае, Монголии, Вьетнаме и Голландии была найдена ассоциация штаммов Beijing с БЦЖ-вакцинацией (van Soolingen *et al.*, 1995; Kremer *et al.*, 2004; Li *et al.*, 2006). При этом, в исследованиях по Южной Америке, Вьетнаму, Индонезии и Тибету, напротив, такой закономерности выявлено не было (Ahn *et al.*, 2000; Shi *et al.*, 2007; van Crevel *et al.*, 2009). Не привнесли ясности и исследования на модельных животных. Как и в популяционных исследованиях, одна часть работ указывает на сниженный протективный эффект БЦЖ-вакцинации при сравнении штаммов Beijing и не Beijing (Castanon-Arreola *et al.*, 2005; Grode *et al.*, 2005), а другая такой закономерности не выявляет (Jeon *et al.*, 2008).

Ассоциация представителей генотипа Beijing с лекарственной устойчивостью показана в достаточно большом количестве эпидемиологических исследований, однако и здесь ярко выражена неоднозначность получаемых данных. Так, например, во многих странах зафиксированы эпидемичные вспышки туберкулеза, ассоциированного с лекарственной устойчивостью (Bifani *et al.*, 2002; Drobniowski *et al.*, 2005; Park *et al.*, 2005; Caws *et al.*, 2008), но также есть и примеры быстрого распространения чувствительных клонов (Bifani *et al.*, 2002). Если рассматривать эндемичные штаммы, то наибольший контраст проявляется при сопоставлении данных из России, где около 20 % представителей генотипа Beijing будут нести лекарственную устойчивость (Casali *et al.*, 2014), и данных из Китая и Японии, где практически все штаммы будут чувствительны к проводимой противотуберкулезной терапии. Так же важно отметить, что, по предположению Rad с соавт., представители генотипа Beijing, а точнее ветвь типичных/современных штаммов, несет мутаторный генотип (Ebrahimi-Rad *et al.*, 2003). Объясняется это найденными мутациями в генах системы репарации *mutT2*, *mutT4* и *ogt*. По мнению авторов это может приводить к изменению скорости

накопления мутаций, в том числе в генах, ассоциированных с лекарственной устойчивостью. Однако, согласно работе Werngren с соавт. в ходе флюктуационного теста Лурии и Дельбрюка уровень спонтанных мутаций в штаммах Beijing не отличался от штаммов других генотипов (Werngren and Hoffner, 2003). При этом в работе исследователей из Гарвардской школы общественного здравоохранения (США), тем же методом было определено, что у штаммов генотипа Beijing частота мутаций выше, чем у штаммов четвертой филогенетической линии (Ford *et al.*, 2013).

Не менее интересными выглядят исследования вирулентности штаммов Beijing. Большое количество исследований сообщают о гипервирулентности штаммов представленного генотипа, как на уровне макрофагальных моделей, так и на уровне эпидемиологических исследований. Однако стоит отметить, гипервирулентность не является характерной особенностью представителей семейства, а лишь присуща определенным генетическим сублиниям внутри Beijing, зачастую ассоциированным со вспышками заболеваний в том или ином регионе (Bifani *et al.*, 2002; Hanekom *et al.*, 2011).

В последние годы, опубликован ряд работ, посвященных сравнению протеомных профилей штаммов семейства Beijing и H37Rv, в которых были описаны существенные различия между ними. Так, белки эффлюксных помп, Rv0341, Rv2688c, Rv3728, были найдены исключительно в штаммах семейства Beijing (de Keijzer *et al.*, 2014). Была продемонстрирована повышенная представленность некоторых факторов вирулентности (Rv0129c, Rv0831c, Rv1096, Rv3117, Rv3804c), в штаммах семейства Beijing (Pheiffer *et al.*, 2005), что так же может свидетельствовать в пользу повышенной вирулентности данных штаммов. «Успешность» семейства Beijing частично может быть обусловлена уменьшением экспрессии белков Hsp65 (Rv0440), Pst1 (Rv0934) и Rv1886c, так как ранее было предложено, что снижение экспрессии определенных основных антигенов может позволить штаммам, избегать иммунных ответов хозяина (Monahan *et al.*, 2001; Sherman *et al.*, 2001). В связи с этим очевидно, что

актуальным является более детальное изучение протеомного профиля штаммов данного семейства.

Однако на сегодняшний день опубликовано только две работы, фокусирующиеся на изучении белкового разнообразия непосредственно внутри семейства Beijing. В первой работе de Souza с соавт. использовали безметочный LC-MS/MS подход для сравнения протеомных профилей гипо- и гипервирулентных штаммов данного семейства и нашли около 50 высокопредставленных белков в каждом штамме, при общем количестве 1668 идентифицированных белков (de Souza *et al.*, 2010). Авторами было показано снижение представленности белка ESAT-6 в гипервирулентных штаммах, в то время как ранние исследования, как было уже сказано, отмечали ESAT-6 как характеристику вирулентных штаммов (He *et al.*, 2003; Tan *et al.*, 2006). Для верификации найденной особенности уровень ESAT-6 в двух штаммах был проверен методом ПЦР в реальном времени, в ходе которой так же был выявлен низкий уровень транскриптов в гипервирулентных штаммах (de Souza *et al.*, 2010). Этот неоднозначный результат доказывает всю сложность роли путей секреции ESAT-6 в патогенности *M. tuberculosis*. Дополнительно, он может свидетельствовать в пользу фенотипических различий между штаммами семейства Beijing и штаммами H37Rv.

Другая работа de Keijzer с соавт. посвящена протеомным особенностям штаммов *M. tuberculosis*, относящихся к древним (атипичным) и современным (типичным) сублиниям семейства Beijing (de Keijzer *et al.*, 2014). В работе описано 2392 белка, полученных с помощью SILAC-LC-MS подхода. Не смотря на то, что по белковому профилю обе сублинии оказались крайне схожи, авторам удалось выявить различия в представленности белков Rv0450c/MmpL4, Rv1269c, Rv3137 и Rv3283/SseA. Так, среди белков, низкопредставленных в современных Beijing, авторы отмечают Rv3283/SseA, уровень транскриптов которого оказался так же значительно ниже в типичных штаммах. При этом представленность белков Rv0450c и Rv3137 была значительно выше в тех же штаммах.

1.5 Характеристика кластера Beijing B0/W148

При описании генетического семейства Beijing в главе выше довольно часто речь заходила об «успешных» представителях генотипа. Одним из таких «успешных» кластеров является Beijing B0/W148. Впервые штаммы кластера были выявлены на рубеже XX и XXI веков. В независимых исследованиях Нарвской, Курепиной и Portaels с использованием IS6110 RFLP типирования обнаружили группы кластеризующихся образцов, названные B0 (отечественная систематика) и W148 (иностранная систематика). Отличительной особенностью этих образцов было наличие двойной полосы (7.1 и 9.2 Kb) в верхней части профиля. В 2008 году штаммы Beijing B0/W148 с *sensu stricto* профилем и характерной двойной полосой были отнесены к Beijing B0/W148 (Mokrousov *et al.*, 2008), а метод IS6110-RFLP типирования был признан «золотым стандартом» для выявления изолятов данной клональной группы. Дополнительными названиями кластера могут считаться CladeB (Casali *et al.*, 2014) и ECDC0002 (de Beer *et al.*, 2014).

По литературным данным клональная группа составляет около одной четверти от всех изолятов генотипа Beijing, циркулирующих в разных частях России и бывшего Советского Союза, а так же среди российских иммигрантов в Соединенных Штатах Америки и Европе (Нарвская с соавт., 1999; Нарвская, 2003; Kruuner *et al.*, 2001; Bifani *et al.*, 2002; Kubica *et al.*, 2004; Cox *et al.*, 2005). Согласно гипотезе Мокроусова с соавт. клон Beijing B0/W148 распространился на территории бывшего Советского Союза исторически недавно из-за его уникальных патогенных особенностей и, таким образом, он может представлять собой успешный клон *M. tuberculosis* (Mokrousov *et al.*, 2008). Тем не менее, полная картина о времени и месте происхождения, эволюции и биологических свойствах штаммов Beijing B0/W148 остается неясна.

Изоляты кластера B0/W148 относятся к ветви типичных/современных Beijing и имеют профиль сполитипирования, ничем не отличающийся от других представителей генотипа (Mokrousov *et al.*, 2002; Kubica *et al.*, 2004; Matrakshin *et al.*, 2004). При более детальном анализе, на основе 68-спейсерного

сполиготипирования, было определено, что CRISPR locus в штаммах Beijing B0/W148 состоит из 15 единиц, что также является общей чертой большинства вариантов генотипа Beijing (Mokrousov *et al.*, 2008). Согласно 12 локусному MIRU-VNTR типированию штаммы кластера соответствуют типу M11 (профиль 223325173533) (Mokrousov *et al.*, 2005; Mokrousov *et al.*, 2012), также называемому MIT17 (база данных SITVIT Web), который является крупнейшим в Восточной Азии, и вторым в России (Mokrousov, 2008; Mokrousov *et al.*, 2012). Лишь на уровне 24-локусного VNTR типирования штаммы характеризуются специфическим паттерном 244233352644425173353723 (локусы перечислены в соответствии с их расположением в геноме по часовой стрелке), отличающим их от других штаммов Beijing (Mokrousov *et al.*, 2008). Также следует отметить, что наличие семи копий MIRU26 и QUB26 первоначально было предложено для специфической идентификации B0/W148 в северо-западном регионе России (Mokrousov *et al.*, 2008).

Полная последовательность первого генома *M. tuberculosis* кластера B0/W148 была опубликована в виде единого скаффолда в 2011 году (штамм W-148; GenBank номер ACSX00000000.1). В конце 2015 года геном был полностью секвенирован и аннотирован (штамм W-148; GenBank CP012090.1). Это послужило чрезвычайно полезным достижением для более глубокого изучения штаммов этого кластера. В частности, открыло новые перспективы для поиска новых молекулярных мишеней для быстрой диагностики. Несмотря на то, что метод IS6110-RFLP типирования считается «золотым стандартом» для идентификации изолятов кластера B0/W148, он является довольно громоздким и длительным. В связи с этим, Мокроусовым с соавт. был разработан более простой метод детекции изолятов на основе мультиплексной ПЦР (Mokrousov *et al.*, 2012). Было доказано, что данный метод является более чувствительным и специфичным, в связи с чем и был рекомендован для простого и надежного обнаружения изолятов B0/W148 в районах их эпидемической циркуляции, например, в России и других странах бывшего Советского Союза, а также в районах проживания иммигрантов из этих стран (Mokrousov *et al.*, 2012).

Относительно вирулентности штаммов Beijing B0/W148 в научном мире продолжается дискуссия. Так, в ходе по-разному спроектированных исследований, основанных на использовании мышиных макрофагальных моделей, разные исследовательские группы получили согласующиеся результаты и показали повышенную вирулентность штаммов Beijing B0/W148 по сравнению с другими вариантами генотипа Beijing (Andreevskaia *et al.*, 2006; Andreevskaia *et al.*, 2007; Lasunskaia *et al.*, 2010). Схожие результаты были продемонстрированы Lasunskaia с соавт. (Lasunskaia *et al.*, 2010) на моноцитоподобной линии клеток THP-1. Штаммы семейства Beijing из России (в том числе B0/W148) были значительно токсичнее, чем другие штаммы Beijing, и индуцировали характерный паттерн экспрессии цитокинов, способствующий иммуносупрессии: снижение продукции TNF и высокий уровень секреции интерлейкина-10.

В отличие от исследований на основе макрофагальных моделей, работы двух российских групп с использованием животных моделей показали контрастирующие результаты (Vishnevskii *et al.*, 2002; Chernousova *et al.*, 2008; Zemskova *et al.*, 2010). Вишневский с соавт. описали увеличение вирулентности изолятов B0/W148 на линии белых неинбредных мышей (Vishnevskii *et al.*, 2002). Напротив, в исследованиях Черноусовой с соавт. и Земсковой с соавт. была показана неоднородность изолятов генотипа Beijing по выживанию инбредной линии мышей C57BL/6 и отсутствие особых вирулентных свойств у изолятов B0/W148 по сравнению с другими Beijing (Chernousova 2008, Zemskova *et al.*, 2010).

На сегодняшний день опубликовано достаточно много данных об ассоциации штаммов кластера с лекарственной устойчивостью, что является достаточным, для оценки степени опасности, исходящей от циркуляции изолятов B0/W148. Первая статья, описывающая российские штаммы *M. tuberculosis*, выделенные в середине 1990х, уже показала широкую распространенность МЛУ среди изолятов данного кластера (Marttila *et al.*, 1998). Недавно изоляты B0/W148 с множественной лекарственной устойчивостью были выявлены в эпидемиологически значимой выборке пациентов с впервые выявленным ТБ в

Ленинградской (Нарвская, 2003), Тульской (Dubiley *et al.*, 2010), Самарской и других областях. Следует отметить, что в исследовании более 1000 штаммов из Самары 119 штаммов относилось к кластеру B0/W148 и все они были лекарственно устойчивы. В Абхазии, 22 из 23 изолятов B0/W148 были МЛУ, в то время как другие включенные в исследование изоляты генотипа Beijing были чувствительны к противотуберкулезным препаратам (10 из 55) (Pardini *et al.*, 2009). В Эстонии изоляты кластера B0/W148 составили 37.2 % от всех лекарственно устойчивых штаммов генотипа Beijing, в то время как ни одного чувствительного штамма Beijing B0/W148 в данном исследовании выявлено не было, включая штаммы, выделенные в 1994 году (Kruuner *et al.*, 2001). Также следует отметить, что в исследовании 2092 образцов из 24 стран Европы методом VNTR кластер B0/W148 был выявлен в 470 случаях (17 стран, преимущественно Восточная Европа). Согласно исследованию этот кластер называется ECDC0002 и в крайней степени ассоциирован с лекарственной устойчивостью.

Довольно интересной является гипотеза о происхождении штаммов и первичном их распространении. По мнению Мокроусова штаммы кластера зародились в Сибири до 1960х годов, что в целом согласуется с исследованием Merker с соавт. (касательно даты возникновения). В дальнейшем, в ходе программы по освоению целины (1955 - 1960 годы), миграционные потоки были направлены в Казахстан и в Сибирь. Следует отметить, что в Казахстане представленность кластера B0/W148 крайне мала и составляет около 4 %. Это подтверждает гипотезу автора о том, что в европейской части России штаммов кластера в те годы еще не было. В свою очередь вторая волна миграции, 1960 - 1980 годы, напротив, из Сибири по всей стране могла повлечь массовое распространение представителей кластера. Согласно Мокроусову, триггером распространения именно устойчивого клона могло послужить повсеместное использование открытого в 1963 году рифампицина.

1.6 Заключение

Анализ эпидемиологических и литературных данных показал, что в настоящее время туберкулез остается одним из социально значимых заболеваний в мире. Разработка лекарственных препаратов и вакцины позволили лишь на время снизить угрозу распространения туберкулеза по миру и количество смертельных случаев. Сегодня же, наблюдается некоторое ухудшение эпидемиологической обстановки в связи с активным распространением лекарственноустойчивых форм *M. tuberculosis*, в том числе устойчивых ко всем известным лекарственным препаратам. Таким образом, исследования возбудителя туберкулеза активно проводятся учеными во всем мире.

Несмотря на большие успехи мирового сообщества в области изучения геномов микобактерий, в настоящее время ученые приходят к выводу о недостаточности и ограниченности получаемых данных. Таким образом, лишней раз была показана актуальность проведения дополнительных анализов, таких как протеомный. Отдельно стоит отметить, что актуальным является изучение наиболее распространенных штаммов, представляющих опасность для того или иного региона. Так, преобладание штаммов кластера Beijing B0/W148 на территории России и стран бывшего СНГ, делает его предпочтительным объектом для исследования.

2. МАТЕРИАЛЫ И МЕТОДЫ

2.1 Бактериальные штаммы

Коллекция клинических образцов включала в себя 10 изолятов *M. tuberculosis* кластера B0/W148, полученных от пациентов, проходивших лечение в поликлинике Санкт-Петербургского НИИ Фтизиопульмонологии с 2005 по 2015 год. Идентификацию и выделение чистой культуры патогена проводили в специализированной лаборатории представленного НИИ согласно соответствующим протоколам (Министерство здравоохранения РФ. О совершенствовании противотуберкулезных мероприятий в Российской Федерации // Приказ № 109 от 21 марта 2003 г).

Лабораторный штамм *M. tuberculosis* H37Rv получен из коллекции Санкт-Петербургского НИИ Фтизиопульмонологии.

Штамм *M. smegmatis* mc² 155 был получен из коллекции Института биохимии им. А. Н. Баха.

2.2 Культивирование микобактерий

Штаммы *M. tuberculosis* кластера B0/W148 и H37Rv культивировали на твердой питательной среде Middlebrook 7H11 (Becton, Dickinson and Company, США) с добавкой OADC (Becton, Dickinson and Company, США) при 35°C 14 - 16 дней до плотности $\pm 2 \cdot 10^8$ клеток в мл. Штамм *M. smegmatis* mc² 155 культивировали на плотной питательной среде Middlebrook 7H9 (HiMedia, Индия) с добавкой OADC (HiMedia, Индия) при 37°C и 2 % CO₂, 1 - 2 дня до плотности $\pm 2 \cdot 10^8$ клеток в мл.

2.3 Генетический анализ *Mycobacterium tuberculosis*

2.3.1 Геномные последовательности

В исследование включили выборку геномных последовательностей 92 образцов *M. tuberculosis* из коллекции *key strains*, представленных в виде чтений с прибора Illumina и находящихся в открытом доступе в базе данных NCBI ERP000124. Дополнительно в работе использовали геномные последовательности

11 образцов генотипа Beijing, представленные в виде протяженных секвенсов (полные геномы, скаффолды, контиги) (Таблица 2)

Таблица 2.

Коллекция геномных секвенсов образцов

Образец	Номер в GenBank
94 M4241A	GCA_000155145.1
02_1987	GCA_000658295.1
CTRI-4	GCA_000256525.1
R1207	CM001045.1
T85	NZ_JLCX01000000
Strain 210	GCA_000176315.1
HN878	CM001043.1
CCDC5180	CP002885.1
X122	CM001044.1
CCDC5079	CP002884.1
W-148	CP012090.1

2.3.2 Амплификация фрагментов генома

Выделение тотальной геномной ДНК образцов проводили в Санкт-Петербургском НИИ Фтизиопульмонологии по методу van Embden с соавт. (van Embden *et al.*, 1993).

Амплификацию фрагментов ДНК проводили в 25 мкл реакционной смеси, содержащей 66 mM Tris-HCl pH 9.0; 16.6 mM (NH₄)₂SO₄; 2.5 mM MgCl₂, по 250 мкМ каждого дНТФ, 1 ед. Taq-полимеразы (ЛИТЕХ, Россия) и по 10 пмоль соответствующих праймеров. Праймеры для амплификации подбирали с помощью программы Oligo 6.71 (Molecular Biology Insights, Inc., Cascade, США). Праймеры, используемые для амплификации инвертированных фрагментов генома, представлены в Таблице 3.

Таблица 3.

Праймеры для амплификации фрагментов генома, участвующих в рекомбинационных событиях

Название праймера	Последовательность (5'→3')
P1	GTGTTGTACATTGGGCATCG
P2	GGTGTACATATCGAAGCTCG

P3	GCTCGACGAAGTGAGATTGC
P4	TTGGCGATCCGATACAGTGC
P5	CTGCCAAGCACTGGACAGC
P6	CAAGTCTCCGGTATTCAAGG
P7	AGCCTTGGCTCGTCCTTACC
P8	CACGGCTCTCCCAACGTGG

Полимеразную цепную реакцию осуществляли на приборе DNAEngineTetrad 2 (MJResearch, США). Общий вид профиля представлен ниже:

94°C – 5 мин.

30–45 циклов 94°C – 15 сек.

 Ta°C – 15 сек.

 72°C – от 15 сек. до 3 мин.

72°C – 10 мин.

4°C – хранение.

Где Ta – температура отжига праймера

Детекцию продуктов амплификации осуществляли методом электрофореза в 1 % агарозном геле. Для окраски гелей использовали бромистый этидий.

Анализ тандемных повторов по 24 локусам генома (VNTR-типирование) для образцов *M. tuberculosis* проводили по методике, описанной ранее (Supply *et al.*, 2006).

2.3.3 Секвенирование фрагментов генома модифицированным методом Сенгера

Для определения нуклеотидных последовательностей фрагментов генома использовали метод Сенгера с модификациями (Sanger *et al.*, 1992) на приборе ABIPrism 3730xl Genetic Analyzer (AppliedBiosystems, США; Hitachi Япония). Для секвенирования использовали амплификационные праймеры (Таблица 3).

Для последующего восстановления полноразмерных последовательностей, а так же выравнивания и сравнительного анализа использовали ContigExpress и AlignX программного пакета Vector NTI 11.0 (Informax Inc, США).

2.3.4 Рестрикционный анализ

Рестрикционный анализ проводили согласно методике van Embden с соавт. (van Embden *et al.*, 1993) с модификациями. Гидролиз геномной ДНК проводили рестриктазой *MluI* (15 ед.) (Thermo Scientific, США) в рекомендованном буфере в течение ночи при 37°C. Далее проводили инактивацию фермента (2 часа при 65°C) и разделяли полученные фрагменты рестрикции в 0.8 % агарозном геле. Зонды для анализа по Саузерну получали с помощью обычной ПЦР с использованием Amersham ECL системы (GE Healthcare, США) с определенным набором праймеров (Таблица 4). Полученные на ECL светочувствительной пленке профили сканировали и обрабатывали программным пакетом BioNumerics версия 5.1 (Applied Maths, Бельгия).

Таблица 4.

Праймеры для RFLP анализа

ДНК зонд	Праймеры для амплификации (5'-3' последовательность)	Длина
А	GGCGTTGAACATCAAAGATCC	412
	CCAGTGCTCTGCTGATGACC	
В	ACGATGTCCCATAATCGTTGG	392
	GAGGCGTCGTGACCGACG	
С	CGAGTGACCGATGAAGAAGG	343
	CGCAGGCAACAACAАСТTCG	
D	TTGGGAGAGCCGTGATTTGC	309
	CGACTACAACCTAGGCAACG	
Е	GATCCACCAACAGCGCAACG	279
	TGAGCAGGCCGTGCTGTCC	
F	CCTCACCCAAATTGGCTTGC	262
	TGGTGATTCTGAATTGGAAGG	
G	AACGCCTCCATCTCCACACC	354
	GAACGACCACGATGGGGTGC	
H	GGTGAАСТTGTСGGGATTTGG	324
	GGTGACCTGGACACGGTGG	

2.3.5 Полногеномное секвенирование

Полногеномные последовательности 10 штаммов, включенных в исследование, определяли с использованием высокопроизводительного

секвенатора GS FLX + (Roche, США). Для получения фрагментных библиотек использовали коммерческий набор GS FLX Titanium Rapid Library Preparation Kit (Roche, США) согласно протоколу фирмы-производителя (Rapid Library Preparation Method Manual GS FLX+ Series – XL+, Roche, США). Средний размер библиотек соответствовал 1400–1800 п.о.

Образец SP21 дополнительно секвенировали на приборе Ion Torrent PGM (Life Technologies, США) с использованием двух библиотек парных фрагментов (размеры библиотек 2000–3000 п.о. и 5000–6000 п.о.). Для приготовления библиотек использовали коммерческий набор 5500 SOLiD Mate-Paired Library Construction Kit (Life Technologies, США) в соответствии с протоколом Ion Mate-Paired Library Preparation (Life Technologies Demonstrated Protocol).

2.3.6 Анализ данных полногеномного секвенирования

Сборку полногеномных последовательностей (ассемблинг) 10 образцов, включенных в исследование, проводили с помощью программного пакета GS *de novo* assembler версии 2.5 (Roche, США). Для оценки перекрытия референсного генома *M. tuberculosis* H37Rv использовали программу 454 GS Reference Mapper (Roche, США). Для визуализации результатов выравнивания использовали программу Mauve версии 2.3.1. (<http://gel.ahabs.wisc.edu/mauve/>) (Darling *et al.*, 2004).

Поиск однонуклеотидных полиморфизмов проводили на высокопроизводительном вычислительном комплексе ФГБУ ФНКЦ ФХМ ФМБА России с использованием открытого программного обеспечения, а также разработанных внутри лаборатории скриптов на языке R. Картирование чтений с прибора (10 исследуемых образцов и 92 образца коллекции *key strains*, описанных в главе 2.3.1 Геномные последовательности) на референсный геном штамма *M. tuberculosis* H37Rv проводили с помощью инструмента bowtie2 (Langmead and Salzberg, 2012) (<http://bowtie-bio.sourceforge.net/bowtie2/index.shtml>). Анализ данных выравнивания проводили в приложении samtools 0.1.19 (Li *et al.*, 2009) (<http://samtools.sourceforge.net/>). Поиск однонуклеотидных полиморфизмов

осуществляли с помощью программы VarScan v 2.3.1 (Koboldt *et al.*, 2012) (<http://varscan.sourceforge.net>).

Выравнивание и поиск замен для 11 образцов, представленных полногеномными последовательностями, осуществляли с помощью программного пакета MUMMER 3 и входящих в него утилит nucmer, show-snps, show-coords (Delcher *et al.*, 2002) (<http://mummer.sourceforge.net/>).

Для оценки разнообразия образцов *M. tuberculosis* вычисляли матрицу расстояний Хэмминга на основании однонуклеотидных замен. Построение филогенетических деревьев проводили исходя из полученной матрицы с применением алгоритма Neighbor-joining. В качестве внешней группы использовали *M. canettii* (NC_015848.1).

Функциональный анализ найденных однонуклеотидных полиморфизмов проводили с использованием аннотации генома *M. tuberculosis* H37Rv из базы данных TubercuList v 2.6 (<http://tuberculist.epfl.ch>)

2.4 Протеомный анализ микобактерий

2.4.1 Экстракция белка из клеток микобактерий

Клетки микобактерий отмывали от питательной среды добавлением Tris-HCl, PBS + 2 % тритон X-100 (pH 7.5-8), затем центрифугировали 15 минут при 4500 g (Jungblut *et al.*, 1999). Супернатант сливали, из клеточного осадка экстрагировали белки.

К осадку клеточной культуры добавляли 10 мкл 100 мМ Tris-HCl pH 7.6 и 5 мкл лизоцима (1 мг/мл), инкубировали 30 мин при 4°C. Дополнительно добавляли 55 мкл 100 мМ Tris-HCl pH 7.6 и 1 мкл смеси ингибиторов протеаз (GE Healthcare, США). Клетки разрушали на Bead Beater (MPBio, США), с использованием 0.5 мм кремний-циркониевых шариков в соотношении 1V образца : 1 V шариков, 2 раза по 2 мин с 5-мин перерывом в ледяной бане. Затем к суспензии добавляли 20 мкл 20 % натрий додецилсульфат (SDS) (Sigma Aldrich, США), 10 мкл 1 М ДТТ (Sigma Aldrich, США); инкубировали 30 мин при 60°C и центрифугировали при

16000 g в течении 5 минут. Супернатант отбирали и использовали для дальнейшего анализа.

Концентрацию белка в белковых экстрактах определяли по методу Брэдфорда (Bradford, 1976) с использованием Bradford Protein Assay Kit (Bio Rad, США).

2.4.2 Денатурирующий электрофорез в полиакриламидном геле

Одномерный гель-электрофорез проводили согласно описанной методике по Лэммли (Laemmli, 1970). Готовили двухкомпонентный гель следующего состава: концентрирующий гель – 5 % смеси акриламид-бисакриламид (соотношение 29:1), 0.1 % SDS, 0.125 M Tris–HCl, pH 6.8; разделяющий гель – 10-7.5 % смеси акриламид-бисакриламид (соотношение 29:1), 0.1 % SDS, 0.375 M Tris–HCl, pH 8.9. Для полимеризации сначала добавляли N,N,N',N'-тетраметилэтилендиамин до концентрации 0.1 %, а затем персульфат аммония до 0.1 %. Верхний электродный Tris-глициновый буфер (150 mM, pH 8.3) содержал 0.1 % SDS.

К раствору белка (белковому экстракту) добавляли 2х кратный раствор Лэммли (1 : 1 V/V) и инкубировали при 95°C в течение 5 минут. Далее центрифугировали при 10000 g в течение 5 минут. Отбирали супернатант и использовали его для нанесения в гель.

Проводили электрофорез с использованием системы PROTEAN II (Bio-Rad, США) 20 мин при 20 mA до перехода красителя в разделяющий гель, после чего выставляли силу тока 40 mA и вели электрофорез до момента выхода краски из разделяющего геля.

Гели окрашивали раствором Кумасси G-250 (Invitrogen, США) в течение 1 - 2 часов и отмывали от краски раствором 10 % уксусной кислоты.

2.4.3 Триптическое расщепление белков

Белки фракционировали с помощью одномерного гель электрофореза (7.5 % ПААГ, 7-см гели) в денатурирующих условиях, как описано выше. Вырезали

нужную часть геля, содержащую белки, нарезали его на кусочки размером 1 x 1 мм. Отмывали гели от раствора Кумасси с помощью 50 % раствора ACN в 50 мМ NH₄HCO₃. Заливали гели раствором 10 мМ ДТТ в 100 мМ NH₄HCO₃ (так, чтобы покрывало все кусочки геля), инкубировали 30 мин при 56°C, раствор сливали. Заливали гели раствором 55 мМ йодацетамида в 100 мМ NH₄HCO₃, инкубировали 20 мин при комнатной температуре в темноте, раствор так же сливали. Высушивали отмывые гели, добавляя в них 100 % ACN, инкубировали 20 мин при комнатной температуре, сливали ACN, оставляли пробирки открытыми до полного высыхания гелей. Непосредственно для гидролиза белков добавляли к высушенным гелям 150 - 300 мкл раствора трипсина (20 нг/мкл) в 40 мМ NH₄HCO₃ и 10 % ACN, инкубировали 1 час на льду, а затем инкубировали ночь (16 часов) при 37°C.

После гидролиза полученные пептиды экстрагировали последовательно в три этапа:

1. Добавляли к образцу раствор 5 % муравьиной кислоты (2 объема, 300 мкл) и инкубировали при комнатной температуре в течение 30 минут. Раствор отбирали в отдельную пробирку;

2. Добавляли к образцу раствор 50 % ACN и 5 % FA (из расчета на 300 мкл раствора: 150 мкл 100 % ACN, 15 мкл 100 % FA и 135 мкл деионизированной воды), инкубировали при комнатной температуре в течение 30 минут. Раствор отбирали в отдельную пробирку.

3. Добавляли к образцу раствор 75 % ACN и 5 % FA (из расчета на 300 мкл раствора: 225 мкл 100 % ACN, 15 мкл 100 % FA и 70 мкл деионизированной воды), инкубировали при комнатной температуре в течение 30 минут. Раствор отбирали в отдельную пробирку.

Экстракты объединяли, полностью высушивали на вакуумной сушке SpeedVac (Thermo Scientific, США).

Далее проводили чистку пептидов с помощью колонок C18 Sep-Pak (Waters, США). Высушенные экстракты растворяли в 1 мл 3 % ACN и 0.01 % TFA и пропускали через колонки C18 Sep-Pak (Waters, США). Элюировали пептиды 700

мл раствора, содержащего: 75 % ACN и 0.01 % TFA. Полученные элюанты так же полностью высушивали на вакуумной сушке SpeedVac (Thermo Scientific, США) и растворяли в минимальном объеме (20 мкл) раствором, содержащим 5 % ACN и 0.1 % FA.

2.4.4 Масс-спектрометрический анализ

Анализ пептидных фракций проводили на масс-спектрометре TripleTOF 5600 с ионным источником NanoSpray III (ABSciex, Канада), оснащенный нанохроматографической системой разделения NanoLC Ultra 2D+ nano-HPLC (Eksigent, Сингапур). Система высокоэффективной жидкостной хроматографии (ВЭЖХ, англ. HPLC, High performance liquid chromatography) была настроена в режиме trap-elute. Для хроматографического разделения использовали следующие буферные растворы: буфер для загрузки образца на предколону - буфер А (98.9 % H₂O, 1 % MeOH, 0.1 % FA (в объемных долях)) и буфер В (99.9 % CH₃CN, 0.1 % FA (в объемных долях)). Образцы наносили на предколону Chrom XP C18 3 мм 120Å 350 мм*0.5 мм (Eksigent, Сингапур) при скорости потока 3 мкл/мин в течение 10 минут и элюировали через разделяющую колонку 3C18-CL-120 (3 мм 120Å) 75 мм*150 мм (Eksigent, Сингапур) при скорости потока 300 нл/мин в линейном повышающемся градиенте буфера В от 5 % до 40 % за 120 мин. Колонку и предколону регенерировали после каждого анализа промывкой 95 % буфера В в течение 7 мин и уравнивали 5 % буфером В в течение 25 минут. Между образцами для исключения вероятности неполной отмывки предыдущего образца через колонку и предколону пропускали 5 коротких (5 мин) градиентов от 5 % до 95 % буфера В, с 3 минутной промывкой 95 % буфером В в конце.

Перед каждым заколом образца закалывали 20 фмоль гидролизата β -галактозидазы в 0.1 % FA с использованием градиентной элюции в течение 15 минут (5 % - 25 % системы В) с выше указанными условиями регенерации колонки и предколони. Каждый цикл анализа включал в себя один ToF-MS спектр в диапазоне от 300 до 1250 Да и один MS/MS скан фрагментов распада

иона-предшественника с молекулярной массой 719.2 Да, согласно спецификациям настройки прибора.

Полученные таким образом спектры β -галактозидазы использовали для калибровки прибора и контроля стабильности, производительности и воспроизводимости работы масс-спектрометра.

Каждый целевой образец пептидной фракции микобактерий анализировали в IDA-режиме (IDA-mode) работы масс-спектрометра в двух технических повторах. IDA эксперимент состоял из одного обзорного MS скана: время накопления сигнала составляет 250 миллисекунд, отбор масс для фрагментации проводится в диапазоне 300 - 1250 m/z , за которым следует 50 зависимых MS/MS сканов. Ионы для MS/MS анализа отбираются на основе интенсивности измеряемого сигнала при превышении порога в 200 cps и при зарядовом состоянии от 2 до 5. Настройки MS/MS эксперимента следующие: разрешение UNIT (0.7 Да), диапазон масс 200 - 1800 m/z , время накопления сигнала 50 миллисекунд для каждого родительского иона. Для фрагментации используется метод диссоциации вызванной соударением (CID) с молекулами азота с энергией соударения возрастающей в диапазоне 25 - 55 Вольт в течение 50 миллисекунд. Проанализированный родительский ион после фрагментации и записи спектра отправляется в динамический лист ионов, не отбираемых для фрагментации на 15 секунд. Это позволяет отбирать родительский ион с вершины хроматографического пика для увеличения интенсивности MS/MS спектра, а также позволяет увеличить глубину инвентаризации белков за счет ограничения повторного анализа одних и тех же пептидов.

2.4.5 Валидация данных методом мониторинга множественных реакций

Мониторинг множественных реакций (MRM анализ) осуществляли на масс-спектрометре ABSciex QTRAP 4500 с ионным источником NanoSpray III (ABSciex, Канада), оснащенный нанохроматографической системой разделения Ekspert NanoLC 425 (Eksigent, Сингапур). ВЭЖХ система использовалась в конфигурации предколонка 94 – разделяющая колонка. Для хроматографического

разделения использовали следующие буферные растворы: буфер для загрузки образца на предколонку и буфер А - 98.9 % H₂O, 1 % метанол, 0.1 % муравьиная кислота (v/v); буфер В - 99.9 % ацетонитрил, 0.1 % муравьиная кислота (v/v). Образцы наносили на предколонку Chrom XP C18 3 120 Å 350 мм (Eksigent, США) при скорости потока 3 мкл/мин в течение 10 мин. Разделение осуществляли на колонке 3C18-CL-120 (3 мкм 120 Å) 75 мкм*150 мм (Eksigent, США) при скорости потока 300 нл/мин в линейном повышающемся градиенте буфера В от 5 до 50 % за 20 мин. Колонку и предколонку регенерировали после каждого анализа промывкой 95 % буфера В в течение 7 мин и уравнивали 5 % буфера В в течение 25 мин. Между образцами для исключения вероятности неполной отмычки предыдущего образца через колонку и предколонку пропускали 5 коротких (5 мин) градиентов от 5 до 95 % буфера В с 3 мин промывкой 95 % буфером В в конце. Масс-спектрометр использовался в режиме анализа положительно-заряженных ионов. Фрагментация методом столкновительной диссоциации с использованием азота. Параметры MRM анализа для анализируемого пептида оптимизировали в процессе прямого ввода раствора пептида в источник ионизации: родительский ион m/z 494.8 (z = +2), потенциал декластеризации 150 В, время накопления сигнала для каждого фрагмента 100 мс, m/z фрагментных ионов (энергия столкновительной диссоциации, В): 745.5 (20), 648.4 (22), 549.3 (20), 450.2 (19), 433.2 (29), 580.3 (26).

2.4.6 Биоинформационный анализ данных

Идентификацию пептидов проводили программными пакетами AB SCIEX ProteinPilot v 4.5 (ABSciex, Канада) и Mascot v 2.2.07 (Matrix Science, Великобритания). При идентификации использовали следующие параметры: точность 10 мд для родительских ионов, 0.5 Да для дочерних ионов, порог достоверности идентификации 95 %. Далее результаты идентификаций экспортировали в формате mascotXML для последующей обработки. Первичную квантификацию пептидов/белков проводили программным пакетом OpenMS v 1.10 (OpenMS, США), использующим mzML входной формат файлов,

разработанный международным консорциумом национальных исследовательских ассоциаций протеомики HUPO (Lange *et al.*, 2008). Файлы mzML получали с помощью ABSciexMSDataConverter, при конвертации из формата wiff с надстройкой «profile».

Для проведения количественного анализа файлы, полученные после масс-спектрометрического анализа (.raw файлы) импортировали в программу Progenesis LC-MS. Для этого проводили выравнивание спектров между экспериментами по времени удерживания, масс-спектрометрический эксперимент с наибольшим количеством спектров выбирался в качестве референсного. Экспортировали по 15 масс-спектров на одну область выхода ионов в Mascot Generic Format (.mgf).

Далее проводили идентификацию пептидов по экспортированным масс-спектрам в программе Mascot Search Engine v 2.5.1 с использованием следующих параметров: способ расщепления - фермент трипсин, модификации - пропионамидирование цистеина (C) и окислирование метионина (M), прибор — ESI-TRAP, ошибка m/z прекурсора 0.05 Да, ошибка m/z дочерних ионов 10 мд, допустим 1 пропущенный сайт ферментативного расщепления, процент ложных identifications с использованием decoy базы данных меньше 5 %. Полученные результаты идентификации пептидов в формате .xml импортировали в программу Progenesis LC-MS (Nonlinear Dynamics, США). Далее отбирали пептиды, значение score которых выше порогового (значение ложных identifications меньше 5 %). Дальнейшую идентификацию проводили согласно инструкции производителя программного пакета Progenesis LC-MS.

Анализ и обработку результатов MRM эксперимента проводили программным пакетом SkyLine v 2.6 (MacCossLab, США).

Локализацию белков *M. smegmatis* и *M. tuberculosis* получали из базы данных PSORTdb (<http://db.psort.org/browse/genome?id=13549>). В свою очередь, аннотацию по геной онтологии (Gene Ontology, GO) загружали из базы данных uniprot (<http://www.uniprot.org/>). Функциональную аннотацию *M. tuberculosis* H37Rv получали из базы данных TubercuList v 2.6 (<http://tuberculist.epfl.ch/>).

Для того чтобы охарактеризовать набор дифференциально экспрессирующихся генов, провели анализ обогащения GO. Для этого использовали пакет TopGO для языка программирования R. Для оценки достоверности обогащения использовали точный тест Фишера. Белки, относящиеся к обогащенным GO категориям, классифицировали с помощью PANTHER (<http://pantherdb.org/>).

3. РЕЗУЛЬТАТЫ

3.1 Формирование экспериментальных групп

Сформирована коллекция из 10 клинических изолятов *M. tuberculosis* кластера B0/W148, выделенных от пациентов СПбНИИФ.

Принадлежность штаммов *M. tuberculosis* к кластеру Beijing B0/W148 была определена методом ПЦР, согласно Мокроусову с соавт. (Mokrousov *et al.*, 2012).

На базе бактериологической лаборатории СПбНИИФ установлен профиль лекарственной устойчивости к препаратам первой (изониазид, рифампицин, этамбутол, стрептомицин) и второй (офлоксацин, канамицин, амикацин, капреомицин) линии противотуберкулезной терапии. Определение чувствительности проводили методом абсолютных концентраций. По данным бактериологического тестирования из 10 образцов *M. tuberculosis* один образец был идентифицирован как чувствительный, 3 – с множественной лекарственной устойчивостью, 6 – с широкой лекарственной устойчивостью (Приложение 1).

Дополнительно в исследование включен лабораторный штамм *M. tuberculosis* H37Rv. В качестве тест-культуры, для апробации методов выделения белка в работу был включен штамм *M. smegmatis* mc² 155.

3.2 Полногеномное секвенирование

Полногеномное секвенирование 10 образцов ДНК представителей кластера Beijing B0/W148 проведено с использованием высокопроизводительного геномного секвенатора GS FLX+ Series — XL+ (Roche, США) (проведено совместно с Карповой Ириной Юрьевной). Чтения с прибора депонированы в Sequence read archive (SRA) (проект № PRJNA181180).

Отчет для каждого исследуемого образца с указанием количества чтений и оснований, вошедших в сборку, количества контигов, а также усредненным покрытием представленных последовательностей был получен по результатам сборки *de novo*. Перекрытие референсного генома штамма *M. tuberculosis* H37Rv проанализировано с использованием 454 GS Reference Mapper (Roche, США). Для поиска однонуклетидных полиморфизмов использован программный пакет

VarScan v 2.3.1. Суммарная отчетная информация по исследуемым геномам представлена в Таблице 5.

Таблица 5.

Характеристика качества считывания последовательности геномной ДНК после обработки программой GS *de novo* assembler (Roche, США)

Образец	Кол. ридов	Кол. контигов >500 п.о. при сборке <i>de novo</i>	Покрытие	Перекрытие реф генома	Однонуклеотидные полиморфизмы относительно генома H37Rv		
					sSNPs	nsSNPs	Межгенные участки
SP1	193266	115	21.0	99.7	615	922	221
SP7	185956	110	20.0	99.7	610	921	220
SP10	196700	116	22.0	99.7	613	924	223
SP13	190826	108	21.0	99.7	620	926	219
SP21	189067	109	20.0	99.7	618	928	218
SP22	198731	111	22.0	99.7	616	929	219
SP23	207250	113	23.0	99.7	615	926	219
SP27	211027	109	23.0	99.7	613	923	220
SP45	192845	108	20.0	99.7	615	926	223
SP54	190822	112	20.0	99.7	622	926	224

3.3 Геномные особенности кластера BeijingB0/W148

Для позиционирования исследуемых штаммов на филогенетическом дереве, а также для определения геномных особенностей штаммов Beijing B0/W148 помимо 10 секвенированных образцов, исследованы 92 образца из коллекции *key strains*, а также 11 образцов генотипа Beijing, представленные в виде протяженных секвенсов. Филогенетический анализ проведен с использованием всех SNPs генома. Названия филогенетических линий были определены на основании исследования Nomolka с соавт. (Nomolka *et al.*, 2012). Полученное филогенетическое дерево представлено на Рисунке 5.

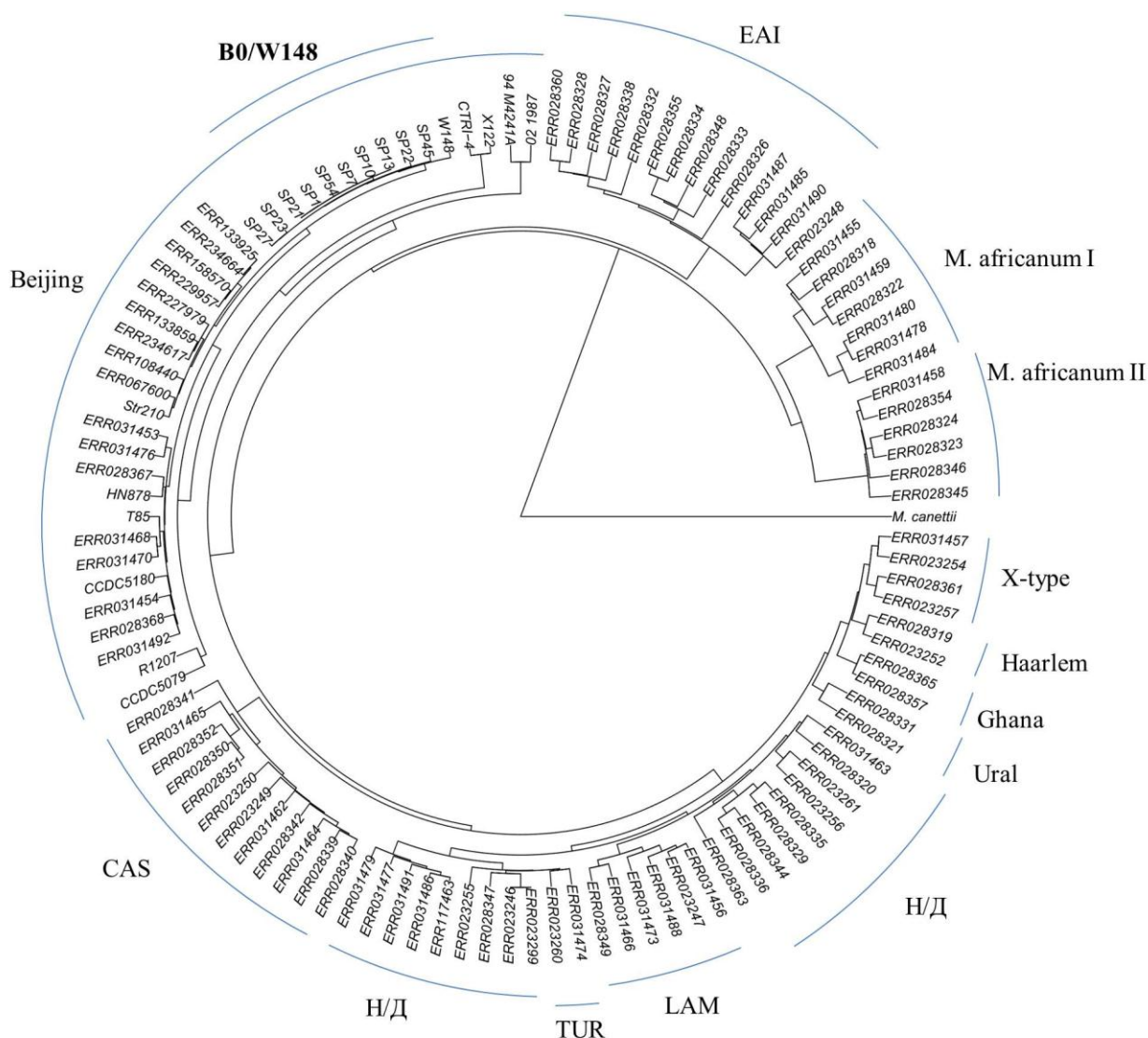


Рисунок 5. Сравнительный филогенетический анализ секвенированных штаммов. Алгоритм построения Neighbor-Joining. Филогенетические линии приведены в соответствии с Homolka с соавт (Homolka *et al.*, 2012). Н/Д – нет данных.

В ходе анализа установлено, что геномы 10 изучаемых образцов относятся к генетическому семейству Beijing и кластеризуются еще с одним представителем данного семейства: образцом W-148, представленным в базе данных GenBank в виде полного генома (CP012090.1). Таким образом, результаты филогенетического анализа показали полную сходимость с результатами первичного генетического анализа, а также указали на родство изучаемых штаммов со штаммом *M. tuberculosis* W-148.

Дополнительно, в ходе анализа литературных данных, для изучаемых образцов был проведен поиск основных молекулярных маркеров согласно

работам, описанным в главе 1.3 Популяционная структура *M. tuberculosis*. Результаты исследования представлены в Таблице 6.

Таблица 6.

Молекулярные маркеры исследуемых образцов

Образец	Сполиготип (Brudey <i>et al.</i> , 2006)	PGG (Sreevatsan <i>et al.</i> , 1997)	SCG (Filliol <i>et al.</i> , 2006)	SNP (Baker <i>et al.</i> , 2004)	SNP (Hershberg <i>et al.</i> , 2008)
SP1, SP7, SP10, SP13, SP21, SP22, SP23, SP27, SP45, SP54	Beijing	Главная генетическая группа 1	Кластер 2	Линия 1	Восточно- Азиатская линия

3.3.1 Кластер-специфические SNPs

В ходе сравнительного анализа данных полногеномного секвенирования для исследуемых штаммов идентифицировано 59 уникальных Beijing B0/W148 кластер-специфических полиморфизмов. Из них 8 полиморфизмов находились в межгенных участках генома, 24 мутации не приводили к изменению аминокислотной последовательности (синонимичные мутации), а 27 мутаций были несинонимичными. При этом 3 мутации из 27 приводили к образованию стоп-кодона. Найденные полиморфизмы, сгруппированные в соответствии с выполняемыми в клетке функциями, представлены в Приложении 2.

Дополнительно проведено сравнение полученных однонуклеотидных полиморфизмов с исследованием Merker с соавт. (Merker *et al.*, 2015). Коллеги в своей работе изучали особенности представителей генетического семейства Beijing на геномном уровне. Несколько образцов из исследования относилось к кластеру Beijing B0/W148. Стоит отметить, что найденные авторами 43 SNPs полностью совпали с полиморфизмами, идентифицированными в ходе настоящего исследования (Приложение 2).

3. 3.2 Структурная организация генома BeijingB0/W148

Установленное в ходе филогенетического анализа сходство изучаемых образцов кластера Beijing B0/W148 со штаммом W-148 позволило исследовать структурную организацию представителей кластера. Так, полногеномное выравнивание последовательностей штаммов H37Rv и W-148 с использованием программного пакета Mauve 2.3.1 показало наличие двух больших инверсий в геноме W-148 (Shitikov *et al.*, 2014). При анализе геном W-148 был разделен на пять коллинеарных блоков. В то время как в геноме W-148 блоки два и четыре были переставлены и инвертированы, блоки один, три и пять сохранили свою ориентацию (Рисунок 6).

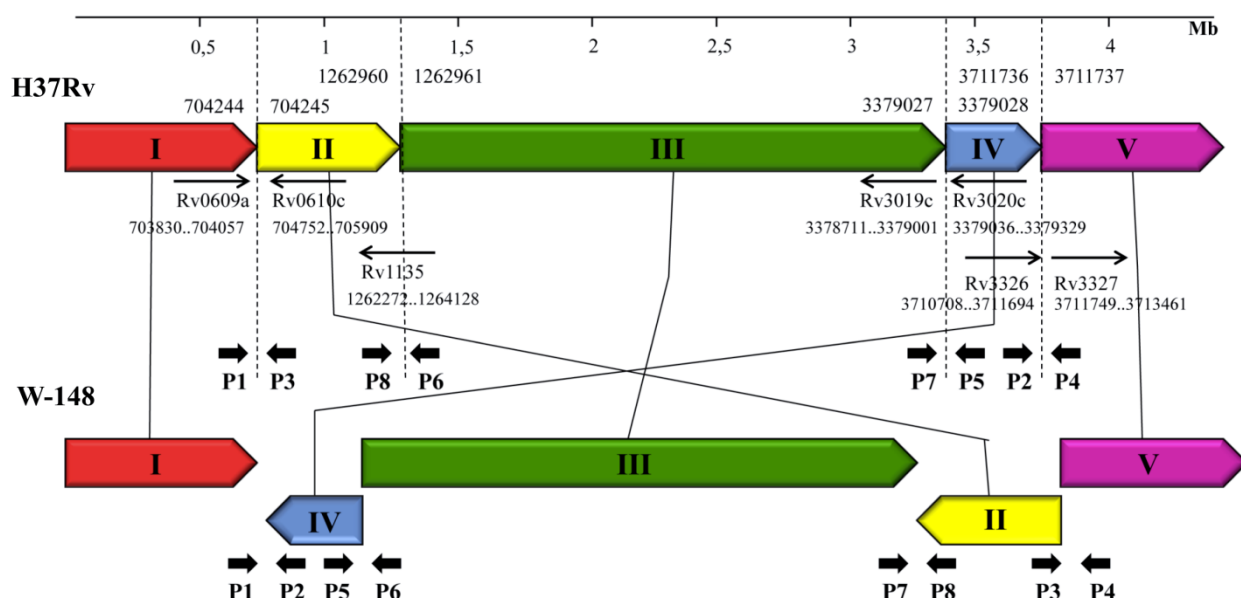


Рисунок 6. Схематическое изображение перестроек в геноме W-148 относительно H37Rv.

Коллинеарные блоки соответствуют таковым в Таблице 17. Блоки II и IV в геноме W-148 инвертированы и переставлены. Тонкими стрелками отмечены гены и их позиции в области инверсии. Жирные стрелки соответствуют праймерам (Таблица 7).

Наличие инверсии в исследуемых образцах было подтверждено результатами полимеразной цепной реакции. Для этого была разработана ПЦР-система, включающая восемь праймеров (Таблица 3 и Рисунок 7). Две пары праймеров (P1-P2 и P3-P4) были подобраны на границы внешней инвертированной области

(между блоками I и IV; II и V, соответственно) и две пары (P5-P6 и P7-P8) на границы внутренней области (между блоками IV и III; III и II, соответственно) (Рисунок 7). Для анализа геномной организации штаммов, не относящихся к кластеру Beijing B0/W148, использованы пары праймеров P1-P3, P2-P4, P6-P8 и P5-P7. Характерные ПЦР-продукты представлены в Таблице 7.

Таблица 7.

Пары праймеров, используемые для детекции рекомбинационных событий

Пары праймеров	Праймеры	Размер ПЦР-продукта		
		B0/W148 Beijing	Другие Beijing	Не Beijing
Внешняя инверсия				
1	P1-P2	1021	-	-
2	P3-P4	2215	-	-
3	P1-P3	-	320	320
4	P2-P4	-	2920	1845
Внутренняя инверсия				
5	P5-P6	1761	-	-
6	P7-P8	2527	-	-
7	P6-P8	-	1841	483
8	P5-P7	-	2447	1090

Апробация разработанной системы осуществлена на 10 образцах кластера Beijing B0/W148 и на референсном штамме H37Rv. По результатам исследования ПЦР-продукты для образцов кластера наблюдались только при использовании пар праймеров 1, 2, 5 и 6 (Таблица 7). При этом для образца H37Rv ни для одной из описанных пар праймеров ПЦР-продукт не был выявлен. В свою очередь использование пар праймеров 3, 4, 7 и 8 (Таблица 7) приводило к появлению ПЦР-продуктов для H37Rv, но не для образцов кластера. Специфичность ПЦР-продуктов доказана секвенированием по методу Сенгера.

Дополнительно наличие инверсий в геноме подтверждено в ходе рестрикционного анализа. Для исследования была выбрана рестриктаза *Mlu*I, так как ее сайты рестрикции во всех случаях находились в непосредственной близости с границами инвертированных регионов (Рисунок 7). Зонды для анализа по Саузерну получены в ходе ПЦР образца H37Rv с использованием Amersham ECL системы (GE Healthcare, США) и праймеров, указанных в Таблице 4.

Рестрикционный анализ проведен для образцов SP21 и H37Rv сотрудниками ФБУН СПб НИИ эпидемиологии и микробиологии имени Пастера. Результаты исследования приведены на Рисунке 8. В случае генома H37Rv гибридационные сигналы от комбинаций зондов А-В, С-Д, Е-Ф и G-Н совпали друг с другом и с *in silico* рассчитанными значениями. При этом для штамма Sp21 паттерны отличались. В свою очередь альтернативная комбинация зондов (А-Г, F-D, Е-С и В-Н), соответствующая произошедшим рекомбинационным событиям, показала совпадение паттернов. Таким образом, рестрикционный анализ также подтвердил наличие инверсий в штаммах кластера Beijing B0/W148.

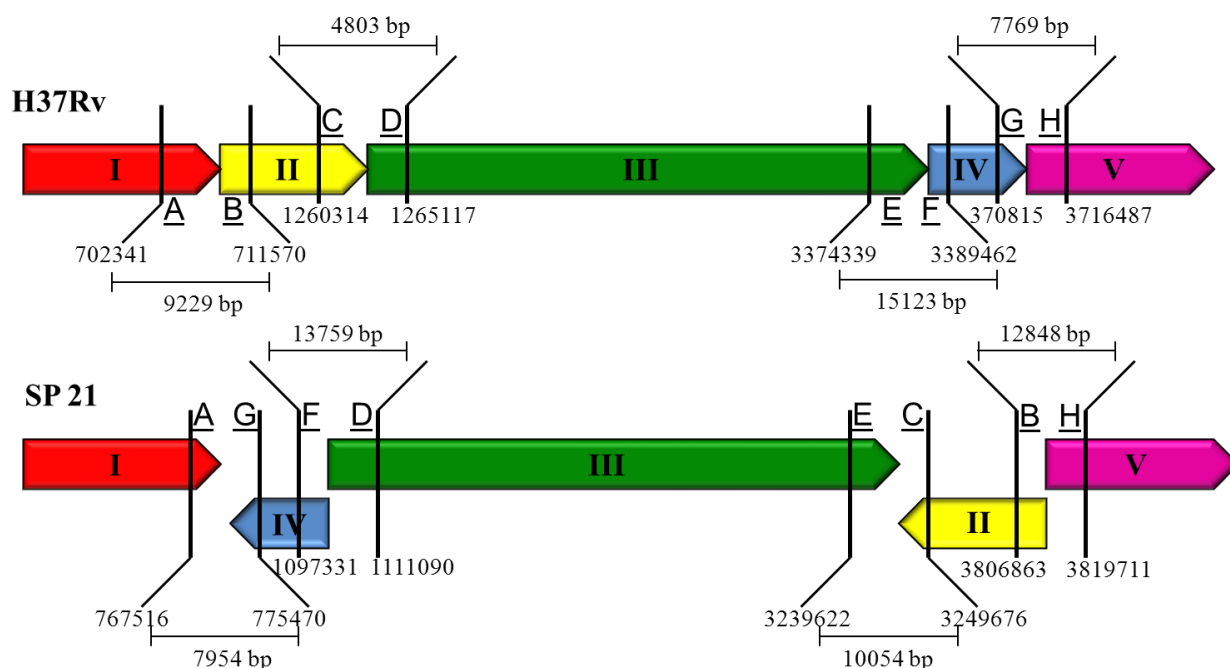


Рисунок 7. Схематическое представление стратегии RFLP анализа.

Вертикальными линиями отмечены сайты узнавания для рестриктазы *Mlu*I. Ожидаемые длины рестрикционных фрагментов представлены между вертикальными линиями. ДНК зонды для гибридизации отмечены буквами с А по Н.

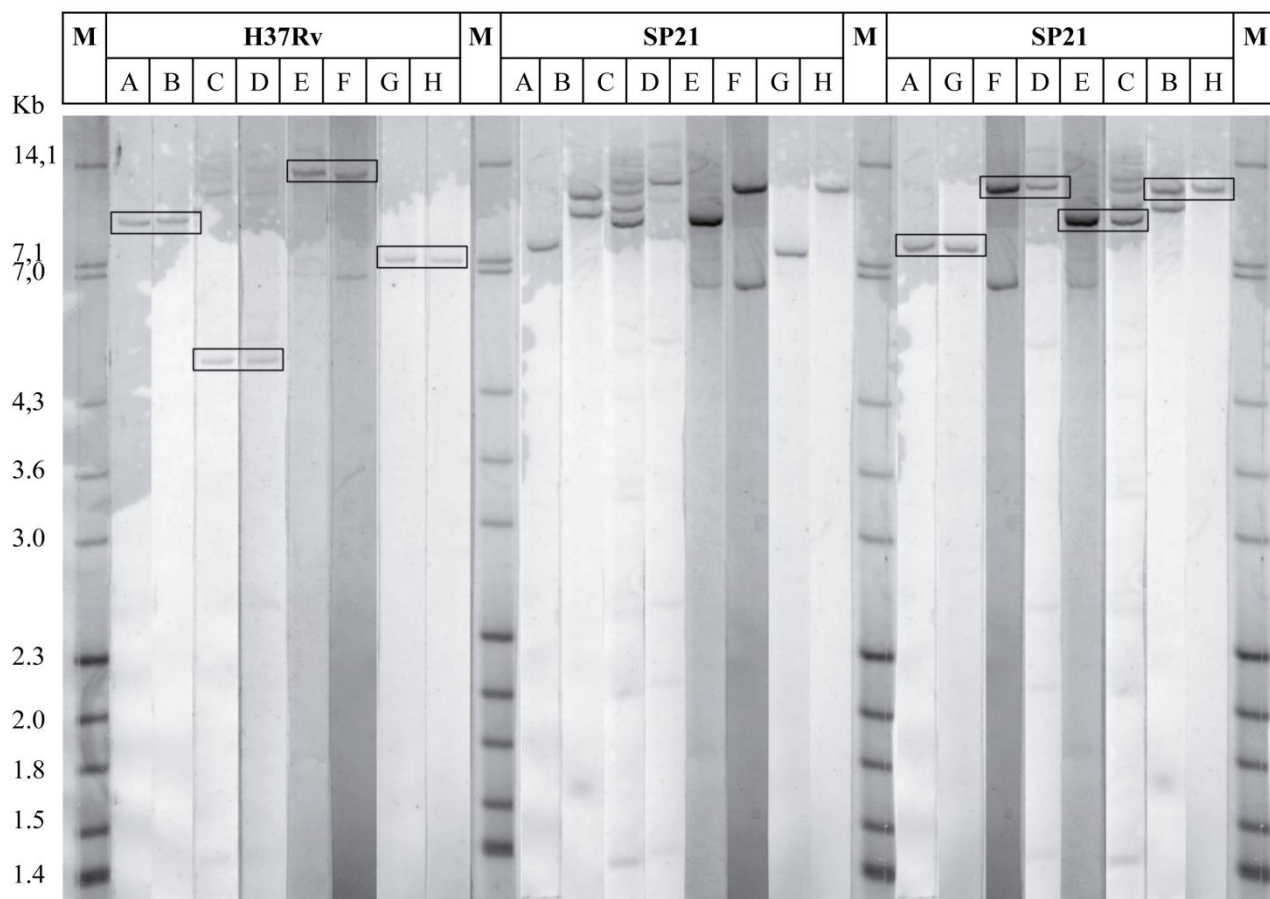


Рисунок 8. Результаты саузерн-блот гибридизации для штаммов H37Rv и SP21. (А) Гибридизационный паттерн штамма H37Rv. Порядок используемых зондов соответствует порядку встречаемости сайтов для генома H37Rv. (В) Гибридизационный паттерн штамма SP21. Порядок используемых зондов соответствует порядку встречаемости сайтов для генома H37Rv. (С) Гибридизационный паттерн штамма SP21. Порядок используемых зондов переорганизован в соответствии с произошедшими рекомбинационными событиями. М, маркерный штамм Mt14323 (Mycobacterial Reference Laboratory, National Public Health Institute (Turku, Финляндия)

Для подтверждения инверсий, а также для поиска новых инвертированных регионов проведено полногеномное секвенирование штамма SP21 с использованием библиотек парных фрагментов. По результатам анализа парных чтений из 391 контига (Таблица 5) было получено 12 скаффолдов общей длиной 4.45 м.п.о. Выравнивание геномов H37Rv, W-148 и SP21 показало наличие одинаковых крупномасштабных инверсии в штаммах SP21 и W-148 (Рисунок 6). При этом границы скаффолдов перекрывали инвертированные регионы. Других рекомбинационных событий выявлено не было.

Стоит отметить, что на границах инвертированных участков во всех случаях были выявлены фрагменты или полные копии инсерционной последовательности *IS6110*. Как было описано в главе 1.2.1 Молекулярно-генетический анализ, данный повторяющийся элемент служит для типирования микобактерий. В ходе исследования штаммов W-148 и SP21 определены сайты интеграции элемента. Результаты также были соотнесены с *IS6110* RFLP-патерном, характерным для штаммов Beijing B0/W148 (Рисунок 9).

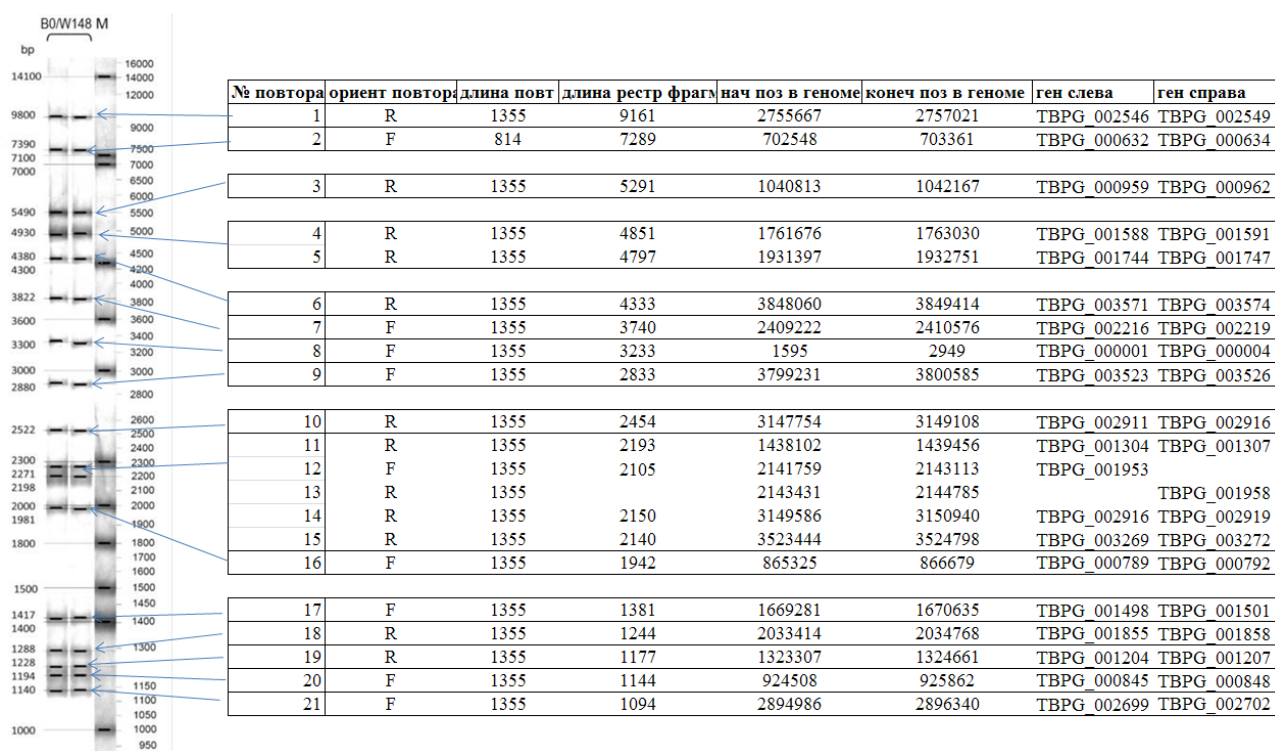


Рисунок 9. Позиционирование повторяющегося элемента *IS6110* в геноме штамма W-148.

Позиции повторяющегося элемента, расчёт длин рестрикционных фрагментов и названия генов приведены в соответствии с геномом штамма W-148 (CP012090v1).

3.4 Отработка протокола экстракции белка из клеток микобактерий

Лабораторный протокол экстракции белка из клеток микобактерий отработан на культуре *M. smegmatis* mc² 155. Культивирование проводили на твердой питательной среде Middlebrook 7H9. Перед экстракцией клетки были подвергнуты стадии отмывки от компонентов питательной среды. Исходя из

поставленных задач, нами протестированы три способа смыва клеток *M. smegmatis* со среды культивирования.

Первый способ согласно (Blackwood *et al.*, 2005) заключался в добавлении раствора 1 % Tween-80 (Sigma Aldrich, США) с последующим центрифугированием. После использования данного способа отмывки осадок клеток оставался вязким и не поддавался дальнейшему анализу. Было сделано предположение, что в отношении микобактерий недостаточно эффективен детергент Tween-80. В связи с этим был протестирован следующий способ согласно AfCS Solution Protocol (Protocol ID: PS00000603) (http://www.signaling-gateway.org/data/cgi-bin/ProtocolFile.cgi/afcs_PS00000603.pdf?pid=PS00000603) включающий 1 % тритон X-100 (Sigma Aldrich, США). Данный способ оказался значительно лучше, согласно визуальной оценке консистенции осадка. Однако некоторая вязкость сохранялась. В связи с этим было предпринято решение увеличить концентрацию детергента. Был опробован третий способ согласно (Jungblut *et al.*, 1999) включающий раствор 2 % тритон X-100 и дополнительно Tris-HCl. Таким образом, основываясь на визуальной оценке цвета и консистенции получаемого бактериального осадка, наиболее эффективным был признан последний метод. Данный метод использовался в последующих экспериментах.

Для отработки протокола экстракции белка суспензию клеток в отмывочном буфере нормировали по мутности. В каждый эксперимент было взято приблизительно 10^{10} клеток. Апробированы пять различных методов экстракции общего белка из микобактериальных клеток.

Каждый метод включал в себя уникальный состав лизирующего буфера и непосредственно способ разрушения клеток. Так основным компонентом лизирующего буфера в первом методе был 0.75 % раствор RapiGest SF (Waters, США). В лизирующих буферах остальных четырех методов использованы различные концентрации ДТТ (Sigma Aldrich, США), SDS (Sigma Aldrich, США) в сочетании с ингибиторами протеаз (GE Healthcare, США) и/или нуклеазами (ДНКаза + РНКаза, Promega, США). Клетки были разрушены на Bead Beater

(MPBio, США) и ультразвуковом гомогенизаторе (И100-6/840, ИНЛАБ, Россия). В ходе экспериментов для обеих систем апробированы различные условия разрушения клеток.

Концентрация белка в белковых экстрактах была определена по методу Брэдфорда (Bradford, 1976) с использованием Bradford Protein Assay Kit (Bio Rad, США). Результаты представлены в Приложении 3 на Рисунке 3.1

Согласно полученным данным, наибольший выход белка зафиксирован при использовании для экстракции реагента RapiGest SF (Waters, США) (Приложение 3 Метод 1) и лизирующего буфера содержащего лизоцим и ингибиторы протеаз в Tris-HCl (Приложение 3 Метод 3). Не смотря на то, что, во втором случае экстрагируется достоверно больше общего белка, чем в первом ($p = 0.0006$) для дальнейшего анализа были использованы белковые экстракты, полученные с использованием обоих методов.

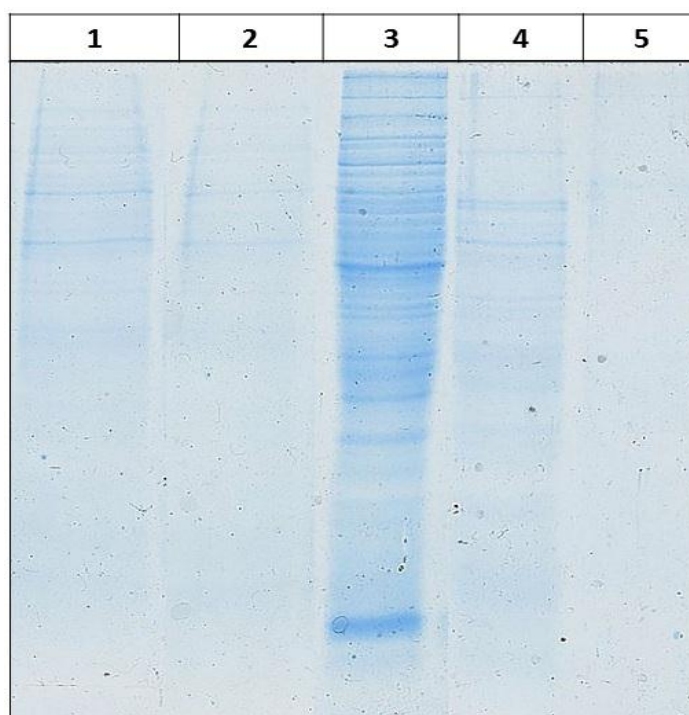


Рисунок 10. Электрофореграмма белковых профилей *M. smegmatis*.

Бэнды соответствуют различным методам экстракции белка из клеток согласно Приложению 3. 1 – экстракция белков методом 2А, 2 – экстракция белков методом 2Б, 3 – экстракция белков методом 2В, 4 – экстракция белков методом 2Г, 5 – экстракция белков методом 2Д.

Так же было апробировано несколько методов триптического расщепления белков. Так, для белков, экстрагированных с использованием реагента RapiGest SF (Waters, США), гидролиз трипсином был проведен согласно описанной методике (Vowinckel *et al.*, 2013). Белки, экстрагированные вторым способом, подвергнуты гидролизу трипсином в растворе по методу FASP (http://www.broadinstitute.org/annotation/tb_sysbio/protocols/proteomics/Proteomics_Protocol_5.pdf; Vowinckel *et al.*, 2013) и в геле согласно описанной ранее методике (Shevchenko *et al.*, 2006). Краткое описание протоколов каждого метода представлено в Приложении 3.

Оценка эффективности триптического расщепления белков проведена методом масс-спектрометрического анализа с подобранными оптимальными условиями. При расщеплении белков трипсином в растворе по методу FASP было идентифицировано 446 белков *M. smegmatis*. Гидролиз белков трипсином в растворе для реагента RapiGest SF позволил выявить 1017 белков *M. smegmatis*. При этом методом гидролиза трипсином в геле удалось идентифицировать 3937 белков. Результаты идентификации белков представлены в Таблице 8. Таким образом, наиболее эффективным был признан метод гидролиза белков в геле, который и был использован для дальнейшего анализа. Не смотря на то, что в данном методе после чистки на колонках количество идентифицированных белков было чуть ниже, в дальнейшие эксперименты так же была включена стадия очистки пептидов. Связано это с тем, что в образцах после трипсинолиза остается достаточное большое количество солей, которые оседают в масс-спектрометрических колонках, при постановке большого количества образцов.

Результаты протеомного анализа *M. smegmatis* для различных тестируемых методов триптического расщепления белков.

Образец/ пробоподготовка	Количество идентифицированных белков			
	1 повтор	2 повтор	3 повтор	ВСЕГО
<i>M. smegmatis</i> (10 ¹⁰) / Трипсинолиз в геле (до чистки на колонках)	3896	3844	3928	3937
<i>M. smegmatis</i> (10 ¹⁰) / Трипсинолиз в геле (после чистки на колонках)	3804	3795	3787	3893
<i>M. smegmatis</i> (10 ¹⁰) / FASP (после чистки на колонках)	439	441	430	446
<i>M. smegmatis</i> (10 ¹⁰) / RapiGest (до чистки на колонках)	978	992	986	1008
<i>M. smegmatis</i> (10 ¹⁰) / RapiGest (после чистки на колонках)	1002	992	988	1017

Таким образом, в ходе проведенного протеомного анализа штамма *M. smegmatis* mc² 155 по отработанному протоколу идентифицировано 17221 уникальных пептидов (с долей ложноположительных идентификаций (FDR) < 5 % по Mascot v 2.2.07), относящихся к 3937 белкам (критерий отбора: не менее двух уникальных пептидов на белок). В качестве референсной базы идентификаций пептидов и белков использован *M. smegmatis* mc² 155 (uid57701). В общей сложности, идентифицированные белки (n = 3937) составляют около 60 % от 6716 белков, аннотированных по геномным данным.

С использованием базы данных PSORTdb (<http://db.psort.org/>) установлена локализация идентифицированных нами белков *M. smegmatis* mc² 155, полученные результаты представлены на Рисунке 11. В целом, не наблюдалось преимущественной потери белков какой-либо локализации, что демонстрирует адекватность примененной схемы экстракции общего белкового пула микобактерий.

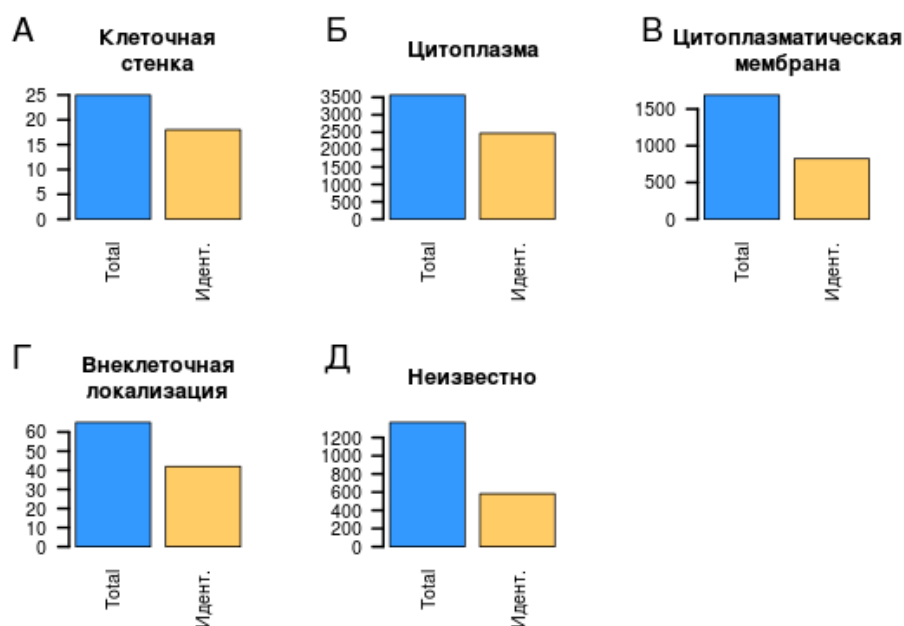


Рисунок 11. Локализация идентифицированных белков *M. smegmatis* mc² 155 согласно базе данных PSORTdb (<http://db.psорт.org/>).

Для каждого типа локализации (А-Д) синим цветом отображено количество белков в геноме с данной локализацией (Total), желтым цветом количество идентифицированных белков с данной локализацией.

3.5 Протеомный анализ *Mycobacterium tuberculosis*

Для протеомного анализа отобрано 7 штаммов *M. tuberculosis* кластера В0/W148 для которых были известны геномные последовательности. Дополнительно в анализ был включен штамм H37Rv. Штаммы *M. tuberculosis* культивированы на питательной среде Middlebrook 7H11, далее проведена экстракция и триптическое расщепление белка согласно разработанному протоколу. Характерный профиль экстрагируемых белков, представлен на электрофореграммах (Рисунок 12). Масс-спектрометрический анализ белковых профилей был проведен по отработанной методике совместно с Иваном Бутенко и Ильей Алтуховым. В результате LC-MS/MS экспериментов накоплено 1098994 MS/MS спектров, которые соответствуют 366621 уникальным пептидам (FDR < 1 %).

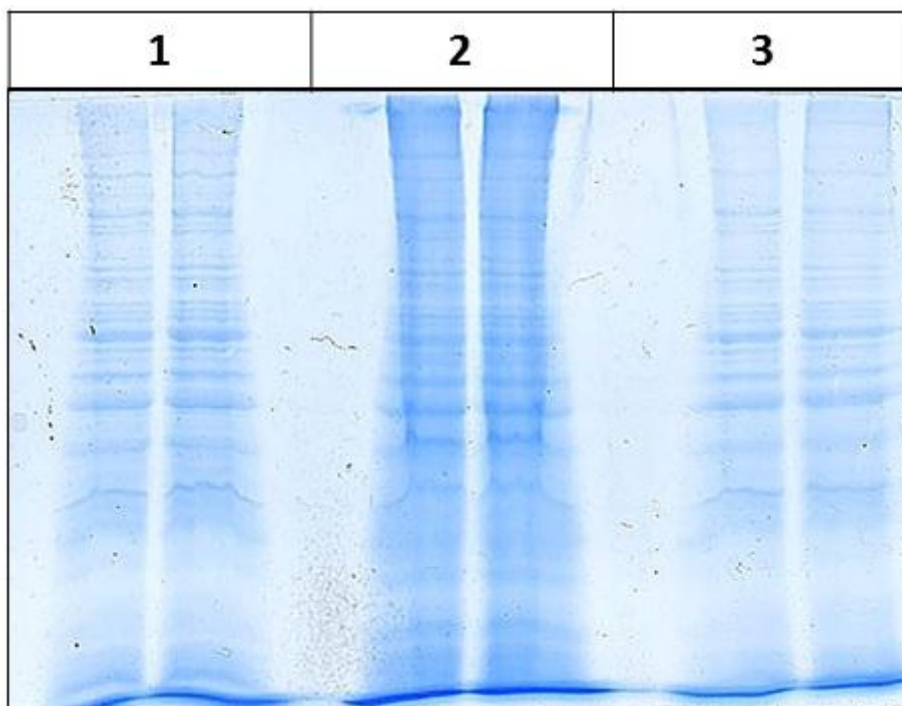


Рисунок 12. Электрофореграмма продуктов белков, выделенных из клеток *M. tuberculosis* кластера B0/W148

1 – Белки, экстрагированные из штамма Sp7 (два биологических повтора); 2 – белки, экстрагированные из штамма Sp10 (два биологических повтора); 3 – белки, экстрагированные из штамма Sp45 (два биологических повтора);

Для штамма H37Rv всего идентифицировано 1560 белков, удовлетворяющих требованию минимум два пептида в двух биологических повторах. Для семи штаммов кластера Beijing B0/W148 идентифицировано 1868 белков, из которых 1176 были найдены во всех штаммах, что составляет > 60 % от всех идентифицированных белков. Списки идентифицированных белков представлены в Приложениях 4 и 5.

Анализ протеомных данных был проведен с использованием аннотации генома *M. tuberculosis* H37Rv из базы данных TubercuList v 2.6 (<http://tuberculist.epfl.ch>) и PSORTdb v 3.0 (<http://db.psort.org/>) (Рисунок 13). Полученные данные свидетельствуют о равномерном распределении белков по всем категориям, на основании чего можно сделать заключение, что применяемые методы протеомного анализа оптимальны как для лабораторного штамма H37Rv, так и клинических изолятов кластера Beijing B0/W148.

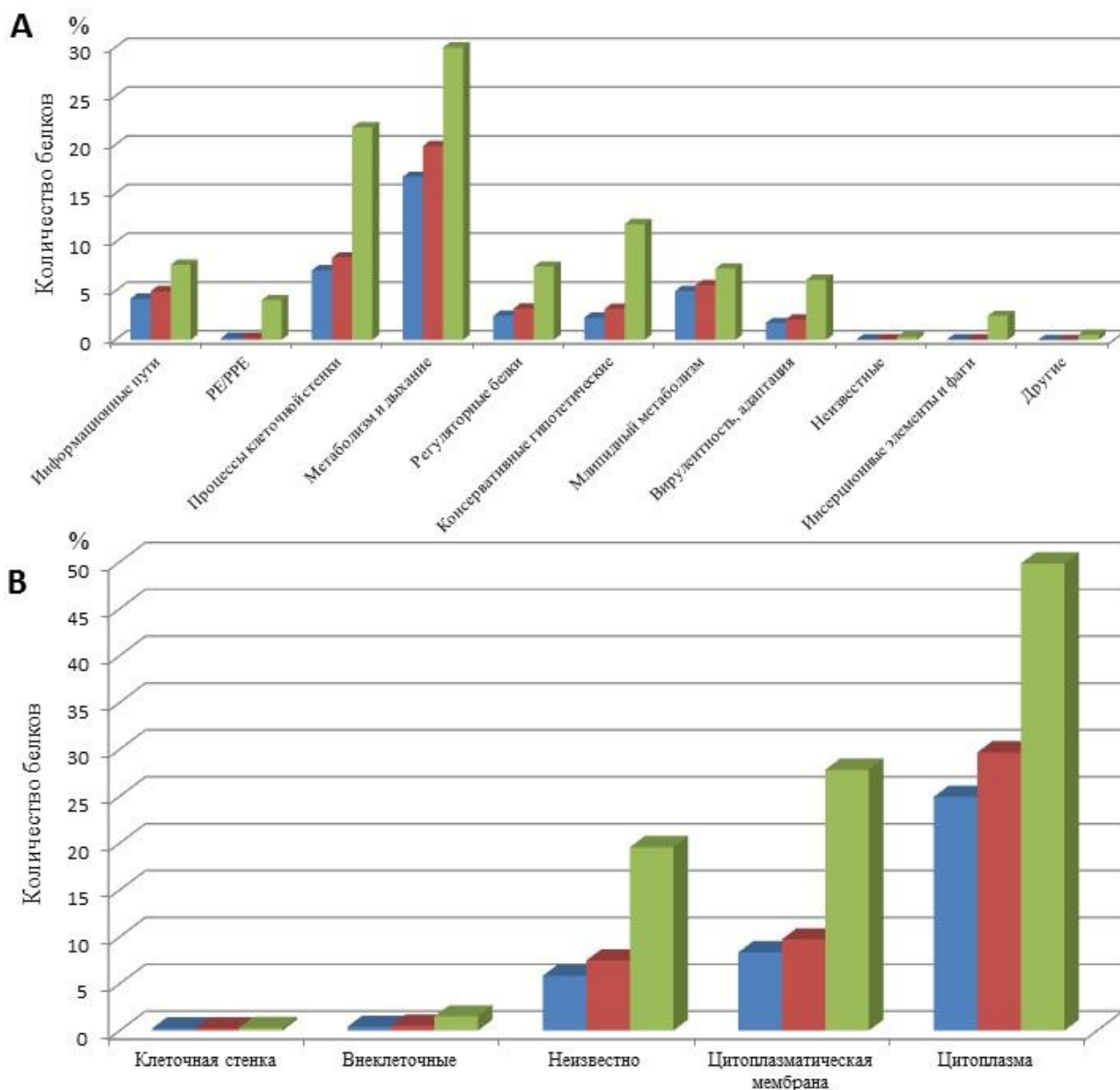


Рисунок 13. Функциональное распределение идентифицированных белков. Белки, идентифицированные для штамма H37Rv (синий столбец) для штаммов кластера Beijing B0/W148 (красный столбец) и все аннотированные согласно базе данных (зеленый столбец) были распределены:

А. по функциональным категориям в соответствии с TubercuList v 2.6 (<http://tuberculist.epfl.ch>);

Б. в соответствии с их локализацией по данным PSORTdb v 3.0 (<http://db.psort.org/>)

Дополнительно были определены мажорные и минорные белки, характерные для штаммов *M. tuberculosis* кластера Beijing B0/W148. Определение белков проведено на основании расчета emPAI (exponential modified protein abundance index), так как показано, что этот индекс пропорционален

концентрации белка (Ishihama *et al.*, 2005). Белки, индекс emPAI которых находился выше 95 % перцентиля, были отнесены к мажорным, а индекс которых ниже 5 % перцентиля к минорным. В частности, было проанализировано распределение данных белков по COG категориям. Результаты распределения белков представлены в Таблице 9.

Таблица 9.

Мажорные и минорные белки, характерные для штаммов *M. tuberculosis* кластера Beijing B0/W148

Мажорные белки			Минорные белки		
Белок	COG-категория	Локализация	Белок	COG-категория	Локализация
iron-regulated peptidyl-prolyl cis-trans isomerase PpiA	COG0652O	Цитоплазма	5-оксипролинэза OplA	COG0146EQ	Цитоплазма
short-chain type dehydrogenase/reductase	COG1028IQR	Цитоплазма	transcriptional regulator	COG3903R	Цитоплазма
3-оксоацил-ACP редуктаза FabG	COG1028IQR	Цитоплазма	exonuclease V subunit beta RecB	COG1074L	Цитоплазма
fructose-bisphosphate aldolase	COG0191G	Цитоплазма	exonuclease V subunit gamma RecC	COG1330L	Цитоплазматическая мембрана
molecular chaperone GroEL	COG0459O	Цитоплазма	hypothetical protein	COG4867R	Цитоплазма
two component sensory transduction protein RegX	COG0745TK	Цитоплазма	FO synthase	COG1060HR	Цитоплазма
hypothetical protein	COG3324R	Неизвестно	ATP-dependent RNA helicase DeaD	COG0513LKJ	Цитоплазма
hypothetical protein	-	Цитоплазматическая мембрана	bifunctional sulfate adenylyltransferase subunit 1 CysN/adenylylsulfate kinase CysC	COG2895P	Цитоплазма
enoyl-CoA hydratase EchA3	COG1024I	Цитоплазма	methyltransferase	COG0144J	Цитоплазма
(3R)-hydroxyacyl-ACP dehydratase subunit HadA	COG2030I	Цитоплазма	macrolide ABC transporter ATP-binding protein	COG0488R	Цитоплазма

elongation factor Tu	COG0050J	Цитоплазма	DNA repair protein RecN	COG0497L	Цитоплазма
hypothetical protein	-	Неизвестно	carboxylase	COG0439I	Цитоплазма
hypothetical protein	-	Неизвестно	hypothetical protein	COG0451MG	Цитоплазма
fatty oxidation protein FadB	COG1250I	Цитоплазма	hypothetical protein	COG0451MG	Цитоплазматическая мембрана
phosphoserine aminotransferase	COG1932HE	Неизвестно	cobalamin biosynthesis protein CobN	COG1429H	Цитоплазма
pterin-4-alpha-carbinolamine dehydratase	COG0521H	Цитоплазма	ATP-dependent DNA helicase HelY	COG4581L	Цитоплазма
beta-ketoacyl CoA thiolase FadA	COG0183I	Цитоплазма	glutamate-ammonia-ligase adenyltransferase	COG1391OT	Цитоплазма
5-methyltetrahydropteroylglutamate--homocysteine methyltransferase	COG0620E	Цитоплазма	glycerol-3-phosphate dehydrogenase	COG0578C	Цитоплазма
ESAT-6 like protein EsxL	-	Неизвестно	DNA primase	COG0358L	Цитоплазма
malate dehydrogenase	COG0039C	Неизвестно	hypothetical protein	COG1948L	Цитоплазма
ATP synthase subunit beta	COG0055C	Цитоплазматическая мембрана	DNA translocase FtsK	COG1674D	Цитоплазматическая мембрана
S-adenosylmethionine synthetase	COG0192H	Цитоплазма	magnesium chelatase	COG1239H	Цитоплазма
iron-regulated aconitate hydratase	COG1048C	Цитоплазма	phthiocerol synthesis polyketide synthase type I PpsD	COG3321Q	Цитоплазматическая мембрана
oxidoreductase	COG0300R	Цитоплазма	ATP-dependent DNA helicase RecG	COG1200LK	Цитоплазма
iron-regulated universal stress protein	COG0589T	Неизвестно	1,4-alpha-glucan branching protein	COG1543S	Цитоплазма
glycogen accumulation regulator GarA	COG1716T	Цитоплазма	hypothetical protein	COG4914R	Неизвестно
inosine-5'-monophosphate dehydrogenase	COG0516F	Цитоплазма	electron transfer protein FdxB	COG3239I	Цитоплазматическая мембрана

L-lactate dehydrogenase	COG1304C	Цитоплазматическая мембрана	arabinosyltransferase C	-	Цитоплазматическая мембрана
alpha-crystallin	COG0071O	Клеточная стенка	excinuclease ABC subunit UvrB	-	Цитоплазма
hypothetical protein	COG1881R	Цитоплазма	helicase HelZ	-	Цитоплазма
hypothetical protein	COG2867I	Цитоплазма	bifunctional penicillin-insensitive transglycosylase/penicillin-sensitive transpeptidase	COG0744M	Цитоплазматическая мембрана
adenosine kinase	COG0524G	Неизвестно	ATP-dependent DNA helicase UvrD	COG0210L	Цитоплазма
glutamine synthetase	COG0174E	Цитоплазма	trehalose-6-phosphate phosphatase OtsB	COG1554G	Цитоплазма
meromycolate extension acyl carrier protein	COG0236IQ	Цитоплазма	bifunctional penicillin-insensitive transglycosylase/penicillin-sensitive transpeptidase	COG0744M	Внеклеточная локализация
3-oxoacyl-ACP synthase 1	COG0304IQ	Цитоплазматическая мембрана			
3-oxoacyl-ACP synthase 2	COG0304IQ	Цитоплазма			
nucleoside diphosphate kinase	COG0105F	Цитоплазма			
hypothetical protein	-	Цитоплазма			
oxidoreductase	COG0656R	Цитоплазма			
electron transfer flavoprotein subunit alpha	COG2025C	Цитоплазма			
electron transfer flavoprotein subunit beta	COG2086C	Цитоплазма			
NADP-dependent alcohol dehydrogenase	COG1064R	Цитоплазма			
iron-regulated short-chain dehydrogenase/reductase	COG1028IQR	Цитоплазма			
chaperonin GroEL	COG0459O	Цитоплазма			
chaperonin GroES	COG0234O	Цитоплазма			

DNA-directed RNA polymerase subunit alpha	COG0202K	Цитоплазма			
hypothetical protein	-	Неизвестно			
enoyl-CoA hydratase EchA21	COG1024I	Неизвестно			
long-chain-fatty-acid--AMP ligase FadD32	COG0318IQ	Цитоплазма			
bacterioferritin BfrB	COG1528P	Цитоплазма			
ESAT-6-like protein EsxB	COG4842S	Внеклеточная локализация			
fatty-acid--CoA ligase FadD31	-	Цитоплазма			
phosphate ABC transporter substrate-binding lipoprotein PstS	COG0226P	Неизвестно			
transcriptional regulator MoxR1	COG0714R	Цитоплазма			
hypothetical protein	COG3360S	Цитоплазма			
ESAT-6 like protein EsxN	-	Цитоплазма			
D-3-phosphoglycerate dehydrogenase	COG0111HE	Цитоплазма			
hypothetical protein	-	Неизвестно			
hypothetical protein	-	Цитоплазма			
ESAT-6 protein EsxA	COG4842S	Внеклеточная локализация			
methyltransferase	-				

Для проведения сравнительного анализа белковых профилей штаммов кластера B0/W148 и H37Rv, полученные данные были расширены списком белков, идентифицированных Schubert с соавт. для H37Rv (Schubert *et al.*, 2013).

Среди спектра идентифицированных белков были обнаружены группы, которые встречались исключительно среди штаммов кластера Beijing B0/W148. Данная группа включала 17 белков (Таблица 10). Так же была обнаружена группа

из 57 белков, встречающихся исключительно в штаммах H37R (Таблица 11). Среди уникальных белков, характерных для клинических изолятов ($n = 17$), стоит отметить белки TatA и TrcR. Белок TatA входит в состав эффлюксной системы Tat, необходимой для секреции белков разного размера и формы. В исследованиях Saint-Joanis с соавт. (Saint-Joanis *et al.*, 2006) показана роль белков данной системы в вирулентности и формировании лекарственной устойчивости. Белок TrcR является одним из членов двухкомпонентной системы TrcRS, задействованной в регуляции синтеза белка Rv1057. Функция последнего недостаточно изучена, однако предполагается, что он задействован в поддержании стабильности клеточной стенки.

Таблица 10.

Белки, идентифицированные исключительно в штаммах
кластера Beijing B0/W148

Белок	Синоним	Аннотация согласно TubercuList	Функциональная категория	Локализация	COG категория	Геномные особенности у Beijing B0/W148
Rv0125	pepA	Probable serine protease PepA (serine proteinase) (MTB32A)	intermediary metabolism and respiration	неизвестна	COG0265 O	без изменений
Rv0386	-	Probable transcriptional regulatory protein (probably LuxR/UhpA-family)	Regulatory proteins	цитоплазма	COG3903 R	sSNP
Rv0509	hemA	Probable glutamyl-tRNA reductase HemA (GLUTR)	intermediary metabolism and respiration	цитоплазма	COG0373 H	без изменений
Rv0945	-	Probable short-chain type dehydrogenase/reductase	intermediary metabolism and respiration	цитоплазма	COG0300 R	Arg180Gly
Rv0976c	-	related to synthesis and oxidation of fatty acids	Lipid metabolism	цитоплазма	-	Upstream region (1090189insG)
Rv1319c	-	Possible adenylate cyclase (ATP pyrophosphate-lyase) (adenylyl cyclase)	intermediary metabolism and respiration	цитоплазматическая мембрана	COG2114 T	Gly106Asp
Rv1397c	vapC10	Possible toxin VapC10	virulence, detoxification, adaptation	неизвестна	COG1487 R	без изменений
Rv1712	cmk	Cytidylate kinase Cmk (CMP kinase) (cytidine monophosphate kinase) (ck)	intermediary metabolism and respiration	цитоплазма	COG0283F	без изменений
Rv2277c	-	Possible glycerolphosphodiesterase	intermediary metabolism and respiration	неизвестна	COG0584 C	IS6110 рядом со старто-кодоном

Rv2351c	plcA	Membrane-associated phospholipase C 1 PlcA (MTP40 antigen)	intermediary metabolism and respiration	цитоплазматическая мембрана	COG3511M	Thr446Ala
Rv2475c	-	Predicted thioesterase	Intermediary metabolism and respiration	цитоплазма	COG0824R	IS6110 is near the start of the gene
Rv2826c	-	Hypothetical protein	conserved hypotheticals	цитоплазма	COG2253S	без изменений
Rv2974c	-	Predicted kinase related to dihydroxyacetone kinase	Intermediary metabolism and respiration	цитоплазма	COG1461R	Upstream region (3101512delCG)
Rv3228	-	Predicted GTPases	Intermediary metabolism and respiration	цитоплазма	COG1162R	sSNP
Rv3323c	moaX	Probable MoaD-MoaE fusion protein MoaX	intermediary metabolism and respiration	цитоплазма	COG0314H	IS6110 рядом со старт-кодоном
Rv3592	-	Possible heme degrading protein MhuD	intermediary metabolism and respiration	цитоплазма	COG2329R	без изменений
Rv3753c	-	Conserved protein	conserved hypotheticals	цитоплазма	-	без изменений

Таблица 11.

Белки, идентифицированные исключительно в штамме H37Rv

Белок	Сининим	Аннотация согласно TubercuList	Функциональная категория	Локализация	COG	Геномные особенности у Beijing B0/W148
Rv0034	-	Ketosteroid isomerase-related protein	Intermediary metabolism and respiration	цитоплазма	COG3631R	sSNP
Rv0045c	-	Possible hydrolase	lipid metabolism	цитоплазма	COG0596R	Val194Ile
Rv0072	-	Probable glutamine-transport transmembrane protein ABC transporter	cell wall and cell processes	цитоплазматическая мембрана	COG4591M	Часть от RD105
Rv0081	-	Probable transcriptional regulatory protein	Regulatory proteins	цитоплазма	COG0640K	без изменений
Rv0143c	-	Probable conserved transmembrane protein	cell wall and cell processes	цитоплазматическая мембрана	COG0038P	без изменений
Rv0311	-	Unknown protein	conserved hypotheticals	цитоплазма	-	Glu119Asp
Rv0495c	-	DNA-associated process	Regulatory proteins	цитоплазма	-	без изменений
Rv0825c	-	Transcriptional regulator	Regulatory proteins	неизвестна	COG1309K	Phe68Cys
Rv0888	-	Probable exported protein	cell wall and cell processes	неизвестна	COG3021S	Rv0888: 987586 insGG
Rv0987	-	Probable adhesion component transport transmembrane protein ABC transporter	cell wall and cell processes	цитоплазматическая мембрана	COG4591M	Ala372Val; Phe717Val
Rv0988	-	Possible conserved exported protein	cell wall and cell processes	цитоплазматическая мембрана	COG5621R	Leu191Ala
Rv1576c	-	Probable PhiRv1 phage	Insertion seqs	цитоплазматическая мембрана	COG4653R	Часть от

		protein	and phages	ая мембрана		RD149
Rv1586c	-	Probable PhiRv1 integrase	Insertion seqs and phages	цитоплазма	COG1961L	Часть от RD149
Rv1591	-	Probable transmembrane protein	cell wall and cell processes	цитоплазматическая мембрана	-	без изменений
Rv1729c	-	Possible S-adenosylmethionine-dependent methyltransferase	lipid metabolism	цитоплазма	COG3315Q	без изменений
Rv1762c	-	Unknown protein	conserved hypotheticals	цитоплазма	COG0393S	Часть от RD152
Rv1835c	-	Predicted acyl esterases	Intermediary metabolism and respiration	цитоплазма	COG2936R	Cys179Arg
Rv1978	-	SAM-dependent methyltransferases	Intermediary metabolism and respiration	цитоплазматическая мембрана	COG2226H	Ser226Thr
Rv2003c	-	SAM-dependent methyltransferases	Intermediary metabolism and respiration	неизвестна	COG2226H	без изменений
Rv2050	-	Conserved protein	conserved hypotheticals	цитоплазма	-	без изменений
Rv2203	-	Possible conserved membrane protein	cell wall and cell processes	неизвестна	-	Upstream region (G2468180A)
Rv2240c	-	Unknown protein	conserved hypotheticals	неизвестна	-	без изменений
Rv2263	-	Possible oxidoreductase	intermediary metabolism and respiration	цитоплазма	COG1028IQR	Часть от RD181
Rv2616	-	Conserved protein	conserved hypotheticals	неизвестна	COG4762S	без изменений
Rv2690c	-	Probable conserved integral membrane alanine and valine and leucine rich protein	cell wall and cell processes	цитоплазматическая мембрана	COG0531E	без изменений
Rv2818c	-	Hypothetical protein	unknown	цитоплазма	-	part of RD207
Rv2952	-	Possible methyltransferase (methylase)	intermediary metabolism and respiration	цитоплазма	COG2226H	Gly176Arg
Rv3272	-	Predicted acyl-CoA transferases/carnitine dehydratase	Intermediary metabolism and respiration	цитоплазма	COG1804C	Val181Met
Rv3482c	-	Probable conserved membrane protein	cell wall and cell processes	неизвестна	-	без изменений
Rv3626c	-	proteolysis/nucleotide biosynthesis	Intermediary metabolism and respiration	цитоплазма	COG5282S	без изменений
Rv3694c	-	Possible conserved transmembrane protein	cell wall and cell processes	цитоплазматическая мембрана	COG1300S	Upstream region (T4137219G)
Rv3802c	-	Probable conserved membrane protein	cell wall and cell processes	цитоплазма	-	Upstream region (4264382ins C)
Rv3805c	aftB	Possible arabinofuranosyltransferase AftB	cell wall and cell processes	цитоплазматическая мембрана	-	Asp397Gly
Rv0768	aldA	Probable aldehyde dehydrogenase NAD	intermediary metabolism and	цитоплазма	COG1012C	Leu102Arg

		dependent AldA (aldehyde dehydrogenase [NAD+])	respiration			
Rv1317c	alkA	Probable bifunctional regulatory protein and DNA repair enzyme AlkA (regulatory protein of adaptative response) (methylphosphotriester-DNA--protein-cysteine S-methyltransferase)	information pathways	цитоплазма	COG0122L	Ile12Val
Rv1589	bioB	Probable biotin synthetase BioB	intermediary metabolism and respiration	цитоплазма	COG0502H	без изменений
Rv0908	ctpE	Probable metal cation transporter ATPase P-type CtpE	cell wall and cell processes	цитоплазматическая мембрана	COG0474P	без изменений
Rv1997	ctpF	Probable metal cation transporter P-type ATPase A CtpF	cell wall and cell processes	цитоплазматическая мембрана	COG0474P	Rv1997: 2241032 delG
Rv3895c	eccB2	ESX conserved component EccB2. ESX-2 type VII secretion system protein. Probable membrane protein.	cell wall and cell processes	неизвестна	-	без изменений
Rv3885c	eccE2	ESX conserved component EccE2. ESX-2 type VII secretion system protein. Possible membrane protein.	cell wall and cell processes	неизвестна	-	sSNP
Rv0035	fadD34	Probable fatty-acid-CoA ligase FadD34 (fatty-acid-CoA synthetase) (fatty-acid-CoA synthase)	lipid metabolism	цитоплазматическая мембрана	COG0318IQ	Ser16Trp; Thr26Ser
Rv3797	fadE35	Probable acyl-CoA dehydrogenase FadE35	lipid metabolism	цитоплазма	COG1960I	Ser328Pro; sSNP
Rv0989c	grcC2	Probable polyprenyl-diphosphate synthase GrcC2 (polyprenyl pyrophosphate synthetase)	intermediary metabolism and respiration	цитоплазма	COG0142H	Ile321Val
Rv0087	hycE	Possible formate hydrogenase HycE (FHL)	intermediary metabolism and respiration	цитоплазма	COG3261C	без изменений
Rv1400c	lipI	Probable lipase LipH	intermediary metabolism and respiration	цитоплазма	COG0657I	Met42Val
Rv0604	lpqO	Probable conserved lipoprotein LpqO	cell wall and cell processes	цитоплазматическая мембрана	-	без изменений
Rv2380c	mbtE	Peptide synthetase MbtE (peptide synthase)	lipid metabolism	цитоплазма	-	sSNP
Rv1183	mmpL10	Probable conserved transmembrane transport protein MmpL10	cell wall and cell processes	цитоплазма	COG2409R	Thr408Arg
Rv0844c	narL	Possible nitrate/nitrite response transcriptional regulatory protein NarL	Regulatory proteins	цитоплазма	COG2197TK	Gly169Arg
Rv3158	nuoN	Probable NADH dehydrogenase I (chain N) NuoN (NADH-ubiquinone oxidoreductase chain N)	intermediary metabolism and respiration	цитоплазматическая мембрана	COG1007C	без изменений
Rv1182	papA3	Probable conserved polyketide synthase associated protein PapA3	lipid metabolism	цитоплазма	-	без изменений
Rv1386	PE15	PE family protein PE15	PE/PPE	цитоплазматическая мембрана	-	без изменений

				ая мембрана		изменений
Rv1822	pgsA2	Probable CDP-diacylglycerol--glycerol-3-phosphate 3-phosphatidyltransferase PgsA2 (PGP synthase) (phosphatidylglycerophosphate synthase) (3-phosphatidyl-1'-glycerol-3'phosphate synthase)	lipid metabolism	цитоплазматическая мембрана	COG0558I	без изменений
Rv0405	pks6	Probable membrane bound polyketide synthase Pks6	lipid metabolism	цитоплазматическая мембрана	COG3321Q	Arg1402Pro
Rv0182c	sigG	Probable alternative RNA polymerase sigma factor SigG (RNA polymerase ECF type sigma factor)	information pathways	цитоплазма	COG1595K	Asp332Tyr; sSNP
Rv0417	thiG	Probable thiamin biosynthesis protein ThiG (thiazole biosynthesis protein)	intermediary metabolism and respiration	цитоплазма	COG2022H	Ser75Cys
Rv2103c	vapC37	Possible toxin VapC37. Contains PIN domain.	virulence, detoxification, adaptation	неизвестна	COG1848R	Thr16Ala

Для определения представленности белков, идентифицированных в штаммах Beijing B0/W148 и H37Rv, использована программа Progenesis LC-MS. Данный анализ был ограничен 1016 белками, присутствующими во всех штаммах кластера Beijing B0/W148 и H37Rv. В общей сложности выявлено 192 белка, представленность которых достоверно различалась ($p < 0.05$). Из них, 24 белка были высокопредставленные в штаммах кластера Beijing B0/W148 и 168 были низкопредставленными.

Далее были сопоставлены результаты масс-спектрометрического и количественного протеомного анализов для выявления возможного общего эффекта дифференциальных белков на функционирование микробной клетки. Белки, идентифицированные только в штаммах кластера, были отнесены в группу высокопредставленных. В свою очередь белки, отсутствующие в штаммах кластера были отнесены в группу низкопредставленных белков. Таким образом, первая группа включала 41 белок, а вторая 225 белков (Приложение 6).

Все дифференциальные белки были классифицированы согласно GO категориям «Биологические процессы» (BP), «Клеточные компоненты» (CC) и «Молекулярные функции» (MF) с использованием системы классификации PANTHER (Рисунок 14). В категории «Молекулярные функции» большинство как

высоко- так и низкопредставленных белков обогащали подгруппы «оксидоредуктазная активность» и «трансферазная активность» в GO группе «каталитическая активность», а так же группу «транспортная активность». Для категории «Биологические процессы» наиболее представленные белки обогащали группу «основные липидные метаболические процессы», в то время как наименее представленные белки равномерно распределились по нескольким группам (Таблица 12, Рисунок 14).

Таблица 12.

Дифференциальные белки, входящие в наиболее значимые обогащенные категории GO анализа.

Белок	Синоним	Представленность в штаммах кластера Beijing B0/W148 относительно H37Rv	Функциональная категория	Продукт	log ₂ fold	p-value
Транспортная активность (GO:0005215)						
Rv2187	fadD15	Высокопредставлен	Lipid metabolism	long-chain-fatty-acid--CoA ligase FadD15	0.64	0.049358 831
Rv2971	-	Высокопредставлен	Intermediary metabolism and respiration	oxidoreductase	0.5	0.041065 417
Rv0073	-	Низкопредставлен	cell wall and cell processes	glutamine ABC transporter ATP-binding protein		
Rv0143c	-	Низкопредставлен	cell wall and cell processes	transmembrane protein		
Rv0933	pstB	Низкопредставлен	cell wall and cell processes	phosphate ABC transporter ATP-binding protein PstB	-2.41	4.65402E-07
Rv0934	pstS1	Низкопредставлен	cell wall and cell processes	phosphate ABC transporter substrate-binding lipoprotein PstS	-1.12	0.001558 62
Rv1281c	oppD	Низкопредставлен	cell wall and cell processes	oligopeptide ABC transporter ATP-binding protein OppD	-0.6	0.045017 563
Rv1308	atpA	Низкопредставлен	Intermediary metabolism and respiration	ATP synthase subunit alpha	-0.46	0.001330 189
Rv1310	atpD	Низкопредставлен	Intermediary metabolism and respiration	ATP synthase subunit beta	-0.66	0.000609 44
Rv2564	glnQ	Низкопредставлен	cell wall and cell	glutamine ABC	-0.76	0.000361

			processes	transporter ATP-binding protein		77
Rv2690c	-	Низкопредставлен	cell wall and cell processes	integral membrane protein		
Липидные метаболические процессы (GO:0006629)						
Rv2187	fadD15	Высокопредставлен	Lipid metabolism	long-chain- fatty-acid--CoA ligase FadD15	0.64	0.049358 831
Rv2277c	-	Высокопредставлен	Intermediary metabolism and respiration	glycerolphospho diesterase		
Rv3107c	agpS	Высокопредставлен	Lipid metabolism	alkyldihydroxya cetonephosphate synthase	3.5	0.022585 521
Rv3570c	hsaA	Высокопредставлен	Intermediary metabolism and respiration	flavin- dependent monoxygenase oxygenase subunit HsaA	1	0.043815 093
Rv0334	rmlA	Низкопредставлен	Intermediary metabolism and respiration	glucose-1- phosphate thymidylyltransf erage	-1.41	7.01184E- 05
Rv0405	pks6	Низкопредставлен	Lipid metabolism	membrane bound polyketide synthase		
Rv0468	fadB2	Низкопредставлен	Lipid metabolism	3- hydroxybutyryl- CoA dehydrogenase	-1.33	2.15157E- 07
Rv0860	fadB	Низкопредставлен	Lipid metabolism	fatty oxidation protein FadB	-0.47	0.008655 458
Rv0905	echA6	Низкопредставлен	Lipid metabolism	enoyl-CoA hydratase EchA6	-0.96	0.002614 821
Rv1185c	fadD21	Низкопредставлен	Lipid metabolism	fatty-acid--CoA ligase FadD21	-1.35	0.004873 026
Rv3230c	-	Низкопредставлен	Intermediary metabolism and respiration	stearoyl-CoA 9- desaturase electron transfer partner	-1.38	0.032334 42
Rv3791	dprE2	Низкопредставлен	Lipid metabolism	decaprenylphos phoryl-D-2-keto erythropentose reductase	-0.85	0.003881 878
Rv3825c	pks2	Низкопредставлен	Lipid metabolism	phthioceranic/hy droxyphtiocera nic acid synthase	-1	0.031606 82

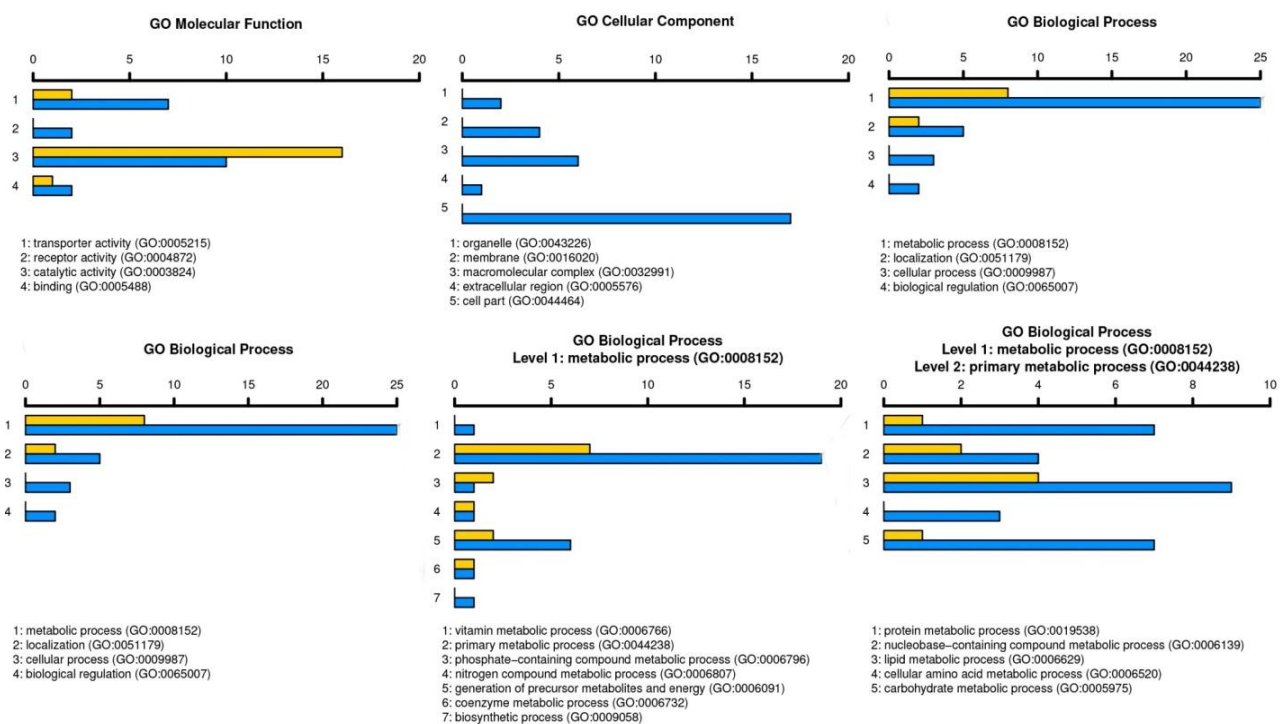


Рисунок 14. Распределение белков по категориям согласно GO анализу. Белки высокопредставленные (желтый) и низкопредставленные (синий) в штаммах кластера Beijing W0/W148 относительно H37Rv

Зная, что прокариотические гены и опероны имеют регулонную организацию, было выдвинуто предположение, что белки, гены которых контролируются одним транскрипционным фактором (ТФ), будут объединяться в «однонаправленно изменяющиеся» группы. Для проверки данного предположения проанализированы 266 дифференциальных белков с использованием генной регуляторной сети, включающей 65 транскрипционных факторов и регулируемые ими 431 ген (Galagan *et al.*, 2013). Полученные результаты визуализированы с помощью программы Cytoscape v 2.8.3. Гены пяти белков из группы наиболее представленных и 38 из группы наименее представленных белков были идентифицированы на данной сети (Таблица 13, Рисунок 15).

Гены дифференциальных белков, наложенные на регуляторную сеть

Ген	Синоним	ТФ для данного гена	Продукт гена	Функциональная категория	log ₂ fold	p-value
Гены, соответствующие высокопредставленным белкам в штаммах кластера Beijing B0/W148 по сравнению с H37Rv						
Rv0509	-	Rv1353c	glutamyl-tRNA reductase	Intermediary metabolism and respiration		
Rv1464	-	Rv1460	cysteine desulfurase	Intermediary metabolism and respiration	1.56	0.0489072
Rv3094c	<i>fadE2</i>	Rv0494; Rv3095	hypothetical protein	Intermediary metabolism and respiration	9.61	0.0386941
Rv3494c	<i>yrbE1A</i>	Rv0022c	Mce family protein Mce4	virulence, detoxification, adaptation	1.46	0.0259927
Rv3570c	<i>mce1A</i>	Rv0678; Rv1353c; Rv3574	flavin-dependent monooxygenase oxygenase subunit HsaA	Intermediary metabolism and respiration	1	0.0438151
Гены, соответствующие низкопредставленным белкам в штаммах кластера Beijing B0/W148 по сравнению с H37Rv						
Rv0034	-	Rv3249c	hypothetical protein	Intermediary metabolism and respiration		
Rv0035	-	Rv3249c	fatty-acid-CoA ligase FadD34	Lipid metabolism		
Rv0081	<i>bioF2</i>	Rv0081	HTH-type transcriptional regulator	Regulatory proteins		
Rv0087	<i>acpA</i>	Rv0081	formate hydrogenase HycE	Intermediary metabolism and respiration		
Rv0101	-	Rv0047c; Rv2069; Rv0324	peptide synthetase Nrp	Lipid metabolism		
Rv0154c	<i>fadD34</i>	Rv1423	acyl-CoA dehydrogenase FadE2	Lipid metabolism	-0.59	0.0008057
Rv0169	-	Rv0023; Rv0081; Rv0757; Rv1049; Rv3416	Mce family protein Mce1A	virulence, detoxification, adaptation	-1.44	5.763E-06
Rv0170	-	Rv0023; Rv0081; Rv1049; Rv3416	Mce family protein Mce1B	virulence, detoxification, adaptation	-1.38	0.0017153
Rv0172	<i>sdaA</i>	Rv0023; Rv0081; Rv1049	Mce family protein Mce1D	virulence, detoxification, adaptation	-1.38	0.0001941
Rv0173	<i>glyA2</i>	Rv1049	Mce family lipoprotein LprK	cell wall and cell processes	-1.29	7.012E-05
Rv0174	-	Rv0023; Rv0081; Rv1049	Mce family protein Mce1F	virulence, detoxification, adaptation	-0.84	7.012E-05
Rv0176	-	Rv1049	Mce associated	cell wall and cell	-1.81	0.0175553

			transmembrane protein	processes		
Rv0241c	-	Rv0238; Rv1049	3-hydroxyacyl-thioester dehydratase HtdX	Intermediary metabolism and respiration	-0.86	0.0386941
Rv0243	-	Rv0238	acetyl-CoA acetyltransferase FadA	Lipid metabolism	-0.91	0.0166606
Rv0675	-	Rv0674	enoyl-CoA hydratase EchA5	Lipid metabolism	-1.1	0.0007597
Rv0768	-	Rv0576; Rv1255c	aldehyde dehydrogenase AldA	Intermediary metabolism and respiration		
Rv0824c	<i>hycD</i>	Rv0472c	acyl-ACP desaturase DesA	Lipid metabolism	-1.98	5.763E-06
Rv0989c	<i>hycP</i>	Rv0767c	polyprenyl-diphosphate synthase GrcC	Intermediary metabolism and respiration		
Rv1094	<i>hycQ</i>	Rv0472c	acyl-ACP desaturase DesA	Lipid metabolism	-1.91	5.196E-09
Rv1192	<i>hycE</i>	Rv2034; Rv0081	hypothetical protein	cell wall and cell processes	-1.91	0.0006327
Rv1386	-	Rv1033c; Rv2359	PE family protein PE15	PE/PPE		
Rv1856c	<i>fcoT</i>	Rv1353c	oxidoreductase	Intermediary metabolism and respiration	-1.06	0.037951
Rv1997	<i>ctpF</i>	Rv3133c	cation transporter ATPase F	cell wall and cell processes		
Rv2004c	-	Rv3133c	hypothetical protein	Regulatory proteins	-1.45	0.0004367
Rv2005c	<i>nrp</i>	Rv3133c	universal stress protein	virulence, detoxification, adaptation	-1.24	0.023446
Rv2029c	<i>gmhB</i>	Rv3133c	6-phosphofruktokinase PfkB	Intermediary metabolism and respiration	-3	0.0006327
Rv2048c	<i>hddA</i>	Rv0767c	polyketide synthase	Lipid metabolism	-1.38	0.0393413
Rv2103c	<i>msrA</i>	Rv1990c	ribonuclease VapC37	virulence, detoxification, adaptation		
Rv2410c	-	Rv0022c	hypothetical protein	Intermediary metabolism and respiration	-1.42	0.0082996
Rv2623	-	Rv3133c	universal stress protein	virulence, detoxification, adaptation	-1.26	0.0157466
Rv2626c	<i>ptbB</i>	Rv3133c	hypoxic response protein	Regulatory proteins	-1.56	0.0216908
Rv3130c	<i>pntAa</i>	Rv3133c	diacylglycerol O-acyltransferase	Lipid metabolism	-2.29	0.0001161
Rv3131	<i>pntAb</i>	Rv3133c	NAD(P)H nitroreductase	Intermediary metabolism and respiration	-1.69	0.0026414
Rv3400	<i>pntB</i>	Rv0135c	hydrolase	Intermediary metabolism and respiration	-0.93	0.0451627
Rv3509c	<i>yrbE1B</i>	Rv0324	acetohydroxyacid synthase large subunit	Intermediary metabolism and respiration	-0.66	0.0080181
Rv3602c	<i>mce1B</i>	Rv1353c	pantothenate synthetase	Intermediary metabolism and respiration	-1.25	0.0014439
Rv3825c	<i>mce1C</i>	Rv0757	phthioceranic/hydroxyphthioceranic acid	Lipid metabolism	-1	0.0316068

			synthase			
--	--	--	----------	--	--	--

Далее внимание было сфокусировано на ко-регулирующихся группах генов, находящихся под регуляцией одного ТФ. Таким образом, выделено 5 ТФ, регулирующих гены 24 дифференциальных белков. Из них, ТФ Rv3133с, Rv1049 и Rv0081 обладали наиболее обширными связями. ТФ Rv3133с, основной компонент DosR регулона, был ассоциирован с восьмью генами наименее представленных белков: Rv1997 [EC: 3.6.3.-], Rv2004с, Rv2005с, Rv2029с [EC: 2.7.1.11], Rv2623, Rv2626с, Rv3130с и Rv3131 (Рисунок 15 А). В связи с этим решено было определить профили представленности всех белков, кодируемых генами DosR регулона. Согласно литературным данным регулон состоит из 53 генов (Schubert *et al.*, 2013). В данной работе идентифицировано 33 из 53 белков (62 %). Одиннадцать из них были наименее представленными в штаммах кластера Beijing B0/W148. Большинство из этих белков участвуют в транспорте и деградации липидов и вероятно так же ассоциированы с ассимиляцией экзогенных липидов из мембран клеток-хозяев. Интересно отметить, что представленность самого белка DosR не изменяется в штаммах кластера Beijing B0/W148 по отношению к H37Rv.

Другая группа включала ТФ Rv1049 и Rv0081, регулирующие синтез восьми генов, кодирующих белки, наименее представленных в штаммах кластера Beijing B0/W148 (Рисунок 15 Б). Из них шесть белков (Rv0169, Rv0170, Rv0172, Rv0173, Rv0174, и Rv0176) вовлечены в мембранный транспорт фосфолипидов и относятся к семейству ABC транспортеров. Стоит отметить, что Rv0081 недавно был описан как регуляторный «hub» (Brugarolas *et al.*, 2012; Galagan *et al.*, 2013). Rv0081 является членом ArsR/SmtB семейства металл-зависимых регуляторов транскрипции и непосредственно регулируется ответом регуляторов DosR/DevR и MprAB (He *et al.*, 2011). Согласно полученным данным, MprA (Rv0981) так же является низкопредставленным белком в штаммах кластера Beijing B0/W148.

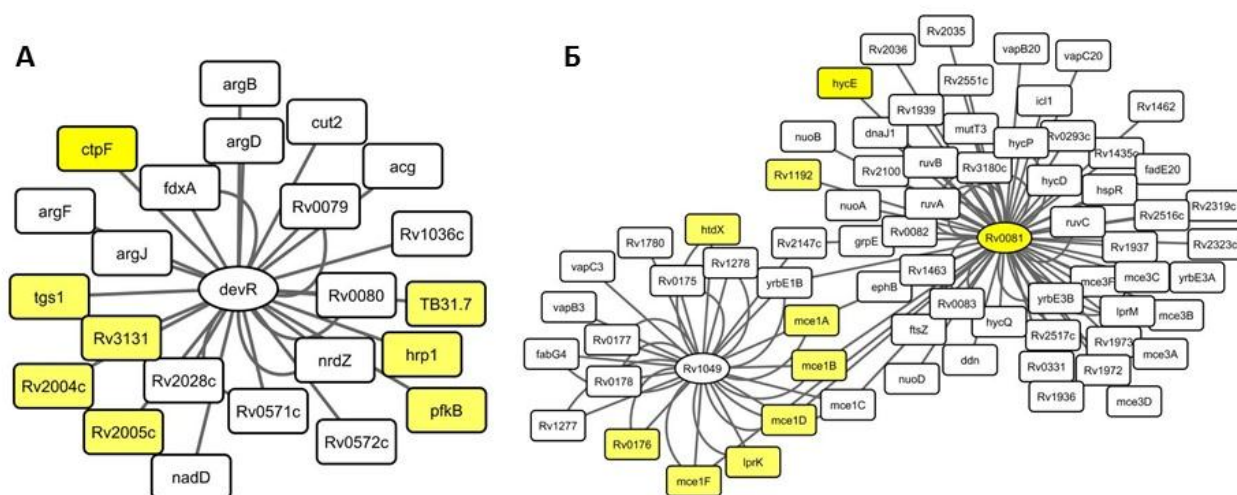


Рисунок 15. Фрагмент регуляторной сети *M. tuberculosis*.

Овалами обозначены транскрипционные факторы, прямоугольниками - регулируемые ими гены. Желтым цветом обозначены гены, кодирующие белки низкопредставленные в штаммах кластера Beijing B0/W148.

А. Группа генов, находящаяся под регуляцией DosR (devR/Rv3133c).

Б. Группа генов, находящаяся под регуляцией Rv0081.

3.5.1. Валидация данных методом MRM анализа

На основании анализа совокупного массива масс-спектров пептидных фракций, накопленных при протеомном анализе штаммов *M. tuberculosis* кластера Beijing B0/W148 ($n = 7$) была сформирована протеомная библиотека переходов (исследование проведено совместно с Бутенко И. О.). Для каждого из идентифицированных пептидов (заряд прекурсора +2 и +3) с FDR меньше 5 % с помощью программного пакета Skyline v 2.6 были отобраны 5 наиболее интенсивных *b*- и *y*-фрагментов. Дополнительным условием были: специфичность гидролиза белков трипсином, заряд ионов-фрагментов +1 или +2. Итого отобрано более 39000 переходов кандидатов. Далее для 759 белков, составляющих «основной» протеом (найденны во всех образцах), были выбраны 840 пептидов. Все переходы (по 5 на пептид) были распределены на 10 запусков, причем в каждом запуске список был оптимизирован по равномерности времени выхода пептидов. В каждом списке присутствовало не более 500 переходов на список, так как большие числа переходов влияют на время съемки MRM спектра (dwell-time). Для 120 пептидов с помощью программы mGen пакета mProphet в списки переходов были добавлены decoy-переходы для дальнейшей оценки FDR.

Для экспериментальной валидации библиотеки переходов проведено 10 LC-MS/MS scheduled-MRM экспериментов смеси 7 образцов *M. tuberculosis* кластера Beijing B0/W148, по одному с каждым из полученных списков. Таким образом получена библиотека переходов для всех 840 пептидов, включающая более 5000 записей о ионном токе (XIC) для отобранных переходов. Полученные результаты были обработаны с помощью программного пакета mProphet. Из 840 пептидов 729 идентифицированы с FDR меньше 0.01. Таким образом, получена библиотека протеотипических пептидов, которую можно использовать в дальнейшем для квантификации. Библиотека размещена в международном депозитарии MRM библиотек (<http://www.peptideatlas.org/PASS/PASS00631>).

Представленность белков, кодируемых генами *DosR* регулона и ассоциированных с ним ТФ в штаммах кластера Beijing B0/W148 и H37Rv дополнительно так же была оценена методом мониторинга множественных реакций. Был сформирован тестовый набор из 40 белков, включающий 35 дифференциальных белков и 5 не идентифицированных. Для проведения анализа использован пул штаммов кластера Beijing B0/W148 (7 образцов в 3х биологических повторах) и штамма H37Rv (1 штамм в 3х биологических повторах). В ходе анализа было идентифицировано 176 пептидов, соответствующих 40 белкам, идентифицированным для обеих групп штаммов. Таким образом, в ходе проведенного анализа 31 белок из 40 идентифицирован с помощью 2 - 11 пептидов. Остальные 9 белков контролировались с помощью одного пептида из-за ограниченности успешного MRM анализа для других триптических пептидов внутри этих белков. Кроме того, данные девять белков не были определены ранее, или были выявлены лишь в одной группе штаммов. Так, удалось подтвердить изменения в представленности наиболее значимых дифференциальных белков, в том числе кодируемых генами *DosR* регулона (Таблица 14).

Результаты MRM анализа белков *M. tuberculosis*

Ген	Синоним	pep_count	Beijing B0/W148	H37Rv	log ₂ _fc
Rv0081		1	335862.818125952	429083.610602908	-0.353386706475704
Rv0169	<i>mce1A</i>	10	7028231.12224762	8828924.08834337	-0.329076003747864
Rv0170	<i>mce1B</i>	11	4403584.36426322	7629552.83768465	-0.792920198839424
Rv0172	<i>mce1D</i>	6	4706467.6524101	6345799.55598646	-0.431157272423716
Rv0410c	<i>pknG</i>	7	2597638.41737314	1617166.89202493	0.683732053429216
Rv0592	<i>mce2D</i>	3	697164.330355706	461279.98642406	0.595856057965263
Rv0675	<i>echA5</i>	5	2057307.08660748	2042579.68417618	0.0103647944585634
Rv0757	<i>phoP</i>	5	4632907.64097685	3548952.84152502	0.384524517907772
Rv0824c	<i>desA1</i>	10	11898610.8843199	33289910.1878298	-1.48429182383199
Rv1071c	<i>echA9</i>	5	4397362.33651898	1312124.65539335	1.7447336271059
Rv1094	<i>desA2</i>	6	3328246.01799548	13380715.2362028	-2.00732125010544
Rv1106c	-	4	3324704.87280775	5968744.56521008	-0.844201233167351
Rv1180	<i>pks3</i>	8	1092841.52715476	1435895.68147992	-0.393866729364285
Rv1221	<i>sigE</i>	1	35830.4830549639	122570.782115351	-1.77435571796492
Rv1492	<i>mutA</i>	6	782861.923981389	339892.548698611	1.20367914221017
Rv1894c	-	1	227570.483281021	505715.048169141	-1.15201126127034
Rv1990c	-	1	32411.873406258	30464.8669639364	0.0893759705283543
Rv2289	<i>cdh</i>	6	3841084.41305565	9747714.47067834	-1.34355032296153
Rv2623	<i>TB31.7</i>	6	11316910.9782926	20100929.3805321	-0.828781987700904
Rv2703	<i>sigA</i>	8	3315868.3276334	4288964.99153673	-0.371242821211329
Rv2710	<i>sigB</i>	2	130048.143732081	110752.485943227	0.231706725050182
Rv2711	<i>ideR</i>	4	2525823.01130485	3588778.25653622	-0.506739234070562
Rv2831	<i>echA16</i>	5	3309104.41235754	2316095.15050904	0.514746290035457
Rv2982c	<i>gpdA2</i>	3	1278509.0995569	1852776.25423598	-0.535226239787549
Rv3010c	<i>pfkA</i>	2	582335.123844435	1352904.94601121	-1.2161389369002
Rv3080c	<i>pknK</i>	4	1209241.2498895	965598.523678702	0.324606722987186
Rv3107c	<i>agpS</i>	3	541280.506875138	683656.054253944	-0.336894257922597
Rv3130c	<i>tgs1</i>	8	1837463.75811705	11344454.6328797	-2.62619955551999
Rv3133c	<i>DosR/DevR</i>	3	3591814.46160309	3898241.83308803	-0.11811076475435
Rv3249c	-	1	206141.315158284	190469.077553339	0.114076884242228
Rv3286c	<i>sigF</i>	1	42122.3272669491	40690.0461464082	0.0499092311864288
Rv3494c	<i>mce4F</i>	2	432592.408819926	333793.824419504	0.374051085564708
Rv3526	<i>kshA</i>	3	600281.433381228	139259.98879206	2.10785823389884
Rv3537	<i>kstD</i>	1	197966.267696383	65923.2006320439	1.58639643087454
Rv3544c	<i>fadE28</i>	1	120834.188567702	114733.520780955	0.0747417517959581
Rv3570c	<i>hsaA</i>	6	2726409.60994089	1780204.66434426	0.614959213153307
Rv3574	<i>kstR</i>	1	54627.2398565	62123.8038090581	-0.185525637905474
Rv3597c	<i>lsr2</i>	2	397742.400578817	294753.401659332	0.432325903120776
Rv3602c	<i>panC</i>	3	843094.632148877	1903353.55139236	-1.17477709078987
Rv3825c	<i>pks2</i>	11	3636459.82324765	5549543.81891844	-0.609834547534889

Серым цветом выделены белки, подробно обсуждающиеся в тексте диссертации

3.6 Сравнительный анализ геномных и протеомных данных

Для сопоставления геномных и протеомных данных использованы данные полногеномного секвенирования штаммов кластера Beijing B0/W148. В ходе анализа были выявлены геномные изменения в 8 из 17 белков, характерных только для штаммов кластера Beijing B0/W148 и отсутствующих в H37Rv. Так, в штамме H37Rv промоторные области генов *Rv2277c*, *Rv2475c* и *Rv3323c* содержали вставку *IS6110*. В то время как в штаммах Beijing B0/W148 она отсутствовала, что возможно сказывается на экспрессии соответствующих генов и как следствие на синтезе соответствующих белков. Так же в геноме H37Rv, в промоторной области гена *Rv2974c* было выявлено три CG повтора, а геноме Beijing B0/W148 только два таких повтора. Дополнительно для Beijing B0/W148 в промоторной области гена *Rv0976c* была выявлена вставка одного нуклеотида, и несинонимичные замены SNPs (nsSNPs) в кодирующих областях *Rv0945*, *Rv1319c* и *Rv2351*.

Среди 57 белков, отсутствующих в штаммах Beijing B0/W148, гены 33х содержали мутации по сравнению с H37Rv. Стоит отметить, 6 из них относились к так называемым регионам различия (RD105, RD149, RD152, RD181 и RD207) (Думова *et al.*, 2014). Отсутствие двух других белков возможно связано со вставкой (*Rv0888*: 987586 insGG) (Рисунок 16) и делецией (*Rv1997*: 2241032 delG) в кодирующих областях соответствующих генов и приводящих к сдвигу рамок считывания. Дополнительно выявлены изменения в промоторных областях 3 генов, а так же 22 гена, содержащих nsSNPs в кодирующих областях (Таблица 10, 11).

Гены, кодирующие количественно различающиеся белки ($n = 192$) были сопоставлены с известными для H37Rv оперонами в соответствии с Microbes Online Operon Predictions (www.microbesonline.org) (Dehal *et al.*, 2010). Таким образом, 30 генов были сгруппированы в 11 оперонов. В большинстве случаев наблюдались однонаправленные изменения в белках, кодируемых генами одного оперона. Однако, изменения в двух генах *Rv1380* и *Rv1384*, относящихся к одному

оперону, привели к разнонаправленным изменениям в представленности соответствующих белков.

Согласно анализу генов, кодирующих дифференциальные белки, были выявлены: вставка (*Rv3234c*: 3610391 insC), делеция (*Rv0927c*: 1034211 delTGC; *Rv1884*: 2094915 delCGTCAG) и 35 nsSNPs в кодирующих областях. Дополнительно, на основании геномных данных был проведен поиск мутаций в сайтах инициации транскрипции (Transcription Start Sites, TSS) и в сайтах связывания транскрипционных факторов (Cortes *et al.*, 2013; Turkarslan *et al.*, 2015). В частности были выявлены мутации в сайтах связывания ТФ генов *Rv0169*, *Rv1129c*, *Rv1872c* и в TSS генов *Rv0169*, *Rv1196*, *Rv1508c*, *Rv1872c*, *Rv2627c*, *Rv2711*.

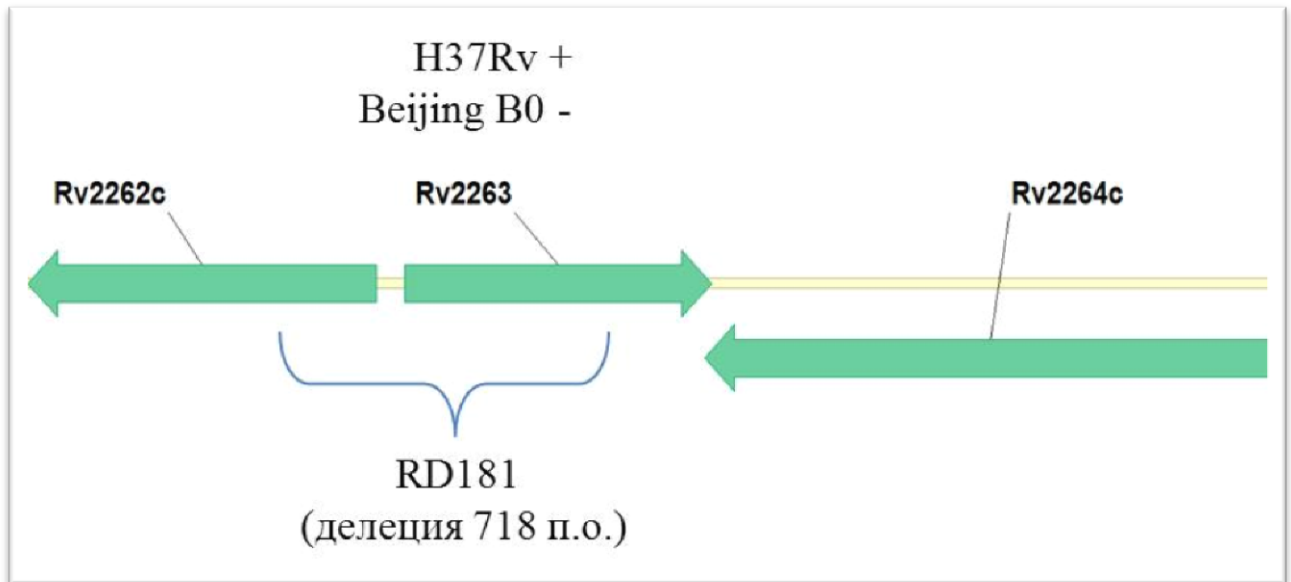


Рисунок 16. Сопоставление протеомных и геномных данных. Стрелками обозначены гены, соответствующих белков.

4. ОБСУЖДЕНИЕ

4.1. Формирование экспериментальных групп штаммов и геномных последовательностей

Исследование проведено на клинических изолятах *M. tuberculosis* кластера Beijing B0/W148, лабораторном штамме *M. tuberculosis* H37Rv и *M. smegmatis* mc² 155, предоставленных соответствующими научно-исследовательскими институтами.

В референсную группу образцов для всех экспериментов был включен штамм *M. tuberculosis* H37Rv. Выбор референсного штамма был обусловлен тем, что на сегодняшний день H37Rv является наиболее изученным штаммом *M. tuberculosis*. Опубликовано большое количество научно-исследовательских работ по изучению лекарственной устойчивости, вирулентности штамма. А главное в свободном доступе имеется как полный геном, так и максимально полный протеом данного штамма.

В свою очередь, в экспериментальную группу вошли 10 штаммов *M. tuberculosis* кластера Beijing B0/W148 бактериологически и генетически хорошо охарактеризованные. Интерес к штаммам данного кластера связан с их широкой распространенностью на территории России и ассоциацией с лекарственной устойчивостью.

Геномный анализ был проведен для всей коллекции образцов *M. tuberculosis*. Однако для протеомного анализа не удалось получить достаточную бактериальную массу для трех штаммов *M. tuberculosis* кластера Beijing B0/W148. Скорость роста данных изолятов была значительно ниже, чем остальных семи. В связи с этим, для получения максимальной сходимости результатов для протеомного анализа были использованы семь штаммов *M. tuberculosis* кластера Beijing B0/W148, скорость роста которых была одинаковая.

Геномные последовательности 92 образцов коллекции *key strains* и 11 образцов генотипа Beijing были использованы для построения филогенетического дерева и определения Beijing B0/W148 кластер-специфических полиморфизмов.

4.2. Генетические особенности штаммов *Mycobacterium tuberculosis* кластера Beijing B0/W148

Для идентификации представителей кластера Beijing B0/W148 могут использоваться классические методы генотипирования, как было описано в главе 1.5 Характеристика кластера Beijing B0/W148. При этом сполиготипирование и 12 локусное MIRU типирование не позволяют отличить образцы кластера от других штаммов Beijing. Однако применение 24 локусного MIRU-VNTR типирования и IS6110 RFLP анализ позволяют получить характерные специфические паттерны (Mokrousov, 2013). В целом стоит отметить, что описанные методы лишь отчасти отражают генетические особенности штаммов и не дают представления об организации всего генома.

Для всестороннего анализа генетических особенностей представителей кластера Beijing B0/W148 в настоящем исследовании было проведено полногеномное секвенирование 10 образцов. Согласно первичному анализу полученных данных было установлено, что геномы имеют большой процент гомологии с референсным штаммом H37Rv (99.9 %), что согласуется с данными мировой литературы (Sreevatsan *et al.*, 1997; Fleischmann *et al.*, 2002). При этом количество однонуклеотидных полиморфизмов между геномами изучаемых штаммов кластера и H37Rv составило 1500 (среднее плюс/минус) (в среднем 3 SNPs на 10 т.п.о.). Последнее в полной мере характеризует объект исследования как генетически мономорфный организм (Achtman, 2008). Стоит отметить, что при попарном сравнении геномы Beijing B0/W148 отличались друг от друга не более чем на 25 полиморфизмов, что говорит о недавней экспансии представителей изучаемого кластера и согласуется с исследованиями Мокроусова и работой Merker (Mokrousov, 2013; Merker *et al.*, 2015).

Для оценки разнообразия секвенированных образцов, определения основных филогенетических маркеров и дальнейшего поиска семейство-специфических полиморфизмов был использован филогенетический анализ. Дополнительно для сравнения полученных результатов с мировыми данными в

анализ было включено 11 образцов *M. tuberculosis*, представленных в GenBank в виде протяженных секвенсов (Таблица 5), и 92 образца коллекции *key strains*.

В ходе анализа было выявлено, что представители кластера образуют узкую монофилитическую линию внутри генетического семейства Beijing. Причем штамм W-148 также соответствовал этому кластеру. Последнее позволило исследовать данный геном совместно с изучаемыми образцами.

В целом же стоит отметить, что штаммы кластера Beijing B0/W148 несли общие для линии Beijing филогенетические маркеры, которые представлены в Таблице 6. Для данных методов SNPs типирования можно провести аналогию со сполиготипированием и 15 локусным VNTR типированием, которые также не дифференцируют штаммы кластера от других представителей данной филогенетической линии. При этом более детальное изучение полиморфизмов геномов дало возможность определить, что Beijing B0/W148 относятся к линии «современных» Beijing (Ebrahimi-Rad *et al.*, 2003). Представители данной линии характеризуются секвенс-типом ST10 по Filliol (Filliol *et al.*, 2006), относятся к группе 3 по Gagneux (Gagneux *et al.*, 2006), секвенс-типу BST19 по Schurch (Schurch *et al.*, 2011), группе 2 по Tsolaki (Tsolaki *et al.*, 2005) и lineage 2.2.1 по Coll (Coll *et al.*, 2014). Представленные в работе данные впервые получены для Beijing B0/W148, однако и они не позволили дифференцировать представителей кластера от других штаммов Beijing, встречаемых по всему миру. В связи с этим был произведен поиск уникальных кластер-специфических полиморфизмов, которые могут служить основой для дальнейших прицельных исследований, а также создания систем генетического мониторинга указанного семейства. Методами сравнительной геномики было выявлено 59 полиморфизмов, которые представлены в Приложении 2 и описаны в диссертационной работе Шитикова Е. А. (Шитиков, 2014). Здесь же следует подчеркнуть, что идентифицированные мутации полностью совпали с исследованием Merker с соавт., также приводящих список Beijing B0/W148 – специфических мутаций (Merker *et al.*, 2015).

Стоит отметить, что в настоящем исследовании была выявлена делеция в гене *kdpD* (с.2541_2542delCA), ведущая к сдвигу рамки считывания. Последнее, в

свою очередь, приводит к образованию единой открытой рамки считывания двух генов *kdpDE* (Рисунок 17). Данная мутация является кластер-специфической для Beijing B0/W148 и в последнее время в мировом сообществе ей уделяют достаточно пристальное внимание, ассоциируя с вирулентностью (Merker *et al.*, 2015).

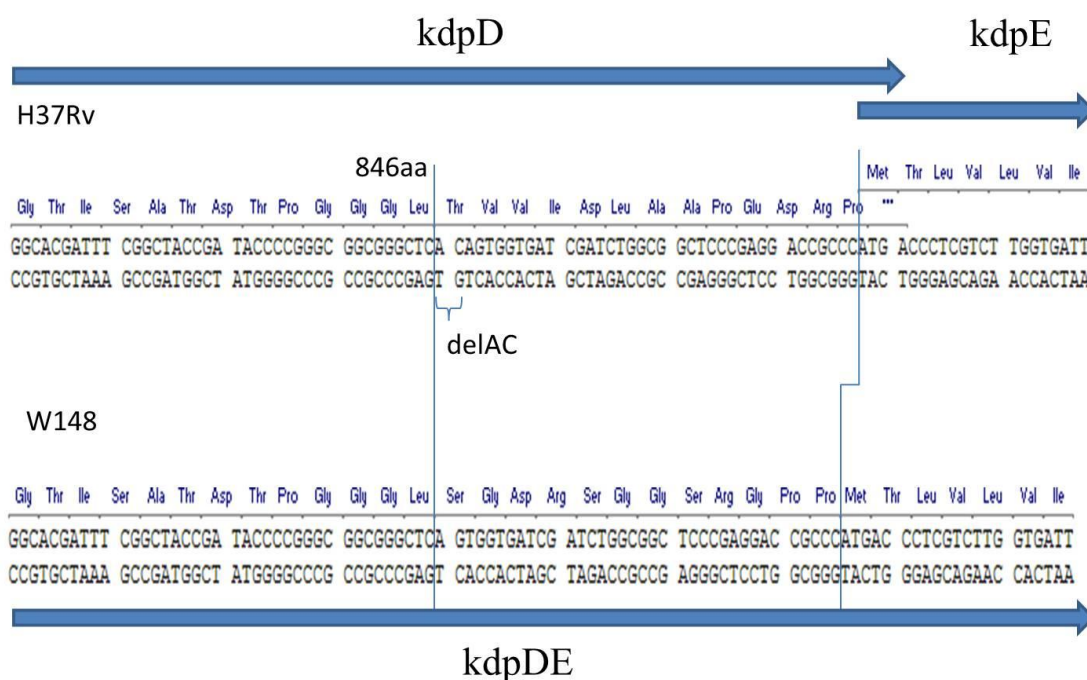


Рисунок 17. Образование единой открытой рамки считывания для генов *kdpD* и *kdpE*.

Рассмотрев особенности представителей кластера Beijing B0/W148 на уровне однонуклеотидных полиморфизмов, следующей задачей было определение структурной организации генома. Как отмечалось ранее, в ходе филогенетического анализа было установлено сходство секвенированных штаммов кластера со штаммом W-148, представленным в GenBank (CP012090.1). Выравнивание геномов W-148 и H37Rv показало наличие двух больших инверсий в геноме W-148 (Рисунок 6, 7), которые впоследствии были подтверждены для всех представителей кластера. Первый инвертированный участок (блок II) затрагивал 559 т.п.о. и располагался между генами *Rv0609a* и *Rv1135c*. Второй инвертированный участок (блок IV) распространялся на 339 т.п.о. и затрагивал участок геномной ДНК между генами *Rv3019c* и *Rv3327*. Следует отметить, что

такие масштабные рекомбинационные события в геноме микобактерий крайне редки и в мировой литературе описаны только два таких случая. Так для трех штаммов микобактерий туберкулеза, выявленных в Южной Африке и секвенированных в Broad Institute, было показано наличие крупной инверсии сегмента хромосомы длиной около 2.5 м.п.о. При этом повторное секвенирование одного из штаммов в другой лаборатории не подтвердило наличие данной перестройки (Ioerger *et al.*, 2009). В свою очередь в геноме *M. avium* была детектирована инверсия между подвидами *hominissuis* и *paratuberculosis* (Hsu *et al.*, 2011). Немного более частым рекомбинационным событием для *M. tuberculosis* является дупликация отдельных сегментов хромосомы. Изначально Domenech с соавт., а позже и Weiner с соавт. описали крупные дупликации, произошедшие в геномах *M. tuberculosis* генотипа Beijing (Domenech *et al.*, 2010; Weiner *et al.*, 2012). Здесь следует подчеркнуть ряд параллелей с настоящим исследованием. Штаммы, описанные в данных работах относились к группам 3, 4 и 5 по Gagneux (Gagneux and Small, 2007). Напомним, что штаммы кластера Beijing B0/W148 относятся к группе 3. Вторым интересным моментом является тот факт, что в большинстве случаев дупликации происходили между сонаправленными копиями повтора IS6110. В описанном в настоящей работе случае инвертированные участки располагались между разнонаправленными копиями того же повторяющегося элемента. Согласно работе Weiner с соавт. граница дублированного участка для одного из образцов находилась в области гена *Rv3327*, что соответствует границе одного из инвертированных участков, описанного в представленном исследовании (Weiner *et al.*, 2012). Дополнительно стоит отметить, что данный регион так же соответствует *RvD5* в геноме H37Rv (Brosch *et al.*, 1999).

Обобщая приведенную информацию можно сделать предположение, что область генома в районе 3 – 4 м.п.о. не стабильна и произошедшие перестройки могут отражать происходящую эволюцию топологии хромосомной ДНК.

Принимая во внимание выявленные рекомбинационные события, а также возможные дупликации геномной информации, было проведено дополнительное

полногеномное секвенирование *M. tuberculosis* штамма SP21 с последующим анализом генома и сопоставлением с результатами IS6110 RFLP типирования. В ходе работы было выявлено сходство геномной организации штамма SP21 с геномом W-148. Дополнительных рекомбинационных событий обнаружить не удалось.

Согласно RFLP анализу штаммы кластера Beijing B0/W148 несут 17 бэндов, соответствующих 17 копиям повторяющегося элемента IS6110 (Рисунок 8, 9). В ходе анализа геномов образцов SP21 и W-148 была выявлена 21 копия элемента. Такое различие количества выявляемых копий IS6110 связано с «маскированием» повторов и описано в мировой литературе (McEvoy *et al.*, 2007). В представленном случае повторы 4 и 5, а также с 11 по 15 имеют схожие длины рестрикционных фрагментов и тем самым становятся неразличимы в ходе RFLP анализа. При этом особое внимание заслуживают повторы 1, 2, 3, 14 и 20 (Рисунок 9). Повтор 1 был изучен в работе Мокроусова с соавт. и расположен между генами *TBPG_002546* и *TBPG_002549* штамма W-148 (*Rv2664* и *Rv2665* по H37Rv) (Mokrousov *et al.*, 2012). Данный повтор может быть использован для типирования штаммов кластера, так как образуемый им рестрикционный фрагмент уникален и не встречается у представителей других генотипов. Повтор 20 располагается в так называемом NTF регионе и характеризует штаммы Beijing как «современные» (Mokrousov *et al.*, 2005). В свою очередь повторы 2, 3 и 14 являются границами инвертированных участков. При этом повторы 3 и 14 были выявлены на границах инвертированных участков между блоками III - IV и II - III, соответственно. Повтор 2 имел усеченную форму и располагался между блоками I и IV. Вторая часть усеченного повтора была выявлена между блоками II и V, однако она не представлена на Рисунке 7, т.к. эта часть не содержит последовательности, комплементарной зонду, используемому для RFLP анализа.

4.3. Разработка метода протеомного анализа микобактерий

На сегодняшний день получение ДНК и РНК микобактерий и любые дальнейшие манипуляции с ней не представляют труда. Доступно достаточно

большое количество различных протоколов (Akhtar *et al.*, 2011; Ghodbane *et al.*, 2015; Tyler *et al.*, 2016) и лабораториями накоплен опыт их применения (Shitikov *et al.*, 2012; Пина *et al.*, 2013; Vespyatykh *et al.*, 2014; Chernyaeva *et al.*, 2014; Shitikov *et al.*, 2014). Однако иначе обстоит дело с получением и анализом белков микобактерий. Ситуация осложняется тем, что для каждого из патогенов необходимо подбирать такую систему лизиса, которая бы позволяла разрушать клетки, не повреждая белки и при этом позволяющая максимально их отделить от остальных компонентов клетки. Отсутствие достаточного количества стандартизированных протоколов протеомного анализа, свидетельствует в пользу того что *M. tuberculosis* является довольно сложным объектом. Стоит так же отметить особенность строения клеточной стенки непосредственно микобактерий, что так же требует более сложных методов ее разрушения.

Для разработки метода экстракции белка, в качестве тест-культуры, был выбран штамм *M. smegmatis* mc² 155. Выбор обусловлен тем, что *M. smegmatis* является близким родственником *M. tuberculosis*, но относиться к быстро растущим и не патогенным микобактериям, что облегчает культивирование данного штамма в обычных лабораторных условиях, а так же позволяет быстрее получать результаты.

Таким образом, все эксперименты по подбору наиболее оптимального метода лизиса клеток проведены на модельном организме *M. smegmatis* mc² 155. Были протестированы несколько лизирующих буферов, описанных в литературе, в сочетании с Bead Beater (MPBio, США) и ультразвуковым (ИНЛАБ, Россия) гомогенизаторами. На *M. smegmatis* буфер 1В (Приложение 3) показал хорошие результаты по разрушению клеток в сочетании с ультразвуковым гомогенизатором. Однако было сделано заключение, что данная система является открытой и небезопасна при использовании для разрушения клеток возбудителя туберкулеза. В связи с этим были подобраны оптимальные условия разрушения клеток в упомянутом выше лизирующем буфере на гомогенизаторе Bead Beater. Согласно полученным данным максимальный выход белка наблюдали при разрушении клеток два раза по 2 минуты с 5-минутными перерывами в ледяной

бане. При отработке методов гидролиза полученных белков трипсином было показано, что наиболее оптимальным является расщепление в геле. Согласно литературным данным в последнее время для многих микроорганизмов применяют метод гидролиза в растворе с реагентом RapiGest (Mishra and Sarkar, 2015). Однако, в нашем случае, при использовании данного метода было детектировано значительно меньшее количество пептидов, нежели после гидролиза в геле. Возможно, данное обстоятельство так же связано с особенностями строения клеточной стенки патогена.

Сравнительный анализ списков идентифицированных белков установил, что три биологических и два технических повтора являются минимально достаточными для описания протеома микобактерий. В общей сложности хромато-масс-спектрометрический анализ триптических пептидов позволил идентифицировать 3937 белков *M. smegmatis*. Полученный результат сопоставим с данными Tagun Chopra с соавт., которые в подобных измерениях идентифицировали 3067 белков *M. smegmatis* (Chopra *et al.*, 2014), из них 2230 белка совпадали с идентифицированными нами.

Стоит отметить, что для протеомного анализа был использован современный масс-спектрометрический подход, называемый так же shotgun, направленный на максимальный охват тотального протеома клетки. Первые работы, использовавшие данный метод для изучения *M. tuberculosis* были опубликованы в 2004 году (Schmidt *et al.*, 2004); на сегодняшний день в PubMed находится около 30 статей, но всего две статьи посвящены семейству Beijing. De Souza с соавт. идентифицировали 1668 белков гипер- и гиповирулентных штаммов семейства Beijing, с использованием безметочной протеомики (de Souza *et al.*, 2010). В другом исследовании авторы идентифицировали 2392 белков семейства с использованием меточного гель-зависимого MS подхода и описали различия между «древними» и «современными» изолятами Beijing (de Keijzer *et al.*, 2014). В свою очередь, в настоящей работе было идентифицировано 1868 белков штаммов кластера Beijing B0/W148, что сопоставимо с количеством получаемых белков в описанных выше работах. Несмотря на успешность

применения методов shotgun протеомного анализа, определить весь белковый пул микобактерий остается практически невозможным. Основная проблема получения информации о полном протеоме связана с различной степенью представленности белков в сложном образце, коим будет являться культура клеток бактерий. Так, в суммарном масс-спектре количество и интенсивность пиков для «мажорных» белков будет значительно выше, чем для «минорных». В связи с этим идентификация именно последних бывает затруднительной. Дополнительно, стоит учитывать, что каждый белок в клетке имеет свою продолжительность жизни.

Как известно разные генетические группы *M. tuberculosis* демонстрируют различные физиологические особенности и характеристики, что может сказываться на эффективности экстракции белка. В связи с этим был проведен сравнительный анализ распределения идентифицированных белков для группы штаммов Beijing B0/W148 и H37Rv в соответствии с их функциями и локализацией, с использованием базы данных TubercuList и PSORTdb (Рисунок 13). Анализ распределения по функциональным категориям показал достаточно гомогенное распределение белков обеих групп штаммов, что так же соответствует результатам опубликованным ранее (de Keijzer *et al.*, 2014). В свою очередь анализ локализации белков так же не выявил различий в распределении идентифицированных белков между группами. Таким образом, было показано, что наблюдаемые белки обеих групп штаммов имеют весьма схожее распределение с точки зрения их функций и локализации, на основании чего сделан вывод, что эффективность разработанного метода протеомного анализа не зависит от физиологии возбудителя. Кроме того, в данной работе было идентифицировано 10 из 12 белков клеточной стенки. Высокий процент идентификации даже сильно гидрофобных белков клеточной стенки свидетельствует о том, что разработанный подход обеспечивает надежную идентификацию протеома *M. tuberculosis* без дискриминации гидрофобных белков.

На сегодняшний день MRM представляет собой один из наиболее чувствительных методов для белковой идентификации. В связи с этим он был использован для дополнительной верификации достоверности получаемых результатов. Так, были отобраны 759 белков, присутствующие во всех штаммах кластера Beijing B0/W148 в трех биологических повторах и определены как основные. Далее для них были отобраны протеотипические пептиды и проведен MRM анализ. В ходе анализа была получена библиотека протеотипических пептидов для 729 основных белков.

Дополнительно, с помощью данного метода были проверены результаты количественного анализа. Был сформирован тестовый набор, включающий 35 дифференциальных белков. В ходе проведенного анализа была показана сходимость получаемых результатов. Дополнительно, с помощью данного метода было показано наличие белка Rv0081 в штаммах кластера Beijing B0/W148. Однако его представленность в них значительно ниже, чем в H37Rv, что согласуется с изначально полученными данными. Таким образом, в ходе валидации данных высокочувствительным методом MRM показана достоверность получаемых протеомных данных.

4.4. Протеомные особенности изолятов *Mycobacterium tuberculosis* кластера Beijing B0/W148

Для детального описания специфических особенностей кластера Beijing B0/W148 был проведен сравнительный анализ штаммов кластера и H37Rv, с использованием данных как масс-спектрометрического, так и количественного протеомного анализа. Для количественного сравнения были рассмотрены все существенные изменения без учета порогов и fold change, для максимального охвата дифференциально синтезируемых белков. В общей сложности было идентифицировано 266 дифференциальных белков, из которых 41 белок был высоко- и 225 низкопредставленными в штаммах кластера Beijing B0/W148.

Для того чтобы предсказать функциональные эффекты найденных особенностей протеома Beijing B0/W148 были использованы два независимых

взаимодополняющих подхода. Первый, GO анализ, был использован для идентификации функциональных категорий, обогащаемых дифференциальными белками ($n = 266$). Очевидно, что распределение белков внутри категории «Биологические процессы» является наиболее показательным. Соответственно, было выявлено, что существенная часть дифференциальных белков относится к группе «Метаболические процессы» (GO:008152). Более детальный анализ показал обогащение подгруппы GO:0006629, «Липидные метаболические процессы», высокопредставленными белками, в то время как низкопредставленные белки были относительно равномерно распределены между подгруппами: «Липидные метаболические процессы» (GO:0006629), «Метаболические процессы клеточных аминокислот» (GO:0006520) и «Углеводные метаболические процессы» (GO:0005975). Среди высокопредставленных белков, задействованных в биосинтезе липидов, длинноцепочечная жирнокислотная-CoA лигаза, FadD15 [EC: 6.2.1.3], известна как одна из семи жирнокислотных-CoA синтаз, индуцирующихся в вирулентных штаммах (Li *et al.*, 2010). Другой белок, AgpS [EC: 2.5.1.26], алкилдигидроксиацетонфосфат синтаза инициирует липидный метаболизм. Суммарно эти результаты позволяют предположить, что синтез липидов увеличен в штаммах кластера Beijing B0/W148. В подтверждение этой гипотезы так же свидетельствует то, что большая часть белков катаболизма жирных кислот низкопредставлена в штаммах кластера по сравнению с H37Rv (Таблица 12). Напротив, оксигеназа HsaA [EC: 1.14.14.12], субъединица флавинзависимой монооксигеназы, кодируемая 1 из 26 генов, необходимых для выживания бактерии в макрофагах (Rengarajan *et al.*, 2005), была наиболее представлена в штаммах Beijing B0/W148. Данный белок участвует в катаболизме стероидов и возможно оказывает важное влияние на инфицированного хозяина путем уменьшения локальной концентрации холестерина мембраны, изменяя иммунорегуляторные стероиды и продуцируя новые вторичные метаболиты (Dresen *et al.*, 2010).

Так же были изучены белки группы «Транспортная активность» (GO:0005215), внутри GO категории «Молекулярные функции». В данную группу вошли пять высокопредставленных и 15 низкопредставленных белков штаммов Beijing B0/W148. Стоит отметить, что пять низкопредставленных белков относились к ABC транспортерам, среди них гены белков PstB [EC: 3.6.3.27] и PstS1 относились к одному оперону, вовлеченному в импорт фосфатов во время голодания, которое характерно для бактерии находящейся внутри фагосомы (Rengarajan *et al.*, 2005). В дополнение, была показана повышенная экспрессия *pstS1* во время фосфатного голодания (Pulido *et al.*, 2014).

Стоит отметить, что GO анализ отражает самостоятельную роль дифференциальных белков и не учитывает, потенциально взаимосвязанных изменений в бактериальной клетке. Для того чтобы оценить эффекты белковых взаимодействий была использована генная регуляторная сеть *M. tuberculosis* из ранее опубликованной работы Peterson с соавт. (Peterson *et al.*, 2014). Анализ показал сниженную представленность 11 белков DosR системы в штаммах Beijing B0/W148. Данная система хорошо описана для *M. tuberculosis*; в H37Rv, она индуцируется при трех условиях, которые ингибируют аэробное дыхание: гипоксия, NO, или CO (Selvaraj *et al.*, 2012; Galagan *et al.*, 2013; Sivaramakrishnan and de Montellano, 2013; Pulido *et al.*, 2014). Так же хорошо известно, что во время экспоненциальной фазы роста уровень транскриптов DosR регулона значительно выше в штаммах семейства Beijing, чем в H37Rv (Domenech *et al.*, 2010; Fallow *et al.*, 2010). Однако, Badillo-López с соавт. продемонстрировали, что во время гипоксии, наоборот, уровень транскриптов DosR регулона значительно ниже в штаммах семейства Beijing по сравнению с H37Rv (Badillo-Lopez *et al.*, 2010). Возможно, это объясняется тем, что штаммы Beijing осуществляют альтернативный ответ на гипоксию, нежели тот, что использует H37Rv. Наши результаты демонстрируют низкую представленность DosR-регулируемых белков в штаммах Beijing B0/W148 во время стационарной фазы роста. Это может быть связано с предварительной адаптацией бактериальных клеток к потенциально низким уровням кислорода при длительном культивировании *in vitro*.

Штаммы Beijing B0/W148 относятся к линии «современных» Beijing (Mokrousov, 2013), которая отличается от «древних» высокой представленностью белков Rv0450c и Rv3137 [EC: 3.1.3.15], а так же низкой представленностью Rv1269c и Rv3283 [EC: 2.8.1.1 2.8.1.2] (de Keijzer *et al.*, 2014). Согласно нашим данным представленность Rv3283 (SseA) [EC: 2.8.1.1 2.8.1.2] была значительно ниже в штаммах Beijing B0/W148, чем в H37Rv (fold change = 0.04). Предположительно данный белок является тиолоксидоредуктазой и совместно с ферментом супероксиддетоксикации SodA (Rv3846) [EC: 1.15.1.1] и интегральным мембранным белком DoxX (Rv3005c), формирует мембранассоциированный оксидоредуктазный комплекс (от англ. membrane-associated oxidoreductase complex (MRC)) (Nambi *et al.*, 2015). Потеря любого компонента MRC ведет к нарушению утилизации микотиола, который является функциональным аналогом глутатиона в клетках *M. tuberculosis*. Таким образом, снижение представленности белка SseA [EC: 2.8.1.1 2.8.1.2] в Beijing B0/W148 может приводить к накоплению окислительных повреждений клетки, вызванных активными формами кислорода (АФК).

4.5 Сравнительный анализ геномных и протеомных данных

В ходе анализа геномов штаммов Beijing B0/W148 предложены объяснения 47 % изменений, обнаруженных в ходе протеомного анализа. В частности, показано, что гены шести белков, отсутствующих в штаммах кластера, входят в состав пяти делетированных регионов (RD), характерных для всего семейства Beijing (Dymova *et al.*, 2014). Дополнительно были найдены nsSNPs в кодирующих областях 60 генов, соответствующих дифференциальным белкам (n = 266). Интересно отметить, что из них семь белков содержали по две и более аминокислотные замены. Большая часть выявленных замен приходилась на серин (13 %) и аргинин (11.6 %). В то время как серин был с одинаковой частотой изменен на ряд различных аминокислот (Ser→Ala (1), Ser→Cys (1), Ser→Gly (1), Ser→Leu (2), Ser→Trp (1), Ser→Thr, Ser→Pro (2)), большинство замен аргинина приходилось на глицин (Arg→Gly (5), Arg→Pro (1), Arg→Cys (1), Arg→Gln (1)).

Были выявлены вставки IS6110 в промоторных областях трех генов, кодирующих белки, высокопредставленные в штаммах кластера Beijing B0/W148. Предположительно данные мутации могут влиять на уровень экспрессии генов и как следствие наблюдается изменение представленности соответствующих белков. Для генов, кодирующих 18 дифференциальных белков были выявлены sSNPs. Хотя, скорее всего вклад в изменение уровня экспрессии они не вносят. Стоит отметить, что большую часть дифференциальных белков, полученных в ходе количественного анализа, не удалось объяснить на геномном уровне.

Отдельно стоит отметить выявленную в данном исследовании делецию в гене *kdpD* (с.2541_2542delCA), приводящую к сдвигу рамки считывания, что приводит к образованию слитого белка KdpDE. При этом не удивительно, что данные белки не были идентифицированы в ходе масс-спектрометрического анализа. Представленные белки относятся к одной из 11 известных двухкомпонентных систем микобактерий туберкулеза и регулируют работу калиевых каналов. При этом белок KdpD [EC: 2.7.13.3] является гистидинкиназой, а белок KdpE – регулятором ответа, влияющим непосредственно на работу KdpFABC комплекса (K⁺-зависимая P-АТФаза) (Рисунок 18). Активность системы увеличивается при нехватке ионов калия в среде, переходе клеток в среднюю и позднюю логарифмические фазы роста при культивировании *in vitro*, а также на вторые сутки при росте в макрофагах. Следует отметить, что в 2003 году Parish с соавт. на мышиных моделях было показано, что частичное делетирование *kdpDE* в *M. tuberculosis* H37Rv приводит к увеличению вирулентности (Parish *et al.*, 2003). Таким образом, можно предположить, что образование слитого белка с измененной функцией с большой вероятностью может служить причиной «успешности» представителей кластера.

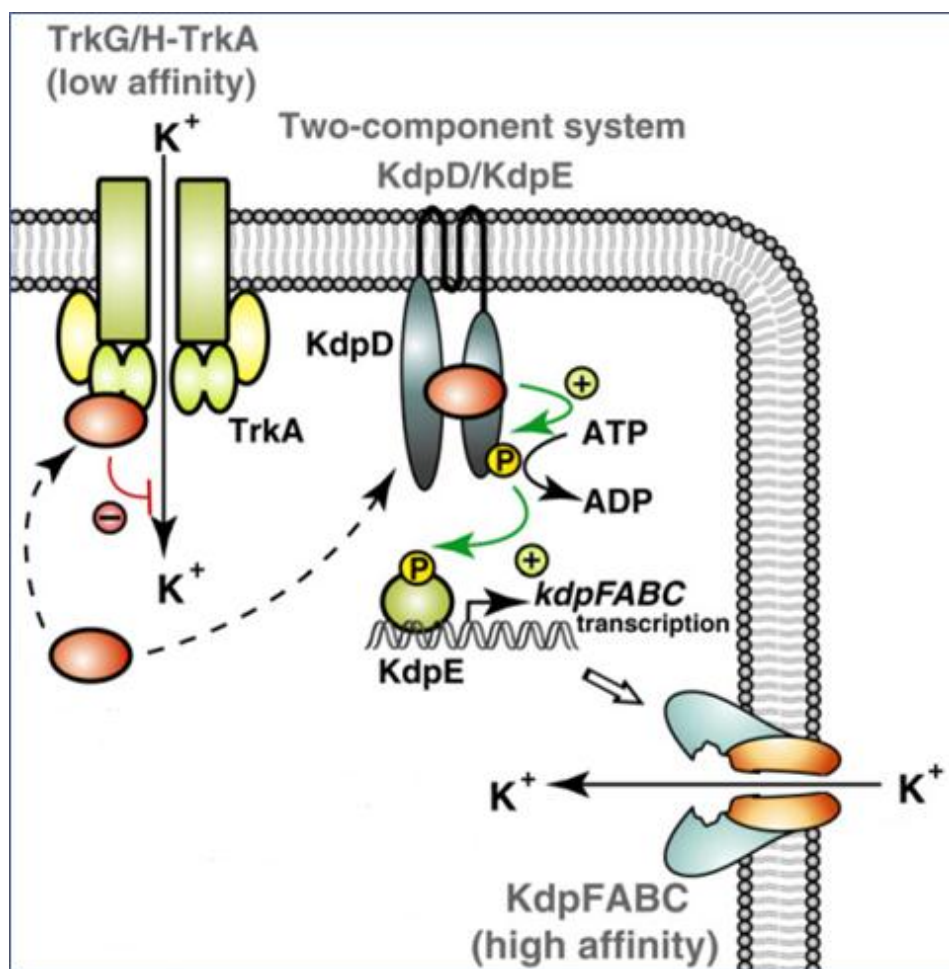


Рисунок 18. Системы регуляции калия в клетках микобактерий.

Наиболее интересные белковые изменения, связаны с DosR регулоном. Данный регулон играет важное значение при переходе бактерии в дормантное состояние (Park *et al.*, 2003) и адаптации к неблагоприятным условиям среды, стрессам (Rustad *et al.*, 2008). Хотя, стоит отметить, что механизмы ответа микобактерий на данные изменения до сих пор до конца не ясны. Экспериментально было показано, что экспрессия генов DosR индуцируется в условиях гипоксии *in vitro* (Rustad *et al.*, 2008), а так же при инфицировании мышей (Karakousis *et al.*, 2004). При выходе бактерий из дормантного состояния так же наблюдается увеличение экспрессии генов регулона (Leistikow *et al.*, 2010). Дополнительно в ряде работ было продемонстрировано, что уровень экспрессии DosR различается в штаммах с разной вирулентностью (Malhotra *et al.*, 2009).

DosR регулон включает около 50 генов, при этом сам ген *Rv3133/dosR* является глобальным регулятором, частью необычной двухкомпонентной

системы DosR–DosS–DosT (Pulido *et al.*, 2014). В настоящей работе выявлены аминокислотные замены в 13 генах данного регулона. При этом шесть из них соответствовали белкам низкопредставленным в штаммах кластера. Интересно так же отметить, что была выявлена замена в промоторной области гена *Rv2003c*, продуктом которого является гидроксистероиддегидрогеназа, не обнаруженная при протеомном профилировании. При этом мы наблюдаем снижение представленности гипотетических белков *Rv2004c*, *Rv2005c*. В настоящее время гены *Rv2003c*, *Rv2004c*, *Rv2005c* не объединяют в один оперон, однако, возможно выявленная замена расположена в промоторе оперона, что связано со сниженной их экспрессией как следствие низкой представленностью соответствующих белков.

Стоит отметить транспортный белок CtpF [EC: 3.6.3.-], отсутствовавший в штаммах кластера Beijing B0/W148, являющийся АТФазой Р-типа и потенциальным транспортером катионов щелочноземельных металлов. CtpF [EC: 3.6.3.-] это единственная АТФаза Р-типа, находящаяся под регуляцией глобального регулятора DosR и сверх экспрессирующаяся в условиях гипоксии. Согласно протеомному анализу данный белок обнаруживается только в штамме H37Rv и не встречается в штаммах кластера. При анализе геномных данных было выявлено, что соответствующий ген содержит однонуклеотидную делецию (*Rv1997*: 2241032 delG), приводящую к сдвигу рамки считывания. Данная мутация объясняет отсутствие белка, стоит отметить, что она не является кластер-специфической, однако, это основная характеристика семейства Beijing в целом.

Представляет интерес то, что ТФ *Rv0081*, входящий в состав DosR регулона, отсутствовал во всех штаммах Beijing B0/W148 согласно масс-спектрометрическому анализу и согласно MRM анализу, был низкопредставлен в штаммах кластера. Ранее это белок был описан Galagan с соавт. как регуляторный центр при ответе на гипоксию (Galagan *et al.*, 2013). В ходе GO анализа было выявлено, что *Rv0081* совместно с *Rv1049*, контролирует экспрессию *mce1* оперона, состоящего из шести генов *mce* (*Rv0169–Rv0174*), двух генов *urbE* (*Rv0167*, *Rv0168*), и четырех генов *mce*-ассоциированных (*Rv0175–Rv0178*).

Согласно полученным результатам четыре белка Mse (Rv0169, Rv0170, Rv0172, и Rv0174), и один белок Mse-ассоциированный (Rv0176) являются низкопредставленными в штаммах Beijing B0/W148. При этом три белка Mse (Rv0169, Rv0172, Rv0174) содержали аминокислотные замены. Белки, кодируемые *mse1* опероном, относятся к ABC транспортерам, участвующим в транспорте фосфолипидов, они так же необходимы микобактериям для выживания в макрофагах или мышинных моделях (Uchida *et al.*, 2007; Forrellad *et al.*, 2013). Было показано, что делеция *mse1* оперона способствует формированию гипервирулентного фенотипа (Forrellad *et al.*, 2013). Примечательно, что эти белки не были определены ранее в ходе GO анализа, подчеркнув тем самым необходимость использования различных подходов к функциональной оценке белков и лучшему пониманию, их совместного функционирования для обеспечения выживания клеток.

Стоит так же отметить, что при этом белок Mse4F, так же относящийся к ABC транспортерам, напротив, был высокопредставлен в штаммах кластера Beijing B0/W148. Данный белок участвует в транспорте холестерина, регулируя его содержание в наружной мембране. Ранее сообщалось, что данный белок необходим микобактериям для выживания в макрофагах, активированных IFN- γ (Pandey and Sasseti, 2008). При этом изменений на геномном уровне для данного белка обнаружено не было.

Согласно опубликованным данным белки эффлюксных помп, Rv0341, Rv2688c [EC: 3.6.3.-] и Rv3728, были найдены исключительно в штаммах семейства Beijing (de Keijzer *et al.*, 2014). Однако в представленной работе белок Rv0341 был обнаружен во всех штаммах, включая H37Rv, причем его представленность в штаммах Beijing B0/W148 и H37Rv так же не отличалась. К сожалению белки Rv2688 [EC: 3.6.3.-] и Rv3728 не были идентифицированы. Следует отметить, что согласно геномному анализу ген Rv3728 содержит B0/W148-специфическую мутацию, приводящую к образованию стоп-кодона. Что согласуется с отсутствием данного белка в общем протеоме штаммов Beijing B0/W148.

Суммарно полученные данные свидетельствуют о том, что протеомные и геномные особенности штаммов Beijing B0/W148, вносят свой вклад в их повышенную вирулентность и успешное географическое распространение. Зафиксировано увеличение представленности ферментов, ответственных за биосинтез длинноцепочечных жирных кислот, что совпало с уменьшением белков, ответственных за их деградацию. Микобактерии могут использовать длинноцепочечные жирные кислоты (в длину до 86 - 95 атомов углерода) для сборки миколовых кислот и различных липидов, которые считаются основными факторами вирулентности *M. tuberculosis*, в частности, на ранней стадии инфицирования, когда бактерии попадают в макрофаги хозяина. Возможно данные изменения так же связаны с использованием бактерией в качестве основного источника углерода липидов вместо углеводов при попадании в макрофаги (Eisenreich *et al.*, 2010). В пользу последнего так же свидетельствует зафиксированное снижение представленности белка PycA [EC: 2.7.1.40], являющегося пируваткиназой, ключевым ферментом гликолиза.

Показано увеличение представленности белка HsaA [EC: 1.14.14.12], участвующего в деградации стероидов. Во внутриклеточной среде *M. tuberculosis* используют холестерол в качестве источника энергии и для биосинтеза липидов клеточной стенки. Наблюдаемая разница в представленности белка HsaA [EC: 1.14.14.12] может свидетельствовать о лучшей выживаемости *M. tuberculosis* в макрофагах, известной характеристике штаммов Beijing B0/W148 (Pardini *et al.*, 2009; Lasunskaja *et al.*, 2010). В дополнение было показано снижение представленности белков, кодируемых генами *mce1* оперона, как известно его делеция приводит к формированию гипервирулентного фенотипа (Forrellad *et al.*, 2013). Полученные данные так же могут служить основой для объяснения механизмов быстрого возникновения лекарственной устойчивости у штаммов Beijing B0/W148. В штаммах B0/W148 отмечен крайне низкий уровень белка SseA [EC: 2.8.1.1 2.8.1.2], который вероятно приводит к накоплению АФК, и последующему повреждению ДНК. Что в свою очередь создает потенциал для создания широкого спектра генетических вариантов, способствующих

выживанию популяций *M. tuberculosis* под действием позитивной селекции, в частности, при лекарственной терапии.

ЗАКЛЮЧЕНИЕ

В России туберкулез является одной из наиболее значимых проблем здравоохранения и социального развития страны. Только за последний год, по данным Всемирной организации здравоохранения, в стране было выявлено около 100 тысяч вновь заболевших (WHO, 2015). Значительную роль в обострении ситуации с туберкулезом играет появление и все большее распространение определенных генетических семейств микобактерий, обладающих рядом уникальных особенностей. Показано, что штаммы Beijing превалируют в структуре популяции возбудителя туберкулеза в России и ассоциированы с лекарственной устойчивостью (Нарвская с соавт., 1999). В свою очередь, кластер Beijing B0/W148 является одним из наиболее распространенным в России вариантом генотипа Beijing (Нарвская, 2003; Toungousova *et al.*, 2002). Представители этого кластера преобладают у больных с тяжелыми формами туберкулеза, демонстрируют высокую вирулентность (Вишневецкий с соавт., 2002; Черноусова с соавт., 2008; Lopez *et al.*, 2003) и трансмиссивность (Mokrousov *et al.*, 2008), ассоциацию с лекарственной устойчивостью (Нарвская, 2003; Kruuner *et al.*, 2001; Toungousova *et al.*, 2002). Очевидно, что в сложившейся ситуации, весьма актуальны исследования, направленные на выявления особенностей данных штаммов, с целью решения проблем коррекции противотуберкулезной терапии, а также проведения адекватных противоэпидемических мероприятий, ориентированных на предупреждение распространения штаммов микобактерий.

В данном исследовании объектом системного подхода, включающего методы геномного и протеомного анализа, выбрана очень узкая группа микобактерий – кластер Beijing B0/W148, представляющая интерес, как с эпидемиологической, так и клинической точек зрения. Приложение технологии полногеномного секвенирования позволило охарактеризовать совокупность кластер-специфических мутаций, которые можно рассматривать как

специализированные маркеры для мониторинга распространения данных изолятов. Кроме того, нами описана уникальная геномная перестройка, характерная исключительно для этой группы. Стоит отметить, что это первый в мире случай регистрации и экспериментального подтверждения рекомбинационных событий для современных штаммов *M. tuberculosis*.

Проведенный сравнительный протеомный анализ штаммов кластера Beijing B0/W148 указал на ряд изменений в их белковом профиле относительно H37Rv. Системный анализ наблюдаемых отличий в представленности 266 дифференциальных белков (41 высоко- и 225 низкопредставленных) позволил предположить молекулярные механизмы известных физиологических особенностях исследуемой группы штаммов – повышенная выживаемость в макрофагах, крайне высокая встречаемость в группе лекарственноустойчивых штаммов. В частности, мы наблюдали увеличение представленности ферментов, ответственных за биосинтез длинноцепочечных жирных кислот, наряду со снижением представленности белков, ответственных за их деградацию. Микобактерии используют длинноцепочечные жирные кислоты для получения микотиоловых кислот и различных липидов, которые рассматривают как основные факторы вирулентности *M. tuberculosis*, проявляющиеся на начальных стадиях инфекции, когда бактерии попадают в макрофаг. Мы так же отметили увеличение представленности белка HsaA [EC: 1.14.14.12], вовлеченного в деградацию стероидов. Было показано, что *M. tuberculosis* используют внеклеточный холестерин в качестве источника энергии и для биосинтеза липидов клеточной стенки. Указанные наблюдения могут свидетельствовать в пользу повышенной выживаемости микобактерий в макрофагах, что является известной характеристикой штаммов кластера Beijing B0/W148. Кроме того, на основании протеомных данных нами зафиксирован очень низкий уровень белка SseA [EC: 2.8.1.1 2.8.1.2] в штаммах B0/W148, что возможно приводит к накоплению активных форм кислорода в клетке и, как следствие, повреждению ДНК. Это в свою очередь может вести к образованию широкого спектра

генетических вариантов, способствующих выживанию бактериальной клетки под действием селективного отбора, в частности во время лекарственной терапии.

Таким образом, анализ совокупно накопленной геномной и протеомной информации приближает нас к пониманию молекулярных механизмов адаптации *M. tuberculosis*. Представленные данные имеют несомненную научную ценность, как в решении прикладных задач, связанных с поиском новых средств элиминации возбудителя, так и фундаментальных, касающихся расширения знаний о физиологии и клеточной организации микобактерий.

В ходе выполнения работы проведено внедрение результатов в практику в виде двух патентов: 1. Патент на изобретение №2551764 (Ильина Е. Н., Шитиков Е. А., Беспятых Ю. А., Способ обнаружения микобактерий туберкулеза генетического кластера Beijing B0/W148). Зарегистрировано в Государственном реестре изобретений РФ 23.04.2015. Приоритет: 04.07.2013 и 2. Патент на изобретение №2548797 (Ильина Е. Н., Шитиков Е. А., Беспятых Ю. А., Способ одновременной амплификации и флуоресцентного маркирования нескольких сегментов генома микобактерий туберкулезного комплекса). Зарегистрировано в Государственном реестре изобретений РФ 24.03.2015. Приоритет: 04.07.2013.

ВЫВОДЫ

1. Кластер Beijing B0/W148 представляет обособленную генетическую группу микобактерий, которая характеризуется 59 кластер-специфическими полиморфизмами и двумя инвертированными участками в структуре генома.

2. Разработанный протокол экстракции и последующей масс-спектрометрической идентификации белков микобактерий позволяет получать данные мирового уровня, пригодные для характеристики протеома.

3. Впервые описанный протеомный профиль *M. tuberculosis* кластера Beijing B0/W148 существенно отличается от H37Rv, часть этих отличий согласуется с отличиями в структуре их геномов.

4. Изменения в представленности белков DosR регулона свидетельствуют о реализации штаммами *M. tuberculosis* кластера Beijing B0/W148 отличного от H37Rv ответа на гипоксию, что предположительно способствует их выживаемости внутри макрофагов.

5. Зафиксированы изменения в ферментах метаболизма липидов и белках транспортных систем представителей *M. tuberculosis* кластера Beijing B0/W148, которые можно рассматривать как потенциальные факторы патогенности.

Список используемых сокращений

CRISPR – Clustered Regularly Interspaced Palindromic Repeats (англ. кластеризованные короткие палиндромные повторы, разделенные спейсерными участками)

DR – Direct Repeat (англ. прямой повтор)

ETR – Exact Tandem Repeats (англ. точные тандемные повторы)

FDR – False discovery rate (англ. доля ложноположительных идентификаций)

IDA – independent data acquisition (англ. получение независимых данных)

INF- γ – interferon gamma (англ. гамма-интерферон)

IS elements – Insertion Sequence elements (англ. инсерционные элементы)

IS6110 RFLP – Restriction Fragment Length Polymorphisms IS6110 (англ. анализ полиморфизма длины рестрикционных фрагментов IS6110)

LAM – Latin American Mediterranean (генетическое семейство вида *M. tuberculosis*)

LSP – Large Sequence Polymorphism (англ. крупный нуклеотидный полиморфизм)

M. smegmatis – *Mycobacterium smegmatis*

M. tuberculosis – *Mycobacterium tuberculosis*, микобактерия туберкулеза

M. tuberculosis complex – *Mycobacterium tuberculosis* complex, микобактерии туберкулезного комплекса

m/z – отношение массы к заряду

MIRU – Mycobacterial Interspersed Repeat Units (англ. микобактериальные рассеянные повторяющиеся единицы)

MRC – membrane-associated oxidoreductase complex (англ. мембран-ассоциированный оксидоредуктазный комплекс)

MRM – multiple reaction monitoring (англ. мониторинг множественных реакций)

MS/MS – тандемная масс-спектрометрия

NGS – next-generation sequencing (англ. технологии секвенирования нового поколения)

NTF – noise transfer function

PBS – фосфатно-солевой буфер (phosphate buffered saline)

PE/PPE – белки с характерными N-терминальными мотивами ProGlu (PE) / ProProGlu (PPE)

PGG – Principal genetic group (англ. главная генетическая группа)

RD – Region of Difference (англ. регион различия)

SCG – SNP cluster groups (англ. кластер – группа на основании SNP)

SDS – sodium dodecyl sulfate (англ. додецилсульфат натрия)

SIT – Spoligotype International Type (англ. международный сполиготип)

SNP – Single Nucleotide Polymorphism (англ. Однонуклеотидный полиморфизм)

SRA – sequence readarchive (англ. архив секвенированных последовательностей)

VNTR – Variable Number of Tandem Repeats (англ. переменное количество тандемных повторов)

АФК – активные формы кислорода

ВОЗ – Всемирная Организация Здравоохранения

ВЭЖХ – высокоэффективная жидкостная хроматография

ДНК – дезоксирибонуклеиновая кислота

дНТФ – дезоксирибонуклеозидтрифосфат

ДТТ – дихлордифенилтрихлорэтан

МАЛДИ – матрично-активированная лазерная десорбция/ионизация

МЛУ – множественная лекарственная устойчивость

п.о / т.п.о. м.п.о – пар оснований / тысяч пар оснований / миллион пар оснований

ПТП – противотуберкулезный препарат

ПЦР – полимеразная цепная реакция

РНК – рибонуклеиновая кислота

РФ – Российская Федерация

ТБ – туберкулез

ТФ – транскрипционный фактор

ШЛУ – широкая лекарственная устойчивость

Список литературы

1. Вишневский Б. И., Нарвская О. В., Васильева С. Н. с соавт. Вирулентность микобактерий туберкулеза. // Проблемы туберкулеза. 2002, № 10. С. 33-36
2. Воробьев А. А., Быков А. С. // Атлас по медицинской микробиологии, вирусологии и иммунологии. - М.; МИА. 2003
3. Драбкина Р. О. // Микробиология туберкулеза, - М.: Медгиз. 1963, С. 225
4. Криг, Н. Определитель бактерий Берджи. В 2-х т. // Н. Криг, П. Снит, С. Уильямс, Э. Бок, Д. Хоулт, Р. Беркли, Д. Бун, Дж. Стейли, П. Стин: пер. с англ. – М. : МАКС Пресс, 2007
5. Министерство здравоохранения РФ. О совершенствовании противотуберкулезных мероприятий в Российской Федерации. // Приказ № 109 от 21 марта 2003
6. Нарвская О. В. Геномный полиморфизм *Mycobacterium tuberculosis* и его роль в эпидемическом процессе. Дисс. докт. мед. наук. С.-Петербург, 2003
7. Нарвская О. В., Мокроусов И. В., Оттен Т. Ф. и др. Генетическое маркирование полирезистентных штаммов *Mycobacterium tuberculosis*, выделенных на Северо-Западе России. // Проблемы туберкулеза 1999; № 3. С. 39-41
8. Норкина, О. В., Киншт, В. Н., Мокроусов, И. В., Курунов, Ю. Н., Краснов, В. А., с соавт. Генетическое разнообразие *Mycobacterium tuberculosis* и оценка факторов риска распространения заболевания туберкулезом в Сибирском регионе России методами молекулярной эпидемиологии // Молекулярная генетика микробиология и вирусология. 2003. № 3. С. 9-18
9. Поздеев О. К., Медицинская микробиология, - М.: ГЕОТАР-МЕД, 2001
10. Черноусова Л. Н., Андреевская С. Н., Смирнова Т. Г. с соавт. Биологические свойства штаммов *M. tuberculosis* кластера W. // Проблемы туберкулеза и болезней легких 2008; № 10. С. 45-50

11. Шитиков Е. А. Геномная вариабельность возбудителей лекарственно-устойчивого туберкулеза, распространенных на территории Российской Федерации. Дисс. канд. биол. наук. Москва, 2014
12. Achtman, M. Evolution, population structure, and phylogeography of genetically monomorphic bacterial pathogens // *Annu Rev Microbiol.* 2008. Vol. 62:53-70. № doi. P. 10.1146/annurev.micro.1162.081307.162832.
13. Affolabi, D., Faihun, F., Sanoussi, N., Anyo, G., Shamputa, I. C., *et al.* Possible outbreak of streptomycin-resistant *Mycobacterium tuberculosis* Beijing in Benin // *Emerg Infect Dis.* 2009. Vol. 15. № 7. P. 1123-1125. doi: 1110.3201/eid1507.080697.
14. Agerton, T. B., Valway, S. E., Blinkhorn, R. J., Shilkret, K. L., Reeves, R., *et al.* Spread of strain W, a highly drug-resistant strain of *Mycobacterium tuberculosis*, across the United States // *Clin Infect Dis.* 1999. Vol. 29. № 1. P. 85-92; discussion 93-85.
15. Ahmad, S., Mokaddas, E. and Jaber, A. A. Rapid detection of ethambutol-resistant *Mycobacterium tuberculosis* strains by PCR-RFLP targeting embB codons 306 and 497 and iniA codon 501 mutations // *Mol Cell Probes.* 2004. Vol. 18. № 5. P. 299-306.
16. Ahn, H. C., Yang, J. H., Lee, H. B., Rhee, Y. K. and Lee, Y. C. Effect of combined therapy of oral anti-tubercular agents on theophylline pharmacokinetics // *Int J Tuberc Lung Dis.* 2000. Vol. 4. № 8. P. 784-787.
17. Akhtar, S., Sarkar, S., Mishra, A. and Sarkar, D. A method to extract intact and pure RNA from mycobacteria // *Anal Biochem.* 2011. Vol. 417. № 2. P. 286-288. doi: 210.1016/j.ab.2011.1006.1028. Epub 2011 Jun 1028.
18. Almeida Da Silva, P. E. and Palomino, J. C. Molecular basis and mechanisms of drug resistance in *Mycobacterium tuberculosis*: classical and new drugs // *J Antimicrob Chemother.* 2011. Vol. 66. № 7. P. 1417-1430. doi: 1410.1093/jac/dkr1173. Epub 2011 May 1419.
19. Alonso-Rodriguez, N., Martinez-Lirola, M., Herranz, M., Sanchez-Benitez, M., Barroso, P., *et al.* Evaluation of the new advanced 15-loci MIRU-VNTR

genotyping tool in *Mycobacterium tuberculosis* molecular epidemiology studies // BMC Microbiol. 2008. Vol. 8:34. № doi. P. 10.1186/1471-2180-1188-1134.

20. Andreevskaia SN, Chernousova LN, Smirnova TG, Larionova EE, Kuz'min AV. 2007. Impact of *Mycobacterium tuberculosis* genotype on survival in mice with experimental tuberculosis. Probl. Tuberk. Bolezn. Legk. 2007(7):45–50. (In Russian.) Andreevskaia SN, Chernousova LN, Smirnova TG, Larionova EE, Kuz'min AV. 2006. Examining the *ex vivo* growth in the macrophages of *Mycobacterium tuberculosis* of various genotypic clusters. Probl. Tuberk. Bolezn. Legk. 2006(12):43–48. (In Russian.)

21. Asgharzadeh, M., Kafil, H. S., Roudsary, A. A. and Hanifi, G. R. Tuberculosis transmission in Northwest of Iran: using MIRU-VNTR, ETR-VNTR and IS6110-RFLP methods // Infect Genet Evol. 2011. Vol. 11. № 1. P. 124-131. doi: 110.1016/j.meegid.2010.1009.1013. Epub 2010 Oct 1014.

22. Azhikina, T., Skvortsov, T., Radaeva, T., Mardanov, A., Ravin, N., *et al.* A new technique for obtaining whole pathogen transcriptomes from infected host tissues // Biotechniques. 2010. Vol. 48. № 2. P. 139-144. doi: 110.2144/000113350.

23. Badillo-Lopez, C., Gonzalez-Mejia, A., Helguera-Repetto, A. C., Salas-Rangel, L. P., Rivera-Gutierrez, S., *et al.* Differential expression of dnaA and dosR genes among members of the *Mycobacterium tuberculosis* complex under oxic and hypoxic conditions // Int Microbiol. 2010. Vol. 13. № 1. P. 9-13.

24. Baker, L., Brown, T., Maiden, M. C. and Drobniowski, F. Silent nucleotide polymorphisms and a phylogeny for *Mycobacterium tuberculosis* // Emerg Infect Dis. 2004. Vol. 10. № 9. P. 1568-1577.

25. Barlow, R. E., Gascoyne-Binzi, D. M., Gillespie, S. H., Dickens, A., Qamer, S., *et al.* Comparison of variable number tandem repeat and IS6110-restriction fragment length polymorphism analyses for discrimination of high- and low-copy-number IS6110 *Mycobacterium tuberculosis* isolates // J Clin Microbiol. 2001. Vol. 39. № 7. P. 2453-2457.

26. Bashyam, M. D., Kaushal, D., Dasgupta, S. K. and Tyagi, A. K. A study of mycobacterial transcriptional apparatus: identification of novel features in promoter elements // *J Bacteriol.* 1996. Vol. 178. № 16. P. 4847-4853.
27. Behr, M. A., Wilson, M. A., Gill, W. P., Salamon, H., Schoolnik, G. K., *et al.* Comparative genomics of BCG vaccines by whole-genome DNA microarray // *Science.* 1999. Vol. 284. № 5419. P. 1520-1523.
28. Bell, C., Smith, G. T., Sweredoski, M. J. and Hess, S. Characterization of the *Mycobacterium tuberculosis* proteome by liquid chromatography mass spectrometry-based proteomics techniques: a comprehensive resource for tuberculosis research // *J Proteome Res.* 2012. Vol. 11. № 1. P. 119-130. doi: 110.1021/pr2007939. Epub 2002011 Nov 2007930.
29. Bespyatykh, J. A., Zimenkov, D. V., Shitikov, E. A., Kulagina, E. V., Lapa, S. A., *et al.* Spoligotyping of *Mycobacterium tuberculosis* complex isolates using hydrogel oligonucleotide microarrays // *Infect Genet Evol.* 2014. Vol. 26:41-6. № doi. P. 10.1016/j.meegid.2014.1004.1024. Epub 2014 May 1019.
30. Bhavsar, A. P., Auweter, S. D. and Finlay, B. B. Proteomics as a probe of microbial pathogenesis and its molecular boundaries // *Future Microbiol.* 2010. Vol. 5. № 2. P. 253-265. doi: 210.2217/fmb.2209.2114.
31. Bifani, P. J., Mathema, B., Kurepina, N. E. and Kreiswirth, B. N. Global dissemination of the *Mycobacterium tuberculosis* W-Beijing family strains // *Trends Microbiol.* 2002. Vol. 10. № 1. P. 45-52.
32. Bifani, P. J., Plikaytis, B. B., Kapur, V., Stockbauer, K., Pan, X., *et al.* Origin and interstate spread of a New York City multidrug-resistant *Mycobacterium tuberculosis* clone family // *Jama.* 1996. Vol. 275. № 6. P. 452-457.
33. Blackwood, K. S., Burdz, T. V., Turenne, C. Y., Sharma, M. K., Kabani, A. M., *et al.* Viability testing of material derived from *Mycobacterium tuberculosis* prior to removal from a containment level-III laboratory as part of a Laboratory Risk Assessment Program // *BMC Infect Dis.* 2005. Vol. 5. № 4.
34. Blackwood, K. S., Wolfe, J. N. and Kabani, A. M. Application of mycobacterial interspersed repetitive unit typing to Manitoba tuberculosis cases: can

restriction fragment length polymorphism be forgotten? // J Clin Microbiol. 2004. Vol. 42. № 11. P. 5001-5006.

35. Boshoff, H. I. and Lun, D. S. Systems biology approaches to understanding mycobacterial survival mechanisms // Drug Discov Today Dis Mech. 2010. Vol. 7. № 1. P. e75-e82.

36. Bouakaze, C., Keyser, C., Gonzalez, A., Sougakoff, W., Veziris, N., *et al.* Matrix-assisted laser desorption ionization-time of flight mass spectrometry-based single nucleotide polymorphism genotyping assay using iPLEX gold technology for identification of *Mycobacterium tuberculosis* complex species and lineages // J Clin Microbiol. 2011. Vol. 49. № 9. P. 3292-3299. doi: 3210.1128/JCM.00744-00711. Epub 02011 Jul 00746.

37. Bradford, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding // Anal Biochem. 1976. Vol. 72. № 248-254.

38. Brites, D. and Gagneux, S. Old and new selective pressures on *Mycobacterium tuberculosis* // Infect Genet Evol. 2012. Vol. 12. № 4. P. 678-685. doi: 610.1016/j.meegid.2011.1008.1010. Epub 2011 Aug 1017.

39. Brosch, R., Philipp, W. J., Stavropoulos, E., Colston, M. J., Cole, S. T., *et al.* Genomic analysis reveals variation between *Mycobacterium tuberculosis* H37Rv and the attenuated *M. tuberculosis* H37Ra strain // Infect Immun. 1999. Vol. 67. № 11. P. 5768-5774.

40. Brosch, R., Pym, A. S., Gordon, S. V. and Cole, S. T. The evolution of mycobacterial pathogenicity: clues from comparative genomics // Trends Microbiol. 2001. Vol. 9. № 9. P. 452-458.

41. Brossier, F., Sola, C., Millot, G., Jarlier, V., Veziris, N., *et al.* Comparison of a semiautomated commercial repetitive-sequence-based PCR method with spoligotyping, 24-locus mycobacterial interspersed repetitive-unit-variable-number tandem-repeat typing, and restriction fragment length polymorphism-based analysis of IS6110 for *Mycobacterium tuberculosis* typing // J Clin Microbiol. 2014. Vol. 52. № 11. P. 4082-4086. doi: 4010.1128/JCM.02226-02214. Epub 02014 Sep 02210.

42. Brown, P. O. and Botstein, D. Exploring the new world of the genome with DNA microarrays // *Nat Genet.* 1999. Vol. 21. № 1 Suppl. P. 33-37.
43. Brudey, K., Driscoll, J. R., Rigouts, L., Prodinger, W. M., Gori, A., *et al.* *Mycobacterium tuberculosis* complex genetic diversity: mining the fourth international spoligotyping database (SpolDB4) for classification, population genetics and epidemiology // *BMC Microbiol.* 2006. Vol. 6. № 23.
44. Brugarolas, P., Movahedzadeh, F., Wang, Y., Zhang, N., Bartek, I. L., *et al.* The oxidation-sensing regulator (MosR) is a new redox-dependent transcription factor in *Mycobacterium tuberculosis* // *J Biol Chem.* 2012. Vol. 287. № 45. P. 37703-37712. doi: 37710.31074/jbc.M37112.388611. Epub 382012 Sep 388618.
45. Caminero, J. A., Pena, M. J., Campos-Herrero, M. I., Rodriguez, J. C., Garcia, I., *et al.* Epidemiological evidence of the spread of a *Mycobacterium tuberculosis* strain of the Beijing genotype on Gran Canaria Island // *Am J Respir Crit Care Med.* 2001. Vol. 164. № 7. P. 1165-1170.
46. Camus, J. C., Pryor, M. J., Medigue, C. and Cole, S. T. Re-annotation of the genome sequence of *Mycobacterium tuberculosis* H37Rv // *Microbiology.* 2002. Vol. 148. № Pt 10. P. 2967-2973.
47. Cappelli, G., Volpe, E., Grassi, M., Liseo, B., Colizzi, V., *et al.* Profiling of *Mycobacterium tuberculosis* gene expression during human macrophage infection: upregulation of the alternative sigma factor G, a group of transcriptional regulators, and proteins with unknown function // *Res Microbiol.* 2006. Vol. 157. № 5. P. 445-455. Epub 2006 Jan 2013.
48. Casali, N., Nikolayevskyy, V., Balabanova, Y., Harris, S. R., Ignatyeva, O., *et al.* Evolution and transmission of drug-resistant tuberculosis in a Russian population // *Nat Genet.* 2014. Vol. 46. № 3. P. 279-286. doi: 210.1038/ng.2878. Epub 2014 Jan 1026.
49. Castanon-Arreola, M., Lopez-Vidal, Y., Espitia-Pinzon, C. and Hernandez-Pando, R. A new vaccine against tuberculosis shows greater protection in a mouse model with progressive pulmonary tuberculosis // *Tuberculosis (Edinb).* 2005. Vol. 85. № 1-2. P. 115-126. Epub 2005 Jan 2022.

50. Caws, M., Thwaites, G., Dunstan, S., Hawn, T. R., Lan, N. T., *et al.* The influence of host and bacterial genotype on the development of disseminated disease with *Mycobacterium tuberculosis* // PLoS Pathog. 2008. Vol. 4. № 3. P. e1000034. doi: 1000010.1001371/journal.ppat.1000034.
51. Chakravorty, S., Kothari, H., Aladegbami, B., Cho, E. J., Lee, J. S., *et al.* Rapid, high-throughput detection of rifampin resistance and heteroresistance in *Mycobacterium tuberculosis* by use of sloppy molecular beacon melting temperature coding // J Clin Microbiol. 2012. Vol. 50. № 7. P. 2194-2202. doi: 2110.1128/JCM.00143-00112. Epub 02012 Apr 00125.
52. Chernousova, L. N., Andreevskaja, S. N., Smirnova, T. G., Zemskova, Z. S. and Larionova, E. E. [Biological properties of M. Tuberculosis W cluster strains] // Probl Tuberk Bolezn Legk. 2008. Vol. № 10. P. 45-50.
53. Chernyaeva, E. N., Shulgina, M. V., Rotkevich, M. S., Dobrynin, P. V., Simonov, S. A., *et al.* Genome-wide *Mycobacterium tuberculosis* variation (GMTV) database: a new tool for integrating sequence variations and epidemiology // BMC Genomics. 2014. Vol. 15:308. № doi. P. 10.1186/1471-2164-1115-1308.
54. Cho, S. H., Goodlett, D. and Franzblau, S. ICAT-based comparative proteomic analysis of non-replicating persistent *Mycobacterium tuberculosis* // Tuberculosis (Edinb). 2006. Vol. 86. № 6. P. 445-460. Epub 2005 Dec 2020.
55. Cho, Y. J., Yi, H., Chun, J., Cho, S. N., Daley, C. L., *et al.* The genome sequence of 'Mycobacterium massiliense' strain CIP 108297 suggests the independent taxonomic status of the *Mycobacterium abscessus* complex at the subspecies level // PLoS One. 2013. Vol. 8. № 11. P. e81560. doi: 81510.81371/journal.pone.0081560. eCollection 0082013.
56. Choo, S. W., Wee, W. Y., Ngeow, Y. F., Mitchell, W., Tan, J. L., *et al.* Genomic reconnaissance of clinical isolates of emerging human pathogen *Mycobacterium abscessus* reveals high evolutionary potential // Sci Rep. 2014. Vol. 4:4061. № doi. P. 10.1038/srep04061.
57. Chopra, T., Hamelin, R., Armand, F., Chiappe, D., Moniatte, M., *et al.* Quantitative mass spectrometry reveals plasticity of metabolic networks in

Mycobacterium smegmatis // Mol Cell Proteomics. 2014. Vol. 13. № 11. P. 3014-3028. doi: 3010.1074/mcp.M3113.034082. Epub 032014 Jul 034085.

58. Cole, S. T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., *et al.* Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence // Nature. 1998. Vol. 393. № 6685. P. 537-544.

59. Coll, F., McNerney, R., Guerra-Assuncao, J. A., Glynn, J. R., Perdigo, J., *et al.* A robust SNP barcode for typing *Mycobacterium tuberculosis* complex strains // Nat Commun. 2014. Vol. 5:4812. № doi. P. 10.1038/ncomms5812.

60. Coll, F., McNerney, R., Preston, M. D., Guerra-Assuncao, J. A., Warry, A., *et al.* Rapid determination of anti-tuberculosis drug resistance from whole-genome sequences // Genome Med. 2015. Vol. 7. № 1. P. 51. doi: 10.1186/s13073-13015-10164-13070. eCollection 12015.

61. Comas, I., Chakravarti, J., Small, P. M., Galagan, J., Niemann, S., *et al.* Human T cell epitopes of *Mycobacterium tuberculosis* are evolutionarily hyperconserved // Nat Genet. 2010. Vol. 42. № 6. P. 498-503. doi: 10.1038/ng.1590. Epub 2010 May 1023.

62. Comas, I., Coscolla, M., Luo, T., Borrell, S., Holt, K. E., *et al.* Out-of-Africa migration and Neolithic coexpansion of *Mycobacterium tuberculosis* with modern humans // Nat Genet. 2013. Vol. 45. № 10. P. 1176-1182. doi: 10.1038/ng.2744. Epub 2013 Sep 1171.

63. Cordone, A., Audrain, B., Calabrese, I., Euphrasie, D. and Reyrat, J. M. Characterization of a *Mycobacterium smegmatis* *uvrA* mutant impaired in dormancy induced by hypoxia and low carbon concentration // BMC Microbiol. 2011. Vol. 11:231. № doi. P. 10.1186/1471-2180-1111-231.

64. Cortes, T., Schubert, O. T., Rose, G., Arnvig, K. B., Comas, I., *et al.* Genome-wide mapping of transcriptional start sites defines an extensive leaderless transcriptome in *Mycobacterium tuberculosis* // Cell Rep. 2013. Vol. 5. № 4. P. 1121-1131. doi: 10.1016/j.celrep.2013.1110.1031. Epub 2013 Nov 1121.

65. Covert, B. A., Spencer, J. S., Orme, I. M. and Belisle, J. T. The application of proteomics in defining the T cell antigens of *Mycobacterium tuberculosis* // Proteomics. 2001. Vol. 1. № 4. P. 574-586.
66. Cowan, L. S., Diem, L., Monson, T., Wand, P., Temporado, D., *et al.* Evaluation of a two-step approach for large-scale, prospective genotyping of *Mycobacterium tuberculosis* isolates in the United States // J Clin Microbiol. 2005. Vol. 43. № 2. P. 688-695.
67. Cowan, L. S., Mosher, L., Diem, L., Massey, J. P. and Crawford, J. T. Variable-number tandem repeat typing of *Mycobacterium tuberculosis* isolates with low copy numbers of IS6110 by using mycobacterial interspersed repetitive units // J Clin Microbiol. 2002. Vol. 40. № 5. P. 1592-1602.
68. Cox, H. S., Kubica, T., Doshetov, D., Kebede, Y., Rusch-Gerdess, S., *et al.* The Beijing genotype and drug resistant tuberculosis in the Aral Sea region of Central Asia // Respir Res. 2005. Vol. 6. № 134.
69. Crofton, J. and Mitchison, D. A. Streptomycin resistance in pulmonary tuberculosis // Br Med J. 1948. Vol. 2. № 4588. P. 1009-1015.
70. Dale, J. W., Brittain, D., Cataldi, A. A., Cousins, D., Crawford, J. T., *et al.* Spacer oligonucleotide typing of bacteria of the *Mycobacterium tuberculosis* complex: recommendations for standardised nomenclature // Int J Tuberc Lung Dis. 2001. Vol. 5. № 3. P. 216-219.
71. Darling, A. C., Mau, B., Blattner, F. R. and Perna, N. T. Mauve: multiple alignment of conserved genomic sequence with rearrangements // Genome Res. 2004. Vol. 14. № 7. P. 1394-1403.
72. de Beer, J. L., Akkerman, O. W., Schurch, A. C., Mulder, A., van der Werf, T. S., *et al.* Optimization of standard in-house 24-locus variable-number tandem-repeat typing for *Mycobacterium tuberculosis* and its direct application to clinical material // J Clin Microbiol. 2014. Vol. 52. № 5. P. 1338-1342. doi: 1310.1128/JCM.03436-03413. Epub 02014 Feb 03435.
73. de Keijzer, J., de Haas, P. E., de Ru, A. H., van Veelen, P. A. and van Soolingen, D. Disclosure of selective advantages in the "modern" sublineage of the

Mycobacterium tuberculosis Beijing genotype family by quantitative proteomics // Mol Cell Proteomics. 2014. Vol. 13. № 10. P. 2632-2645. doi: 2610.1074/mcp.M2114.038380. Epub 032014 Jul 038314.

74. de Souza, G. A., Fortuin, S., Aguilar, D., Pando, R. H., McEvoy, C. R., *et al.* Using a label-free proteomics method to identify differentially abundant proteins in closely related hypo- and hypervirulent clinical *Mycobacterium tuberculosis* Beijing isolates // Mol Cell Proteomics. 2010. Vol. 9. № 11. P. 2414-2423. doi: 2410.1074/mcp.M900422-MCP900200. Epub 902010 Feb 900426.

75. Dehal, P. S., Joachimiak, M. P., Price, M. N., Bates, J. T., Baumohl, J. K., *et al.* MicrobesOnline: an integrated portal for comparative and functional genomics // Nucleic Acids Res. 2010. Vol. 38. № Database issue. P. D396-400. doi: 310.1093/nar/gkp1919. Epub 2009 Nov 1011.

76. Delcher, A. L., Phillippy, A., Carlton, J. and Salzberg, S. L. Fast algorithms for large-scale genome alignment and comparison // Nucleic Acids Res. 2002. Vol. 30. № 11. P. 2478-2483.

77. Demay, C., Liens, B., Burguiere, T., Hill, V., Couvin, D., *et al.* SITVITWEB--a publicly available international multimarker database for studying *Mycobacterium tuberculosis* genetic diversity and molecular epidemiology // Infect Genet Evol. 2012. Vol. 12. № 4. P. 755-766. doi: 710.1016/j.meegid.2012.1002.1004. Epub 2012 Feb 1017.

78. Domenech, P., Kolly, G. S., Leon-Solis, L., Fallow, A. and Reed, M. B. Massive gene duplication event among clinical isolates of the *Mycobacterium tuberculosis* W/Beijing family // J Bacteriol. 2010. Vol. 192. № 18. P. 4562-4570. doi: 4510.1128/JB.00536-00510. Epub 02010 Jul 00516.

79. Dormans, J., Burger, M., Aguilar, D., Hernandez-Pando, R., Kremer, K., *et al.* Correlation of virulence, lung pathology, bacterial load and delayed type hypersensitivity responses after infection with different *Mycobacterium tuberculosis* genotypes in a BALB/c mouse model // Clin Exp Immunol. 2004. Vol. 137. № 3. P. 460-468.

80. Dresen, C., Lin, L. Y., D'Angelo, I., Tocheva, E. I., Strynadka, N., *et al.* A flavin-dependent monooxygenase from *Mycobacterium tuberculosis* involved in cholesterol catabolism // *J Biol Chem.* 2010. Vol. 285. № 29. P. 22264-22275. doi: 22210.21074/jbc.M22109.099028. Epub 092010 May 099026.
81. Drobniewski, F., Balabanova, Y., Nikolayevsky, V., Ruddy, M., Kuznetsov, S., *et al.* Drug-resistant tuberculosis, clinical virulence, and the dominance of the Beijing strain family in Russia // *Jama.* 2005. Vol. 293. № 22. P. 2726-2731.
82. Dubiley, S., Ignatova, A., Mukhina, T., Nizova, A., Blagodatskikh, S., *et al.* Molecular epidemiology of tuberculosis in the Tula area, Central Russia, before the introduction of the Directly Observed Therapy Strategy // *Clin Microbiol Infect.* 2010. Vol. 16. № 9. P. 1421-1426. doi: 1410.1111/j.1469-0691.2009.03102.x.
83. Dymova, M. A., Cherednichenko, A. G., Alkhovik, O. I., Khrapov, E. A., Petrenko, T. I., *et al.* Characterization of extensively drug-resistant *Mycobacterium tuberculosis* isolates circulating in Siberia // *BMC Infect Dis.* 2014. Vol. 14:478. № doi. P. 10.1186/1471-2334-1114-478.
84. Ebrahimi-Rad, M., Bifani, P., Martin, C., Kremer, K., Samper, S., *et al.* Mutations in putative mutator genes of *Mycobacterium tuberculosis* strains of the W-Beijing family // *Emerg Infect Dis.* 2003. Vol. 9. № 7. P. 838-845.
85. Eisenach, K. D., Crawford, J. T. and Bates, J. H. Genetic relatedness among strains of the *Mycobacterium tuberculosis* complex. Analysis of restriction fragment heterogeneity using cloned DNA probes // *Am Rev Respir Dis.* 1986. Vol. 133. № 6. P. 1065-1068. PMID- 3013055
86. Eisenach, K. D., Crawford, J. T. and Bates, J. H. Repetitive DNA sequences as probes for *Mycobacterium tuberculosis* // *J Clin Microbiol.* 1988. Vol. 26. № 11. P. 2240-2245.
87. Eisenreich, W., Dandekar, T., Heesemann, J. and Goebel, W. Carbon metabolism of intracellular bacterial pathogens and possible links to virulence // *Nat Rev Microbiol.* 2010. Vol. 8. № 6. P. 401-412. doi: 410.1038/nrmicro2351. Epub 2010 May 1010.

88. Fallow, A., Domenech, P. and Reed, M. B. Strains of the East Asian (W/Beijing) lineage of *Mycobacterium tuberculosis* are DosS/DosT-DosR two-component regulatory system natural mutants // J Bacteriol. 2010. Vol. 192. № 8. P. 2228-2238. doi: 2210.1128/JB.01597-01509. Epub 02010 Feb 01512.
89. Farhat, M. R., Shapiro, B. J., Kieser, K. J., Sultana, R., Jacobson, K. R., *et al.* Genomic analysis identifies targets of convergent positive selection in drug-resistant *Mycobacterium tuberculosis* // Nat Genet. 2013. Vol. 45. № 10. P. 1183-1189. doi: 1110.1038/ng.2747. Epub 2013 Sep 1181.
90. Feltcher, M. E., Gunawardena, H. P., Zulauf, K. E., Malik, S., Griffin, J. E., *et al.* Label-free Quantitative Proteomics Reveals a Role for the *Mycobacterium tuberculosis* SecA2 Pathway in Exporting Solute Binding Proteins and Mce Transporters to the Cell Wall // Mol Cell Proteomics. 2015. Vol. 14. № 6. P. 1501-1516. doi: 1510.1074/mcp.M1114.044685. Epub 042015 Mar 044626.
91. Feng, Y., Zhang, Y., Ying, C., Wang, D. and Du, C. Nanopore-based fourth-generation DNA sequencing technology // Genomics Proteomics Bioinformatics. 2015. Vol. 13. № 1. P. 4-16. doi: 10.1016/j.gpb.2015.1001.1009. Epub 2015 Mar 1012.
92. Filliol, I., Motiwala, A. S., Cavatore, M., Qi, W., Hazbon, M. H., *et al.* Global phylogeny of *Mycobacterium tuberculosis* based on single nucleotide polymorphism (SNP) analysis: insights into tuberculosis evolution, phylogenetic accuracy of other DNA fingerprinting systems, and recommendations for a minimal standard SNP set // J Bacteriol. 2006. Vol. 188. № 2. P. 759-772.
93. Fleischmann, R. D., Alland, D., Eisen, J. A., Carpenter, L., White, O., *et al.* Whole-genome comparison of *Mycobacterium tuberculosis* clinical and laboratory strains // J Bacteriol. 2002. Vol. 184. № 19. P. 5479-5490.
94. Ford, C. B., Shah, R. R., Maeda, M. K., Gagneux, S., Murray, M. B., *et al.* *Mycobacterium tuberculosis* mutation rate estimates from different lineages predict substantial differences in the emergence of drug-resistant tuberculosis // Nat Genet. 2013. Vol. 45. № 7. P. 784-790. doi: 710.1038/ng.2656. Epub 2013 Jun 1039.

95. Forrellad, M. A., Klepp, L. I., Gioffre, A., Sabio y Garcia, J., Morbidoni, H. R., *et al.* Virulence factors of the *Mycobacterium tuberculosis* complex // *Virulence*. 2013. Vol. 4. № 1. P. 3-66. doi: 10.4161/viru.22329. Epub 22012 Oct 22317.
96. Frieden, T. R., Sherman, L. F., Maw, K. L., Fujiwara, P. I., Crawford, J. T., *et al.* A multi-institutional outbreak of highly drug-resistant tuberculosis: epidemiology and clinical outcomes // *Jama*. 1996. Vol. 276. № 15. P. 1229-1235.
97. Frothingham, R. and Meeker-O'Connell, W. A. Genetic diversity in the *Mycobacterium tuberculosis* complex based on variable numbers of tandem DNA repeats // *Microbiology*. 1998. Vol. 144. № Pt 5. P. 1189-1196.
98. Fu, L. M. and Shinnick, T. M. Understanding the action of INH on a highly INH-resistant *Mycobacterium tuberculosis* strain using Genechips // *Tuberculosis (Edinb)*. 2007. Vol. 87. № 1. P. 63-70. Epub 2006 Aug 2002.
99. Gagneux, S. and Small, P. M. Global phylogeography of *Mycobacterium tuberculosis* and implications for tuberculosis product development // *Lancet Infect Dis*. 2007. Vol. 7. № 5. P. 328-337.
100. Gagneux, S., DeRiemer, K., Van, T., Kato-Maeda, M., de Jong, B. C., *et al.* Variable host-pathogen compatibility in *Mycobacterium tuberculosis* // *Proc Natl Acad Sci U S A*. 2006. Vol. 103. № 8. P. 2869-2873. Epub 2006 Feb 2813.
101. Galagan, J. E., Minch, K., Peterson, M., Lyubetskaya, A., Azizi, E., *et al.* The *Mycobacterium tuberculosis* regulatory network and hypoxia // *Nature*. 2013. Vol. 499. № 7457. P. 178-183. doi: 110.1038/nature12337. Epub 12013 Jul 12333.
102. Gazi, M. A., Islam, M. R., Kibria, M. G. and Mahmud, Z. General and advanced diagnostic tools to detect *Mycobacterium tuberculosis* and their drug susceptibility: a review // *Eur J Clin Microbiol Infect Dis*. 2015. Vol. 34. № 5. P. 851-861. doi: 810.1007/s10096-10014-12306-10095. Epub 12015 Jan 10096.
103. Ghodbane, R., Asmar, S., Betzner, M., Linet, M., Pierquin, J., *et al.* Rapid Diagnosis of Tuberculosis by Real-Time High-Resolution Imaging of *Mycobacterium tuberculosis* Colonies // *J Clin Microbiol*. 2015. Vol. 53. № 8. P. 2693-2696. doi: 2610.1128/JCM.00684-00615. Epub 02015 Jun 00617.

104. Gillet, L. C., Navarro, P., Tate, S., Rost, H., Selevsek, N., *et al.* Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis // *Mol Cell Proteomics*. 2012. Vol. 11. № 6. P. O111.016717. doi: 016710.011074/mcp.O016111.016717. Epub 012012 Jan 016718.
105. Godreuil, S., Torrea, G., Terru, D., Chevenet, F., Diagbouga, S., *et al.* First molecular epidemiology study of *Mycobacterium tuberculosis* in Burkina Faso // *J Clin Microbiol*. 2007. Vol. 45. № 3. P. 921-927. Epub 2007 Jan 2024.
106. Grode, L., Seiler, P., Baumann, S., Hess, J., Brinkmann, V., *et al.* Increased vaccine efficacy against tuberculosis of recombinant *Mycobacterium bovis* bacille Calmette-Guerin mutants that secrete listeriolysin // *J Clin Invest*. 2005. Vol. 115. № 9. P. 2472-2479. Epub 2005 Aug 2418.
107. Gryadunov, D., Mikhailovich, V., Lapa, S., Roudinskii, N., Donnikov, M., *et al.* Evaluation of hybridisation on oligonucleotide microarrays for analysis of drug-resistant *Mycobacterium tuberculosis* // *Clin Microbiol Infect*. 2005. Vol. 11. № 7. P. 531-539.
108. Gunawardena, H. P., Feltcher, M. E., Wrobel, J. A., Gu, S., Braunstein, M., *et al.* Comparison of the membrane proteome of virulent *Mycobacterium tuberculosis* and the attenuated *Mycobacterium bovis* BCG vaccine strain by label-free quantitative proteomics // *J Proteome Res*. 2013. Vol. 12. № 12. P. 5463-5474. doi: 5410.1021/pr400334k. Epub 402013 Oct 400328.
109. Gutacker, M. M., Mathema, B., Soini, H., Shashkina, E., Kreiswirth, B. N., *et al.* Single-nucleotide polymorphism-based population genetic analysis of *Mycobacterium tuberculosis* strains from 4 geographic sites // *J Infect Dis*. 2006. Vol. 193. № 1. P. 121-128. Epub 2005 Nov 2028.
110. Gutierrez, M. C., Brisse, S., Brosch, R., Fabre, M., Omais, B., *et al.* Ancient origin and gene mosaicism of the progenitor of *Mycobacterium tuberculosis* // *PLoS Pathog*. 2005. Vol. 1. № 1. P. e5. Epub 2005 Aug 2019.

111. Gygi, S. P., Rist, B., Gerber, S. A., Turecek, F., Gelb, M. H., *et al.* Quantitative analysis of complex protein mixtures using isotope-coded affinity tags // *Nat Biotechnol.* 1999. Vol. 17. № 10. P. 994-999.
112. Hakkinen, L., Uitto, V. J. and Larjava, H. Cell biology of gingival wound healing // *Periodontol.* 2000. Vol. 24. № 127-152.
113. Hanekom, M., Gey van Pittius, N. C., McEvoy, C., Victor, T. C., Van Helden, P. D., *et al.* *Mycobacterium tuberculosis* Beijing genotype: a template for success // *Tuberculosis (Edinb).* 2011. Vol. 91. № 6. P. 510-523. doi: 510.1016/j.tube.2011.1007.1005. Epub 2011 Aug 1010.
114. Hanekom, M., van der Spuy, G. D., Gey van Pittius, N. C., McEvoy, C. R., Hoek, K. G., *et al.* Discordance between mycobacterial interspersed repetitive-unit-variable-number tandem-repeat typing and IS6110 restriction fragment length polymorphism genotyping for analysis of *Mycobacterium tuberculosis* Beijing strains in a setting of high incidence of tuberculosis // *J Clin Microbiol.* 2008. Vol. 46. № 10. P. 3338-3345. doi: 3310.1128/JCM.00770-00708. Epub 02008 Aug 00720.
115. Hawkey, P. M., Smith, E. G., Evans, J. T., Monk, P., Bryan, G., *et al.* Mycobacterial interspersed repetitive unit typing of *Mycobacterium tuberculosis* compared to IS6110-based restriction fragment length polymorphism analysis for investigation of apparently clustered cases of tuberculosis // *J Clin Microbiol.* 2003. Vol. 41. № 8. P. 3514-3520.
116. He, H., Bretl, D. J., Penoske, R. M., Anderson, D. M. and Zahrt, T. C. Components of the Rv0081-Rv0088 locus, which encodes a predicted formate hydrogenlyase complex, are coregulated by Rv0081, MprA, and DosR in *Mycobacterium tuberculosis* // *J Bacteriol.* 2011. Vol. 193. № 19. P. 5105-5118. doi: 5110.1128/JB.05562-05511. Epub 02011 Aug 05565.
117. He, X. Y., Zhuang, Y. H., Zhang, X. G. and Li, G. L. Comparative proteome analysis of culture supernatant proteins of *Mycobacterium tuberculosis* H37Rv and H37Ra // *Microbes Infect.* 2003. Vol. 5. № 10. P. 851-856.
118. Herrera-Leon, L., Molina, T., Saiz, P., Saez-Nieto, J. A. and Jimenez, M. S. New multiplex PCR for rapid detection of isoniazid-resistant *Mycobacterium*

tuberculosis clinical isolates // Antimicrob Agents Chemother. 2005. Vol. 49. № 1. P. 144-147.

119. Hirsh, A. E., Tsolaki, A. G., DeRiemer, K., Feldman, M. W. and Small, P. M. Stable association between strains of *Mycobacterium tuberculosis* and their human host populations // Proc Natl Acad Sci U S A. 2004. Vol. 101. № 14. P. 4871-4876. Epub 2004 Mar 4823.

120. Homolka, S., Niemann, S., Russell, D. G. and Rohde, K. H. Functional genetic diversity among *Mycobacterium tuberculosis* complex clinical isolates: delineation of conserved core and lineage-specific transcriptomes during intracellular survival // PLoS Pathog. 2010. Vol. 6. № 7. P. e1000988. doi: 1000910.1001371/journal.ppat.1000988.

121. Homolka, S., Projahn, M., Feuerriegel, S., Ubben, T., Diel, R., *et al.* High resolution discrimination of clinical *Mycobacterium tuberculosis* complex strains based on single nucleotide polymorphisms // PLoS One. 2012. Vol. 7. № 7. P. e39855. doi: 39810.31371/journal.pone.0039855. Epub 0032012 Jul 0039852.

122. Honisch, C., Mosko, M., Arnold, C., Gharbia, S. E., Diel, R., *et al.* Replacing reverse line blot hybridization spoligotyping of the *Mycobacterium tuberculosis* complex // J Clin Microbiol. 2010. Vol. 48. № 5. P. 1520-1526. doi: 1510.1128/JCM.02299-02209. Epub 02010 Mar 02293.

123. Hsu, C. Y., Wu, C. W. and Talaat, A. M. Genome-Wide Sequence Variation among *Mycobacterium avium* Subspecies paratuberculosis Isolates: A Better Understanding of Johne's Disease Transmission Dynamics // Front Microbiol. 2011. Vol. 2:236. № doi. P. 10.3389/fmicb.2011.00236. eCollection 02011.

124. Ilina, E. N., Shitikov, E. A., Ikryannikova, L. N., Alekseev, D. G., Kamashev, D. E., *et al.* Comparative genomic analysis of *Mycobacterium tuberculosis* drug resistant strains from Russia // PLoS One. 2013. Vol. 8. № 2. P. e56577. doi: 56510.51371/journal.pone.0056577. Epub 0052013 Feb 0056520.

125. Ioerger, T. R., Koo, S., No, E. G., Chen, X., Larsen, M. H., *et al.* Genome analysis of multi- and extensively-drug-resistant tuberculosis from KwaZulu-Natal,

South Africa // PLoS One. 2009. Vol. 4. № 11. P. e7778. doi: 7710.1371/journal.pone.0007778.

126. Ishihama, Y., Oda, Y., Tabata, T., Sato, T., Nagasu, T., *et al.* Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein // Mol Cell Proteomics. 2005. Vol. 4. № 9. P. 1265-1272. Epub 2005 Jun 12 14.

127. Jain, R., Kumar, P. and Varshney, U. A distinct role of formamidopyrimidine DNA glycosylase (MutM) in down-regulation of accumulation of G, C mutations and protection against oxidative stress in mycobacteria // DNA Repair (Amst). 2007. Vol. 6. № 12. P. 1774-1785. Epub 2007 Aug 17 16.

128. James, P. Protein identification in the post-genome era: the rapid rise of proteomics // Q Rev Biophys. 1997. Vol. 30. № 4. P. 279-331.

129. Jeon, B. Y., Derrick, S. C., Lim, J., Kolibab, K., Dheenadhayalan, V., *et al.* *Mycobacterium bovis* BCG immunization induces protective immunity against nine different *Mycobacterium tuberculosis* strains in mice // Infect Immun. 2008. Vol. 76. № 11. P. 5173-5180. doi: 5110.1128/IAI.00019-00008. Epub 02008 Aug 00018.

130. Jungblut, P. R., Schaible, U. E., Mollenkopf, H. J., Zimny-Arndt, U., Raupach, B., *et al.* Comparative proteome analysis of *Mycobacterium tuberculosis* and *Mycobacterium bovis* BCG strains: towards functional genomics of microbial pathogens // Mol Microbiol. 1999. Vol. 33. № 6. P. 1103-1117.

131. Kamerbeek, J., Schouls, L., Kolk, A., van Agterveld, M., van Soolingen, D., *et al.* Simultaneous detection and strain differentiation of *Mycobacterium tuberculosis* for diagnosis and epidemiology // J Clin Microbiol. 1997. Vol. 35. № 4. P. 907-914.

132. Kanduma, E., McHugh, T. D. and Gillespie, S. H. Molecular methods for *Mycobacterium tuberculosis* strain typing: a users guide // J Appl Microbiol. 2003. Vol. 94. № 5. P. 781-791

133. Karakousis, P. C., Yoshimatsu, T., Lamichhane, G., Woolwine, S. C., Nueremberger, E. L., *et al.* Dormancy phenotype displayed by extracellular

Mycobacterium tuberculosis within artificial granulomas in mice // J Exp Med. 2004. Vol. 200. № 5. P. 647-657.

134. Karboul, A., Gey van Pittius, N. C., Namouchi, A., Vincent, V., Sola, C., *et al.* Insights into the evolutionary history of tubercle bacilli as disclosed by genetic rearrangements within a PE_PGRS duplicated gene pair // BMC Evol Biol. 2006. Vol. 6. № 107.

135. Koboldt, D. C., Zhang, Q., Larson, D. E., Shen, D., McLellan, M. D., *et al.* VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing // Genome Res. 2012. Vol. 22. № 3. P. 568-576. doi: 510.1101/gr.129684.129111. Epub 122012 Feb 129682.

136. Kremer, K., Glynn, J. R., Lillebaek, T., Niemann, S., Kurepina, N. E., *et al.* Definition of the Beijing/W lineage of *Mycobacterium tuberculosis* on the basis of genetic markers // J Clin Microbiol. 2004. Vol. 42. № 9. P. 4040-4049.

137. Kremer, K., van Soolingen, D., Frothingham, R., Haas, W. H., Hermans, P. W., *et al.* Comparison of methods based on different molecular epidemiological markers for typing of *Mycobacterium tuberculosis* complex strains: interlaboratory study of discriminatory power and reproducibility // J Clin Microbiol. 1999. Vol. 37. № 8. P. 2607-2618.

138. Kruuner, A., Hoffner, S. E., Sillastu, H., Danilovits, M., Levina, K., *et al.* Spread of drug-resistant pulmonary tuberculosis in Estonia // J Clin Microbiol. 2001. Vol. 39. № 9. P. 3339-3345.

139. Kubica, T., Rusch-Gerdes, S. and Niemann, S. The Beijing genotype is emerging among multidrug-resistant *Mycobacterium tuberculosis* strains from Germany // Int J Tuberc Lung Dis. 2004. Vol. 8. № 9. P. 1107-1113.

140. Kumar, B., Sharma, D., Sharma, P., Katoch, V. M., Venkatesan, K., *et al.* Proteomic analysis of *Mycobacterium tuberculosis* isolates resistant to kanamycin and amikacin // J Proteomics. 2013. Vol. 94:68-77. № doi. P. 10.1016/j.jprot.2013.1008.1025. Epub 2013 Sep 1012.

141. Kurepina, N. E., Sreevatsan, S., Plikaytis, B. B., Bifani, P. J., Connell, N. D., *et al.* Characterization of the phylogenetic distribution and chromosomal insertion

sites of five IS6110 elements in *Mycobacterium tuberculosis*: non-random integration in the dnaA-dnaN region // *Tuber Lung Dis.* 1998. Vol. 79. № 1. P. 31-42.

142. Laemmli, U. K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4 // *Nature.* 1970. Vol. 227. № 5259. P. 680-685.

143. Lange, V., Picotti, P., Domon, B. and Aebersold, R. Selected reaction monitoring for quantitative proteomics: a tutorial // *Mol Syst Biol.* 2008. Vol. 4:222. № doi. P. 10.1038/msb.2008.1061. Epub 2008 Oct 1014.

144. Langmead, B. and Salzberg, S. L. Fast gapped-read alignment with Bowtie 2 // *Nat Methods.* 2012. Vol. 9. № 4. P. 357-359. doi: 310.1038/nmeth.1923.

145. Lasunskaja, E., Ribeiro, S. C., Manicheva, O., Gomes, L. L., Suffys, P. N., *et al.* Emerging multidrug resistant *Mycobacterium tuberculosis* strains of the Beijing genotype circulating in Russia express a pattern of biological properties associated with enhanced virulence // *Microbes Infect.* 2010. Vol. 12. № 6. P. 467-475. doi: 410.1016/j.micinf.2010.1002.1008. Epub 2010 Mar 1017.

146. Lee, A. S., Tang, L. L., Lim, I. H., Bellamy, R. and Wong, S. Y. Discrimination of single-copy IS6110 DNA fingerprints of *Mycobacterium tuberculosis* isolates by high-resolution minisatellite-based typing // *J Clin Microbiol.* 2002. Vol. 40. № 2. P. 657-659.

147. Leistikow, R. L., Morton, R. A., Bartek, I. L., Frimpong, I., Wagner, K., *et al.* The *Mycobacterium tuberculosis* DosR regulon assists in metabolic homeostasis and enables rapid recovery from nonrespiring dormancy // *J Bacteriol.* 2010. Vol. 192. № 6. P. 1662-1670. doi: 1610.1128/JB.00926-00909. Epub 02009 Dec 00918.

148. Li, A. H., Waddell, S. J., Hinds, J., Malloff, C. A., Bains, M., *et al.* Contrasting transcriptional responses of a virulent and an attenuated strain of *Mycobacterium tuberculosis* infecting macrophages // *PLoS One.* 2010. Vol. 5. № 6. P. e11066. doi: 11010.11371/journal.pone.0011066.

149. Li, H. T., Zhang, T. T., Zhou, Y. Q., Huang, Q. H. and Huang, J. SLC11A1 (formerly NRAMP1) gene polymorphisms and tuberculosis susceptibility: a meta-analysis // *Int J Tuberc Lung Dis.* 2006. Vol. 10. № 1. P. 3-12.

150. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., *et al.* The Sequence Alignment/Map format and SAMtools // *Bioinformatics*. 2009. Vol. 25. № 16. P. 2078-2079. doi: 2010.1093/bioinformatics/btp2352. Epub 2009 Jun 2078.
151. Li, Z., Kelley, C., Collins, F., Rouse, D. and Morris, S. Expression of katG in *Mycobacterium tuberculosis* is associated with its growth and persistence in mice and guinea pigs // *J Infect Dis*. 1998. Vol. 177. № 4. P. 1030-1035.
152. Linger, Y., Kukhtin, A., Golova, J., Perov, A., Lambarqui, A., *et al.* Simplified microarray system for simultaneously detecting rifampin, isoniazid, ethambutol, and streptomycin resistance markers in *Mycobacterium tuberculosis* // *J Clin Microbiol*. 2014. Vol. 52. № 6. P. 2100-2107. doi: 2110.1128/JCM.00238-00214. Epub 02014 Apr 00239.
153. Lipin, M. Y., Stepanshina, V. N., Shemyakin, I. G. and Shinnick, T. M. Association of specific mutations in katG, rpoB, rpsL and rrs genes with spoligotypes of multidrug-resistant *Mycobacterium tuberculosis* isolates in Russia // *Clin Microbiol Infect*. 2007. Vol. 13. № 6. P. 620-626. Epub 2007 Apr 2002.
154. Liu, L., Li, Y., Li, S., Hu, N., He, Y., *et al.* Comparison of next-generation sequencing systems // *J Biomed Biotechnol*. 2012. Vol. 2012:251364. № doi. P. 10.1155/2012/251364. Epub 252012 Jul 251365.
155. Liu, Q., Chen, X., Hu, C., Zhang, R., Yue, J., *et al.* Serum protein profiling of smear-positive and smear-negative pulmonary tuberculosis using SELDI-TOF mass spectrometry // *Lung*. 2010. Vol. 188. № 1. P. 15-23. doi: 10.1007/s00408-00009-09199-00406. Epub 02009 Dec 00409.
156. Lopez, B., Aguilar, D., Orozco, H., Burger, M., Espitia, C., *et al.* A marked difference in pathogenesis and immune response induced by different *Mycobacterium tuberculosis* genotypes // *Clin Exp Immunol*. 2003. Vol. 133. № 1. P. 30-37.
157. Malen, H., De Souza, G. A., Pathak, S., Softeland, T. and Wiker, H. G. Comparison of membrane proteins of *Mycobacterium tuberculosis* H37Rv and H37Ra strains // *BMC Microbiol*. 2011. Vol. 11:18. № doi. P. 10.1186/1471-2180-1111-1118.
158. Malhotra, V., Tyagi, J. S. and Clark-Curtiss, J. E. DevR-mediated adaptive response in *Mycobacterium tuberculosis* H37Ra: links to asparagine metabolism //

Tuberculosis (Edinb). 2009. Vol. 89. № 2. P. 169-174. doi: 110.1016/j.tube.2008.1012.1003. Epub 2009 Feb 1013.

159. Manca, C., Paul, S., Barry, C. E., 3rd, Freedman, V. H. and Kaplan, G. *Mycobacterium tuberculosis* catalase and peroxidase activities and resistance to oxidative killing in human monocytes in vitro // Infect Immun. 1999. Vol. 67. № 1. P. 74-79.

160. Marttila, H. J., Soini, H., Eerola, E., Vyshnevskaya, E., Vyshnevskiy, B. I., *et al.* A Ser315Thr substitution in KatG is predominant in genetically heterogeneous multidrug-resistant *Mycobacterium tuberculosis* isolates originating from the St. Petersburg area in Russia // Antimicrob Agents Chemother. 1998. Vol. 42. № 9. P. 2443-2445.

161. Mascarello, A., Mori, M., Chiaradia-Delatorre, L. D., Menegatti, A. C., Delle Monache, F., *et al.* Discovery of *Mycobacterium tuberculosis* protein tyrosine phosphatase B (PtpB) inhibitors from natural products // PLoS One. 2013. Vol. 8. № 10. P. e77081. doi: 77010.71371/journal.pone.0077081. eCollection 0072013.

162. Matrakshin, A. G., Mes'ko, E. M., Beliakova, N. K., Andreevskaya, S. N., Smirnova, T. G., *et al.* [Genotypic characteristics of *Mycobacterium tuberculosis* strains from the Republic of Tyva] // Probl Tuberk Bolezn Legk. 2004. Vol. № 3. P. 37-40.

163. Mattow, J., Jungblut, P. R., Schaible, U. E., Mollenkopf, H. J., Lamer, S., *et al.* Identification of proteins from *Mycobacterium tuberculosis* missing in attenuated *Mycobacterium bovis* BCG strains // Electrophoresis. 2001. Vol. 22. № 14. P. 2936-2946.

164. Mattow, J., Schaible, U. E., Schmidt, F., Hagens, K., Siejak, F., *et al.* Comparative proteome analysis of culture supernatant proteins from virulent *Mycobacterium tuberculosis* H37Rv and attenuated M. bovis BCG Copenhagen // Electrophoresis. 2003. Vol. 24. № 19-20. P. 3405-3420.

165. Mattow, J., Siejak, F., Hagens, K., Becher, D., Albrecht, D., *et al.* Proteins unique to intraphagosomally grown *Mycobacterium tuberculosis* // Proteomics. 2006. Vol. 6. № 8. P. 2485-2494.

166. Mawuenyega, K. G., Forst, C. V., Dobos, K. M., Belisle, J. T., Chen, J., *et al.* *Mycobacterium tuberculosis* functional network analysis by global subcellular protein profiling // *Mol Biol Cell*. 2005. Vol. 16. № 1. P. 396-404. Epub 2004 Nov 2003.
167. Mazars, E., Lesjean, S., Banuls, A. L., Gilbert, M., Vincent, V., *et al.* High-resolution minisatellite-based typing as a portable approach to global analysis of *Mycobacterium tuberculosis* molecular epidemiology // *Proc Natl Acad Sci U S A*. 2001. Vol. 98. № 4. P. 1901-1906.
168. McEvoy, C. R., Falmer, A. A., Gey van Pittius, N. C., Victor, T. C., van Helden, P. D., *et al.* The role of IS6110 in the evolution of *Mycobacterium tuberculosis* // *Tuberculosis (Edinb)*. 2007. Vol. 87. № 5. P. 393-404. Epub 2007 Jul 2012.
169. Mehaffy, C., Hess, A., Prenni, J. E., Mathema, B., Kreiswirth, B., *et al.* Descriptive proteomic analysis shows protein variability between closely related clinical isolates of *Mycobacterium tuberculosis* // *Proteomics*. 2010. Vol. 10. № 10. P. 1966-1984. doi: 1910.1002/pmic.200900836.
170. Merker, M., Blin, C., Mona, S., Duforet-Frebourg, N., Lecher, S., *et al.* Evolutionary history and global spread of the *Mycobacterium tuberculosis* Beijing lineage // *Nat Genet*. 2015. Vol. 47. № 3. P. 242-249. doi: 210.1038/ng.3195. Epub 2015 Jan 1019.
171. Mishra, A. and Sarkar, D. Qualitative and quantitative proteomic analysis of Vitamin C induced changes in *Mycobacterium smegmatis* // *Front Microbiol*. 2015. Vol. 6:451. № doi. P. 10.3389/fmicb.2015.00451. eCollection 02015.
172. Mokrousov, I. Genetic geography of *Mycobacterium tuberculosis* Beijing genotype: a multifacet mirror of human history? // *Infect Genet Evol*. 2008. Vol. 8. № 6. P. 777-785. doi: 710.1016/j.meegid.2008.1007.1003. Epub 2008 Jul 1022.
173. Mokrousov, I. Insights into the origin, emergence, and current spread of a successful Russian clone of *Mycobacterium tuberculosis* // *Clin Microbiol Rev*. 2013. Vol. 26. № 2. P. 342-360. doi: 310.1128/CMR.00087-00012.
174. Mokrousov, I., Ly, H. M., Otten, T., Lan, N. N., Vyshnevskiy, B., *et al.* Origin and primary dispersal of the *Mycobacterium tuberculosis* Beijing genotype:

clues from human phylogeography // *Genome Res.* 2005. Vol. 15. № 10. P. 1357-1364. Epub 2005 Sep 1316.

175. Mokrousov, I., Narvskaya, O., Otten, T., Vyazovaya, A., Limeschenko, E., *et al.* Phylogenetic reconstruction within *Mycobacterium tuberculosis* Beijing genotype in northwestern Russia // *Res Microbiol.* 2002. Vol. 153. № 10. P. 629-637.

176. Mokrousov, I., Narvskaya, O., Vyazovaya, A., Millet, J., Otten, T., *et al.* *Mycobacterium tuberculosis* Beijing genotype in Russia: in search of informative variable-number tandem-repeat loci // *J Clin Microbiol.* 2008. Vol. 46. № 11. P. 3576-3584. doi: 3510.1128/JCM.00414-00408. Epub 02008 Aug 00427.

177. Mokrousov, I., Narvskaya, O., Vyazovaya, A., Otten, T., Jiao, W. W., *et al.* Russian "successful" clone B0/W148 of *Mycobacterium tuberculosis* Beijing genotype: a multiplex PCR assay for rapid detection and global screening // *J Clin Microbiol.* 2012. Vol. 50. № 11. P. 3757-3759. doi: 3710.1128/JCM.02001-02012. Epub 02012 Aug 02029.

178. Mokrousov, I., Otten, T., Vyazovaya, A., Limeschenko, E., Filipenko, M. L., *et al.* PCR-based methodology for detecting multidrug-resistant strains of *Mycobacterium tuberculosis* Beijing family circulating in Russia // *Eur J Clin Microbiol Infect Dis.* 2003. Vol. 22. № 6. P. 342-348. Epub 2003 Jun 2003.

179. Molloy, M. P. and Witzmann, F. A. Proteomics: technologies and applications // *Brief Funct Genomic Proteomic.* 2002. Vol. 1. № 1. P. 23-39.

180. Monahan, I. M., Betts, J., Banerjee, D. K. and Butcher, P. D. Differential expression of mycobacterial proteins following phagocytosis by macrophages // *Microbiology.* 2001. Vol. 147. № Pt 2. P. 459-471.

181. Monteoliva, L. and Albar, J. P. Differential proteomics: an overview of gel and non-gel based approaches // *Brief Funct Genomic Proteomic.* 2004. Vol. 3. № 3. P. 220-239.

182. Musser, J. M., Amin, A. and Ramaswamy, S. Negligible genetic diversity of *Mycobacterium tuberculosis* host immune system protein targets: evidence of limited selective pressure // *Genetics.* 2000. Vol. 155. № 1. P. 7-16.

183. Nambi, S., Long, J. E., Mishra, B. B., Baker, R., Murphy, K. C., *et al.* The Oxidative Stress Network of *Mycobacterium tuberculosis* Reveals Coordination between Radical Detoxification Systems // *Cell Host Microbe*. 2015. Vol. 17. № 6. P. 829-837. doi: 810.1016/j.chom.2015.1005.1008.
184. Naya, H., Romero, H., Zavala, A., Alvarez, B. and Musto, H. Aerobiosis increases the genomic guanine plus cytosine content (GC%) in prokaryotes // *J Mol Evol*. 2002. Vol. 55. № 3. P. 260-264.
185. Niemann, S., Diel, R., Khechinashvili, G., Gegia, M., Mdivani, N., *et al.* *Mycobacterium tuberculosis* Beijing lineage favors the spread of multidrug-resistant tuberculosis in the Republic of Georgia // *J Clin Microbiol*. 2010. Vol. 48. № 10. P. 3544-3550. doi: 3510.1128/JCM.00715-00710. Epub 02010 Aug 00711.
186. Niemann, S., Koser, C. U., Gagneux, S., Plinke, C., Homolka, S., *et al.* Genomic diversity among drug sensitive and multidrug resistant isolates of *Mycobacterium tuberculosis* with identical DNA fingerprints // *PLoS One*. 2009. Vol. 4. № 10. P. e7407. doi: 7410.1371/journal.pone.0007407.
187. Oelemann, M. C., Diel, R., Vatin, V., Haas, W., Rusch-Gerdes, S., *et al.* Assessment of an optimized mycobacterial interspersed repetitive- unit-variable-number tandem-repeat typing system combined with spoligotyping for population-based molecular epidemiology studies of tuberculosis // *J Clin Microbiol*. 2007. Vol. 45. № 3. P. 691-697. Epub 2006 Dec 2027.
188. Ong, S. E., Blagoev, B., Kratchmarova, I., Kristensen, D. B., Steen, H., *et al.* Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics // *Mol Cell Proteomics*. 2002. Vol. 1. № 5. P. 376-386.
189. Otal, I., Gomez, A. B., Kremer, K., de Haas, P., Garcia, M. J., *et al.* Mapping of IS6110 insertion sites in *Mycobacterium bovis* isolates in relation to adaptation from the animal to human host // *Vet Microbiol*. 2008. Vol. 129. № 3-4. P. 333-341. doi: 310.1016/j.vetmic.2007.1011.1038. Epub 2007 Dec 1018.
190. Palittapongarnpim, P., Rienthong, S. and Panbangred, W. Comparison of restriction fragment length polymorphism of *Mycobacterium tuberculosis* isolated from

cerebrospinal fluid and sputum: a preliminary report // *Tuber Lung Dis.* 1993. Vol. 74. № 3. P. 204-207.

191. Pandey, A. K. and Sasseti, C. M. Mycobacterial persistence requires the utilization of host cholesterol // *Proc Natl Acad Sci U S A.* 2008. Vol. 105. № 11. P. 4376-4380. doi: 4310.1073/pnas.0711159105. Epub 0711152008 Mar 0711159111.

192. Pandey, R. and Rodriguez, G. M. A ferritin mutant of *Mycobacterium tuberculosis* is highly susceptible to killing by antibiotics and is unable to establish a chronic infection in mice // *Infect Immun.* 2012. Vol. 80. № 10. P. 3650-3659. doi: 3610.1128/IAI.00229-00212. Epub 02012 Jul 00216.

193. Parandhaman, D. K., Sharma, P., Bisht, D. and Narayanan, S. Proteome and phosphoproteome analysis of the serine/threonine protein kinase E mutant of *Mycobacterium tuberculosis* // *Life Sci.* 2014. Vol. 109. № 2. P. 116-126. doi: 110.1016/j.lfs.2014.1006.1013. Epub 2014 Jun 1024.

194. Pardini, M., Niemann, S., Varaine, F., Iona, E., Meacci, F., *et al.* Characteristics of drug-resistant tuberculosis in Abkhazia (Georgia), a high-prevalence area in Eastern Europe // *Tuberculosis (Edinb).* 2009. Vol. 89. № 4. P. 317-324. doi: 310.1016/j.tube.2009.1004.1002. Epub 2009 Jun 1017.

195. Parish, T., Smith, D. A., Kendall, S., Casali, N., Bancroft, G. J., *et al.* Deletion of two-component regulatory systems increases the virulence of *Mycobacterium tuberculosis* // *Infect Immun.* 2003. Vol. 71. № 3. P. 1134-1140.

196. Park, H. D., Guinn, K. M., Harrell, M. I., Liao, R., Voskuil, M. I., *et al.* Rv3133c/dosR is a transcription factor that mediates the hypoxic response of *Mycobacterium tuberculosis* // *Mol Microbiol.* 2003. Vol. 48. № 3. P. 833-843.

197. Park, Y. K., Shin, S., Ryu, S., Cho, S. N., Koh, W. J., *et al.* Comparison of drug resistance genotypes between Beijing and non-Beijing family strains of *Mycobacterium tuberculosis* in Korea // *J Microbiol Methods.* 2005. Vol. 63. № 2. P. 165-172.

198. Peterson, E. J., Reiss, D. J., Turkarslan, S., Minch, K. J., Rustad, T., *et al.* A high-resolution network model for global gene regulation in *Mycobacterium*

tuberculosis // Nucleic Acids Res. 2014. Vol. 42. № 18. P. 11291-11303. doi: 11210.11093/nar/gku11777. Epub 12014 Sep 11217.

199. Pheiffer, C., Betts, J. C., Flynn, H. R., Lukey, P. T. and van Helden, P. Protein expression by a Beijing strain differs from that of another clinical isolate and *Mycobacterium tuberculosis* H37Rv // Microbiology. 2005. Vol. 151. № Pt 4. P. 1139-1150.

200. Piatek, A. S., Telenti, A., Murray, M. R., El-Hajj, H., Jacobs, W. R., Jr., *et al.* Genotypic analysis of *Mycobacterium tuberculosis* in two distinct populations using molecular beacons: implications for rapid susceptibility testing // Antimicrob Agents Chemother. 2000. Vol. 44. № 1. P. 103-110.

201. Poulet, S. and Cole, S. T. Characterization of the highly abundant polymorphic GC-rich-repetitive sequence (PGRS) present in *Mycobacterium tuberculosis* // Arch Microbiol. 1995. Vol. 163. № 2. P. 87-95.

202. Poulet, S. and Cole, S. T. Repeated DNA sequences in mycobacteria // Arch Microbiol. 1995. Vol. 163. № 2. P. 79-86.

203. Poynten, M., Andresen, D. N. and Gottlieb, T. Laboratory cross-contamination of *Mycobacterium tuberculosis*: an investigation and analysis of causes and consequences // Intern Med J. 2002. Vol. 32. № 11. P. 512-519.

204. Pulido, P. A., Novoa-Aponte, L., Villamil, N. and Soto, C. Y. The DosR dormancy regulator of *Mycobacterium tuberculosis* stimulates the Na(+)/K (+) and Ca (2+) ATPase activities in plasma membrane vesicles of mycobacteria // Curr Microbiol. 2014. Vol. 69. № 5. P. 604-610. doi: 610.1007/s00284-00014-00632-00286. Epub 02014 Jun 00218.

205. Raman, K., Yeturu, K. and Chandra, N. targetTB: a target identification pipeline for *Mycobacterium tuberculosis* through an interactome, reactome and genome-scale structural analysis // BMC Syst Biol. 2008. Vol. 2:109. № doi. P. 10.1186/1752-0509-1182-1109.

206. Reddy, T. B., Riley, R., Wymore, F., Montgomery, P., DeCaprio, D., *et al.* TB database: an integrated platform for tuberculosis research // Nucleic Acids Res.

2009. Vol. 37. № Database issue. P. D499-508. doi: 410.1093/nar/gkn1652. Epub 2008 Oct 1093.

207. Rengarajan, J., Bloom, B. R. and Rubin, E. J. Genome-wide requirements for *Mycobacterium tuberculosis* adaptation and survival in macrophages // Proc Natl Acad Sci U S A. 2005. Vol. 102. № 23. P. 8327-8332. Epub 2005 May 8331.

208. Reyrat, J. M. and Kahn, D. *Mycobacterium smegmatis*: an absurd model for tuberculosis? // Trends Microbiol. 2001. Vol. 9. № 10. P. 472-474.

209. Rodriguez, J. G., Hernandez, A. C., Helguera-Repetto, C., Aguilar Ayala, D., Guadarrama-Medina, R., *et al.* Global adaptation to a lipid environment triggers the dormancy-related phenotype of *Mycobacterium tuberculosis* // MBio. 2014. Vol. 5. № 3. P. e01125-01114. doi: 01110.01128/mBio.01125-01114.

210. Roetzer, A., Diel, R., Kohl, T. A., Ruckert, C., Nubel, U., *et al.* Whole genome sequencing versus traditional genotyping for investigation of a *Mycobacterium tuberculosis* outbreak: a longitudinal molecular epidemiological study // PLoS Med. 2013. Vol. 10. № 2. P. e1001387. doi: 1001310.1001371/journal.pmed.1001387. Epub 1002013 Feb 1001312.

211. Roring, S., Scott, A., Brittain, D., Walker, I., Hewinson, G., *et al.* Development of variable-number tandem repeat typing of *Mycobacterium bovis*: comparison of results with those obtained by using existing exact tandem repeats and spoligotyping // J Clin Microbiol. 2002. Vol. 40. № 6. P. 2126-2133.

212. Rosenkrands, I., King, A., Weldingh, K., Moniatte, M., Moertz, E., *et al.* Towards the proteome of *Mycobacterium tuberculosis* // Electrophoresis. 2000. Vol. 21. № 17. P. 3740-3756.

213. Ross, B. C., Raios, K., Jackson, K., Sievers, A. and Dwyer, B. Differentiation of *Mycobacterium tuberculosis* strains by use of a nonradioactive Southern blot hybridization method // J Infect Dis. 1991. Vol. 163. № 4. P. 904-907.

214. Rustad, T. R., Harrell, M. I., Liao, R. and Sherman, D. R. The enduring hypoxic response of *Mycobacterium tuberculosis* // PLoS One. 2008. Vol. 3. № 1. P. e1502. doi: 1510.1371/journal.pone.0001502.

215. Saint-Joanis, B., Demangel, C., Jackson, M., Brodin, P., Marsollier, L., *et al.* Inactivation of Rv2525c, a substrate of the twin arginine translocation (Tat) system of *Mycobacterium tuberculosis*, increases beta-lactam susceptibility and virulence // *J Bacteriol.* 2006. Vol. 188. № 18. P. 6669-6679.
216. Samanich, K. M., Belisle, J. T., Sonnenberg, M. G., Keen, M. A., Zolla-Pazner, S., *et al.* Delineation of human antibody responses to culture filtrate antigens of *Mycobacterium tuberculosis* // *J Infect Dis.* 1998. Vol. 178. № 5. P. 1534-1538.
217. Sanger, F., Nicklen, S. and Coulson, A. R. DNA sequencing with chain-terminating inhibitors. 1977 // *Biotechnology.* 1992. Vol. 24. № 104-108.
218. Schmidt, F., Donahoe, S., Hagens, K., Mattow, J., Schaible, U. E., *et al.* Complementary analysis of the *Mycobacterium tuberculosis* proteome by two-dimensional electrophoresis and isotope-coded affinity tag technology // *Mol Cell Proteomics.* 2004. Vol. 3. № 1. P. 24-42. Epub 2003 Oct 2013.
219. Schmidt-Schultz, T. H. and Schultz, M. AG 85, a major secretion protein of *Mycobacterium tuberculosis*, can be identified in ancient bone // *Tuberculosis (Edinb).* 2015. Vol. 95. № Suppl 1. P. S87-92. doi: 10.1016/j.tube.2015.1002.1034. Epub 2015 Feb 1018.
220. Schnappinger, D., Ehrt, S., Voskuil, M. I., Liu, Y., Mangan, J. A., *et al.* Transcriptional Adaptation of *Mycobacterium tuberculosis* within Macrophages: Insights into the Phagosomal Environment // *J Exp Med.* 2003. Vol. 198. № 5. P. 693-704.
221. Schubert, O. T., Mouritsen, J., Ludwig, C., Rost, H. L., Rosenberger, G., *et al.* The Mtb proteome library: a resource of assays to quantify the complete proteome of *Mycobacterium tuberculosis* // *Cell Host Microbe.* 2013. Vol. 13. № 5. P. 602-612. doi: 610.1016/j.chom.2013.1004.1008.
222. Schurch, A. C., Kremer, K., Hendriks, A. C., Freyee, B., McEvoy, C. R., *et al.* SNP/RD typing of *Mycobacterium tuberculosis* Beijing strains reveals local and worldwide disseminated clonal complexes // *PLoS One.* 2011. Vol. 6. № 12. P. e28365. doi: 28310.21371/journal.pone.0028365. Epub 0022011 Dec 0028365.

223. Scott, A. N., Menzies, D., Tannenbaum, T. N., Thibert, L., Kozak, R., *et al.* Sensitivities and specificities of spoligotyping and mycobacterial interspersed repetitive unit-variable-number tandem repeat typing methods for studying molecular epidemiology of tuberculosis // *J Clin Microbiol.* 2005. Vol. 43. № 1. P. 89-94.
224. Selvaraj, S., Sambandam, V., Sardar, D. and Anishetty, S. In silico analysis of DosR regulon proteins of *Mycobacterium tuberculosis* // *Gene.* 2012. Vol. 506. № 1. P. 233-241. doi: 210.1016/j.gene.2012.1006.1033. Epub 2012 Jul 1011.
225. Shalon, D., Smith, S. J. and Brown, P. O. A DNA microarray system for analyzing complex DNA samples using two-color fluorescent probe hybridization // *Genome Res.* 1996. Vol. 6. № 7. P. 639-645.
226. Sharma, D., Kumar, B., Lata, M., Joshi, B., Venkatesan, K., *et al.* Comparative Proteomic Analysis of Aminoglycosides Resistant and Susceptible *Mycobacterium tuberculosis* Clinical Isolates for Exploring Potential Drug Targets // *PLoS One.* 2015. Vol. 10. № 10. P. e0139414. doi: 0139410.0131371/journal.pone.0139414. eCollection 0132015.
227. Sherman, D. R., Voskuil, M., Schnappinger, D., Liao, R., Harrell, M. I., *et al.* Regulation of the *Mycobacterium tuberculosis* hypoxic response gene encoding alpha -crystallin // *Proc Natl Acad Sci U S A.* 2001. Vol. 98. № 13. P. 7534-7539.
228. Shevchenko, A., Tomas, H., Havlis, J., Olsen, J. V. and Mann, M. In-gel digestion for mass spectrometric characterization of proteins and proteomes // *Nat Protoc.* 2006. Vol. 1. № 6. P. 2856-2860.
229. Shi, C., Wang, X., Zhang, H., Xu, Z., Li, Y., *et al.* Immune responses and protective efficacy induced by 85B antigen and early secreted antigenic target-6 kDa antigen fusion protein secreted by recombinant bacille Calmette-Guerin // *Acta Biochim Biophys Sin (Shanghai).* 2007. Vol. 39. № 4. P. 290-296.
230. Shitikov, E. A., Bespyatykh, J. A., Ischenko, D. S., Alexeev, D. G., Karpova, I. Y., *et al.* Unusual large-scale chromosomal rearrangements in *Mycobacterium tuberculosis* Beijing B0/W148 cluster isolates // *PLoS One.* 2014. Vol. 9. № 1. P. e84971. doi: 84910.81371/journal.pone.0084971. eCollection 0082014.

231. Shitikov, E., Ilina, E., Chernousova, L., Borovskaya, A., Rukin, I., *et al.* Mass spectrometry based methods for the discrimination and typing of mycobacteria // *Infect Genet Evol.* 2012. Vol. 12. № 4. P. 838-845. doi: 810.1016/j.meegid.2011.1012.1013. Epub 2011 Dec 1030.
232. Shui, W., Petzold, C. J., Redding, A., Liu, J., Pitcher, A., *et al.* Organelle membrane proteomics reveals differential influence of mycobacterial lipoglycans on macrophage phagosome maturation and autophagosome accumulation // *J Proteome Res.* 2011. Vol. 10. № 1. P. 339-348. doi: 310.1021/pr100688h. Epub 102010 Dec 100616.
233. Singh, A. K., Dutta, D., Singh, V., Srivastava, V., Biswas, R. K., *et al.* Characterization of *Mycobacterium smegmatis* sigF mutant and its regulon: overexpression of SigF antagonist (MSMEG_1803) in *M. smegmatis* mimics sigF mutant phenotype, loss of pigmentation, and sensitivity to oxidative stress // *Microbiologyopen.* 2015. Vol. 4. № 6. P. 896-916. doi: 810.1002/mbo1003.1288. Epub 2015 Oct 1005.
234. Singhal, N., Sharma, P., Kumar, M., Joshi, B. and Bisht, D. Analysis of intracellular expressed proteins of *Mycobacterium tuberculosis* clinical isolates // *Proteome Sci.* 2012. Vol. 10. № 1. P. 14. doi: 10.1186/1477-5956-1110-1114.
235. Sivaramakrishnan, S. and de Montellano, P. R. The DosS-DosT/DosR Mycobacterial Sensor System // *Biosensors (Basel).* 2013. Vol. 3. № 3. P. 259-282.
236. Snapper, S. B., Melton, R. E., Mustafa, S., Kieser, T. and Jacobs, W. R., Jr. Isolation and characterization of efficient plasmid transformation mutants of *Mycobacterium smegmatis* // *Mol Microbiol.* 1990. Vol. 4. № 11. P. 1911-1919.
237. Sreevatsan, S., Pan, X., Stockbauer, K. E., Connell, N. D., Kreiswirth, B. N., *et al.* Restricted structural gene polymorphism in the *Mycobacterium tuberculosis* complex indicates evolutionarily recent global dissemination // *Proc Natl Acad Sci U S A.* 1997. Vol. 94. № 18. P. 9869-9874.
238. Stinear, T. P., Seemann, T., Pidot, S., Frigui, W., Reysset, G., *et al.* Reductive evolution and niche adaptation inferred from the genome of *Mycobacterium*

ulcerans, the causative agent of Buruli ulcer // *Genome Res.* 2007. Vol. 17. № 2. P. 192-200. Epub 2007 Jan 2008.

239. Stover, C. K., de la Cruz, V. F., Fuerst, T. R., Burlein, J. E., Benson, L. A., *et al.* New use of BCG for recombinant vaccines // *Nature.* 1991. Vol. 351. № 6326. P. 456-460.

240. Supply, P., Allix, C., Lesjean, S., Cardoso-Oelemann, M., Rusch-Gerdes, S., *et al.* Proposal for standardization of optimized mycobacterial interspersed repetitive unit-variable-number tandem repeat typing of *Mycobacterium tuberculosis* // *J Clin Microbiol.* 2006. Vol. 44. № 12. P. 4498-4510. Epub 2006 Sep 4427.

241. Supply, P., Marceau, M., Mangenot, S., Roche, D., Rouanet, C., *et al.* Genomic analysis of smooth tubercle bacilli provides insights into ancestry and pathoadaptation of *Mycobacterium tuberculosis* // *Nat Genet.* 2013. Vol. 45. № 2. P. 172-179. doi: 110.1038/ng.2517. Epub 2013 Jan 1036.

242. Supply, P., Mazars, E., Lesjean, S., Vincent, V., Gicquel, B., *et al.* Variable human minisatellite-like regions in the *Mycobacterium tuberculosis* genome // *Mol Microbiol.* 2000. Vol. 36. № 3. P. 762-771.

243. Tailleux, L., Waddell, S. J., Pelizzola, M., Mortellaro, A., Withers, M., *et al.* Probing host pathogen cross-talk by transcriptional profiling of both *Mycobacterium tuberculosis* and infected human dendritic cells and macrophages // *PLoS One.* 2008. Vol. 3. № 1. P. e1403. doi: 1410.1371/journal.pone.0001403.

244. Talaat, A. M., Lyons, R., Howard, S. T. and Johnston, S. A. The temporal expression profile of *Mycobacterium tuberculosis* infection in mice // *Proc Natl Acad Sci U S A.* 2004. Vol. 101. № 13. P. 4602-4607. Epub 2004 Mar 4618.

245. Talaat, A. M., Ward, S. K., Wu, C. W., Rondon, E., Tavano, C., *et al.* Mycobacterial bacilli are metabolically active during chronic tuberculosis in murine lungs: insights from genome-wide transcriptional profiling // *J Bacteriol.* 2007. Vol. 189. № 11. P. 4265-4274. Epub 2007 Mar 4223.

246. Tan, T., Lee, W. L., Alexander, D. C., Grinstein, S. and Liu, J. The ESAT-6/CFP-10 secretion system of *Mycobacterium marinum* modulates phagosome maturation // *Cell Microbiol.* 2006. Vol. 8. № 9. P. 1417-1429.

247. Thong-On, A., Smittipat, N., Juthayothin, T., Yanai, H., Yamada, N., *et al.* Variable-number tandem repeats typing of *Mycobacterium tuberculosis* isolates with low copy numbers of IS6110 in Thailand // Tuberculosis (Edinb). 2010. Vol. 90. № 1. P. 9-15. doi: 10.1016/j.tube.2009.1010.1006. Epub 2009 Nov 1012.
248. Toungousova, O. S., Sandven, P., Mariandyshv, A. O., Nizovtseva, N. I., Bjune, G., *et al.* Spread of drug-resistant *Mycobacterium tuberculosis* strains of the Beijing genotype in the Archangel Oblast, Russia // J Clin Microbiol. 2002. Vol. 40. № 6. P. 1930-1937.
249. Treumann, A. and Thiede, B. Isobaric protein and peptide quantification: perspectives and issues // Expert Rev Proteomics. 2010. Vol. 7. № 5. P. 647-653. doi: 610.1586/epr.1510.1529.
250. Triccas, J. A., Roche, P. W., Winter, N., Feng, C. G., Butlin, C. R., *et al.* A 35-kilodalton protein is a major target of the human immune response to *Mycobacterium leprae* // Infect Immun. 1996. Vol. 64. № 12. P. 5171-5177.
251. Tsolaki, A. G., Gagneux, S., Pym, A. S., Goguet de la Salmoniere, Y. O., Kreiswirth, B. N., *et al.* Genomic deletions classify the Beijing/W strains as a distinct genetic lineage of *Mycobacterium tuberculosis* // J Clin Microbiol. 2005. Vol. 43. № 7. P. 3185-3191.
252. Tsolaki, A. G., Hirsh, A. E., DeRiemer, K., Enciso, J. A., Wong, M. Z., *et al.* Functional and evolutionary genomics of *Mycobacterium tuberculosis*: insights from genomic deletions in 100 strains // Proc Natl Acad Sci U S A. 2004. Vol. 101. № 14. P. 4865-4870. Epub 2004 Mar 4815.
253. Turkarslan, S., Peterson, E. J., Rustad, T. R., Minch, K. J., Reiss, D. J., *et al.* A comprehensive map of genome-wide gene regulation in *Mycobacterium tuberculosis* // Sci Data. 2015. Vol. 2:150010. № doi. P. 10.1038/sdata.2015.1010. eCollection 2015.
254. Tyler, A. D., Christianson, S., Knox, N. C., Mabon, P., Wolfe, J., *et al.* Comparison of Sample Preparation Methods Used for the Next-Generation Sequencing of *Mycobacterium tuberculosis* // PLoS One. 2016. Vol. 11. № 2. P. e0148676. doi: 0148610.0141371/journal.pone.0148676. eCollection 0142016.

255. Uchida, Y., Casali, N., White, A., Morici, L., Kendall, L. V., *et al.* Accelerated immunopathological response of mice infected with *Mycobacterium tuberculosis* disrupted in the *mce1* operon negative transcriptional regulator // *Cell Microbiol.* 2007. Vol. 9. № 5. P. 1275-1283. Epub 2007 Jan 1211.
256. Vadwai, V., Shetty, A., Supply, P. and Rodrigues, C. Evaluation of 24-locus MIRU-VNTR in extrapulmonary specimens: study from a tertiary centre in Mumbai // *Tuberculosis (Edinb).* 2012. Vol. 92. № 3. P. 264-272. doi: 210.1016/j.tube.2012.1001.1002. Epub 2012 Feb 1018.
257. van Crevel, R., Parwati, I., Sahiratmadja, E., Marzuki, S., Ottenhoff, T. H., *et al.* Infection with *Mycobacterium tuberculosis* Beijing genotype strains is associated with polymorphisms in SLC11A1/NRAMPI in Indonesian patients with tuberculosis // *J Infect Dis.* 2009. Vol. 200. № 11. P. 1671-1674. doi: 1610.1086/648477.
258. van Embden, J. D., Cave, M. D., Crawford, J. T., Dale, J. W., Eisenach, K. D., *et al.* Strain identification of *Mycobacterium tuberculosis* by DNA fingerprinting: recommendations for a standardized methodology // *J Clin Microbiol.* 1993. Vol. 31. № 2. P. 406-409.
259. Van Soolingen, D. Molecular epidemiology of tuberculosis and other mycobacterial infections: main methodologies and achievements // *J Intern Med.* 2001. Vol. 249. № 1. P. 1-26.
260. van Soolingen, D., Qian, L., de Haas, P. E., Douglas, J. T., Traore, H., *et al.* Predominance of a single genotype of *Mycobacterium tuberculosis* in countries of east Asia // *J Clin Microbiol.* 1995. Vol. 33. № 12. P. 3234-3238.
261. Varma-Basil, M., Kumar, S., Arora, J., Angrup, A., Zozio, T., *et al.* Comparison of spoligotyping, mycobacterial interspersed repetitive units typing and IS6110-RFLP in a study of genotypic diversity of *Mycobacterium tuberculosis* in Delhi, North India // *Mem Inst Oswaldo Cruz.* 2011. Vol. 106. № 5. P. 524-535.
262. Viana-Niero, C., Rodriguez, C. A., Bigi, F., Zanini, M. S., Ferreira-Neto, J. S., *et al.* Identification of an IS6110 insertion site in *plcD*, the unique phospholipase C gene of *Mycobacterium bovis* // *J Med Microbiol.* 2006. Vol. 55. № Pt 4. P. 451-457.

263. Vishnevskii, B. I., Narvskaiia, O. V., Vasil'eva, S. N., Sapozhnikova, N. V., Mukrousov, I. V., *et al.* [Virulence of *Mycobacteria tuberculosis*] // Probl Tuberk. 2002. Vol. № 10. P. 33-36.
264. Vowinckel, J., Capuano, F., Campbell, K., Deery, M. J., Lilley, K. S., *et al.* The beauty of being (label)-free: sample preparation methods for SWATH-MS and next-generation targeted proteomics // 1000Res. 2013. Vol. 2:272. № doi. P. 10.12688/f11000research.12682-12272.v12682. eCollection 12013.
265. Wang, H., Dong, D., Tang, S., Chen, X. and Gao, Q. PPE38 of *Mycobacterium marinum* triggers the cross-talk of multiple pathways involved in the host response, as revealed by subcellular quantitative proteomics // J Proteome Res. 2013. Vol. 12. № 5. P. 2055-2066. doi: 2010.1021/pr301017e. Epub 302013 Apr 301018.
266. Wang, H., Yue, J., Han, M., Yang, J. and Zhao, Y. Rapid method for identification of six common species of mycobacteria based on multiplex SNP analysis // J Clin Microbiol. 2010. Vol. 48. № 1. P. 247-250. doi: 210.1128/JCM.01084-01009. Epub 02009 Nov 01084.
267. Warner, D. F. and Mizrahi, V. Complex genetics of drug resistance in *Mycobacterium tuberculosis* // Nat Genet. 2013. Vol. 45. № 10. P. 1107-1108. doi: 1110.1038/ng.2769.
268. Weiner, B., Gomez, J., Victor, T. C., Warren, R. M., Sloutsky, A., *et al.* Independent large scale duplications in multiple *M. tuberculosis* lineages overlapping the same genomic region // PLoS One. 2012. Vol. 7. № 2. P. e26038. doi: 26010.21371/journal.pone.0026038. Epub 0022012 Feb 0026037.
269. Wellman Ribón (2012). Biochemical Isolation and Identification of Mycobacteria, Biochemical Testing, Dr. Jose C. Jimenez-Lopez (Ed.), InTech, DOI: 10.5772/34309.
270. Werngren, J. and Hoffner, S. E. Drug-susceptible *Mycobacterium tuberculosis* Beijing genotype does not develop mutation-conferred resistance to rifampin at an elevated rate // J Clin Microbiol. 2003. Vol. 41. № 4. P. 1520-1524.

271. WHO. Extensively drug-resistant tuberculosis (XDR-TB): recommendations for prevention and control // *Wkly Epidemiol Rec.* 2015. Vol. 81. № 45. P. 430-432.
272. Wilkins, M. R., Pasquali, C., Appel, R. D., Ou, K., Golaz, O., *et al.* From proteins to proteomes: large scale protein identification by two-dimensional electrophoresis and amino acid analysis // *Biotechnology (N Y)*. 1996. Vol. 14. № 1. P. 61-65.
273. Wirth, T., Hildebrand, F., Allix-Beguec, C., Wolbeling, F., Kubica, T., *et al.* Origin, spread and demography of the *Mycobacterium tuberculosis* complex // *PLoS Pathog.* 2008. Vol. 4. № 9. P. e1000160. doi: 1000110.1001371/journal.ppat.1000160.
274. Wolfe, L. M., Mahaffey, S. B., Kruh, N. A. and Dobos, K. M. Proteomic definition of the cell wall of *Mycobacterium tuberculosis* // *J Proteome Res.* 2010. Vol. 9. № 11. P. 5816-5826. doi: 5810.1021/pr1005873. Epub 1002010 Sep 1005829.
275. Xiong, Y., Chalmers, M. J., Gao, F. P., Cross, T. A. and Marshall, A. G. Identification of *Mycobacterium tuberculosis* H37Rv integral membrane proteins by one-dimensional gel electrophoresis and liquid chromatography electrospray ionization tandem mass spectrometry // *J Proteome Res.* 2005. Vol. 4. № 3. P. 855-861.
276. Yadeta K.A., Elmore J.M., and Coaker G. // *Front Plant Sci.* 2013. V. 4 № 86
277. Yuksel, P. and Tansel, O. Characterization of *pncA* mutations of pyrazinamide-resistant *Mycobacterium tuberculosis* in Turkey // *New Microbiol.* 2009. Vol. 32. № 2. P. 153-158.
278. Zemskova ZS, Andreyevskaya SN, Smirnova TG, Larionova YY, Chernousova LN. 2010. Experimental tuberculosis caused by *Mycobacterium tuberculosis* strains of genotypical W, AI, and HD clusters. *Probl. Tuberk. Bolezn. Legk.* 2010(3):41– 46. (In Russian.)
279. Zhang, H., Li, D., Zhao, L., Fleming, J., Lin, N., *et al.* Genome sequencing of 161 *Mycobacterium tuberculosis* isolates from China identifies genes and intergenic regions associated with drug resistance // *Nat Genet.* 2013. Vol. 45. № 10. P. 1255-1260. doi: 1210.1038/ng.2735. Epub 2013 Sep 1251.

280. Zhang, J., Abadia, E., Refregier, G., Tafaj, S., Boschioli, M. L., *et al.* *Mycobacterium tuberculosis* complex CRISPR genotyping: improving efficiency, throughput and discriminative power of 'spoligotyping' with new spacers and a microbead-based hybridization assay // J Med Microbiol. 2010. Vol. 59. № Pt 3. P. 285-294. doi: 210.1099/jmm.1090.016949-016940. Epub 012009 Dec 016943.

281. Zhang, L., Wang, Q., Wang, W., Liu, Y., Wang, J., *et al.* Identification of putative biomarkers for the serodiagnosis of drug-resistant *Mycobacterium tuberculosis* // Proteome Sci. 2012. Vol. 10:12. № doi. P. 10.1186/1477-5956-1110-1112.

282. Zhang, M., Yue, J., Yang, Y. P., Zhang, H. M., Lei, J. Q., *et al.* Detection of mutations associated with isoniazid resistance in *Mycobacterium tuberculosis* isolates from China // J Clin Microbiol. 2005. Vol. 43. № 11. P. 5477-5482.

283. Zhu, W., Smith, J. W. and Huang, C. M. Mass spectrometry-based label-free quantitative proteomics // J Biomed Biotechnol. 2010. Vol. 2010:840518. № doi. P. 10.1155/2010/840518. Epub 842009 Nov 840510

ПРИЛОЖЕНИЯ

ПРИЛОЖЕНИЕ 1. Молекулярно-генетическая и фенотипическая характеристика образцов коллекции

№	Образец	Clade	SIT №	24-VNTR*	Профиль резистентности	INH	RIF	EMB	STR	CPM	AM	KAN	OFX	ETH	PZA	Ac Numb
1	Sp1	Beijing	SIT-1	223325173533424672454433	ШЛУ	R	R	R	R	S	S	R	R	R	R	SRX216883
2	Sp7	Beijing	SIT-1	223325173533424672454433	ШЛУ	R	R	R	R	R	R	R	R	R	R	SRX216889
3	Sp10	Beijing	SIT-1	223325173533424672444433	МЛУ	R	R	R	R	S	S	S	R	R	R	SRX216892
4	Sp13	Beijing	SIT-1	223325173533424672444433	МЛУ	R	R	R	R	S	S	R	S	S	R	SRX216895
5	Sp21	Beijing	SIT-1	223325173533424672454433	ШЛУ	R	R	R	R	R	S	R	R	R	R	SRX216899
6	Sp22	Beijing	SIT-1	223325173533424672444433	ШЛУ	R	R	S	R	R	S	R	R	S	S	SRX216900
7	Sp23	Beijing	SIT-1	223325173533424672444433	ШЛУ	R	R	R	R	R	S	R	R	R	R	-
8	Sp27	Beijing	SIT-1	223325173533424672444433	ШЛУ	R	R	R	R	R	R	S	R	R	R	-
9	Sp45	Beijing	SIT-1	223325173533424672444433	чувствительный	S	S	S	S	S	S	S	S	S	S	-
10	Sp54	Beijing	SIT-1	223325173533424672444433	МЛУ	R	R	R	R	R	R	S	S	R	R	-

11	H37R v	H37Rv	SIT- 451	233'22613332124253423355 2	чувствительный	S	S	S	S	S	S	S	S	S	S	NC_0009 62.3
----	-----------	-------	-------------	-------------------------------	----------------	---	---	---	---	---	---	---	---	---	---	-----------------

INH – изониазид, RIF – рифампицин, EMB – этамбутол, STR – стрептомицин, CPM – капреомицин, AM – амикацин, KAN – канамицин, OFX – офлоксацин, ETH – этионамид, PZA – пипразинамид.

*24 – VNTR: s154, s580, s960, s1644, s2059, s2531, s2687, s2996, s3007, s3192, s4348, s802, s2165, s2461, s577, s2163, s4052, s4156, s424, s1955, s2347, s2401, s3171, s3690 (Supply *et al.*, 2006).

Серым цветом отмечены образцы для протеомного анализа

ПРИЛОЖЕНИЕ 2. *M. tuberculosis* Beijing B0/W148 кластер - специфические полиморфизмы

Серым цветом отмечены позиции, вываленные в работе (Merker *et al.*, 2015)

Позиция по H37Rv	мутация	ген	аминокислотная замена	Продукт гена	Функциональная категория согласно TubercuList
1035426	T>G	pstS3	F175C	phosphate ABC transporter substrate-binding lipoprotein PstS	Cell wall and cell processes
1155884	T>C	kdpC	M11T	potassium-transporting ATPase subunit C	Cell wall and cell processes
2127011	T>G	Rv1877	W370G	MFS-type transporter	Cell wall and cell processes
3031090	G>A	Rv2719c	Q150*	membrane protein	Cell wall and cell processes
4175847	G>A	Rv3728	W325*	membrane protein	Cell wall and cell processes
4349982	G>T	eccCb1	A386S	ESX-1 secretion system protein EccCb	Cell wall and cell processes
4358998	A>G	espK	I262T	ESX-1 secretion-associated protein EspK	Cell wall and cell processes
4367633	T>C	eccE2	T297A	ESX-2 secretion system protein EccE	Cell wall and cell processes
751999	C>T	mkl	P161P	ribonucleotide ABC transporter ATP-binding protein	Cell wall and cell processes
1154618	C>T	kdpB	L299L	potassium-transporting ATPase subunit B	Cell wall and cell processes
1272899	T>G	Rv1145	R159R	transmembrane transport protein	Cell wall and cell processes
3014343	C>T	Rv2698	V57V	transmembrane protein	Cell wall and cell processes
4243346	A>G	embA	Q38Q	arabinosyltransferase A	Cell wall and cell processes
4399683	C>G	Rv3910	A1029A	peptidoglycan biosynthesis protein	Cell wall and cell processes
1710601	T>A	Rv1518	*320K	hypothetical protein	Conserved hypotheticals
2335638	T>C	Rv2079	L95S	hypothetical protein	Conserved hypotheticals
2604740	A>G	Rv2331A	M1V	hypothetical protein	Conserved hypotheticals
2964454	C>T	Rv2638	P17L	hypothetical protein	Conserved hypotheticals
3566107	C>A	Rv3196	A107D	hypothetical protein	Conserved hypotheticals
1353462	G>A	Rv1209	E102E	hypothetical protein	Conserved hypotheticals
2267125	T>C	Rv2022c	K200K	hypothetical protein	Conserved hypotheticals
2569288	G>C	Rv2298	G69G	oxidoreductase	Conserved hypotheticals
2954197	G>A	Rv2627c	T184T	hypothetical protein	Conserved hypotheticals
1203693	C>T	Rv1078	S127S	hypothetical protein	Conserved hypotheticals
4086	G>T	recF	G269G	DNA replication/repair protein RecF	Information pathways
1838646	C>T	uvrB	A524A	excinuclease ABC subunit UvrB	Information pathways
143120	C>A	oxcA	A253S	oxalyl-CoA decarboxylase OxcA	Intermediary metabolism and respiration
404130	T>C	Rv0338c	E571G	iron-sulfur-binding reductase	Intermediary metabolism and respiration
1482185	C>T	Rv1319c	G106D	adenylate cyclase	Intermediary metabolism and respiration
2078967	C>A	Rv1833c	D275Y	haloalkane dehalogenase	Intermediary metabolism and respiration
2099129	C>T	ureC	A390V	urease subunit alpha	Intermediary metabolism and respiration
2544135	A>C	cyp128	S48A	cytochrome P450 Cyp128	Intermediary metabolism and

					respiration
3109512	G>A	Rv2800	W366*	hydrolase	Intermediary metabolism and respiration
527895	T>C	Rv0439c	V140V	dehydrogenase/reductase	Intermediary metabolism and respiration
826756	C>T	mapA	A29A	methionine aminopeptidase	Intermediary metabolism and respiration
895164	G>A	Rv0802c	I155I	succinyl-CoA transferase	Intermediary metabolism and respiration
907046	G>A	Rv0812	P208P	amino acid aminotransferase	Intermediary metabolism and respiration
1584762	G>A	rpe	A88A	ribulose-phosphate 3-epimerase	Intermediary metabolism and respiration
2132640	G>C	Rv1882c	L188L	short-chain type dehydrogenase/reductase	Intermediary metabolism and respiration
2908357	G>A	relA	G614G	bifunctional (p)ppGpp synthase/hydrolase RelA	Intermediary metabolism and respiration
3161858	T>G	mgo	P68P	malate:quinone oxidoreductase	Intermediary metabolism and respiration
3203157	C>T	Rv2893	C246C	oxidoreductase	Intermediary metabolism and respiration
3697016	G>C	upp	A26A	uracil phosphoribosyltransferase	Intermediary metabolism and respiration
2517129	A>G	fabD	T115A	malonyl CoA-acyl carrier protein transacylase	Lipid metabolism
1619513	A>C	PE_PGRS26	Y58D	PE-PGRS family protein PE_PGRS26	PE/PPE
4338371	T>G	whiB6	T51P	transcriptional regulator WhiB6	Regulatory proteins
3086731	A>G	Rv2779c	H8H	Lrp/AsnC family transcriptional regulator	Regulatory proteins
4017448	G>A	Rv3575c	L241L	LacI family transcriptional regulator	Regulatory proteins
1668082	G>C	ripA	A365P	peptidoglycan endopeptidase RipA	Virulence, detoxification, adaption
1947325	T>C	vapC12	H32R	ribonuclease VapC12	Virulence, detoxification, adaption
2210745	T>G	mce3B	S49A	Mce family protein Mce3B	Virulence, detoxification, adaption
886661	C>T	intergenic			
1023883	G>A	intergenic			
2205550	A>G	intergenic			
2468180	G>A	intergenic			
3557244	C>G	intergenic			
3610335	G>T	intergenic			
4216783	A>C	intergenic			
4321479	G>A	intergenic			

ПРИЛОЖЕНИЕ 3. Разработка протокола экстракции белка из клеток микобактерий

Исходя из поставленных задач, нами протестированы три способа смыва клеток *M. smegmatis* со среды культивирования.

1) Добавление к клеткам раствора 1 % Tween-80 (Sigma Aldrich, США) в фосфатно-солевом буфере (PBS) (Sigma Aldrich, США) и центрифугирование 20 минут при 4500 g (Blackwood *et al.*, 2005);

1) Добавление к клеткам PBS + 1 % тритон X-100 (Sigma Aldrich, США) и центрифугирование 15 минут при 4000 g (http://www.signaling-gateway.org/data/cgi-bin/ProtocolFile.cgi/afcs_PS00000603.pdf?pid=PS00000603);

1) Добавление к клеткам Tris-HCl, PBS + 2 % тритон X-100 (pH 7.5 - 8) и центрифугирование 15 минут при 4500 g (Jungblut *et al.*, 1999).

Согласно визуальной оценке цвета и консистенции получаемого бактериального осадка наиболее эффективным был признан метод 1В. Данный метод использовался в последующих экспериментах.

Для отработки протокола экстракции белка суспензию клеток в отмывочном буфере нормировали по мутности. В каждый эксперимент брали приблизительно 10^{10} клеток. Апробированы пять различных методов экстракции общего белка из микобактериальных клеток.

Метод 2А

К клеточному осадку добавляли 10 мкл 100 мМ раствора NH_4HCO_3 , 10 мкл лизоцима (1 мг/мл) (Sigma Aldrich, США) и инкубировали 30 мин при 4°C. К клеточной суспензии добавляли 1 мкл смеси нуклеаз (GE Healthcare, США), 50 мкл 0.75 % раствора RapiGest SF (Waters, США) в 100 мМ NH_4HCO_3 и 1 мкл смеси ингибиторов протеаз (GE Healthcare, США). Клетки разрушали на Bead Beater (MPBio, США), с использованием 0.5 мм кремний-циркониевых шариков в соотношении 1V образца : 1 V шариков, или ультразвуковом гомогенизаторе (И100-6/840, ИНЛАБ, Россия), 3 раза по 1 мин с 5-мин перерывами в ледяной бане. Затем нагревали в течение 5 мин при 100°C, охлаждали до комнатной

температуры и центрифугировали в течение 5 мин при 4°C и 16500 g. Супернатант отбирали и использовали для дальнейшего анализа.

К клеточному лизату (200 мкг белка) добавляли ДТТ (Sigma Aldrich, США) (использовали сток – 200 мМ раствор в 100 мМ NH_4HCO_3) до конечной концентрации 10 мМ, инкубировали 30 мин при 60°C. Затем добавляли йодацетамид (BioRad, США) до конечной концентрации 30 мМ, инкубировали 30 мин при комнатной температуре в темноте. К клеточному лизату снова добавляли ДТТ до конечной концентрации 10 мМ, инкубировали 30 мин при комнатной температуре. Далее к полученной суспензии добавляли трипсин (Trypsin Gold, Promega, США) в соотношении трипсин : белок, равному 1 : 50 (по объему) (pH < 8.0), инкубировали ночь (16 часов) при 37°C. Для остановки гидролиза добавляли трифторуксусную кислоту до конечной концентрации 0.5 % (pH < 2.0), инкубировали 45 мин при 37°C и центрифугировали 10 мин при 16500 g, отбирали супернатант (белковый экстракт).

Метод 2Б

Бактериальный осадок растворяли в 1 мл буфера pH 7.5 - 8.0 (Tris-HCl + 1 % Triton X-100). Полученную суспензию нагревали 15 минут при 95°C, затем остужали до комнатной температуры. Добавляли ингибиторы протеаз 1 мкл (GE Healthcare, США), 1 мкл нуклеаз (ДНКаза + РНКаза, Promega, США) и инкубировали 30 минут при комнатной температуре. В лизат добавляли гуанидин гидрохлорид до конечной концентрации 6 М и растворяли без нагревания. Клетки разрушали на Bead Beater (MPBio, США) с использованием 0.1 мм кремний-циркониевых шариков в соотношении 1V образца : 1V шариков или ультразвуковом гомогенизаторе (И100-6/840, ИНЛАБ, Россия), в течение 5 минут. Лизат инкубировали 40 минут при комнатной температуре при покачивании и затем центрифугировали 10 минут при 16100 g. Супернатант отбирали и использовали для дальнейшего анализа.

Метод 2В

К осадку клеточной культуры добавляли 10 мкл 100 мМ Tris-HCl pH 7.6 и 5 мкл лизоцима (1 мг/мл), инкубировали 30 мин при 4°C. Дополнительно добавляли

55 мкл 100 мМ Tris-HCl pH 7.6 и 1 мкл смеси ингибиторов протеаз (GE Healthcare, США). Клетки разрушали на Bead Beater (MPBio, США), с использованием 0.5 мм кремний-циркониевых шариков в соотношении 1V образца : 1 V шариков, или ультразвуковом гомогенизаторе (И100-6/840, ИНЛАБ, Россия), 2 раза по 2 мин с 5-мин перерывом в ледяной бане. Затем к суспензии добавляли 20 мкл 20 % SDS (Sigma Aldrich, США), 10 мкл 1 М ДТТ (Sigma Aldrich, США); инкубировали 30 мин при 60°C и центрифугировали при 16000 g в течении 5 минут. Супернатант отбирали и использовали для дальнейшего анализа.

Метод 2Г

Клеточный осадок растворяли в 300 мкл 4 % SDS (Sigma Aldrich, США), 100 мкл 100 мМ NH₄HCO₃ (Sigma Aldrich, США), 40 мкл 1М ДТТ (Sigma Aldrich, США), перемешивали на вортексе 20 секунд. Инкубировали смесь 30 минут при 80°C, остужали до комнатной температуры. Клетки разрушали на Bead Beater (MPBio, США), с использованием 0.1 мм кремний-циркониевых шариков в соотношении 1 V образца : 1 V шариков 4 минуты, центрифугировали в течение 15 минут при 15000 g. Супернатант отбирали и использовали для дальнейшего анализа.

Метод 2Д

Клеточный осадок растворяли в 300 мкл 4 % SDS (Sigma Aldrich, США), 100 мкл мочевины (Sigma Aldrich, США), 40 мкл 1М ДТТ (Sigma Aldrich, США), 30 мкл CHAPS (Sigma Aldrich, США). Клетки разрушали на Bead Beater (MPBio, США), с использованием 0.1 и 0.5 мм кремний-циркониевых шариков в соотношении 1 V образца : 1V шариков 5 мин, центрифугировали в течение 15 минут при 15000 g. Супернатант отбирали и использовали для дальнейшего анализа.

Концентрацию белка в белковых экстрактах определяли по методу Брэдфорда (Bradford *et al.*, 1976) с использованием Bradford Protein Assay Kit (Bio Rad, США). Результаты представлены на Рисунке 3.1

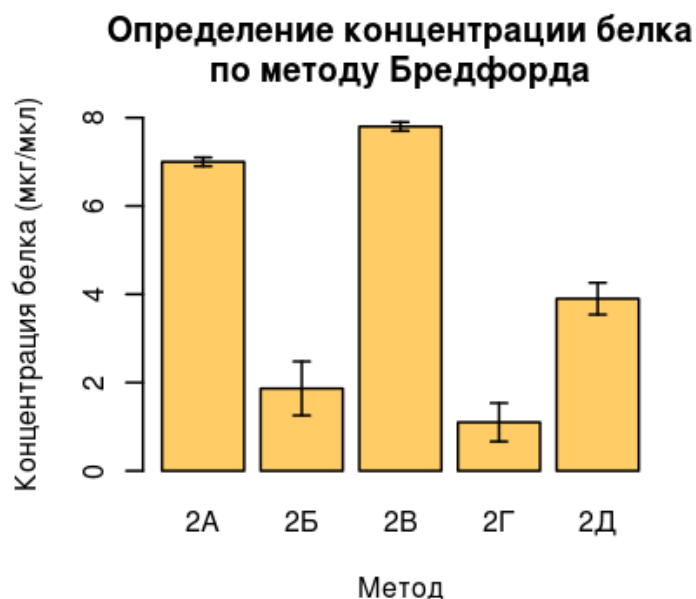


Рисунок 3.1 Определение концентрации белка по методу Бредфорда.

Гидролиз белков трипсином в растворе по методу Filter-Aided Sample Preparation (FASP). Клеточный лизат (200 мкг белка) разводили в 100 мМ Tris-HCl pH 8.5 до объема 30 мкл и наносили на фильтр 10 кДа (Amicon Ultra, США). Далее последовательно наносили на фильтр растворы:

а) добавляли 200 мкл 8 М мочевины и 100 мМ Tris-HCl pH 8.5 (UA буфер), центрифугировали при 14000 g 30 мин при 20°C;

б) добавляли 200 мкл буфера UA, центрифугировали при 14000 g 30 мин при 20°C;

в) добавляли 100 мкл буфера раствор 50 мМ йодацетамида в буфере UA, центрифугировали при 14000 g 20 мин при 20°C;

г) добавляли 100 мкл буфер UA, центрифугировали при 14000 g 20 мин при 20°C;

д) добавляли 100 мкл буфер UA, центрифугировали при 14000 g 20 мин при 20°C;

е) добавляли 100 мкл 50 мМ NH_4HCO_3 , центрифугировали при 14000 g 20 мин при 20°C;

ж) добавляли 100 мкл 50 мМ NH_4HCO_3 , центрифугировали при 14000 g 20 мин при 20°C.

Затем переносили фильтр в чистый коллектор и добавляли 40 мкл буфера, содержащего 10 % ацетонитрил (ACN), 50 мМ NH_4HCO_3 и трипсин (Trypsin Gold, Promega, США) в соотношении трипсин: белок, равному 1 : 50 (по объему), инкубировали ночь (16 часов) при 37°C.

После гидролиза пробы центрифугировали 20 мин при 14,000 g и 20°C. Элюировали пептиды в три стадии:

1. Добавляли 200 мкл элюирующего буфера 1 (0.1 % муравьиная кислота (FA)) и центрифугировали 20 мин при 14000 g, 20°C.

2. Добавляли 200 мкл элюирующего буфера 2 (0.1 % FA, 50 % ACN) и центрифугировали 30 мин при 14000 g, 20°C.

3. Добавляли 200 мкл элюирующего буфера 3 (0.1 % FA, 75 % ACN) и центрифугировали 30 мин при 14000 g, 20°C.

Гидролиз белков трипсином в геле. К клеточному лизату (200 мкг белка) добавляли равный объем 2х кратного раствора Лэммли, прогревали 5 мин при 95°C, центрифугировали 5 мин при 10000 g и отбирали супернатант. Белки фракционировали с помощью одномерного геля электрофореза (7.5 % ПААГ, 7-см гели) в денатурирующих условиях. Белки разгоняли 20 мин при 20 мА и затем до конца при 40 мА с использованием системы PROTEAN II (Bio-Rad, США). Гель окрашивали раствором Кумасси G-250 (Invitrogen, США) в течение 1 - 2 часов и отмывали от краски раствором 10 % уксусной кислоты. Вырезали нужную часть геля, содержащую белки, нарезали его на кусочки размером 1 x 1 мм. Отмывали гели от раствора Кумасси с помощью 50 % раствора ACN в 50 мМ NH_4HCO_3 . Заливали гели раствором 10 мМ ДТТ в 100 мМ NH_4HCO_3 (так, чтобы покрывало все кусочки геля), инкубировали 30 мин при 56°C, раствор сливали. Заливали гели раствором 55 мМ йодацетамида в 100 мМ NH_4HCO_3 , инкубировали 20 мин при комнатной температуре в темноте, раствор так же сливали. Высушивали отмые гели, добавляя в них 100 % ACN, инкубировали 20 мин при комнатной температуре, сливали ACN, оставляли пробирки открытыми до полного высыхания гелей. Непосредственно для гидролиза белков добавляли к высушенным гелям 150-300 мкл раствора трипсина (20 нг/мкл) в 40 мМ NH_4HCO_3

и 10 % ACN, инкубировали 1 час на льду, а затем инкубировали ночь (16 часов) при 37°C.

После гидролиза полученные пептиды экстрагировали последовательно в три этапа:

1. Добавляли к образцу 5 % раствор муравьиной кислоты (2 объема, 300 мкл) и инкубировали при комнатной температуре в течение 30 минут. Раствор отбирали в отдельную пробирку;

2. Добавляли к образцу раствор 50 % ACN и 5 % FA (из расчета на 300 мкл раствора: 150 мкл 100 % ACN, 15 мкл 100 % FA и 135 мкл деионизированной воды), инкубировали при комнатной температуре в течение 30 минут. Раствор отбирали в отдельную пробирку.

3. Добавляли к образцу раствор 75 % ACN и 5 % FA (из расчета на 300 мкл раствора: 225 мкл 100 % ACN, 15 мкл 100 % FA и 70 мкл деионизированной воды), инкубировали при комнатной температуре в течение 30 минут. Раствор отбирали в отдельную пробирку.

Экстракты объединяли, полностью высушивали на вакуумной сушке (SpeedVac) и растворяли в минимальном объеме (20 мкл) раствором, содержащим 5 % ACN и 0.1 % FA.

Гидролиз белков трипсином в растворе для реагента RapiGest SF. К клеточному лизату (200 мкг белка) добавляли ДТТ (использовали сток – 200 мМ раствор в 100 мМ NH_4HCO_3) до конечной концентрации 10 мМ, инкубировали 30 мин при 60°C. Затем добавляли йодацетамид (BioRad, США) до конечной концентрации 30 мМ, инкубировали 30 мин при комнатной температуре в темноте. К клеточному лизату снова добавляли ДТТ до конечной концентрации 10 мМ, инкубировали 30 мин при комнатной температуре. Далее к полученной суспензии добавляли трипсин (Trypsin Gold, Promega, США) в соотношении трипсин : белок, равному 1 : 50 (по объему) (pH < 8.0), инкубировали ночь (16 часов) при 37°C. Для остановки гидролиза добавляли трифторуксусную кислоту до конечной концентрации 0.5 % (pH < 2.0), инкубировали 45 мин при 37°C и

центрифугировали 10 мин при 16500 g, отбирали супернатант (пептидный экстракт).

ПРИЛОЖЕНИЕ 4. Белки, идентифицированные для штамма *M. tuberculosis* H37Rv

Синоним	Ген	Молекулярная масса	Локализация	Аннотация NCBI	Аннотация Tuberculist	Функциональная категория	COG категория
Rv0001	dnaA	56548.6	Cytoplasmic	chromosomal replication initiator protein DnaA	Chromosomal replication initiator protein DnaA	information pathways	COG0593L
Rv0002	dnaN	42113.1	Cytoplasmic	DNA polymerase III subunit beta	DNA polymerase III (beta chain) DnaN (DNA nucleotidyltransferase)	information pathways	COG0592L
Rv0005	gyrB	74058.7	Cytoplasmic	DNA gyrase subunit B	DNA gyrase (subunit B) GyrB (DNA topoisomerase (ATP-hydrolysing)) (DNA topoisomerase II) (type II DNA topoisomerase)	information pathways	-
Rv0006	gyrA	92274.3	Cytoplasmic	DNA gyrase subunit A	DNA gyrase (subunit A) GyrA (DNA topoisomerase (ATP-hydrolysing)) (DNA topoisomerase II) (type II DNA topoisomerase)	information pathways	COG0188L
Rv0008c	-	15670.9	Cytoplasmic Membrane	cell wall synthesis protein CwsA	Possible membrane protein	cell wall and cell processes	-
Rv0009	ppiA	19239.4	Cytoplasmic	iron-regulated peptidyl-prolyl cis-trans isomerase PpiA	Probable iron-regulated peptidyl-prolyl cis-trans isomerase A PpiA (PPIase A) (rotamase A)	information pathways	COG0652O
Rv0020c	fhaA	56880.3	Cytoplasmic	FHA domain-containing protein FhaA	Conserved protein with FHA domain, FhaA	Regulatory proteins	COG1716T
Rv0036c	-	27546.1	Cytoplasmic Membrane	hypothetical protein Rv0036c	Conserved protein	conserved hypotheticals	-
Rv0041	leuS	107563	Cytoplasmic	leucine--tRNA ligase	Probable leucyl-tRNA synthetase LeuS (leucine--tRNA ligase) (LEURS)	information pathways	COG0495J
Rv0046c	ino1	40093.6	Cytoplasmic	inositol-3-phosphate synthase	myo-inositol-1-phosphate synthase Ino1 (inositol 1-phosphate synthetase) (D-glucose 6-phosphate cycloaldolase) (glucose 6-phosphate cyclase) (glucocycloaldolase)	intermediary metabolism and respiration	COG1260I
Rv0054	ssb	17321.1	Cytoplasmic	single-strand DNA-binding protein	Single-strand binding protein Ssb (helix-destabilizing protein)	information pathways	COG0629L
Rv0056	rplI	16169.6	Cytoplasmic	50S ribosomal protein L9	50S ribosomal protein L9 RplI	information pathways	COG0359J
Rv0066c	icd2	82550	Cytoplasmic	isocitrate dehydrogenase	Probable isocitrate dehydrogenase [NADP] Icd2 (oxalosuccinate decarboxylase) (IDH) (NADP+-specific ICDH) (IDP)	intermediary metabolism and respiration	COG2838C
Rv0079	-	29474.7	Cytoplasmic	hypothetical protein Rv0079	Unknown protein	conserved hypotheticals	-
Rv0088	-	24617.3	Unknown	polyketide cyclase/dehydrase	Possible polyketide cyclase/dehydrase	lipid metabolism	COG2867I
Rv0091	mtn	27339.6	Cytoplasmic Membrane	bifunctional 5'-methylthioadenosine nucleosidase /S-adenosylhomocysteine nucleosidase	Probable bifunctional MTA/SAH nucleosidase Mtn: 5'-methylthioadenosine nucleosidase (methylthioadenosine methylthioribohydrolase) + S-adenosylhomocysteine nucleosidase (S-adenosyl-L-homocysteine homocysteinylribohydrolase)	intermediary metabolism and respiration	COG0775F
Rv0119	fadD7	55158.8	Cytoplasmic	fatty-acid--CoA ligase FadD7	Probable fatty-acid-CoA ligase FadD7 (fatty-acid-CoA synthetase) (fatty-acid-CoA synthase)	lipid metabolism	COG0318IQ

Rv0120c	fusA2	75630.5	Cytoplasmic	elongation factor G	Probable elongation factor G FusA2 (EF-G)	information pathways	COG0480J
Rv0126	treS	68593	Cytoplasmic	trehalose synthase/amylase TreS	Trehalose synthase TreS	virulence, detoxification, adaptation	COG0366G
Rv0127	mak	49879.1	Cytoplasmic	maltokinase	Maltokinase Mak	virulence, detoxification, adaptation	COG3281G
Rv0131c	fadE1	50223.6	Cytoplasmic	acyl-CoA dehydrogenase FadE1	Probable acyl-CoA dehydrogenase FadE1	lipid metabolism	COG1960I
Rv0134	ephF	33796.7	Cytoplasmic	epoxide hydrolase EphF	Possible epoxide hydrolase EphF (epoxide hydratase) (arene-oxide hydratase)	virulence, detoxification, adaptation	COG0596R
Rv0136	cyp138	49260.6	Cytoplasmic	cytochrome P450 Cyp138	Probable cytochrome P450 138 Cyp138	intermediary metabolism and respiration	COG2124Q
Rv0146	-	34016.4	Cytoplasmic	S-adenosylmethionine-dependent methyltransferase	Possible S-adenosylmethionine-dependent methyltransferase	lipid metabolism	COG3315Q
Rv0147	-	55035.1	Cytoplasmic	aldehyde dehydrogenase	Probable aldehyde dehydrogenase (NAD ⁺) dependent	intermediary metabolism and respiration	COG1012C
Rv0148	-	29778.7	Cytoplasmic	short-chain type dehydrogenase/reductase	Probable short-chain type dehydrogenase/reductase	intermediary metabolism and respiration	COG1028IQR
Rv0149	-	33278.5	Cytoplasmic Membrane	quinone oxidoreductase	Possible quinone oxidoreductase (NADPH:quinone oxidoreductase) (zeta-crystallin)	intermediary metabolism and respiration	COG0604CR
Rv0154c	fadE2	44343.4	Cytoplasmic	acyl-CoA dehydrogenase FadE2	Probable acyl-CoA dehydrogenase FadE2	lipid metabolism	COG1960I
Rv0155	pntAa	37693.9	Cytoplasmic Membrane	NAD(P) transhydrogenase subunit alpha PntAa	Probable NAD(P) transhydrogenase (subunit alpha) PntAa [first part; catalytic part] (pyridine nucleotide transhydrogenase subunit alpha) (nicotinamide nucleotide transhydrogenase subunit alpha)	intermediary metabolism and respiration	COG3288C
Rv0157	pntB	48858.3	Cytoplasmic Membrane	NAD(P) transhydrogenase subunit beta PntB	Probable NAD(P) transhydrogenase (subunit beta) PntB [integral membrane protein] (pyridine nucleotide transhydrogenase subunit beta) (nicotinamide nucleotide transhydrogenase subunit beta)	intermediary metabolism and respiration	COG1282C
Rv0161	-	47355.9	Unknown	oxidoreductase	Possible oxidoreductase	intermediary metabolism and respiration	COG0277C
Rv0169	mce1A	47786.4	Cytoplasmic Membrane	Mce family protein Mce1A	Mce-family protein Mce1A	virulence, detoxification, adaptation	COG1463Q
Rv0170	mce1B	37727.2	Unknown	Mce family protein Mce1B	Mce-family protein Mce1B	virulence, detoxification, adaptation	COG1463Q
Rv0171	mce1C	53817.3	Cytoplasmic	Mce family protein Mce1C	Mce-family protein Mce1C	virulence, detoxification,	COG1463Q

						adaptation	
Rv0172	mce1D	56498.4	Unknown	Mce family protein Mce1D	Mce-family protein Mce1D	virulence, detoxification, adaptation	COG1463Q
Rv0173	lprK	42239.6	Cytoplasmic Membrane	Mce family lipoprotein LprK	Possible Mce-family lipoprotein LprK (Mce-family lipoprotein Mce1E)	cell wall and cell processes	COG1463Q
Rv0174	mce1F	54005.4	Unknown	Mce family protein Mce1F	Mce-family protein Mce1F	virulence, detoxification, adaptation	COG1463Q
Rv0180c	-	47593.4	Cytoplasmic Membrane	transmembrane protein	Probable conserved transmembrane protein	cell wall and cell processes	COG1511S
Rv0183	-	30261.6	Cytoplasmic Membrane	lysophospholipase	Possible lysophospholipase	intermediary metabolism and respiration	COG2267I
Rv0189c	ilvD	59351.7	Cytoplasmic	dihydroxy-acid dehydratase	Probable dihydroxy-acid dehydratase IlvD (dad)	intermediary metabolism and respiration	COG0129EG
Rv0206c	mmpL3	100872	Cytoplasmic Membrane	transmembrane transport protein MmpL3	Possible conserved transmembrane transport protein MmpL3	cell wall and cell processes	COG2409R
Rv0211	pckA	67253.1	Cytoplasmic	phosphoenolpyruvate carboxykinase	Probable iron-regulated phosphoenolpyruvate carboxykinase [GTP] PckA (phosphoenolpyruvate carboxylase) (PEPCK)(pep carboxykinase)	intermediary metabolism and respiration	COG1274C
Rv0222	echA1	27280.3	Cytoplasmic Membrane	enoyl-CoA hydratase EchA1	Probable enoyl-CoA hydratase EchA1 (enoyl hydratase) (unsaturated acyl-CoA hydratase) (crotonase)	lipid metabolism	COG1024I
Rv0223c	-	51157.4	Cytoplasmic	aldehyde dehydrogenase	Probable aldehyde dehydrogenase	intermediary metabolism and respiration	COG1012C
Rv0227c	-	45527.8	Unknown	membrane protein	Probable conserved membrane protein	cell wall and cell processes	-
Rv0231	fadE4	63494.8	Cytoplasmic	acyl-CoA dehydrogenase FadE4	Probable acyl-CoA dehydrogenase FadE4	lipid metabolism	COG1960I
Rv0234c	gabD1	54299.6	Cytoplasmic	succinate-semialdehyde dehydrogenase	Succinate-semialdehyde dehydrogenase [NADP+] dependent (SSDH) GabD1	intermediary metabolism and respiration	COG1012C
Rv0241c	htdX	30162.9	Cytoplasmic	3-hydroxyacyl-thioester dehydratase HtdX	Probable 3-hydroxyacyl-thioester dehydratase HtdX	intermediary metabolism and respiration	COG2030I
Rv0242c	fabG4	46798.1	Cytoplasmic	3-oxoacyl-ACP reductase FabG	Probable 3-oxoacyl-[acyl-carrier protein] reductase FabG4 (3-ketoacyl-acyl carrier protein reductase)	lipid metabolism	COG1028IQR
Rv0243	fadA2	46075.2	Cytoplasmic	acetyl-CoA acetyltransferase FadA	Probable acetyl-CoA acyltransferase FadA2 (3-ketoacyl-CoA thiolase) (beta-ketothiolase)	lipid metabolism	COG0183I
Rv0244c	fadE5	66014.2	Cytoplasmic	acyl-CoA dehydrogenase FadE5	Probable acyl-CoA dehydrogenase FadE5	lipid metabolism	COG1960I
Rv0248c	-	70680.6	Cytoplasmic Membrane	succinate dehydrogenase flavoprotein subunit	Probable succinate dehydrogenase [iron-sulfur subunit] (succinic dehydrogenase)	intermediary metabolism and respiration	COG1053C
Rv0266c	oplA	129618	Cytoplasmic	5-oxoprolinase OplA	Probable 5-oxoprolinase OplA (5-oxo-L-prolinase)	intermediary	COG0146EQ

					(pyroglutamase) (5-OPASE)	metabolism and respiration	
Rv0270	fadD2	59892.6	Cytoplasmic	fatty-acid--CoA ligase FadD2	Probable fatty-acid-CoA ligase FadD2 (fatty-acid-CoA synthetase) (fatty-acid-CoA synthase)	lipid metabolism	COG0318IQ
Rv0271c	fadE6	78477.2	Cytoplasmic	acyl-CoA dehydrogenase FadE6	Probable acyl-CoA dehydrogenase FadE6	lipid metabolism	COG1960I
Rv0281	-	32965.9	Cytoplasmic	S-adenosylmethionine-dependent methyltransferase	Possible S-adenosylmethionine-dependent methyltransferase	lipid metabolism	COG3315Q
Rv0282	eccA3	68106.6	Cytoplasmic	ESX-3 secretion system protein EccA	ESX conserved component EccA3. ESX-3 type VII secretion system protein.	cell wall and cell processes	COG0464O
Rv0283	eccB3	55942.5	Cytoplasmic Membrane	ESX-3 secretion system protein EccB3	ESX conserved component EccB3. ESX-3 type VII secretion system protein. Possible membrane protein.	cell wall and cell processes	-
Rv0284	eccC3	145138	Cytoplasmic Membrane	ESX-3 secretion system protein EccC3	ESX conserved component EccC3. ESX-3 type VII secretion system protein. Possible membrane protein.	cell wall and cell processes	COG1674D
Rv0295c	-	29774.6	Cytoplasmic	hypothetical protein Rv0295c	Conserved protein	conserved hypotheticals	COG4424S
Rv0296c	-	51814	Cytoplasmic	sulfatase	Probable sulfatase	intermediary metabolism and respiration	COG3119P
Rv0315	-	32186.1	Unknown	beta-1,3-glucanase	Possible beta-1,3-glucanase precursor	intermediary metabolism and respiration	COG2273G
Rv0334	rmlA	31492	Cytoplasmic	glucose-1-phosphate thymidyltransferase	Alpha-D-glucose-1-phosphate thymidyltransferase RmlA (dTDP-glucose synthase) (dTDP-glucose pyrophosphorylase)	intermediary metabolism and respiration	COG1209M
Rv0338c	-	95453.7	Cytoplasmic Membrane	iron-sulfur-binding reductase	Probable iron-sulfur-binding reductase	intermediary metabolism and respiration	COG0247C
Rv0342	iniA	70082.9	Cytoplasmic	isoniazid inductible protein IniA	Isoniazid inductible gene protein IniA	cell wall and cell processes	COG0699R
Rv0343	iniC	52986.9	Cytoplasmic Membrane	iIsoniazid inductible protein IniC	Isoniazid inductible gene protein IniC	cell wall and cell processes	COG0699R
Rv0350	dnaK	66830.9	Cytoplasmic	chaperone protein DnaK	Probable chaperone protein DnaK (heat shock protein 70) (heat shock 70 kDa protein) (HSP70)	virulence, detoxification, adaptation	COG0443O
Rv0352	dnaJ1	41345.1	Cytoplasmic	chaperone protein DnaJ	Probable chaperone protein DnaJ1	virulence, detoxification, adaptation	COG0484O
Rv0357c	purA	46822.5	Cytoplasmic	adenylosuccinate synthetase	Probable adenylosuccinate synthetase PurA (imp--aspartate ligase) (ADSS) (ampase)	intermediary metabolism and respiration	COG0104F
Rv0363c	fba	36544.5	Cytoplasmic	fructose-bisphosphate aldolase	Probable fructose-bisphosphate aldolase Fba	intermediary metabolism and respiration	COG0191G
Rv0379	secE2	7966.14	Cytoplasmic	protein translocase subunit SecE	Possible protein transport protein SecE2	cell wall and cell processes	COG3360S

Rv0383c	-	31801.4	Cytoplasmic	hypothetical protein Rv0383c	Possible conserved secreted protein	cell wall and cell processes	-
Rv0384c	clpB	92536.2	Cytoplasmic	chaperone protein ClpB	Probable endopeptidase ATP binding protein (chain B) ClpB (ClpB protein) (heat shock protein F84.1)	virulence, detoxification, adaptation	COG0542O
Rv0389	purT	43645.7	Cytoplasmic Membrane	phosphoribosylglycinamide formyltransferase PurT	Probable phosphoribosylglycinamide formyltransferase 2 PurT (GART 2) (gar transformylase 2) (5'-phosphoribosylglycinamide transformylase 2) (formate-dependent gar transformylase)	intermediary metabolism and respiration	COG0027F
Rv0391	metZ	43345.2	Cytoplasmic	O-succinylhomoserine sulfhydrylase	Probable O-succinylhomoserine sulfhydrylase MetZ (OSH sulfhydrylase)	intermediary metabolism and respiration	COG0626E
Rv0392c	ndhA	50385.3	Cytoplasmic Membrane	NADH dehydrogenase NdhA	Probable membrane NADH dehydrogenase NdhA	intermediary metabolism and respiration	COG1252C
Rv0400c	fadE7	42297.4	Cytoplasmic	acyl-CoA dehydrogenase FadE7	Acyl-CoA dehydrogenase FadE7	lipid metabolism	COG1960I
Rv0407	fgd1	36956.8	Cytoplasmic	F420-dependent glucose-6-phosphate dehydrogenase	F420-dependent glucose-6-phosphate dehydrogenase Fgd1	intermediary metabolism and respiration	COG2141C
Rv0410c	pknG	81577.4	Cytoplasmic Membrane	serine/threonine-protein kinase PknG	Serine/threonine-protein kinase PknG (protein kinase G) (STPK G)	Regulatory proteins	COG0515RTKL
Rv0411c	glnH	35405.7	Cytoplasmic Membrane	glutamine-binding lipoprotein GlnH	Probable glutamine-binding lipoprotein GlnH (GLNBP)	cell wall and cell processes	COG0834ET
Rv0419	lpqM	52949.1	Cytoplasmic Membrane	lipoprotein peptidase LpqM	Possible lipoprotein peptidase LpqM	cell wall and cell processes	COG2321R
Rv0423c	thiC	59897.7	Cytoplasmic	phosphomethylpyrimidine synthase	Probable thiamine biosynthesis protein ThiC	intermediary metabolism and respiration	COG0422H
Rv0432	sodC	23844.5	Cytoplasmic Membrane	superoxide dismutase	Periplasmic superoxide dismutase [Cu-Zn] SodC	virulence, detoxification, adaptation	COG2032P
Rv0433	-	42277.2	Cytoplasmic	carboxylate-amine ligase	Conserved hypothetical protein	conserved hypotheticals	COG2170S
Rv0435c	-	75314	Cytoplasmic Membrane	ATPase	Putative conserved ATPase	cell wall and cell processes	COG0464O
Rv0439c	-	33650.1	Unknown	dehydrogenase/reductase	Probable dehydrogenase/reductase	intermediary metabolism and respiration	COG1028IQR
Rv0440	groEL ₂	56726.7	Cytoplasmic	molecular chaperone GroEL	60 kDa chaperonin 2 GroEL2 (protein CPN60-2) (GroEL protein 2) (65 kDa antigen) (heat shock protein 65) (cell wall protein A) (antigen A)	virulence, detoxification, adaptation	COG0459O
Rv0444c	rskA	23883.1	Cytoplasmic Membrane	anti-sigma-K factor RskA	Anti-sigma factor RskA (regulator of sigma K)	information pathways	COG5343S
Rv0445c	sigK	21034.7	Cytoplasmic Membrane	ECF RNA polymerase sigma factor SigK	Alternative RNA polymerase sigma factor SigK	information pathways	COG1595K

Rv0450c	mmpL4	105202	Cytoplasmic Membrane	transmembrane transport protein MmpL4	Probable conserved transmembrane transport protein MmpL4	cell wall and cell processes	COG2409R
Rv0458	-	54574.8	Cytoplasmic	aldehyde dehydrogenase	Probable aldehyde dehydrogenase	intermediary metabolism and respiration	COG1012C
Rv0462	lpdC	49207.1	Cytoplasmic	dihydrolipoamide dehydrogenase	Dihydrolipoamide dehydrogenase LpdC (lipoamide reductase (NADH)) (lipoyl dehydrogenase) (dihydrolipoyl dehydrogenase) (diaphorase)	intermediary metabolism and respiration	COG1249C
Rv0464c	-	21304.6	Unknown	hypothetical protein Rv0464c	Conserved protein	conserved hypotheticals	COG2128S
Rv0466	-	30153.4	Cytoplasmic	hypothetical protein Rv0466	Acyl-ACP thioesterase	Intermediary metabolism and respiration	COG3884I
Rv0468	fadB2	30696.4	Cytoplasmic	3-hydroxybutyryl-CoA dehydrogenase	3-hydroxybutyryl-CoA dehydrogenase FadB2 (beta-hydroxybutyryl-CoA dehydrogenase) (BHBD)	lipid metabolism	COG1250I
Rv0469	umaA	33095.7	Cytoplasmic	mycolic acid synthase UmaA	Possible mycolic acid synthase UmaA	lipid metabolism	COG2230M
Rv0470c	pcaA	33027.7	Cytoplasmic	cyclopropane mycolic acid synthase	Mycolic acid synthase PcaA (cyclopropane synthase)	lipid metabolism	COG2230M
Rv0475	hbhA	21534.5	Cellwall	heparin binding hemagglutinin HbhA	Iron-regulated heparin binding hemagglutinin HbhA (adhesin)	cell wall and cell processes	COG1754R
Rv0479c	-	37016	Unknown	membrane protein	Probable conserved membrane protein	cell wall and cell processes	-
Rv0484c	-	26416.1	Cytoplasmic	short-chain type oxidoreductase	Probable short-chain type oxidoreductase	intermediary metabolism and respiration	COG4221R
Rv0489	gpm1	27215.8	Unknown	2,3-bisphosphoglycerate-dependent phosphoglycerate mutase	Probable phosphoglycerate mutase 1 Gpm1 (phosphoglyceromutase) (PGAM) (BPG-dependent PGAM)	intermediary metabolism and respiration	COG0588G
Rv0500	proC	30171.6	Cytoplasmic	pyrroline-5-carboxylate reductase	Probable pyrroline-5-carboxylate reductase ProC (P5CR) (P5C reductase)	intermediary metabolism and respiration	COG0345E
Rv0503c	cmaA2	34660.1	Cytoplasmic	cyclopropane mycolic acid synthase	Cyclopropane-fatty-acyl-phospholipid synthase 2 CmaA2 (cyclopropane fatty acid synthase) (CFA synthase) (cyclopropane mycolic acid synthase 2) (mycolic acid trans-cyclopropane synthetase)	lipid metabolism	COG2230M
Rv0505c	serB1	39508	Cytoplasmic	phosphoserine phosphatase SerB	Possible phosphoserine phosphatase SerB1 (PSP) (O-phosphoserine phosphohydrolase) (pspase)	intermediary metabolism and respiration	COG0560E
Rv0511	hemD	58835.6	Cytoplasmic	uroporphyrin-III C-methyltransferase	Probable uroporphyrin-III C-methyltransferase HemD (uroporphyrinogen III methylase) (urogen III methylase) (SUMT) (urogen III methylase) (UROM)	intermediary metabolism and respiration	COG0007H
Rv0524	hemL	47515.9	Cytoplasmic	glutamate-1-semialdehyde 2,1-aminomutase	Probable glutamate-1-semialdehyde 2,1-aminomutase HemL (GSA) (glutamate-1-semialdehyde aminotransferase) (GSA-at)	intermediary metabolism and respiration	COG0001H
Rv0530	-	43022.5	Cytoplasmic	hypothetical protein Rv0530	ATPases involved in chromosome partitioning	cell wall and cell processes	COG0455D

Rv0547c	-	31834.9	Cytoplasmic	oxidoreductase	Possible oxidoreductase	intermediary metabolism and respiration	COG0300R
Rv0548c	menB	34656.9	Cytoplasmic	1,4-dihydroxy-2-naphthoyl-CoA synthase	Naphthoate synthase MenB (dihydroxynaphthoic acid synthetase) (DHNA synthetase)	intermediary metabolism and respiration	COG0447H
Rv0554	bpoC	28349.4	Cytoplasmic	non-heme bromoperoxidase BpoC	Possible peroxidase BpoC (non-haem peroxidase)	virulence, detoxification, adaptation	COG0596R
Rv0558	menH	25298.9	Unknown	demethylmenaquinone methyltransferase	Probable ubiquinone/menaquinone biosynthesis methyltransferase MenH (2-heptaprenyl-1,4-naphthoquinone methyltransferase)	intermediary metabolism and respiration	COG2226H
Rv0562	grcC1	35528.3	Cytoplasmic	polyprenyl-diphosphate synthase GrcC	Probable polyprenyl-diphosphate synthase GrcC1 (polyprenyl pyrophosphate synthetase)	intermediary metabolism and respiration	COG0142H
Rv0564c	gpdA1	36153.4	Cytoplasmic	glycerol-3-phosphate dehydrogenase	Probable glycerol-3-phosphate dehydrogenase [NAD(P)+] GpdA1 (NAD(P)H-dependent glycerol-3-phosphate dehydrogenase) (NAD(P)H-dependent dihydroxyacetone-phosphate reductase)	lipid metabolism	COG0240C
Rv0576	-	46407.7	Cytoplasmic	transcriptional regulator	Probable transcriptional regulatory protein (possibly ArsR-family)	Regulatory proteins	COG3832S
Rv0577	TB27.3	27342.9	Unknown	hypothetical protein Rv0577	Predicted enzyme related to lactoylglutathione lyase	Intermediary metabolism and respiration	COG3324R
Rv0580c	-	18034.8	Cytoplasmic Membrane	hypothetical protein Rv0580c	Conserved protein	conserved hypotheticals	-
Rv0613c	-	92721.4	Cytoplasmic	hypothetical protein Rv0613c	Unknown protein	conserved hypotheticals	-
Rv0631c	recC	119501	Cytoplasmic Membrane	exonuclease V subunit gamma RecC	Probable exonuclease V (gamma chain) RecC (exodeoxyribonuclease V gamma chain)(exodeoxyribonuclease V polypeptide)	information pathways	COG1330L
Rv0632c	echA3	24354.9	Cytoplasmic	enoyl-CoA hydratase EchA3	Probable enoyl-CoA hydratase EchA3 (enoyl hydratase) (unsaturated acyl-CoA hydratase) (crotonase)	lipid metabolism	COG1024I
Rv0635	hadA	17448.5	Cytoplasmic	(3R)-hydroxyacyl-ACP dehydratase subunit HadA	(3R)-hydroxyacyl-ACP dehydratase subunit HadA	intermediary metabolism and respiration	COG2030I
Rv0636	hadB	14934	Unknown	(3R)-hydroxyacyl-ACP dehydratase subunit HadB	(3R)-hydroxyacyl-ACP dehydratase subunit HadB	intermediary metabolism and respiration	COG2030I
Rv0637	hadC	18929.7	Cytoplasmic	(3R)-hydroxyacyl-ACP dehydratase subunit HadC	(3R)-hydroxyacyl-ACP dehydratase subunit HadC	intermediary metabolism and respiration	COG2030I
Rv0639	nusG	25414.5	Unknown	transcription termination/antitermination protein NusG	Probable transcription antitermination protein NusG	information pathways	COG0250K

Rv0641	rplA	24756.4	Cytoplasmic	50S ribosomal protein L1	50S ribosomal protein L1 RplA	information pathways	COG0081J
Rv0642c	mmaA4	34670.4	Cytoplasmic	hydroxymycolate synthase MmaA4	Methoxy mycolic acid synthase 4 MmaA4 (methyl mycolic acid synthase 4) (MMA4) (hydroxy mycolic acid synthase)	lipid metabolism	COG2230M
Rv0643c	mmaA3	33262.9	Cytoplasmic	methoxy mycolic acid synthase MmaA3	Methoxy mycolic acid synthase 3 MmaA3 (methyl mycolic acid synthase 3) (MMA3) (hydroxy mycolic acid synthase)	lipid metabolism	COG2230M
Rv0644c	mmaA2	32724.3	Cytoplasmic	cyclopropane mycolic acid synthase MmaA2	Methoxy mycolic acid synthase 2 MmaA2 (methyl mycolic acid synthase 2) (MMA2) (hydroxy mycolic acid synthase)	lipid metabolism	COG2230M
Rv0645c	mmaA1	33149.4	Cytoplasmic	mycolic acid methyltransferase MmaA1	Methoxy mycolic acid synthase 1 MmaA1 (methyl mycolic acid synthase 1) (MMA1) (hydroxy mycolic acid synthase)	lipid metabolism	COG2230M
Rv0646c	lipG	32931	Unknown	lipase/esterase LipG	Probable lipase/esterase LipG	intermediary metabolism and respiration	COG0596R
Rv0647c	-	54824	Cytoplasmic Membrane	hypothetical protein Rv0647c	Predicted unusual protein kinase	Intermediary metabolism and respiration	COG0661R
Rv0651	rplJ	18478.3	Cytoplasmic	50S ribosomal protein L10	50S ribosomal protein L10 RplJ	information pathways	COG0244J
Rv0652	rplL	13440.5	Cytoplasmic	50S ribosomal protein L7/L12	50S ribosomal protein L7/L12 RplL (SA1)	information pathways	COG0222J
Rv0655	mkl	39362.2	Cytoplasmic Membrane	ribonucleotide ABC transporter ATP-binding protein	Possible ribonucleotide-transport ATP-binding protein ABC transporter Mkl	cell wall and cell processes	COG1127Q
Rv0667	rpoB	129218	Cytoplasmic	DNA-directed RNA polymerase subunit beta	DNA-directed RNA polymerase (beta chain) RpoB (transcriptase beta chain) (RNA polymerase beta subunit)	information pathways	COG0085K
Rv0668	rpoC	146737	Cytoplasmic	DNA-directed RNA polymerase subunit beta'	DNA-directed RNA polymerase (beta' chain) RpoC (transcriptase beta' chain) (RNA polymerase beta' subunit).	information pathways	COG0086K
Rv0669c	-	69490.1	Cytoplasmic	neutral ceramidase	Possible hydrolase	intermediary metabolism and respiration	-
Rv0672	fadE8	58534.5	Cytoplasmic	acyl-CoA dehydrogenase FadE8	Probable acyl-CoA dehydrogenase FadE8	lipid metabolism	COG1960I
Rv0675	echA5	27441.4	Cytoplasmic	enoyl-CoA hydratase EchA5	Probable enoyl-CoA hydratase EchA5 (enoyl hydrase) (unsaturated acyl-CoA hydratase) (crotonase)	lipid metabolism	COG1024I
Rv0683	rpsG	17600.4	Cytoplasmic	30S ribosomal protein S7	30S ribosomal protein S7 RpsG	information pathways	COG0049J
Rv0684	fusA1	77170.5	Cytoplasmic	elongation factor G	Probable elongation factor G FusA1 (EF-G)	information pathways	COG0480J
Rv0685	tuf	43561.5	Cytoplasmic	elongation factor Tu	Probable iron-regulated elongation factor TU Tuf (EF-TU)	information pathways	COG0050J
Rv0686	-	28543.2	Cytoplasmic Membrane	membrane protein	Probable membrane protein	cell wall and cell processes	-
Rv0688	-	43017.8	Cytoplasmic	ferredoxin reductase	Putative ferredoxin reductase	intermediary metabolism and respiration	COG1251C
Rv0694	-	42129.2	Cytoplasmic	mycofactocin system heme/flavin	Possible L-lactate dehydrogenase (cytochrome) LldD1	intermediary	COG1304C

			Membrane	oxidoreductase MftD		metabolism and respiration	
Rv0700	rpsJ	11399.2	Cytoplasmic	30S ribosomal protein S10	30S ribosomal protein S10 RpsJ (transcription antitermination factor NusE)	information pathways	COG0051J
Rv0701	rplC	23090.5	Cytoplasmic	50S ribosomal protein L3	50S ribosomal protein L3 RplC	information pathways	COG0087J
Rv0702	rplD	23742.9	Cytoplasmic	50S ribosomal protein L4	50S ribosomal protein L4 RplD	information pathways	COG0088J
Rv0704	rplB	30576.8	Cytoplasmic	50S ribosomal protein L2	50S ribosomal protein L2 RplB	information pathways	COG0090J
Rv0706	rplV	20380.5	Cytoplasmic	50S ribosomal protein L22	50S ribosomal protein L22 RplV	information pathways	COG0091J
Rv0707	rpsC	29987.9	Cytoplasmic	30S ribosomal protein S3	30S ribosomal protein S3 RpsC	information pathways	COG0092J
Rv0708	rplP	15692.2	Cytoplasmic	50S ribosomal protein L16	50S ribosomal protein L16 RplP	information pathways	COG0197J
Rv0711	atsA	86214.5	Cytoplasmic	arylsulfatase AtsA	Possible arylsulfatase AtsA (aryl-sulfate sulphohydrolase) (arylsulphatase)	intermediary metabolism and respiration	COG3119P
Rv0716	rplE	21019.3	Cytoplasmic	50S ribosomal protein L5	50S ribosomal protein L5 RplE	information pathways	COG0094J
Rv0719	rplF	19377.3	Cytoplasmic	50S ribosomal protein L6	50S ribosomal protein L6 RplF	information pathways	COG0097J
Rv0721	rpsE	22888.2	Cytoplasmic	30S ribosomal protein S5	30S ribosomal protein S5 RpsE	information pathways	COG0098J
Rv0724	sppA	65935.9	Cytoplasmic Membrane	protease IV SppA	Possible protease IV SppA (endopeptidase IV) (signal peptide peptidase)	cell wall and cell processes	COG0616OU
Rv0730	-	25980.1	Cytoplasmic	GCN5-like N-acetyltransferase	GCN5-related N-acetyltransferase	intermediary metabolism and respiration	COG0188L
Rv0731c	-	34950	Cytoplasmic	S-adenosylmethionine-dependent methyltransferase	Possible S-adenosylmethionine-dependent methyltransferase	lipid metabolism	COG3315Q
Rv0733	adk	20092.8	Cytoplasmic	adenylate kinase	Adenylate kinase Adk (ATP-AMP transphosphorylase)	intermediary metabolism and respiration	COG0563F
Rv0753c	mmsA	54453.9	Cytoplasmic	methylmalonate-semialdehyde dehydrogenase	Probable methylmalonate-semialdehyde dehydrogenase MmsA (methylmalonic acid semialdehyde dehydrogenase) (MMSDH)	intermediary metabolism and respiration	COG1012C
Rv0757	phoP	27513.5	Cytoplasmic	two component system response transcriptional positive regulator PhoP	Possible two component system response transcriptional positive regulator PhoP	Regulatory proteins	COG0745TK
Rv0761c	adhB	39715.6	Cytoplasmic	alcohol dehydrogenase B	Possible zinc-containing alcohol dehydrogenase NAD dependent AdhB	intermediary metabolism and respiration	COG1062C
Rv0772	purD	43476.6	Cytoplasmic	phosphoribosylamine--glycine ligase	Probable phosphoribosylamine--glycine ligase PurD (GARS) (glycinamide ribonucleotide synthetase) (phosphoribosylglycinamide synthetase) (5'-phosphoribosylglycinamide synthetase)	intermediary metabolism and respiration	COG0151F
Rv0777	purB	51008.4	Cytoplasmic	adenylosuccinate lyase PurB	Probable adenylosuccinate lyase PurB (adenylosuccinase) (ASL) (ASASE)	intermediary metabolism and respiration	COG0015F
Rv0780	purC	32930.2	Cytoplasmic	phosphoribosylaminoimidazole-succinocarboxamide synthase	Phosphoribosylaminoimidazole- succinocarboxamide synthase PurC (SAICAR synthetase)	intermediary metabolism and	COG0152F

						respiration	
Rv0785	-	61219.7	Cytoplasmic	KsdD-like steroid dehydrogenase	Predicted oxidoreductase	Intermediary metabolism and respiration	COG3573R
Rv0788	purQ	23600.8	Cytoplasmic	phosphoribosylformylglycinamide synthase	Probable phosphoribosylformylglycinamide synthase I PURG (FGAM synthase I)	intermediary metabolism and respiration	COG0047F
Rv0794c	-	52476.4	Cytoplasmic	oxidoreductase	Probable oxidoreductase	intermediary metabolism and respiration	COG1249C
Rv0798c	cfp29	28830.4	Unknown	hypothetical protein Rv0798c	29 KDa antigen CFP29	virulence, detoxification, adaptation	COG1659S
Rv0803	purL	79475	Cytoplasmic	phosphoribosylformylglycinamide synthase 2	Phosphoribosylformylglycinamide synthase II PurL (FGAM synthase II)	intermediary metabolism and respiration	COG0046F
Rv0809	purM	38393.3	Cytoplasmic	phosphoribosylformylglycinamide cyclo-ligase PurM	Probable phosphoribosylformylglycinamide CYCLO-ligase PurM (AIRS) (phosphoribosyl-aminoimidazole synthetase) (air synthase)	intermediary metabolism and respiration	COG0150F
Rv0818	-	27579.4	Cytoplasmic	transcriptional regulator	Transcriptional regulatory protein	Regulatory proteins	COG0745TK
Rv0821c	phoY2	23513.8	Cytoplasmic	phosphate-transport system transcriptional regulator PhoY2	Probable phosphate-transport system transcriptional regulatory protein PhoY2	cell wall and cell processes	COG0704P
Rv0824c	desA1	38770.2	Cytoplasmic	acyl-ACP desaturase DesA	Probable acyl-[acyl-carrier protein] desaturase DesA1 (acyl-[ACP] desaturase) (stearoyl-ACP desaturase) (protein Des)	lipid metabolism	-
Rv0831c	-	30189.2	Unknown	hypothetical protein Rv0831c	Conserved protein	conserved hypotheticals	-
Rv0853c	pdc	59751.1	Cytoplasmic Membrane	alpha-keto-acid decarboxylase	Probable pyruvate or indole-3-pyruvate decarboxylase Pdc	intermediary metabolism and respiration	COG3961GHR
Rv0859	-	42414.5	Cytoplasmic	acyltransferase	Possible acyl-CoA thiolase FadA	lipid metabolism	COG0183I
Rv0860	fadB	76104.1	Cytoplasmic	fatty oxidation protein FadB	Probable fatty oxidation protein FadB	lipid metabolism	COG1250I
Rv0861c	ercc3	59772.2	Cytoplasmic	DNA helicase Ercc3	DNA helicase Ercc3	information pathways	-
Rv0862c	-	79666.5	Cytoplasmic	hypothetical protein Rv0862c	Conserved protein	conserved hypotheticals	COG2378K
Rv0871	cspB	14989	Cytoplasmic	cold shock-like protein CspB	Probable cold shock-like protein B CspB	virulence, detoxification, adaptation	COG1278K
Rv0873	fadE10	70744.6	Cytoplasmic	acyl-CoA dehydrogenase FadE10	Probable acyl-CoA dehydrogenase FadE10	lipid metabolism	COG1960I
Rv0884c	serC	40233.4	Unknown	phosphoserine aminotransferase	Possible phosphoserine aminotransferase SerC (PSAT)	intermediary metabolism and respiration	COG1932HE
Rv0889c	citA	40146.8	Cytoplasmic	citrate synthase 2	Probable citrate synthase II CitA	intermediary metabolism and	COG0372C

						respiration	
Rv0896	gltA2	47946.4	Cytoplasmic	citrate synthase 1	Probable citrate synthase I GltA2	intermediary metabolism and respiration	COG0372C
Rv0899	arfA	33542.1	Cytoplasmic Membrane	peptidoglycan-binding protein ArfA	Outer membrane protein A OmpA	cell wall and cell processes	COG2885M
Rv0902c	prfB	47829.3	Cytoplasmic Membrane	two component sensor histidine kinase PrrB	Two component sensor histidine kinase PrrB	Regulatory proteins	COG0642T
Rv0903c	prfA	25253.9	Cytoplasmic	two component transcriptional regulator PrrA	Two component response transcriptional regulatory protein PrrA	Regulatory proteins	COG0745TK
Rv0905	echA6	26028.8	Cytoplasmic	enoyl-CoA hydratase EchA6	Possible enoyl-CoA hydratase EchA6 (enoyl hydratase) (unsaturated acyl-CoA hydratase) (crotonase)	lipid metabolism	COG1024I
Rv0906	-	40641.2	Unknown	hypothetical protein Rv0906	Predicted Zn-dependent hydrolases of the beta-lactamase fold	Intermediary metabolism and respiration	COG2220R
Rv0925c	-	26944.3	Unknown	hypothetical protein Rv0925c	Multimeric flavodoxin WrbA	Intermediary metabolism and respiration	COG0655R
Rv0926c	-	37751.3	Cytoplasmic	hypothetical protein Rv0926c	Uncharacterized conserved protein related to dihydrodipicolinate reductase	Intermediary metabolism and respiration	COG3804S
Rv0927c	-	26745.7	Cytoplasmic	oxidoreductase	Probable short-chain type dehydrogenase/reductase	intermediary metabolism and respiration	COG1028IQR
Rv0928	pstS3	37935.1	Unknown	phosphate ABC transporter substrate-binding lipoprotein PstS	Periplasmic phosphate-binding lipoprotein PstS3 (PBP-3) (PstS3) (PHOS1)	cell wall and cell processes	COG0226P
Rv0931c	pknD	69513.2	Cytoplasmic Membrane	serine/threonine-protein kinase PknD	Transmembrane serine/threonine-protein kinase D PknD (protein kinase D) (STPK D)	Regulatory proteins	COG0515RTKL
Rv0932c	pstS2	37832	Extracellular	phosphate ABC transporter substrate-binding lipoprotein PstS	Periplasmic phosphate-binding lipoprotein PstS2 (PBP-2) (PstS2)	cell wall and cell processes	COG0226P
Rv0933	pstB	29996.4	Cytoplasmic Membrane	phosphate ABC transporter ATP-binding protein PstB	Phosphate-transport ATP-binding protein ABC transporter PstB	cell wall and cell processes	COG1117P
Rv0934	pstS1	38211.1	Unknown	phosphate ABC transporter substrate-binding lipoprotein PstS	Periplasmic phosphate-binding lipoprotein PstS1 (PBP-1) (PstS1)	cell wall and cell processes	COG0226P
Rv0946c	pgi	59974.4	Cytoplasmic	glucose-6-phosphate isomerase	Probable glucose-6-phosphate isomerase Pgi (GPI) (phosphoglucose isomerase) (phosphohexose isomerase) (phi)	intermediary metabolism and respiration	COG0166G
Rv0951	sucC	40925.7	Cytoplasmic	succinyl-CoA ligase subunit beta	Probable succinyl-CoA synthetase (beta chain) SucC (SCS-beta)	intermediary metabolism and respiration	COG0045C
Rv0952	sucD	31228.9	Unknown	succinyl-CoA ligase subunit alpha	Probable succinyl-CoA synthetase (alpha chain) SucD (SCS-alpha)	intermediary metabolism and respiration	COG0074C
Rv0957	purH	55026.2	Cytoplasmic	bifunctional phosphoribosylaminoimidazolecarbox	Probable bifunctional purine biosynthesis protein PurH: phosphoribosylaminoimidazolecarboxamide	intermediary metabolism and	COG0138F

				amide formyltransferase/inosinemonospho hate cyclohydrolase	formyltransferase (AICAR transformylase) (5'- phosphoribosyl-5-aminoimidazole-4-carboxamide formyltransferase) + inosinemonophosphate cyclohydrolase (imp cyclohydrolase) (inosinicase) (imp synthetase) (ATIC)	respiration	
Rv0958	-	49581.4	Cytoplasmic	magnesium chelatase	Possible magnesium chelatase	intermediary metabolism and respiration	COG1239H
Rv0968	-	10263.8	Unknown	hypothetical protein Rv0968	Conserved protein	conserved hypotheticals	-
Rv0969	ctpV	80102.5	Cytoplasmic Membrane	copper transporter ATPase V	Probable metal cation transporter P-type ATPase CtpV	cell wall and cell processes	-
Rv0981	mprA	25675.4	Cytoplasmic	two-component response regulator MrpA	Mycobacterial persistence regulator MRPA (two component response transcriptional regulatory protein)	Regulatory proteins	-
Rv0984	moaB 2	18441	Cytoplasmic	pterin-4-alpha-carbinolamine dehydratase	Possible pterin-4-alpha-carbinolamine dehydratase MoaB2 (PHS) (4-alpha- hydroxy-tetrahydropterin dehydratase) (pterin-4-a-carbinolamine dehydratase) (phenylalanine hydroxylase-stimulating protein) (PHS) (pterin carbinolamine dehydratase) (PCD)	intermediary metabolism and respiration	COG0521H
Rv0994	moeA 1	44304.8	Cytoplasmic	molybdopterin molybdenumtransferase 1	Probable molybdopterin biosynthesis protein MoeA1	intermediary metabolism and respiration	COG0303H
Rv1001	arcA	43057.1	Cytoplasmic	arginine deiminase	Probable arginine deiminase ArcA (adi) (ad) (arginine dihydrolase)	intermediary metabolism and respiration	COG2235E
Rv1006	-	61297.6	Unknown	hypothetical protein Rv1006	Unknown protein	conserved hypotheticals	-
Rv1007c	metS	58092.6	Cytoplasmic	methionine--tRNA ligase	Methionyl-tRNA synthetase MetS (MetRS) (methionine- tRNA ligase)	information pathways	COG0143J
Rv1015c	rplY	22441.2	Cytoplasmic	50S ribosomal protein L25	50S ribosomal protein L25 RplY	information pathways	COG1825J
Rv1017c	prsA	35459.4	Cytoplasmic	ribose-phosphate pyrophosphokinase	Probable ribose-phosphate pyrophosphokinase PrsA (phosphoribosyl pyrophosphate synthetase) (PRPP synthetase)	intermediary metabolism and respiration	COG0462FE
Rv1018c	glmU	51584	Cytoplasmic	bifunctional UDP-N- acetylglucosamine pyrophosphorylase/glucosamine-1- phosphate N-acetyltransferase	Probable UDP-N-acetylglucosamine pyrophosphorylase GlmU	cell wall and cell processes	COG1207M
Rv1019	-	21701.5	Unknown	transcriptional regulator	Probable transcriptional regulatory protein (probably TetR-family)	Regulatory proteins	COG1309K
Rv1023	eno	44929.5	Cytoplasmic	enolase	Probable enolase Eno	intermediary metabolism and respiration	COG0148G
Rv1056	-	28945.6	Cytoplasmic	hypothetical protein Rv1056	Conserved protein	conserved hypotheticals	COG2343S
Rv1058	fadD1	59624.6	Cytoplasmic	fatty-acid--CoA ligase FadD14	Probable medium chain fatty-acid-CoA ligase FadD14	lipid metabolism	COG0318IQ

	4				(fatty-acid-CoA synthetase) (fatty-acid-CoA synthase)		
Rv1069c	-	64341.5	Cytoplasmic Membrane	hypothetical protein Rv1069c	Predicted membrane protein	cell wall and cell processes	COG4425S
Rv1070c	echA8	27273.1	Cytoplasmic Membrane	enoyl-CoA hydratase EchA8	Probable enoyl-CoA hydratase EchA8 (enoyl hydratase) (unsaturated acyl-CoA hydratase) (crotonase)	lipid metabolism	COG1024I
Rv1071c	echA9	36322	Cytoplasmic	enoyl-CoA hydratase EchA9	Possible enoyl-CoA hydratase EchA9 (enoyl hydratase) (unsaturated acyl-CoA hydratase) (crotonase)	lipid metabolism	COG1024I
Rv1074c	fadA3	42655.4	Cytoplasmic	beta-ketoacyl CoA thiolase FadA	Probable beta-ketoacyl CoA thiolase FadA3	lipid metabolism	COG0183I
Rv1077	cbs	48635.2	Cytoplasmic	cystathionine beta-synthase	Probable cystathionine beta-synthase Cbs (serine sulfhydrase) (beta-thionase) (hemoprotein H-450)	intermediary metabolism and respiration	COG0031E
Rv1078	-	25076	Cytoplasmic Membrane	hypothetical protein Rv1078	cell wall and cell processes , Predicted membrane protein/domain, proline-rich antigen homolog	cell wall and cell processes	COG5178A
Rv1079	metB	40982.3	Cytoplasmic	cystathionine gamma-synthase	Cystathionine gamma-synthase MetB (CGS) (O-succinylhomoserine [thiol]-lyase)	intermediary metabolism and respiration	COG0626E
Rv1093	glyA1	45029.4	Cytoplasmic	serine hydroxymethyltransferase	Serine hydroxymethyltransferase 1 GlyA1	intermediary metabolism and respiration	-
Rv1094	desA2	31359.5	Cytoplasmic	acyl-ACP desaturase DesA	Possible acyl-[acyl-carrier protein] desaturase DesA2 (acyl-[ACP] desaturase) (stearoyl-ACP desaturase)	lipid metabolism	COG0208F
Rv1095	phoH2	46881.5	Cytoplasmic	phosphate starvation-inducible protein PsiH	Probable PHOH-like protein PhoH2 (phosphate starvation-inducible protein PSIH)	intermediary metabolism and respiration	COG1875T
Rv1098c	fum	50141	Cytoplasmic	fumarate hydratase	Probable fumarase Fum (fumarate hydratase)	intermediary metabolism and respiration	COG0114C
Rv1099c	glpX	38084	Cytoplasmic	fructose 1,6-bisphosphatase	Fructose 1,6-bisphosphatase GlpX	intermediary metabolism and respiration	-
Rv1106c	-	40741.6	Unknown	3 beta-hydroxysteroid dehydrogenase/delta 5--_4-isomerase	3-beta-hydroxysteroid dehydrogenase	intermediary metabolism and respiration	COG0451MG
Rv1109c	-	22957.2	Unknown	hypothetical protein Rv1109c	Conserved protein	conserved hypotheticals	-
Rv1112	-	38065.2	Cytoplasmic	GTP-binding protein	Predicted GTPase, probable translation factor	Information pathways	COG0012J
Rv1118c	-	32157.5	Cytoplasmic	hypothetical protein Rv1118c	Conserved protein	conserved hypotheticals	-
Rv1121	zwf1	52155.3	Cytoplasmic	glucose-6-phosphate 1-dehydrogenase	Probable glucose-6-phosphate 1-dehydrogenase Zwf1 (G6PD)	intermediary metabolism and respiration	COG0364G
Rv1122	gnd2	36358.1	Cytoplasmic	6-phosphogluconate dehydrogenase	Probable 6-phosphogluconate dehydrogenase, decarboxylating Gnd2	intermediary metabolism and respiration	COG1023G
Rv1124	ephC	34694.6	Cytoplasmic	epoxide hydrolase EphC	Probable epoxide hydrolase EphC (epoxide hydratase)	virulence,	COG0596R

						detoxification, adaptation	
Rv1130	prpD	57750.6	Cytoplasmic	methylcitrate dehydratase PrpD	Possible methylcitrate dehydratase PrpD	intermediary metabolism and respiration	COG2079R
Rv1133c	metE	81549.5	Cytoplasmic	5-methyltetrahydropteroyltrimethylglutamate--homocysteine methyltransferase	Probable 5-methyltetrahydropteroyltrimethylglutamate--homocysteine methyltransferase MetE (methionine synthase, vitamin-B12 independent isozyme)	intermediary metabolism and respiration	COG0620E
Rv1144	-	25787.8	Cytoplasmic	oxidoreductase	Probable short-chain type dehydrogenase/reductase	intermediary metabolism and respiration	COG1028IQR
Rv1161	narG	136924	Cytoplasmic Membrane	nitrate reductase subunit alpha	Respiratory nitrate reductase (alpha chain) NarG	intermediary metabolism and respiration	COG5013C
Rv1162	narH	62037.8	Cytoplasmic Membrane	nitrate reductase subunit beta	Probable respiratory nitrate reductase (beta chain) NarH	intermediary metabolism and respiration	COG1140C
Rv1165	typA	67548.2	Cytoplasmic	GTP-binding translation elongation factor	Possible GTP-binding translation elongation factor TypA (tyrosine phosphorylated protein A) (GTP-binding protein)	information pathways	COG1217T
Rv1170	mshB	31742.3	Unknown	1D-myo-inositol 2-acetamido-2-deoxy-alpha-D-glucopyranoside deacetylase	N-acetyl-1-D-myo-inositol-2-amino-2-deoxy-alpha-D-glucopyranoside deacetylase MshB (GlcNAc-Ins deacetylase)	intermediary metabolism and respiration	COG2120S
Rv1173	fbiC	92474.2	Cytoplasmic	FO synthase	Probable F420 biosynthesis protein FbiC	intermediary metabolism and respiration	COG1060HR
Rv1175c	fadH	72901.5	Cytoplasmic	NADPH dependent 2,4-dienoyl-CoA reductase FadH	Probable NADPH dependent 2,4-dienoyl-CoA reductase FadH (2,4-dienoyl coenzyme A reductase) (4-enoyl-CoA reductase)	lipid metabolism	COG1902C
Rv1179c	-	100642	Cytoplasmic Membrane	hypothetical protein Rv1179c	DNA or RNA helicases of superfamily II	Information pathways	COG1061KL
Rv1180	pks3	50566.5	Cytoplasmic Membrane	polyketide beta-ketoacyl synthase	Probable polyketide beta-ketoacyl synthase Pks3	lipid metabolism	COG3321Q
Rv1181	pks4	168084	Cytoplasmic Membrane	polyketide beta-ketoacyl synthase	Probable polyketide beta-ketoacyl synthase Pks4	lipid metabolism	COG3321Q
Rv1185c	fadD21	62756.9	Cytoplasmic	fatty-acid--CoA ligase FadD21	Probable fatty-acid-AMP ligase FadD21 (fatty-acid-AMP synthetase) (fatty-acid-AMP synthase)	lipid metabolism	COG0318IQ
Rv1191	-	32611.6	Cytoplasmic Membrane	hypothetical protein Rv1191	Predicted hydrolases or acyltransferases (alpha/beta hydrolase superfamily)	Intermediary metabolism and respiration	COG0596R
Rv1192	-	30212.8	Cytoplasmic	hypothetical protein Rv1192	Protein involved in GPI anchor metabolic process	cell wall and cell processes	COG0596R
Rv1194c	-	46066	Cytoplasmic	hypothetical protein Rv1194c	Regulator of polyketide synthase expression	Lipid metabolism	COG2508TQ
Rv1196	PPE18	39158.4	Cytoplasmic Membrane	PPE family protein PPE18	PPE family protein PPE18	PE/PPE	COG5651N

Rv1198	esxL	9927.97	Unknown	ESAT-6 like protein EsxL	Putative ESAT-6 like protein EsxL (ESAT-6 like protein 4)	cell wall and cell processes	-
Rv1201c	dapD	32609.8	Unknown	2,3,4,5-tetrahydropyridine-2,6-dicarboxylate N-succinyltransferase	Tetrahydrodipicolinate N-succinyltransferase DapD	intermediary metabolism and respiration	COG2171E
Rv1206	fadD6	64305	Cytoplasmic	fatty-acid--CoA ligase FadD6	Probable fatty-acid-CoA ligase FadD6 (fatty-acid-CoA synthetase) (fatty-acid-CoA synthase)	lipid metabolism	COG0318IQ
Rv1215c	-	62578.9	Cytoplasmic	hypothetical protein Rv1215c	Predicted acyl esterases	Intermediary metabolism and respiration	COG2936R
Rv1218c	-	33448.1	Cytoplasmic Membrane	tetronasin ABC transporter ATP-binding protein	Probable tetronasin-transport ATP-binding protein ABC transporter	cell wall and cell processes	COG1131V
Rv1220c	-	22107.1	Cytoplasmic	methyltransferase	Probable methyltransferase	intermediary metabolism and respiration	COG4122R
Rv1223	htrA	54157.6	Unknown	serine protease HtrA	Probable serine protease HtrA (DEGP protein)	intermediary metabolism and respiration	COG0265O
Rv1229c	mrp	41063.3	Cytoplasmic	Mrp-like protein	Probable Mrp-related protein Mrp	intermediary metabolism and respiration	COG0489D
Rv1232c	-	47374.4	Cytoplasmic	hypothetical protein Rv1232c	Mg/Co/Ni transporter MgtE (contains CBS domain)	cell wall and cell processes	COG2239P
Rv1234	-	18604.5	Cytoplasmic Membrane	transmembrane protein	Probable transmembrane protein	cell wall and cell processes	-
Rv1235	lpqY	49761.4	Unknown	trehalose ABC transporter substrate-binding lipoprotein LpqY	Probable sugar-binding lipoprotein LpqY	cell wall and cell processes	COG1653G
Rv1238	sugC	42915.2	Cytoplasmic Membrane	sugar ABC transporter ATP-binding protein SugC	Probable sugar-transport ATP-binding protein ABC transporter SugC	cell wall and cell processes	COG3839G
Rv1239c	corA	41468.8	Cytoplasmic Membrane	magnesium and cobalt transport transmembrane protein CorA	Possible magnesium and cobalt transport transmembrane protein CorA	cell wall and cell processes	COG0598P
Rv1240	mdh	34289.7	Unknown	malate dehydrogenase	Probable malate dehydrogenase Mdh	intermediary metabolism and respiration	COG0039C
Rv1245c	-	29215.3	Cytoplasmic	short-chain type dehydrogenase/reductase	Probable short-chain type dehydrogenase/reductase	intermediary metabolism and respiration	COG4221R
Rv1248c	-	135870	Cytoplasmic	multifunctional 2-oxoglutarate dehydrogenase E1 component /2-oxoglutarate dehydrogenase dihydrolipoyllysine-residue succinyltransferase	Multifunctional alpha-ketoglutarate metabolic enzyme	intermediary metabolism and respiration	COG0567C
Rv1257c	-	47237.9	Cytoplasmic	oxidoreductase	Probable oxidoreductase	intermediary metabolism and respiration	COG0277C

Rv1260	-	41297.2	Cytoplasmic	oxidoreductase	Probable oxidoreductase	intermediary metabolism and respiration	COG0654HC
Rv1265	-	25229.7	Cytoplasmic	hypothetical protein Rv1265	Unknown protein	conserved hypotheticals	-
Rv1267c	embR	41933.5	Cytoplasmic	transcriptional regulator EmbR	Probable transcriptional regulatory protein EmbR	Regulatory proteins	COG3629T
Rv1270c	lprA	24874	Cytoplasmic Membrane	lipoprotein LprA	Possible lipoprotein LprA	cell wall and cell processes	-
Rv1279	-	57331.2	Cytoplasmic Membrane	GMC-type oxidoreductase	Probable dehydrogenase FAD flavoprotein GMC oxidoreductase	intermediary metabolism and respiration	COG2303E
Rv1280c	oppA	63444.5	Unknown	oligopeptide ABC transporter substrate-binding lipoprotein OppA	Probable periplasmic oligopeptide-binding lipoprotein OppA	cell wall and cell processes	COG0747E
Rv1281c	oppD	65293	Cytoplasmic Membrane	oligopeptide ABC transporter ATP-binding protein OppD	Probable oligopeptide-transport ATP-binding protein ABC transporter OppD	cell wall and cell processes	COG0444EP
Rv1284	canA	18156.7	Cytoplasmic	beta-carbonic anhydrase	Beta-carbonic anhydrase	intermediary metabolism and respiration	COG0288P
Rv1288	-	49618.9	Cytoplasmic Membrane	hypothetical protein Rv1288	FOG- LysM repeat	Regulatory proteins	COG0627R
Rv1292	argS	59677.4	Cytoplasmic	arginine--tRNA ligase	Probable arginyl-tRNA synthetase ArgS (ARGRS) (arginine--tRNA ligase)	information pathways	COG0018J
Rv1294	thrA	45520.7	Cytoplasmic	homoserine dehydrogenase	Probable homoserine dehydrogenase ThrA	intermediary metabolism and respiration	COG0460E
Rv1295	thrC	37321.9	Cytoplasmic	threonine synthase	Threonine synthase ThrC (ts)	intermediary metabolism and respiration	COG0498E
Rv1297	rho	65101.4	Cytoplasmic	transcription termination factor Rho	Probable transcription termination factor Rho homolog	information pathways	COG1158K
Rv1301	-	22552.7	Unknown	threonylcarbamoyl-AMP synthase	Putative translation factor (SUA5)	Information pathways	COG0009J
Rv1306	atpF	18324.8	Unknown	ATP synthase subunit B	Probable ATP synthase B chain AtpF	intermediary metabolism and respiration	COG0711C
Rv1307	atpH	48805.8	Cytoplasmic	ATP synthase subunit b/delta	Probable ATP synthase delta chain AtpH	intermediary metabolism and respiration	COG0712C
Rv1308	atpA	59288.3	Cytoplasmic	ATP synthase subunit alpha	Probable ATP synthase alpha chain AtpA	intermediary metabolism and respiration	COG0056C
Rv1309	atpG	33890.4	Cytoplasmic	ATP synthase subunit gamma	Probable ATP synthase gamma chain AtpG	intermediary metabolism and respiration	COG0224C
Rv1310	atpD	53094.2	Cytoplasmic Membrane	ATP synthase subunit beta	Probable ATP synthase beta chain AtpD	intermediary metabolism and respiration	COG0055C

Rv1314c	-	20694.4	Cytoplasmic	cob(I)yrinic acid a,c-diamide adenosyltransferase	Conserved protein	conserved hypotheticals	COG2096S
Rv1315	murA	44031.4	Cytoplasmic	UDP-N-acetylglucosamine 1-carboxyvinyltransferase	Probable UDP-N-acetylglucosamine 1-carboxyvinyltransferase MurA	cell wall and cell processes	COG0766M
Rv1323	fadA4	40048.5	Cytoplasmic	acetyl-CoA acetyltransferase	Probable acetyl-CoA acetyltransferase FadA4 (acetoacetyl-CoA thiolase)	lipid metabolism	COG0183I
Rv1324	-	32138.2	Cytoplasmic	thioredoxin	Possible thioredoxin	intermediary metabolism and respiration	COG3118O
Rv1327c	glgE	78607.9	Cytoplasmic	alpha-1,4-glucan:maltose-1-phosphate maltosyltransferase	Probable glucanase GlgE	intermediary metabolism and respiration	COG0366G
Rv1328	glgP	95483.3	Cytoplasmic	glycogen phosphorylase	Probable glycogen phosphorylase GlgP	intermediary metabolism and respiration	COG0058G
Rv1340	rphA	27319	Cytoplasmic	ribonuclease PH	Probable ribonuclease RphA (RNase PH) (tRNA nucleotidyltransferase)	information pathways	COG0689J
Rv1368	lprF	26851.6	Cytoplasmic Membrane	lipoprotein LprF	Probable conserved lipoprotein LprF	cell wall and cell processes	-
Rv1380	pyrB	33818.6	Cytoplasmic	aspartate carbamoyltransferase	Probable aspartate carbamoyltransferase PyrB (ATCase) (aspartate transcarbamylase)	intermediary metabolism and respiration	COG0540F
Rv1381	pyrC	45160.3	Cytoplasmic	dihydroorotase	Probable dihydroorotase PyrC (DHOase)	intermediary metabolism and respiration	COG0044F
Rv1384	carB	118929	Unknown	carbamoyl-phosphate synthase large subunit	Probable carbamoyl-phosphate synthase large chain CarB (carbamoyl-phosphate synthetase ammonia chain)	intermediary metabolism and respiration	COG0458EF
Rv1388	mihF	20817.1	Cytoplasmic	integration host factor MihF	Putative integration host factor MihF	information pathways	-
Rv1391	dfp	43544.9	Cytoplasmic	bifunctional phosphopantothenoylcysteine decarboxylase/phosphopantothenate--cysteine ligase	Probable DNA/pantothenate metabolism flavoprotein homolog Dfp	intermediary metabolism and respiration	COG0452H
Rv1392	metK	43014.6	Cytoplasmic	S-adenosylmethionine synthetase	Probable S-adenosylmethionine synthetase MetK (mat) (AdoMet synthetase) (methionine adenosyltransferase)	intermediary metabolism and respiration	COG0192H
Rv1409	ribG	35366.5	Cytoplasmic	bifunctional riboflavin biosynthesis diaminohydroxyphosphoribosylamino pyrimidine deaminase/5-amino-6-(5-phosphoribosylamino) uracil reductase	Probable bifunctional riboflavin biosynthesis protein RibG : diaminohydroxyphosphoribosylaminopyrimidine deaminase (riboflavin-specific deaminase) + 5-amino-6-(5-phosphoribosylamino) uracil reductase (HTP reductase)	intermediary metabolism and respiration	COG0117H
Rv1410c	-	54688.5	Cytoplasmic Membrane	aminoglycosides/tetracycline-transport integral membrane protein	Aminoglycosides/tetracycline-transport integral membrane protein	cell wall and cell processes	COG2814G
Rv1415	ribA2	46017.3	Cytoplasmic	bifunctional riboflavin biosynthesis GTP cyclohydrolase II/3,4-	Probable riboflavin biosynthesis protein RibA2 : GTP cyclohydrolase II + 3,4-dihydroxy-2-butanone 4-	intermediary metabolism and	COG0108H

				dihydroxy-2-butanone 4-phosphate synthase	phosphate synthase (DHBP synthase)	respiration	
Rv1423	whiA	34933	Cytoplasmic	transcriptional regulator WhiA	Probable transcriptional regulatory protein WhiA	Regulatory proteins	COG1481S
Rv1427c	fadD12	58064.1	Cytoplasmic	long-chain-fatty-acid--CoA ligase FadD12	Possible long-chain-fatty-acid--CoA ligase FadD12 (fatty-acid-CoA synthetase) (fatty-acid-CoA synthase)	lipid metabolism	COG0318IQ
Rv1431	-	65322.2	Cytoplasmic Membrane	hypothetical protein Rv1431	Conserved membrane protein	cell wall and cell processes	-
Rv1436	gap	35923.9	Cytoplasmic	glyceraldehyde 3-phosphate dehydrogenase	Probable glyceraldehyde 3-phosphate dehydrogenase Gap (GAPDH)	intermediary metabolism and respiration	COG0057G
Rv1437	pgk	42512.5	Cytoplasmic	phosphoglycerate kinase	Probable phosphoglycerate kinase Pgk	intermediary metabolism and respiration	COG0126G
Rv1438	tpi	27371	Cytoplasmic	triosephosphate isomerase	Probable triosephosphate isomerase Tpi (TIM)	intermediary metabolism and respiration	COG0149G
Rv1443c	-	18284.1	Cytoplasmic Membrane	hypothetical protein Rv1443c	Unknown protein	conserved hypotheticals	COG4276S
Rv1445c	devB	25772.2	Unknown	6-phosphogluconolactonase	Probable 6-phosphogluconolactonase DevB (6PGL)	intermediary metabolism and respiration	COG0363G
Rv1447c	zwf2	57343	Cytoplasmic	glucose-6-phosphate 1-dehydrogenase	Probable glucose-6-phosphate 1-dehydrogenase Zwf2 (G6PD)	intermediary metabolism and respiration	COG0364G
Rv1448c	tal	40720.7	Cytoplasmic	transaldolase	Probable transaldolase Tal	intermediary metabolism and respiration	COG0176G
Rv1449c	tkt	75556.7	Cytoplasmic	transketolase	Transketolase Tkt (TK)	intermediary metabolism and respiration	COG0021G
Rv1454c	qor	34046.5	Cytoplasmic	quinone reductase	Probable quinone reductase Qor (NADPH:quinone reductase) (zeta-crystallin homolog protein)	intermediary metabolism and respiration	COG0604CR
Rv1462	-	42522.6	Cytoplasmic	hypothetical protein Rv1462	ABC-type transport system involved in Fe-S cluster assembly, permease component	cell wall and cell processes	COG0719O
Rv1463	-	28822.7	Cytoplasmic	ABC transporter ATP-binding protein	Probable conserved ATP-binding protein ABC transporter	cell wall and cell processes	COG0396O
Rv1464	csd	44596.5	Cytoplasmic	cysteine desulfurase	Probable cysteine desulfurase Csd	intermediary metabolism and respiration	COG0520E
Rv1467c	fadE15	65905.9	Cytoplasmic	acyl-CoA dehydrogenase FadE15	Probable acyl-CoA dehydrogenase FadE15	lipid metabolism	COG1960I
Rv1475c	acn	102417	Cytoplasmic	iron-regulated aconitate hydratase	Probable iron-regulated aconitate hydratase Acn (citrate hydro-lyase) (aconitase)	intermediary metabolism and respiration	COG1048C

Rv1479	moxR1	40762.9	Cytoplasmic	transcriptional regulator MoxR1	Probable transcriptional regulatory protein MoxR1	Regulatory proteins	COG0714R
Rv1480	-	34301.3	Cytoplasmic	hypothetical protein Rv1480	Uncharacterized conserved protein (some members contain a von Willebrand factor type A (vWA) domain)	Regulatory proteins	COG1721R
Rv1483	fabG1	25665.2	Cytoplasmic	3-oxoacyl-ACP reductase FabG1	3-oxoacyl-[acyl-carrier protein] reductase FabG1 (3-ketoacyl-acyl carrier protein reductase) (mycolic acid biosynthesis a protein)	lipid metabolism	COG1028IQR
Rv1484	inhA	28527.8	Cytoplasmic Membrane	NADH-dependent enoyl-[ACP] reductase	NADH-dependent enoyl-[acyl-carrier-protein] reductase InhA (NADH-dependent enoyl-ACP reductase)	lipid metabolism	COG0623I
Rv1488	-	41250.2	Cytoplasmic Membrane	hypothetical protein Rv1488	Possible exported conserved protein	cell wall and cell processes	COG0330O
Rv1489	-	12305.7	Cytoplasmic Membrane	hypothetical protein Rv1489	Conserved protein	conserved hypotheticals	-
Rv1507c	-	27072.6	Unknown	hypothetical protein Rv1507c	cell wall biosynthesis	cell wall and cell processes	-
Rv1511	gmdA	38310.4	Cytoplasmic	GDP-D-mannose dehydratase GmdA	GDP-D-mannose dehydratase GmdA (GDP-mannose 4,6 dehydratase) (GMD)	intermediary metabolism and respiration	COG1089M
Rv1513	-	26684.3	Cytoplasmic	hypothetical protein Rv1513	SAM-dependent methyltransferases	Intermediary metabolism and respiration	-
Rv1521	fadD25	63141.6	Cytoplasmic	fatty-acid--CoA ligase FadD25	Probable fatty-acid-AMP ligase FadD25 (fatty-acid-AMP synthetase) (fatty-acid-AMP synthase)	lipid metabolism	COG0318IQ
Rv1527c	pks5	223857	Unknown	polyketide synthase	Probable polyketide synthase Pks5	lipid metabolism	COG3321Q
Rv1536	ileS	117308	Cytoplasmic	isoleucine--tRNA ligase	Isoleucyl-tRNA synthetase IleS	information pathways	COG0060J
Rv1543	-	36820.9	Cytoplasmic	oxidoreductase	Possible fatty acyl-CoA reductase	lipid metabolism	COG0300R
Rv1544	-	28199.6	Cytoplasmic	ketoacyl reductase	Possible ketoacyl reductase	lipid metabolism	COG0300R
Rv1559	ilvA	45008.9	Cytoplasmic	threonine dehydratase IlvA	Probable threonine dehydratase IlvA	intermediary metabolism and respiration	COG1171E
Rv1593c	-	25907.8	Cytoplasmic	hypothetical protein Rv1593c	ADP-ribose pyrophosphatase	Intermediary metabolism and respiration	COG1051F
Rv1594	nadA	37408.4	Cytoplasmic	quinolinate synthetase A	Probable quinolinate synthetase NadA	intermediary metabolism and respiration	COG0379H
Rv1596	nadC	29951.1	Cytoplasmic	nicotinate-nucleotide pyrophosphatase	Probable nicotinate-nucleotide pyrophosphatase NadC	intermediary metabolism and respiration	COG0157H
Rv1609	trpE	55816.9	Cytoplasmic	anthranilate synthase component I	Anthranilate synthase component I TrpE (glutamine amidotransferase)	intermediary metabolism and respiration	COG0147EH
Rv1611	trpC	28023.1	Cytoplasmic	indole-3-glycerol phosphate synthase	Probable indole-3-glycerol phosphate synthase TrpC	intermediary metabolism and respiration	COG0134E

Rv1612	trpB	43286.7	Cytoplasmic	tryptophan synthase subunit beta	Tryptophan synthase, beta subunit TrpB	intermediary metabolism and respiration	COG0133E
Rv1613	trpA	27727.7	Unknown	tryptophan synthase subunit alpha	Probable tryptophan synthase, alpha subunit TrpA	intermediary metabolism and respiration	COG0159E
Rv1614	lgt	50392.2	Cytoplasmic Membrane	prolipoprotein diacylglyceryl transferase	Possible prolipoprotein diacylglyceryl transferases Lgt	cell wall and cell processes	COG0682M
Rv1617	pykA	50667.1	Cytoplasmic	pyruvate kinase	Probable pyruvate kinase PykA	intermediary metabolism and respiration	COG0469G
Rv1621c	cydD	54766.1	Cytoplasmic Membrane	cytochrome biosynthesis ABC transporter ATP-binding protein/permease CydD	Probable 'component linked with the assembly of cytochrome' transport transmembrane ATP-binding protein ABC transporter CydD	intermediary metabolism and respiration	COG4988CO
Rv1623c	cydA	53826.5	Cytoplasmic Membrane	cytochrome D ubiquinol oxidase subunit I CydA	Probable integral membrane cytochrome D ubiquinol oxidase (subunit I) CydA (cytochrome BD-I oxidase subunit I)	intermediary metabolism and respiration	COG1271C
Rv1626	-	22669.2	Cytoplasmic	two-component system transcriptional regulator	Probable two-component system transcriptional regulator	Regulatory proteins	COG3707T
Rv1627c	-	42386	Cytoplasmic	nonspecific lipid-transfer protein	Probable nonspecific lipid-transfer protein	lipid metabolism	COG0183I
Rv1629	polA	98440	Cytoplasmic	DNA polymerase I	Probable DNA polymerase I PolA	information pathways	COG0749L
Rv1630	rpsA	53201.8	Cytoplasmic	30S ribosomal protein S1	30S ribosomal protein S1 RpsA	information pathways	COG0539J
Rv1631	coaE	44668.6	Cytoplasmic	dephospho-CoA kinase CoaE	Probable dephospho-CoA kinase CoaE (dephosphocoenzyme a kinase)	intermediary metabolism and respiration	COG2320S
Rv1636	TB15.3	15312.4	Unknown	iron-regulated universal stress protein	Iron-regulated universal stress protein family protein TB15.3	virulence, detoxification, adaptation	COG0589T
Rv1638	uvrA	106099	Cytoplasmic	excinuclease ABC subunit UvrA	Probable excinuclease ABC (subunit A - DNA-binding ATPase) UvrA	information pathways	COG0178L
Rv1650	pheT	88374	Cytoplasmic	phenylalanine--tRNA ligase subunit beta	Probable phenylalanyl-tRNA synthetase, beta chain PheT	information pathways	COG0072J
Rv1652	argC	36304.3	Unknown	N-acetyl-gamma-glutamyl-phosphate reductase	Probable N-acetyl-gamma-glutamyl-phosphate reductase ArgC	intermediary metabolism and respiration	COG0002E
Rv1654	argB	30936.8	Cytoplasmic	acetylglutamate kinase	Probable acetylglutamate kinase ArgB	intermediary metabolism and respiration	COG0548E
Rv1655	argD	40909.9	Cytoplasmic	acetylornithine aminotransferase	Probable acetylornithine aminotransferase ArgD	intermediary metabolism and respiration	COG4992E
Rv1658	argG	43682.4	Cytoplasmic	argininosuccinate synthase	Probable argininosuccinate synthase ArgG	intermediary metabolism and respiration	COG0137E
Rv1659	argH	49742.2	Unknown	argininosuccinate lyase	Probable argininosuccinate lyase ArgH	intermediary	COG0165E

						metabolism and respiration	
Rv1679	fadE16	37876.2	Cytoplasmic	acyl-CoA dehydrogenase FadE16	Possible acyl-CoA dehydrogenase FadE16	lipid metabolism	COG1960I
Rv1680	-	30214	Cytoplasmic	hypothetical protein Rv1680	ABC-type phosphate/phosphonate transport system, periplasmic component	cell wall and cell processes	COG3221P
Rv1683	-	107362	Cytoplasmic	bifunctional long-chain acyl-CoA synthase/lipase	Possible bifunctional enzyme; long-chain acyl-CoA synthase and lipase.	lipid metabolism	COG3243I
Rv1689	tyrS	46330.4	Cytoplasmic	tyrosine--tRNA ligase	Probable tyrosyl-tRNA synthase TyrS (TYRRS)	information pathways	COG0162J
Rv1699	pyrG	63603	Cytoplasmic	CTP synthase	Probable CTP synthase PyrG	intermediary metabolism and respiration	COG0504F
Rv1708	-	34362.4	Cytoplasmic Membrane	initiation inhibitor protein	Putative initiation inhibitor protein	cell wall and cell processes	COG1192D
Rv1713	engA	49925.7	Cytoplasmic Membrane	GTPase Der	Probable GTP-binding protein EngA	intermediary metabolism and respiration	COG1160R
Rv1722	-	55428.8	Cytoplasmic	carboxylase	Possible carboxylase	lipid metabolism	COG0439I
Rv1731	gabD2	55324.5	Cytoplasmic	succinate-semialdehyde dehydrogenase	Possible succinate-semialdehyde dehydrogenase [NADP+] dependent (SSDH) GabD2	intermediary metabolism and respiration	COG1012C
Rv1751	-	50412.1	Cytoplasmic	oxidoreductase	Probable oxidoreductase	intermediary metabolism and respiration	COG0654HC
Rv1754c	-	60608.4	Cytoplasmic Membrane	hypothetical protein Rv1754c	Conserved protein	conserved hypotheticals	COG5373S
Rv1769	-	45380.9	Cytoplasmic	hypothetical protein Rv1769	Predicted amino acid aldolase or racemase	Intermediary metabolism and respiration	COG3616E
Rv1770	-	45972.6	Cytoplasmic	hypothetical protein Rv1770	Predicted aminopeptidases	Intermediary metabolism and respiration	COG2234R
Rv1771	-	48045.4	Cytoplasmic	L-gulono-1,4-lactone dehydrogenase	L-gulono-1,4-lactone dehydrogenase	intermediary metabolism and respiration	COG0277C
Rv1782	eccB5	53689.2	Unknown	ESX-5 type VII secretion system protein EccB5	ESX conserved component EccB5. ESX-5 type VII secretion system protein. Probable membrane protein.	cell wall and cell processes	-
Rv1783	eccC5	152740	Cytoplasmic Membrane	ESX-5 type VII secretion system protein EccC5	ESX conserved component EccC5. ESX-5 type VII secretion system protein.	cell wall and cell processes	-
Rv1794	-	32399.6	Cytoplasmic	hypothetical protein Rv1794	cell wall processes	cell wall and cell processes	-
Rv1798	eccA5	67756.1	Cytoplasmic	ESX-5 type VII secretion system protein EccA	ESX conserved component EccA5. ESX-5 type VII secretion system protein.	cell wall and cell processes	COG1223R
Rv1808	PPE32	39899.3	Cytoplasmic Membrane	PPE family protein PPE32	PPE family protein PPE32	PE/PPE	COG5651N

Rv1820	ilvG	57520.5	Cytoplasmic	acetolactate synthase large subunit IlvG	Probable acetolactate synthase IlvG (acetohydroxy-acid synthase)(ALS)	intermediary metabolism and respiration	COG0028EH
Rv1821	secA2	88919.7	Cytoplasmic	protein translocase subunit SecA	Possible preprotein translocase ATPase SecA2	cell wall and cell processes	COG0653U
Rv1827	garA	17218.8	Cytoplasmic	glycogen accumulation regulator GarA	FOG- FHA domain	Regulatory proteins	COG1716T
Rv1832	gcvB	99478.8	Cytoplasmic	glycine dehydrogenase	Probable glycine dehydrogenase GcvB (glycine decarboxylase) (glycine cleavage system P-protein)	intermediary metabolism and respiration	COG1003E
Rv1833c	-	32151.1	Cytoplasmic	haloalkane dehalogenase	Possible haloalkane dehalogenase	intermediary metabolism and respiration	COG0596R
Rv1836c	-	69676.8	Cytoplasmic Membrane	hypothetical protein Rv1836c	Von Willebrand factor (vWF) type A domain-containing protein	Regulatory proteins	-
Rv1837c	glcB	80403	Cytoplasmic	malate synthase	Malate synthase G GlcB	intermediary metabolism and respiration	COG2225C
Rv1843c	guaB1	49932.8	Cytoplasmic	inosine-5'-monophosphate dehydrogenase	Probable inosine-5'-monophosphate dehydrogenase GuaB1(imp dehydrogenase) (IMPDH) (IMPD)	intermediary metabolism and respiration	COG0516F
Rv1852	ureG	23348.4	Unknown	urease accessory protein UreG	Urease accessory protein UreG	intermediary metabolism and respiration	COG0378OK
Rv1854c	ndh	49619	Cytoplasmic Membrane	NADH dehydrogenase	Probable NADH dehydrogenase Ndh	intermediary metabolism and respiration	COG1252C
Rv1855c	-	33188.2	Cytoplasmic	oxidoreductase	Possible oxidoreductase	intermediary metabolism and respiration	COG2141C
Rv1856c	-	23228.1	Cytoplasmic	oxidoreductase	Possible oxidoreductase	intermediary metabolism and respiration	COG1028IQR
Rv1864c	-	27207.8	Cytoplasmic	hypothetical protein Rv1864c	Conserved protein	conserved hypotheticals	COG2258S
Rv1869c	-	43629.3	Cytoplasmic	reductase	Probable reductase	intermediary metabolism and respiration	COG1251C
Rv1871c	-	14662.9	Cytoplasmic	hypothetical protein Rv1871c	DNA-repair	Information pathways	-
Rv1872c	lldD2	45341.6	Cytoplasmic Membrane	L-lactate dehydrogenase	Possible L-lactate dehydrogenase (cytochrome) LldD2	intermediary metabolism and respiration	COG1304C
Rv1876	bfrA	18340.7	Cytoplasmic	bacterioferritin BfrA	Probable bacterioferritin BfrA	intermediary metabolism and respiration	COG2193P

Rv1880c	cyp140	48839.9	Cytoplasmic Membrane	cytochrome P450 Cyp140	Probable cytochrome P450 140 Cyp140	intermediary metabolism and respiration	-
Rv1886c	fbpB	34580.9	Extracellular	diacylglycerol acyltransferase/mycolyltransferase Ag85B	Secreted antigen 85-B FbpB (85B) (antigen 85 complex B) (mycolyl transferase 85B) (fibronectin-binding protein B) (extracellular alpha-antigen)	lipid metabolism	COG0627R
Rv1899c	lppD	35071.8	Cytoplasmic	lipoprotein LppD	Possible lipoprotein LppD	cell wall and cell processes	COG2110R
Rv1901	cinA	45512	Cytoplasmic	protein CinA	Probable CinA-like protein CinA	virulence, detoxification, adaptation	COG1058R
Rv1908c	katG	80572.8	Cytoplasmic	catalase-peroxidase	Catalase-peroxidase-peroxynitritase T KatG	virulence, detoxification, adaptation	COG0376P
Rv1925	fadD31	66310.1	Cytoplasmic	fatty-acid--CoA ligase FadD31	Probable acyl-CoA ligase FadD31 (acyl-CoA synthetase) (acyl-CoA synthase)	lipid metabolism	-
Rv1932	tpx	16896.1	Unknown	thiol peroxidase	Probable thiol peroxidase Tpx	virulence, detoxification, adaptation	COG2077O
Rv1996	-	33879.7	Cytoplasmic Membrane	universal stress protein	Universal stress protein family protein	virulence, detoxification, adaptation	COG0589T
Rv2004c	-	54422.5	Cytoplasmic	hypothetical protein Rv2004c	Predicted kinase	Regulatory proteins	COG2187S
Rv2005c	-	30985.4	Cytoplasmic	universal stress protein	Universal stress protein family protein	virulence, detoxification, adaptation	COG0589T
Rv2006	otsB1	145752	Cytoplasmic	trehalose-6-phosphate phosphatase OtsB	Probable trehalose-6-phosphate phosphatase OtsB1 (trehalose-phosphatase) (TPP)	virulence, detoxification, adaptation	COG1554G
Rv2030c	-	74898.2	Cytoplasmic	hypothetical protein Rv2030c	Erythromycin esterase homolog	Intermediary metabolism and respiration	COG2312R
Rv2031c	hspX	16227.3	Cellwall	alpha-crystallin	Heat shock protein HspX (alpha-crystallin homolog) (14 kDa antigen) (HSP16.3)	virulence, detoxification, adaptation	COG0071O
Rv2032	acg	36559.3	Cytoplasmic	NAD(P)H nitroreductase acg	nitroreductase	Lipid metabolism	-
Rv2048c	pks12	431607	Unknown	polyketide synthase	Polyketide synthase Pks12	lipid metabolism	-
Rv2051c	ppm1	93792.4	Cytoplasmic Membrane	polyprenol-monophosphomannose synthase	Polyprenol-monophosphomannose synthase Ppm1	cell wall and cell processes	COG0815M
Rv2054	-	25183.1	Cytoplasmic Membrane	hypothetical protein Rv2054	Dienelactone hydrolase and related enzymes	Intermediary metabolism and respiration	COG0412Q
Rv2061c	-	14782	Unknown	hypothetical protein Rv2061c	Predicted flavin-nucleotide-binding protein structurally related to pyridoxine 5-phosphate oxidase	Information pathways	COG3576R
Rv2066	cobIJ	53911	Unknown	bifunctional S-adenosyl-L-	Probable bifunctional protein, CobI-COBj fusion protein:	intermediary	COG1010H

				methionine-precorrin-2 methyl transferase/precorrin-3 methylase	S-adenosyl-L-methionine-precorrin-2 methyl transferase + precorrin-3 methylase	metabolism and respiration	
Rv2067c	-	45898.4	Cytoplasmic	hypothetical protein Rv2067c	SAM-dependent methyltransferases	Intermediary metabolism and respiration	COG2227H
Rv2091c	-	26019.8	Cytoplasmic Membrane	membrane protein	Probable membrane protein	cell wall and cell processes	-
Rv2097c	pafA	51352	Cytoplasmic	proteasome accessory factor PafA	Proteasome accessory factor a PafA	intermediary metabolism and respiration	-
Rv2101	helZ	111612	Cytoplasmic	helicase HelZ	Probable helicase HelZ	information pathways	-
Rv2109c	prcA	26849	Cytoplasmic Membrane	proteasome subunit alpha	Proteasome alpha subunit PrcA; assembles with beta subunit PrcB.	intermediary metabolism and respiration	COG0638O
Rv2110c	prcB	30273	Unknown	proteasome subunit beta	Proteasome beta subunit PrcB; assembles with alpha subunit PrcA.	intermediary metabolism and respiration	COG0638O
Rv2111c	pup	6944.31	Cytoplasmic	ubiquitin-like protein Pup	Prokaryotic ubiquitin-like protein Pup	intermediary metabolism and respiration	-
Rv2112c	dop	60129.2	Cytoplasmic	pup deamidase/depupylase	Deamidase of pup Dop	intermediary metabolism and respiration	-
Rv2115c	mpa	67401.1	Cytoplasmic Membrane	proteasome-associated ATPase	Mycobacterial proteasome ATPase Mpa	cell wall and cell processes	COG0464O
Rv2118c	-	30090.4	Cytoplasmic	tRNA (adenine(58)-N(1))-methyltransferase	RNA methyltransferase	intermediary metabolism and respiration	COG2519J
Rv2121c	hisG	30480.8	Cytoplasmic	ATP phosphoribosyltransferase	ATP phosphoribosyltransferase HisG	intermediary metabolism and respiration	COG0040E
Rv2129c	-	31331.3	Cytoplasmic	oxidoreductase	Probable oxidoreductase	intermediary metabolism and respiration	COG4221R
Rv2131c	cysQ	28415.1	Cytoplasmic	3'-phosphoadenosine 5'-phosphate phosphatase	Monophosphatase CysQ	intermediary metabolism and respiration	COG1218P
Rv2140c	TB18.6	18634	Cytoplasmic	hypothetical protein Rv2140c	Phospholipid-binding protein	Intermediary metabolism and respiration	COG1881R
Rv2145c	wag31	28277.2	Cytoplasmic	cell wall synthesis protein Wag31	Diviva family protein Wag31	cell wall and cell processes	COG3599D
Rv2150c	ftsZ	38755.9	Cytoplasmic	cell division protein FtsZ	Cell division protein FtsZ	cell wall and cell processes	COG0206D
Rv2155c	murD	49313.4	Cytoplasmic	UDP-N-acetylmuramoylalanine--D-	Probable UDP-N-acetylmuramoylalanine-D-glutamate	cell wall and cell	-

				glutamate ligase	ligase MurD	processes	
Rv2157c	murF	51632.6	Cytoplasmic	UDP-N-acetylmuramoyl-tripeptide--D-alanyl-D-alanine ligase	Probable UDP-N-acetylmuramoylalanyl-D-glutamyl-2,6-diaminopimelate-D-alanyl-D-alanyl ligase MurF	cell wall and cell processes	COG0770M
Rv2158c	murE	55308.7	Cytoplasmic	UDP-N-acetylmuramoylalanyl-D-glutamate--2,6-diaminopimelate ligase	Probable UDP-N-acetylmuramoylalanyl-D-glutamate-2,6-diaminopimelate ligase MurE	cell wall and cell processes	COG0769M
Rv2159c	-	36376.5	Cytoplasmic Membrane	hypothetical protein Rv2159c	Peroxidase	Intermediary metabolism and respiration	COG0599S
Rv2161c	-	30768.3	Cytoplasmic	hypothetical protein Rv2161c	Conserved protein	intermediary metabolism and respiration	COG2141C
Rv2172c	-	33006.9	Cytoplasmic	hypothetical protein Rv2172c	membrane associated process	cell wall and cell processes	-
Rv2178c	aroG	50641.4	Cytoplasmic	phospho-2-dehydro-3-deoxyheptonate aldolase AroG	3-deoxy-D-arabino-heptulosonate 7-phosphate synthase AroG (DAHP synthetase, phenylalanine-repressible)	intermediary metabolism and respiration	COG3200E
Rv2185c	TB16.3	16292.7	Cytoplasmic	hypothetical protein Rv2185c	Cyclase/Dehydrase	Intermediary metabolism and respiration	COG2867I
Rv2187	fadD15	64002.9	Cytoplasmic	long-chain-fatty-acid--CoA ligase FadD15	Long-chain-fatty-acid-CoA ligase FadD15 (fatty-acid-CoA synthetase) (fatty-acid-CoA synthase)	lipid metabolism	COG1022I
Rv2192c	trpD	37707.8	Cytoplasmic Membrane	anthranilate phosphoribosyltransferase	Probable anthranilate phosphoribosyltransferase TrpD	intermediary metabolism and respiration	COG0547E
Rv2195	qcrA	46922.6	Cytoplasmic Membrane	ubiquinol-cytochrome C reductase rieske iron-sulfur subunit	Probable rieske iron-sulfur protein QcrA	intermediary metabolism and respiration	COG0723C
Rv2200c	ctaC	40447.6	Cytoplasmic Membrane	cytochrome C oxidase subunit II	Probable transmembrane cytochrome C oxidase (subunit II) CtaC	intermediary metabolism and respiration	COG1622C
Rv2202c	adoK	34440	Unknown	adenosine kinase	Adenosine kinase	intermediary metabolism and respiration	COG0524G
Rv2204c	-	12543.9	Cytoplasmic	hypothetical protein Rv2204c	Conserved protein	conserved hypotheticals	COG0316S
Rv2210c	ilvE	39666.1	Cytoplasmic	branched-chain amino acid aminotransferase	Branched-chain amino acid transaminase IlvE	intermediary metabolism and respiration	COG0115EH
Rv2211c	gcvT	39597	Cytoplasmic	aminomethyltransferase	Probable aminomethyltransferase GcvT (glycine cleavage system T protein)	intermediary metabolism and respiration	COG0404E
Rv2212	-	39743	Cytoplasmic	adenylyl cyclase	Adenylyl cyclase (ATP pyrophosphate-lyase) (adenylate cyclase)	intermediary metabolism and respiration	COG2114T

Rv2213	pepB	53449.3	Cytoplasmic	cytosol aminopeptidase	Probable aminopeptidase PepB	intermediary metabolism and respiration	COG0260E
Rv2214c	ephD	64029.8	Cytoplasmic	oxidoreductase EphD	Possible short-chain dehydrogenase EphD	virulence, detoxification, adaptation	COG0300R
Rv2215	dlaT	57087.7	Cytoplasmic	pyruvate dehydrogenase E2 component dihydrolipoamide acyltransferase	DlaT, dihydrolipoamide acyltransferase, E2 component of pyruvate dehydrogenase	intermediary metabolism and respiration	COG0508C
Rv2216	-	31671.9	Cytoplasmic Membrane	epimerase family protein	Predicted nucleoside-diphosphate sugar epimerase	Intermediary metabolism and respiration	COG1090R
Rv2220	glnA1	53537.9	Cytoplasmic	glutamine synthetase	Glutamine synthetase GlnA1 (glutamine synthase) (GS-I)	intermediary metabolism and respiration	COG0174E
Rv2222c	glnA2	49607.8	Cytoplasmic	glutamine synthetase	Probable glutamine synthetase GlnA2 (glutamine synthase) (GS-II)	intermediary metabolism and respiration	COG0174E
Rv2224c	caeA	55892.4	Cytoplasmic	carboxylesterase A	Probable carboxylesterase CaeA	cell wall and cell processes	COG0596R
Rv2226	-	56333.2	Cytoplasmic	hypothetical protein Rv2226	Protein involved in cAMP biosynthetic process	Intermediary metabolism and respiration	COG5607S
Rv2228c	-	39113.1	Cellwall	multifunctional RNASE H/alpha-ribazole phosphatase/acid phosphatase	Multifunctional protein. Has RNASE H, alpha-ribazole phosphatase, and acid phosphatase activities.	information pathways	COG0406G
Rv2241	aceE	100185	Cytoplasmic	pyruvate dehydrogenase E1 component	Pyruvate dehydrogenase E1 component AceE (pyruvate decarboxylase) (pyruvate dehydrogenase) (pyruvic dehydrogenase)	intermediary metabolism and respiration	-
Rv2243	fabD	30756.1	Cytoplasmic	malonyl CoA-acyl carrier protein transacylase	Malonyl CoA-acyl carrier protein transacylase FabD (malonyl CoA:ACPM acyltransferase) (MCT)	lipid metabolism	COG0331I
Rv2244	acpM	12491.9	Cytoplasmic	meromycolate extension acyl carrier protein	Meromycolate extension acyl carrier protein AcpM	lipid metabolism	COG0236IQ
Rv2245	kasA	43284.1	Cytoplasmic Membrane	3-oxoacyl-ACP synthase 1	3-oxoacyl-[acyl-carrier protein] synthase 1 KasA (beta-ketoacyl-ACP synthase) (KAS I)	lipid metabolism	COG0304IQ
Rv2246	kasB	46388.7	Cytoplasmic	3-oxoacyl-ACP synthase 2	3-oxoacyl-[acyl-carrier protein] synthase 2 KasB (beta-ketoacyl-ACP synthase) (KAS I)	lipid metabolism	COG0304IQ
Rv2247	accD6	50136	Cytoplasmic	acetyl-/propionyl-CoA carboxylase subunit beta	Acetyl/propionyl-CoA carboxylase (beta subunit) AccD6	lipid metabolism	COG4799I
Rv2251	-	49782.5	Cytoplasmic	flavoprotein	Possible flavoprotein	intermediary metabolism and respiration	COG0277C
Rv2258c	-	37513.6	Cytoplasmic	transcriptional regulator	Possible transcriptional regulatory protein	Regulatory proteins	COG2226H
Rv2259	mscR	37888.2	Cytoplasmic	S-nitrosomycothioliol reductase MscR	S-nitrosomycothioliol reductase MscR	intermediary metabolism and	COG1062C

						respiration	
Rv2289	cdh	28576.4	Cytoplasmic Membrane	CDP-diacylglycerol pyrophosphatase	Probable CDP-diacylglycerol pyrophosphatase Cdh (CDP-diacylglycerol diphosphatase) (CDP-diacylglycerol phosphatidylhydrolase)	lipid metabolism	COG2134I
Rv2291	sseB	29367.9	Cytoplasmic	thiosulfate sulfurtransferase SseB	Probable thiosulfate sulfurtransferase SseB	intermediary metabolism and respiration	COG2897P
Rv2294	-	44251.1	Cytoplasmic	cystathionine beta-lyase	Probable aminotransferase	intermediary metabolism and respiration	COG1168E
Rv2296	-	33357.9	Cytoplasmic	haloalkane dehalogenase	Probable haloalkane dehalogenase	intermediary metabolism and respiration	COG0596R
Rv2298	-	34986.1	Cytoplasmic	oxidoreductase	Predicted oxidoreductases (related to aryl-alcohol dehydrogenases)	Intermediary metabolism and respiration	COG0667C
Rv2299c	htpG	72961.1	Cytoplasmic	chaperone protein HtpG	Probable chaperone protein HtpG (heat shock protein) (HSP90 family protein) (high temperature protein G)	virulence, detoxification, adaptation	COG0326O
Rv2305	-	46124	Cytoplasmic	hypothetical protein Rv2305	Uncharacterized protein containing a NRPS condensation (elongation) domain	cell wall and cell processes	COG4908R
Rv2314c	-	48712	Cytoplasmic Membrane	hypothetical protein Rv2314c	Predicted Zn-dependent proteases and their inactivated homologs	Intermediary metabolism and respiration	COG0312R
Rv2315c	-	54546.1	Unknown	hypothetical protein Rv2315c	Predicted Zn-dependent proteases and their inactivated homologs	Intermediary metabolism and respiration	COG0312R
Rv2345	-	70029.7	Cytoplasmic Membrane	transmembrane protein	Possible conserved transmembrane protein	cell wall and cell processes	COG1512R
Rv2350c	plcB	55623.3	Cytoplasmic Membrane	membrane-associated phospholipase B	Membrane-associated phospholipase C 2 PlcB	intermediary metabolism and respiration	COG3511M
Rv2357c	glyS	52905.7	Cytoplasmic	glycine--tRNA ligase	Probable glycyl-tRNA synthetase GlyS (glycine--tRNA ligase) (GLYRS)	information pathways	COG0423J
Rv2363	amiA2	50851.7	Cytoplasmic	amidase	Probable amidase AmiA2 (aminohydrolase)	intermediary metabolism and respiration	COG0154J
Rv2373c	dnaJ2	40456.8	Cytoplasmic	chaperone protein DnaJ	Probable chaperone protein DnaJ2	virulence, detoxification, adaptation	COG0484O
Rv2391	sirA	62997.5	Cytoplasmic	sulfite reductase	Ferredoxin-dependent sulfite reductase SirA	intermediary metabolism and respiration	COG0155P
Rv2402	-	72734	Cytoplasmic Membrane	trehalase	Glucoamylase and related glycosyl hydrolases	Intermediary metabolism and	-

						respiration	
Rv2410c	-	36020.6	Cytoplasmic Membrane	hypothetical protein Rv2410c	proteasome	Intermediary metabolism and respiration	COG2307S
Rv2411c	-	61383.4	Cytoplasmic	hypothetical protein Rv2411c	proteasome	Intermediary metabolism and respiration	COG2308S
Rv2416c	eis	43771.8	Unknown	enhanced intracellular survival protein	Enhanced intracellular survival protein Eis, GCN5-related N-acetyltransferase	virulence, detoxification, adaptation	COG4552R
Rv2427c	proA	43744.7	Cytoplasmic	gamma-glutamyl phosphate reductase	Probable gamma-glutamyl phosphate reductase protein ProA (GPR) (glutamate-5-semialdehyde dehydrogenase) (glutamyl-gamma-semialdehyde dehydrogenase)	intermediary metabolism and respiration	COG0014E
Rv2428	ahpC	21566.3	Cytoplasmic	alkyl hydroperoxide reductase subunit C	Alkyl hydroperoxide reductase C protein AhpC (alkyl hydroperoxidase C)	virulence, detoxification, adaptation	COG04500
Rv2438c	nadE	74682.5	Unknown	glutamine-dependent NAD(+) synthetase	Glutamine-dependent NAD(+) synthetase NadE (NAD(+) synthase [glutamine-hydrolysing])	intermediary metabolism and respiration	COG0171H
Rv2439c	proB	38788.3	Unknown	glutamate 5-kinase protein	Probable glutamate 5-kinase protein ProB (gamma-glutamyl kinase) (GK)	intermediary metabolism and respiration	COG0263E
Rv2442c	rplU	11151.1	Unknown	50S ribosomal protein L21	50S ribosomal protein L21 RplU	information pathways	COG0261J
Rv2444c	rne	103358	Cytoplasmic	ribonuclease E	Possible ribonuclease E Rne	information pathways	COG1530J
Rv2445c	ndkA	14475.5	Cytoplasmic	nucleoside diphosphate kinase	Probable nucleoside diphosphate kinase NdkA (NDK) (NDP kinase) (nucleoside-2-P kinase)	intermediary metabolism and respiration	COG0105F
Rv2448c	valS	97820.5	Cytoplasmic	valine--tRNA ligase	Probable valyl-tRNA synthase protein ValS (valyl-tRNA synthetase) (valine--tRNA ligase) (valine translase)	information pathways	COG0525J
Rv2449c	-	44346.5	Cytoplasmic	trans-acting enoyl reductase	Conserved protein	conserved hypotheticals	COG3268S
Rv2454c	-	40143.2	Cytoplasmic	2-oxoglutarate oxidoreductase subunit KorB	Probable oxidoreductase (beta subunit)	intermediary metabolism and respiration	COG1013C
Rv2455c	-	69150.4	Cytoplasmic	2-oxoglutarate oxidoreductase subunit KorA	Probable oxidoreductase (alpha subunit)	intermediary metabolism and respiration	COG0674C
Rv2457c	clpX	46782.7	Cytoplasmic	ATP-dependent CLP protease ATP-binding subunit ClpX	Probable ATP-dependent CLP protease ATP-binding subunit ClpX	intermediary metabolism and respiration	COG1219O
Rv2460c	clpP2	23507.7	Cytoplasmic	ATP-dependent CLP protease proteolytic subunit 2	Probable ATP-dependent CLP protease proteolytic subunit 2 ClpP2 (endopeptidase CLP 2)	intermediary metabolism and respiration	COG0740OU
Rv2461c	clpP1	21675.6	Cytoplasmic	ATP-dependent CLP protease proteolytic subunit 1	Probable ATP-dependent CLP protease proteolytic subunit 1 ClpP1 (endopeptidase CLP)	intermediary metabolism and	COG0740OU

						respiration	
Rv2462c	tig	50584	Cytoplasmic	trigger factor	Probable trigger factor (TF) protein Tig	cell wall and cell processes	COG0544O
Rv2465c	rpiB	17277.5	Cytoplasmic	ribose-5-phosphate isomerase B	Ribose-5-phosphate isomerase	intermediary metabolism and respiration	COG0698G
Rv2467	pepN	94252.7	Cytoplasmic	aminopeptidase PepN	Probable aminopeptidase N PepN (Lysyl aminopeptidase) (LYS-AP) (alanine aminopeptidase)	intermediary metabolism and respiration	COG0308E
Rv2476c	gdh	176900	Cytoplasmic Membrane	NAD-dependent glutamate dehydrogenase	Probable NAD-dependent glutamate dehydrogenase Gdh (NAD-Gdh) (NAD-dependent glutamic dehydrogenase)	intermediary metabolism and respiration	COG2902E
Rv2477c	-	61893.1	Cytoplasmic	macrolide ABC transporter ATP-binding protein	Probable macrolide-transport ATP-binding protein ABC transporter	cell wall and cell processes	COG0488R
Rv2482c	plsB2	88283	Cytoplasmic Membrane	glycerol-3-phosphate acyltransferase	Probable glycerol-3-phosphate acyltransferase PlsB2 (GPAT)	lipid metabolism	COG2937I
Rv2483c	plsC	61788.1	Cytoplasmic Membrane	bifunctional L-3-phosphoserine phosphatase/1-acyl-sn-glycerol-3-phosphate acyltransferase	Possible transmembrane phospholipid biosynthesis bifunctional enzyme PlsC: putative L-3-phosphoserine phosphatase (O-phosphoserine phosphohydrolase) (PSP) (pspase) + 1-acyl-SN-glycerol-3-phosphate acyltransferase (1-AGP acyltransferase) (1-AGPAT) (lysophosphatidic acid acyltransferase) (LPAAT)	lipid metabolism	COG0560E
Rv2495c	bkdC	41061	Cytoplasmic	branched-chain keto acid dehydrogenase E2 component	Probable branched-chain keto acid dehydrogenase E2 component BkdC	intermediary metabolism and respiration	COG0508C
Rv2496c	bkdB	38064.3	Cytoplasmic	branched-chain keto acid dehydrogenase E1 component subunit beta	Probable branched-chain keto acid dehydrogenase E1 component, beta subunit BkdB	intermediary metabolism and respiration	COG0022C
Rv2500c	fadE1 ₉	41916.4	Cytoplasmic	acyl-CoA dehydrogenase FadE19	Possible acyl-CoA dehydrogenase FadE19 (MMGC)	lipid metabolism	COG1960I
Rv2501c	accA1	70592.2	Cytoplasmic	acetyl-/propionyl-CoA carboxylase subunit alpha	Probable acetyl-/propionyl-coenzyme A carboxylase alpha chain (alpha subunit) AccA1: biotin carboxylase + biotin carboxyl carrier protein (BCCP)	lipid metabolism	COG4770I
Rv2509	-	28013.2	Cytoplasmic	short-chain type dehydrogenase/reductase	Probable short-chain type dehydrogenase/reductase	intermediary metabolism and respiration	COG0300R
Rv2510c	-	56546.7	Unknown	hypothetical protein Rv2510c	Predicted ATPase	Intermediary metabolism and respiration	COG0433R
Rv2524c	fas	326222	Cytoplasmic Membrane	fatty acid synthase	Probable fatty acid synthase Fas (fatty acid synthetase)	lipid metabolism	COG4982I
Rv2528c	mrr	33648.2	Cytoplasmic	restriction system protein	Probable restriction system protein Mrr	information pathways	COG1715V
Rv2531c	-	106029	Cytoplasmic	amino acid decarboxylase	Probable amino acid decarboxylase	intermediary metabolism and respiration	COG1982E

Rv2535c	pepQ	38758.9	Cytoplasmic	cytoplasmic peptidase PepQ	Probable cytoplasmic peptidase PepQ	intermediary metabolism and respiration	COG0006E
Rv2538c	aroB	38119.1	Cytoplasmic	3-dehydroquinate synthase	3-dehydroquinate synthase AroB	intermediary metabolism and respiration	COG0337E
Rv2555c	alaS	97324.8	Cytoplasmic	alanine--tRNA ligase	Probable alanyl-tRNA synthetase AlaS (alanine--tRNA ligase) (alanine transase) (ALARS)	information pathways	COG0013J
Rv2563	-	36025.4	Cytoplasmic Membrane	glutamine ABC transporter permease	Probable glutamine-transport transmembrane protein ABC transporter	cell wall and cell processes	COG4591M
Rv2564	glnQ	35428.7	Cytoplasmic Membrane	glutamine ABC transporter ATP-binding protein	Probable glutamine-transport ATP-binding protein ABC transporter GlnQ	cell wall and cell processes	COG1136V
Rv2565	-	62123.8	Cytoplasmic	NTE family protein	cAMP-binding proteins - catabolite gene activator and regulatory subunit of cAMP-dependent protein kinases	Regulatory proteins	COG1752R
Rv2566	-	124504	Cytoplasmic	hypothetical protein Rv2566	Transglutaminase-like enzymes, putative cysteine proteases	Intermediary metabolism and respiration	COG4196S
Rv2572c	aspS	65194.8	Cytoplasmic	aspartate--tRNA ligase	Probable aspartyl-tRNA synthetase AspS (aspartate--tRNA ligase) (ASPRS) (aspartic acid transase)	information pathways	COG0173J
Rv2582	ppiB	32370.9	Cytoplasmic	peptidyl-prolyl cis-trans isomerase B	Probable peptidyl-prolyl cis-trans isomerase B PpiB (cyclophilin) (PPIase) (rotamase) (peptidylprolyl isomerase)	information pathways	COG0652O
Rv2589	gabT	46780.6	Cytoplasmic	4-aminobutyrate aminotransferase	4-aminobutyrate aminotransferase GabT (gamma-amino-N-butyrate transaminase) (GABA transaminase) (glutamate:succinic semialdehyde transaminase) (GABA aminotransferase) (GABA-at)	intermediary metabolism and respiration	COG0160E
Rv2590	fadD9	127328	Cytoplasmic	fatty-acid--CoA ligase FadD9	Probable fatty-acid-CoA ligase FadD9 (fatty-acid-CoA synthetase) (fatty-acid-CoA synthase)	lipid metabolism	COG1022I
Rv2605c	tesB2	31536.7	Cytoplasmic	acyl-CoA thioesterase TesB	Probable acyl-CoA thioesterase II TesB2 (TEII)	lipid metabolism	COG1946I
Rv2606c	snzP	31365.1	Cytoplasmic	pyridoxine biosynthesis protein	Possible pyridoxine biosynthesis protein SnzP	intermediary metabolism and respiration	COG0214H
Rv2614c	thrS	77121.9	Cytoplasmic	threonine--tRNA ligase	Probable threonyl-tRNA synthetase ThrS (threonine-tRNA synthetase)(ThrRS) (threonine-tRNA ligase)	information pathways	COG0441J
Rv2623	TB31.7	31652.1	Cytoplasmic	universal stress protein	Universal stress protein family protein TB31.7	virulence, detoxification, adaptation	COG0589T
Rv2624c	-	29400.5	Cytoplasmic	universal stress protein	Universal stress protein family protein	virulence, detoxification, adaptation	COG0589T
Rv2626c	hrp1	15517.7	Unknown	hypoxic response protein	FOG- CBS domain	Regulatory proteins	COG3620K
Rv2627c	-	46252.3	Cytoplasmic	hypothetical protein Rv2627c	Conserved protein	conserved hypotheticals	-
Rv2629	-	40839.9	Cytoplasmic	hypothetical protein Rv2629	Conserved protein	conserved hypotheticals	COG1503J

Rv2676c	-	26249.8	Cytoplasmic	hypothetical protein Rv2676c	porphyrin biosynthetic process	Regulatory proteins	COG3253S
Rv2691	ceoB	24239.5	Cytoplasmic Membrane	TRK system potassium uptake protein CeoB	TRK system potassium uptake protein CeoB	cell wall and cell processes	COG0569P
Rv2692	ceoC	23947.4	Cytoplasmic Membrane	TRK system potassium uptake protein CeoC	TRK system potassium uptake protein CeoC	cell wall and cell processes	COG0569P
Rv2703	sigA	57768.2	Cytoplasmic	RNA polymerase sigma factor SigA	RNA polymerase sigma factor SigA (sigma-A)	information pathways	COG0568K
Rv2711	ideR	25232.9	Cytoplasmic	iron-dependent repressor and activator IdeR	Iron-dependent repressor and activator IdeR	Regulatory proteins	COG1321K
Rv2713	sthA	50754	Cytoplasmic	pyridine nucleotide transhydrogenase	Probable soluble pyridine nucleotide transhydrogenase SthA (STH) (NAD(P)(+) transhydrogenase [B-specific]) (nicotinamide nucleotide transhydrogenase)	intermediary metabolism and respiration	COG1249C
Rv2714	-	35519.8	Cytoplasmic	alanine/leucine-rich protein	Conserved alanine and leucine rich protein	conserved hypotheticals	-
Rv2715	-	36932.6	Cytoplasmic	hydrolase	Possible hydrolase	intermediary metabolism and respiration	COG0596R
Rv2721c	-	72339.7	Cellwall	transmembrane alanine/glycine-rich protein	Possible conserved transmembrane alanine and glycine rich protein	cell wall and cell processes	COG5479M
Rv2731	-	49796.1	Cytoplasmic	alanine/arginine-rich protein	Conserved alanine and arginine rich protein	conserved hypotheticals	COG4717S
Rv2737c	recA	85389.1	Cytoplasmic	recombinase A	RecA protein (recombinase A) [contains: endonuclease PI-MTUI (MTU RecA intein)].	information pathways	COG0468L
Rv2740	ephG	16593	Unknown	epoxide hydrolase	Epoxide hydrolase	virulence, detoxification, adaptation	COG4308Q
Rv2744c	35kd_ag	29257.9	Cytoplasmic	alanine-rich protein	Phage shock protein A (IM30), suppresses sigma54-dependent transcription	Insertion seqs and phages	COG1842KT
Rv2751	-	33610.4	Cytoplasmic Membrane	hypothetical protein Rv2751	O-Methyltransferase involved in polyketide biosynthesis	Lipid metabolism	COG3315Q
Rv2753c	dapA	30826.2	Cytoplasmic	4-hydroxy-tetrahydrodipicolinate synthase	Probable dihydrodipicolinate synthase DapA (DHDPS) (dihydrodipicolinate synthetase)	intermediary metabolism and respiration	COG0329EM
Rv2773c	dapB	25733.4	Cytoplasmic	4-hydroxy-tetrahydrodipicolinate reductase	Dihydrodipicolinate reductase DapB (DHPR)	intermediary metabolism and respiration	COG0289E
Rv2780	ald	38713.2	Extracellular	L-alanine dehydrogenase	Secreted L-alanine dehydrogenase Ald (40 kDa antigen) (TB43)	intermediary metabolism and respiration	COG0686E
Rv2783c	gpsI	79734.7	Cytoplasmic	bifunctional guanosine pentaphosphate synthetase/polyribonucleotide nucleotidyltransferase	Bifunctional protein polyribonucleotide nucleotidyltransferase GpsI: guanosine pentaphosphate synthetase + polyribonucleotide nucleotidyltransferase (polynucleotide phosphorylase) (pnpase)	information pathways	COG1185J
Rv2785c	rpsO	10443.2	Cytoplasmic	30S ribosomal protein S15	30S ribosomal protein S15 RpsO	information pathways	COG0184J
Rv2789c	fadE2_1	44743.4	Cytoplasmic	acyl-CoA dehydrogenase FadE21	Probable acyl-CoA dehydrogenase FadE21	lipid metabolism	COG1960I

Rv2831	echA16	26630.4	Unknown	enoyl-CoA hydratase EchA16	Probable enoyl-CoA hydratase EchA16 (enoyl hydratase) (unsaturated acyl-CoA hydratase) (crotonase)	lipid metabolism	COG1024I
Rv2839c	infB	94008.9	Cytoplasmic	translation initiation factor IF-2	Probable translation initiation factor if-2 InfB	information pathways	COG0532J
Rv2845c	proS	63269.6	Cytoplasmic	proline--tRNA ligase	Probable prolyl-tRNA synthetase ProS (proline--tRNA ligase) (PRORS) (global RNA synthesis factor) (proline transase)	information pathways	COG0442J
Rv2852c	mqo	53563.2	Cellwall	malate:quinone oxidoreductase	Probable malate:quinone oxidoreductase Mqo (malate dehydrogenase [acceptor])	intermediary metabolism and respiration	COG0579R
Rv2855	mtr	49945.7	Cytoplasmic	mycothione reductase	NADPH-dependent mycothiol reductase Mtr	intermediary metabolism and respiration	COG1249C
Rv2858c	aldC	48258.7	Cytoplasmic	aldehyde dehydrogenase AldC	Probable aldehyde dehydrogenase AldC	intermediary metabolism and respiration	COG1012C
Rv2860c	glnA4	49715.4	Cytoplasmic	glutamine synthetase	Probable glutamine synthetase GlnA4 (glutamine synthase) (GS-II)	intermediary metabolism and respiration	COG0174E
Rv2868c	gcpE	40450.3	Cytoplasmic	4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase	Enzyme involved in the deoxyxylulose pathway of isoprenoid biosynthesis	Intermediary metabolism and respiration	COG0821I
Rv2882c	frr	20827.6	Cytoplasmic	ribosome recycling factor	Ribosome recycling factor Frr (ribosome releasing factor) (RRF)	information pathways	COG0233J
Rv2883c	pyrH	27429.6	Cytoplasmic	uridylate kinase	Probable uridylate kinase PyrH (UK) (uridine monophosphate kinase) (UMP kinase)	intermediary metabolism and respiration	COG0528F
Rv2888c	amiC	50916.1	Cytoplasmic	amidase AmiC	Probable amidase AmiC (aminohydrolase)	intermediary metabolism and respiration	COG0154J
Rv2889c	tsf	28754.7	Cytoplasmic	elongation factor EF-Ts	Probable elongation factor Tsf (EF-ts)	information pathways	COG0264J
Rv2890c	rpsB	31089.3	Cytoplasmic	30S ribosomal protein S2	30S ribosomal protein S2 RpsB	information pathways	COG0052J
Rv2895c	viuB	30570.6	Cytoplasmic	mycobactin utilization protein ViuB	Possible mycobactin utilization protein ViuB	intermediary metabolism and respiration	COG2375P
Rv2904c	rplS	13013.1	Cytoplasmic	50S ribosomal protein L19	50S ribosomal protein L19 RplS	information pathways	COG0335J
Rv2916c	ffh	54969.5	Cytoplasmic Membrane	signal recognition particle protein	Probable signal recognition particle protein Ffh (fifty-four homolog) (SRP protein)	cell wall and cell processes	COG0541U
Rv2921c	ftsY	43999.2	Cytoplasmic Membrane	signal recognition particle receptor FtsY	Probable cell division protein FtsY (SRP receptor) (signal recognition particle receptor)	cell wall and cell processes	COG0552U
Rv2922c	smc	130605	Cytoplasmic	chromosome partition protein Smc	Probable chromosome partition protein Smc	cell wall and cell processes	COG1196D
Rv2927c	-	26953.8	Cytoplasmic	hypothetical protein Rv2927c	Cell division initiation protein	Regulatory proteins	COG3599D
Rv2928	tesA	29151.8	Cytoplasmic	thioesterase TesA	Probable thioesterase TesA	lipid metabolism	COG3208Q
Rv2930	fadD2	63044	Cytoplasmic	fatty-acid--CoA ligase FadD26	Fatty-acid-AMP ligase FadD26 (fatty-acid-AMP)	lipid metabolism	COG0318IQ

	6				synthetase) (fatty-acid-AMP synthase)		
Rv2931	ppsA	198835	Unknown	phthiocerol synthesis polyketide synthase type I PpsA	Phenolphthiocerol synthesis type-I polyketide synthase PpsA	lipid metabolism	COG3321Q
Rv2932	ppsB	162496	Cytoplasmic Membrane	phthiocerol synthesis polyketide synthase type I PpsB	Phenolphthiocerol synthesis type-I polyketide synthase PpsB	lipid metabolism	COG3321Q
Rv2933	ppsC	230622	Cytoplasmic Membrane	phthiocerol synthesis polyketide synthase type I PpsC	Phenolphthiocerol synthesis type-I polyketide synthase PpsC	lipid metabolism	COG3321Q
Rv2934	ppsD	193315	Cytoplasmic Membrane	phthiocerol synthesis polyketide synthase type I PpsD	Phenolphthiocerol synthesis type-I polyketide synthase PpsD	lipid metabolism	COG3321Q
Rv2935	ppsE	158713	Cytoplasmic Membrane	phthiocerol synthesis polyketide synthase type I PpsE	Phenolphthiocerol synthesis type-I polyketide synthase PpsE	lipid metabolism	COG3321Q
Rv2940c	mas	224380	Cytoplasmic Membrane	multifunctional mycocerosic acid synthase	Probable multifunctional mycocerosic acid synthase membrane-associated Mas	lipid metabolism	COG3321Q
Rv2941	fadD28	62640.7	Cytoplasmic	long-chain-fatty-acid--AMP ligase FadD28	Fatty-acid-AMP ligase FadD28 (fatty-acid-AMP synthetase) (fatty-acid-AMP synthase)	lipid metabolism	COG03181Q
Rv2945c	lppX	24140.2	Cytoplasmic Membrane	lipoprotein LppX	Probable conserved lipoprotein LppX	cell wall and cell processes	-
Rv2950c	fadD29	67446.5	Cytoplasmic	long-chain-fatty-acid--AMP ligase FadD29	Fatty-acid-AMP ligase FadD29 (fatty-acid-AMP synthetase) (fatty-acid-AMP synthase)	lipid metabolism	-
Rv2951c	-	41315.6	Cytoplasmic	phthiodiolone/phenolphthiodiolone dimycocerosates ketoreductase	Possible oxidoreductase	intermediary metabolism and respiration	COG2141C
Rv2953	-	45104.2	Unknown	trans-acting enoyl reductase	Enoyl reductase	lipid metabolism	COG3268S
Rv2967c	pca	120392	Cytoplasmic	pyruvate carboxylase	Probable pyruvate carboxylase Pca (pyruvic carboxylase)	intermediary metabolism and respiration	COG1038C
Rv2969c	-	26829.5	Cytoplasmic	hypothetical protein Rv2969c	Possible conserved membrane or secreted protein	cell wall and cell processes	COG1651O
Rv2970c	lipN	40134.7	Cytoplasmic	lipase/esterase LipN	Probable lipase/esterase LipN	intermediary metabolism and respiration	COG0657I
Rv2971	-	30332	Cytoplasmic	oxidoreductase	Probable oxidoreductase	intermediary metabolism and respiration	COG0656R
Rv2981c	ddlA	39678.2	Cytoplasmic	D-alanine--D-alanine ligase	Probable D-alanine--D-alanine ligase DdlA (D-alanylalanine synthetase) (D-ala-D-ala ligase)	cell wall and cell processes	COG1181M
Rv2984	ppk1	83039.3	Cytoplasmic Membrane	polyphosphate kinase	Polyphosphate kinase PPK (polyphosphoric acid kinase) (ATP-polyphosphate phosphotransferase)	intermediary metabolism and respiration	COG0855P
Rv2986c	hupB	22187.3	Cytoplasmic	DNA-binding protein HU	DNA-binding protein HU homolog HupB (histone-like protein) (HLP) (21-kDa laminin-2-binding protein)	information pathways	COG0776L
Rv2987c	leuD	21779.8	Cytoplasmic	3-isopropylmalate dehydratase small subunit	Probable 3-isopropylmalate dehydratase (small subunit) LeuD (isopropylmalate isomerase) (alpha-IPM isomerase) (IPMI)	intermediary metabolism and respiration	COG0066E

Rv2988c	leuC	50185	Cytoplasmic	3-isopropylmalate dehydratase large subunit	Probable 3-isopropylmalate dehydratase (large subunit) LeuC (isopropylmalate isomerase) (alpha-IPM isomerase) (IPMI)	intermediary metabolism and respiration	COG0065E
Rv2992c	gltS	53832	Cytoplasmic	glutamate--tRNA ligase	Glutamyl-tRNA synthetase GltS (glutamate--tRNA ligase) (glutamyl-tRNA synthase) (GLURS)	information pathways	COG0008J
Rv2995c	leuB	35306.1	Cytoplasmic	3-isopropylmalate dehydrogenase	Probable 3-isopropylmalate dehydrogenase LeuB (beta-IPM dehydrogenase) (IMDH) (3-IPM-DH)	intermediary metabolism and respiration	COG0473CE
Rv2996c	serA1	54522.2	Cytoplasmic	D-3-phosphoglycerate dehydrogenase	Probable D-3-phosphoglycerate dehydrogenase SerA1 (PGDH)	intermediary metabolism and respiration	COG0111HE
Rv3001c	ilvC	36090.1	Cytoplasmic	ketol-acid reductoisomerase	Probable KETOL-acid reductoisomerase IlvC (acetohydroxy-acid isomeroeductase) (alpha-keto-beta-hydroxylacil reductoisomerase)	intermediary metabolism and respiration	-
Rv3002c	ilvN	18186.9	Cytoplasmic	acetolactate synthase small subunit	Probable acetolactate synthase (small subunit) IlvN (acetohydroxy-acid synthase) (AHAS) (ALS)	intermediary metabolism and respiration	COG0440E
Rv3003c	ilvB1	66090.7	Cytoplasmic	acetolactate synthase large subunit IlvB	Acetolactate synthase (large subunit) IlvB1 (acetohydroxy-acid synthase)	intermediary metabolism and respiration	COG0028EH
Rv3006	lppZ	38752.1	Unknown	lipoprotein LppZ	Probable conserved lipoprotein LppZ	cell wall and cell processes	COG2133G
Rv3009c	gatB	54632.9	Cytoplasmic	aspartyl/glutamyl-tRNA(Asn/Gln) amidotransferase subunit B	Probable glutamyl-tRNA(GLN) amidotransferase (subunit B) GatB (Glu-ADT subunit B)	information pathways	COG0064J
Rv3010c	pfkA	36880	Cytoplasmic	6-phosphofructokinase	Probable 6-phosphofructokinase PfkA (phosphohexokinase) (phosphofructokinase)	intermediary metabolism and respiration	COG0205G
Rv3011c	gatA	51388.2	Cytoplasmic	glutamyl-tRNA(GLN) amidotransferase subunit A	Probable glutamyl-tRNA(GLN) amidotransferase (subunit A) GatA (Glu-ADT subunit A)	information pathways	-
Rv3014c	ligA	75257.1	Cytoplasmic	DNA ligase A	DNA ligase [NAD dependent] LigA (polydeoxyribonucleotide synthase [NAD+])	information pathways	COG0272L
Rv3015c	-	34211.9	Unknown	hypothetical protein Rv3015c	B12 independent protein	Regulatory proteins	COG0620E
Rv3028c		31691	Cytoplasmic	electron transfer flavoprotein subunit alpha	Probable electron transfer flavoprotein (alpha-subunit) FixB (alpha-ETF) (electron transfer flavoprotein large subunit) (ETFLS)	intermediary metabolism and respiration	COG2025C
Rv3029c	fixA	28080.8	Cytoplasmic	electron transfer flavoprotein subunit beta	Probable electron transfer flavoprotein (beta-subunit) FixA (beta-ETF) (electron transfer flavoprotein small subunit) (ETFSS)	intermediary metabolism and respiration	COG2086C
Rv3040c	-	31485.2	Cytoplasmic	hypothetical protein Rv3040c	NTP pyrophosphohydrolases including oxidative damage repair enzymes	Information pathways	-
Rv3042c	serB2	43058.7	Cytoplasmic	phosphoserine phosphatase SerB	Probable phosphoserine phosphatase SerB2 (PSP) (O-phosphoserine phosphohydrolase) (pspase)	intermediary metabolism and respiration	COG0560E
Rv3044	fecB	36910.4	Unknown	FeIII-dicitrate-binding periplasmic lipoprotein	Probable FEIII-dicitrate-binding periplasmic lipoprotein FecB	cell wall and cell processes	COG0614P

Rv3045	adhC	37075.3	Cytoplasmic	NADP-dependent alcohol dehydrogenase	Probable NADP-dependent alcohol dehydrogenase AdhC	intermediary metabolism and respiration	COG1064R
Rv3050c	-	27586.6	Cytoplasmic	AsnC family transcriptional regulator	Probable transcriptional regulatory protein (probably AsnC-family)	Regulatory proteins	COG1309K
Rv3051c	nrdE	79290.8	Cytoplasmic	ribonucleoside-diphosphate reductase subunit alpha	Ribonucleoside-diphosphate reductase (alpha chain) NrdE (ribonucleotide reductase small subunit) (R1F protein)	information pathways	COG0209F
Rv3058c	-	23803	Cytoplasmic	TetR family transcriptional regulator	Possible transcriptional regulatory protein (probably TetR-family)	Regulatory proteins	COG1309K
Rv3061c	fadE2 2	76741.1	Cytoplasmic	acyl-CoA dehydrogenase FadE22	Probable acyl-CoA dehydrogenase FadE22	lipid metabolism	COG1960I
Rv3075c	-	33100.2	Cytoplasmic	hypothetical protein Rv3075c	Citrate lyase beta subunit	Intermediary metabolism and respiration	COG2301G
Rv3088	tgs4	50886.5	Unknown	diacylglycerol O-acyltransferase	Putative triacylglycerol synthase (diacylglycerol acyltransferase) Tgs4	lipid metabolism	-
Rv3090	-	32270.9	Unknown	alanine/valine-rich protein	Membrane protease subunits, stomatin/prohibitin homologs	cell wall and cell processes	COG0330O
Rv3091	-	61645.5	Cytoplasmic	hypothetical protein Rv3091	Conserved protein	conserved hypotheticals	COG1752R
Rv3099c	-	30475.7	Cytoplasmic	hypothetical protein Rv3099c	Conserved protein	conserved hypotheticals	-
Rv3102c	ftsE	25596.5	Cytoplasmic Membrane	cell division ATP-binding protein FtsE	Putative cell division ATP-binding protein FtsE (septation component-transport ATP-binding protein ABC transporter)	cell wall and cell processes	COG2884D
Rv3105c	prfB	42125	Cytoplasmic	peptide chain release factor PrfB	Probable peptide chain release factor 2 PrfB (RF-2)	information pathways	COG1186J
Rv3106	fprA	49341.1	Cytoplasmic	NADPH-ferredoxin reductase FprA	NADPH:adrenodoxin oxidoreductase FprA (NADPH-ferredoxin reductase)	intermediary metabolism and respiration	COG0493ER
Rv3107c	agpS	56520.5	Cytoplasmic	alkyldihydroxyacetonephosphate synthase	Possible alkyldihydroxyacetonephosphate synthase AgpS (alkyl-DHAP synthase) (alkylglycerone-phosphate synthase)	lipid metabolism	COG0277C
Rv3127	-	38488.6	Cytoplasmic	hypothetical protein Rv3127	Conserved protein	conserved hypotheticals	-
Rv3130c	tgs1	50720.6	Cytoplasmic	diacylglycerol O-acyltransferase	Triacylglycerol synthase (diacylglycerol acyltransferase) Tgs1	lipid metabolism	-
Rv3131	-	35978.4	Cytoplasmic Membrane	NAD(P)H nitroreductase	diacylglycerol O-acyltransferase activity	Intermediary metabolism and respiration	-
Rv3132c	devS	62240.6	Cytoplasmic Membrane	two component sensor histidine kinase DevS	Two component sensor histidine kinase DevS	Regulatory proteins	COG4585T
Rv3133c	devR	23261.9	Cytoplasmic	two component transcriptional regulator DevR	Two component transcriptional regulatory protein DevR (probably LuxR/UhpA-family)	Regulatory proteins	COG2197TK
Rv3136	PPE51	37979.9	Cytoplasmic	PPE family protein PPE51	PPE family protein PPE51	PE/PPE	COG5651N

			Membrane				
Rv3139	fadE2 4	49646.4	Cytoplasmic	acyl-CoA dehydrogenase FadE24	Probable acyl-CoA dehydrogenase FadE24	lipid metabolism	COG1960I
Rv3140	fadE2 3	43344.2	Cytoplasmic	acyl-CoA dehydrogenase FadE23	Probable acyl-CoA dehydrogenase FadE23	lipid metabolism	COG1960I
Rv3141	fadB4	33852.8	Cytoplasmic	NADPH quinone oxidoreductase FadB	Probable NADPH quinone oxidoreductase FadB4 (NADPH:quinone reductase) (zeta-crystallin)	lipid metabolism	COG0604CR
Rv3147	nuoC	26932.1	Cytoplasmic	NADH-quinone oxidoreductase subunit C	Probable NADH dehydrogenase I (chain C) NuoC (NADH-ubiquinone oxidoreductase chain C)	intermediary metabolism and respiration	COG0852C
Rv3148	nuoD	48164.5	Cytoplasmic	NADH-quinone oxidoreductase subunit D	Probable NADH dehydrogenase I (chain D) NuoD (NADH-ubiquinone oxidoreductase chain D)	intermediary metabolism and respiration	COG0649C
Rv3149	nuoE	27165.6	Cytoplasmic	NADH-quinone oxidoreductase subunit E	Probable NADH dehydrogenase I (chain E) NuoE (NADH-ubiquinone oxidoreductase chain E)	intermediary metabolism and respiration	COG1905C
Rv3151	nuoG	85392.1	Cytoplasmic	NADH-quinone oxidoreductase subunit G	Probable NADH dehydrogenase I (chain G) NuoG (NADH-ubiquinone oxidoreductase chain G)	intermediary metabolism and respiration	COG1034C
Rv3161c	-	42504.2	Unknown	dioxygenase	Possible dioxygenase	intermediary metabolism and respiration	COG4638PR
Rv3168	-	42529.2	Cytoplasmic	aminoglycoside phosphotransferase	Putative aminoglycoside phosphotransferase	virulence, detoxification, adaptation	COG3173R
Rv3169	-	41759.8	Cytoplasmic	hypothetical protein Rv3169	Conserved protein	conserved hypotheticals	-
Rv3193c	-	107470	Cytoplasmic Membrane	transmembrane protein	Probable conserved transmembrane protein	cell wall and cell processes	COG1615S
Rv3197	-	49121.9	Cytoplasmic Membrane	ABC transporter ATP-binding protein	Probable conserved ATP-binding protein ABC transporter	cell wall and cell processes	COG0661R
Rv3200c	-	38099.9	Cytoplasmic Membrane	transmembrane cation transporter	Possible transmembrane cation transporter	cell wall and cell processes	COG0569P
Rv3205c	-	31351.9	Cytoplasmic	hypothetical protein Rv3205c	membrane associated process	cell wall and cell processes	-
Rv3206c	moeB 1	42141.4	Cytoplasmic	adenylyltransferase/sulfurtransferase	Probable molybdenum cofactor biosynthesis protein MoeB1 (MPT-synthase sulfurylase) (molybdopterin synthase sulphurylase)	intermediary metabolism and respiration	COG0476H
Rv3208A	TB9.4	9400.53	Cytoplasmic	hypothetical protein Rv3208A	Conserved protein TB9.4	conserved hypotheticals	-
Rv3211	rhIE	56703.1	Cytoplasmic	ATP-dependent RNA helicase RhIE	Probable ATP-dependent RNA helicase RhIE	information pathways	COG0513LKJ
Rv3215	entC	38905.7	Cytoplasmic	isochorismate synthase	Probable isochorismate synthase EntC (isochorismate hydroxymutase) (enterochelin biosynthesis)	intermediary metabolism and respiration	COG1169HQ
Rv3224	-	29814.4	Cytoplasmic	iron-regulated short-chain	Possible iron-regulated short-chain	intermediary	COG1028IQR

				dehydrogenase/reductase	dehydrogenase/reductase	metabolism and respiration	
Rv3227	aroA	46393.8	Unknown	3-phosphoshikimate 1-carboxyvinyltransferase	3-phosphoshikimate 1-carboxyvinyltransferase AroA (5-enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS)	intermediary metabolism and respiration	COG0128E
Rv3230c	-	40794.7	Cytoplasmic	stearoyl-CoA 9-desaturase electron transfer partner	Hypothetical oxidoreductase	intermediary metabolism and respiration	COG1018C
Rv3240c	secA1	105990	Cytoplasmic	protein translocase subunit SecA	Probable preprotein translocase SecA1 1 subunit	cell wall and cell processes	COG0653U
Rv3246c	mtrA	25279.3	Cytoplasmic	two component DNA-binding response regulator MtrA	Two component sensory transduction transcriptional regulatory protein MtrA	Regulatory proteins	COG0745TK
Rv3248c	sahH	54323.6	Cytoplasmic	adenosylhomocysteinase	Probable adenosylhomocysteinase SahH (S-adenosyl-L-homocysteine hydrolase) (adohcysase)	intermediary metabolism and respiration	COG0499H
Rv3254	-	51518.1	Cytoplasmic Membrane	hypothetical protein Rv3254	Monooxygenase	Intermediary metabolism and respiration	COG0654HC
Rv3255c	manA	43339.2	Cytoplasmic	mannose-6-phosphate isomerase	Probable mannose-6-phosphate isomerase ManA (phosphomannose isomerase) (phosphomannoisomerase) (PMI) (phosphohexoisomerase) (phosphohexomutase)	intermediary metabolism and respiration	COG1482G
Rv3256c	-	35277.4	Unknown	hypothetical protein Rv3256c	Sugar binding protein	Intermediary metabolism and respiration	-
Rv3257c	pmmA	49041	Cytoplasmic	phosphomannomutase PmmA	Probable phosphomannomutase PmmA (PMM) (phosphomannose mutase)	intermediary metabolism and respiration	COG1109G
Rv3262	fbiB	47560.3	Cytoplasmic	coenzyme F420:L-glutamate ligase	Probable F420 biosynthesis protein FbiB	intermediary metabolism and respiration	COG1478S
Rv3264c	manB	37839	Cytoplasmic	D-alpha-D-mannose-1-phosphate guanylyltransferase ManB	D-alpha-D-mannose-1-phosphate guanylyltransferase ManB (D-alpha-D-heptose-1-phosphate guanylyltransferase)	cell wall and cell processes	COG1208MJ
Rv3273	-	80533.6	Cytoplasmic Membrane	transmembrane carbonic anhydrase	Probable transmembrane carbonic anhydrase (carbonate dehydratase) (carbonic dehydratase)	intermediary metabolism and respiration	COG0659P
Rv3274c	fadE25	41723.4	Cytoplasmic	acyl-CoA dehydrogenase FadE25	Probable acyl-CoA dehydrogenase FadE25	lipid metabolism	COG1960I
Rv3276c	purK	45695.1	Cytoplasmic Membrane	N5-carboxyaminoimidazole ribonucleotide synthase	Probable phosphoribosylaminoimidazole carboxylase ATPase subunit PurK (air carboxylase) (AIRC)	intermediary metabolism and respiration	COG0026F
Rv3280	accD5	59354.2	Cytoplasmic	propionyl-CoA carboxylase subunit beta	Probable propionyl-CoA carboxylase beta chain 5 AccD5 (pccase) (propanoyl-CoA:carbon dioxide ligase)	lipid metabolism	COG4799I
Rv3283	sseA	33288.4	Cytoplasmic	thiosulfate sulfurtransferase SseA	Probable thiosulfate sulfurtransferase SseA (rhodanese) (thiosulfate cyanide transsulfurase) (thiosulfate	intermediary metabolism and	COG2897P

					thiotransferase)	respiration	
Rv3285	accA3	63751	Cytoplasmic	bifunctional protein acetyl-/propionyl-CoA carboxylase subunit alpha AccA	Probable bifunctional protein acetyl-/propionyl-coenzyme A carboxylase (alpha chain) AccA3: biotin carboxylase + biotin carboxyl carrier protein (BCCP)	lipid metabolism	COG4770I
Rv3290c	lat	49012.1	Cytoplasmic	L-lysine-epsilon aminotransferase	Probable L-lysine-epsilon aminotransferase Lat (L-lysine aminotransferase) (lysine 6-aminotransferase)	intermediary metabolism and respiration	COG0160E
Rv3293	pcd	51312.2	Cytoplasmic	piperideine-6-carboxylic acid dehydrogenase	Probable piperideine-6-carboxylic acid dehydrogenase Pcd (piperideine-6-carboxylate dehydrogenase)	intermediary metabolism and respiration	COG1012C
Rv3295	-	24805.3	Cytoplasmic Membrane	TetR family transcriptional regulator	Probable transcriptional regulatory protein (probably TetR-family)	Regulatory proteins	COG1309K
Rv3302c	glpD2	62745.8	Cytoplasmic	glycerol-3-phosphate dehydrogenase	Probable glycerol-3-phosphate dehydrogenase GlpD2	intermediary metabolism and respiration	COG0578C
Rv3305c	amiA1	41227.1	Cytoplasmic	N-acyl-L-amino acid amidohydrolase AmiA	Possible N-acyl-L-amino acid amidohydrolase AmiA1 (N-acyl-L-amino acid aminohydrolase)	intermediary metabolism and respiration	COG1473R
Rv3311	-	45732.3	Cytoplasmic	hypothetical protein Rv3311	Conserved protein	conserved hypotheticals	-
Rv3314c	deoA	44454	Cytoplasmic	thymidine phosphorylase	Probable thymidine phosphorylase DeoA (tdrpase) (pyrimidine phosphorylase)	intermediary metabolism and respiration	COG0213F
Rv3318	sdhA	64807.5	Cytoplasmic Membrane	succinate dehydrogenase flavoprotein subunit	Probable succinate dehydrogenase (flavoprotein subunit) SdhA (succinic dehydrogenase) (fumarate reductase) (fumarate dehydrogenase) (fumaric hydrogenase)	intermediary metabolism and respiration	COG1053C
Rv3319	sdhB	29334.9	Cytoplasmic Membrane	succinate dehydrogenase iron-sulphur protein subunit	Probable succinate dehydrogenase (iron-sulphur protein subunit) SdhB (succinic dehydrogenase) (fumarate reductase) (fumarate dehydrogenase) (fumaric hydrogenase)	intermediary metabolism and respiration	COG0479C
Rv3340	metC	47371.2	Cytoplasmic	O-acetylhomoserine sulfhydrylase	Probable O-acetylhomoserine sulfhydrylase MetC (homocysteine synthase) (O-acetylhomoserine (thiol)-lyase) (OAH sulfhydrylase) (O-acetyl-L-homoserine sulfhydrylase)	intermediary metabolism and respiration	COG2873E
Rv3365c	-	93409.6	Cytoplasmic Membrane	hypothetical protein Rv3365c	Signal transduction histidine kinase	Regulatory proteins	COG0642T
Rv3389c	htdY	30296.3	Cytoplasmic	3-hydroxyacyl-thioester dehydratase HtdY	Probable 3-hydroxyacyl-thioester dehydratase HtdY	intermediary metabolism and respiration	COG2030I
Rv3392c	cmaA1	32461.2	Cytoplasmic	cyclopropane mycolic acid synthase CmaA	Cyclopropane-fatty-acyl-phospholipid synthase 1 CmaA1 (cyclopropane fatty acid synthase) (CFA synthase) (cyclopropane mycolic acid synthase 1)	lipid metabolism	COG2230M
Rv3396c	guaA	56027.8	Cytoplasmic	GMP synthase	Probable GMP synthase [glutamine-hydrolyzing] GuaA (glutamine amidotransferase) (GMP synthetase)	intermediary metabolism and respiration	COG0519F

Rv3399	-	38140.2	Cytoplasmic	S-adenosylmethionine-dependent methyltransferase	Possible S-adenosylmethionine-dependent methyltransferase	lipid metabolism	COG3315Q
Rv3400	-	28227.7	Cytoplasmic	hydrolase	Probable hydrolase	intermediary metabolism and respiration	COG0637R
Rv3401	-	87315.5	Cytoplasmic	glycosyl hydrolase	Conserved protein	intermediary metabolism and respiration	COG1554G
Rv3410c	guaB3	38990.6	Cytoplasmic	oxidoreductase	Probable inosine-5'-monophosphate dehydrogenase GuaB3 (imp dehydrogenase) (inosinic acid dehydrogenase) (inosinate dehydrogenase) (imp oxidoreductase) (inosine-5'-monophosphate oxidoreductase) (IMPDH) (IMPD)	intermediary metabolism and respiration	COG0516F
Rv3411c	guaB2	54867	Unknown	inosine-5'-monophosphate dehydrogenase	Probable inosine-5'-monophosphate dehydrogenase GuaB2 (imp dehydrogenase) (inosinic acid dehydrogenase) (inosinate dehydrogenase) (imp oxidoreductase) (inosine-5'-monophosphate oxidoreductase) (IMPDH) (IMPD)	intermediary metabolism and respiration	COG0516F
Rv3417c	groEL ₁	55877.5	Cytoplasmic	chaperonin GroEL	60 kDa chaperonin 1 GroEL1 (protein CPN60-1) (GroEL protein 1)	virulence, detoxification, adaptation	COG0459O
Rv3418c	groES	10772.2	Cytoplasmic	chaperonin GroES	10 kDa chaperonin GroES (protein CPN10) (protein GroES) (BCG-a heat shock protein) (10 kDa antigen)	virulence, detoxification, adaptation	COG0234O
Rv3432c	gadB	50747.7	Cytoplasmic	glutamate decarboxylase GadB	Probable glutamate decarboxylase GadB	intermediary metabolism and respiration	COG0076E
Rv3436c	glmS	67539.7	Cytoplasmic	glucosamine--fructose-6-phosphate aminotransferase	Probable glucosamine--fructose-6-phosphate aminotransferase [isomerizing] GlmS (hexosephosphate aminotransferase) (D-fructose-6-phosphate amidotransferase) (GFAT) (L-glutamine-D-fructose-6-phosphate amidotransferase) (glucosamine-6-phosphate synthase)	intermediary metabolism and respiration	COG0449M
Rv3438	-	29208.9	Unknown	hypothetical protein Rv3438	membrane associated process	cell wall and cell processes	COG0412Q
Rv3443c	rplM	16304.9	Cytoplasmic	50S ribosomal protein L13	50S ribosomal protein L13 RplM	information pathways	COG0102J
Rv3456c	rplQ	19475.2	Cytoplasmic	50S ribosomal protein L17	50S ribosomal protein L17 RplQ	information pathways	COG0203J
Rv3457c	rpoA	37706.5	Cytoplasmic	DNA-directed RNA polymerase subunit alpha	Probable DNA-directed RNA polymerase (alpha chain) RpoA (transcriptase alpha chain) (RNA polymerase alpha subunit) (DNA-directed RNA nucleotidyltransferase)	information pathways	COG0202K
Rv3458c	rpsD	23475.8	Cytoplasmic	30S ribosomal protein S4	30S ribosomal protein S4 RpsD	information pathways	COG0522J
Rv3460c	rpsM	14350.7	Cytoplasmic	30S ribosomal protein S13	30S ribosomal protein S13 RpsM	information pathways	COG0099J
Rv3464	rmlB	37557.7	Cytoplasmic	dTDP-glucose 4,6-dehydratase	dTDP-glucose 4,6-dehydratase RmlB	intermediary metabolism and respiration	COG1088M

Rv3485c	-	33193.8	Cytoplasmic	short-chain type dehydrogenase/reductase	Probable short-chain type dehydrogenase/reductase	intermediary metabolism and respiration	COG1028IQR
Rv3495c	lprN	41293	Unknown	Mce family lipoprotein LprN	Possible Mce-family lipoprotein LprN (Mce-family lipoprotein Mce4E)	cell wall and cell processes	COG1463Q
Rv3496c	mce4D	47201.4	Unknown	Mce family protein Mce4D	Mce-family protein Mce4D	virulence, detoxification, adaptation	COG1463Q
Rv3497c	mce4C	38132.4	Cytoplasmic Membrane	Mce family protein Mce4C	Mce-family protein Mce4C	virulence, detoxification, adaptation	COG1463Q
Rv3498c	mce4B	36880.3	Unknown	Mce family protein Mce4B	Mce-family protein Mce4B	virulence, detoxification, adaptation	COG1463Q
Rv3509c	ilvX	52040.3	Cytoplasmic Membrane	acetohydroxyacid synthase large subunit	Probable acetohydroxyacid synthase IlvX (acetolactate synthase)	intermediary metabolism and respiration	COG0028EH
Rv3519	-	25960	Cytoplasmic	hypothetical protein Rv3519	Carboxy-Lyase	Intermediary metabolism and respiration	-
Rv3520c	-	37584.9	Cytoplasmic	coenzyme F420-dependent oxidoreductase	Possible coenzyme F420-dependent oxidoreductase	intermediary metabolism and respiration	COG2141C
Rv3528c	-	27440.4	Cytoplasmic	hypothetical protein Rv3528c	Unknown protein	conserved hypotheticals	-
Rv3547	ddn	17371	Unknown	deazaflavin-dependent nitroreductase	Deazaflavin-dependent nitroreductase Ddn	intermediary metabolism and respiration	-
Rv3556c	fadA6	40902.3	Cytoplasmic	acetyl-CoA acetyltransferase FadA	Probable acetyl-CoA acetyltransferase FadA6 (acetoacetyl-CoA thiolase)	lipid metabolism	COG0183I
Rv3570c	hsaA	43110.5	Cytoplasmic Membrane	flavin-dependent monooxygenase oxygenase subunit HsaA	Possible oxidoreductase. Possible 3-hydroxy-9,10-seconandrost-1,3,5(10)-triene-9,17-dione hydroxylase.	intermediary metabolism and respiration	COG1960I
Rv3575c	-	38177.7	Cytoplasmic	LacI family transcriptional regulator	Transcriptional regulatory protein (probably LacI-family)	Regulatory proteins	COG1609K
Rv3580c	cysS1	51822.5	Cytoplasmic	cysteine--tRNA ligase	Cysteinyl-tRNA synthetase 1 CysS1 (cysteine--tRNA ligase 1) (CYSRS 1) (cysteine transase)	information pathways	COG0215J
Rv3582c	ispD	24041.5	Cytoplasmic	2-C-methyl-D-erythritol 4-phosphate cytidyltransferase	4-diphosphocytidyl-2C-methyl-D-erythritol synthase IspD (MEP cytidyltransferase) (MCT)	intermediary metabolism and respiration	COG1211I
Rv3584	lpqE	18786.6	Unknown	lipoprotein LpqE	Possible conserved lipoprotein LpqE	cell wall and cell processes	-
Rv3586	-	39031.1	Cytoplasmic	DNA integrity scanning protein DisA	Predicted nucleic-acid-binding protein (contains the HHH domain)	Regulatory proteins	COG1623R
Rv3596c	clpC1	93552.4	Cytoplasmic	ATP-dependent protease ATP-binding subunit ClpC	Probable ATP-dependent protease ATP-binding subunit ClpC1	intermediary metabolism and	COG0542O

						respiration	
Rv3597c	lsr2	12098.4	Cytoplasmic	iron-regulated H-NS-like protein	Iron-regulated H-NS-like protein Lsr2	information pathways	-
Rv3598c	lysS	55677.3	Cytoplasmic	lysine--tRNA ligase	Lysyl-tRNA synthetase 1 LysS (lysine--tRNA ligase 1) (LysRS 1) (lysine transase)	information pathways	COG1190J
Rv3602c	panC	32677.5	Cytoplasmic	pantothenate synthetase	Pantoate--beta-alanine ligase PanC (pantothenate synthetase) (pantoate activating enzyme)	intermediary metabolism and respiration	COG0414H
Rv3616c	espA	39888.3	Cytoplasmic	ESX-1 secretion-associated protein EspA	ESX-1 secretion-associated protein A, EspA	cell wall and cell processes	-
Rv3623	lpqG	24836.8	Unknown	lipoprotein LpqG	Probable conserved lipoprotein LpqG	cell wall and cell processes	COG2968S
Rv3634c	galE1	33549.6	Cytoplasmic	UDP-glucose 4-epimerase	UDP-glucose 4-epimerase GalE1 (galactowaldenase) (UDP-galactose 4-epimerase) (uridine diphosphate galactose 4-epimerase) (uridine diphospho-galactose 4-epimerase)	intermediary metabolism and respiration	COG0451MG
Rv3646c	topA	102318	Cytoplasmic	DNA topoisomerase I	DNA topoisomerase I TopA (omega-protein) (relaxing enzyme) (untwisting enzyme) (swivelase) (type I DNA topoisomerase) (nicking-closing enzyme) (TOPO I)	information pathways	COG0550L
Rv3651	-	37967.3	Unknown	hypothetical protein Rv3651	Conserved hypothetical protein	conserved hypotheticals	-
Rv3667	acs	71443.9	Cytoplasmic	acetyl-CoAsynthetase	Acetyl-coenzyme A synthetase Acs (acetate--CoA ligase) (acetyl-CoA synthetase) (acetyl-CoA synthase) (acyl-activating enzyme) (acetate thiokinase) (acetyl-activating enzyme) (acetate--coenzyme A ligase) (acetyl-coenzyme A synthase)	lipid metabolism	COG0365I
Rv3671c	-	40721.4	Cytoplasmic Membrane	serine protease	Membrane-associated serine protease	intermediary metabolism and respiration	COG0265O
Rv3676	crp	24759.3	Cytoplasmic	cAMP receptor protein	Transcriptional regulatory protein Crp (Crp/Fnr-family)	Regulatory proteins	COG0664T
Rv3678c	-	15245.6	Cytoplasmic	hypothetical protein Rv3678c	Putative translation initiation inhibitor, yjgF family	Information pathways	COG0251J
Rv3679	-	35856.3	Cytoplasmic	anion transporter ATPase	Probable anion transporter ATPase	cell wall and cell processes	COG0003P
Rv3680	-	41404.6	Cytoplasmic Membrane	anion transporter ATPase	Probable anion transporter ATPase	cell wall and cell processes	COG0003P
Rv3683	-	34394.4	Unknown	hypothetical protein Rv3683	Predicted phosphohydrolases	Intermediary metabolism and respiration	COG1408R
Rv3684	-	37623.3	Unknown	lyase	Probable lyase	intermediary metabolism and respiration	-
Rv3692	moxR ₂	37904.4	Cytoplasmic	methanol dehydrogenase transcriptional regulator MoxR	Probable methanol dehydrogenase transcriptional regulatory protein MoxR2	Regulatory proteins	COG0714R
Rv3699	-	24998.1	Cytoplasmic	hypothetical protein Rv3699	SAM-dependent methyltransferases	Intermediary metabolism and respiration	COG2227H

Rv3708c	asd	36230.2	Cytoplasmic	aspartate-semialdehyde dehydrogenase	Aspartate-semialdehyde dehydrogenase Asd (ASA dehydrogenase) (ASADH) (aspartic semialdehyde dehydrogenase) (L-aspartate-beta-semialdehyde dehydrogenase)	intermediary metabolism and respiration	COG0136E
Rv3709c	ask	44429.6	Cytoplasmic	aspartokinase	Aspartokinase Ask (aspartate kinase) [contains: aspartokinase alpha subunit (Ask-alpha); and aspartokinase beta subunit (Ask-beta)]	intermediary metabolism and respiration	COG0527E
Rv3710	leuA	70081.6	Cytoplasmic	2-isopropylmalate synthase	2-isopropylmalate synthase LeuA (alpha-isopropylmalate synthase) (alpha-IPM synthetase) (IPMS)	intermediary metabolism and respiration	COG0119E
Rv3712	-	43389.4	Cytoplasmic	ligase	Possible ligase	intermediary metabolism and respiration	COG0769M
Rv3716c	-	13357.5	Unknown	hypothetical protein Rv3716c	DNA-repair	Information pathways	COG0718S
Rv3718c	-	15661.8	Unknown	hypothetical protein Rv3718c	Conserved protein	conserved hypotheticals	-
Rv3719	-	53007	Unknown	hypothetical protein Rv3719	FAD/FMN-containing dehydrogenases	Intermediary metabolism and respiration	COG0277C
Rv3720	-	46852.8	Cytoplasmic	fatty acid synthase	Possible fatty acid synthase	lipid metabolism	-
Rv3721c	dnaZ X	61891.3	Cytoplasmic	DNA polymerase III subunit gamma/tau	DNA polymerase III (subunit gamma/tau) DnaZ/X	information pathways	COG2812L
Rv3722c	-	47309	Unknown	hypothetical protein Rv3722c	Transcriptional regulators containing a DNA-binding HTH domain and an aminotransferase domain (MocR family) and their eukaryotic orthologs	Regulatory proteins	COG1167KE
Rv3726	-	42035.8	Cytoplasmic	dehydrogenase	Possible dehydrogenase	intermediary metabolism and respiration	COG1063ER
Rv3732	-	37532.8	Unknown	hypothetical protein Rv3732	Conserved protein	conserved hypotheticals	COG4402S
Rv3734c	tgs2	49303.8	Cytoplasmic	diacylglycerol O-acyltransferase	Putative triacylglycerol synthase (diacylglycerol acyltransferase) Tgs2	lipid metabolism	-
Rv3763	lpqH	15114.8	Unknown	lipoprotein LpqH	19 kDa lipoprotein antigen precursor LpqH	cell wall and cell processes	-
Rv3774	echA2 1	29101.2	Unknown	enoyl-CoA hydratase EchA21	Possible enoyl-CoA hydratase EchA21 (enoyl hydratase) (unsaturated acyl-CoA hydratase) (crotonase)	lipid metabolism	COG1024I
Rv3775	lipE	45273.7	Unknown	lipase LipE	Probable lipase LipE	intermediary metabolism and respiration	COG1680V
Rv3778c	-	41851.7	Cytoplasmic	aminotransferase	Possible aminotransferase	intermediary metabolism and respiration	COG0520E
Rv3790	dprE1	50163.2	Unknown	decaprenylphosphoryl-beta-D-ribose oxidase	Decaprenylphosphoryl-beta-D-ribose 2'-oxidase	lipid metabolism	COG0277C
Rv3791	dprE2	27468.8	Cytoplasmic	decaprenylphosphoryl-D-2-keto	Decaprenylphosphoryl-D-2-keto erythro pentose	lipid metabolism	COG0300R

				erythrose reductase	reductase		
Rv3793	embC	117490	Cytoplasmic Membrane	arabinosyltransferase C	Integral membrane indolylacetylaminositol arabinosyltransferase EmbC (arabinosylindolylacetylaminositol synthase)	cell wall and cell processes	-
Rv3794	embA	115692	Cytoplasmic Membrane	arabinosyltransferase A	Integral membrane indolylacetylaminositol arabinosyltransferase EmbA (arabinosylindolylacetylaminositol synthase)	cell wall and cell processes	COG1807M
Rv3799c	accD4	56647.7	Cytoplasmic	propionyl-CoA carboxylase subunit beta AccD	Probable propionyl-CoA carboxylase beta chain 4 AccD4 (pccase) (propanoyl-CoA:carbon dioxide ligase)	lipid metabolism	COG4799I
Rv3800c	pks13	186446	Cytoplasmic Membrane	polyketide synthase	Polyketide synthase Pks13	lipid metabolism	COG3321Q
Rv3801c	fadD3 ₂	69232	Cytoplasmic	long-chain-fatty-acid--AMP ligase FadD32	Fatty-acid-AMP ligase FadD32 (fatty-acid-AMP synthetase) (fatty-acid-AMP synthase). Also shown to have acyl-ACP ligase activity.	lipid metabolism	COG0318IQ
Rv3803c	fbpD	31088.9	Extracellular	MPT51/MPB51 antigen	Secreted MPT51/MPB51 antigen protein FbpD (MPT51/MPB51 antigen 85 complex C) (AG58C) (mycolyl transferase 85C) (fibronectin-binding protein C) (85C)	lipid metabolism	COG0627R
Rv3804c	fbpA	35686.2	Extracellular	diacylglycerol acyltransferase/mycolyltransferase Ag85A	Secreted antigen 85-a FbpA (mycolyl transferase 85A) (fibronectin-binding protein A) (antigen 85 complex A)	lipid metabolism	COG0627R
Rv3808c	glfT2	71506.5	Cytoplasmic	galactofuranosyl transferase GlfT	Bifunctional UDP-galactofuranosyl transferase GlfT2	cell wall and cell processes	COG1216R
Rv3809c	glf	45814.4	Cytoplasmic	UDP-galactopyranose mutase	UDP-galactopyranose mutase Glf (UDP-GALP mutase) (NAD ⁺ -flavin adenine dinucleotide-requiring enzyme)	cell wall and cell processes	COG0562M
Rv3818	-	57615.2	Cytoplasmic	hypothetical protein Rv3818	Predicted Zn-dependent hydrolases of the beta-lactamase fold	Intermediary metabolism and respiration	COG2220R
Rv3825c	pks2	225758	Cytoplasmic Membrane	phthioceranic/hydroxyphthioceranic acid synthase	Polyketide synthase Pks2	lipid metabolism	COG3321Q
Rv3826	fadD2 ₃	62840.8	Cytoplasmic	long-chain-fatty-acid--CoA ligase FadD23	Probable fatty-acid-AMP ligase FadD23 (fatty-acid-AMP synthetase) (fatty-acid-AMP synthase)	lipid metabolism	COG0318IQ
Rv3834c	serS	45292.3	Cytoplasmic	serine--tRNA ligase	SERYL-tRNA synthetase SerS (serine--tRNA ligase) (SERRS) (serine translase)	information pathways	COG0172J
Rv3838c	pheA	33601.2	Cytoplasmic	prephenate dehydratase	Prephenate dehydratase PheA	intermediary metabolism and respiration	COG0077E
Rv3841	bfrB	20442	Cytoplasmic	bacterioferritin BfrB	Bacterioferritin BfrB	intermediary metabolism and respiration	COG1528P
Rv3846	sodA	23001.9	Extracellular	superoxide dismutase	Superoxide dismutase [FE] SodA	virulence, detoxification, adaptation	COG0605P
Rv3849	espR	14708.7	Cytoplasmic	ESX-1 transcriptional regulator EspR	ESX-1 transcriptional regulatory protein EspR	Regulatory proteins	-
Rv3850	-	23810.9	Cytoplasmic	hypothetical protein Rv3850	Conserved protein	conserved	-

						hypotheticals	
Rv3858c	gltD	53451.8	Cytoplasmic	glutamate synthase small subunit	Probable NADH-dependent glutamate synthase (small subunit) GltD (L-glutamate synthase) (L-glutamate synthetase) (NADH-glutamate synthase) (glutamate synthase (NADH)) (GLTS beta chain) (NADPH-GOGAT)	intermediary metabolism and respiration	COG0493ER
Rv3859c	gltB	165943	Cytoplasmic Membrane	glutamate synthase large subunit	Probable ferredoxin-dependent glutamate synthase [NADPH] (large subunit) GltB (L-glutamate synthase) (L-glutamate synthetase) (NADH-glutamate synthase) (glutamate synthase (NADH))(NADPH-GOGAT)	intermediary metabolism and respiration	COG0069E
Rv3864	espE	42068.8	Cytoplasmic	ESX-1 secretion-associated protein EspE	ESX-1 secretion-associated protein EspE	cell wall and cell processes	-
Rv3865	espF	10618	Unknown	ESX-1 secretion-associated protein EspF	ESX-1 secretion-associated protein EspF	cell wall and cell processes	-
Rv3868	eccA1	62425.6	Cytoplasmic	ESX-1 secretion system protein EccA1	ESX conserved component EccA1. ESX-1 type VII secretion system protein.	cell wall and cell processes	COG0464O
Rv3870	eccCa1	80913	Cytoplasmic Membrane	ESX-1 secretion system protein EccCa	ESX conserved component EccCa1. ESX-1 type VII secretion system protein. Possible transmembrane protein.	cell wall and cell processes	COG1674D
Rv3871	eccCb1	64560.8	Cytoplasmic	ESX-1 secretion system protein EccCb	ESX conserved component EccCb1. ESX-1 type VII secretion system protein.	cell wall and cell processes	COG1674D
Rv3875	esxA	9903.87	Extracellular	ESAT-6 protein EsxA	6 kDa early secretory antigenic target EsxA (ESAT-6)	cell wall and cell processes	COG4842S
Rv3876	espI	70645.1	Cytoplasmic	ESX-1 secretion-associated protein EspI	ESX-1 secretion-associated protein EspI. Conserved proline and alanine rich protein.	cell wall and cell processes	COG0455D
Rv3888c	-	37691.8	Cytoplasmic	membrane protein	Probable conserved membrane protein	cell wall and cell processes	COG0455D
Rv3894c	eccC2	153700	Cytoplasmic Membrane	ESX-2 type VII secretion system protein EccC	ESX conserved component EccC2. ESX-2 type VII secretion system protein. Possible membrane protein.	cell wall and cell processes	COG1674D
Rv3907c	pcnA	53055.4	Cytoplasmic	poly(A) polymerase PcnA	Probable poly(A) polymerase PcnA (polynucleotide adenylyltransferase) (NTP polymerase) (RNA adenylyating enzyme) (poly(A) polymerase)	information pathways	COG0617J
Rv3909	-	83877.4	Cytoplasmic Membrane	hypothetical protein Rv3909	cell wall	cell wall and cell processes	-
Rv3910	-	123564	Cytoplasmic Membrane	peptidoglycan biosynthesis protein	Probable conserved transmembrane protein	cell wall and cell processes	COG0728R
Rv3913	trxB2	35643	Cytoplasmic	thioredoxin reductase	Probable thioredoxin reductase TrxB2 (TRXR) (TR)	intermediary metabolism and respiration	COG0492O
Rv3914	trxC	12544.5	Cytoplasmic	thioredoxin TrxC	Thioredoxin TrxC (TRX) (MPT46)	intermediary metabolism and respiration	COG3118O
Rv3917c	parB	37018.3	Cytoplasmic	chromosome partitioning protein ParB	Probable chromosome partitioning protein ParB	cell wall and cell processes	COG1475K
Rv3920c	-	20558.6	Cytoplasmic	hypothetical protein Rv3920c	Predicted RNA-binding protein	Regulatory proteins	COG1847R

Rv0013	trpG	24626.9	Cytoplasmic	anthranilate synthase component II	Possible anthranilate synthase component II TrpG (glutamine amidotransferase)	intermediary metabolism and respiration	COG0512EH
Rv0014c	pknB	66509.8	Cytoplasmic Membrane	serine/threonine-protein kinase PknB	Transmembrane serine/threonine-protein kinase B PknB (protein kinase B) (STPK B)	Regulatory proteins	COG2815S
Rv0015c	pknA	45597	Cytoplasmic Membrane	serine/threonine-protein kinase PknA	Transmembrane serine/threonine-protein kinase A PknA (protein kinase A) (STPK A)	Regulatory proteins	COG0515RTKL
Rv0016c	pbpA	51575.6	Cytoplasmic Membrane	penicillin-binding protein PbpA	Probable penicillin-binding protein PbpA	cell wall and cell processes	COG0768M
Rv0018c	pstP	53779.9	Cytoplasmic	phosphoserine/threonine phosphatase PstP	Phosphoserine/threonine phosphatase PstP	Regulatory proteins	COG0631T
Rv0019c	fhaB	17152.1	Cytoplasmic Membrane	FHA domain-containing protein FhaB	Conserved protein with FHA domain, FhaB	Regulatory proteins	COG1716T
Rv0044c	-	28524.3	Unknown	oxidoreductase	Possible oxidoreductase	intermediary metabolism and respiration	COG2141C
Rv0045c	-	32099.1	Cytoplasmic	hydrolase	Possible hydrolase	lipid metabolism	COG0596R
Rv0048c	-	30785	Unknown	membrane protein	Possible membrane protein	cell wall and cell processes	-
Rv0050	ponA1	71118.3	Cytoplasmic Membrane	bifunctional penicillin-insensitive transglycosylase/penicillin-sensitive transpeptidase	Probable bifunctional penicillin-binding protein 1A/1B PonA1 (murein polymerase) (PBP1): penicillin-insensitive transglycosylase (peptidoglycan TGASE) + penicillin-sensitive transpeptidase (DD-transpeptidase)	cell wall and cell processes	COG0744M
Rv0053	rpsF	10934.7	Cytoplasmic	30S ribosomal protein S6	30S ribosomal protein S6 RpsF	information pathways	COG0360J
Rv0055	rpsR1	9543.21	Cytoplasmic	30S ribosomal protein S18	30S ribosomal protein S18-1 RpsR1	information pathways	COG0238J
Rv0068	-	32398.9	Cytoplasmic	oxidoreductase	Probable oxidoreductase	intermediary metabolism and respiration	COG1028IQR
Rv0072	-	36410.1	Cytoplasmic Membrane	glutamine ABC transporter permease	Probable glutamine-transport transmembrane protein ABC transporter	cell wall and cell processes	COG4591M
Rv0073	-	35768.9	Cytoplasmic Membrane	glutamine ABC transporter ATP-binding protein	Probable glutamine-transport ATP-binding protein ABC transporter	cell wall and cell processes	COG1136V
Rv0075	-	42766	Cytoplasmic	aminotransferase	Probable aminotransferase	intermediary metabolism and respiration	COG1168E
Rv0077c	-	29686.3	Cytoplasmic	oxidoreductase	Probable oxidoreductase	intermediary metabolism and respiration	COG0596R
Rv0087	hycE	53457.7	Cytoplasmic	formate hydrogenase HycE	Possible formate hydrogenase HycE (FHL)	intermediary metabolism and respiration	COG3261C
Rv0101	nrp	269377	Cytoplasmic	peptide synthetase Nrp	Probable peptide synthetase Nrp (peptide synthase)	lipid metabolism	COG1020Q
Rv0107c	ctpI	169575	Cytoplasmic Membrane	cation-transporter ATPase I	Probable cation-transporter ATPase I CtpI	cell wall and cell processes	COG0474P

Rv0129c	fbpC	36771.3	Extracellular	diacylglycerol acyltransferase/mycolyltransferase Ag85C	Secreted antigen 85-C FbpC (85C) (antigen 85 complex C) (AG58C) (mycolyl transferase 85C) (fibronectin-binding protein C)	lipid metabolism	COG0627R
Rv0139	-	36976.2	Cytoplasmic Membrane	oxidoreductase	Possible oxidoreductase	intermediary metabolism and respiration	COG0451MG
Rv0144	-	30610.9	Cytoplasmic	transcriptional regulator	Probable transcriptional regulatory protein (possibly TetR-family)	Regulatory proteins	COG1309K
Rv0158	-	24128.1	Cytoplasmic Membrane	transcriptional regulator	Probable transcriptional regulatory protein (possibly TetR-family)	Regulatory proteins	COG1309K
Rv0164	TB18.5	17738.3	Cytoplasmic	hypothetical protein Rv0164	Cyclase/Dehydrase	Intermediary metabolism and respiration	COG2867I
Rv0166	fadD5	59887.8	Cytoplasmic	fatty-acid--CoA ligase FadD5	Probable fatty-acid-CoA ligase FadD5 (fatty-acid-CoA synthetase) (fatty-acid-CoA synthase)	lipid metabolism	COG0318IQ
Rv0176	-	35405.1	Cytoplasmic Membrane	Mce associated transmembrane protein	Probable conserved Mce associated transmembrane protein	cell wall and cell processes	COG1714S
Rv0198c	zmp1	73881.6	Cytoplasmic	zinc metalloprotease	Probable zinc metalloprotease Zmp1	intermediary metabolism and respiration	COG3590O
Rv0208c	-	28546.9	Cytoplasmic	tRNA (guanine-N(7)-)-methyltransferase	Hypothetical methyltransferase (methylase)	intermediary metabolism and respiration	COG0220R
Rv0216	-	35756.5	Unknown	hydratase	Double hotdog hydratase	intermediary metabolism and respiration	COG2030I
Rv0217c	lipW	32184.8	Cytoplasmic	esterase LipW	Possible esterase LipW	intermediary metabolism and respiration	COG0657I
Rv0221	-	51829.6	Cytoplasmic	diacylglycerol O-acyltransferase	Possible triacylglycerol synthase (diacylglycerol acyltransferase)	lipid metabolism	COG4908R
Rv0237	lpqI	39518.9	Cytoplasmic	lipoprotein LpqI	Probable conserved lipoprotein LpqI	cell wall and cell processes	COG1472G
Rv0247c	-	28625.9	Cytoplasmic	succinate dehydrogenase iron-sulfur subunit	Probable succinate dehydrogenase [iron-sulfur subunit] (succinic dehydrogenase)	intermediary metabolism and respiration	COG0479C
Rv0251c	hsp	17785.9	Cellwall	heat shock protein	Heat shock protein Hsp (heat-stress-induced ribosome-binding protein A)	virulence, detoxification, adaptation	COG0071O
Rv0265c	-	35252.3	Cytoplasmic Membrane	iron ABC transporter substrate-binding lipoprotein	Probable periplasmic iron-transport lipoprotein	cell wall and cell processes	COG0614P
Rv0290	eccD3	47944	Cytoplasmic Membrane	ESX-3 secretion system protein EccD	ESX conserved component EccD3. ESX-3 type VII secretion system protein. Probable transmembrane protein.	cell wall and cell processes	-
Rv0291	mycP	46127.1	Cytoplasmic	membrane-anchored mycosin MycP	Probable membrane-anchored mycosin MycP3 (serine	intermediary	COG1404O

	3		Membrane		protease) (subtilisin-like protease) (subtilase-like) (mycosin-3)	metabolism and respiration	
Rv0292	eccE3	35932.1	Cytoplasmic Membrane	ESX-3 secretion system protein EccE	ESX conserved component EccE3. ESX-3 type VII secretion system protein. Probable transmembrane protein.	cell wall and cell processes	-
Rv0309	-	22528.8	Unknown	hypothetical protein Rv0309	Possible conserved exported protein	cell wall and cell processes	COG3786S
Rv0311	-	44610.9	Cytoplasmic	hypothetical protein Rv0311	Unknown protein	conserved hypotheticals	-
Rv0341	iniB	43921.2	Cytoplasmic Membrane	isoniazid inducible protein IniB	Isoniazid inductible gene protein IniB	cell wall and cell processes	-
Rv0347	-	36565	Cytoplasmic	membrane protein	Probable conserved membrane protein	cell wall and cell processes	-
Rv0351	grpE	24500.6	Unknown	protein GrpE	Probable GrpE protein (HSP-70 cofactor)	virulence, detoxification, adaptation	COG0576O
Rv0360c	-	15297.3	Unknown	hypothetical protein Rv0360c	Conserved protein	conserved hypotheticals	-
Rv0409	ackA	41318.2	Cytoplasmic	acetate kinase	Probable acetate kinase AckA (acetokinase)	intermediary metabolism and respiration	COG0282C
Rv0418	lpqL	52042.4	Extracellular	lipoprotein aminopeptidase LpqL	Probable lipoprotein aminopeptidase LpqL	cell wall and cell processes	COG2234R
Rv0427c	xthA	32108.3	Cytoplasmic	exodeoxyribonuclease III protein XthA	Probable exodeoxyribonuclease III protein XthA (exonuclease III) (EXO III) (AP endonuclease VI)	information pathways	COG0708L
Rv0431	-	16905.2	Cytoplasmic Membrane	tuberculin-like peptide	Putative tuberculin related peptide	cell wall and cell processes	-
Rv0437c	psd	24227.9	Cytoplasmic Membrane	phosphatidylserine decarboxylase	Possible phosphatidylserine decarboxylase Psd (PS decarboxylase)	lipid metabolism	COG0688I
Rv0457c	-	74464.1	Cytoplasmic	peptidase	Probable peptidase	intermediary metabolism and respiration	COG1505E
Rv0467	icl1	47086.6	Cytoplasmic	isocitrate lyase	Isocitrate lyase Icl (isocitrase) (isocitratase)	intermediary metabolism and respiration	COG2224C
Rv0472c	-	26385.5	Cytoplasmic	HTH-type transcriptional regulator	Probable transcriptional regulatory protein (possibly TetR-family)	Regulatory proteins	COG1309K
Rv0490	senX3	44792.7	Cytoplasmic Membrane	two component sensor histidine kinase SenX3	Putative two component sensor histidine kinase SenX3	Regulatory proteins	COG5002T
Rv0496	-	34843.5	Cytoplasmic	hypothetical protein Rv0496	Exopolyphosphatase	cell wall and cell processes	-
Rv0501	galE2	41032.1	Cytoplasmic	UDP-glucose 4-epimerase GalE	Possible UDP-glucose 4-epimerase GalE2 (galactowaldenase) (UDP-galactose 4-epimerase) (uridine diphosphate galactose 4-epimerase) (uridine diphospho-galactose 4-epimerase)	intermediary metabolism and respiration	COG0451MG

Rv0504c	-	18359.9	Cytoplasmic	hypothetical protein Rv0504c	UPF0336 protein	cell wall and cell processes	COG2030I
Rv0551c	fadD8	61065.9	Cytoplasmic	fatty-acid--CoA ligase FadD8	Probable fatty-acid-CoA ligase FadD8 (fatty-acid-CoA synthetase) (fatty-acid-CoA synthase)	lipid metabolism	COG0318IQ
Rv0568	cyp135B1	50687.5	Cytoplasmic	cytochrome P450 Cyp135B1	Possible cytochrome P450 135B1 Cyp135B1	intermediary metabolism and respiration	COG2124Q
Rv0570	nrdZ	74489	Cytoplasmic	vitamin B12-dependent ribonucleoside-diphosphate reductase	Probable ribonucleoside-diphosphate reductase (large subunit) NrdZ (ribonucleotide reductase)	information pathways	COG0209F
Rv0571c	-	46655.3	Cytoplasmic	hypothetical protein Rv0571c	Predicted phosphoribosyltransferases	Intermediary metabolism and respiration	COG1926R
Rv0638	secE1	16924.3	Cytoplasmic Membrane	preprotein translocase SecE	Probable preprotein translocase SecE1	cell wall and cell processes	COG0690U
Rv0654	-	54943.8	Cytoplasmic	carotenoid cleavage oxygenase	Probable dioxygenase	intermediary metabolism and respiration	COG3670Q
Rv0670	end	26813.1	Cytoplasmic	endonuclease IV	Probable endonuclease IV End (endodeoxyribonuclease IV) (apurinase)	information pathways	COG0648L
Rv0676c	mmpL5	104785	Cytoplasmic Membrane	transmembrane transport protein MmpL5	Probable conserved transmembrane transport protein MmpL5	cell wall and cell processes	COG2409R
Rv0681	-	21230	Cytoplasmic	transcriptional regulator	Probable transcriptional regulatory protein (possibly TetR-family)	Regulatory proteins	COG1309K
Rv0703	rplW	10957.8	Cytoplasmic	50S ribosomal protein L23	50S ribosomal protein L23 RplW	information pathways	COG0089J
Rv0714	rplN	13395.6	Cytoplasmic	50S ribosomal protein L14	50S ribosomal protein L14 RplN	information pathways	COG0093J
Rv0718	rpsH	14411.6	Cytoplasmic	30S ribosomal protein S8	30S ribosomal protein S8 RpsH	information pathways	COG0096J
Rv0723	rplO	15501	Unknown	50S ribosomal protein L15	50S ribosomal protein L15 RplO	information pathways	COG0200J
Rv0734	mapA	27276.9	Cytoplasmic	methionine aminopeptidase	Methionine aminopeptidase MapA (map) (peptidase M) (MetAP)	intermediary metabolism and respiration	COG0024J
Rv0791c	-	37714.8	Cytoplasmic	hypothetical protein Rv0791c	Coenzyme F420-dependent N5,N10-methylene tetrahydromethanopterin reductase and related flavin-dependent oxidoreductases	Intermediary metabolism and respiration	COG2141C
Rv0800	pepC	46013	Cytoplasmic	M18 family aminopeptidase	Probable aminopeptidase PepC	intermediary metabolism and respiration	COG1362E
Rv0814c	-	10165.3	Unknown	hypothetical protein Rv0814c	Conserved protein SseC2	intermediary metabolism and respiration	-
Rv0815c	cysA2	31014.6	Cytoplasmic	thiosulfate sulfurtransferase CysA	Probable thiosulfate sulfurtransferase CysA2 (rhodanese-like protein) (thiosulfate cyanide transsulfurase) (thiosulfate thiotransferase)	intermediary metabolism and respiration	COG2897P
Rv0819	mshD	33567.1	Cytoplasmic Membrane	mycothiol acetyltransferase	GCN5-related N-acetyltransferase, MshD	intermediary metabolism and respiration	COG0456R

Rv0825c	-	23409.8	Unknown	hypothetical protein Rv0825c	Transcriptional regulator	Regulatory proteins	COG1309K
Rv0854	-	16345.6	Cytoplasmic	hypothetical protein Rv0854	Cyclase/Dehydrase	Intermediary metabolism and respiration	-
Rv0858c	dapC	42209.1	Cytoplasmic	N-succinyldiaminopimelate aminotransferase DapC	Probable N-succinyldiaminopimelate aminotransferase DapC (DAP-at)	intermediary metabolism and respiration	COG0436E
Rv0885	-	39798	Cytoplasmic Membrane	hypothetical protein Rv0885	Conserved hypothetical protein	conserved hypotheticals	-
Rv0888	-	52034.2	Unknown	hypothetical protein Rv0888	Probable exported protein	cell wall and cell processes	COG3021S
Rv0910	-	15754.5	Unknown	toxin	Conserved hypothetical protein	conserved hypotheticals	COG3427S
Rv0911	-	27594.7	Cytoplasmic	hypothetical protein Rv0911	Predicted enzyme related to lactoylglutathione lyase	Intermediary metabolism and respiration	COG3324R
Rv0939	-	69353.6	Cytoplasmic	bifunctional 2-hydroxyhepta-2,4-diene-1,7-dioate isomerase/cyclase/dehydrase	Possible bifunctional enzyme: 2-hydroxyhepta-2,4-diene-1,7-dioate isomerase (HHDD isomerase) + cyclase/dehydrase	intermediary metabolism and respiration	COG0179Q
Rv0940c	-	31756.5	Cytoplasmic	oxidoreductase	Possible oxidoreductase	intermediary metabolism and respiration	COG2141C
Rv0972c	fadE1 ₂	41483.3	Cytoplasmic	acyl-CoA dehydrogenase FadE12	Acyl-CoA dehydrogenase FadE12	lipid metabolism	COG1960I
Rv0983	pepD	46452.6	Unknown	serine protease PepD	Probable serine protease PepD (serine proteinase) (MTB32B)	intermediary metabolism and respiration	COG0265O
Rv1013	pks16	58449.3	Cytoplasmic Membrane	polyketide synthase	Putative polyketide synthase Pks16	lipid metabolism	COG0318IQ
Rv1020	mfd	132908	Cytoplasmic	transcription-repair coupling factor	Probable transcription-repair coupling factor Mfd (TRCF)	information pathways	COG1197LK
Rv1021	-	35413.2	Cytoplasmic	nucleoside triphosphate pyrophosphohydrolase	Predicted pyrophosphatase	Intermediary metabolism and respiration	COG3956R
Rv1037c	esxI	9832.78	Unknown	ESAT-6 like protein EsxI	Putative ESAT-6 like protein EsxI (ESAT-6 like protein 1)	cell wall and cell processes	-
Rv1082	mca	32699.7	Cytoplasmic	mycothiol S-conjugate amidase	Mycothiol conjugate amidase Mca (mycothiol S-conjugate amidase)	virulence, detoxification, adaptation	COG2120S
Rv1092c	coaA	35656.9	Cytoplasmic	pantothenate kinase	Probable pantothenate kinase CoaA (pantothenic acid kinase)	intermediary metabolism and respiration	COG1072H
Rv1108c	xseA	44596.3	Cytoplasmic	exodeoxyribonuclease VII large subunit	Probable exodeoxyribonuclease VII (large subunit) XseA (exonuclease VII large subunit)	information pathways	COG1570L
Rv1110	lytB2	36297.1	Cytoplasmic	4-hydroxy-3-methylbut-2-enyl	Probable LYTB-related protein LytB2	cell wall and cell	COG0761IM

				diphosphate reductase		processes	
Rv1178	-	37702.2	Unknown	aminotransferase	Probable aminotransferase	intermediary metabolism and respiration	COG0436E
Rv1187	rocA	58841.4	Cytoplasmic	pyrroline-5-carboxylate dehydrogenase RocA	Probable pyrroline-5-carboxylate dehydrogenase RocA	intermediary metabolism and respiration	COG1012C
Rv1188	-	36237.1	Cytoplasmic	proline dehydrogenase	Probable proline dehydrogenase	intermediary metabolism and respiration	COG0506E
Rv1193	fadD36	49488.3	Cytoplasmic	fatty-acid--CoA ligase FadD36	Probable fatty-acid-CoA ligase FadD36 (fatty-acid-CoA synthetase) (fatty-acid-CoA synthase)	lipid metabolism	COG0318IQ
Rv1202	dapE	37241	Cytoplasmic	succinyl-diaminopimelate desuccinylase DapE	Probable succinyl-diaminopimelate desuccinylase DapE	intermediary metabolism and respiration	COG0624E
Rv1261c	-	16756	Cytoplasmic Membrane	hypothetical protein Rv1261c	Conserved protein	conserved hypotheticals	-
Rv1263	amiB2	49048.9	Cytoplasmic	amidase AmiB2	Probable amidase AmiB2 (aminohydrolase)	intermediary metabolism and respiration	COG0154J
Rv1286	-	67838.9	Cytoplasmic	bifunctional sulfate adenylyltransferase subunit 1 CysN/adenylylsulfate kinase CysC	Probable bifunctional enzyme CysN/CysC: sulfate adenylyltransferase (subunit 1) + adenylylsulfate kinase	intermediary metabolism and respiration	COG2895P
Rv1293	lysA	47425.9	Cytoplasmic	diaminopimelate decarboxylase	Diaminopimelate decarboxylase LysA (DAP decarboxylase)	intermediary metabolism and respiration	COG0019E
Rv1299	prfA	39036.5	Cytoplasmic	peptide chain release factor PrfA	Probable peptide chain release factor 1 PrfA (RF-1)	information pathways	COG0216J
Rv1311	atpC	13134.8	Cytoplasmic	ATP synthase subunit epsilon	Probable ATP synthase epsilon chain AtpC	intermediary metabolism and respiration	COG0355C
Rv1326c	glgB	81729.5	Cytoplasmic	1,4-alpha-glucan branching protein	1,4-alpha-glucan branching enzyme GlgB (glycogen branching enzyme)	intermediary metabolism and respiration	COG0296G
Rv1361c	PPE19	39983.5	Cytoplasmic Membrane	PPE family protein PPE19	PPE family protein PPE19	PE/PPE	COG5651N
Rv1363c	-	28308.3	Cytoplasmic Membrane	membrane protein	Possible membrane protein	cell wall and cell processes	-
Rv1367c	-	41325.4	Cytoplasmic Membrane	hypothetical protein Rv1367c	Beta-lactamase class C and other penicillin binding proteins	cell wall and cell processes	COG1680V
Rv1383	carA	39748.7	Unknown	carbamoyl-phosphate synthase small subunit	Probable carbamoyl-phosphate synthase small chain CarA (carbamoyl-phosphate synthetase glutamine chain)	intermediary metabolism and respiration	COG0505EF
Rv1387	PPE20	55403.6	Cytoplasmic	PPE family protein PPE20	PPE family protein PPE20	PE/PPE	COG5651N
Rv1400c	lipI	34052.7	Cytoplasmic	lipase	Probable lipase LipH	intermediary metabolism and	COG0657I

						respiration	
Rv1404	-	17521.1	Cytoplasmic	transcriptional regulator	Probable transcriptional regulatory protein	Regulatory proteins	COG1846K
Rv1407	fmu	48493.4	Cytoplasmic	methyltransferase	Probable Fmu protein (sun protein)	information pathways	COG0144J
Rv1411c	lprG	24547.8	Cytoplasmic Membrane	lipoprotein LprG	Conserved lipoprotein LprG	cell wall and cell processes	-
Rv1412	ribC	21320.2	Cytoplasmic	riboflavin synthase	Probable riboflavin synthase alpha chain RibC (RibE)	intermediary metabolism and respiration	COG0307H
Rv1421	-	32912.5	Cytoplasmic	hypothetical protein Rv1421	Predicted P-loop-containing kinase	Intermediary metabolism and respiration	COG1660R
Rv1429	-	46866.7	Cytoplasmic	hypothetical protein Rv1429	Regulator of polyketide synthase expression	Lipid metabolism	COG2508TQ
Rv1446c	opcA	32717.6	Cytoplasmic	OXPP cycle protein OpcA	Putative OXPP cycle protein OpcA	intermediary metabolism and respiration	COG3429G
Rv1458c	-	33726.6	Cytoplasmic Membrane	antibiotic ABC transporter ATP-binding protein	Probable unidentified antibiotic-transport ATP-binding protein ABC transporter	cell wall and cell processes	COG1131V
Rv1461	-	94171.4	Cytoplasmic	hypothetical protein Rv1461	ABC-type transport system involved in Fe-S cluster assembly, permease component	cell wall and cell processes	COG0719O
Rv1481	-	36029.9	Unknown	membrane protein	Probable membrane protein	cell wall and cell processes	COG1240H
Rv1492	mutA	64712.3	Cytoplasmic	methylmalonyl-CoA mutase small subunit	Probable methylmalonyl-CoA mutase small subunit MutA (MCM)	lipid metabolism	COG1884I
Rv1498A	-	7628.64	Cytoplasmic	hypothetical protein Rv1498A	Conserved protein	conserved hypotheticals	COG3360S
Rv1508c	-	63856.9	Cytoplasmic Membrane	membrane protein	Probable membrane protein	cell wall and cell processes	-
Rv1512	epiA	34537.2	Cytoplasmic	nucleotide-sugar epimerase EpiA	Probable nucleotide-sugar epimerase EpiA	intermediary metabolism and respiration	COG0451MG
Rv1516c	-	37932.4	Cytoplasmic	sugar transferase	Probable sugar transferase	intermediary metabolism and respiration	COG0463M
Rv1523	-	37915	Unknown	methyltransferase	Probable methyltransferase	intermediary metabolism and respiration	COG2226H
Rv1529	fadD24	62797.4	Cytoplasmic	fatty-acid--CoA ligase FadD24	Probable fatty-acid-AMP ligase FadD24 (fatty-acid-AMP synthetase) (fatty-acid-AMP synthase)	lipid metabolism	COG0318IQ
Rv1531	-	20795.7	Cytoplasmic Membrane	hypothetical protein Rv1531	Conserved protein	conserved hypotheticals	COG2128S
Rv1533	-	39620	Cytoplasmic	monooxygenase	Dioxygenases related to 2-nitropropane dioxygenase	Intermediary metabolism and respiration	COG2070R
Rv1547	dnaE1	129323	Cytoplasmic	DNA polymerase III subunit alpha	Probable DNA polymerase III (alpha chain) DnaE1	information pathways	COG0587L

					(DNA nucleotidyltransferase)		
Rv1556	-	22329.5	Cytoplasmic	HTH-type transcriptional regulator	Possible regulatory protein	Regulatory proteins	COG1309K
Rv1592c	-	48043.1	Cytoplasmic	hypothetical protein Rv1592c	Secretory lipase	cell wall and cell processes	COG0657I
Rv1598c	-	14892.4	Cytoplasmic	hypothetical protein Rv1598c	Conserved protein	conserved hypotheticals	-
Rv1599	hisD	45346.2	Unknown	histidinol dehydrogenase	Probable histidinol dehydrogenase HisD (HDH)	intermediary metabolism and respiration	COG0141E
Rv1605	hisF	27220.9	Cytoplasmic	imidazole glycerol phosphate synthase subunit HisF	Probable cyclase HisF	intermediary metabolism and respiration	COG0107E
Rv1608c	bcpB	16894.5	Cytoplasmic	peroxidoxin	Probable peroxidoxin BcpB	virulence, detoxification, adaptation	COG1225O
Rv1620c	cydC	59496.9	Cytoplasmic Membrane	cytochrome biosynthesis ATP-binding protein/permease CydC	Probable 'component linked with the assembly of cytochrome' transport transmembrane ATP-binding protein ABC transporter CydC	intermediary metabolism and respiration	COG4987CO
Rv1633	uvrB	78038.3	Cytoplasmic	excinuclease ABC subunit UvrB	Probable excinuclease ABC (subunit B - helicase) UvrB	information pathways	-
Rv1637c	-	27936.3	Unknown	hypothetical protein Rv1637c	Zn-dependent hydrolases, including glyoxylases	Intermediary metabolism and respiration	COG0491R
Rv1640c	lysX	128208	Cytoplasmic	bifunctional lysine--tRNA ligase/phosphatidylglycerol lysyltransferase	Lysyl-tRNA synthetase 2 LysX	information pathways	COG1190J
Rv1661	pks7	221046	Cytoplasmic Membrane	polyketide synthase	Probable polyketide synthase Pks7	lipid metabolism	COG3321Q
Rv1676	-	25859.2	Cytoplasmic	hypothetical protein Rv1676	Antioxidant protein	virulence, detoxification, adaptation	COG0450O
Rv1703c	-	21537.5		methyltransferase	Probable catechol-O-methyltransferase	intermediary metabolism and respiration	-
Rv1710	scpB	25244.6	Cytoplasmic	segregation and condensation protein ScpB	Possible segregation and condensation protein ScpB	cell wall and cell processes	COG1386K
Rv1736c	narX	72793.3	Cytoplasmic Membrane	nitrate reductase-like protein NarX	Probable nitrate reductase NarX	intermediary metabolism and respiration	COG5013C
Rv1738	-	10606.1	Cytoplasmic	hypothetical protein Rv1738	Conserved protein	conserved hypotheticals	-
Rv1747	-	92121.2	Cytoplasmic Membrane	ABC transporter ATP-binding protein/permease	Probable conserved transmembrane ATP-binding protein ABC transporter	cell wall and cell processes	COG1131V
Rv1779c	-	63511.7	Cytoplasmic Membrane	integral membrane protein	Possible integral membrane protein	cell wall and cell processes	-
Rv1780	-	20456.5	Cytoplasmic	hypothetical protein Rv1780	Conserved protein	conserved	-

						hypotheticals	
Rv1793	esxN	9941.95	Cytoplasmic	ESAT-6 like protein EsxN	Putative ESAT-6 like protein EsxN (ESAT-6 like protein 5)	cell wall and cell processes	-
Rv1797	eccE5	44178.1	Cytoplasmic Membrane	ESX-5 type VII secretion system protein EccE	ESX conserved component EccE5. ESX-5 type VII secretion system protein. Probable membrane protein.	cell wall and cell processes	-
Rv1812c	-	41938	Cytoplasmic Membrane	dehydrogenase	Probable dehydrogenase	intermediary metabolism and respiration	COG1252C
Rv1819c	bacA	71299.1	Cytoplasmic Membrane	vitamin B12 transport ATP-binding protein BacA	Probable drug-transport transmembrane ATP-binding protein ABC transporter BacA	cell wall and cell processes	COG4178R
Rv1842c	-	48092	Cytoplasmic Membrane	hypothetical protein Rv1842c	Conserved hypothetical membrane protein	cell wall and cell processes	COG1253R
Rv1860	apa	32720.8	Extracellular	alanine/proline-rich secreted protein Apa	Alanine and proline rich secreted protein Apa (fibronectin attachment protein) (immunogenic protein MPT32) (antigen MPT-32) (45-kDa glycoprotein) (45/47 kDa antigen)	cell wall and cell processes	COG5373S
Rv1866	-	83182.5	Cytoplasmic	hypothetical protein Rv1866	Conserved protein	lipid metabolism	-
Rv1874	-	24899.2	Cytoplasmic Membrane	hypothetical protein Rv1874	link to sulfotransferase activity	Intermediary metabolism and respiration	-
Rv1875	-	15979.1	Unknown	hypothetical protein Rv1875	Pyridoxamine protein	Intermediary metabolism and respiration	-
Rv1878	glnA3	46720.1	Cytoplasmic	glutamine synthetase GlnA	Probable glutamine synthetase GlnA3 (glutamine synthase) (GS-I)	intermediary metabolism and respiration	COG0174E
Rv1896c	-	33267.4	Cytoplasmic	S-adenosyl-L-methionine-dependent methyltransferase	O-Methyltransferase involved in polyketide biosynthesis	Lipid metabolism	COG3315Q
Rv1900c	lipJ	49684.4	Cytoplasmic	lignin peroxidase LipJ	Probable lignin peroxidase LipJ	intermediary metabolism and respiration	COG2114T
Rv1914c	-	14719.3	Cytoplasmic Membrane	hypothetical protein Rv1914c	Unknown protein	conserved hypotheticals	-
Rv1978	-	30584.1	Cytoplasmic Membrane	hypothetical protein Rv1978	SAM-dependent methyltransferases	Intermediary metabolism and respiration	COG2226H
Rv2000	-	59894.8	Cytoplasmic	hypothetical protein Rv2000	Unknown protein	conserved hypotheticals	COG1249C
Rv2028c	-	29441.6	Cytoplasmic Membrane	universal stress protein	Universal stress protein family protein	virulence, detoxification, adaptation	COG0589T
Rv2029c	pfkB	35401.4	Cytoplasmic Membrane	6-phosphofructokinase PfkB	6-phosphofructokinase PfkB (phosphohexokinase) (phosphofructokinase)	intermediary metabolism and respiration	COG1105G
Rv2037c	-	34656.5	Cytoplasmic	transmembrane protein	Conserved transmembrane protein	cell wall and cell	COG1752R

			Membrane				processes	
Rv2042c	-	29091.2	Unknown	hypothetical protein Rv2042c	Nuclear transport factor 2	Regulatory proteins	-	
Rv2062c	cobN	129178	Cytoplasmic	cobalamin biosynthesis protein CobN	Cobalamin biosynthesis protein CobN	intermediary metabolism and respiration	COG1429H	
Rv2068c	blaC	32567.9	Extracellular	beta-lactamase	Class a beta-lactamase BlaC	intermediary metabolism and respiration	COG2367V	
Rv2095c	pafC	33764.4	Cytoplasmic	proteasome accessory factor C	Proteasome accessory factor C PafC	intermediary metabolism and respiration	COG2378K	
Rv2096c	pafB	35300	Cytoplasmic	proteasome accessory factor B	Proteasome accessory factor B PafB	intermediary metabolism and respiration	COG2378K	
Rv2124c	metH	130291	Cytoplasmic	methionine synthase	5-methyltetrahydrofolate--homocystein methyltransferase MetH (methionine synthase, vitamin-B12 dependent isozyme) (ms)	intermediary metabolism and respiration	COG1410E	
Rv2130c	mshC	45594.5	Cytoplasmic	cysteine:1D-myo-inosityl 2-amino-2-deoxy--D-glucopyranoside ligase	Cysteine:1D-myo-inosityl 2-amino-2-deoxy--D-glucopyranoside ligase MshC	intermediary metabolism and respiration	COG0215J	
Rv2153c	murG	41828.1	Cytoplasmic Membrane	UDP-N-acetylglucosamine--N-acetylmuramyl-(pentapeptide) pyrophosphoryl-undecaprenol-N-acetylglucosamine transferase	Probable UPD-N-acetylglucosamine-N-acetylmuramyl-(pentapeptide) pyrophosphoryl-undecaprenol-N-acetylglucosamine transferase MurG	cell wall and cell processes	COG0707M	
Rv2165c	-	42497.2	Cytoplasmic	rRNA small subunit methyltransferase H	Predicted S-adenosylmethionine-dependent methyltransferase involved in cell envelope biogenesis	cell wall and cell processes	-	
Rv2182c	-	26955.1	Unknown	1-acylglycerol-3-phosphate O-acyltransferase	1-acyl-sn-glycerol-3-phosphate acyltransferase	Intermediary metabolism and respiration	COG0204I	
Rv2201	asnB	72117.5	Cytoplasmic	asparagine synthetase	Probable asparagine synthetase AsnB	intermediary metabolism and respiration	COG0367E	
Rv2221c	glnE	109106	Cytoplasmic	glutamate-ammonia-ligase adenylyltransferase	Glutamate-ammonia-ligase adenylyltransferase GlnE (glutamine-synthetase adenylyltransferase)	intermediary metabolism and respiration	COG1391OT	
Rv2229c	-	26850.9	Cytoplasmic	hypothetical protein Rv2229c	Zn-ribbon protein, possibly nucleic acid-binding	Regulatory proteins	COG1579R	
Rv2263	-	33965.9	Cytoplasmic	oxidoreductase	Possible oxidoreductase	intermediary metabolism and respiration	COG1028IQR	
Rv2284	lipM	46681.4	Cytoplasmic Membrane	esterase LipM	Probable esterase LipM	intermediary metabolism and respiration	COG0657I	
Rv2285	-	47675	Cytoplasmic Membrane	diacylglycerol acyltransferase	Possible triacylglycerol synthase (diacylglycerol acyltransferase)	lipid metabolism	COG4908R	
Rv2301	cut2	23894.1	Unknown	cutinase	Probable cutinase Cut2	cell wall and cell	-	

						processes	
Rv2326c	-	72835.5	Cytoplasmic Membrane	ABC transporter ATP-binding protein	Possible transmembrane ATP-binding protein ABC transporter	cell wall and cell processes	COG1123R
Rv2334	cysK1	32752.7	Cytoplasmic	O-acetylserine sulfhydrylase	Cysteine synthase a CysK1 (O-acetylserine sulfhydrylase A) (O-acetylserine (thiol)-lyase A) (CSASE A)	intermediary metabolism and respiration	COG0031E
Rv2338c	moeW	35271.3	Cytoplasmic	molybdopterin biosynthesis protein MoeW	Possible molybdopterin biosynthesis protein MoeW	intermediary metabolism and respiration	COG0476H
Rv2343c	dnaG	69593	Cytoplasmic	DNA primase	Probable DNA primase DnaG	information pathways	COG0358L
Rv2346c	esxO	9954.01	Cytoplasmic	ESAT-6 like protein EsxO	Putative ESAT-6 like protein EsxO (ESAT-6 like protein 6)	cell wall and cell processes	-
Rv2361c	-	33759.3	Cytoplasmic	decaprenyl diphosphate synthase	Long (C50) chain Z-isoprenyl diphosphate synthase (Z-decaprenyl diphosphate synthase)	cell wall and cell processes	COG0020I
Rv2374c	hrcA	36474.5	Cytoplasmic	heat-inducible transcription repressor HrcA	Probable heat shock protein transcriptional repressor HrcA	virulence, detoxification, adaptation	COG1420K
Rv2394	ggtB	66527.5	Extracellular	gamma-glutamyltranspeptidase precursor GgtB	Probable gamma-glutamyltranspeptidase precursor GgtB (gamma-glutamyltransferase) (glutamyl transpeptidase)	intermediary metabolism and respiration	COG0405E
Rv2409c	-	31028.8	Cytoplasmic	hypothetical protein Rv2409c	Transglutaminase-like enzymes, putative cysteine proteases	Intermediary metabolism and respiration	COG1305E
Rv2413c	-	33113.7	Cellwall	hypothetical protein Rv2413c	DNA polymerase III, delta subunit	Information pathways	-
Rv2429	ahpD	18780.6	Unknown	alkyl hydroperoxide reductase AphD	Alkyl hydroperoxide reductase D protein AhpD (alkyl hydroperoxidase D)	virulence, detoxification, adaptation	COG2128S
Rv2447c	folC	50779.3	Cytoplasmic	folylpolyglutamate synthase FolC	Probable folylpolyglutamate synthase protein FolC (folylpoly-gamma-glutamate synthetase) (FPGS)	intermediary metabolism and respiration	COG0285H
Rv2453c	mobA	21052.4	Cytoplasmic	molybdenum cofactor guanylyltransferase	Probable molybdopterin-guanine dinucleotide biosynthesis protein A MobA	intermediary metabolism and respiration	COG0746H
Rv2468c	-	17288.6	Cytoplasmic	hypothetical protein Rv2468c	Conserved protein	conserved hypotheticals	-
Rv2473	-	25101.3	Unknown	alanine/proline-rich membrane protein	Possible alanine and proline rich membrane protein	cell wall and cell processes	-
Rv2484c	-	52309.1	Cytoplasmic	diacylglycerol O-acyltransferase	Possible triacylglycerol synthase (diacylglycerol acyltransferase)	lipid metabolism	-
Rv2486	echA14	26280.1	Cytoplasmic	enoyl-CoA hydratase EchA14	Probable enoyl-CoA hydratase EchA14 (enoyl hydratase) (unsaturated acyl-CoA hydratase) (crotonase)	lipid metabolism	COG1024I
Rv2497c	bkdA	40615.8	Cytoplasmic	branched-chain keto acid dehydrogenase E1 component subunit alpha	Probable branched-chain keto acid dehydrogenase E1 component, alpha subunit BkdA	intermediary metabolism and respiration	COG1071C
Rv2502c	accD1	56745.7	Cytoplasmic	acetyl-/propionyl-CoA carboxylase	Probable acetyl-/propionyl-CoA carboxylase (beta	lipid metabolism	COG4799I

				subunit beta	subunit) AccD1		
Rv2518c	ldtB	43365.9	Unknown	L,D-transpeptidase LdtB	Probable L,D-transpeptidase LdtB	cell wall and cell processes	COG1376S
Rv2521	bcp	17045.3	Cellwall	peroxiredoxin	Probable bacterioferritin comigratory protein Bcp	virulence, detoxification, adaptation	COG1225O
Rv2533c	nusB	16740.2	Cytoplasmic	N utilization substance protein B	N utilization substance protein NusB (NusB protein)	information pathways	COG0781K
Rv2534c	efp	20375.1	Cytoplasmic	elongation factor P	Probable elongation factor P Efp	information pathways	COG0231J
Rv2540c	aroF	41760.4	Cytoplasmic	chorismate synthase	Probable chorismate synthase AroF (5-enolpyruvylshikimate-3-phosphate phospholyase)	intermediary metabolism and respiration	COG0082E
Rv2567	-	95448.6	Cytoplasmic Membrane	alanine/leucine-rich protein	proteasome	Intermediary metabolism and respiration	COG2308S
Rv2574	-	19001.9	Cytoplasmic Membrane	hypothetical protein Rv2574	Conserved protein	conserved hypotheticals	COG3427S
Rv2579	dhaA	33728.2	Cytoplasmic	haloalkane dehalogenase	Possible haloalkane dehalogenase DhaA (1-chlorohexane halidohydrolase)	intermediary metabolism and respiration	COG0596R
Rv2586c	secF	47013.6	Cytoplasmic Membrane	protein translocase subunit SecF	Probable protein-export membrane protein SecF	cell wall and cell processes	COG0341U
Rv2588c	yajC	12966.6	Cytoplasmic	membrane protein secretion factor YajC	Probable conserved membrane protein secretion factor YajC	cell wall and cell processes	COG1862U
Rv2622	-	29542	Cytoplasmic	methyltransferase	Possible methyltransferase (methylase)	intermediary metabolism and respiration	COG2226H
Rv2625c	-	41480.2	Cytoplasmic Membrane	zinc metalloprotease Rip3	Probable conserved transmembrane alanine and leucine rich protein	cell wall and cell processes	COG1994R
Rv2672	-	54015.7	Unknown	protease	Possible secreted protease	intermediary metabolism and respiration	COG0596R
Rv2673	aftC	48882.7	Cytoplasmic Membrane	alpha-(1-3)-arabinofuranosyltransferase	Possible arabinofuranosyltransferase AftC	cell wall and cell processes	-
Rv2679	echA15	29634.9	Cytoplasmic	enoyl-CoA hydratase EchA15	Probable enoyl-CoA hydratase EchA15 (enoyl hydrase) (unsaturated acyl-CoA hydratase) (crotonase)	lipid metabolism	COG1024I
Rv2682c	dxs1	67885.4	Cytoplasmic	1-deoxy-D-xylulose 5-phosphate synthase	Probable 1-deoxy-D-xylulose 5-phosphate synthase Dxs1 (1-deoxyxylulose-5-phosphate synthase) (DXP synthase) (DXPS)	intermediary metabolism and respiration	COG1154HI
Rv2701c	suhB	29995	Cytoplasmic	inositol-1-monophosphatase SuhB	Inositol-1-monophosphatase SuhB	intermediary metabolism and respiration	COG0483G
Rv2716	-	24574.3	Cytoplasmic	hypothetical protein Rv2716	Predicted epimerase, PhzC/PhzF homolog	Intermediary metabolism and respiration	COG0384R
Rv2718c	nrdR	17177.3	Cytoplasmic	transcriptional regulator NrdR	Probable transcriptional regulatory protein NrdR	Regulatory proteins	COG1327K

Rv2720	lexA	24841.2	Unknown	repressor LexA	Repressor LexA	Regulatory proteins	COG1974KT
Rv2725c	hflX	53327.7	Cytoplasmic	GTP-binding protein HflX	Probable GTP-binding protein HflX	intermediary metabolism and respiration	COG2262R
Rv2732c	-	21989.4	Cytoplasmic Membrane	transmembrane protein	Probable conserved transmembrane protein	cell wall and cell processes	-
Rv2766c	-	27140.1	Cytoplasmic	short-chain type dehydrogenase/reductase	Probable short-chain type dehydrogenase/reductase	intermediary metabolism and respiration	COG1028IQR
Rv2778c	-	17268.5	Cytoplasmic	hypothetical protein Rv2778c	Conserved protein	conserved hypotheticals	-
Rv2790c	ltp1	42904.1	Cytoplasmic	lipid-transfer protein	Probable lipid-transfer protein Ltp1	cell wall and cell processes	-
Rv2837c	-	35383.2	Unknown	bifunctional oligoribonuclease/PAP phosphatase NrnA	Exopolyphosphatase-related proteins	Intermediary metabolism and respiration	COG0618R
Rv2841c	nusA	37641.6	Cytoplasmic	transcription termination/antitermination protein NusA	Probable N utilization substance protein A NusA	information pathways	COG0195K
Rv2847c	cysG	41938.4	Cytoplasmic	multifunctional uroporphyrin-III C-methyltransferase/precorrin-2 oxidase/ferrochelatase	Possible multifunctional enzyme siroheme synthase CysG: uroporphyrin-III C-methyltransferase (urogen III methylase) (SUMT) (uroporphyrinogen III methylase) (UROM) + precorrin-2 oxidase + ferrochelatase	intermediary metabolism and respiration	COG0007H
Rv2861c	mapB	30891	Cytoplasmic	methionine aminopeptidase	Methionine aminopeptidase MapB (map) (peptidase M)	intermediary metabolism and respiration	COG0024J
Rv2900c	fdhF	84566.1	Cytoplasmic	formate dehydrogenase subunit alpha FdhF	Possible formate dehydrogenase H FdhF (formate-hydrogen-lyase-linked, selenocysteine-containing polypeptide) (formate dehydrogenase-H alpha subunit) (FDH-H)	intermediary metabolism and respiration	COG0243C
Rv2909c	rpsP	17437.1	Cytoplasmic	30S ribosomal protein S16	30S ribosomal protein S16 RpsP	information pathways	COG0228J
Rv2925c	rnc	25398.9	Cytoplasmic	ribonuclease III	Probable ribonuclease III Rnc (RNase III)	information pathways	COG0571K
Rv2936	drvA	35818.2	Cytoplasmic Membrane	daunorubicin ABC transporter ATP-binding protein DrrA	Daunorubicin-dim-transport ATP-binding protein ABC transporter DrrA	cell wall and cell processes	COG1131V
Rv2942	mmpL7	95121.9	Cytoplasmic Membrane	transmembrane transport protein MmpL7	Conserved transmembrane transport protein MmpL7	cell wall and cell processes	COG2409R
Rv2948c	fadD22	75198.2	Cytoplasmic Membrane	p-hydroxybenzoyl--AMP ligase	P-hydroxybenzoyl-AMP ligase FadD22	lipid metabolism	COG0318IQ
Rv2949c	-	22587	Cytoplasmic	chorismate pyruvate-lyase	Chorismate pyruvate lyase	intermediary metabolism and respiration	COG3161H
Rv2952	-	30651.6	Cytoplasmic	phthiotriol/phenolphthiotriol dimycocerosates methyltransferase	Possible methyltransferase (methylase)	intermediary metabolism and respiration	COG2226H
Rv2954c	-	27144.1	Cytoplasmic	hypothetical protein Rv2954c	SAM-dependent methyltransferases	Intermediary	COG0030J

						metabolism and respiration	
Rv2955c	-	35918.3	Cytoplasmic	hypothetical protein Rv2955c	SAM-dependent methyltransferases	Intermediary metabolism and respiration	COG4123R
Rv2956	-	26534.2	Cytoplasmic	hypothetical protein Rv2956	SAM-dependent methyltransferases	Intermediary metabolism and respiration	-
Rv2959c	-	27844.5	Cytoplasmic	rhamnosyl O-methyltransferase	Possible methyltransferase (methylase)	intermediary metabolism and respiration	COG3510V
Rv2982c	gpdA2	33992	Cytoplasmic Membrane	glycerol-3-phosphate dehydrogenase	Probable glycerol-3-phosphate dehydrogenase [NAD(P)+] GpdA2 (NAD(P)H- dependent glycerol-3-phosphate dehydrogenase)	lipid metabolism	COG0240C
Rv2985	mutT1	34748.3	Cytoplasmic	8-oxo-dGTP diphosphatase	Possible hydrolase MutT1	information pathways	COG1051F
Rv2993c	-	25341.1	Unknown	2-hydroxyhepta-2,4-diene-1,7-dioate isomerase	Possible 2-hydroxyhepta-2,4-diene-1,7-dioate isomerase (HHDD isomerase)	intermediary metabolism and respiration	-
Rv3013	-	22967.2	Cytoplasmic	hypothetical protein Rv3013	Amino acid-binding ACT protein	Regulatory proteins	COG0317TK
Rv3024c	trmU	38120.1	Cytoplasmic	tRNA-specific 2-thiouridylase	Probable tRNA (5-methylaminomethyl-2-thiouridylate)-methyltransferase TrmU	information pathways	COG0482J
Rv3041c	-	31249.4	Cytoplasmic Membrane	ABC transporter ATP-binding protein	Probable conserved ATP-binding protein ABC transporter	cell wall and cell processes	COG1119P
Rv3043c	ctaD	63654.7	Cytoplasmic Membrane	cytochrome C oxidase cytochrome 1	Probable cytochrome C oxidase polypeptide I CtaD (cytochrome AA3 subunit 1)	intermediary metabolism and respiration	COG0843C
Rv3068c	pgmA	58265.7	Cytoplasmic	phosphoglucomutase PgmA	Probable phosphoglucomutase PgmA (glucose phosphomutase) (PGM)	intermediary metabolism and respiration	COG0033G
Rv3080c	pknK	119418	Cytoplasmic Membrane	serine/threonine-protein kinase PknK	Serine/threonine-protein kinase transcriptional regulatory protein PknK (protein kinase K) (STPK K)	Regulatory proteins	COG2909K
Rv3089	fadD13	54459.5	Cytoplasmic	long chain-fatty-acid--CoA ligase FadD13	Probable chain -fatty-acid-CoA ligase FadD13 (fatty-acyl-CoA synthetase)	lipid metabolism	COG0318IQ
Rv3134c	-	28007.9	Cytoplasmic	universal stress protein	Universal stress protein family protein	virulence, detoxification, adaptation	COG0589T
Rv3137	-	27661.1	Unknown	histidinol-phosphatase	Probable monophosphatase	intermediary metabolism and respiration	COG0483G
Rv3146	nuoB	20186.8	Cytoplasmic Membrane	NADH-quinone oxidoreductase subunit B	Probable NADH dehydrogenase I (chain B) NuoB (NADH-ubiquinone oxidoreductase chain B)	intermediary metabolism and respiration	COG0377C
Rv3150	nuoF	48101.6	Cytoplasmic	NADH-quinone oxidoreductase subunit F	Probable NADH dehydrogenase I (chain F) NuoF (NADH-ubiquinone oxidoreductase chain F)	intermediary metabolism and respiration	COG1894C

Rv3153	nuoI	23434.4	Unknown	NADH-quinone oxidoreductase subunit I	Probable NADH dehydrogenase I (chain I) NuoI (NADH-ubiquinone oxidoreductase chain I)	intermediary metabolism and respiration	COG1143C
Rv3157	nuoM	59175.1	Cytoplasmic Membrane	NADH-quinone oxidoreductase subunit M	Probable NADH dehydrogenase I (chain M) NUOK (NADH-ubiquinone oxidoreductase chain M)	intermediary metabolism and respiration	COG1008C
Rv3170	aofH	48407.6	Cytoplasmic	flavin-containing monoamine oxidase	Probable flavin-containing monoamine oxidase AofH (amine oxidase) (MAO)	intermediary metabolism and respiration	COG1231E
Rv3194c	-	35440.6	Unknown	secreted protein	Possible conserved secreted protein	cell wall and cell processes	COG3480T
Rv3203	lipV	23624.1	Cytoplasmic	lipase LipV	Possible lipase LipV	intermediary metabolism and respiration	COG0596R
Rv3210c	-	25121.4	Cytoplasmic	hypothetical protein Rv3210c	Conserved protein	conserved hypotheticals	COG4445FJ
Rv3212	-	42506.1	Unknown	alanine/valine-rich protein	intermediary metabolism and respiration , Valine rich protein	Intermediary metabolism and respiration	-
Rv3213c	-	28575.3	Cytoplasmic Membrane	SOJ/ParA-like protein	Possible SOJ/para-related protein	cell wall and cell processes	COG1192D
Rv3214	gpm2	21948.8	Cytoplasmic	phosphoglycerate mutase	Possible phosphoglycerate mutase Gpm2 (phosphoglyceromutase) (PGAM) (BPG-dependent PGAM)	intermediary metabolism and respiration	COG0406G
Rv3220c	-	54012.5	Cytoplasmic	two component sensor kinase	Probable two component sensor kinase	Regulatory proteins	COG3920T
Rv3223c	sigH	24225.1	Cytoplasmic	ECF RNA polymerase sigma factor SigH	Alternative RNA polymerase sigma-E factor (sigma-24) SigH (RPOE)	information pathways	COG1595K
Rv3232c	ppk2	34092	Cytoplasmic	polyphosphate kinase	Polyphosphate kinase Ppk2 (polyphosphoric acid kinase)	intermediary metabolism and respiration	COG2326S
Rv3241c	-	24534.7	Cytoplasmic	hypothetical protein Rv3241c	Conserved protein	information pathways	-
Rv3270	ctpC	76495.1	Cytoplasmic Membrane	manganese/zinc-exporting P-type ATPase	Probable metal cation-transporting P-type ATPase C CtpC	cell wall and cell processes	COG2217P
Rv3272	-	42458.6	Cytoplasmic	hypothetical protein Rv3272	Predicted acyl-CoA transferases/carnitine dehydratase	Intermediary metabolism and respiration	COG1804C
Rv3292	-	45332	Cytoplasmic	hypothetical protein Rv3292	Conserved hypothetical protein	conserved hypotheticals	COG5383S
Rv3296	lhr	161348	Cytoplasmic Membrane	ATP-dependent helicase	Probable ATP-dependent helicase Lhr (large helicase-related protein)	information pathways	COG1201R
Rv3299c	atsB	105648	Cytoplasmic	arylsulfatase AtsB	Probable arylsulfatase AtsB (aryl-sulfate sulphohydrolase) (sulfatase)	intermediary metabolism and respiration	COG3119P
Rv3301c	phoY1	24826.5	Cytoplasmic	phosphate transport system transcriptional regulator PhoY	Probable phosphate-transport system transcriptional regulatory protein PhoU homolog 1 PhoY1	cell wall and cell processes	COG0704P

Rv3303c	lpdA	51456.9	Cytoplasmic	NAD(P)H quinone reductase LpdA	NAD(P)H quinone reductase LpdA	intermediary metabolism and respiration	COG1249C
Rv3313c	add	39743.2	Cytoplasmic	adenosine deaminase	Probable adenosine deaminase Add (adenosine aminohydrolase)	intermediary metabolism and respiration	COG1816F
Rv3329	-	47699.2	Cytoplasmic	aminotransferase	Probable aminotransferase	intermediary metabolism and respiration	-
Rv3336c	trpS	36272	Cytoplasmic	tryptophan--tRNA ligase	Probable tryptophanyl-tRNA synthetase TrpS (tryptophan--tRNA ligase) (TRPRS) (tryptophan translase)	information pathways	COG0180J
Rv3342	-	26608.6	Unknown	methyltransferase	Possible methyltransferase (methylase)	intermediary metabolism and respiration	COG2226H
Rv3356c	folD	29451.7	Cytoplasmic	bifunctional methylenetetrahydrofolate dehydrogenase/methenyltetrahydrofolate cyclohydrolase	Probable bifunctional protein FolD: methylenetetrahydrofolate dehydrogenase + methenyltetrahydrofolate cyclohydrolase	intermediary metabolism and respiration	COG0190H
Rv3368c	-	23733	Cytoplasmic	oxidoreductase	Possible oxidoreductase	intermediary metabolism and respiration	COG0778C
Rv3369	-	15717.8	Cytoplasmic	hypothetical protein Rv3369	FMN binding protein	Intermediary metabolism and respiration	COG3871R
Rv3376	-	23645.8	Cytoplasmic	phosphatase	Predicted hydrolase (HAD superfamily)	Intermediary metabolism and respiration	COG1011R
Rv3390	lpqD	25056.4	Cytoplasmic Membrane	lipoprotein LpqD	Probable conserved lipoprotein LpqD	cell wall and cell processes	COG0406G
Rv3404c	-	26483	Cytoplasmic	hypothetical protein Rv3404c	Methionyl-tRNA formyltransferase	Information pathways	COG0223J
Rv3409c	choD	63023.8	Unknown	cholesterol oxidase	Cholesterol oxidase ChoD (cholesterol-O2 oxidoreductase)	lipid metabolism	COG2303E
Rv3441c	mrsA	45873.1	Cytoplasmic	phosphoglucosamine mutase	Probable phospho-sugar mutase / MrsA protein homolog	intermediary metabolism and respiration	COG1109G
Rv3478	PPE60	39381.6	Cytoplasmic Membrane	PE family protein PPE60	PE family protein PPE60	PE/PPE	COG5651N
Rv3490	otsA	55862.5	Cytoplasmic	trehalose-phosphate synthase	Alpha, alpha-trehalose-phosphate synthase [UDP-forming] OtsA (trehalose-6-phosphate synthase) (UDP-glucose-glucosephosphate glucosyltransferase) (trehalosephosphate-UDP glucosyltransferase) (trehalose-6-phosphate synthetase) (trehalose-phosphate synthase) (trehalose-phosphate synthetase) (transglucosylase) (trehalosephosphate-UDP glucosyl transferase)	virulence, detoxification, adaptation	COG0380G

Rv3494c	mce4F	59636.8	Unknown	Mce family protein Mce4	Mce-family protein Mce4F	virulence, detoxification, adaptation	COG1463Q
Rv3499c	mce4A	42418.3	Unknown	Mce family protein Mce4A	Mce-family protein Mce4A	virulence, detoxification, adaptation	COG1463Q
Rv3502c	-	32734.8	Cytoplasmic	short-chain type dehydrogenase/reductase	Probable short-chain type dehydrogenase/reductase. Possible 17-beta-hydroxysteroid dehydrogenase.	intermediary metabolism and respiration	COG1028IQR
Rv3515c	fadD19	59708.5	Cytoplasmic	long-chain-fatty-acid--CoA ligase FadD19	Fatty-acid-CoA ligase FadD19 (fatty-acid-CoA synthetase) (fatty-acid-CoA synthase)	lipid metabolism	COG0318IQ
Rv3518c	cyp142	44398.5	Cytoplasmic Membrane	cytochrome P450 monooxygenase Cyp142	Probable cytochrome P450 monooxygenase 142 Cyp142	intermediary metabolism and respiration	COG2124Q
Rv3548c	-	31070.1	Unknown	short-chain type dehydrogenase/reductase	Probable short-chain type dehydrogenase/reductase	intermediary metabolism and respiration	COG1028IQR
Rv3561	fadD3	54076.5	Cytoplasmic	fatty-acid--CoA ligase FadD3	Probable fatty-acid-CoA ligase FadD3 (fatty-acid-CoA synthetase) (fatty-acid-CoA synthase)	lipid metabolism	COG0318IQ
Rv3563	fadE32	33455.1	Cytoplasmic	acyl-CoA dehydrogenase FadE32	Probable acyl-CoA dehydrogenase FadE32	lipid metabolism	COG1960I
Rv3569c	hsaD	31875.4	Cytoplasmic	4,5:9,10-diseco-3-hydroxy-5,9,17-trioxoandrosta-1(10),2-diene-4-oate hydrolase	4,9-DHSA hydrolase	intermediary metabolism and respiration	COG0596R
Rv3579c	-	34009.8	Unknown	tRNA/rRNA methyltransferase	Possible tRNA/rRNA methyltransferase	intermediary metabolism and respiration	COG0566J
Rv3581c	ispF	16425.8	Cytoplasmic Membrane	2C-methyl-D-erythritol 2,4-cyclodiphosphate synthase	Probable 2C-methyl-D-erythritol 2,4-cyclodiphosphate synthase IspF (MECPS)	intermediary metabolism and respiration	COG0245I
Rv3583c	-	17907.4	Cytoplasmic	RNA polymerase-binding transcription factor CarD	Possible transcription factor	Regulatory proteins	COG1329K
Rv3600c	-	29272.6	Cytoplasmic	type III pantothenate kinase	Putative transcriptional regulator, homolog of Bvg accessory factor	Regulatory proteins	COG1521K
Rv3610c	ftsH	81985.5	Cytoplasmic Membrane	zinc metalloprotease FtsH	Membrane-bound protease FtsH (cell division protein)	cell wall and cell processes	COG0465O
Rv3627c	-	46834.3	Unknown	hypothetical protein Rv3627c	D-alanyl-D-alanine carboxypeptidase (penicillin-binding protein 4)	cell wall and cell processes	COG2027M
Rv3628	ppa	18296.6	Cytoplasmic	inorganic pyrophosphatase	Inorganic pyrophosphatase Ppa (pyrophosphate phosphohydrolase) (PPASE) (inorganic diphosphatase) (diphosphate phosphohydrolase)	intermediary metabolism and respiration	COG0221C
Rv3633	-	32447	Cytoplasmic	hypothetical protein Rv3633	Protein involved in biosynthesis of mitomycin antibiotics/polyketide fumonisins	Lipid metabolism	COG5285Q
Rv3644c	-	41784.5	Cytoplasmic	DNA polymerase	Possible DNA polymerase	information pathways	COG0470L
Rv3674c	nth	26998.3	Cytoplasmic	endonuclease III	Probable endonuclease III Nth (DNA-(apurinic or	information pathways	COG0177L

			Membrane		aprimidinic site)lyase) (AP lyase) (AP endonuclease class I) (endodeoxyribonuclease (apurinic or apyrimidinic)) (deoxyribonuclease (apurinic or apyrimidinic))		
Rv3696c	glpK	55827.8	Cytoplasmic	glycerol kinase	Probable glycerol kinase GlpK (ATP:glycerol 3-phosphotransferase) (glycerokinase) (GK)	intermediary metabolism and respiration	COG0554C
Rv3758c	proV	39552.3	Cytoplasmic Membrane	glycine betaine/carnitine/choline/L-proline ABC transporter ATP-binding protein ProV	Possible osmoprotectant (glycine betaine/carnitine/choline/L-proline) transport ATP-binding protein ABC transporter ProV	virulence, detoxification, adaptation	COG1125E
Rv3767c	-	34723.2	Unknown	S-adenosylmethionine-dependent methyltransferase	Possible S-adenosylmethionine-dependent methyltransferase	lipid metabolism	COG3315Q
Rv3772	hisC2	38007.5	Unknown	histidinol-phosphate aminotransferase	Probable histidinol-phosphate aminotransferase HisC2 (imidazole acetol-phosphate transaminase) (imidazolylacetolphosphate aminotransferase)	intermediary metabolism and respiration	COG0079E
Rv3777	-	33931.3	Cytoplasmic	oxidoreductase	Probable oxidoreductase	intermediary metabolism and respiration	COG0604CR
Rv3781	rfbE	29960.3	Cytoplasmic Membrane	O-antigen/lipopolysaccharide ABC transporter ATP-binding protein RfbE	Probable O-antigen/lipopolysaccharide transport ATP-binding protein ABC transporter RfbE	cell wall and cell processes	COG1134GM
Rv3795	embB	118021	Cytoplasmic Membrane	arabinoxyltransferase B	Integral membrane indolylacetylaminositol arabinoxyltransferase EmbB (arabinoxylindolylacetylaminositol synthase)	cell wall and cell processes	-
Rv3796	-	39443.6	Cytoplasmic Membrane	hypothetical protein Rv3796	Metal-dependent hydrolases of the beta-lactamase superfamily III	Intermediary metabolism and respiration	COG1234R
Rv3797	fadE35	65322.2	Cytoplasmic	acyl-CoA dehydrogenase FadE35	Probable acyl-CoA dehydrogenase FadE35	lipid metabolism	COG1960I
Rv3816c	-	28421.8	Cytoplasmic	acyltransferase	Possible acyltransferase	intermediary metabolism and respiration	COG0204I
Rv3820c	papA2	52118	Cytoplasmic	trehalose-2-sulfate acyltransferase	Possible conserved polyketide synthase associated protein PapA2	lipid metabolism	-
Rv3853	rraA	16203.4	Unknown	RNase E regulator RraA	Regulator of RNase E activity a RraA	Regulatory proteins	COG0684H
Rv3874	esxB	10793.8	Extracellular	ESAT-6-like protein EsxB	10 kDa culture filtrate antigen EsxB (LHP) (CFP10)	cell wall and cell processes	COG4842S
Rv3877	eccD1	53981.3	Cytoplasmic Membrane	ESX-1 secretion system protein EccD1	ESX conserved component EccD1. ESX-1 type VII secretion system protein. Probable transmembrane protein.	cell wall and cell processes	-
Rv3879c	espK	74492.3	Unknown	ESX-1 secretion-associated protein EspK	ESX-1 secretion-associated protein EspK. Alanine and proline rich protein.	cell wall and cell processes	COG0810M
Rv3881c	espB	47593.8	Cytoplasmic	ESX-1 secretion-associated protein EspB	Secreted ESX-1 substrate protein B, EspB. Conserved alanine and glycine rich protein	cell wall and cell processes	COG0038P
Rv3882c	eccE1	50396.7	Cytoplasmic	ESX-1 secretion system protein EccE1	ESX conserved component EccE1. ESX-1 type VII secretion system protein. Possible membrane protein.	cell wall and cell processes	-

Rv3895c	eccB2	51587	Unknown	ESX-2 secretion system protein EccB	ESX conserved component EccB2. ESX-2 type VII secretion system protein. Probable membrane protein.	cell wall and cell processes	-
Rv3915	-	43911.5	Cellwall	peptidoglycan hydrolase	Probable peptidoglycan hydrolase	intermediary metabolism and respiration	COG0860M
Rv3921c	-	40884.4	Cytoplasmic Membrane	membrane protein insertase YidC	Probable conserved transmembrane protein	cell wall and cell processes	COG0706U
Rv0003	recF	42180.2	Cytoplasmic	DNA replication/repair protein RecF	DNA replication and repair protein RecF (single-strand DNA binding protein)	information pathways	COG1195L
Rv0029	-	39559.2	Cytoplasmic Membrane	hypothetical protein Rv0029	Conserved hypothetical protein	conserved hypotheticals	-
Rv0034	-	13967.7	Cytoplasmic	hypothetical protein Rv0034	Ketosteroid isomerase-related protein	Intermediary metabolism and respiration	COG3631R
Rv0035	fadD34	60961.6	Cytoplasmic Membrane	fatty-acid--CoA ligase FadD34	Probable fatty-acid-CoA ligase FadD34 (fatty-acid-CoA synthetase) (fatty-acid-CoA synthase)	lipid metabolism	COG0318IQ
Rv0042c	-	22410.9	Cytoplasmic	transcriptional regulator	Possible transcriptional regulatory protein (probably MarR-family)	Regulatory proteins	COG1846K
Rv0069c	sdaA	48576.2	Cytoplasmic	L-serine dehydratase	Probable L-serine dehydratase SdaA (L-serine deaminase) (SDH) (L-SD)	intermediary metabolism and respiration	COG1760E
Rv0070c	glyA2	45525.7	Cytoplasmic	serine hydroxymethyltransferase	Serine hydroxymethyltransferase GlyA2 (serine methylase 2) (SHMT 2)	intermediary metabolism and respiration	COG0112E
Rv0074	-	42791.3	Cytoplasmic	hypothetical protein Rv0074	Imidazolonepropionase and related amidohydrolases	Intermediary metabolism and respiration	COG1228Q
Rv0081	-	12324.3	Cytoplasmic	HTH-type transcriptional regulator	Probable transcriptional regulatory protein	Regulatory proteins	COG0640K
Rv0099	fadD10	56589.7	Cytoplasmic	fatty-acid--CoA ligase FadD10	Possible fatty-acid-CoA ligase FadD10 (fatty-acid-CoA synthetase) (fatty-acid-CoA synthase)	lipid metabolism	COG0318IQ
Rv0117	oxyS	34191.2	Cytoplasmic	oxidative stress response regulatory protein OxyS	Oxidative stress response regulatory protein OxyS	Regulatory proteins	COG0583K
Rv0118c	oxcA	61171.3	Cytoplasmic	oxalyl-CoA decarboxylase OxcA	Probable oxalyl-CoA decarboxylase OxcA	intermediary metabolism and respiration	COG0028EH
Rv0133	-	22827.1	Cytoplasmic	GCN5-like N-acetyltransferase	GCN5-related N-acetyltransferase	intermediary metabolism and respiration	COG3153R
Rv0143c	-	50808.1	Cytoplasmic Membrane	transmembrane protein	Probable conserved transmembrane protein	cell wall and cell processes	COG0038P
Rv0145	-	34349.6	Unknown	S-adenosylmethionine-dependent methyltransferase	Possible S-adenosylmethionine-dependent methyltransferase	lipid metabolism	COG3315Q
Rv0167	yrbE1A	27743.1	Cytoplasmic Membrane	membrane protein	Conserved integral membrane protein YrbE1A	virulence, detoxification, adaptation	COG0767Q

Rv0178	-	25879.7	Unknown	Mce associated membrane protein	Probable conserved Mce associated membrane protein	cell wall and cell processes	-
Rv0179c	lprO	39160.3	Unknown	lipoprotein LprO	Possible lipoprotein LprO	cell wall and cell processes	-
Rv0182c	sigG	40982.7	Cytoplasmic	ECF RNA polymerase sigma factor SigG	Probable alternative RNA polymerase sigma factor SigG (RNA polymerase ECF type sigma factor)	information pathways	COG1595K
Rv0186	bglS	73520.2	Cytoplasmic	beta-glucosidase BglS	Probable beta-glucosidase BglS (gentiobiase) (cellobiase) (beta-D-glucoside glucohydrolase)	intermediary metabolism and respiration	COG1472G
Rv0187	-	23079.4	Cytoplasmic	O-methyltransferase	Probable O-methyltransferase	intermediary metabolism and respiration	COG4122R
Rv0190	-	10364.8	Cytoplasmic	hypothetical protein Rv0190	Conserved protein	conserved hypotheticals	COG1937S
Rv0202c	mmpL11	103502	Cytoplasmic Membrane	transmembrane transport protein MmpL11	Probable conserved transmembrane transport protein MmpL11	cell wall and cell processes	COG2409R
Rv0207c	-	26175.6	Unknown	hypothetical protein Rv0207c	membrane associated process	cell wall and cell processes	-
Rv0230c	php	35853.8	Cytoplasmic Membrane	phosphotriesterase	Probable phosphotriesterase Php (parathion hydrolase) (PTE) (aryldialkylphosphatase) (paraoxonase) (a-esterase) (aryltriphosphatase) (paraoxon hydrolase)	lipid metabolism	COG1735R
Rv0255c	cobQ1	52134.6	Cytoplasmic	cobyrinic acid synthase	Probable cobyrinic acid synthase CobQ1	intermediary metabolism and respiration	COG1492H
Rv0262c	aac	20005.6	Unknown	aminoglycoside 2'-N-acetyltransferase	Aminoglycoside 2'-N-acetyltransferase Aac (Aac(2')-IC)	virulence, detoxification, adaptation	COG3393R
Rv0289	espG3	31558.8	Cytoplasmic Membrane	ESX-3 secretion-associated protein EspG3	ESX-3 secretion-associated protein EspG3	cell wall and cell processes	-
Rv0293c	-	43970.5	Cytoplasmic	hypothetical protein Rv0293c	Conserved protein	conserved hypotheticals	COG1647R
Rv0317c	glpQ2	28286.6	Cytoplasmic	glycerophosphoryl diester phosphodiesterase GlpQ	Possible glycerophosphoryl diester phosphodiesterase GlpQ2 (glycerophosphodiester phosphodiesterase)	intermediary metabolism and respiration	COG0584C
Rv0331	-	40788.4	Cytoplasmic Membrane	dehydrogenase/reductase	Possible dehydrogenase/reductase	intermediary metabolism and respiration	COG0446R
Rv0332	-	28602.2	Cytoplasmic	hypothetical protein Rv0332	Conserved protein	conserved hypotheticals	-
Rv0333	-	13074.4	Cytoplasmic	hypothetical protein Rv0333	Unknown protein	conserved hypotheticals	-
Rv0337c	aspC	47318	Cytoplasmic	aspartate aminotransferase	Probable aspartate aminotransferase AspC (transaminase A) (ASPAT)	intermediary metabolism and respiration	COG0436E
Rv0339c	-	87743.9	Cytoplasmic	transcriptional regulator	Possible transcriptional regulatory protein	Regulatory proteins	COG2197TK

Rv0356c	-	22878.9	Unknown	hypothetical protein Rv0356c	Thioesterase	Intermediary metabolism and respiration	COG2050Q
Rv0365c	-	41504	Cytoplasmic Membrane	hypothetical protein Rv0365c	Predicted glycosyl hydrolase	Intermediary metabolism and respiration	COG4833G
Rv0394c	-	25278.7	Cytoplasmic Membrane	hypothetical protein Rv0394c	Possible secreted protein	cell wall and cell processes	-
Rv0404	fadD30	63843.4	Unknown	long-chain-fatty-acid--AMP ligase FadD30	Fatty-acid-AMP ligase FadD30 (fatty-acid-AMP synthetase) (fatty-acid-AMP synthase)	lipid metabolism	COG0318IQ
Rv0405	pks6	147613	Cytoplasmic Membrane	membrane bound polyketide synthase	Probable membrane bound polyketide synthase Pks6	lipid metabolism	COG3321Q
Rv0408	pta	72930.5	Cytoplasmic	phosphate acetyltransferase	Probable phosphate acetyltransferase Pta (phosphotransacetylase)	intermediary metabolism and respiration	COG0280C
Rv0412c	-	47083.6	Cytoplasmic Membrane	membrane protein	Possible conserved membrane protein	cell wall and cell processes	-
Rv0417	thiG	25845.7	Cytoplasmic	thiazole synthase	Probable thiamin biosynthesis protein ThiG (thiazole biosynthesis protein)	intermediary metabolism and respiration	COG2022H
Rv0428c	-	32650.1	Unknown	GCN5-like N-acetyltransferase	GCN5-related N-acetyltransferase	intermediary metabolism and respiration	-
Rv0429c	def	20938.7	Cytoplasmic Membrane	polypeptide deformylase	Probable polypeptide deformylase Def (PDF) (formylmethionine deformylase)	information pathways	COG0242J
Rv0438c	moeA2	41989	Cytoplasmic	molybdopterin molybdenumtransferase	Probable molybdopterin biosynthesis protein MoeA2	intermediary metabolism and respiration	COG0303H
Rv0443	-	18958.3	Cytoplasmic	hypothetical protein Rv0443	link to FMN-binding protein	Intermediary metabolism and respiration	-
Rv0456c	echA2	33085.8	Cytoplasmic	enoyl-CoA hydratase EchA2	enoyl-CoA hydratase EchA2 (enoyl hydratase) (unsaturated acyl-CoA hydratase) (crotonase)	lipid metabolism	COG1024I
Rv0474	-	15364.2	Cytoplasmic	HTH-type transcriptional regulator	Probable transcriptional regulatory protein	Regulatory proteins	COG1396K
Rv0480c	-	28995	Unknown	amidohydrolase	Possible amidohydrolase	intermediary metabolism and respiration	COG0388R
Rv0483	lprQ	47858.8	Unknown	lipoprotein LprQ	Probable conserved lipoprotein LprQ	cell wall and cell processes	COG1376S
Rv0491	regX3	24849.4	Cytoplasmic	two component sensory transduction protein RegX	Two component sensory transduction protein RegX3 (transcriptional regulatory protein) (probably LuxR-family)	Regulatory proteins	COG0745TK
Rv0492c	-	66154.2	Cytoplasmic	GMC-type oxidoreductase	Probable oxidoreductase GMC-type	intermediary metabolism and respiration	COG2303E

Rv0495c	-	32960.2	Cytoplasmic	hypothetical protein Rv0495c	DNA-associated process	Regulatory proteins	-
Rv0497	-	33092.1	Cytoplasmic Membrane	transmembrane protein	Probable conserved transmembrane protein	cell wall and cell processes	-
Rv0502	-	39287.9	Cytoplasmic Membrane	hypothetical protein Rv0502	1-acyl-sn-glycerol-3-phosphate acyltransferase	Intermediary metabolism and respiration	COG0204I
Rv0507	mmpL2	106202	Cytoplasmic Membrane	transmembrane transport protein MmpL2	Probable conserved transmembrane transport protein MmpL2	cell wall and cell processes	COG2409R
Rv0510	hemC	31906.3	Cytoplasmic	porphobilinogen deaminase	Probable porphobilinogen deaminase HemC (PBG) (hydroxymethylbilane synthase) (HMBS) (pre-uroporphyrinogen synthase)	intermediary metabolism and respiration	COG0181H
Rv0523c	-	14827.2	Cytoplasmic Membrane	hypothetical protein Rv0523c	Conserved protein	conserved hypotheticals	-
Rv0525	-	22212.1	Cytoplasmic Membrane	hypothetical protein Rv0525	Fructose-2,6-bisphosphatase	Intermediary metabolism and respiration	COG0406G
Rv0531	-	11436.5	Cytoplasmic Membrane	membrane protein	Possible conserved membrane protein	cell wall and cell processes	-
Rv0533c	fabH	34872.5	Cytoplasmic	3-oxoacyl-[acyl-carrier-protein] synthase III	3-oxoacyl-[acyl-carrier-protein] synthase III FabH (beta-ketoacyl-ACP synthase III) (KAS III)	lipid metabolism	COG0332I
Rv0534c	menA	29982.3	Cytoplasmic Membrane	1,4-dihydroxy-2-naphthoate octaprenyltransferase	1,4-dihydroxy-2-naphthoate octaprenyltransferase MenA (DHNA-octaprenyltransferase)	intermediary metabolism and respiration	COG1575H
Rv0555	menD	57803.8	Unknown	bifunctional 2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate synthase/2-oxoglutarate decarboxylase	Probable bifunctional menaquinone biosynthesis protein MenD : 2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate synthase (SHCHC synthase) + 2-oxoglutarate decarboxylase (alpha- ketoglutarate decarboxylase) (KDC)	intermediary metabolism and respiration	COG1165H
Rv0565c	-	54525.8	Cytoplasmic	monooxygenase	Probable monooxygenase	intermediary metabolism and respiration	COG2072P
Rv0566c	-	18077.6	Cytoplasmic	hypothetical protein Rv0566c	Conserved protein	conserved hypotheticals	COG1666S
Rv0586	mce2R	26509.2	Cytoplasmic	HTH-type transcriptional regulator Mce2R	Probable transcriptional regulatory protein Mce2R (GntR-family)	Regulatory proteins	COG2186K
Rv0587	yrbE2A	27327.3	Cytoplasmic Membrane	hypothetical protein Rv0587	Conserved hypothetical integral membrane protein YrbE2A	virulence, detoxification, adaptation	COG0767Q
Rv0592	mce2D	54539.8	Unknown	Mce family protein Mce2D	Mce-family protein Mce2D	virulence, detoxification, adaptation	COG1463Q
Rv0593	lprL	43142.3	Cytoplasmic Membrane	Mce family lipoprotein LprL	Possible Mce-family lipoprotein LprL (Mce-family lipoprotein Mce2E)	cell wall and cell processes	COG1463Q
Rv0597c	-	44694.2	Cytoplasmic	hypothetical protein Rv0597c	Predicted ATPase (AAA+ superfamily)	Intermediary metabolism and	COG1373R

						respiration	
Rv0604	lpqO	32887.1	Cytoplasmic Membrane	lipoprotein LpqO	Probable conserved lipoprotein LpqO	cell wall and cell processes	-
Rv0630c	recB	118722	Cytoplasmic	exonuclease V subunit beta RecB	Probable exonuclease V (beta chain) RecB (exodeoxyribonuclease V beta chain)(exodeoxyribonuclease V polypeptide) (chi-specific endonuclease)	information pathways	COG1074L
Rv0633c	-	32261.7	Cytoplasmic	hypothetical protein Rv0633c	Possible exported protein	cell wall and cell processes	-
Rv0640	rplK	15003.4	Cytoplasmic	50S ribosomal protein L11	50S ribosomal protein L11 RplK	information pathways	COG0080J
Rv0663	atsD	86127.6	Cytoplasmic	arylsulfatase AtsD	Possible arylsulfatase AtsD (aryl-sulfate sulphohydrolase) (arylsulphatase)	intermediary metabolism and respiration	COG3119P
Rv0673	echA4	33541.1	Cytoplasmic	enoyl-CoA hydratase EchA4	Possible enoyl-CoA hydratase EchA4 (enoyl hydrase) (unsaturated acyl-CoA hydratase) (crotonase)	lipid metabolism	COG1024I
Rv0678	-	18346.7	Cytoplasmic	hypothetical protein Rv0678	Transcriptional regulator protein	Regulatory proteins	COG1510K
Rv0687	-	28995.9	Cytoplasmic	NAD-dependent oxidoreductase	Probable short-chain type dehydrogenase/reductase	intermediary metabolism and respiration	COG1028IQR
Rv0696	-	50666.6	Cytoplasmic	mycofactocin biosynthesis glycosyltransferase MftF	Probable membrane sugar transferase	intermediary metabolism and respiration	COG1215M
Rv0705	rpsS	10803.5	Cytoplasmic	30S ribosomal protein S19	30S ribosomal protein S19 RpsS	information pathways	COG0185J
Rv0710	rpsQ	14840.4	Cytoplasmic	30S ribosomal protein S17	30S ribosomal protein S17 RpsQ	information pathways	COG0186J
Rv0726c	-	40875.9	Cytoplasmic	S-adenosylmethionine-dependent methyltransferase	Possible S-adenosylmethionine-dependent methyltransferase	lipid metabolism	COG3315Q
Rv0732	secY	47579.3	Cytoplasmic Membrane	preprotein translocase SecY	Probable preprotein translocase SecY	cell wall and cell processes	COG0201U
Rv0738	-	19115.7	Unknown	hypothetical protein Rv0738	Conserved protein	conserved hypotheticals	-
Rv0751c	mmsB	29679	Cytoplasmic	3-hydroxyisobutyrate dehydrogenase	Probable 3-hydroxyisobutyrate dehydrogenase MmsB (hibadh)	intermediary metabolism and respiration	COG2084I
Rv0752c	fadE9	41451.2	Cytoplasmic	acyl-CoA dehydrogenase FadE9	Probable acyl-CoA dehydrogenase FadE9	lipid metabolism	COG1960I
Rv0758	phoR	52016.4	Cytoplasmic Membrane	two component system response sensor kinase PhoR	Possible two component system response sensor kinase membrane associated PhoR	Regulatory proteins	COG0642T
Rv0768	aldA	51183.9	Cytoplasmic	aldehyde dehydrogenase AldA	Probable aldehyde dehydrogenase NAD dependent AldA (aldehyde dehydrogenase [NAD+])	intermediary metabolism and respiration	COG1012C
Rv0769	-	25965.9	Cytoplasmic	oxidoreductase	Probable dehydrogenase/reductase	intermediary metabolism and respiration	COG1028IQR
Rv0778	cyp126	45954	Cytoplasmic Membrane	cytochrome P450 Cyp126	Possible cytochrome P450 126 Cyp126	intermediary metabolism and	COG2124Q

						respiration	
Rv0799c	-	36015.7	Cytoplasmic Membrane	hypothetical protein Rv0799c	Predicted iron-dependent peroxidase	Intermediary metabolism and respiration	COG2837P
Rv0802c	-	24983.3	Cytoplasmic Membrane	succinyl-CoA transferase	Possible succinyltransferase in the GCN5-related N-acetyltransferase family	intermediary metabolism and respiration	COG1670J
Rv0808	purF	56129.6	Cytoplasmic	amidophosphoribosyltransferase	Amidophosphoribosyltransferase PurF (glutamine phosphoribosylpyrophosphate amidotransferase) (ATASE) (gpatase)	intermediary metabolism and respiration	COG0034F
Rv0811c	-	39414.8	Cytoplasmic	hypothetical protein Rv0811c	Predicted aminomethyltransferase related to GcvT	Intermediary metabolism and respiration	COG0354R
Rv0813c	-	23868.3	Cytoplasmic Membrane	hypothetical protein Rv0813c	UPF0678 fatty acid-binding protein-like protein	Lipid metabolism	-
Rv0823c	-	41388.4	Cytoplasmic	tRNA-dihydrouridine synthase	Possible transcriptional regulatory protein	Regulatory proteins	COG0042J
Rv0830	-	33416.5	Cytoplasmic	S-adenosylmethionine-dependent methyltransferase	Possible S-adenosylmethionine-dependent methyltransferase	lipid metabolism	COG3315Q
Rv0844c	narL	22916.4	Cytoplasmic	nitrate/nitrite response transcriptional regulator NarL	Possible nitrate/nitrite response transcriptional regulatory protein NarL	Regulatory proteins	COG2197TK
Rv0865	mog	16188.3	Unknown	molybdopterin biosynthesis protein	Probable molybdopterin biosynthesis Mog protein	intermediary metabolism and respiration	COG0521H
Rv0886	fprB	61337.2	Cytoplasmic	ferredoxin/ferredoxin--NADP reductase	Probable NADPH:adrenodoxin oxidoreductase FprB (adrenodoxin reductase) (AR) (ferredoxin-NADP(+) reductase)	intermediary metabolism and respiration	COG0493ER
Rv0897c	-	56233.9	Unknown	oxidoreductase	Probable oxidoreductase	intermediary metabolism and respiration	COG1233Q
Rv0904c	accD3	51740.3	Cytoplasmic	acetyl-CoA carboxylase carboxyl transferase subunit beta	Putative acetyl-coenzyme A carboxylase carboxyl transferase (subunit beta) AccD3 (accase beta chain)	lipid metabolism	COG0825I
Rv0908	ctpE	84973.5	Cytoplasmic Membrane	metal cation transporter ATPase E	Probable metal cation transporter ATPase P-type CtpE	cell wall and cell processes	COG0474P
Rv0913c	-	56639.1	Unknown	dioxygenase	Possible dioxygenase	intermediary metabolism and respiration	COG3670Q
Rv0914c	-	43874.9	Cytoplasmic	lipid carrier protein or keto acyl-CoA thiolase	Possible lipid carrier protein or keto acyl-CoA thiolase	lipid metabolism	COG0183I
Rv0966c	-	22210.2	Unknown	hypothetical protein Rv0966c	Conserved protein	conserved hypotheticals	-
Rv0973c	accA2	70744.8	Cytoplasmic	acetyl-/propionyl-CoA carboxylase subunit alpha	Probable acetyl-/propionyl-coenzyme A carboxylase alpha chain (alpha subunit) AccA2: biotin carboxylase + biotin carboxyl carrier protein (BCCP)	lipid metabolism	COG4770I
Rv0974c	accD2	56224.3	Cytoplasmic Membrane	acetyl-/propionyl-CoA carboxylase subunit beta	Probable acetyl-/propionyl-CoA carboxylase (beta subunit) AccD2	lipid metabolism	COG4799I

Rv0982	mprB	54407.7	Cytoplasmic Membrane	two component histidine-protein kinase/phosphatase MprB	Two component sensor kinase MprB	Regulatory proteins	COG0642T
Rv0985c	mscL	16023.5	Cytoplasmic Membrane	large-conductance ion mechanosensitive channel	Possible large-conductance ion mechanosensitive channel MscL	cell wall and cell processes	COG1970M
Rv0987	-	93605.8	Cytoplasmic Membrane	adhesion component ABC transporter permease	Probable adhesion component transport transmembrane protein ABC transporter	cell wall and cell processes	COG4591M
Rv0988	-	42776	Cytoplasmic Membrane	hypothetical protein Rv0988	Possible conserved exported protein	cell wall and cell processes	COG5621R
Rv0989c	grcC2	34757.3	Cytoplasmic	polyprenyl-diphosphate synthase GrcC	Probable polyprenyl-diphosphate synthase GrcC2 (polyprenyl pyrophosphate synthetase)	intermediary metabolism and respiration	COG0142H
Rv0993	galU	32378.2	Cytoplasmic	UTP--glucose-1-phosphate uridylyltransferase	UTP--glucose-1-phosphate uridylyltransferase GalU (UDP-glucose pyrophosphorylase) (UDPGP) (alpha-D-glucosyl-1-phosphate uridylyltransferase) (uridine diphosphoglucose pyrophosphorylase)	intermediary metabolism and respiration	COG1210M
Rv1011	ispE	31349.6	Unknown	4-diphosphocytidyl-2-C-methyl-D-erythritol kinase	Probable 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase IspE (CMK) (4-(cytidine-5'-diphospho)-2-C-methyl-D-erythritol kinase)	intermediary metabolism and respiration	COG1947I
Rv1059	-	36878.2	Cytoplasmic	hypothetical protein Rv1059	Uncharacterized conserved protein related to dihydrodipicolinate reductase	Intermediary metabolism and respiration	COG3804S
Rv1063c	-	37522.7	Cytoplasmic	NTE family protein	Predicted esterase of the alpha-beta hydrolase superfamily	Intermediary metabolism and respiration	COG1752R
Rv1072	-	29111.4	Cytoplasmic Membrane	transmembrane protein	Probable conserved transmembrane protein	cell wall and cell processes	COG4760S
Rv1084	-	71050.3	Cytoplasmic Membrane	hypothetical protein Rv1084	Highly conserved protein containing a thioredoxin domain	Intermediary metabolism and respiration	COG1331O
Rv1096	-	31077.1	Unknown	glycosyl hydrolase	Possible glycosyl hydrolase	intermediary metabolism and respiration	COG0726G
Rv1097c	-	29873.6	Unknown	membrane glycine/proline-rich protein	Probable membrane glycine and proline rich protein	cell wall and cell processes	COG5178A
Rv1101c	-	41698.6	Cytoplasmic Membrane	hypothetical protein Rv1101c	Conserved membrane protein	cell wall and cell processes	COG0628R
Rv1126c	-	22197.3	Cytoplasmic	hypothetical protein Rv1126c	Conserved protein	conserved hypotheticals	-
Rv1129c	-	54087.4	Cytoplasmic	transcriptional regulator	Probable transcriptional regulator protein	Regulatory proteins	COG3800R
Rv1132	-	64019.7	Cytoplasmic Membrane	hypothetical protein Rv1132	Conserved membrane protein	cell wall and cell processes	-
Rv1141c	echA11	27615.8	Cytoplasmic Membrane	enoyl-CoA hydratase EchA11	Probable enoyl-CoA hydratase EchA11 (enoyl hydratase) (unsaturated acyl-CoA hydratase) (crotonase)	lipid metabolism	COG1024I
Rv1142c	echA10	27648.9	Unknown	enoyl-CoA hydratase EchA10	Probable enoyl-CoA hydratase EchA10 (enoyl hydratase) (unsaturated acyl-CoA hydratase) (crotonase)	lipid metabolism	COG1024I

Rv1143	mcr	38684.9	Cytoplasmic	alpha-methylacyl-CoA racemase	Probable alpha-methylacyl-CoA racemase Mcr (2-methylacyl-CoA racemase) (2-arylpropionyl-CoA epimerase)	lipid metabolism	COG1804C
Rv1182	papA3	51575.4	Cytoplasmic	acyltransferase PapA	Probable conserved polyketide synthase associated protein PapA3	lipid metabolism	-
Rv1183	mmpL10	106383	Cytoplasmic Membrane	transmembrane transport protein MmpL10	Probable conserved transmembrane transport protein MmpL10	cell wall and cell processes	COG2409R
Rv1190	-	31804.4	Unknown	hypothetical protein Rv1190	Predicted hydrolases or acyltransferases (alpha/beta hydrolase superfamily)	Intermediary metabolism and respiration	COG0596R
Rv1213	glgC	43800.4	Unknown	glucose-1-phosphate adenylyltransferase	Glucose-1-phosphate adenylyltransferase GlgC (ADP-glucose synthase) (ADP-glucose pyrophosphorylase)	intermediary metabolism and respiration	COG0448G
Rv1253	deaD	61452.6	Cytoplasmic	ATP-dependent RNA helicase DeaD	Probable cold-shock DeaD-box protein A homolog DeaD (ATP-dependent RNA helicase dead homolog)	information pathways	COG0513LKJ
Rv1264	-	42200	Cytoplasmic	adenylyl cyclase	Adenylyl cyclase (ATP pyrophosphate-lyase) (adenylate cyclase)	intermediary metabolism and respiration	COG2114T
Rv1266c	pknH	66753.9	Cytoplasmic Membrane	serine/threonine-protein kinase PknH	Probable transmembrane serine/threonine-protein kinase H PknH (protein kinase H) (STPK H)	Regulatory proteins	COG0515R TKL
Rv1275	lprC	18870.7	Unknown	lipoprotein LprC	Possible lipoprotein LprC	cell wall and cell processes	-
Rv1278	-	93318.8	Cytoplasmic	hypothetical protein Rv1278	DNA-repair	Information pathways	COG4717S
Rv1285	cysD	37512.5	Cytoplasmic	sulfate adenylyltransferase subunit 2	Probable sulfate adenylyltransferase subunit 2 CysD	intermediary metabolism and respiration	COG0175EH
Rv1290c	-	55995.7	Cytoplasmic Membrane	hypothetical protein Rv1290c	Predicted membrane protein	cell wall and cell processes	COG4325S
Rv1296	thrB	32346.4	Cytoplasmic Membrane	homoserine kinase	Probable homoserine kinase ThrB	intermediary metabolism and respiration	COG0083E
Rv1298	rpmE	8753.03	Cytoplasmic	50S ribosomal protein L31	50S ribosomal protein L31 RpmE	information pathways	COG0254J
Rv1317c	alkA	53710.5	Cytoplasmic	bifunctional regulatory protein/DNA repair enzyme AlkA	Probable bifunctional regulatory protein and DNA repair enzyme AlkA (regulatory protein of adaptative response) (methylphosphotriester-DNA--protein-cysteine S-methyltransferase)	information pathways	COG0122L
Rv1333	-	33952.8	Cytoplasmic Membrane	hydrolase	Probable hydrolase	intermediary metabolism and respiration	COG3191EQ
Rv1339	-	29154	Unknown	hypothetical protein Rv1339	Metal-dependent hydrolases of the beta-lactamase superfamily III	Intermediary metabolism and respiration	COG1234R
Rv1360	-	37251.9	Cytoplasmic	oxidoreductase	Probable oxidoreductase	intermediary metabolism and respiration	COG2141C

Rv1362c	-	23471.2	Unknown	membrane protein	Possible membrane protein	cell wall and cell processes	-
Rv1386	PE15	9830.04	Cytoplasmic Membrane	PE family protein PE15	PE family protein PE15	PE/PPE	-
Rv1390	rpoZ	11810.2	Cytoplasmic	DNA-directed RNA polymerase subunit omega	Probable DNA-directed RNA polymerase (omega chain) RpoZ (transcriptase omega chain) (RNA polymerase omega subunit)	information pathways	COG1758K
Rv1393c	-	55251.7	Cytoplasmic	monooxygenase	Probable monooxygenase	intermediary metabolism and respiration	COG2072P
Rv1416	ribH	16338.5	Unknown	6,7-dimethyl-8-ribityllumazine synthase	Probable riboflavin synthase beta chain RibH (6,7-dimethyl-8-ribityllumazine synthase) (DMRL synthase) (lumazine synthase)	intermediary metabolism and respiration	-
Rv1420	uvrC	71582.1	Cytoplasmic	excinuclease ABC subunit UvrC	Probable excinuclease ABC (subunit C - nuclease) UvrC	information pathways	COG0322L
Rv1425	-	50062.7	Cytoplasmic	diacylglycerol O-acyltransferase	Possible triacylglycerol synthase (diacylglycerol acyltransferase)	lipid metabolism	COG4908R
Rv1426c	lipO	46096.9	Cytoplasmic	esterase LipO	Probable esterase LipO	intermediary metabolism and respiration	COG0657I
Rv1432	-	50476.4	Cytoplasmic	dehydrogenase	Probable dehydrogenase	intermediary metabolism and respiration	COG1233Q
Rv1455	-	31426.7	Cytoplasmic Membrane	hypothetical protein Rv1455	Conserved protein	conserved hypotheticals	-
Rv1493	mutB	80604.2	Cytoplasmic	methylmalonyl-CoA mutase large subunit	Probable methylmalonyl-CoA mutase large subunit MutB (MCM)	lipid metabolism	COG1884I
Rv1496	-	36256.4	Cytoplasmic Membrane	transport system kinase	Possible transport system kinase	cell wall and cell processes	COG1703E
Rv1498c	-	23110	Cytoplasmic	methyltransferase	Probable methyltransferase	intermediary metabolism and respiration	COG2226H
Rv1502	-	33472.6	Cytoplasmic	hypothetical protein Rv1502	Hypothetical protein	unknown	-
Rv1530	adh	37996.5	Cytoplasmic	alcohol dehydrogenase	Probable alcohol dehydrogenase Adh	intermediary metabolism and respiration	COG1063ER
Rv1534	-	24468.8	Cytoplasmic	transcriptional regulator	Probable transcriptional regulator	Regulatory proteins	COG1309K
Rv1538c	ansA	32946.5	Cytoplasmic	L-aporaginase	Probable L-aporaginase AnsA	intermediary metabolism and respiration	COG0252EJ
Rv1546	-	15324.5	Unknown	hypothetical protein Rv1546	Conserved protein	conserved hypotheticals	-
Rv1558	-	16347.5	Unknown	hypothetical protein Rv1558	Conserved protein	conserved hypotheticals	-
Rv1564c	treX	81081.1	Cytoplasmic	maltooligosyl trehalose synthase	Probable maltooligosyltrehalose synthase TreX	virulence, detoxification,	COG1523G

						adaptation	
Rv1566c	-	23964.6	Cytoplasmic Membrane	hypothetical protein Rv1566c	Possible Inv protein	virulence, detoxification, adaptation	COG0791M
Rv1576c	-	50474.4	Cytoplasmic Membrane	phage capsid protein	Probable PhiRv1 phage protein	Insertion seqs and phages	COG4653R
Rv1586c	-	51710.2	Cytoplasmic	phage integrase	Probable PhiRv1 integrase	Insertion seqs and phages	COG1961L
Rv1589	bioB	37517.8	Cytoplasmic	biotin synthetase	Probable biotin synthetase BioB	intermediary metabolism and respiration	COG0502H
Rv1591	-	23150.8	Cytoplasmic Membrane	transmembrane protein	Probable transmembrane protein	cell wall and cell processes	-
Rv1595	nadB	53784.8	Cytoplasmic	L-aspartate oxidase	Probable L-aspartate oxidase NadB	intermediary metabolism and respiration	COG0029H
Rv1597	-	27930.6	Cytoplasmic	hypothetical protein Rv1597	SAM-dependent methyltransferases	Intermediary metabolism and respiration	COG2226H
Rv1600	hisC1	40581.1	Cytoplasmic	histidinol-phosphate aminotransferase	Probable histidinol-phosphate aminotransferase HisC1	intermediary metabolism and respiration	COG0079E
Rv1602	hisH	21386.4	Cytoplasmic	imidazole glycerol phosphate synthase subunit HisH	Probable amidotransferase HisH	intermediary metabolism and respiration	COG0118E
Rv1603	hisA	25730.9	Cytoplasmic	phosphoribosyl isomerase A	Probable phosphoribosylformimino-5-aminoimidazole carboxamide ribotide isomerase HisA	intermediary metabolism and respiration	COG0106E
Rv1641	infC	22315.4	Cytoplasmic	initiation factor IF-3	Probable initiation factor if-3 InfC	information pathways	COG0290J
Rv1643	rplT	14526.7	Cytoplasmic	50S ribosomal protein L20	50S ribosomal protein L20 RplT	information pathways	COG0292J
Rv1647	-	33938.7	Cytoplasmic	adenylate cyclase	Adenylate cyclase (ATP pyrophosphate-lyase) (adenylyl cyclase)	intermediary metabolism and respiration	COG2114T
Rv1649	pheS	37371.4	Cytoplasmic	phenylalanine--tRNA ligase subunit alpha	Probable phenylalanyl-tRNA synthetase, alpha chain PheS	information pathways	-
Rv1653	argJ	41115.3	Cytoplasmic	bifunctional glutamate N-acetyltransferase/amino-acid acetyltransferase	Probable glutamate N-acetyltransferase ArgJ	intermediary metabolism and respiration	COG1364E
Rv1656	argF	33025.2	Cytoplasmic	ornithine carbamoyltransferase	Probable ornithine carbamoyltransferase, anabolic ArgF	intermediary metabolism and respiration	COG0078E
Rv1662	pks8	167096	Cytoplasmic Membrane	polyketide synthase	Probable polyketide synthase Pks8	lipid metabolism	COG3321Q
Rv1665	pks11	37639.2	Cytoplasmic	chalcone synthase	Chalcone synthase Pks11	lipid metabolism	COG3424Q
Rv1711	-	27578.8	Cytoplasmic	RNA pseudouridine synthase	16S rRNA uridine-516 pseudouridylylate synthase and	Information pathways	COG1187J

					related pseudouridylate synthases		
Rv1719	-	27872.7	Cytoplasmic	transcriptional regulator	Probable transcriptional regulatory protein	Regulatory proteins	COG1414K
Rv1729c	-	33622.1	Cytoplasmic	S-adenosylmethionine-dependent methyltransferase	Possible S-adenosylmethionine-dependent methyltransferase	lipid metabolism	COG3315Q
Rv1732c	-	19398.9	Unknown	hypothetical protein Rv1732c	Thiol-disulfide isomerase and thioredoxins	Intermediary metabolism and respiration	COG1225O
Rv1737c	narK2	41109.1	Cytoplasmic Membrane	nitrate/nitrite transporter	Possible nitrate/nitrite transporter NarK2	cell wall and cell processes	COG2223P
Rv1743	pknE	60512.2	Cytoplasmic Membrane	serine/threonine-protein kinase PknE	Probable transmembrane serine/threonine-protein kinase E PknE (protein kinase E) (STPK E)	Regulatory proteins	COG0515RTKL
Rv1746	pknF	50668.1	Cytoplasmic Membrane	serine/threonine-protein kinase PknF	Anchored-membrane serine/threonine-protein kinase PknF (protein kinase F) (STPK F)	Regulatory proteins	COG0515RTKL
Rv1750c	fadD1	57413.5	Cytoplasmic Membrane	fatty-acid--CoA ligase FadD1	Possible fatty-acid-CoA ligase FadD1 (fatty-acid-CoA synthetase) (fatty-acid-CoA synthase)	lipid metabolism	COG0318IQ
Rv1762c	-	28754.5	Cytoplasmic	hypothetical protein Rv1762c	Unknown protein	conserved hypotheticals	COG0393S
Rv1778c	-	15828	Unknown	hypothetical protein Rv1778c	Unknown protein	conserved hypotheticals	-
Rv1795	eccD5	53435.2	Cytoplasmic Membrane	ESX-5 type VII secretion system protein EccD	ESX conserved component EccD5. ESX-5 type VII secretion system protein. Probable membrane protein.	cell wall and cell processes	-
Rv1822	pgsA2	22911.3	Cytoplasmic Membrane	CDP-diacylglycerol--glycerol-3-phosphate 3-phosphatidyltransferase	Probable CDP-diacylglycerol--glycerol-3-phosphate 3-phosphatidyltransferase PgsA2 (PGP synthase) (phosphatidylglycerophosphate synthase) (3-phosphatidyl-1'-glycerol-3'phosphate synthase)	lipid metabolism	COG0558I
Rv1825	-	30612.8	Unknown	hypothetical protein Rv1825	membrane associated process	cell wall and cell processes	COG3879S
Rv1835c	-	69931.6	Cytoplasmic	serine esterase	Predicted acyl esterases	Intermediary metabolism and respiration	COG2936R
Rv1841c	-	36685	Cytoplasmic Membrane	hypothetical protein Rv1841c	Conserved hypothetical membrane protein	cell wall and cell processes	COG1253R
Rv1844c	gnd1	51547.1	Cytoplasmic	6-phosphogluconate dehydrogenase	Probable 6-phosphogluconate dehydrogenase Gnd1	intermediary metabolism and respiration	COG0362G
Rv1850	ureC	60825.1	Cytoplasmic	urease subunit alpha	Urease alpha subunit UreC (urea amidohydrolase)	intermediary metabolism and respiration	COG0804E
Rv1865c	-	30372.9	Cytoplasmic	short-chain type dehydrogenase	Probable short-chain type dehydrogenase	intermediary metabolism and respiration	COG4221R
Rv1867	-	53560.5	Unknown	hypothetical protein Rv1867	Conserved protein	lipid metabolism	COG0183I
Rv1882c	-	30010.4	Cytoplasmic	short-chain type dehydrogenase/reductase	Probable short-chain type dehydrogenase/reductase	intermediary metabolism and respiration	COG4221R

Rv1883c	-	17280.7	Unknown	hypothetical protein Rv1883c	Conserved hypothetical protein	conserved hypotheticals	-
Rv1894c	-	40185.7	Cytoplasmic	hypothetical protein Rv1894c	Dioxygenases related to 2-nitropropane dioxygenase	Intermediary metabolism and respiration	COG2070R
Rv1912c	fadB5	35813.2	Cytoplasmic	oxidoreductase FadB	Possible oxidoreductase FadB5	lipid metabolism	COG0604CR
Rv1915	aceAa	40489.8	Cytoplasmic	isocitrate lyase AceAa	Probable isocitrate lyase AceAa [first part] (isocitrase) (isocitratase) (Icl)	intermediary metabolism and respiration	COG2224C
Rv1916	aceAb	44581.8	Cytoplasmic	isocitrate lyase AceAb	Probable isocitrate lyase AceAb [second part] (isocitrase) (isocitratase) (Icl)	intermediary metabolism and respiration	COG2224C
Rv1919c	-	16803.1	Unknown	hypothetical protein Rv1919c	Conserved protein	conserved hypotheticals	-
Rv1920	-	31948.8	Cytoplasmic	membrane protein	Probable membrane protein	cell wall and cell processes	COG0204I
Rv1923	lipD	47151.8	Unknown	lipase LipD	Probable lipase LipD	intermediary metabolism and respiration	COG1680V
Rv1926c	mpt63	16514	Extracellular	immunogenic protein Mpt63	Immunogenic protein Mpt63 (antigen Mpt63/MPB63) (16 kDa immunoprotective extracellular protein)	cell wall and cell processes	-
Rv1963c	mce3R	44341.7	Cytoplasmic	transcriptional repressor Mce3R	Probable transcriptional repressor (probably TetR-family) Mce3R	Regulatory proteins	COG1309K
Rv1977	-	38307.9	Cytoplasmic	hypothetical protein Rv1977	Zn-dependent protease with chaperone function	Information pathways	COG0501O
Rv1992c	ctpG	79277.4	Cytoplasmic Membrane	cation transporter ATPase G	Probable metal cation transporter P-type ATPase G CtpG	cell wall and cell processes	COG2217P
Rv1997	ctpF	95014.6	Cytoplasmic Membrane	cation transporter ATPase F	Probable metal cation transporter P-type ATPase A CtpF	cell wall and cell processes	COG0474P
Rv2001	-	28701.5	Cytoplasmic	hypothetical protein Rv2001	Acyl-ACP thioesterase	Intermediary metabolism and respiration	COG3884I
Rv2003c	-	31017.4	Unknown	hypothetical protein Rv2003c	SAM-dependent methyltransferases	Intermediary metabolism and respiration	COG2226H
Rv2007c	fdxA	12031.7	Unknown	ferredoxin	Ferredoxin FdxA	intermediary metabolism and respiration	COG1146C
Rv2008c	-	48392.8	Cytoplasmic Membrane	hypothetical protein Rv2008c	Predicted ATPase (AAA+ superfamily)	Intermediary metabolism and respiration	COG1373R
Rv2024c	-	57273.9	Cytoplasmic	hypothetical protein Rv2024c	Predicted helicase	Information pathways	COG4889R
Rv2027c	dosT	62136.5	Cytoplasmic Membrane	two component sensor histidine kinase DosT	Two component sensor histidine kinase DosT	Regulatory proteins	COG4585T
Rv2045c	lipT	56055.8	Cytoplasmic	carboxylesterase LipT	Carboxylesterase LipT	intermediary metabolism and	COG2272I

						respiration	
Rv2047c	-	91352.8	Cytoplasmic Membrane	hypothetical protein Rv2047c	Nucleoside-diphosphate-sugar epimerases	Intermediary metabolism and respiration	COG0451MG
Rv2050	-	12971.7	Cytoplasmic	RNA polymerase-binding protein RbpA	Conserved protein	conserved hypotheticals	-
Rv2052c	-	55821.8	Unknown	hypothetical protein Rv2052c	Predicted metal-dependent hydrolase with the TIM-barrel fold	Intermediary metabolism and respiration	COG1574R
Rv2065	cobH	21581.7	Cytoplasmic	precorrin-8X methylmutase	Precorrin-8X methylmutase CobH (aka precorrin isomerase)	intermediary metabolism and respiration	COG2082H
Rv2088	pknJ	61563.3	Cytoplasmic Membrane	transmembrane serine/threonine-protein kinase PknJ	Transmembrane serine/threonine-protein kinase J PknJ (protein kinase J) (STPK J)	Regulatory proteins	COG0515RTKL
Rv2092c	helY	99574	Cytoplasmic	ATP-dependent DNA helicase HelY	ATP-dependent DNA helicase HelY	information pathways	COG4581L
Rv2103c	vapC37	15662.1	Unknown	ribonuclease VapC37	Possible toxin VapC37. Contains PIN domain.	virulence, detoxification, adaptation	COG1848R
Rv2113	-	42394.2	Cytoplasmic Membrane	integral membrane protein	Probable integral membrane protein	cell wall and cell processes	-
Rv2114	-	22774.7	Unknown	hypothetical protein Rv2114	membrane associated process	cell wall and cell processes	-
Rv2141c	-	48110.3	Unknown	hypothetical protein Rv2141c	Conserved protein	intermediary metabolism and respiration	COG0624E
Rv2149c	yfiH	25946.5	Unknown	laccase domain-containing protein	cell division	cell wall and cell processes	COG1496S
Rv2151c	ftsQ	33933.5	Unknown	cell division protein FtsQ	Possible cell division protein FtsQ	cell wall and cell processes	COG1589M
Rv2152c	murC	51144.5	Cytoplasmic	UDP-N-acetylmuramate--alanine ligase	Probable UDP-N-acetylmuramate-alanine ligase MurC	cell wall and cell processes	COG0773M
Rv2184c	-	40838.9	Cytoplasmic	hypothetical protein Rv2184c	Oxyanion-translocating ATPase	Intermediary metabolism and respiration	COG0003P
Rv2194	qcrC	29120.4	Cytoplasmic Membrane	ubiquinol-cytochrome C reductase cytochrome subunit C	Probable ubiquinol-cytochrome C reductase QcrC (cytochrome C subunit)	intermediary metabolism and respiration	COG2010C
Rv2196	qcrB	61014.4	Cytoplasmic Membrane	ubiquinol-cytochrome C reductase cytochrome subunit B	Probable ubiquinol-cytochrome C reductase QcrB (cytochrome B subunit)	intermediary metabolism and respiration	COG1290C
Rv2198c	mmpS3	30955	Unknown	membrane protein MmpS3	Probable conserved membrane protein MmpS3	cell wall and cell processes	COG3889R
Rv2203	-	24371.6	Unknown	membrane protein	Possible conserved membrane protein	cell wall and cell processes	-
Rv2207	cobT	36411.3	Cytoplasmic	nicotinate-nucleotide-	Probable nicotinate-nucleotide-dimethylbenzimidazol	intermediary	COG2038H

				dimethylbenzimidazol phosphoribosyltransferase	phosphoribosyltransferase CobT	metabolism and respiration	
Rv2219	-	26863.6	Cytoplasmic Membrane	transmembrane protein	Probable conserved transmembrane protein	cell wall and cell processes	-
Rv2223c	-	55077.8	Unknown	carboxylesterase B	Probable exported protease	cell wall and cell processes	COG0596R
Rv2225	panB	29336.4	Unknown	3-methyl-2-oxobutanoate hydroxymethyltransferase	3-methyl-2-oxobutanoate hydroxymethyltransferase PanB	intermediary metabolism and respiration	COG0413H
Rv2230c	-	39597.8	Unknown	GTP cyclohydrolase	DNA-associated processes	Regulatory proteins	COG3323S
Rv2237	-	29067.5	Cytoplasmic Membrane	hypothetical protein Rv2237	Conserved protein	conserved hypotheticals	COG3662S
Rv2238c	ahpE	16819	Cytoplasmic	peroxiredoxin	Probable peroxiredoxin AhpE	virulence, detoxification, adaptation	COG1225O
Rv2240c	-	20051.6	Unknown	hypothetical protein Rv2240c	Unknown protein	conserved hypotheticals	-
Rv2242	-	44605.6	Cytoplasmic	hypothetical protein Rv2242	Regulator of polyketide synthase expression	Lipid metabolism	COG2508TQ
Rv2249c	glpD1	54143.9	Cytoplasmic	glycerol-3-phosphate dehydrogenase	Probable glycerol-3-phosphate dehydrogenase GlpD1	intermediary metabolism and respiration	COG0578C
Rv2252	-	32805.4	Cytoplasmic	diacylglycerol kinase	Diacylglycerol kinase	lipid metabolism	COG1597IR
Rv2257c	-	28385.1	Unknown	hypothetical protein Rv2257c	Beta-lactamase class C and other penicillin binding proteins	cell wall and cell processes	COG1680V
Rv2264c	-	59847.6	Unknown	proline-rich protein	Molecular chaperone	Information pathways	COG0443O
Rv2271	-	11534.9	Cytoplasmic	hypothetical protein Rv2271	Conserved hypothetical protein	conserved hypotheticals	-
Rv2297	-	16467.9	Unknown	hypothetical protein Rv2297	Unknown protein	conserved hypotheticals	-
Rv2303c	-	34550.5	Cytoplasmic	antibiotic-resistance protein	Probable antibiotic-resistance protein	virulence, detoxification, adaptation	COG2159R
Rv2319c	-	31834.6	Unknown	universal stress protein	Universal stress protein family protein	virulence, detoxification, adaptation	COG0589T
Rv2332	mez	59422.6	Cytoplasmic Membrane	malate oxidoreductase	Probable [NAD] dependent malate oxidoreductase Mez (malic enzyme) (NAD-malic enzyme) (malate dehydrogenase (oxaloacetate decarboxylating)) (pyruvic-malic carboxylase) (NAD-me)	intermediary metabolism and respiration	COG0281C
Rv2368c	phoH1	37767.3	Cytoplasmic	protein PhoH	Probable PHOH-like protein PhoH1 (phosphate starvation-inducible protein PSIH)	intermediary metabolism and respiration	COG1702T
Rv2380c	mbtE	183282	Cytoplasmic	peptide synthetase	Peptide synthetase MbtE (peptide synthase)	lipid metabolism	-
Rv2383c	mbtB	151602	Cytoplasmic	phenyloxazoline synthase	Phenyloxazoline synthase MbtB (phenyloxazoline	lipid metabolism	COG1020Q

					synthetase)		
Rv2392	cysH	27422.9	Cytoplasmic	phosphoadenosine phosphosulfate reductase	Probable 3'-phosphoadenosine 5'-phosphosulfate reductase CysH (PAPS reductase, thioredoxin DEP.) (padops reductase) (3'- phosphoadenylylsulfate reductase) (PAPS sulfotransferase)	intermediary metabolism and respiration	COG0175EH
Rv2400c	subI	37418.3	Unknown	sulfate ABC transporter substrate-binding lipoprotein SubI	Probable sulfate-binding lipoprotein SubI	cell wall and cell processes	COG1613P
Rv2404c	lepA	72363.7	Cytoplasmic Membrane	GTP-binding protein LepA	Probable GTP-binding protein LepA (GTP-binding elongation factor)	intermediary metabolism and respiration	COG0481M
Rv2406c	-	15053.5	Cytoplasmic	hypothetical protein Rv2406c	FOG- CBS domain	Regulatory proteins	COG2905T
Rv2412	rpsT	9372.81	Cytoplasmic	30S ribosomal protein S20	30S ribosomal protein S20 RpsT	information pathways	COG0268J
Rv2417c	-	28455.7	Unknown	DegV domain-containing protein	Conserved protein	conserved hypotheticals	COG1307S
Rv2418c	-	27225.2	Cytoplasmic	hypothetical protein Rv2418c	Unknown protein	conserved hypotheticals	-
Rv2441c	rpmA	8968.23	Cytoplasmic	50S ribosomal protein L27	50S ribosomal protein L27 RpmA	information pathways	COG0211J
Rv2498c	citE	28885.8	Cytoplasmic	citrate (pro-3S)-lyase subunit beta	Probable citrate (pro-3S)-lyase (beta subunit) CitE (citrase) (citratase) (citritase) (citridesmolase) (citrase aldolase)	intermediary metabolism and respiration	COG2301G
Rv2504c	scoA	26275.7	Cytoplasmic	succinyl-CoA:3-ketoacid-CoA transferase subunit A	Probable succinyl-CoA:3-ketoacid-coenzyme A transferase (alpha subunit) ScoA (3-oxo acid:CoA transferase) (OXCT A) (succinyl-CoA:3-oxoacid-coenzyme A transferase)	lipid metabolism	COG1788I
Rv2520c	-	8341.7	Cytoplasmic	membrane protein	Possible conserved membrane protein	cell wall and cell processes	-
Rv2525c	-	25369.6	Extracellular	hypothetical protein Rv2525c	Conserved hypothetical protein. Secreted; predicted to be a substrate of the twin arginine translocation (tat) export system.	conserved hypotheticals	-
Rv2529	-	50160.1	Cytoplasmic	hypothetical protein Rv2529	Nuclease	Information pathways	COG1948L
Rv2536	-	24626.3	Cytoplasmic Membrane	transmembrane protein	Probable conserved transmembrane protein	cell wall and cell processes	-
Rv2553c	-	45473.8	Cytoplasmic	membrane protein	Probable conserved membrane protein	cell wall and cell processes	COG1559R
Rv2554c	-	18033.4	Cytoplasmic	Holliday junction resolvase	Predicted endonuclease involved in recombination (possible Holliday junction resolvase in Mycoplasmas and B. subtilis)	Information pathways	COG0816L
Rv2571c	-	36789.9	Cytoplasmic Membrane	transmembrane protein	Probable transmembrane alanine and valine and leucine rich protein	cell wall and cell processes	COG4129S
Rv2580c	hisS	45117.1	Cytoplasmic	histidine--tRNA ligase	Probable histidyl-tRNA synthetase HisS (histidine--tRNA ligase) (HISRS) (histidine--translase)	information pathways	COG0124J
Rv2581c	-	24165.6	Cytoplasmic	glyoxalase II	Possible glyoxalase II (hydroxyacylglutathione hydrolase) (GLX II)	virulence, detoxification, adaptation	COG0491R

Rv2583c	relA	87321.8	Cytoplasmic	bifunctional (p)ppGpp synthase/hydrolase RelA	Probable GTP pyrophosphokinase RelA (ATP:GTP 3'-pyrophosphotransferase) (PPGPP synthetase I) ((P)PPGPP synthetase) (GTP diphosphokinase)	intermediary metabolism and respiration	COG0317TK
Rv2584c	apt	23155.7	Cytoplasmic	adenine phosphoribosyltransferase	Adenine phosphoribosyltransferase Apt (APRT) (AMP diphosphorylase) (AMP pyrophosphorylase) (transphosphoribosidase)	intermediary metabolism and respiration	COG0503F
Rv2585c	-	58602.6	Unknown	lipoprotein	Possible conserved lipoprotein	cell wall and cell processes	COG0747E
Rv2587c	secD	60235.1	Cytoplasmic Membrane	protein translocase subunit SecD	Probable protein-export membrane protein SecD	cell wall and cell processes	COG0342U
Rv2597	-	21910.9	Unknown	membrane protein	Probable membrane protein	cell wall and cell processes	-
Rv2607	pdxH	25186.3	Cytoplasmic	pyridoxine/pyridoxamine 5'-phosphate oxidase	Probable pyridoxamine 5'-phosphate oxidase PdxH (PNP/PMP oxidase) (pyridoxinephosphate oxidase) (PNPOX) (pyridoxine 5'-phosphate oxidase)	intermediary metabolism and respiration	COG0259H
Rv2609c	-	38096.4	Cytoplasmic	membrane protein	Probable conserved membrane protein	cell wall and cell processes	COG0494LR
Rv2611c	-	35094.3	Unknown	phosphatidylinositol mannoside acyltransferase	Probable acyltransferase	lipid metabolism	COG1560M
Rv2613c	-	21836.7	Cytoplasmic Membrane	AP-4-A phosphorylase	Conserved protein	lipid metabolism	COG0537FGR
Rv2616	-	17801.3	Unknown	hypothetical protein Rv2616	Conserved protein	conserved hypotheticals	COG4762S
Rv2632c	-	10083.3	Cytoplasmic	hypothetical protein Rv2632c	Conserved protein	conserved hypotheticals	-
Rv2670c	-	39659.2	Cytoplasmic	hypothetical protein Rv2670c	Predicted ATPase	Intermediary metabolism and respiration	COG1485R
Rv2677c	hemY	46846.4	Cytoplasmic	protoporphyrinogen oxidase	Probable protoporphyrinogen oxidase HemY (protoporphyrinogen-IX oxidase) (protoporphyrinogenase) (PPO)	intermediary metabolism and respiration	COG1232H
Rv2678c	hemE	37606.3	Cytoplasmic	uroporphyrinogen decarboxylase	Probable uroporphyrinogen decarboxylase HemE (uroporphyrinogen III decarboxylase) (URO-D) (UPD)	intermediary metabolism and respiration	COG0407H
Rv2680	-	22555.1	Cytoplasmic	hypothetical protein Rv2680	DNA/RNA-associated process	Regulatory proteins	-
Rv2690c	-	70770	Cytoplasmic Membrane	integral membrane protein	Probable conserved integral membrane alanine and valine and leucine rich protein	cell wall and cell processes	COG0531E
Rv2696c	-	27216.7	Cytoplasmic	hypothetical protein Rv2696c	nucleotide metabolism	Regulatory proteins	-
Rv2697c	dut	15770.9	Cytoplasmic Membrane	deoxyuridine 5'-triphosphate nucleotidohydrolase	Probable deoxyuridine 5'-triphosphate nucleotidohydrolase Dut (dUTPase) (dUTP pyrophosphatase) (deoxyuridine 5'-triphosphatase) (dUTP diphosphatase) (deoxyuridine-triphosphatase)	intermediary metabolism and respiration	COG0756F
Rv2702	ppgK	27429.1	Cytoplasmic	polyphosphate glucokinase	Polyphosphate glucokinase PpgK (polyphosphate-glucose phosphotransferase)	intermediary metabolism and respiration	COG1940KG

Rv2710	sigB	36343.4	Cytoplasmic	RNA polymerase sigma factor SigB	RNA polymerase sigma factor SigB	information pathways	COG0568K
Rv2724c	fadE20	42844.7	Cytoplasmic	acyl-CoA dehydrogenase FadE20	Probable acyl-CoA dehydrogenase FadE20	lipid metabolism	COG1960I
Rv2726c	dapF	29697.6	Cytoplasmic	diaminopimelate epimerase	Probable diaminopimelate epimerase DapF (DAP epimerase)	intermediary metabolism and respiration	COG0253E
Rv2739c	-	39962.3	Unknown	alanine-rich transferase	Possible alanine rich transferase	intermediary metabolism and respiration	COG1819GC
Rv2748c	ftsK	94406.5	Cytoplasmic Membrane	DNA translocase FtsK	Possible cell division transmembrane protein FtsK	cell wall and cell processes	COG1674D
Rv2752c	-	59500.4	Cytoplasmic	ribonuclease J	Predicted hydrolase of the metallo-beta-lactamase superfamily	Intermediary metabolism and respiration	COG0595R
Rv2756c	hsdM	60084.1	Cytoplasmic	type I restriction/modification system DNA methylase HsdM	Possible type I restriction/modification system DNA methylase HsdM (M protein) (DNA methyltransferase)	information pathways	COG0286V
Rv2765	-	26159.5	Unknown	alanine-rich hydrolase	Probable alanine rich hydrolase	intermediary metabolism and respiration	COG0412Q
Rv2775	-	17405.9	Cytoplasmic	GCN5-like N-acetyltransferase	GCN5-related N-acetyltransferase	intermediary metabolism and respiration	-
Rv2786c	ribF	36058.1	Cytoplasmic	bifunctional riboflavin kinase /FMN adenylyltransferase	Probable bifunctional FAD synthetase/riboflavin biosynthesis protein RibF: riboflavin kinase (flavokinase) + FMN adenylyltransferase (FAD pyrophosphorylase) (FAD synthetase)(FAD diphosphorylase) (flavin adenine dinucleotide synthetase)	intermediary metabolism and respiration	COG0196H
Rv2793c	truB	31819.5	Cytoplasmic	tRNA pseudouridine synthase B	Probable tRNA pseudouridine synthase B TruB (tRNA pseudouridine 55 synthase) (PSI55 synthase) (pseudouridylate synthase) (uracil hydrolyase)	information pathways	COG0130J
Rv2795c	-	37568.5	Cytoplasmic	hypothetical protein Rv2795c	Predicted phosphohydrolases	Intermediary metabolism and respiration	COG1409R
Rv2802c	-	38793.9	Cytoplasmic	arginine/alanine-rich protein	Hypothetical arginine and alanine rich protein	conserved hypotheticals	COG5586S
Rv2818c	-	42648.6	Cytoplasmic	CRISPR-associated protein Csm6	Hypothetical protein	unknown	-
Rv2850c	-	66955	Cytoplasmic	magnesium chelatase	Possible magnesium chelatase	intermediary metabolism and respiration	COG1239H
Rv2854	-	38191.2	Cytoplasmic	hypothetical protein Rv2854	Lysophospholipase	Lipid metabolism	COG2267I
Rv2873	mpt83	22070.1	Cytoplasmic Membrane	cell surface lipoprotein	Cell surface lipoprotein Mpt83 (lipoprotein P23)	cell wall and cell processes	COG2335M
Rv2901c	-	12225.8	Cytoplasmic	hypothetical protein Rv2901c	Conserved protein	conserved hypotheticals	-
Rv2903c	lepB	31848.2	Cytoplasmic	signal peptidase	Probable signal peptidase I LepB (SPASE I) (leader	cell wall and cell	COG0681U

			Membrane		peptidase I).	processes	
Rv2908c	-	8520.68	Cytoplasmic	hypothetical protein Rv2908c	Predicted RNA-binding protein (contains KH domain)	Regulatory proteins	COG1837R
Rv2911	dacB2	29746.7	Cytoplasmic Membrane	penicillin-binding protein DacB2	Probable penicillin-binding protein DacB2 (D-alanyl-D-alanine carboxypeptidase) (DD-peptidase) (DD-carboxypeptidase) (PBP) (DD-transpeptidase) (serine-type D-ala-D-ala carboxypeptidase) (D-amino acid hydrolase)	cell wall and cell processes	COG1686M
Rv2913c	-	67205.2	Cytoplasmic	D-amino acid aminohydrolase	Possible D-amino acid aminohydrolase (D-amino acid hydrolase)	intermediary metabolism and respiration	COG3653Q
Rv2918c	glnD	86437.8	Cytoplasmic	bifunctional uridylyltransferase/uridylyl-removing enzyme	Probable [protein-PII] uridylyltransferase GlnD (PII uridylyl-transferase) (uridylyl removing enzyme) (UTASE)	intermediary metabolism and respiration	COG2844O
Rv2946c	pks1	166640	Cytoplasmic Membrane	polyketide synthase	Probable polyketide synthase Pks1	lipid metabolism	COG3321Q
Rv2957	-	31046.1	Cytoplasmic Membrane	PGL/p-HBAD biosynthesis glycosyltransferase	Possible glycosyl transferase	intermediary metabolism and respiration	COG0463M
Rv2958c	-	46795.9	Cytoplasmic Membrane	PGL/p-HBAD biosynthesis glycosyltransferase	Possible glycosyl transferase	intermediary metabolism and respiration	COG1819GC
Rv2973c	recG	80329	Cytoplasmic	ATP-dependent DNA helicase RecG	Probable ATP-dependent DNA helicase RecG	information pathways	COG1200LK
Rv2976c	ung	24449.1	Cytoplasmic	uracil-DNA glycosylase	Probable uracil-DNA glycosylase Ung (UDG)	information pathways	COG0692L
Rv2977c	thiL	34298.5	Cytoplasmic	thiamine-monophosphate kinase	Probable thiamine-monophosphate kinase ThiL (thiamine-phosphate kinase)	intermediary metabolism and respiration	COG0611H
Rv2989	-	24283.9	Cytoplasmic Membrane	transcriptional regulator	Probable transcriptional regulatory protein	Regulatory proteins	COG1414K
Rv2999	lppY	33949.6	Unknown	lipoprotein LppY	Probable conserved lipoprotein LppY	cell wall and cell processes	-
Rv3031	-	57797.3	Cytoplasmic	1,4-alpha-glucan branching protein	linked to sugar transferase activity	Intermediary metabolism and respiration	COG1543S
Rv3038c	-	36049.7	Cytoplasmic	hypothetical protein Rv3038c	SAM-dependent methyltransferases	Intermediary metabolism and respiration	COG2226H
Rv3039c	echA17	26505.9	Unknown	enoyl-CoA hydratase EchA17	Probable enoyl-CoA hydratase EchA17 (crotonase) (unsaturated acyl-CoA hydratase) (enoyl hydratase)	lipid metabolism	COG1024I
Rv3048c	nrdF2	36959.5	Unknown	ribonucleoside-diphosphate reductase subunit beta NrdF2	Ribonucleoside-diphosphate reductase (beta chain) NrdF2 (ribonucleotide reductase small subunit) (R2F protein)	information pathways	COG0208F
Rv3049c	-	58744.6	Cytoplasmic Membrane	monooxygenase	Probable monooxygenase	intermediary metabolism and respiration	COG2072P
Rv3057c	-	31153.8	Cytoplasmic	short chain alcohol	Probable short chain alcohol dehydrogenase/reductase	intermediary	COG4221R

				dehydrogenase/reductase		metabolism and respiration	
Rv3059	cyp136	56227.7	Cytoplasmic	cytochrome P450 Cyp136	Probable cytochrome P450 136 Cyp136	intermediary metabolism and respiration	COG2124Q
Rv3062	ligB	53704.6	Unknown	DNA ligase	Probable ATP-dependent DNA ligase LigB (polydeoxyribonucleotide synthase [ATP]) (polynucleotide ligase [ATP]) (sealase) (DNA repair protein) (DNA joinase)	information pathways	COG1793L
Rv3071	-	40466.2	Cytoplasmic	hypothetical protein Rv3071	membrane associated process	cell wall and cell processes	COG1993S
Rv3077	-	66740.9	Cytoplasmic	hydrolase	Possible hydrolase	intermediary metabolism and respiration	COG3119P
Rv3079c	-	30064.4	Cytoplasmic	hypothetical protein Rv3079c	Coenzyme F420-dependent N5,N10-methylene tetrahydromethanopterin reductase and related flavin-dependent oxidoreductases	Intermediary metabolism and respiration	COG2141C
Rv3081	-	45788.1	Unknown	hypothetical protein Rv3081	Conserved hypothetical protein	conserved hypotheticals	COG4914R
Rv3094c	-	40296.8	Cytoplasmic	hypothetical protein Rv3094c	Acyl-CoA dehydrogenases	Intermediary metabolism and respiration	COG1960I
Rv3117	cysA3	31014.6	Cytoplasmic	thiosulfate sulfurtransferase	Probable thiosulfate sulfurtransferase CysA3 (rhodanese-like protein) (thiosulfate cyanide transsulfurase) (thiosulfate thiotransferase)	intermediary metabolism and respiration	COG2897P
Rv3118	-	10165.3	Unknown	hypothetical protein Rv3118	Conserved hypothetical protein SseC1	intermediary metabolism and respiration	-
Rv3143	-	14353.6	Unknown	response regulator	Probable response regulator	Regulatory proteins	COG0745TK
Rv3154	nuoJ	27589.3	Cytoplasmic Membrane	NADH-quinone oxidoreductase subunit J	Probable NADH dehydrogenase I (chain J) NuoJ (NADH-ubiquinone oxidoreductase chain J)	intermediary metabolism and respiration	COG0839C
Rv3158	nuoN	55342.2	Cytoplasmic Membrane	NADH-quinone oxidoreductase subunit N	Probable NADH dehydrogenase I (chain N) NuoN (NADH-ubiquinone oxidoreductase chain N)	intermediary metabolism and respiration	COG1007C
Rv3160c	-	23120.7	Cytoplasmic	TetR family transcriptional regulator	Possible transcriptional regulatory protein (probably TetR-family)	Regulatory proteins	COG1309K
Rv3179	-	48021	Cytoplasmic	hypothetical protein Rv3179	Predicted ATPase (AAA+ superfamily)	Intermediary metabolism and respiration	COG1373R
Rv3198A	-	8963.21	Unknown	glutaredoxin protein	Possible glutaredoxin protein	intermediary metabolism and respiration	COG0695O
Rv3198c	uvrD2	75603.7	Cytoplasmic	ATP-dependent DNA helicase UvrD	Probable ATP-dependent DNA helicase II UvrD2	information pathways	COG0210L
Rv3208	-	25166.3	Cytoplasmic	TetR family transcriptional regulator	Probable transcriptional regulatory protein (probably	Regulatory proteins	COG1309K

					TetR-family)		
Rv3218	-	34474.5	Unknown	hypothetical protein Rv3218	Sphingosine kinase and enzymes related to eukaryotic diacylglycerol kinase	cell wall and cell processes	COG1597IR
Rv3234c	tgs3	30346.8	Cytoplasmic	diacylglycerol O-acyltransferase	Putative triacylglycerol synthase (diacylglycerol acyltransferase) Tgs3	lipid metabolism	-
Rv3245c	mtrB	61636.6	Cytoplasmic Membrane	two component sensory histidine kinase MtrB	Two component sensory transduction histidine kinase MtrB	Regulatory proteins	COG0642T
Rv3261	fbiA	35303.1	Cytoplasmic	2-phospho-L-lactate transferase	Probable F420 biosynthesis protein FbiA	intermediary metabolism and respiration	COG0391S
Rv3265c	wbbL1	33262.8	Unknown	N-acetylglucosaminyl-diphosphodecaprenol L-rhamnosyltransferase	dTDP-RHA:a-D-GlcNAc-diphosphoryl polyprenol, a-3-L-rhamnosyl transferase WbbL1 (alpha-L-rhamnose-(1->3)-alpha-D-GlcNAc(1->P)-P-decaprenyl)	cell wall and cell processes	-
Rv3269	-	9750.17	Cytoplasmic	hypothetical protein Rv3269	Conserved protein	virulence, detoxification, adaptation	-
Rv3278c	-	19820	Cytoplasmic Membrane	transmembrane protein	Probable conserved transmembrane protein	cell wall and cell processes	COG3428S
Rv3291c	lrpA	16510.6	Unknown	transcriptional regulator LrpA	Probable transcriptional regulatory protein LrpA (Lrp/AsnC-family)	Regulatory proteins	COG1522K
Rv3298c	lpqC	31368.3	Unknown	esterase lipoprotein LpqC	Possible esterase lipoprotein LpqC	cell wall and cell processes	COG3509Q
Rv3306c	amiB1	40740.3	Cytoplasmic	amidase AmiB	Probable amidohydrolase AmiB1 (aminohydrolase)	intermediary metabolism and respiration	COG1473R
Rv3308	pmmB	56163.8	Cytoplasmic	phosphomannomutase PmmB	Probable phosphomannomutase PmmB (phosphomannose mutase)	intermediary metabolism and respiration	COG1109G
Rv3322c	-	21614.6	Cytoplasmic	methyltransferase	Possible methyltransferase	intermediary metabolism and respiration	COG2890J
Rv3324c	moaC3	18523.2	Cytoplasmic	cyclic pyranopterin monophosphate synthase accessory protein	Probable molybdenum cofactor biosynthesis protein C 3 MoaC3	intermediary metabolism and respiration	-
Rv3328c	sigJ	33530.2	Cytoplasmic	ECF RNA polymerase sigma factor SigJ	Probable alternative RNA polymerase sigma factor (fragment) SigJ	information pathways	COG1595K
Rv3339c	icd1	45513.8	Unknown	isocitrate dehydrogenase	Probable isocitrate dehydrogenase [NADP] Icd1 (oxalosuccinate decarboxylase) (IDH) (NADP+-specific ICDH) (IDP)	intermediary metabolism and respiration	COG0538C
Rv3364c	-	13440.5	Unknown	hypothetical protein Rv3364c	Uncharacterized distant relative of homeotic protein bithoraxoid	cell wall and cell processes	COG2018R
Rv3373	echA18	22662.2	Cytoplasmic	enoyl-CoA hydratase	Probable enoyl-CoA hydratase EchA18 (enoyl hydratase) (unsaturated acyl-CoA hydratase) (crotonase)	lipid metabolism	COG1024I
Rv3391	acrA1	70940.2	Cytoplasmic	acyl-CoA-reductase AcrA	Possible multi-functional enzyme with acyl-CoA-reductase activity AcrA1	lipid metabolism	COG3320Q

Rv3423c	alr	43355.6	Cytoplasmic	alanine racemase	Alanine racemase Alr	intermediary metabolism and respiration	COG0787M
Rv3442c	rpsI	16436	Cytoplasmic	30S ribosomal protein S9	30S ribosomal protein S9 RpsI	information pathways	COG0103J
Rv3455c	truA	28059.9	Cytoplasmic	tRNA pseudouridine synthase A	Probable tRNA pseudouridine synthase A TruA (pseudouridylylase synthase I) (pseudouridine synthase I) (uracil hydrolyase)	information pathways	-
Rv3459c	rpsK	14784.9	Cytoplasmic	30S ribosomal protein S11	30S ribosomal protein S11 RpsK	information pathways	COG0100J
Rv3463	-	30650.8	Cytoplasmic	hypothetical protein Rv3463	Coenzyme F420-dependent N5,N10-methylene tetrahydromethanopterin reductase and related flavin-dependent oxidoreductases	Intermediary metabolism and respiration	COG2141C
Rv3465	rmlC	22314	Unknown	dTDP-4-dehydrorhamnose 3,5-epimerase	dTDP-4-dehydrorhamnose 3,5-epimerase RmlC (dTDP-4-keto-6-deoxyglucose 3,5-epimerase) (dTDP-L-rhamnose synthetase) (thymidine diphospho-4-keto-rhamnose 3,5-epimerase)	intermediary metabolism and respiration	COG1898M
Rv3479	-	109260	Cytoplasmic Membrane	transmembrane protein	Possible transmembrane protein	cell wall and cell processes	COG1752R
Rv3482c	-	28501.2	Unknown	membrane protein	Probable conserved membrane protein	cell wall and cell processes	-
Rv3484	-	54354.6	Cytoplasmic Membrane	hypothetical protein Rv3484	Transcriptional regulator	Regulatory proteins	COG1316K
Rv3492c	-	18006.9	Cytoplasmic Membrane	Mce associated protein	Conserved hypothetical Mce associated protein	conserved hypotheticals	-
Rv3505	fadE27	39007.9	Cytoplasmic	acyl-CoA dehydrogenase FadE27	Probable acyl-CoA dehydrogenase FadE27	lipid metabolism	COG1960I
Rv3516	echA19	28216.5	Cytoplasmic	enoyl-CoA hydratase EchA19	Possible enoyl-CoA hydratase EchA19 (enoyl hydratase) (unsaturated acyl-CoA hydratase) (crotonase)	lipid metabolism	COG1024I
Rv3523	ltp3	41421.3	Cytoplasmic	lipid carrier protein	Probable lipid carrier protein or keto acyl-CoA thiolase Ltp3	lipid metabolism	COG0183I
Rv3526	kshA	44235.5	Cytoplasmic	3-ketosteroid-9-alpha-hydroxylase oxygenase subunit	Oxygenase component of 3-ketosteroid-9-alpha-hydroxylase KshA	intermediary metabolism and respiration	COG4638PR
Rv3554	fdxB	76533.7	Cytoplasmic Membrane	electron transfer protein FdxB	Possible electron transfer protein FdxB	intermediary metabolism and respiration	COG3239I
Rv3560c	fadE30	42732.1	Cytoplasmic	acyl-CoA dehydrogenase FadE30	Probable acyl-CoA dehydrogenase FadE30	lipid metabolism	COG1960I
Rv3565	aspB	41048.7	Unknown	aspartate aminotransferase AspB	Possible aspartate aminotransferase AspB (transaminase A) (ASPAT) (glutamic--oxaloacetic transaminase) (glutamic--aspartic transaminase)	intermediary metabolism and respiration	COG0436E
Rv3568c	hsaC	33582.3	Cytoplasmic	extradiol dioxygenase	3,4-DHSA dioxygenase	intermediary metabolism and respiration	COG2514R
Rv3573c	fadE34	74546.2	Cytoplasmic	acyl-CoA dehydrogenase FadE34	Probable acyl-CoA dehydrogenase FadE34	lipid metabolism	COG1960I

Rv3591c	-	28510.4	Cytoplasmic	hydrolase	Possible hydrolase	intermediary metabolism and respiration	COG0596R
Rv3603c	-	31104.6	Cytoplasmic	alanine/leucine-rich protein	folic acid-containing compound biosynthetic process	Intermediary metabolism and respiration	COG5495S
Rv3608c	folP1	28810.9	Unknown	dihydropteroate synthase	Dihydropteroate synthase 1 FolP (DHPS 1) (dihydropteroate pyrophosphorylase 1) (dihydropteroate diphosphorylase 1)	intermediary metabolism and respiration	COG0294H
Rv3609c	folE	22394.7	Unknown	GTP cyclohydrolase I	GTP cyclohydrolase I FolE (GTP-ch-I)	intermediary metabolism and respiration	COG0302H
Rv3614c	espD	19802.5	Unknown	ESX-1 secretion-associated protein EspD	ESX-1 secretion-associated protein EspD	cell wall and cell processes	-
Rv3619c	esxV	9832.78	Unknown	ESAT-6 like protein EsxV	Putative ESAT-6 like protein EsxV (ESAT-6 like protein 1)	cell wall and cell processes	-
Rv3625c	mesJ	34129.9	Cytoplasmic	tRNA(Ile)-lysidine synthase	Possible cell cycle protein MesJ	cell wall and cell processes	COG0037D
Rv3626c	-	38407	Cytoplasmic	hypothetical protein Rv3626c	proteolysis/nucleotide biosynthesis	Intermediary metabolism and respiration	COG5282S
Rv3631	-	26212.9	Cytoplasmic	transferase	Possible transferase (possibly glycosyltransferase)	intermediary metabolism and respiration	COG0463M
Rv3648c	cspA	7370.2	Cytoplasmic	cold shock protein A	Probable cold shock protein A CspA	virulence, detoxification, adaptation	COG1278K
Rv3661	-	30679.1	Unknown	hypothetical protein Rv3661	Conserved hypothetical protein	virulence, detoxification, adaptation	COG0560E
Rv3666c	dppA	58381.5	Cellwall	dipeptide ABC transporter substrate-binding lipoprotein DppA	Probable periplasmic dipeptide-binding lipoprotein DppA	cell wall and cell processes	COG4166E
Rv3669	-	18887.1	Cytoplasmic Membrane	transmembrane protein	Probable conserved transmembrane protein	cell wall and cell processes	-
Rv3682	ponA2	84635.8	Extracellular	bifunctional penicillin-insensitive transglycosylase/penicillin-sensitive transpeptidase	Probable bifunctional membrane-associated penicillin-binding protein 1A/1B PonA2 (murein polymerase) [includes: penicillin-insensitive transglycosylase (peptidoglycan TGASE) + penicillin-sensitive transpeptidase (DD-transpeptidase)]	cell wall and cell processes	COG0744M
Rv3688c	-	16701.1	Cytoplasmic	hypothetical protein Rv3688c	Conserved protein	conserved hypotheticals	COG1610S
Rv3694c	-	35439	Cytoplasmic Membrane	transmembrane protein	Possible conserved transmembrane protein	cell wall and cell processes	COG1300S
Rv3695	-	33167.9	Cytoplasmic Membrane	membrane protein	Possible conserved membrane protein	cell wall and cell processes	COG1714S

Rv3700c	egtE	40446.1	Cytoplasmic	pyridoxal-phosphate-dependent protein EgtE	Conserved hypothetical protein	intermediary metabolism and respiration	COG0520E
Rv3702c	egtC	24627	Unknown	amidohydrolase EgtC	Predicted glutamine amidotransferase	Intermediary metabolism and respiration	COG0121R
Rv3713	cobQ2	24125.7	Cytoplasmic	cobyrinic acid synthase CobQ	Possible cobyrinic acid synthase CobQ2	intermediary metabolism and respiration	COG3442R
Rv3725	-	32725.7	Unknown	oxidoreductase	Possible oxidoreductase	intermediary metabolism and respiration	COG0451MG
Rv3747	-	13507.5	Unknown	hypothetical protein Rv3747	link to PE family	Regulatory proteins	-
Rv3754	tyrA	30991.1	Unknown	prephenate dehydrogenase TyrA	Prephenate dehydrogenase TyrA (PDH) (hydroxyphenylpyruvate synthase)	intermediary metabolism and respiration	-
Rv3755c	-	21969.8	Cytoplasmic	hypothetical protein Rv3755c	Conserved protein	conserved hypotheticals	COG3554S
Rv3761c	fadE36	38532.7	Cytoplasmic	acyl-CoA dehydrogenase FadE36	Possible acyl-CoA dehydrogenase FadE36	lipid metabolism	COG3173R
Rv3782	glfT1	33864	Cytoplasmic Membrane	galactofuranosyl transferase GlfT	UDP-galactofuranosyl transferase GlfT1	cell wall and cell processes	COG1216R
Rv3786c	-	44944.1	Cytoplasmic	hypothetical protein Rv3786c	Membrane proteins related to metalloendopeptidases	Intermediary metabolism and respiration	COG0739M
Rv3802c	-	35448.3	Cytoplasmic	membrane protein	Probable conserved membrane protein	cell wall and cell processes	-
Rv3805c	aftB	68709.6	Cytoplasmic Membrane	terminal beta-(1-2)-arabinofuranosyltransferase	Possible arabinofuranosyltransferase AftB	cell wall and cell processes	-
Rv3807c	-	17218.4	Cytoplasmic Membrane	decaprenylphosphoryl-5-phosphoribose phosphatase	Possible conserved transmembrane protein	cell wall and cell processes	COG0671I
Rv3814c	-	29290.9	Cytoplasmic	acyltransferase	Possible acyltransferase	intermediary metabolism and respiration	COG0204I
Rv3819	-	12825.1	Cytoplasmic	hypothetical protein Rv3819	Unknown protein	conserved hypotheticals	-
Rv3829c	-	57073.2	Cytoplasmic	dehydrogenase	Probable dehydrogenase	intermediary metabolism and respiration	COG1233Q
Rv3831	-	17867.6	Cytoplasmic Membrane	hypothetical protein Rv3831	Hypothetical protein	conserved hypotheticals	-
Rv3852	hns	13823.1	Unknown	histone-like protein Hns	Possible histone-like protein Hns	information pathways	-
Rv3854c	ethA	55326.1	Cytoplasmic	monooxygenase EthA	Monooxygenase EthA	intermediary metabolism and respiration	COG2072P

Rv3866	espG1	30064.1	Cytoplasmic	ESX-1 secretion-associated protein EspG	ESX-1 secretion-associated protein EspG1	cell wall and cell processes	COG4170V
Rv3867	espH	19945.6	Cytoplasmic	ESX-1 secretion-associated protein EspH	ESX-1 secretion-associated protein EspH	cell wall and cell processes	-
Rv3869	eccB1	51092.7	Cytoplasmic Membrane	ESX-1 secretion system protein EccB	ESX conserved component EccB1. ESX-1 type VII secretion system protein. Possible membrane protein.	cell wall and cell processes	-
Rv3873	PPE68	37330	Cytoplasmic Membrane	PPE family protein PPE68	PPE family protein PPE68	PE/PPE	COG5651N
Rv3883c	mycP1	45086	Extracellular	membrane-anchored mycosin	Membrane-anchored mycosin MycP1 (serine protease) (subtilisin-like protease) (subtilase-like) (mycosin-1)	intermediary metabolism and respiration	COG1404O
Rv3885c	eccE2	57638.1	Unknown	ESX-2 secretion system protein EccE	ESX conserved component EccE2. ESX-2 type VII secretion system protein. Possible membrane protein.	cell wall and cell processes	-
Rv0001	dnaA	56548.6	Cytoplasmic	chromosomal replication initiator protein DnaA	Chromosomal replication initiator protein DnaA	information pathways	COG0593L

ПРИЛОЖЕНИЕ 5. Белки, идентифицированные для *M. tuberculosis* кластера Beijing B0/W148

Белок	Синоним	Молекулярная масса	Локализация	Аннотация NCBI	Аннотация Tuberculist	Функциональная категория	COG категория
Rv0001	dnaA	56548.6	Cytoplasmic	chromosomal replication initiator protein DnaA	Chromosomal replication initiator protein DnaA	information pathways	COG0593L
Rv0002	dnaN	42113.1	Cytoplasmic	DNA polymerase III subunit beta	DNA polymerase III (beta chain) DnaN (DNA nucleotidyltransferase)	information pathways	COG0592L
Rv0003	recF	42180.2	Cytoplasmic	DNA replication/repair protein RecF	DNA replication and repair protein RecF (single-strand DNA binding protein)	information pathways	COG1195L
Rv0005	gyrB	74058.7	Cytoplasmic	DNA gyrase subunit B	DNA gyrase (subunit B) GyrB (DNA topoisomerase (ATP-hydrolysing)) (DNA topoisomerase II) (type II DNA topoisomerase)	information pathways	-
Rv0006	gyrA	92274.3	Cytoplasmic	DNA gyrase subunit A	DNA gyrase (subunit A) GyrA (DNA topoisomerase (ATP-hydrolysing)) (DNA topoisomerase II) (type II DNA topoisomerase)	information pathways	COG0188L
Rv0007	-	31040.9	Cytoplasmic Membrane	membrane protein	Possible conserved membrane protein	cell wall and cell processes	-
Rv0008c	-	15670.9	Cytoplasmic Membrane	cell wall synthesis protein CwsA	Possible membrane protein	cell wall and cell processes	-
Rv0009	ppiA	19239.4	Cytoplasmic	iron-regulated peptidyl-prolyl cis-trans isomerase PpiA	Probable iron-regulated peptidyl-prolyl cis-trans isomerase A PpiA (PPIase A) (rotamase A)	information pathways	COG0652O
Rv0013	trpG	24626.9	Cytoplasmic	anthranilate synthase component II	Possible anthranilate synthase component II TrpG (glutamine amidotransferase)	intermediary metabolism and respiration	COG0512EH
Rv0014c	pknB	66509.8	Cytoplasmic Membrane	serine/threonine-protein kinase PknB	Transmembrane serine/threonine-protein kinase B PknB (protein kinase B) (STPK B)	Regulatory proteins	COG2815S
Rv0015c	pknA	45597	Cytoplasmic Membrane	serine/threonine-protein kinase PknA	Transmembrane serine/threonine-protein kinase A PknA (protein kinase A) (STPK A)	Regulatory proteins	COG0515RTKL
Rv0016c	pbpA	51575.6	Cytoplasmic Membrane	penicillin-binding protein PbpA	Probable penicillin-binding protein PbpA	cell wall and cell processes	COG0768M
Rv0018c	pstP	53779.9	Cytoplasmic	phosphoserine/threonine phosphatase PstP	Phosphoserine/threonine phosphatase PstP	Regulatory proteins	COG0631T
Rv0019c	fhaB	17152.1	Cytoplasmic Membrane	FHA domain-containing protein FhaB	Conserved protein with FHA domain, FhaB	Regulatory proteins	COG1716T
Rv0020c	fhaA	56880.3	Cytoplasmic	FHA domain-containing protein FhaA	Conserved protein with FHA domain, FhaA	Regulatory proteins	COG1716T
Rv0029	-	39559.2	Cytoplasmic Membrane	hypothetical protein Rv0029	Conserved hypothetical protein	conserved hypotheticals	-
Rv0036c	-	27546.1	Cytoplasmic Membrane	hypothetical protein Rv0036c	Conserved protein	conserved hypotheticals	-
Rv0038	-	21776.1	Unknown	hypothetical protein Rv0038	Putative transcriptional regulator	Regulatory proteins	COG1678K
Rv0041	leuS	107563	Cytoplasmic	leucine--tRNA ligase	Probable leucyl-tRNA synthetase LeuS (leucine--tRNA ligase) (LEURS)	information pathways	COG0495J

Rv0042c	-	22410.9	Cytoplasmic	transcriptional regulator	Possible transcriptional regulatory protein (probably MarR-family)	Regulatory proteins	COG1846K
Rv0044c	-	28524.3	Unknown	oxidoreductase	Possible oxidoreductase	intermediary metabolism and respiration	COG2141C
Rv0046c	ino1	40093.6	Cytoplasmic	inositol-3-phosphate synthase	myo-inositol-1-phosphate synthase Ino1 (inositol 1-phosphate synthetase) (D-glucose 6-phosphate cycloaldolase) (glucose 6-phosphate cyclase) (glucocycloaldolase)	intermediary metabolism and respiration	COG1260I
Rv0048c	-	30785	Unknown	membrane protein	Possible membrane protein	cell wall and cell processes	-
Rv0050	ponA1	71118.3	Cytoplasmic Membrane	bifunctional penicillin-insensitive transglycosylase/penicillin-sensitive transpeptidase	Probable bifunctional penicillin-binding protein 1A/1B PonA1 (murein polymerase) (PBP1): penicillin-insensitive transglycosylase (peptidoglycan TGASE) + penicillin-sensitive transpeptidase (DD-transpeptidase)	cell wall and cell processes	COG0744M
Rv0053	rpsF	10934.7	Cytoplasmic	30S ribosomal protein S6	30S ribosomal protein S6 RpsF	information pathways	COG0360J
Rv0054	ssb	17321.1	Cytoplasmic	single-strand DNA-binding protein	Single-strand binding protein Ssb (helix-destabilizing protein)	information pathways	COG0629L
Rv0055	rpsR1	9543.21	Cytoplasmic	30S ribosomal protein S18	30S ribosomal protein S18-1 RpsR1	information pathways	COG0238J
Rv0056	rplI	16169.6	Cytoplasmic	50S ribosomal protein L9	50S ribosomal protein L9 RplI	information pathways	COG0359J
Rv0059	-	25580.1	Unknown	hypothetical protein Rv0059	DNA-associated process	Regulatory proteins	-
Rv0060	-	38563.2	Cytoplasmic	hypothetical protein Rv0060	Predicted phosphatase homologous to the C-terminal domain of histone macroH2A1	Intermediary metabolism and respiration	COG2110R
Rv0065	vapC1	14339.4	Unknown	ribonuclease VapC1	Possible toxin VapC1	virulence, detoxification, adaptation	COG4113R
Rv0066c	icd2	82550	Cytoplasmic	isocitrate dehydrogenase	Probable isocitrate dehydrogenase [NADP] Icd2 (oxalosuccinate decarboxylase) (IDH) (NADP ⁺ -specific ICDH) (IDP)	intermediary metabolism and respiration	COG2838C
Rv0068	-	32398.9	Cytoplasmic	oxidoreductase	Probable oxidoreductase	intermediary metabolism and respiration	COG1028IQR
Rv0069c	sdaA	48576.2	Cytoplasmic	L-serine dehydratase	Probable L-serine dehydratase SdaA (L-serine deaminase) (SDH) (L-SD)	intermediary metabolism and respiration	COG1760E
Rv0070c	glyA2	45525.7	Cytoplasmic	serine hydroxymethyltransferase	Serine hydroxymethyltransferase GlyA2 (serine methylase 2) (SHMT 2)	intermediary metabolism and respiration	COG0112E
Rv0073	-	35768.9	Cytoplasmic Membrane	glutamine ABC transporter ATP-binding protein	Probable glutamine-transport ATP-binding protein ABC transporter	cell wall and cell processes	COG1136V
Rv0075	-	42766	Cytoplasmic	aminotransferase	Probable aminotransferase	intermediary metabolism and respiration	COG1168E
Rv0077c	-	29686.3	Cytoplasmic	oxidoreductase	Probable oxidoreductase	intermediary metabolism and	COG0596R

						respiration	
Rv0079	-	29474.7	Cytoplasmic	hypothetical protein Rv0079	Unknown protein	conserved hypotheticals	-
Rv0088	-	24617.3	Unknown	polyketide cyclase/dehydrase	Possible polyketide cyclase/dehydrase	lipid metabolism	COG2867I
Rv0091	mtn	27339.6	Cytoplasmic Membrane	bifunctional 5'- methylthioadenosine nucleosidase /S-adenosylhomocysteine nucleosidase	Probable bifunctional MTA/SAH nucleosidase Mtn: 5'- methylthioadenosine nucleosidase (methylthioadenosine methylthioribohydrolase) + S-adenosylhomocysteine nucleosidase (S-adenosyl-L-homocysteine homocysteinylribohydrolase)	intermediary metabolism and respiration	COG0775F
Rv0101	nrp	269377	Cytoplasmic	peptide synthetase Nrp	Probable peptide synthetase Nrp (peptide synthase)	lipid metabolism	COG1020Q
Rv0106	-	43699.8	Cytoplasmic	hypothetical protein Rv0106	Putative GTPases (G3E family)	Intermediary metabolism and respiration	COG0523R
Rv0107c	ctpI	169575	Cytoplasmic Membrane	cation-transporter ATPase I	Probable cation-transporter ATPase I CtpI	cell wall and cell processes	COG0474P
Rv0117	oxyS	34191.2	Cytoplasmic	oxidative stress response regulatory protein OxyS	Oxidative stress response regulatory protein OxyS	Regulatory proteins	COG0583K
Rv0118c	oxcA	61171.3	Cytoplasmic	oxalyl-CoA decarboxylase OxcA	Probable oxalyl-CoA decarboxylase OxcA	intermediary metabolism and respiration	COG0028EH
Rv0119	fadD7	55158.8	Cytoplasmic	fatty-acid--CoA ligase FadD7	Probable fatty-acid-CoA ligase FadD7 (fatty-acid-CoA synthetase) (fatty-acid-CoA synthase)	lipid metabolism	COG0318IQ
Rv0120c	fusA2	75630.5	Cytoplasmic	elongation factor G	Probable elongation factor G FusA2 (EF-G)	information pathways	COG0480J
Rv0123	-	13623.5	Cytoplasmic	hypothetical protein Rv0123	Unknown protein	conserved hypotheticals	-
Rv0125	pepA	34926.1	Unknown	serine protease PepA	Probable serine protease PepA (serine proteinase) (MTB32A)	intermediary metabolism and respiration	COG0265O
Rv0126	treS	68593	Cytoplasmic	trehalose synthase/amylase TreS	Trehalose synthase TreS	virulence, detoxification, adaptation	COG0366G
Rv0127	mak	49879.1	Cytoplasmic	maltokinase	Maltokinase Mak	virulence, detoxification, adaptation	COG3281G
Rv0129c	fbpC	36771.3	Extracellular	diacylglycerol acyltransferase/mycolyltransferas e Ag85C	Secreted antigen 85-C FbpC (85C) (antigen 85 complex C) (AG58C) (mycolyl transferase 85C) (fibronectin-binding protein C)	lipid metabolism	COG0627R
Rv0130	htdZ	16007.2	Unknown	3-hydroxyl-thioester dehydratase	Probable 3-hydroxyl-thioester dehydratase	intermediary metabolism and respiration	COG2030I
Rv0131c	fadE1	50223.6	Cytoplasmic	acyl-CoA dehydrogenase FadE1	Probable acyl-CoA dehydrogenase FadE1	lipid metabolism	COG1960I
Rv0133	-	22827.1	Cytoplasmic	GCN5-like N-acetyltransferase	GCN5-related N-acetyltransferase	intermediary metabolism and respiration	COG3153R
Rv0134	ephF	33796.7	Cytoplasmic	epoxide hydrolase EphF	Possible epoxide hydrolase EphF (epoxide hydratase) (arene-	virulence,	COG0596R

					oxide hydratase)	detoxification, adaptation	
Rv0135c	-	21943.1	Cytoplasmic	transcriptional regulator	Possible transcriptional regulatory protein	Regulatory proteins	COG1309K
Rv0136	cyp138	49260.6	Cytoplasmic	cytochrome P450 Cyp138	Probable cytochrome P450 138 Cyp138	intermediary metabolism and respiration	COG2124Q
Rv0137c	msrA	20484.8	Unknown	peptide methionine sulfoxide reductase MsrA	Probable peptide methionine sulfoxide reductase MsrA (protein-methionine-S-oxide reductase) (peptide met(O) reductase)	intermediary metabolism and respiration	COG0225O
Rv0139	-	36976.2	Cytoplasmic Membrane	oxidoreductase	Possible oxidoreductase	intermediary metabolism and respiration	COG0451MG
Rv0141c	-	15490.5	Unknown	hypothetical protein Rv0141c	Ketosteroid isomerase-related protein	Intermediary metabolism and respiration	COG3631R
Rv0144	-	30610.9	Cytoplasmic	transcriptional regulator	Probable transcriptional regulatory protein (possibly TetR-family)	Regulatory proteins	COG1309K
Rv0145	-	34349.6	Unknown	S-adenosylmethionine-dependent methyltransferase	Possible S-adenosylmethionine-dependent methyltransferase	lipid metabolism	COG3315Q
Rv0146	-	34016.4	Cytoplasmic	S-adenosylmethionine-dependent methyltransferase	Possible S-adenosylmethionine-dependent methyltransferase	lipid metabolism	COG3315Q
Rv0147	-	55035.1	Cytoplasmic	aldehyde dehydrogenase	Probable aldehyde dehydrogenase (NAD ⁺) dependent	intermediary metabolism and respiration	COG1012C
Rv0148	-	29778.7	Cytoplasmic	short-chain type dehydrogenase/reductase	Probable short-chain type dehydrogenase/reductase	intermediary metabolism and respiration	COG1028IQR
Rv0149	-	33278.5	Cytoplasmic Membrane	quinone oxidoreductase	Possible quinone oxidoreductase (NADPH:quinone oxidoreductase) (zeta-crystallin)	intermediary metabolism and respiration	COG0604CR
Rv0153c	ptbB	30153.8	Cytoplasmic	phosphotyrosine protein phosphatase	Phosphotyrosine protein phosphatase PTPB (protein-tyrosine-phosphatase) (PTPase)	Regulatory proteins	COG2365T
Rv0154c	fadE2	44343.4	Cytoplasmic	acyl-CoA dehydrogenase FadE2	Probable acyl-CoA dehydrogenase FadE2	lipid metabolism	COG1960I
Rv0155	pntAa	37693.9	Cytoplasmic Membrane	NAD(P) transhydrogenase subunit alpha PntAa	Probable NAD(P) transhydrogenase (subunit alpha) PntAa [first part; catalytic part] (pyridine nucleotide transhydrogenase subunit alpha) (nicotinamide nucleotide transhydrogenase subunit alpha)	intermediary metabolism and respiration	COG3288C
Rv0157	pntB	48858.3	Cytoplasmic Membrane	NAD(P) transhydrogenase subunit beta PntB	Probable NAD(P) transhydrogenase (subunit beta) PntB [integral membrane protein] (pyridine nucleotide transhydrogenase subunit beta) (nicotinamide nucleotide transhydrogenase subunit beta)	intermediary metabolism and respiration	COG1282C
Rv0158	-	24128.1	Cytoplasmic Membrane	transcriptional regulator	Probable transcriptional regulatory protein (possibly TetR-family)	Regulatory proteins	COG1309K
Rv0161	-	47355.9	Unknown	oxidoreductase	Possible oxidoreductase	intermediary metabolism and	COG0277C

						respiration	
Rv0162c	adhE1	39333.3	Cytoplasmic	zinc-type alcohol dehydrogenase subunit E	Probable zinc-type alcohol dehydrogenase (E subunit) AdhE1	intermediary metabolism and respiration	COG1062C
Rv0164	TB18.5	17738.3	Cytoplasmic	hypothetical protein Rv0164	Cyclase/Dehydrase	Intermediary metabolism and respiration	COG2867I
Rv0166	fadD5	59887.8	Cytoplasmic	fatty-acid--CoA ligase FadD5	Probable fatty-acid-CoA ligase FadD5 (fatty-acid-CoA synthetase) (fatty-acid-CoA synthase)	lipid metabolism	COG0318IQ
Rv0167	yrbE1A	27743.1	Cytoplasmic Membrane	membrane protein	Conserved integral membrane protein YrbE1A	virulence, detoxification, adaptation	COG0767Q
Rv0169	mce1A	47786.4	Cytoplasmic Membrane	Mce family protein Mce1A	Mce-family protein Mce1A	virulence, detoxification, adaptation	COG1463Q
Rv0170	mce1B	37727.2	Unknown	Mce family protein Mce1B	Mce-family protein Mce1B	virulence, detoxification, adaptation	COG1463Q
Rv0171	mce1C	53817.3	Cytoplasmic	Mce family protein Mce1C	Mce-family protein Mce1C	virulence, detoxification, adaptation	COG1463Q
Rv0172	mce1D	56498.4	Unknown	Mce family protein Mce1D	Mce-family protein Mce1D	virulence, detoxification, adaptation	COG1463Q
Rv0173	lprK	42239.6	Cytoplasmic Membrane	Mce family lipoprotein LprK	Possible Mce-family lipoprotein LprK (Mce-family lipoprotein Mce1E)	cell wall and cell processes	COG1463Q
Rv0174	mce1F	54005.4	Unknown	Mce family protein Mce1F	Mce-family protein Mce1F	virulence, detoxification, adaptation	COG1463Q
Rv0175	-	22323.9	Cytoplasmic	Mce associated membrane protein	Probable conserved Mce associated membrane protein	cell wall and cell processes	-
Rv0176	-	35405.1	Cytoplasmic Membrane	Mce associated transmembrane protein	Probable conserved Mce associated transmembrane protein	cell wall and cell processes	COG1714S
Rv0177	-	20164	Cytoplasmic Membrane	Mce associated protein	membrane associated process	cell wall and cell processes	-
Rv0178	-	25879.7	Unknown	Mce associated membrane protein	Probable conserved Mce associated membrane protein	cell wall and cell processes	-
Rv0179c	lprO	39160.3	Unknown	lipoprotein LprO	Possible lipoprotein LprO	cell wall and cell processes	-
Rv0180c	-	47593.4	Cytoplasmic Membrane	transmembrane protein	Probable conserved transmembrane protein	cell wall and cell processes	COG1511S
Rv0183	-	30261.6	Cytoplasmic Membrane	lysophospholipase	Possible lysophospholipase	intermediary metabolism and respiration	COG2267I
Rv0184	-	26826.1	Unknown	hypothetical protein Rv0184	Conserved hypothetical protein	conserved	-

						hypotheticals	
Rv0186	bglS	73520.2	Cytoplasmic	beta-glucosidase BglS	Probable beta-glucosidase BglS (gentiobiase) (cellobiase) (beta-D-glucoside glucohydrolase)	intermediary metabolism and respiration	COG1472G
Rv0187	-	23079.4	Cytoplasmic	O-methyltransferase	Probable O-methyltransferase	intermediary metabolism and respiration	COG4122R
Rv0189c	ilvD	59351.7	Cytoplasmic	dihydroxy-acid dehydratase	Probable dihydroxy-acid dehydratase IlvD (dad)	intermediary metabolism and respiration	COG0129EG
Rv0190	-	10364.8	Cytoplasmic	hypothetical protein Rv0190	Conserved protein	conserved hypotheticals	COG1937S
Rv0192	-	38903.7	Unknown	hypothetical protein Rv0192	Conserved hypothetical protein	conserved hypotheticals	COG1376S
Rv0198c	zmp1	73881.6	Cytoplasmic	zinc metalloprotease	Probable zinc metalloprotease Zmp1	intermediary metabolism and respiration	COG3590O
Rv0202c	mmpL11	103502	Cytoplasmic Membrane	transmembrane transport protein MmpL11	Probable conserved transmembrane transport protein MmpL11	cell wall and cell processes	COG2409R
Rv0206c	mmpL3	100872	Cytoplasmic Membrane	transmembrane transport protein MmpL3	Possible conserved transmembrane transport protein MmpL3	cell wall and cell processes	COG2409R
Rv0207c	-	26175.6	Unknown	hypothetical protein Rv0207c	membrane associated process	cell wall and cell processes	-
Rv0208c	-	28546.9	Cytoplasmic	tRNA (guanine-N(7)-)-methyltransferase	Hypothetical methyltransferase (methylase)	intermediary metabolism and respiration	COG0220R
Rv0211	pckA	67253.1	Cytoplasmic	phosphoenolpyruvate carboxykinase	Probable iron-regulated phosphoenolpyruvate carboxykinase [GTP] PckA (phosphoenolpyruvate carboxylase) (PEPCK)(pep carboxykinase)	intermediary metabolism and respiration	COG1274C
Rv0214	fadD4	58871	Cytoplasmic	fatty-acid--CoA ligase FadD4	Probable fatty-acid-CoA ligase FadD4 (fatty-acid-CoA synthetase) (fatty-acid-CoA synthase)	lipid metabolism	COG0318IQ
Rv0216	-	35756.5	Unknown	hydratase	Double hotdog hydratase	intermediary metabolism and respiration	COG2030I
Rv0217c	lipW	32184.8	Cytoplasmic	esterase LipW	Possible esterase LipW	intermediary metabolism and respiration	COG0657I
Rv0220	lipC	44307.6	Cytoplasmic Membrane	esterase LipC	Probable esterase LipC	intermediary metabolism and respiration	COG0657I
Rv0221	-	51829.6	Cytoplasmic	diacylglycerol O-acyltransferase	Possible triacylglycerol synthase (diacylglycerol acyltransferase)	lipid metabolism	COG4908R
Rv0222	echA1	27280.3	Cytoplasmic Membrane	enoyl-CoA hydratase EchA1	Probable enoyl-CoA hydratase EchA1 (enoyl hydrase) (unsaturated acyl-CoA hydratase) (crotonase)	lipid metabolism	COG1024I
Rv0223c	-	51157.4	Cytoplasmic	aldehyde dehydrogenase	Probable aldehyde dehydrogenase	intermediary metabolism and	COG1012C

						respiration	
Rv0227c	-	45527.8	Unknown	membrane protein	Probable conserved membrane protein	cell wall and cell processes	-
Rv0230c	php	35853.8	Cytoplasmic Membrane	phosphotriesterase	Probable phosphotriesterase Php (parathion hydrolase) (PTE) (aryldialkylphosphatase) (paraoxonase) (a-esterase) (aryltriphosphatase) (paraoxon hydrolase)	lipid metabolism	COG1735R
Rv0231	fadE4	63494.8	Cytoplasmic	acyl-CoA dehydrogenase FadE4	Probable acyl-CoA dehydrogenase FadE4	lipid metabolism	COG1960I
Rv0233	nrdB	35656.1	Cytoplasmic	ribonucleoside-diphosphate reductase subunit beta NrdB	Ribonucleoside-diphosphate reductase (beta chain) NrdB (ribonucleotide reductase small chain)	information pathways	COG0208F
Rv0234c	gabD1	54299.6	Cytoplasmic	succinate-semialdehyde dehydrogenase	Succinate-semialdehyde dehydrogenase [NADP+] dependent (SSDH) GabD1	intermediary metabolism and respiration	COG1012C
Rv0237	lpqI	39518.9	Cytoplasmic	lipoprotein LpqI	Probable conserved lipoprotein LpqI	cell wall and cell processes	COG1472G
Rv0238	-	22496.7	Unknown	transcriptional regulator	Possible transcriptional regulatory protein (probably TetR-family)	Regulatory proteins	COG1309K
Rv0241c	htdX	30162.9	Cytoplasmic	3-hydroxyacyl-thioester dehydratase HtdX	Probable 3-hydroxyacyl-thioester dehydratase HtdX	intermediary metabolism and respiration	COG2030I
Rv0242c	fabG4	46798.1	Cytoplasmic	3-oxoacyl-ACP reductase FabG	Probable 3-oxoacyl-[acyl-carrier protein] reductase FabG4 (3-ketoacyl-acyl carrier protein reductase)	lipid metabolism	COG1028IQR
Rv0243	fadA2	46075.2	Cytoplasmic	acetyl-CoA acetyltransferase FadA	Probable acetyl-CoA acyltransferase FadA2 (3-ketoacyl-CoA thiolase) (beta-ketothiolase)	lipid metabolism	COG0183I
Rv0244c	fadE5	66014.2	Cytoplasmic	acyl-CoA dehydrogenase FadE5	Probable acyl-CoA dehydrogenase FadE5	lipid metabolism	COG1960I
Rv0247c	-	28625.9	Cytoplasmic	succinate dehydrogenase iron-sulfur subunit	Probable succinate dehydrogenase [iron-sulfur subunit] (succinic dehydrogenase)	intermediary metabolism and respiration	COG0479C
Rv0248c	-	70680.6	Cytoplasmic Membrane	succinate dehydrogenase flavoprotein subunit	Probable succinate dehydrogenase [iron-sulfur subunit] (succinic dehydrogenase)	intermediary metabolism and respiration	COG1053C
Rv0250c	-	10878.1	Cytoplasmic	hypothetical protein Rv0250c	Conserved protein	conserved hypotheticals	-
Rv0251c	hsp	17785.9	Cellwall	heat shock protein	Heat shock protein Hsp (heat-stress-induced ribosome-binding protein A)	virulence, detoxification, adaptation	COG0071O
Rv0255c	cobQ1	52134.6	Cytoplasmic	cobyric acid synthase	Probable cobyric acid synthase CobQ1	intermediary metabolism and respiration	COG1492H
Rv0260c	-	40732.7	Cytoplasmic	transcriptional regulator	Possible transcriptional regulatory protein	Regulatory proteins	COG1587H
Rv0262c	aac	20005.6	Unknown	aminoglycoside 2'-N-acetyltransferase	Aminoglycoside 2'-N-acetyltransferase Aac (Aac(2')-IC)	virulence, detoxification, adaptation	COG3393R
Rv0265c	-	35252.3	Cytoplasmic Membrane	iron ABC transporter substrate-binding lipoprotein	Probable periplasmic iron-transport lipoprotein	cell wall and cell processes	COG0614P

Rv0266c	oplA	129618	Cytoplasmic	5-oxoprolinase OplA	Probable 5-oxoprolinase OplA (5-oxo-L-prolinase) (pyroglutamase) (5-OPASE)	intermediary metabolism and respiration	COG0146EQ
Rv0270	fadD2	59892.6	Cytoplasmic	fatty-acid--CoA ligase FadD2	Probable fatty-acid-CoA ligase FadD2 (fatty-acid-CoA synthetase) (fatty-acid-CoA synthase)	lipid metabolism	COG0318IQ
Rv0271c	fadE6	78477.2	Cytoplasmic	acyl-CoA dehydrogenase FadE6	Probable acyl-CoA dehydrogenase FadE6	lipid metabolism	COG1960I
Rv0272c	-	41059.4	Unknown	hypothetical protein Rv0272c	Unknown protein	conserved hypotheticals	COG1506E
Rv0281	-	32965.9	Cytoplasmic	S-adenosylmethionine-dependent methyltransferase	Possible S-adenosylmethionine-dependent methyltransferase	lipid metabolism	COG3315Q
Rv0282	eccA3	68106.6	Cytoplasmic	ESX-3 secretion system protein EccA	ESX conserved component EccA3. ESX-3 type VII secretion system protein.	cell wall and cell processes	COG0464O
Rv0283	eccB3	55942.5	Cytoplasmic Membrane	ESX-3 secretion system protein EccB3	ESX conserved component EccB3. ESX-3 type VII secretion system protein. Possible membrane protein.	cell wall and cell processes	-
Rv0284	eccC3	145138	Cytoplasmic Membrane	ESX-3 secretion system protein EccC3	ESX conserved component EccC3. ESX-3 type VII secretion system protein. Possible membrane protein.	cell wall and cell processes	COG1674D
Rv0289	espG3	31558.8	Cytoplasmic Membrane	ESX-3 secretion-associated protein EspG3	ESX-3 secretion-associated protein EspG3	cell wall and cell processes	-
Rv0290	eccD3	47944	Cytoplasmic Membrane	ESX-3 secretion system protein EccD	ESX conserved component EccD3. ESX-3 type VII secretion system protein. Probable transmembrane protein.	cell wall and cell processes	-
Rv0291	mycP ₃	46127.1	Cytoplasmic Membrane	membrane-anchored mycosin MycP	Probable membrane-anchored mycosin MycP3 (serine protease) (subtilisin-like protease) (subtilase-like) (mycosin-3)	intermediary metabolism and respiration	COG1404O
Rv0292	eccE3	35932.1	Cytoplasmic Membrane	ESX-3 secretion system protein EccE	ESX conserved component EccE3. ESX-3 type VII secretion system protein. Probable transmembrane protein.	cell wall and cell processes	-
Rv0293c	-	43970.5	Cytoplasmic	hypothetical protein Rv0293c	Conserved protein	conserved hypotheticals	COG1647R
Rv0294	tam	29268.3	Cytoplasmic	trans-aconitate methyltransferase	Probable trans-aconitate methyltransferase Tam	intermediary metabolism and respiration	COG4106R
Rv0295c	-	29774.6	Cytoplasmic	hypothetical protein Rv0295c	Conserved protein	conserved hypotheticals	COG4424S
Rv0296c	-	51814	Cytoplasmic	sulfatase	Probable sulfatase	intermediary metabolism and respiration	COG3119P
Rv0309	-	22528.8	Unknown	hypothetical protein Rv0309	Possible conserved exported protein	cell wall and cell processes	COG3786S
Rv0315	-	32186.1	Unknown	beta-1,3-glucanase	Possible beta-1,3-glucanase precursor	intermediary metabolism and respiration	COG2273G
Rv0317c	glpQ2	28286.6	Cytoplasmic	glycerophosphoryl diester phosphodiesterase GlpQ	Possible glycerophosphoryl diester phosphodiesterase GlpQ2 (glycerophosphodiester phosphodiesterase)	intermediary metabolism and respiration	COG0584C
Rv0321	dcd	20837.6	Cytoplasmic	deoxycytidine triphosphate deaminase	Probable deoxycytidine triphosphate deaminase Dcd (dCTP deaminase)	intermediary metabolism and	COG0717F

						respiration	
Rv0322	udgA	47330.2	Cytoplasmic	UDP-glucose 6-dehydrogenase UdgA	Probable UDP-glucose 6-dehydrogenase UdgA (UDP-GLC dehydrogenase) (UDP-GLCDH) (UDPGDH)	intermediary metabolism and respiration	COG1004M
Rv0328	-	21379.1	Unknown	transcriptional regulator	Possible transcriptional regulatory protein (possibly TetR/AcrR-family)	Regulatory proteins	COG1309K
Rv0331	-	40788.4	Cytoplasmic Membrane	dehydrogenase/reductase	Possible dehydrogenase/reductase	intermediary metabolism and respiration	COG0446R
Rv0332	-	28602.2	Cytoplasmic	hypothetical protein Rv0332	Conserved protein	conserved hypotheticals	-
Rv0333	-	13074.4	Cytoplasmic	hypothetical protein Rv0333	Unknown protein	conserved hypotheticals	-
Rv0334	rmlA	31492	Cytoplasmic	glucose-1-phosphate thymidyltransferase	Alpha-D-glucose-1-phosphate thymidyltransferase RmlA (dTDP-glucose synthase) (dTDP-glucose pyrophosphorylase)	intermediary metabolism and respiration	COG1209M
Rv0337c	aspC	47318	Cytoplasmic	aspartate aminotransferase	Probable aspartate aminotransferase AspC (transaminase A) (ASPAT)	intermediary metabolism and respiration	COG0436E
Rv0338c	-	95453.7	Cytoplasmic Membrane	iron-sulfur-binding reductase	Probable iron-sulfur-binding reductase	intermediary metabolism and respiration	COG0247C
Rv0339c	-	87743.9	Cytoplasmic	transcriptional regulator	Possible transcriptional regulatory protein	Regulatory proteins	COG2197TK
Rv0341	iniB	43921.2	Cytoplasmic Membrane	isoniazid inducible protein IniB	Isoniazid inducible gene protein IniB	cell wall and cell processes	-
Rv0342	iniA	70082.9	Cytoplasmic	isoniazid inducible protein IniA	Isoniazid inducible gene protein IniA	cell wall and cell processes	COG0699R
Rv0343	iniC	52986.9	Cytoplasmic Membrane	iisoniazid inducible protein IniC	Isoniazid inducible gene protein IniC	cell wall and cell processes	COG0699R
Rv0347	-	36565	Cytoplasmic	membrane protein	Probable conserved membrane protein	cell wall and cell processes	-
Rv0349	-	24670.7	Cytoplasmic	hypothetical protein Rv0349	DNA-associated process	Regulatory proteins	-
Rv0350	dnaK	66830.9	Cytoplasmic	chaperone protein DnaK	Probable chaperone protein DnaK (heat shock protein 70) (heat shock 70 kDa protein) (HSP70)	virulence, detoxification, adaptation	COG0443O
Rv0351	grpE	24500.6	Unknown	protein GrpE	Probable GrpE protein (HSP-70 cofactor)	virulence, detoxification, adaptation	COG0576O
Rv0352	dnaJ1	41345.1	Cytoplasmic	chaperone protein DnaJ	Probable chaperone protein DnaJ1	virulence, detoxification, adaptation	COG0484O
Rv0356c	-	22878.9	Unknown	hypothetical protein Rv0356c	Thioesterase	Intermediary metabolism and respiration	COG2050Q
Rv0357c	purA	46822.5	Cytoplasmic	adenylosuccinate synthetase	Probable adenylosuccinate synthetase PurA (imp--aspartate	intermediary	COG0104F

					ligase) (ADSS) (ampase)	metabolism and respiration	
Rv0358	-	23102.5	Cytoplasmic Membrane	hypothetical protein Rv0358	Conserved protein	conserved hypotheticals	-
Rv0360c	-	15297.3	Unknown	hypothetical protein Rv0360c	Conserved protein	conserved hypotheticals	-
Rv0363c	fba	36544.5	Cytoplasmic	fructose-bisphosphate aldolase	Probable fructose-bisphosphate aldolase Fba	intermediary metabolism and respiration	COG0191G
Rv0365c	-	41504	Cytoplasmic Membrane	hypothetical protein Rv0365c	Predicted glycosyl hydrolase	Intermediary metabolism and respiration	COG4833G
Rv0379	secE2	7966.14	Cytoplasmic	protein translocase subunit SecE	Possible protein transport protein SecE2	cell wall and cell processes	COG3360S
Rv0382c	pyrE	18860.4	Unknown	orotate phosphoribosyltransferase	Probable orotate phosphoribosyltransferase PyrE (OPRT) (oprtae)	intermediary metabolism and respiration	COG0461F
Rv0383c	-	31801.4	Cytoplasmic	hypothetical protein Rv0383c	Possible conserved secreted protein	cell wall and cell processes	-
Rv0384c	clpB	92536.2	Cytoplasmic	chaperone protein ClpB	Probable endopeptidase ATP binding protein (chain B) ClpB (ClpB protein) (heat shock protein F84.1)	virulence, detoxification, adaptation	COG0542O
Rv0385	-	43245.1	Cytoplasmic	monooxygenase	Probable monooxygenase	intermediary metabolism and respiration	COG0543HC
Rv0386	-	116792	Cytoplasmic	transcriptional regulator	Probable transcriptional regulatory protein (probably LuxR/UhpA-family)	Regulatory proteins	COG3903R
Rv0389	purT	43645.7	Cytoplasmic Membrane	phosphoribosylglycinamide formyltransferase PurT	Probable phosphoribosylglycinamide formyltransferase 2 PurT (GART 2) (gar transformylase 2) (5'-phosphoribosylglycinamide transformylase 2) (formate-dependent gar transformylase)	intermediary metabolism and respiration	COG0027F
Rv0390	-	15360	Cytoplasmic	hypothetical protein Rv0390	Rhodanese-related sulfurtransferase	Intermediary metabolism and respiration	COG0607P
Rv0391	metZ	43345.2	Cytoplasmic	O-succinylhomoserine sulfhydrylase	Probable O-succinylhomoserine sulfhydrylase MetZ (OSH sulfhydrylase)	intermediary metabolism and respiration	COG0626E
Rv0392c	ndhA	50385.3	Cytoplasmic Membrane	NADH dehydrogenase NdhA	Probable membrane NADH dehydrogenase NdhA	intermediary metabolism and respiration	COG1252C
Rv0394c	-	25278.7	Cytoplasmic Membrane	hypothetical protein Rv0394c	Possible secreted protein	cell wall and cell processes	-
Rv0400c	fadE7	42297.4	Cytoplasmic	acyl-CoA dehydrogenase FadE7	Acyl-CoA dehydrogenase FadE7	lipid metabolism	COG1960I
Rv0404	fadD30	63843.4	Unknown	long-chain-fatty-acid--AMP ligase FadD30	Fatty-acid-AMP ligase FadD30 (fatty-acid-AMP synthetase) (fatty-acid-AMP synthase)	lipid metabolism	COG0318IQ
Rv0406c	-	29220.4	Cytoplasmic	beta lactamase-like protein	Zn-dependent hydrolases, including glyoxyalases	Intermediary	COG0491R

			Membrane			metabolism and respiration	
Rv0407	fgd1	36956.8	Cytoplasmic	F420-dependent glucose-6-phosphate dehydrogenase	F420-dependent glucose-6-phosphate dehydrogenase Fgd1	intermediary metabolism and respiration	COG2141C
Rv0408	pta	72930.5	Cytoplasmic	phosphate acetyltransferase	Probable phosphate acetyltransferase Pta (phosphotransacetylase)	intermediary metabolism and respiration	COG0280C
Rv0409	ackA	41318.2	Cytoplasmic	acetate kinase	Probable acetate kinase AckA (acetokinase)	intermediary metabolism and respiration	COG0282C
Rv0410c	pknG	81577.4	Cytoplasmic Membrane	serine/threonine-protein kinase PknG	Serine/threonine-protein kinase PknG (protein kinase G) (STPK G)	Regulatory proteins	COG0515RTKL
Rv0411c	glnH	35405.7	Cytoplasmic Membrane	glutamine-binding lipoprotein GlnH	Probable glutamine-binding lipoprotein GlnH (GLNBP)	cell wall and cell processes	COG0834ET
Rv0412c	-	47083.6	Cytoplasmic Membrane	membrane protein	Possible conserved membrane protein	cell wall and cell processes	-
Rv0414c	thiE	23239.3	Cytoplasmic	thiamine-phosphate synthase	Thiamine-phosphate pyrophosphorylase ThiE (TMP pyrophosphorylase) (TMP-PPASE) (thiamine-phosphate synthase)	intermediary metabolism and respiration	COG0352H
Rv0418	lpqL	52042.4	Extracellular	lipoprotein aminopeptidase LpqL	Probable lipoprotein aminopeptidase LpqL	cell wall and cell processes	COG2234R
Rv0419	lpqM	52949.1	Cytoplasmic Membrane	lipoprotein peptidase LpqM	Possible lipoprotein peptidase LpqM	cell wall and cell processes	COG2321R
Rv0422c	thiD	27510.5	Unknown	hydroxymethylpyrimidine/phosphomethylpyrimidine kinase	Probable phosphomethylpyrimidine kinase ThiD (HMP-phosphate kinase) (HMP-P kinase)	intermediary metabolism and respiration	COG0351H
Rv0423c	thiC	59897.7	Cytoplasmic	phosphomethylpyrimidine synthase	Probable thiamine biosynthesis protein ThiC	intermediary metabolism and respiration	COG0422H
Rv0425c	ctpH	159773	Cytoplasmic Membrane	metal cation transporting ATPase H	Possible metal cation transporting P-type ATPase CtpH	cell wall and cell processes	COG0474P
Rv0427c	xthA	32108.3	Cytoplasmic	exodeoxyribonuclease III protein XthA	Probable exodeoxyribonuclease III protein XthA (exonuclease III) (EXO III) (AP endonuclease VI)	information pathways	COG0708L
Rv0428c	-	32650.1	Unknown	GCN5-like N-acetyltransferase	GCN5-related N-acetyltransferase	intermediary metabolism and respiration	-
Rv0429c	def	20938.7	Cytoplasmic Membrane	polypeptide deformylase	Probable polypeptide deformylase Def (PDF) (formylmethionine deformylase)	information pathways	COG0242J
Rv0431	-	16905.2	Cytoplasmic Membrane	tuberculin-like peptide	Putative tuberculin related peptide	cell wall and cell processes	-
Rv0432	sodC	23844.5	Cytoplasmic Membrane	superoxide dismutase	Periplasmic superoxide dismutase [Cu-Zn] SodC	virulence, detoxification, adaptation	COG2032P
Rv0433	-	42277.2	Cytoplasmic	carboxylate-amine ligase	Conserved hypothetical protein	conserved	COG2170S

						hypotheticals	
Rv0435c	-	75314	Cytoplasmic Membrane	ATPase	Putative conserved ATPase	cell wall and cell processes	COG0464O
Rv0437c	psd	24227.9	Cytoplasmic Membrane	phosphatidylserine decarboxylase	Possible phosphatidylserine decarboxylase Psd (PS decarboxylase)	lipid metabolism	COG0688I
Rv0438c	moeA ₂	41989	Cytoplasmic	molybdopterin molybdenumtransferase	Probable molybdopterin biosynthesis protein MoeA2	intermediary metabolism and respiration	COG0303H
Rv0439c	-	33650.1	Unknown	dehydrogenase/reductase	Probable dehydrogenase/reductase	intermediary metabolism and respiration	COG1028IQR
Rv0440	groEL ₂	56726.7	Cytoplasmic	molecular chaperone GroEL	60 kDa chaperonin 2 GroEL2 (protein CPN60-2) (GroEL protein 2) (65 kDa antigen) (heat shock protein 65) (cell wall protein A) (antigen A)	virulence, detoxification, adaptation	COG0459O
Rv0441c	-	15004	Cytoplasmic	hypothetical protein Rv0441c	FMN binding protein	Regulatory proteins	-
Rv0443	-	18958.3	Cytoplasmic	hypothetical protein Rv0443	link to FMN-binding protein	Intermediary metabolism and respiration	-
Rv0444c	rskA	23883.1	Cytoplasmic Membrane	anti-sigma-K factor RskA	Anti-sigma factor RskA (regulator of sigma K)	information pathways	COG5343S
Rv0445c	sigK	21034.7	Cytoplasmic Membrane	ECF RNA polymerase sigma factor SigK	Alternative RNA polymerase sigma factor SigK	information pathways	COG1595K
Rv0450c	mmpL ₄	105202	Cytoplasmic Membrane	transmembrane transport protein MmpL4	Probable conserved transmembrane transport protein MmpL4	cell wall and cell processes	COG2409R
Rv0451c	mmpS ₄	15371.5	Unknown	membrane protein MmpS4	Probable conserved membrane protein MmpS4	cell wall and cell processes	-
Rv0452	-	25325.1	Cytoplasmic	transcriptional regulator	Possible transcriptional regulatory protein	Regulatory proteins	COG1309K
Rv0455c	-	16639.8	Cytoplasmic Membrane	hypothetical protein Rv0455c	Conserved protein	conserved hypotheticals	-
Rv0456c	echA2	33085.8	Cytoplasmic	enoyl-CoA hydratase EchA2	enoyl-CoA hydratase EchA2 (enoyl hydratase) (unsaturated acyl-CoA hydratase) (crotonase)	lipid metabolism	COG1024I
Rv0457c	-	74464.1	Cytoplasmic	peptidase	Probable peptidase	intermediary metabolism and respiration	COG1505E
Rv0458	-	54574.8	Cytoplasmic	aldehyde dehydrogenase	Probable aldehyde dehydrogenase	intermediary metabolism and respiration	COG1012C
Rv0462	lpdC	49207.1	Cytoplasmic	dihydrolipoamide dehydrogenase	Dihydrolipoamide dehydrogenase LpdC (lipoamide reductase (NADH)) (lipoyl dehydrogenase) (dihydrolipoyl dehydrogenase) (diaphorase)	intermediary metabolism and respiration	COG1249C
Rv0464c	-	21304.6	Unknown	hypothetical protein Rv0464c	Conserved protein	conserved hypotheticals	COG2128S
Rv0466	-	30153.4	Cytoplasmic	hypothetical protein Rv0466	Acyl-ACP thioesterase	Intermediary metabolism and respiration	COG3884I

Rv0467	icl1	47086.6	Cytoplasmic	isocitrate lyase	Isocitrate lyase Icl (isocitrase) (isocitratase)	intermediary metabolism and respiration	COG2224C
Rv0468	fadB2	30696.4	Cytoplasmic	3-hydroxybutyryl-CoA dehydrogenase	3-hydroxybutyryl-CoA dehydrogenase FadB2 (beta-hydroxybutyryl-CoA dehydrogenase) (BHBD)	lipid metabolism	COG1250I
Rv0469	umaA	33095.7	Cytoplasmic	mycolic acid synthase UmaA	Possible mycolic acid synthase UmaA	lipid metabolism	COG2230M
Rv0470c	pcaA	33027.7	Cytoplasmic	cyclopropane mycolic acid synthase	Mycolic acid synthase PcaA (cyclopropane synthase)	lipid metabolism	COG2230M
Rv0472c	-	26385.5	Cytoplasmic	HTH-type transcriptional regulator	Probable transcriptional regulatory protein (possibly TetR-family)	Regulatory proteins	COG1309K
Rv0473	-	50556.9	Cytoplasmic Membrane	transmembrane protein	Possible conserved transmembrane protein	cell wall and cell processes	-
Rv0474	-	15364.2	Cytoplasmic	HTH-type transcriptional regulator	Probable transcriptional regulatory protein	Regulatory proteins	COG1396K
Rv0475	hbhA	21534.5	Cellwall	heparin binding hemagglutinin HbhA	Iron-regulated heparin binding hemagglutinin HbhA (adhesin)	cell wall and cell processes	COG1754R
Rv0479c	-	37016	Unknown	membrane protein	Probable conserved membrane protein	cell wall and cell processes	-
Rv0480c	-	28995	Unknown	amidohydrolase	Possible amidohydrolase	intermediary metabolism and respiration	COG0388R
Rv0481c	-	19063.6	Unknown	hypothetical protein Rv0481c	membrane associated process	cell wall and cell processes	-
Rv0482	murB	38521	Cytoplasmic	UDP-N-acetylenolpyruvoylglucosamine reductase	Probable UDP-N-acetylenolpyruvoylglucosamine reductase MurB (UDP-N-acetylmuramate dehydrogenase)	cell wall and cell processes	COG0812M
Rv0483	lprQ	47858.8	Unknown	lipoprotein LprQ	Probable conserved lipoprotein LprQ	cell wall and cell processes	COG1376S
Rv0484c	-	26416.1	Cytoplasmic	short-chain type oxidoreductase	Probable short-chain type oxidoreductase	intermediary metabolism and respiration	COG4221R
Rv0489	gpm1	27215.8	Unknown	2,3-bisphosphoglycerate-dependent phosphoglycerate mutase	Probable phosphoglycerate mutase 1 Gpm1 (phosphoglyceromutase) (PGAM) (BPG-dependent PGAM)	intermediary metabolism and respiration	COG0588G
Rv0490	senX3	44792.7	Cytoplasmic Membrane	two component sensor histidine kinase SenX3	Putative two component sensor histidine kinase SenX3	Regulatory proteins	COG5002T
Rv0491	regX3	24849.4	Cytoplasmic	two component sensory transduction protein RegX	Two component sensory transduction protein RegX3 (transcriptional regulatory protein) (probably LuxR-family)	Regulatory proteins	COG0745TK
Rv0492c	-	66154.2	Cytoplasmic	GMC-type oxidoreductase	Probable oxidoreductase GMC-type	intermediary metabolism and respiration	COG2303E
Rv0493c	-	35395	Unknown	hypothetical protein Rv0493c	Conserved protein	conserved hypotheticals	-
Rv0496	-	34843.5	Cytoplasmic	hypothetical protein Rv0496	Exopolyphosphatase	cell wall and cell processes	-

Rv0497	-	33092.1	Cytoplasmic Membrane	transmembrane protein	Probable conserved transmembrane protein	cell wall and cell processes	-
Rv0498	-	30433.6	Cytoplasmic	hypothetical protein Rv0498	Sugar phosphate isomerases/epimerases	Intermediary metabolism and respiration	COG1082G
Rv0500	proC	30171.6	Cytoplasmic	pyrroline-5-carboxylate reductase	Probable pyrroline-5-carboxylate reductase ProC (P5CR) (P5C reductase)	intermediary metabolism and respiration	COG0345E
Rv0500A	-	8507.73	Cytoplasmic	DNA-binding protein	DNA binding protein	Regulatory proteins	-
Rv0501	galE2	41032.1	Cytoplasmic	UDP-glucose 4-epimerase GalE	Possible UDP-glucose 4-epimerase GalE2 (galactowaldenase) (UDP-galactose 4-epimerase) (uridine diphosphate galactose 4-epimerase) (uridine diphospho-galactose 4-epimerase)	intermediary metabolism and respiration	COG0451MG
Rv0502	-	39287.9	Cytoplasmic Membrane	hypothetical protein Rv0502	1-acyl-sn-glycerol-3-phosphate acyltransferase	Intermediary metabolism and respiration	COG0204I
Rv0503c	cmaA 2	34660.1	Cytoplasmic	cyclopropane mycolic acid synthase	Cyclopropane-fatty-acyl-phospholipid synthase 2 CmaA2 (cyclopropane fatty acid synthase) (CFA synthase) (cyclopropane mycolic acid synthase 2) (mycolic acid trans-cyclopropane synthetase)	lipid metabolism	COG2230M
Rv0504c	-	18359.9	Cytoplasmic	hypothetical protein Rv0504c	UPF0336 protein	cell wall and cell processes	COG2030I
Rv0505c	serB1	39508	Cytoplasmic	phosphoserine phosphatase SerB	Possible phosphoserine phosphatase SerB1 (PSP) (O-phosphoserine phosphohydrolase) (pspase)	intermediary metabolism and respiration	COG0560E
Rv0507	mmpL 2	106202	Cytoplasmic Membrane	transmembrane transport protein MmpL2	Probable conserved transmembrane transport protein MmpL2	cell wall and cell processes	COG2409R
Rv0509	hemA	49329.3	Cytoplasmic	glutamyl-tRNA reductase	Probable glutamyl-tRNA reductase HemA (GLUTR)	intermediary metabolism and respiration	COG0373H
Rv0510	hemC	31906.3	Cytoplasmic	porphobilinogen deaminase	Probable porphobilinogen deaminase HemC (PBG) (hydroxymethylbilane synthase) (HMBS) (pre-uroporphyrinogen synthase)	intermediary metabolism and respiration	COG0181H
Rv0511	hemD	58835.6	Cytoplasmic	uroporphyrin-III C-methyltransferase	Probable uroporphyrin-III C-methyltransferase HemD (uroporphyrinogen III methylase) (urogen III methylase) (SUMT) (urogen III methylase) (UROM)	intermediary metabolism and respiration	COG0007H
Rv0512	hemB	34871.5	Cytoplasmic	delta-aminolevulinic acid dehydratase	Probable delta-aminolevulinic acid dehydratase HemB (porphobilinogen synthase) (ALAD) (ALADH)	intermediary metabolism and respiration	COG0113H
Rv0523c	-	14827.2	Cytoplasmic Membrane	hypothetical protein Rv0523c	Conserved protein	conserved hypotheticals	-
Rv0524	hemL	47515.9	Cytoplasmic	glutamate-1-semialdehyde 2,1-aminomutase	Probable glutamate-1-semialdehyde 2,1-aminomutase HemL (GSA) (glutamate-1-semialdehyde aminotransferase) (GSA-at)	intermediary metabolism and respiration	COG0001H
Rv0525	-	22212.1	Cytoplasmic Membrane	hypothetical protein Rv0525	Fructose-2,6-bisphosphatase	Intermediary metabolism and	COG0406G

						respiration	
Rv0530	-	43022.5	Cytoplasmic	hypothetical protein Rv0530	ATPases involved in chromosome partitioning	cell wall and cell processes	COG0455D
Rv0533c	fabH	34872.5	Cytoplasmic	3-oxoacyl-[acyl-carrier-protein] synthase III	3-oxoacyl-[acyl-carrier-protein] synthase III FabH (beta-ketoacyl-ACP synthase III) (KAS III)	lipid metabolism	COG0332I
Rv0535	-	27963.7	Cytoplasmic	S-methyl-5'-thioadenosine phosphorylase	Probable 5'-methylthioadenosine phosphorylase Pnp (MTA phosphorylase)	intermediary metabolism and respiration	COG0005F
Rv0538	-	55076	Cytoplasmic Membrane	membrane protein	Possible conserved membrane protein	cell wall and cell processes	COG3889R
Rv0546c	-	14346.2	Cytoplasmic	hypothetical protein Rv0546c	Lactoylglutathione lyase and related lyases	Intermediary metabolism and respiration	COG3324R
Rv0547c	-	31834.9	Cytoplasmic	oxidoreductase	Possible oxidoreductase	intermediary metabolism and respiration	COG0300R
Rv0548c	menB	34656.9	Cytoplasmic	1,4-dihydroxy-2-naphthoyl-CoA synthase	Naphthoate synthase MenB (dihydroxynaphthoic acid synthetase) (DHNA synthetase)	intermediary metabolism and respiration	COG0447H
Rv0551c	fadD8	61065.9	Cytoplasmic	fatty-acid--CoA ligase FadD8	Probable fatty-acid-CoA ligase FadD8 (fatty-acid-CoA synthetase) (fatty-acid-CoA synthase)	lipid metabolism	COG0318IQ
Rv0552	-	57340.1	Cytoplasmic	hypothetical protein Rv0552	Predicted metal-dependent hydrolase with the TIM-barrel fold	Intermediary metabolism and respiration	COG1574R
Rv0554	bpoC	28349.4	Cytoplasmic	non-heme bromoperoxidase BpoC	Possible peroxidase BpoC (non-haem peroxidase)	virulence, detoxification, adaptation	COG0596R
Rv0555	menD	57803.8	Unknown	bifunctional 2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate synthase/2-oxoglutarate decarboxylase	Probable bifunctional menaquinone biosynthesis protein MenD : 2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate synthase (SHCHC synthase) + 2-oxoglutarate decarboxylase (alpha-ketoglutarate decarboxylase) (KDC)	intermediary metabolism and respiration	COG1165H
Rv0557	mgtA	41205.4	Cytoplasmic	GDP-mannose-dependent alpha-mannosyltransferase	Mannosyltransferase MgtA	lipid metabolism	COG0438M
Rv0558	menH	25298.9	Unknown	demethylmenaquinone methyltransferase	Probable ubiquinone/menaquinone biosynthesis methyltransferase MenH (2-heptaprenyl-1,4-naphthoquinone methyltransferase)	intermediary metabolism and respiration	COG2226H
Rv0560c	-	25945.3	Cytoplasmic	benzoquinone methyltransferase	Possible benzoquinone methyltransferase (methylase)	intermediary metabolism and respiration	COG2890J
Rv0561c	-	43837.9	Unknown	oxidoreductase	Possible oxidoreductase	intermediary metabolism and respiration	COG0644C
Rv0562	grcC1	35528.3	Cytoplasmic	polyprenyl-diphosphate synthase GrcC	Probable polyprenyl-diphosphate synthase GrcC1 (polyprenyl pyrophosphate synthetase)	intermediary metabolism and respiration	COG0142H

Rv0564c	gpdA1	36153.4	Cytoplasmic	glycerol-3-phosphate dehydrogenase	Probable glycerol-3-phosphate dehydrogenase [NAD(P)+] GpdA1 (NAD(P)H-dependent glycerol-3-phosphate dehydrogenase) (NAD(P)H-dependent dihydroxyacetone-phosphate reductase)	lipid metabolism	COG0240C
Rv0565c	-	54525.8	Cytoplasmic	monooxygenase	Probable monooxygenase	intermediary metabolism and respiration	COG2072P
Rv0566c	-	18077.6	Cytoplasmic	hypothetical protein Rv0566c	Conserved protein	conserved hypotheticals	COG1666S
Rv0568	cyp135B1	50687.5	Cytoplasmic	cytochrome P450 Cyp135B1	Possible cytochrome P450 135B1 Cyp135B1	intermediary metabolism and respiration	COG2124Q
Rv0569	-	9522.73	Unknown	hypothetical protein Rv0569	Conserved protein	conserved hypotheticals	-
Rv0570	nrdZ	74489	Cytoplasmic	vitamin B12-dependent ribonucleoside-diphosphate reductase	Probable ribonucleoside-diphosphate reductase (large subunit) NrdZ (ribonucleotide reductase)	information pathways	COG0209F
Rv0571c	-	46655.3	Cytoplasmic	hypothetical protein Rv0571c	Predicted phosphoribosyltransferases	Intermediary metabolism and respiration	COG1926R
Rv0576	-	46407.7	Cytoplasmic	transcriptional regulator	Probable transcriptional regulatory protein (possibly ArsR-family)	Regulatory proteins	COG3832S
Rv0577	TB27.3	27342.9	Unknown	hypothetical protein Rv0577	Predicted enzyme related to lactoylglutathione lyase	Intermediary metabolism and respiration	COG3324R
Rv0580c	-	18034.8	Cytoplasmic Membrane	hypothetical protein Rv0580c	Conserved protein	conserved hypotheticals	-
Rv0583c	lpqN	23682.4	Unknown	lipoprotein LpqN	Probable conserved lipoprotein LpqN	cell wall and cell processes	-
Rv0586	mce2R	26509.2	Cytoplasmic	HTH-type transcriptional regulator Mce2R	Probable transcriptional regulatory protein Mce2R (GntR-family)	Regulatory proteins	COG2186K
Rv0587	yrbE2A	27327.3	Cytoplasmic Membrane	hypothetical protein Rv0587	Conserved hypothetical integral membrane protein YrbE2A	virulence, detoxification, adaptation	COG0767Q
Rv0591	mce2C	50764.1	Unknown	Mce family protein Mce2C	Mce-family protein Mce2C	virulence, detoxification, adaptation	COG1463Q
Rv0592	mce2D	54539.8	Unknown	Mce family protein Mce2D	Mce-family protein Mce2D	virulence, detoxification, adaptation	COG1463Q
Rv0593	lprL	43142.3	Cytoplasmic Membrane	Mce family lipoprotein LprL	Possible Mce-family lipoprotein LprL (Mce-family lipoprotein Mce2E)	cell wall and cell processes	COG1463Q
Rv0597c	-	44694.2	Cytoplasmic	hypothetical protein Rv0597c	Predicted ATPase (AAA+ superfamily)	Intermediary metabolism and respiration	COG1373R

Rv0613c	-	92721.4	Cytoplasmic	hypothetical protein Rv0613c	Unknown protein	conserved hypotheticals	-
Rv0629c	recD	61714.7	Cytoplasmic	exonuclease V subunit alpha RecD	Probable exonuclease V (alpha chain) RecD (exodeoxyribonuclease V alpha chain) (exodeoxyribonuclease V polypeptide)	information pathways	COG0507L
Rv0630c	recB	118722	Cytoplasmic	exonuclease V subunit beta RecB	Probable exonuclease V (beta chain) RecB (exodeoxyribonuclease V beta chain)(exodeoxyribonuclease V polypeptide) (chi-specific endonuclease)	information pathways	COG1074L
Rv0631c	recC	119501	Cytoplasmic Membrane	exonuclease V subunit gamma RecC	Probable exonuclease V (gamma chain) RecC (exodeoxyribonuclease V gamma chain)(exodeoxyribonuclease V polypeptide)	information pathways	COG1330L
Rv0632c	echA3	24354.9	Cytoplasmic	enoyl-CoA hydratase EchA3	Probable enoyl-CoA hydratase EchA3 (enoyl hydratase) (unsaturated acyl-CoA hydratase) (crotonase)	lipid metabolism	COG1024I
Rv0633c	-	32261.7	Cytoplasmic	hypothetical protein Rv0633c	Possible exported protein	cell wall and cell processes	-
Rv0634A	-	9408.44	Cytoplasmic	hypothetical protein Rv0634A	Transcription regulator of the Arc/MetJ class	Regulatory proteins	COG5450K
Rv0635	hadA	17448.5	Cytoplasmic	(3R)-hydroxyacyl-ACP dehydratase subunit HadA	(3R)-hydroxyacyl-ACP dehydratase subunit HadA	intermediary metabolism and respiration	COG2030I
Rv0636	hadB	14934	Unknown	(3R)-hydroxyacyl-ACP dehydratase subunit HadB	(3R)-hydroxyacyl-ACP dehydratase subunit HadB	intermediary metabolism and respiration	COG2030I
Rv0637	hadC	18929.7	Cytoplasmic	(3R)-hydroxyacyl-ACP dehydratase subunit HadC	(3R)-hydroxyacyl-ACP dehydratase subunit HadC	intermediary metabolism and respiration	COG2030I
Rv0638	secE1	16924.3	Cytoplasmic Membrane	preprotein translocase SecE	Probable preprotein translocase SecE1	cell wall and cell processes	COG0690U
Rv0639	nusG	25414.5	Unknown	transcription termination/antitermination protein NusG	Probable transcription antitermination protein NusG	information pathways	COG0250K
Rv0640	rplK	15003.4	Cytoplasmic	50S ribosomal protein L11	50S ribosomal protein L11 RplK	information pathways	COG0080J
Rv0641	rplA	24756.4	Cytoplasmic	50S ribosomal protein L1	50S ribosomal protein L1 RplA	information pathways	COG0081J
Rv0642c	mmaA 4	34670.4	Cytoplasmic	hydroxymycolate synthase MmaA4	Methoxy mycolic acid synthase 4 MmaA4 (methyl mycolic acid synthase 4) (MMA4) (hydroxy mycolic acid synthase)	lipid metabolism	COG2230M
Rv0643c	mmaA 3	33262.9	Cytoplasmic	methoxy mycolic acid synthase MmaA3	Methoxy mycolic acid synthase 3 MmaA3 (methyl mycolic acid synthase 3) (MMA3) (hydroxy mycolic acid synthase)	lipid metabolism	COG2230M
Rv0644c	mmaA 2	32724.3	Cytoplasmic	cyclopropane mycolic acid synthase MmaA2	Methoxy mycolic acid synthase 2 MmaA2 (methyl mycolic acid synthase 2) (MMA2) (hydroxy mycolic acid synthase)	lipid metabolism	COG2230M
Rv0645c	mmaA 1	33149.4	Cytoplasmic	mycolic acid methyltransferase MmaA1	Methoxy mycolic acid synthase 1 MmaA1 (methyl mycolic acid synthase 1) (MMA1) (hydroxy mycolic acid synthase)	lipid metabolism	COG2230M
Rv0646c	lipG	32931	Unknown	lipase/esterase LipG	Probable lipase/esterase LipG	intermediary metabolism and respiration	COG0596R
Rv0647c	-	54824	Cytoplasmic Membrane	hypothetical protein Rv0647c	Predicted unusual protein kinase	Intermediary metabolism and	COG0661R

						respiration	
Rv0648	-	129949	Cytoplasmic	alpha-mannosidase	Alpha-mannosidase	intermediary metabolism and respiration	COG0383G
Rv0651	rplJ	18478.3	Cytoplasmic	50S ribosomal protein L10	50S ribosomal protein L10 RplJ	information pathways	COG0244J
Rv0652	rplL	13440.5	Cytoplasmic	50S ribosomal protein L7/L12	50S ribosomal protein L7/L12 RplL (SA1)	information pathways	COG0222J
Rv0654	-	54943.8	Cytoplasmic	carotenoid cleavage oxygenase	Probable dioxygenase	intermediary metabolism and respiration	COG3670Q
Rv0655	mkl	39362.2	Cytoplasmic Membrane	ribonucleotide ABC transporter ATP-binding protein	Possible ribonucleotide-transport ATP-binding protein ABC transporter Mkl	cell wall and cell processes	COG1127Q
Rv0663	atsD	86127.6	Cytoplasmic	arylsulfatase AtsD	Possible arylsulfatase AtsD (aryl-sulfate sulphohydrolase) (arylsulphatase)	intermediary metabolism and respiration	COG3119P
Rv0667	rpoB	129218	Cytoplasmic	DNA-directed RNA polymerase subunit beta	DNA-directed RNA polymerase (beta chain) RpoB (transcriptase beta chain) (RNA polymerase beta subunit)	information pathways	COG0085K
Rv0668	rpoC	146737	Cytoplasmic	DNA-directed RNA polymerase subunit beta'	DNA-directed RNA polymerase (beta' chain) RpoC (transcriptase beta' chain) (RNA polymerase beta' subunit).	information pathways	COG0086K
Rv0669c	-	69490.1	Cytoplasmic	neutral ceramidase	Possible hydrolase	intermediary metabolism and respiration	-
Rv0670	end	26813.1	Cytoplasmic	endonuclease IV	Probable endonuclease IV End (endodeoxyribonuclease IV) (apurinase)	information pathways	COG0648L
Rv0672	fadE8	58534.5	Cytoplasmic	acyl-CoA dehydrogenase FadE8	Probable acyl-CoA dehydrogenase FadE8	lipid metabolism	COG1960I
Rv0673	echA4	33541.1	Cytoplasmic	enoyl-CoA hydratase EchA4	Possible enoyl-CoA hydratase EchA4 (enoyl hydrase) (unsaturated acyl-CoA hydratase) (crotonase)	lipid metabolism	COG1024I
Rv0675	echA5	27441.4	Cytoplasmic	enoyl-CoA hydratase EchA5	Probable enoyl-CoA hydratase EchA5 (enoyl hydrase) (unsaturated acyl-CoA hydratase) (crotonase)	lipid metabolism	COG1024I
Rv0676c	mmpL5	104785	Cytoplasmic Membrane	transmembrane transport protein MmpL5	Probable conserved transmembrane transport protein MmpL5	cell wall and cell processes	COG2409R
Rv0677c	mmpS5	15248.5	Unknown	membrane protein MmpS5	Possible conserved membrane protein MmpS5	cell wall and cell processes	-
Rv0678	-	18346.7	Cytoplasmic	hypothetical protein Rv0678	Transcriptional regulator protein	Regulatory proteins	COG1510K
Rv0681	-	21230	Cytoplasmic	transcriptional regulator	Probable transcriptional regulatory protein (possibly TetR-family)	Regulatory proteins	COG1309K
Rv0683	rpsG	17600.4	Cytoplasmic	30S ribosomal protein S7	30S ribosomal protein S7 RpsG	information pathways	COG0049J
Rv0684	fusA1	77170.5	Cytoplasmic	elongation factor G	Probable elongation factor G FusA1 (EF-G)	information pathways	COG0480J
Rv0685	tuf	43561.5	Cytoplasmic	elongation factor Tu	Probable iron-regulated elongation factor TU Tuf (EF-TU)	information pathways	COG0050J
Rv0686	-	28543.2	Cytoplasmic Membrane	membrane protein	Probable membrane protein	cell wall and cell processes	-
Rv0687	-	28995.9	Cytoplasmic	NAD-dependent oxidoreductase	Probable short-chain type dehydrogenase/reductase	intermediary metabolism and respiration	COG1028IQR
Rv0688	-	43017.8	Cytoplasmic	ferredoxin reductase	Putative ferredoxin reductase	intermediary	COG1251C

						metabolism and respiration	
Rv0690c	-	38098.3	Cytoplasmic	hypothetical protein Rv0690c	Conserved hypothetical protein	conserved hypotheticals	COG4427S
Rv0694	-	42129.2	Cytoplasmic Membrane	mycofactocin system heme/flavin oxidoreductase MftD	Possible L-lactate dehydrogenase (cytochrome) LldD1	intermediary metabolism and respiration	COG1304C
Rv0696	-	50666.6	Cytoplasmic	mycofactocin biosynthesis glycosyltransferase MftF	Probable membrane sugar transferase	intermediary metabolism and respiration	COG1215M
Rv0700	rpsJ	11399.2	Cytoplasmic	30S ribosomal protein S10	30S ribosomal protein S10 RpsJ (transcription antitermination factor NusE)	information pathways	COG0051J
Rv0701	rplC	23090.5	Cytoplasmic	50S ribosomal protein L3	50S ribosomal protein L3 RplC	information pathways	COG0087J
Rv0702	rplD	23742.9	Cytoplasmic	50S ribosomal protein L4	50S ribosomal protein L4 RplD	information pathways	COG0088J
Rv0703	rplW	10957.8	Cytoplasmic	50S ribosomal protein L23	50S ribosomal protein L23 RplW	information pathways	COG0089J
Rv0704	rplB	30576.8	Cytoplasmic	50S ribosomal protein L2	50S ribosomal protein L2 RplB	information pathways	COG0090J
Rv0705	rpsS	10803.5	Cytoplasmic	30S ribosomal protein S19	30S ribosomal protein S19 RpsS	information pathways	COG0185J
Rv0706	rplV	20380.5	Cytoplasmic	50S ribosomal protein L22	50S ribosomal protein L22 RplV	information pathways	COG0091J
Rv0707	rpsC	29987.9	Cytoplasmic	30S ribosomal protein S3	30S ribosomal protein S3 RpsC	information pathways	COG0092J
Rv0708	rplP	15692.2	Cytoplasmic	50S ribosomal protein L16	50S ribosomal protein L16 RplP	information pathways	COG0197J
Rv0710	rpsQ	14840.4	Cytoplasmic	30S ribosomal protein S17	30S ribosomal protein S17 RpsQ	information pathways	COG0186J
Rv0711	atsA	86214.5	Cytoplasmic	arylsulfatase AtsA	Possible arylsulfatase AtsA (aryl-sulfate sulphohydrolase) (arylsulphatase)	intermediary metabolism and respiration	COG3119P
Rv0712	-	32711.7	Unknown	hypothetical protein Rv0712	Conserved protein	conserved hypotheticals	COG1262S
Rv0714	rplN	13395.6	Cytoplasmic	50S ribosomal protein L14	50S ribosomal protein L14 RplN	information pathways	COG0093J
Rv0715	rplX	11475.2	Cytoplasmic	50S ribosomal protein L24	50S ribosomal protein L24 RplX	information pathways	COG0198J
Rv0716	rplE	21019.3	Cytoplasmic	50S ribosomal protein L5	50S ribosomal protein L5 RplE	information pathways	COG0094J
Rv0718	rpsH	14411.6	Cytoplasmic	30S ribosomal protein S8	30S ribosomal protein S8 RpsH	information pathways	COG0096J
Rv0719	rplF	19377.3	Cytoplasmic	50S ribosomal protein L6	50S ribosomal protein L6 RplF	information pathways	COG0097J
Rv0720	rplR	13184.1	Cytoplasmic	50S ribosomal protein L18	50S ribosomal protein L18 RplR	information pathways	COG0256J
Rv0721	rpsE	22888.2	Cytoplasmic	30S ribosomal protein S5	30S ribosomal protein S5 RpsE	information pathways	COG0098J
Rv0723	rplO	15501	Unknown	50S ribosomal protein L15	50S ribosomal protein L15 RplO	information pathways	COG0200J
Rv0724	sppA	65935.9	Cytoplasmic Membrane	protease IV SppA	Possible protease IV SppA (endopeptidase IV) (signal peptide peptidase)	cell wall and cell processes	COG0616OU
Rv0726c	-	40875.9	Cytoplasmic	S-adenosylmethionine-dependent methyltransferase	Possible S-adenosylmethionine-dependent methyltransferase	lipid metabolism	COG3315Q
Rv0729	xylB	46543.7	Cytoplasmic	D-xylulose kinase XylB	Possible D-xylulose kinase XylB (xylulokinase) (xylulose kinase)	intermediary metabolism and respiration	COG1070G
Rv0730	-	25980.1	Cytoplasmic	GCN5-like N-acetyltransferase	GCN5-related N-acetyltransferase	intermediary metabolism and respiration	COG0188L

Rv0731c	-	34950	Cytoplasmic	S-adenosylmethionine-dependent methyltransferase	Possible S-adenosylmethionine-dependent methyltransferase	lipid metabolism	COG3315Q
Rv0732	secY	47579.3	Cytoplasmic Membrane	preprotein translocase SecY	Probable preprotein translocase SecY	cell wall and cell processes	COG0201U
Rv0733	adk	20092.8	Cytoplasmic	adenylate kinase	Adenylate kinase Adk (ATP-AMP transphosphorylase)	intermediary metabolism and respiration	COG0563F
Rv0734	mapA	27276.9	Cytoplasmic	methionine aminopeptidase	Methionine aminopeptidase MapA (map) (peptidase M) (MetAP)	intermediary metabolism and respiration	COG0024J
Rv0738	-	19115.7	Unknown	hypothetical protein Rv0738	Conserved protein	conserved hypotheticals	-
Rv0751c	mmsB	29679	Cytoplasmic	3-hydroxyisobutyrate dehydrogenase	Probable 3-hydroxyisobutyrate dehydrogenase MmsB (hibadh)	intermediary metabolism and respiration	COG2084I
Rv0752c	fadE9	41451.2	Cytoplasmic	acyl-CoA dehydrogenase FadE9	Probable acyl-CoA dehydrogenase FadE9	lipid metabolism	COG1960I
Rv0753c	mmsA	54453.9	Cytoplasmic	methylmalonate-semialdehyde dehydrogenase	Probable methylmalonate-semialdehyde dehydrogenase MmsA (methylmalonic acid semialdehyde dehydrogenase) (MMSDH)	intermediary metabolism and respiration	COG1012C
Rv0756c	-	24884.5	Cytoplasmic	hypothetical protein Rv0756c	Unknown protein	conserved hypotheticals	-
Rv0757	phoP	27513.5	Cytoplasmic	two component system response transcriptional positive regulator PhoP	Possible two component system response transcriptional positive regulator PhoP	Regulatory proteins	COG0745TK
Rv0758	phoR	52016.4	Cytoplasmic Membrane	two component system response sensor kinase PhoR	Possible two component system response sensor kinase membrane associated PhoR	Regulatory proteins	COG0642T
Rv0761c	adhB	39715.6	Cytoplasmic	alcohol dehydrogenase B	Possible zinc-containing alcohol dehydrogenase NAD dependent AdhB	intermediary metabolism and respiration	COG1062C
Rv0764c	cyp51	50877.8	Cytoplasmic	lanosterol 14-alpha demethylase	Cytochrome P450 51 Cyp51 (CYPL1) (P450-L1A1) (sterol 14-alpha demethylase) (lanosterol 14-alpha demethylase) (P450-14DM)	intermediary metabolism and respiration	COG2124Q
Rv0769	-	25965.9	Cytoplasmic	oxidoreductase	Probable dehydrogenase/reductase	intermediary metabolism and respiration	COG1028IQR
Rv0772	purD	43476.6	Cytoplasmic	phosphoribosylamine--glycine ligase	Probable phosphoribosylamine--glycine ligase PurD (GARS) (glycinamide ribonucleotide synthetase) (phosphoribosylglycinamide synthetase) (5'-phosphoribosylglycinamide synthetase)	intermediary metabolism and respiration	COG0151F
Rv0775	-	22381.8	Cytoplasmic	hypothetical protein Rv0775	Conserved hypothetical protein	conserved hypotheticals	COG3226S
Rv0777	purB	51008.4	Cytoplasmic	adenylosuccinate lyase PurB	Probable adenylosuccinate lyase PurB (adenylosuccinase) (ASL) (ASASE)	intermediary metabolism and respiration	COG0015F
Rv0778	cyp12	45954	Cytoplasmic	cytochrome P450 Cyp126	Possible cytochrome P450 126 Cyp126	intermediary	COG2124Q

	6		Membrane			metabolism and respiration	
Rv0780	purC	32930.2	Cytoplasmic	phosphoribosylaminoimidazole-succinocarboxamide synthase	Phosphoribosylaminoimidazole- succinocarboxamide synthase PurC (SAICAR synthetase)	intermediary metabolism and respiration	COG0152F
Rv0784	-	25020.1	Cytoplasmic	hypothetical protein Rv0784	Predicted deacetylase	Intermediary metabolism and respiration	COG3233R
Rv0785	-	61219.7	Cytoplasmic	KsdD-like steroid dehydrogenase	Predicted oxidoreductase	Intermediary metabolism and respiration	COG3573R
Rv0786c	-	13917		hypothetical protein Rv0786c	Predicted Zn-dependent hydrolases of the beta-lactamase fold	Intermediary metabolism and respiration	-
Rv0788	purQ	23600.8	Cytoplasmic	phosphoribosylformylglycinamide synthase	Probable phosphoribosylformylglycinamide synthase I PURG (FGAM synthase I)	intermediary metabolism and respiration	COG0047F
Rv0791c	-	37714.8	Cytoplasmic	hypothetical protein Rv0791c	Coenzyme F420-dependent N5,N10-methylene tetrahydromethanopterin reductase and related flavin-dependent oxidoreductases	Intermediary metabolism and respiration	COG2141C
Rv0793	-	11191.7	Cytoplasmic	monooxygenase	Possible monooxygenase	intermediary metabolism and respiration	COG1359S
Rv0794c	-	52476.4	Cytoplasmic	oxidoreductase	Probable oxidoreductase	intermediary metabolism and respiration	COG1249C
Rv0798c	cfp29	28830.4	Unknown	hypothetical protein Rv0798c	29 KDa antigen CFP29	virulence, detoxification, adaptation	COG1659S
Rv0799c	-	36015.7	Cytoplasmic Membrane	hypothetical protein Rv0799c	Predicted iron-dependent peroxidase	Intermediary metabolism and respiration	COG2837P
Rv0800	pepC	46013	Cytoplasmic	M18 family aminopeptidase	Probable aminopeptidase PepC	intermediary metabolism and respiration	COG1362E
Rv0801	-	12512.2	Unknown	hypothetical protein Rv0801	Glyoxalase	Intermediary metabolism and respiration	COG3324R
Rv0802c	-	24983.3	Cytoplasmic Membrane	succinyl-CoA transferase	Possible succinyltransferase in the GCN5-related N-acetyltransferase family	intermediary metabolism and respiration	COG1670J
Rv0803	purL	79475	Cytoplasmic	phosphoribosylformylglycinamide synthase 2	Phosphoribosylformylglycinamide synthase II PurL (FGAM synthase II)	intermediary metabolism and respiration	COG0046F
Rv0806c	cpsY	60268	Cytoplasmic	exopolysaccharide	Possible UDP-glucose-4-epimerase CpsY (galactowaldenase)	cell wall and cell	-

				phosphotransferase CpsY	(UDP-galactose-4-epimerase) (uridine diphosphate galactose-4-epimerase) (uridine diphospho-galactose-4-epimerase)	processes	
Rv0808	purF	56129.6	Cytoplasmic	amidophosphoribosyltransferase	Amidophosphoribosyltransferase PurF (glutamine phosphoribosylpyrophosphate amidotransferase) (ATASE) (gpatase)	intermediary metabolism and respiration	COG0034F
Rv0809	purM	38393.3	Cytoplasmic	phosphoribosylformylglycinamide cyclo-ligase PurM	Probable phosphoribosylformylglycinamide CYCLO-ligase PurM (AIRS) (phosphoribosyl-aminoimidazole synthetase) (air synthase)	intermediary metabolism and respiration	COG0150F
Rv0811c	-	39414.8	Cytoplasmic	hypothetical protein Rv0811c	Predicted aminomethyltransferase related to GcvT	Intermediary metabolism and respiration	COG0354R
Rv0812	-	31090.6	Cytoplasmic	amino acid aminotransferase	Probable amino acid aminotransferase	intermediary metabolism and respiration	COG0115EH
Rv0813c	-	23868.3	Cytoplasmic Membrane	hypothetical protein Rv0813c	UPF0678 fatty acid-binding protein-like protein	Lipid metabolism	-
Rv0814c	-	10165.3	Unknown	hypothetical protein Rv0814c	Conserved protein SseC2	intermediary metabolism and respiration	-
Rv0815c	cysA2	31014.6	Cytoplasmic	thiosulfate sulfurtransferase CysA	Probable thiosulfate sulfurtransferase CysA2 (rhodanese-like protein) (thiosulfate cyanide transsulfurase) (thiosulfate thiotransferase)	intermediary metabolism and respiration	COG2897P
Rv0817c	-	28567.7	Cytoplasmic	exported protein	Probable conserved exported protein	cell wall and cell processes	-
Rv0818	-	27579.4	Cytoplasmic	transcriptional regulator	Transcriptional regulatory protein	Regulatory proteins	COG0745TK
Rv0819	mshD	33567.1	Cytoplasmic Membrane	mycothiol acetyltransferase	GCN5-related N-acetyltransferase, MshD	intermediary metabolism and respiration	COG0456R
Rv0820	phoT	28029.1	Cytoplasmic Membrane	phosphate ABC transporter ATP-binding protein PhoT	Probable phosphate-transport ATP-binding protein ABC transporter PhoT	cell wall and cell processes	COG1117P
Rv0821c	phoY2	23513.8	Cytoplasmic	phosphate-transport system transcriptional regulator PhoY2	Probable phosphate-transport system transcriptional regulatory protein PhoY2	cell wall and cell processes	COG0704P
Rv0823c	-	41388.4	Cytoplasmic	tRNA-dihydrouridine synthase	Possible transcriptional regulatory protein	Regulatory proteins	COG0042J
Rv0824c	desA1	38770.2	Cytoplasmic	acyl-ACP desaturase DesA	Probable acyl-[acyl-carrier protein] desaturase DesA1 (acyl-[ACP] desaturase) (stearyl-ACP desaturase) (protein Des)	lipid metabolism	-
Rv0826	-	40249	Cytoplasmic	hypothetical protein Rv0826	Conserved hypothetical protein	conserved hypotheticals	COG3662S
Rv0830	-	33416.5	Cytoplasmic	S-adenosylmethionine-dependent methyltransferase	Possible S-adenosylmethionine-dependent methyltransferase	lipid metabolism	COG3315Q
Rv0831c	-	30189.2	Unknown	hypothetical protein Rv0831c	Conserved protein	conserved hypotheticals	-
Rv0839	-	29608.3	Cytoplasmic	hypothetical protein Rv0839	SAM-dependent methyltransferases	Intermediary metabolism and respiration	COG2226H
Rv0846c	-	53795.9	Unknown	oxidase	Probable oxidase	intermediary	COG2132Q

						metabolism and respiration	
Rv0853c	pdc	59751.1	Cytoplasmic Membrane	alpha-keto-acid decarboxylase	Probable pyruvate or indole-3-pyruvate decarboxylase Pdc	intermediary metabolism and respiration	COG3961GHR
Rv0854	-	16345.6	Cytoplasmic	hypothetical protein Rv0854	Cyclase/Dehydrase	Intermediary metabolism and respiration	-
Rv0857	-	17617.1	Cytoplasmic	hypothetical protein Rv0857	link to Polyketide_cyclase/dehydratase.	Lipid metabolism	-
Rv0858c	dapC	42209.1	Cytoplasmic	N-succinyldiaminopimelate aminotransferase DapC	Probable N-succinyldiaminopimelate aminotransferase DapC (DAP-at)	intermediary metabolism and respiration	COG0436E
Rv0859	-	42414.5	Cytoplasmic	acyltransferase	Possible acyl-CoA thiolase FadA	lipid metabolism	COG0183I
Rv0860	fadB	76104.1	Cytoplasmic	fatty oxidation protein FadB	Probable fatty oxidation protein FadB	lipid metabolism	COG1250I
Rv0861c	ercc3	59772.2	Cytoplasmic	DNA helicase Ercc3	DNA helicase Ercc3	information pathways	-
Rv0862c	-	79666.5	Cytoplasmic	hypothetical protein Rv0862c	Conserved protein	conserved hypotheticals	COG2378K
Rv0865	mog	16188.3	Unknown	molybdopterin biosynthesis protein	Probable molybdopterin biosynthesis Mog protein	intermediary metabolism and respiration	COG0521H
Rv0866	moaE2	15048.1	Unknown	molybdopterin synthase catalytic subunit 2	Probable molybdenum cofactor biosynthesis protein E2 MoaE2 (molybdopterin converting factor large subunit) (molybdopterin [MPT] converting factor, subunit 2)	intermediary metabolism and respiration	COG0314H
Rv0870c	-	13687.8	Cytoplasmic Membrane	integral membrane protein	Possible conserved integral membrane protein	cell wall and cell processes	COG3304S
Rv0871	cspB	14989	Cytoplasmic	cold shock-like protein CspB	Probable cold shock-like protein B CspB	virulence, detoxification, adaptation	COG1278K
Rv0873	fadE10	70744.6	Cytoplasmic	acyl-CoA dehydrogenase FadE10	Probable acyl-CoA dehydrogenase FadE10	lipid metabolism	COG1960I
Rv0877	-	27437.4	Cytoplasmic	hypothetical protein Rv0877	cell wall processes	cell wall and cell processes	-
Rv0880	-	15543.9	Cytoplasmic	HTH-type transcriptional regulator	Possible transcriptional regulatory protein (possibly MarR-family)	Regulatory proteins	COG1846K
Rv0884c	serC	40233.4	Unknown	phosphoserine aminotransferase	Possible phosphoserine aminotransferase SerC (PSAT)	intermediary metabolism and respiration	COG1932HE
Rv0885	-	39798	Cytoplasmic Membrane	hypothetical protein Rv0885	Conserved hypothetical protein	conserved hypotheticals	-
Rv0886	fprB	61337.2	Cytoplasmic	ferredoxin/ferredoxin--NADP reductase	Probable NADPH:adrenodoxin oxidoreductase FprB (adrenodoxin reductase) (AR) (ferredoxin-NADP(+) reductase)	intermediary metabolism and respiration	COG0493ER
Rv0887c	-	16098.2	Unknown	hypothetical protein Rv0887c	Conserved hypothetical protein	conserved hypotheticals	COG2764S

Rv0889c	citA	40146.8	Cytoplasmic	citrate synthase 2	Probable citrate synthase II CitA	intermediary metabolism and respiration	COG0372C
Rv0896	gltA2	47946.4	Cytoplasmic	citrate synthase 1	Probable citrate synthase I GltA2	intermediary metabolism and respiration	COG0372C
Rv0897c	-	56233.9	Unknown	oxidoreductase	Probable oxidoreductase	intermediary metabolism and respiration	COG1233Q
Rv0899	arfA	33542.1	Cytoplasmic Membrane	peptidoglycan-binding protein ArfA	Outer membrane protein A OmpA	cell wall and cell processes	COG2885M
Rv0902c	prfB	47829.3	Cytoplasmic Membrane	two component sensor histidine kinase PrrB	Two component sensor histidine kinase PrrB	Regulatory proteins	COG0642T
Rv0903c	prfA	25253.9	Cytoplasmic	two component transcriptional regulator PrrA	Two component response transcriptional regulatory protein PrrA	Regulatory proteins	COG0745TK
Rv0904c	accD3	51740.3	Cytoplasmic	acetyl-CoA carboxylase carboxyl transferase subunit beta	Putative acetyl-coenzyme A carboxylase carboxyl transferase (subunit beta) AccD3 (accase beta chain)	lipid metabolism	COG0825I
Rv0905	echA6	26028.8	Cytoplasmic	enoyl-CoA hydratase EchA6	Possible enoyl-CoA hydratase EchA6 (enoyl hydratase) (unsaturated acyl-CoA hydratase) (crotonase)	lipid metabolism	COG1024I
Rv0906	-	40641.2	Unknown	hypothetical protein Rv0906	Predicted Zn-dependent hydrolases of the beta-lactamase fold	Intermediary metabolism and respiration	COG2220R
Rv0910	-	15754.5	Unknown	toxin	Conserved hypothetical protein	conserved hypotheticals	COG3427S
Rv0911	-	27594.7	Cytoplasmic	hypothetical protein Rv0911	Predicted enzyme related to lactoylglutathione lyase	Intermediary metabolism and respiration	COG3324R
Rv0913c	-	56639.1	Unknown	dioxygenase	Possible dioxygenase	intermediary metabolism and respiration	COG3670Q
Rv0914c	-	43874.9	Cytoplasmic	lipid carrier protein or keto acyl-CoA thiolase	Possible lipid carrier protein or keto acyl-CoA thiolase	lipid metabolism	COG0183I
Rv0925c	-	26944.3	Unknown	hypothetical protein Rv0925c	Multimeric flavodoxin WrbA	Intermediary metabolism and respiration	COG0655R
Rv0926c	-	37751.3	Cytoplasmic	hypothetical protein Rv0926c	Uncharacterized conserved protein related to dihydrodipicolinate reductase	Intermediary metabolism and respiration	COG3804S
Rv0927c	-	26745.7	Cytoplasmic	oxidoreductase	Probable short-chain type dehydrogenase/reductase	intermediary metabolism and respiration	COG1028IQR
Rv0928	pstS3	37935.1	Unknown	phosphate ABC transporter substrate-binding lipoprotein PstS	Periplasmic phosphate-binding lipoprotein PstS3 (PBP-3) (PstS3) (PHOS1)	cell wall and cell processes	COG0226P
Rv0931c	pknD	69513.2	Cytoplasmic Membrane	serine/threonine-protein kinase PknD	Transmembrane serine/threonine-protein kinase D PknD (protein kinase D) (STPK D)	Regulatory proteins	COG0515R TKL

Rv0932c	pstS2	37832	Extracellular	phosphate ABC transporter substrate-binding lipoprotein PstS	Periplasmic phosphate-binding lipoprotein PstS2 (PBP-2) (PstS2)	cell wall and cell processes	COG0226P
Rv0933	pstB	29996.4	Cytoplasmic Membrane	phosphate ABC transporter ATP-binding protein PstB	Phosphate-transport ATP-binding protein ABC transporter PstB	cell wall and cell processes	COG1117P
Rv0934	pstS1	38211.1	Unknown	phosphate ABC transporter substrate-binding lipoprotein PstS	Periplasmic phosphate-binding lipoprotein PstS1 (PBP-1) (PstS1)	cell wall and cell processes	COG0226P
Rv0937c	mku	30904	Cytoplasmic	non-homologous end joining protein Ku	DNA end-binding protein, Mku	information pathways	COG1273S
Rv0938	ligD	83572.4	Cytoplasmic	multifunctional non-homologous end joining DNA repair protein/ATP dependent DNA ligase LigD	ATP dependent DNA ligase LigD (ATP dependent polydeoxyribonucleotide synthase) (thermostable DNA ligase) (ATP dependent polynucleotide ligase) (sealase) (DNA repair enzyme) (DNA joinase)	information pathways	COG3285L
Rv0939	-	69353.6	Cytoplasmic	bifunctional 2-hydroxyhepta-2,4-diene-1,7-dioate isomerase/cyclase/dehydrase	Possible bifunctional enzyme: 2-hydroxyhepta-2,4-diene-1,7-dioate isomerase (HHDD isomerase) + cyclase/dehydrase	intermediary metabolism and respiration	COG0179Q
Rv0940c	-	31756.5	Cytoplasmic	oxidoreductase	Possible oxidoreductase	intermediary metabolism and respiration	COG2141C
Rv0945	-	27138.6	Cytoplasmic	oxidoreductase	Probable short-chain type dehydrogenase/reductase	intermediary metabolism and respiration	COG0300R
Rv0946c	pgi	59974.4	Cytoplasmic	glucose-6-phosphate isomerase	Probable glucose-6-phosphate isomerase Pgi (GPI) (phosphoglucose isomerase) (phosphohexose isomerase) (phi)	intermediary metabolism and respiration	COG0166G
Rv0949	uvrD1	85049.9	Cytoplasmic	ATP-dependent DNA helicase UvrD	Probable ATP-dependent DNA helicase II UvrD1	information pathways	COG0210L
Rv0951	sucC	40925.7	Cytoplasmic	succinyl-CoA ligase subunit beta	Probable succinyl-CoA synthetase (beta chain) SucC (SCS-beta)	intermediary metabolism and respiration	COG0045C
Rv0952	sucD	31228.9	Unknown	succinyl-CoA ligase subunit alpha	Probable succinyl-CoA synthetase (alpha chain) SucD (SCS-alpha)	intermediary metabolism and respiration	COG0074C
Rv0956	purN	22383.7	Unknown	phosphoribosylglycinamide formyltransferase PurN	Probable 5'-phosphoribosylglycinamide formyltransferase PurN (GART) (gar transformylase) (5'-phosphoribosylglycinamide transformylase)	intermediary metabolism and respiration	COG0299F
Rv0957	purH	55026.2	Cytoplasmic	bifunctional phosphoribosylaminoimidazolecarboxamide formyltransferase/inosinemonophosphate cyclohydrolase	Probable bifunctional purine biosynthesis protein PurH: phosphoribosylaminoimidazolecarboxamide formyltransferase (AICAR transformylase) (5'-phosphoribosyl-5-aminoimidazole-4-carboxamide formyltransferase) + inosinemonophosphate cyclohydrolase (imp cyclohydrolase) (inosinase) (imp synthetase) (ATIC)	intermediary metabolism and respiration	COG0138F
Rv0958	-	49581.4	Cytoplasmic	magnesium chelatase	Possible magnesium chelatase	intermediary metabolism and respiration	COG1239H
Rv0959	-	74635.7	Cytoplasmic	hypothetical protein Rv0959	Uncharacterized protein with a von Willebrand factor type A	Regulatory proteins	COG4867R

					(vWA) domain		
Rv0966c	-	22210.2	Unknown	hypothetical protein Rv0966c	Conserved protein	conserved hypotheticals	-
Rv0968	-	10263.8	Unknown	hypothetical protein Rv0968	Conserved protein	conserved hypotheticals	-
Rv0969	ctpV	80102.5	Cytoplasmic Membrane	copper transporter ATPase V	Probable metal cation transporter P-type ATPase CtpV	cell wall and cell processes	-
Rv0971c	echA7	27604.5	Unknown	enoyl-CoA hydratase EchA7	Probable enoyl-CoA hydratase EchA7 (enoyl hydratase) (unsaturated acyl-CoA hydratase) (crotonase)	lipid metabolism	COG1024I
Rv0972c	fadE1 ₂	41483.3	Cytoplasmic	acyl-CoA dehydrogenase FadE12	Acyl-CoA dehydrogenase FadE12	lipid metabolism	COG1960I
Rv0973c	accA2	70744.8	Cytoplasmic	acetyl-/propionyl-CoA carboxylase subunit alpha	Probable acetyl-/propionyl-coenzyme A carboxylase alpha chain (alpha subunit) AccA2: biotin carboxylase + biotin carboxyl carrier protein (BCCP)	lipid metabolism	COG4770I
Rv0974c	accD2	56224.3	Cytoplasmic Membrane	acetyl-/propionyl-CoA carboxylase subunit beta	Probable acetyl-/propionyl-CoA carboxylase (beta subunit) AccD2	lipid metabolism	COG4799I
Rv0975c	fadE1 ₃	40715.1	Cytoplasmic	acyl-CoA dehydrogenase FadE13	Probable acyl-CoA dehydrogenase FadE13	lipid metabolism	COG1960I
Rv0976c	-	59091	Cytoplasmic	hypothetical protein Rv0976c	related to synthesis and oxidation of fatty acids	Lipid metabolism	-
Rv0981	mprA	25675.4	Cytoplasmic	two-component response regulator MrpA	Mycobacterial persistence regulator MRPA (two component response transcriptional regulatory protein)	Regulatory proteins	-
Rv0982	mprB	54407.7	Cytoplasmic Membrane	two component histidine-protein kinase/phosphatase MprB	Two component sensor kinase MprB	Regulatory proteins	COG0642T
Rv0983	pepD	46452.6	Unknown	serine protease PepD	Probable serine protease PepD (serine proteinase) (MTB32B)	intermediary metabolism and respiration	COG0265O
Rv0984	moaB ₂	18441	Cytoplasmic	pterin-4-alpha-carbinolamine dehydratase	Possible pterin-4-alpha-carbinolamine dehydratase MoaB2 (PHS) (4-alpha-hydroxy-tetrahydropterin dehydratase) (pterin-4-alpha-carbinolamine dehydratase) (phenylalanine hydroxylase-stimulating protein) (PHS) (pterin carbinolamine dehydratase) (PCD)	intermediary metabolism and respiration	COG0521H
Rv0985c	mscL	16023.5	Cytoplasmic Membrane	large-conductance ion mechanosensitive channel	Possible large-conductance ion mechanosensitive channel MscL	cell wall and cell processes	COG1970M
Rv0986	-	27373.1	Cytoplasmic Membrane	adhesion component ABC transporter ATP-binding protein	Probable adhesion component transport ATP-binding protein ABC transporter	cell wall and cell processes	COG1136V
Rv0993	galU	32378.2	Cytoplasmic	UTP--glucose-1-phosphate uridylyltransferase	UTP--glucose-1-phosphate uridylyltransferase GalU (UDP-glucose pyrophosphorylase) (UDPGP) (alpha-D-glucosyl-1-phosphate uridylyltransferase) (uridine diphosphoglucose pyrophosphorylase)	intermediary metabolism and respiration	COG1210M
Rv0994	moeA ₁	44304.8	Cytoplasmic	molybdopterin molybdenumtransferase 1	Probable molybdopterin biosynthesis protein MoeA1	intermediary metabolism and respiration	COG0303H
Rv0998	-	35607.8	Cytoplasmic	acetyltransferase	cAMP-binding proteins - catabolite gene activator and regulatory subunit of cAMP-dependent protein kinases	Regulatory proteins	COG0664T
Rv0999	-	25937.6	Unknown	hypothetical protein Rv0999	Unknown protein	conserved	-

						hypotheticals	
Rv1001	arcA	43057.1	Cytoplasmic	arginine deiminase	Probable arginine deiminase ArcA (adi) (ad) (arginine dihydrolase)	intermediary metabolism and respiration	COG2235E
Rv1006	-	61297.6	Unknown	hypothetical protein Rv1006	Unknown protein	conserved hypotheticals	-
Rv1007c	metS	58092.6	Cytoplasmic	methionine--tRNA ligase	Methionyl-tRNA synthetase MetS (MetRS) (methionine--tRNA ligase)	information pathways	COG0143J
Rv1008	tatD	29065.1	Cytoplasmic	deoxyribonuclease TatD	Probable deoxyribonuclease TatD (YJIV protein)	cell wall and cell processes	COG0084L
Rv1011	ispE	31349.6	Unknown	4-diphosphocytidyl-2-C-methyl-D-erythritol kinase	Probable 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase IspE (CMK) (4-(cytidine-5'-diphospho)-2-C-methyl-D-erythritol kinase)	intermediary metabolism and respiration	COG1947I
Rv1013	pks16	58449.3	Cytoplasmic Membrane	polyketide synthase	Putative polyketide synthase Pks16	lipid metabolism	COG0318IQ
Rv1015c	rplY	22441.2	Cytoplasmic	50S ribosomal protein L25	50S ribosomal protein L25 RplY	information pathways	COG1825J
Rv1017c	prsA	35459.4	Cytoplasmic	ribose-phosphate pyrophosphokinase	Probable ribose-phosphate pyrophosphokinase PrsA (phosphoribosyl pyrophosphate synthetase) (PRPP synthetase)	intermediary metabolism and respiration	COG0462FE
Rv1018c	glmU	51584	Cytoplasmic	bifunctional UDP-N-acetylglucosamine pyrophosphorylase/glucosamine-1-phosphate N-acetyltransferase	Probable UDP-N-acetylglucosamine pyrophosphorylase GlmU	cell wall and cell processes	COG1207M
Rv1019	-	21701.5	Unknown	transcriptional regulator	Probable transcriptional regulatory protein (probably TetR-family)	Regulatory proteins	COG1309K
Rv1020	mfd	132908	Cytoplasmic	transcription-repair coupling factor	Probable transcription-repair coupling factor Mfd (TRCF)	information pathways	COG1197LK
Rv1021	-	35413.2	Cytoplasmic	nucleoside triphosphate pyrophosphohydrolase	Predicted pyrophosphatase	Intermediary metabolism and respiration	COG3956R
Rv1023	eno	44929.5	Cytoplasmic	enolase	Probable enolase Eno	intermediary metabolism and respiration	COG0148G
Rv1030	kdpB	74601.5	Cytoplasmic Membrane	potassium-transporting ATPase subunit B	Probable potassium-transporting P-type ATPase B chain KdpB (potassium-translocating ATPase B chain) (ATP phosphohydrolase [potassium-transporting] B chain) (potassium binding and translocating subunit B)	cell wall and cell processes	COG2216P
Rv1037c	esxI	9832.78	Unknown	ESAT-6 like protein EsxI	Putative ESAT-6 like protein EsxI (ESAT-6 like protein 1)	cell wall and cell processes	-
Rv1056	-	28945.6	Cytoplasmic	hypothetical protein Rv1056	Conserved protein	conserved hypotheticals	COG2343S
Rv1058	fadD14	59624.6	Cytoplasmic	fatty-acid--CoA ligase FadD14	Probable medium chain fatty-acid-CoA ligase FadD14 (fatty-acid-CoA synthetase) (fatty-acid-CoA synthase)	lipid metabolism	COG0318IQ
Rv1059	-	36878.2	Cytoplasmic	hypothetical protein Rv1059	Uncharacterized conserved protein related to dihydrodipicolinate reductase	Intermediary metabolism and	COG3804S

						respiration	
Rv1060	-	16866.5	Cytoplasmic Membrane	hypothetical protein Rv1060	Unknown protein	conserved hypotheticals	-
Rv1061	-	32037.8	Cytoplasmic	hypothetical protein Rv1061	Predicted glutamine amidotransferase	Intermediary metabolism and respiration	COG0121R
Rv1063c	-	37522.7	Cytoplasmic	NTE family protein	Predicted esterase of the alpha-beta hydrolase superfamily	Intermediary metabolism and respiration	COG1752R
Rv1069c	-	64341.5	Cytoplasmic Membrane	hypothetical protein Rv1069c	Predicted membrane protein	cell wall and cell processes	COG4425S
Rv1070c	echA8	27273.1	Cytoplasmic Membrane	enoyl-CoA hydratase EchA8	Probable enoyl-CoA hydratase EchA8 (enoyl hydratase) (unsaturated acyl-CoA hydratase) (crotonase)	lipid metabolism	COG1024I
Rv1071c	echA9	36322	Cytoplasmic	enoyl-CoA hydratase EchA9	Possible enoyl-CoA hydratase EchA9 (enoyl hydratase) (unsaturated acyl-CoA hydratase) (crotonase)	lipid metabolism	COG1024I
Rv1072	-	29111.4	Cytoplasmic Membrane	transmembrane protein	Probable conserved transmembrane protein	cell wall and cell processes	COG4760S
Rv1074c	fadA3	42655.4	Cytoplasmic	beta-ketoacyl CoA thiolase FadA	Probable beta-ketoacyl CoA thiolase FadA3	lipid metabolism	COG0183I
Rv1077	cbs	48635.2	Cytoplasmic	cystathionine beta-synthase	Probable cystathionine beta-synthase Cbs (serine sulfhydrase) (beta-thionase) (hemoprotein H-450)	intermediary metabolism and respiration	COG0031E
Rv1078	-	25076	Cytoplasmic Membrane	hypothetical protein Rv1078	cell wall and cell processes , Predicted membrane protein/domain, proline-rich antigen homolog	cell wall and cell processes	COG5178A
Rv1079	metB	40982.3	Cytoplasmic	cystathionine gamma-synthase	Cystathionine gamma-synthase MetB (CGS) (O-succinylhomoserine [thiol]-lyase)	intermediary metabolism and respiration	COG0626E
Rv1080c	greA	17854.7	Cytoplasmic	transcription elongation factor GreA	Probable transcription elongation factor GreA (transcript cleavage factor GreA)	information pathways	COG0782K
Rv1082	mca	32699.7	Cytoplasmic	mycothiol S-conjugate amidase	Mycothioli conjugate amidase Mca (mycothiol S-conjugate amidase)	virulence, detoxification, adaptation	COG2120S
Rv1084	-	71050.3	Cytoplasmic Membrane	hypothetical protein Rv1084	Highly conserved protein containing a thioredoxin domain	Intermediary metabolism and respiration	COG1331O
Rv1092c	coaA	35656.9	Cytoplasmic	pantothenate kinase	Probable pantothenate kinase CoaA (pantothenic acid kinase)	intermediary metabolism and respiration	COG1072H
Rv1093	glyA1	45029.4	Cytoplasmic	serine hydroxymethyltransferase	Serine hydroxymethyltransferase 1 GlyA1	intermediary metabolism and respiration	-
Rv1094	desA2	31359.5	Cytoplasmic	acyl-ACP desaturase DesA	Possible acyl-[acyl-carrier protein] desaturase DesA2 (acyl-[ACP] desaturase) (stearyl-ACP desaturase)	lipid metabolism	COG0208F
Rv1095	phoH2	46881.5	Cytoplasmic	phosphate starvation-inducible protein PsiH	Probable PHOH-like protein PhoH2 (phosphate starvation-inducible protein PSIH)	intermediary metabolism and respiration	COG1875T

Rv1096	-	31077.1	Unknown	glycosyl hydrolase	Possible glycosyl hydrolase	intermediary metabolism and respiration	COG0726G
Rv1097c	-	29873.6	Unknown	membrane glycine/proline-rich protein	Probable membrane glycine and proline rich protein	cell wall and cell processes	COG5178A
Rv1098c	fum	50141	Cytoplasmic	fumarate hydratase	Probable fumarase Fum (fumarate hydratase)	intermediary metabolism and respiration	COG0114C
Rv1099c	glpX	38084	Cytoplasmic	fructose 1,6-bisphosphatase	Fructose 1,6-bisphosphatase GlpX	intermediary metabolism and respiration	-
Rv1101c	-	41698.6	Cytoplasmic Membrane	hypothetical protein Rv1101c	Conserved membrane protein	cell wall and cell processes	COG0628R
Rv1106c	-	40741.6	Unknown	3 beta-hydroxysteroid dehydrogenase/delta 5--_4-isomerase	3-beta-hydroxysteroid dehydrogenase	intermediary metabolism and respiration	COG0451MG
Rv1108c	xseA	44596.3	Cytoplasmic	exodeoxyribonuclease VII large subunit	Probable exodeoxyribonuclease VII (large subunit) XseA (exonuclease VII large subunit)	information pathways	COG1570L
Rv1109c	-	22957.2	Unknown	hypothetical protein Rv1109c	Conserved protein	conserved hypotheticals	-
Rv1110	lytB2	36297.1	Cytoplasmic	4-hydroxy-3-methylbut-2-enyl diphosphate reductase	Probable LYTB-related protein LytB2	cell wall and cell processes	COG0761IM
Rv1112	-	38065.2	Cytoplasmic	GTP-binding protein	Predicted GTPase, probable translation factor	Information pathways	COG0012J
Rv1118c	-	32157.5	Cytoplasmic	hypothetical protein Rv1118c	Conserved protein	conserved hypotheticals	-
Rv1121	zwf1	52155.3	Cytoplasmic	glucose-6-phosphate 1-dehydrogenase	Probable glucose-6-phosphate 1-dehydrogenase Zwf1 (G6PD)	intermediary metabolism and respiration	COG0364G
Rv1122	gnd2	36358.1	Cytoplasmic	6-phosphogluconate dehydrogenase	Probable 6-phosphogluconate dehydrogenase, decarboxylating Gnd2	intermediary metabolism and respiration	COG1023G
Rv1124	ephC	34694.6	Cytoplasmic	epoxide hydrolase EphC	Probable epoxide hydrolase EphC (epoxide hydratase)	virulence, detoxification, adaptation	COG0596R
Rv1126c	-	22197.3	Cytoplasmic	hypothetical protein Rv1126c	Conserved protein	conserved hypotheticals	-
Rv1129c	-	54087.4	Cytoplasmic	transcriptional regulator	Probable transcriptional regulator protein	Regulatory proteins	COG3800R
Rv1130	prpD	57750.6	Cytoplasmic	methylcitrate dehydratase PrpD	Possible methylcitrate dehydratase PrpD	intermediary metabolism and respiration	COG2079R
Rv1131	prpC	42969.4	Cytoplasmic	methylcitrate synthase PrpC	Probable methylcitrate synthase PrpC	intermediary metabolism and respiration	COG0372C
Rv1132	-	64019.7	Cytoplasmic Membrane	hypothetical protein Rv1132	Conserved membrane protein	cell wall and cell processes	-

Rv1133c	metE	81549.5	Cytoplasmic	5-methyltetrahydropteroyltri-glutamate--homocysteine methyltransferase	Probable 5-methyltetrahydropteroyltri-glutamate--homocysteine methyltransferase MetE (methionine synthase, vitamin-B12 independent isozyme)	intermediary metabolism and respiration	COG0620E
Rv1138c	-	36345.6	Cytoplasmic Membrane	oxidoreductase	Possible oxidoreductase	intermediary metabolism and respiration	COG0644C
Rv1141c	echA11	27615.8	Cytoplasmic Membrane	enoyl-CoA hydratase EchA11	Probable enoyl-CoA hydratase EchA11 (enoyl hydratase) (unsaturated acyl-CoA hydratase) (crotonase)	lipid metabolism	COG1024I
Rv1142c	echA10	27648.9	Unknown	enoyl-CoA hydratase EchA10	Probable enoyl-CoA hydratase EchA10 (enoyl hydratase) (unsaturated acyl-CoA hydratase) (crotonase)	lipid metabolism	COG1024I
Rv1143	mcr	38684.9	Cytoplasmic	alpha-methylacyl-CoA racemase	Probable alpha-methylacyl-CoA racemase Mcr (2-methylacyl-CoA racemase) (2-arylpropionyl-CoA epimerase)	lipid metabolism	COG1804C
Rv1144	-	25787.8	Cytoplasmic	oxidoreductase	Probable short-chain type dehydrogenase/reductase	intermediary metabolism and respiration	COG1028IQR
Rv1151c	-	25661	Cytoplasmic Membrane	NAD-dependent protein deacylase	Transcriptional regulatory protein	Regulatory proteins	COG0846K
Rv1153c	omt	30854.9	Cytoplasmic Membrane	O-methyltransferase	Probable O-methyltransferase Omt	lipid metabolism	COG3315Q
Rv1155	-	16300.5	Cytoplasmic Membrane	pyridoxine/pyridoxamine 5'-phosphate oxidase	Possible pyridoxamine 5'-phosphate oxidase (PNP/PMP oxidase) (pyridoxinephosphate oxidase) (PNPOX) (pyridoxine 5'-phosphate oxidase)	intermediary metabolism and respiration	COG3871R
Rv1156	-	20839.9	Cytoplasmic	hypothetical protein Rv1156	Base excision DNA repair protein	Information pathways	COG0177L
Rv1161	narG	136924	Cytoplasmic Membrane	nitrate reductase subunit alpha	Respiratory nitrate reductase (alpha chain) NarG	intermediary metabolism and respiration	COG5013C
Rv1162	narH	62037.8	Cytoplasmic Membrane	nitrate reductase subunit beta	Probable respiratory nitrate reductase (beta chain) NarH	intermediary metabolism and respiration	COG1140C
Rv1165	typA	67548.2	Cytoplasmic	GTP-binding translation elongation factor	Possible GTP-binding translation elongation factor TypA (tyrosine phosphorylated protein A) (GTP-binding protein)	information pathways	COG1217T
Rv1166	lpqW	66130.2	Unknown	monoacyl phosphatidylinositol tetramannoside-binding protein LpqW	Probable conserved lipoprotein LpqW	cell wall and cell processes	COG0747E
Rv1170	mshB	31742.3	Unknown	1D-myo-inositol 2-acetamido-2-deoxy-alpha-D-glucopyranoside deacetylase	N-acetyl-1-D-myo-inositol-2-amino-2-deoxy-alpha-D-glucopyranoside deacetylase MshB (GlcNAc-Ins deacetylase)	intermediary metabolism and respiration	COG2120S
Rv1173	fbiC	92474.2	Cytoplasmic	FO synthase	Probable F420 biosynthesis protein FbiC	intermediary metabolism and respiration	COG1060HR
Rv1175c	fadH	72901.5	Cytoplasmic	NADPH dependent 2,4-dienoyl-CoA reductase FadH	Probable NADPH dependent 2,4-dienoyl-CoA reductase FadH (2,4-dienoyl coenzyme A reductase) (4-enoyl-CoA reductase)	lipid metabolism	COG1902C
Rv1178	-	37702.2	Unknown	aminotransferase	Probable aminotransferase	intermediary metabolism and	COG0436E

						respiration	
Rv1179c	-	100642	Cytoplasmic Membrane	hypothetical protein Rv1179c	DNA or RNA helicases of superfamily II	Information pathways	COG1061KL
Rv1180	pks3	50566.5	Cytoplasmic Membrane	polyketide beta-ketoacyl synthase	Probable polyketide beta-ketoacyl synthase Pks3	lipid metabolism	COG3321Q
Rv1181	pks4	168084	Cytoplasmic Membrane	polyketide beta-ketoacyl synthase	Probable polyketide beta-ketoacyl synthase Pks4	lipid metabolism	COG3321Q
Rv1185c	fadD21	62756.9	Cytoplasmic	fatty-acid--CoA ligase FadD21	Probable fatty-acid-AMP ligase FadD21 (fatty-acid-AMP synthetase) (fatty-acid-AMP synthase)	lipid metabolism	COG0318IQ
Rv1187	rocA	58841.4	Cytoplasmic	pyrroline-5-carboxylate dehydrogenase RocA	Probable pyrroline-5-carboxylate dehydrogenase RocA	intermediary metabolism and respiration	COG1012C
Rv1188	-	36237.1	Cytoplasmic	proline dehydrogenase	Probable proline dehydrogenase	intermediary metabolism and respiration	COG0506E
Rv1190	-	31804.4	Unknown	hypothetical protein Rv1190	Predicted hydrolases or acyltransferases (alpha/beta hydrolase superfamily)	Intermediary metabolism and respiration	COG0596R
Rv1191	-	32611.6	Cytoplasmic Membrane	hypothetical protein Rv1191	Predicted hydrolases or acyltransferases (alpha/beta hydrolase superfamily)	Intermediary metabolism and respiration	COG0596R
Rv1192	-	30212.8	Cytoplasmic	hypothetical protein Rv1192	Protein involved in GPI anchor metabolic process	cell wall and cell processes	COG0596R
Rv1193	fadD36	49488.3	Cytoplasmic	fatty-acid--CoA ligase FadD36	Probable fatty-acid-CoA ligase FadD36 (fatty-acid-CoA synthetase) (fatty-acid-CoA synthase)	lipid metabolism	COG0318IQ
Rv1194c	-	46066	Cytoplasmic	hypothetical protein Rv1194c	Regulator of polyketide synthase expression	Lipid metabolism	COG2508TQ
Rv1196	PPE18	39158.4	Cytoplasmic Membrane	PPE family protein PPE18	PPE family protein PPE18	PE/PPE	COG5651N
Rv1198	esxL	9927.97	Unknown	ESAT-6 like protein EsxL	Putative ESAT-6 like protein EsxL (ESAT-6 like protein 4)	cell wall and cell processes	-
Rv1201c	dapD	32609.8	Unknown	2,3,4,5-tetrahydropyridine-2,6-dicarboxylate N-succinyltransferase	Tetrahydrodipicolinate N-succinyltransferase DapD	intermediary metabolism and respiration	COG2171E
Rv1202	dapE	37241	Cytoplasmic	succinyl-diaminopimelate desuccinylase DapE	Probable succinyl-diaminopimelate desuccinylase DapE	intermediary metabolism and respiration	COG0624E
Rv1206	fadD6	64305	Cytoplasmic	fatty-acid--CoA ligase FadD6	Probable fatty-acid-CoA ligase FadD6 (fatty-acid-CoA synthetase) (fatty-acid-CoA synthase)	lipid metabolism	COG0318IQ
Rv1207	folP2	33013.7	Cytoplasmic	dihydropteroate synthase	Dihydropteroate synthase 2 FolP2 (DHPS 2) (dihydropteroate pyrophosphorylase 2)	intermediary metabolism and respiration	COG0294H
Rv1213	glgC	43800.4	Unknown	glucose-1-phosphate adenylyltransferase	Glucose-1-phosphate adenylyltransferase GlgC (ADP-glucose synthase) (ADP-glucose pyrophosphorylase)	intermediary metabolism and respiration	COG0448G
Rv1215c	-	62578.9	Cytoplasmic	hypothetical protein Rv1215c	Predicted acyl esterases	Intermediary	COG2936R

						metabolism and respiration	
Rv1218c	-	33448.1	Cytoplasmic Membrane	tetronasin ABC transporter ATP-binding protein	Probable tetronasin-transport ATP-binding protein ABC transporter	cell wall and cell processes	COG1131V
Rv1220c	-	22107.1	Cytoplasmic	methyltransferase	Probable methyltransferase	intermediary metabolism and respiration	COG4122R
Rv1223	htrA	54157.6	Unknown	serine protease HtrA	Probable serine protease HtrA (DEGP protein)	intermediary metabolism and respiration	COG0265O
Rv1229c	mrp	41063.3	Cytoplasmic	Mrp-like protein	Probable Mrp-related protein Mrp	intermediary metabolism and respiration	COG0489D
Rv1232c	-	47374.4	Cytoplasmic	hypothetical protein Rv1232c	Mg/Co/Ni transporter MgtE (contains CBS domain)	cell wall and cell processes	COG2239P
Rv1234	-	18604.5	Cytoplasmic Membrane	transmembrane protein	Probable transmembrane protein	cell wall and cell processes	-
Rv1235	lpqY	49761.4	Unknown	trehalose ABC transporter substrate-binding lipoprotein LpqY	Probable sugar-binding lipoprotein LpqY	cell wall and cell processes	COG1653G
Rv1238	sugC	42915.2	Cytoplasmic Membrane	sugar ABC transporter ATP-binding protein SugC	Probable sugar-transport ATP-binding protein ABC transporter SugC	cell wall and cell processes	COG3839G
Rv1239c	corA	41468.8	Cytoplasmic Membrane	magnesium and cobalt transport transmembrane protein CorA	Possible magnesium and cobalt transport transmembrane protein CorA	cell wall and cell processes	COG0598P
Rv1240	mdh	34289.7	Unknown	malate dehydrogenase	Probable malate dehydrogenase Mdh	intermediary metabolism and respiration	COG0039C
Rv1244	lpqZ	29569.9	Unknown	lipoprotein LpqZ	Probable lipoprotein LpqZ	cell wall and cell processes	COG1732M
Rv1245c	-	29215.3	Cytoplasmic	short-chain type dehydrogenase/reductase	Probable short-chain type dehydrogenase/reductase	intermediary metabolism and respiration	COG4221R
Rv1248c	-	135870	Cytoplasmic	multifunctional 2-oxoglutarate dehydrogenase E1 component /2-oxoglutarate dehydrogenase dihydrolipoyllysine-residue succinyltransferase	Multifunctional alpha-ketoglutarate metabolic enzyme	intermediary metabolism and respiration	COG0567C
Rv1251c	-	123117	Cytoplasmic	hypothetical protein Rv1251c	Predicted nuclease (RecB family)	Information pathways	COG2251R
Rv1253	deaD	61452.6	Cytoplasmic	ATP-dependent RNA helicase DeaD	Probable cold-shock DeaD-box protein A homolog DeaD (ATP-dependent RNA helicase dead homolog)	information pathways	COG0513LKJ
Rv1257c	-	47237.9	Cytoplasmic	oxidoreductase	Probable oxidoreductase	intermediary metabolism and respiration	COG0277C
Rv1260	-	41297.2	Cytoplasmic	oxidoreductase	Probable oxidoreductase	intermediary metabolism and	COG0654HC

						respiration	
Rv1261c	-	16756	Cytoplasmic Membrane	hypothetical protein Rv1261c	Conserved protein	conserved hypotheticals	-
Rv1263	amiB2	49048.9	Cytoplasmic	amidase AmiB2	Probable amidase AmiB2 (aminohydrolase)	intermediary metabolism and respiration	COG0154J
Rv1264	-	42200	Cytoplasmic	adenyl cyclase	Adenyl cyclase (ATP pyrophosphate-lyase) (adenylate cyclase)	intermediary metabolism and respiration	COG2114T
Rv1265	-	25229.7	Cytoplasmic	hypothetical protein Rv1265	Unknown protein	conserved hypotheticals	-
Rv1266c	pknH	66753.9	Cytoplasmic Membrane	serine/threonine-protein kinase PknH	Probable transmembrane serine/threonine-protein kinase H PknH (protein kinase H) (STPK H)	Regulatory proteins	COG0515RTKL
Rv1267c	embR	41933.5	Cytoplasmic	transcriptional regulator EmbR	Probable transcriptional regulatory protein EmbR	Regulatory proteins	COG3629T
Rv1270c	lprA	24874	Cytoplasmic Membrane	lipoprotein LprA	Possible lipoprotein LprA	cell wall and cell processes	-
Rv1272c	-	68272.6	Cytoplasmic Membrane	drug ABC transporter ATP-binding protein	Probable drugs-transport transmembrane ATP-binding protein ABC transporter	cell wall and cell processes	COG1132V
Rv1275	lprC	18870.7	Unknown	lipoprotein LprC	Possible lipoprotein LprC	cell wall and cell processes	-
Rv1278	-	93318.8	Cytoplasmic	hypothetical protein Rv1278	DNA-repair	Information pathways	COG4717S
Rv1279	-	57331.2	Cytoplasmic Membrane	GMC-type oxidoreductase	Probable dehydrogenase FAD flavoprotein GMC oxidoreductase	intermediary metabolism and respiration	COG2303E
Rv1280c	oppA	63444.5	Unknown	oligopeptide ABC transporter substrate-binding lipoprotein OppA	Probable periplasmic oligopeptide-binding lipoprotein OppA	cell wall and cell processes	COG0747E
Rv1281c	oppD	65293	Cytoplasmic Membrane	oligopeptide ABC transporter ATP-binding protein OppD	Probable oligopeptide-transport ATP-binding protein ABC transporter OppD	cell wall and cell processes	COG0444EP
Rv1284	canA	18156.7	Cytoplasmic	beta-carbonic anhydrase	Beta-carbonic anhydrase	intermediary metabolism and respiration	COG0288P
Rv1285	cysD	37512.5	Cytoplasmic	sulfate adenylyltransferase subunit 2	Probable sulfate adenylyltransferase subunit 2 CysD	intermediary metabolism and respiration	COG0175EH
Rv1286	-	67838.9	Cytoplasmic	bifunctional sulfate adenylyltransferase subunit 1 CysN/adenylylsulfate kinase CysC	Probable bifunctional enzyme CysN/CysC: sulfate adenylyltransferase (subunit 1) + adenylylsulfate kinase	intermediary metabolism and respiration	COG2895P
Rv1287	-	17209.4	Cytoplasmic	HTH-type transcriptional regulator	Predicted transcriptional regulator	Regulatory proteins	COG1959K
Rv1288	-	49618.9	Cytoplasmic Membrane	hypothetical protein Rv1288	FOG- LysM repeat	Regulatory proteins	COG0627R
Rv1289	-	23549.8	Cytoplasmic	hypothetical protein Rv1289	Unknown protein	conserved hypotheticals	-

Rv1290c	-	55995.7	Cytoplasmic Membrane	hypothetical protein Rv1290c	Predicted membrane protein	cell wall and cell processes	COG4325S
Rv1292	argS	59677.4	Cytoplasmic	arginine--tRNA ligase	Probable arginyl-tRNA synthetase ArgS (ARGRS) (arginine--tRNA ligase)	information pathways	COG0018J
Rv1293	lysA	47425.9	Cytoplasmic	diaminopimelate decarboxylase	Diaminopimelate decarboxylase LysA (DAP decarboxylase)	intermediary metabolism and respiration	COG0019E
Rv1294	thrA	45520.7	Cytoplasmic	homoserine dehydrogenase	Probable homoserine dehydrogenase ThrA	intermediary metabolism and respiration	COG0460E
Rv1295	thrC	37321.9	Cytoplasmic	threonine synthase	Threonine synthase ThrC (ts)	intermediary metabolism and respiration	COG0498E
Rv1296	thrB	32346.4	Cytoplasmic Membrane	homoserine kinase	Probable homoserine kinase ThrB	intermediary metabolism and respiration	COG0083E
Rv1297	rho	65101.4	Cytoplasmic	transcription termination factor Rho	Probable transcription termination factor Rho homolog	information pathways	COG1158K
Rv1299	prfA	39036.5	Cytoplasmic	peptide chain release factor PrfA	Probable peptide chain release factor 1 PrfA (RF-1)	information pathways	COG0216J
Rv1301	-	22552.7	Unknown	threonylcarbamoyl-AMP synthase	Putative translation factor (SUA5)	Information pathways	COG0009J
Rv1304	atpB	27466.6	Cytoplasmic Membrane	ATP synthase subunit A	Probable ATP synthase a chain AtpB (protein 6)	intermediary metabolism and respiration	COG0356C
Rv1306	atpF	18324.8	Unknown	ATP synthase subunit B	Probable ATP synthase B chain AtpF	intermediary metabolism and respiration	COG0711C
Rv1307	atpH	48805.8	Cytoplasmic	ATP synthase subunit b/delta	Probable ATP synthase delta chain AtpH	intermediary metabolism and respiration	COG0712C
Rv1308	atpA	59288.3	Cytoplasmic	ATP synthase subunit alpha	Probable ATP synthase alpha chain AtpA	intermediary metabolism and respiration	COG0056C
Rv1309	atpG	33890.4	Cytoplasmic	ATP synthase subunit gamma	Probable ATP synthase gamma chain AtpG	intermediary metabolism and respiration	COG0224C
Rv1310	atpD	53094.2	Cytoplasmic Membrane	ATP synthase subunit beta	Probable ATP synthase beta chain AtpD	intermediary metabolism and respiration	COG0055C
Rv1311	atpC	13134.8	Cytoplasmic	ATP synthase subunit epsilon	Probable ATP synthase epsilon chain AtpC	intermediary metabolism and respiration	COG0355C
Rv1312	-	16594.4	Cytoplasmic Membrane	hypothetical protein Rv1312	Conserved hypothetical secreted protein	cell wall and cell processes	-
Rv1314c	-	20694.4	Cytoplasmic	cob(I)yrinic acid a,c-diamide adenosyltransferase	Conserved protein	conserved hypotheticals	COG2096S

Rv1315	murA	44031.4	Cytoplasmic	UDP-N-acetylglucosamine 1-carboxyvinyltransferase	Probable UDP-N-acetylglucosamine 1-carboxyvinyltransferase MurA	cell wall and cell processes	COG0766M
Rv1319c	-	58097.6	Cytoplasmic Membrane	adenylate cyclase	Possible adenylate cyclase (ATP pyrophosphate-lyase) (adenylyl cyclase)	intermediary metabolism and respiration	COG2114T
Rv1321	-	25061.7	Cytoplasmic	endonuclease NucS	Predicted nuclease of the RecB family	Information pathways	COG1637L
Rv1322A	-	16594.1	Cytoplasmic	hypothetical protein Rv1322A	Lactoylglutathione lyase and related lyases	Intermediary metabolism and respiration	COG0346E
Rv1323	fadA4	40048.5	Cytoplasmic	acetyl-CoA acetyltransferase	Probable acetyl-CoA acetyltransferase FadA4 (acetoacetyl-CoA thiolase)	lipid metabolism	COG0183I
Rv1324	-	32138.2	Cytoplasmic	thioredoxin	Possible thioredoxin	intermediary metabolism and respiration	COG3118O
Rv1326c	glgB	81729.5	Cytoplasmic	1,4-alpha-glucan branching protein	1,4-alpha-glucan branching enzyme GlgB (glycogen branching enzyme)	intermediary metabolism and respiration	COG0296G
Rv1327c	glgE	78607.9	Cytoplasmic	alpha-1,4-glucan:maltose-1-phosphate maltosyltransferase	Probable glucanase GlgE	intermediary metabolism and respiration	COG0366G
Rv1328	glgP	95483.3	Cytoplasmic	glycogen phosphorylase	Probable glycogen phosphorylase GlgP	intermediary metabolism and respiration	COG0058G
Rv1330c	pncB1	47194.6	Cytoplasmic	nicotinic acid phosphoribosyltransferase PncB1	Nicotinic acid phosphoribosyltransferase PncB1	intermediary metabolism and respiration	COG1488H
Rv1333	-	33952.8	Cytoplasmic Membrane	hydrolase	Probable hydrolase	intermediary metabolism and respiration	COG3191EQ
Rv1335	cysO	9556.7	Unknown	sulfur carrier protein CysO	Sulfur carrier protein CysO	intermediary metabolism and respiration	COG1977H
Rv1338	murI	28642.8	Cytoplasmic	glutamate racemase	Probable glutamate racemase MurI	cell wall and cell processes	COG0796M
Rv1339	-	29154	Unknown	hypothetical protein Rv1339	Metal-dependent hydrolases of the beta-lactamase superfamily III	Intermediary metabolism and respiration	COG1234R
Rv1340	rphA	27319	Cytoplasmic	ribonuclease PH	Probable ribonuclease RphA (RNase PH) (tRNA nucleotidyltransferase)	information pathways	COG0689J
Rv1341	-	20983.8	Unknown	non-canonical purine NTP pyrophosphatase	Xanthosine triphosphate pyrophosphatase	Intermediary metabolism and respiration	COG0127F
Rv1350	fabG2	25870.9	Cytoplasmic	3-oxoacyl-ACP reductase FabG	Probable 3-oxoacyl-[acyl-carrier protein] reductase FabG2 (3-ketoacyl-acyl carrier protein reductase)	lipid metabolism	COG1028IQR
Rv1360	-	37251.9	Cytoplasmic	oxidoreductase	Probable oxidoreductase	intermediary	COG2141C

						metabolism and respiration	
Rv1361c	PPE19	39983.5	Cytoplasmic Membrane	PPE family protein PPE19	PPE family protein PPE19	PE/PPE	COG5651N
Rv1362c	-	23471.2	Unknown	membrane protein	Possible membrane protein	cell wall and cell processes	-
Rv1363c	-	28308.3	Cytoplasmic Membrane	membrane protein	Possible membrane protein	cell wall and cell processes	-
Rv1364c	-	69524.1	Cytoplasmic	sigma factor regulatory protein	Possible sigma factor regulatory protein	information pathways	COG2208TK
Rv1367c	-	41325.4	Cytoplasmic Membrane	hypothetical protein Rv1367c	Beta-lactamase class C and other penicillin binding proteins	cell wall and cell processes	COG1680V
Rv1368	lprF	26851.6	Cytoplasmic Membrane	lipoprotein LprF	Probable conserved lipoprotein LprF	cell wall and cell processes	-
Rv1373	-	37005.8	Cytoplasmic Membrane	glycolipid sulfotransferase	Glycolipid sulfotransferase	intermediary metabolism and respiration	-
Rv1377c	-	22815.5	Cytoplasmic	transferase	Putative transferase	intermediary metabolism and respiration	COG2227H
Rv1379	pyrR	20626.5	Cytoplasmic	bifunctional pyrimidine operon regulatory protein/uracil phosphoribosyltransferase	Probable pyrimidine operon regulatory protein PyrR	Regulatory proteins	COG2065F
Rv1380	pyrB	33818.6	Cytoplasmic	aspartate carbamoyltransferase	Probable aspartate carbamoyltransferase PyrB (ATCase) (aspartate transcarbamylase)	intermediary metabolism and respiration	COG0540F
Rv1381	pyrC	45160.3	Cytoplasmic	dihydroorotase	Probable dihydroorotase PyrC (DHOase)	intermediary metabolism and respiration	COG0044F
Rv1382	-	18189.1	Cytoplasmic Membrane	hypothetical protein Rv1382	Probable export or membrane protein	cell wall and cell processes	-
Rv1383	carA	39748.7	Unknown	carbamoyl-phosphate synthase small subunit	Probable carbamoyl-phosphate synthase small chain CarA (carbamoyl-phosphate synthetase glutamine chain)	intermediary metabolism and respiration	COG0505EF
Rv1384	carB	118929	Unknown	carbamoyl-phosphate synthase large subunit	Probable carbamoyl-phosphate synthase large chain CarB (carbamoyl-phosphate synthetase ammonia chain)	intermediary metabolism and respiration	COG0458EF
Rv1387	PPE20	55403.6	Cytoplasmic	PPE family protein PPE20	PPE family protein PPE20	PE/PPE	COG5651N
Rv1388	mihF	20817.1	Cytoplasmic	integration host factor MihF	Putative integration host factor MihF	information pathways	-
Rv1389	gmk	22063.2	Cytoplasmic	guanylate kinase	Probable guanylate kinase Gmk	intermediary metabolism and respiration	COG0194F
Rv1390	rpoZ	11810.2	Cytoplasmic	DNA-directed RNA polymerase subunit omega	Probable DNA-directed RNA polymerase (omega chain) RpoZ (transcriptase omega chain) (RNA polymerase omega subunit)	information pathways	COG1758K
Rv1391	dfp	43544.9	Cytoplasmic	bifunctional phosphopantothenoyleysteine	Probable DNA/pantothenate metabolism flavoprotein homolog Dfp	intermediary metabolism and	COG0452H

				decarboxylase/phosphopantothenate--cysteine ligase		respiration	
Rv1392	metK	43014.6	Cytoplasmic	S-adenosylmethionine synthetase	Probable S-adenosylmethionine synthetase MetK (mat) (AdoMet synthetase) (methionine adenosyltransferase)	intermediary metabolism and respiration	COG0192H
Rv1393c	-	55251.7	Cytoplasmic	monoxygenase	Probable monoxygenase	intermediary metabolism and respiration	COG2072P
Rv1397c	vapC10	14952.4	Unknown	ribonuclease VapC10	Possible toxin VapC10	virulence, detoxification, adaptation	COG1487R
Rv1399c	nlhH	33906.5	Cytoplasmic	carboxylesterase NlhH	Probable non lipolytic carboxylesterase NlhH	intermediary metabolism and respiration	COG0657I
Rv1402	priA	69839.1	Cytoplasmic	primosomal protein N'	Putative primosomal protein N' PriA (replication factor Y)	information pathways	COG1198L
Rv1404	-	17521.1	Cytoplasmic	transcriptional regulator	Probable transcriptional regulatory protein	Regulatory proteins	COG1846K
Rv1406	fmt	32659.6	Cytoplasmic	methionyl-tRNA formyltransferase	Probable methionyl-tRNA formyltransferase Fmt	information pathways	COG0223J
Rv1407	fmu	48493.4	Cytoplasmic	methyltransferase	Probable Fmu protein (sun protein)	information pathways	COG0144J
Rv1408	rpe	24073.4	Unknown	ribulose-phosphate 3-epimerase	Probable ribulose-phosphate 3-epimerase Rpe (PPE) (R5P3E) (pentose-5-phosphate 3-epimerase)	intermediary metabolism and respiration	COG0036G
Rv1409	ribG	35366.5	Cytoplasmic	bifunctional riboflavin biosynthesis diaminohydroxyphosphoribosylaminopyrimidine deaminase/5-amino-6-(5-phosphoribosylamino) uracil reductase	Probable bifunctional riboflavin biosynthesis protein RibG : diaminohydroxyphosphoribosylaminopyrimidine deaminase (riboflavin-specific deaminase) + 5-amino-6-(5-phosphoribosylamino) uracil reductase (HTP reductase)	intermediary metabolism and respiration	COG0117H
Rv1410c	-	54688.5	Cytoplasmic Membrane	aminoglycosides/tetracycline-transport integral membrane protein	Aminoglycosides/tetracycline-transport integral membrane protein	cell wall and cell processes	COG2814G
Rv1411c	lprG	24547.8	Cytoplasmic Membrane	lipoprotein LprG	Conserved lipoprotein LprG	cell wall and cell processes	-
Rv1412	ribC	21320.2	Cytoplasmic	riboflavin synthase	Probable riboflavin synthase alpha chain RibC (RibE)	intermediary metabolism and respiration	COG0307H
Rv1415	ribA2	46017.3	Cytoplasmic	bifunctional riboflavin biosynthesis GTP cyclohydrolase II/3,4-dihydroxy-2-butanone 4-phosphate synthase	Probable riboflavin biosynthesis protein RibA2 : GTP cyclohydrolase II + 3,4-dihydroxy-2-butanone 4-phosphate synthase (DHBP synthase)	intermediary metabolism and respiration	COG0108H
Rv1416	ribH	16338.5	Unknown	6,7-dimethyl-8-ribityllumazine synthase	Probable riboflavin synthase beta chain RibH (6,7-dimethyl-8-ribityllumazine synthase) (DMRL synthase) (lumazine synthase)	intermediary metabolism and respiration	-
Rv1420	uvrC	71582.1	Cytoplasmic	excinuclease ABC subunit UvrC	Probable excinuclease ABC (subunit C - nuclease) UvrC	information pathways	COG0322L

Rv1421	-	32912.5	Cytoplasmic	hypothetical protein Rv1421	Predicted P-loop-containing kinase	Intermediary metabolism and respiration	COG1660R
Rv1422	-	35823.2	Cytoplasmic	hypothetical protein Rv1422	coenzyme biosynthetic process	Intermediary metabolism and respiration	COG0391S
Rv1423	whiA	34933	Cytoplasmic	transcriptional regulator WhiA	Probable transcriptional regulatory protein WhiA	Regulatory proteins	COG1481S
Rv1425	-	50062.7	Cytoplasmic	diacylglycerol O-acyltransferase	Possible triacylglycerol synthase (diacylglycerol acyltransferase)	lipid metabolism	COG4908R
Rv1426c	lipO	46096.9	Cytoplasmic	esterase LipO	Probable esterase LipO	intermediary metabolism and respiration	COG0657I
Rv1427c	fadD12	58064.1	Cytoplasmic	long-chain-fatty-acid--CoA ligase FadD12	Possible long-chain-fatty-acid--CoA ligase FadD12 (fatty-acid-CoA synthetase) (fatty-acid-CoA synthase)	lipid metabolism	COG0318IQ
Rv1428c	-	29894.2	Cytoplasmic	hypothetical protein Rv1428c	1-acyl-sn-glycerol-3-phosphate acyltransferase	Intermediary metabolism and respiration	COG0204I
Rv1429	-	46866.7	Cytoplasmic	hypothetical protein Rv1429	Regulator of polyketide synthase expression	Lipid metabolism	COG2508TQ
Rv1431	-	65322.2	Cytoplasmic Membrane	hypothetical protein Rv1431	Conserved membrane protein	cell wall and cell processes	-
Rv1432	-	50476.4	Cytoplasmic	dehydrogenase	Probable dehydrogenase	intermediary metabolism and respiration	COG1233Q
Rv1436	gap	35923.9	Cytoplasmic	glyceraldehyde 3-phosphate dehydrogenase	Probable glyceraldehyde 3-phosphate dehydrogenase Gap (GAPDH)	intermediary metabolism and respiration	COG0057G
Rv1437	pgk	42512.5	Cytoplasmic	phosphoglycerate kinase	Probable phosphoglycerate kinase Pkg	intermediary metabolism and respiration	COG0126G
Rv1438	tpi	27371	Cytoplasmic	triosephosphate isomerase	Probable triosephosphate isomerase Tpi (TIM)	intermediary metabolism and respiration	COG0149G
Rv1442	bisC	83390.1	Cytoplasmic	biotin sulfoxide reductase BisC	Probable biotin sulfoxide reductase BisC (BDS reductase) (BSO reductase)	intermediary metabolism and respiration	COG0243C
Rv1443c	-	18284.1	Cytoplasmic Membrane	hypothetical protein Rv1443c	Unknown protein	conserved hypotheticals	COG4276S
Rv1444c	-	15023.1	Cytoplasmic	hypothetical protein Rv1444c	Unknown protein	conserved hypotheticals	-
Rv1445c	devB	25772.2	Unknown	6-phosphogluconolactonase	Probable 6-phosphogluconolactonase DevB (6PGL)	intermediary metabolism and respiration	COG0363G
Rv1446c	opcA	32717.6	Cytoplasmic	OXPP cycle protein OpcA	Putative OXPP cycle protein OpcA	intermediary metabolism and respiration	COG3429G
Rv1447c	zwf2	57343	Cytoplasmic	glucose-6-phosphate 1-	Probable glucose-6-phosphate 1-dehydrogenase Zwf2 (G6PD)	intermediary	COG0364G

				dehydrogenase		metabolism and respiration	
Rv1448c	tal	40720.7	Cytoplasmic	transaldolase	Probable transaldolase Tal	intermediary metabolism and respiration	COG0176G
Rv1449c	tkt	75556.7	Cytoplasmic	transketolase	Transketolase Tkt (TK)	intermediary metabolism and respiration	COG0021G
Rv1454c	qor	34046.5	Cytoplasmic	quinone reductase	Probable quinone reductase Qor (NADPH:quinone reductase) (zeta-crystallin homolog protein)	intermediary metabolism and respiration	COG0604CR
Rv1455	-	31426.7	Cytoplasmic Membrane	hypothetical protein Rv1455	Conserved protein	conserved hypotheticals	-
Rv1458c	-	33726.6	Cytoplasmic Membrane	antibiotic ABC transporter ATP-binding protein	Probable unidentified antibiotic-transport ATP-binding protein ABC transporter	cell wall and cell processes	COG1131V
Rv1461	-	94171.4	Cytoplasmic	hypothetical protein Rv1461	ABC-type transport system involved in Fe-S cluster assembly, permease component	cell wall and cell processes	COG0719O
Rv1462	-	42522.6	Cytoplasmic	hypothetical protein Rv1462	ABC-type transport system involved in Fe-S cluster assembly, permease component	cell wall and cell processes	COG0719O
Rv1463	-	28822.7	Cytoplasmic	ABC transporter ATP-binding protein	Probable conserved ATP-binding protein ABC transporter	cell wall and cell processes	COG0396O
Rv1464	csd	44596.5	Cytoplasmic	cysteine desulfurase	Probable cysteine desulfurase Csd	intermediary metabolism and respiration	COG0520E
Rv1466	-	12383.8	Cytoplasmic	hypothetical protein Rv1466	Predicted metal-sulfur cluster biosynthetic enzyme	Intermediary metabolism and respiration	COG2151R
Rv1467c	fadE15	65905.9	Cytoplasmic	acyl-CoA dehydrogenase FadE15	Probable acyl-CoA dehydrogenase FadE15	lipid metabolism	COG1960I
Rv1471	trxB1	13308.9	Cytoplasmic	thioredoxin	Probable thioredoxin TrxB1	intermediary metabolism and respiration	COG3118O
Rv1473	-	58392.2	Cytoplasmic	macrolide ABC transporter ATP-binding protein	Probable macrolide-transport ATP-binding protein ABC transporter	cell wall and cell processes	COG0488R
Rv1475c	acn	102417	Cytoplasmic	iron-regulated aconitate hydratase	Probable iron-regulated aconitate hydratase Acn (citrate hydro-lyase) (aconitase)	intermediary metabolism and respiration	COG1048C
Rv1477	ripA	49808.1	Extracellular	peptidoglycan endopeptidase RipA	Peptidoglycan hydrolase	virulence, detoxification, adaptation	COG0791M
Rv1478	-	24624	Extracellular	peptidoglycan endopeptidase RipB	Possible invasion protein	virulence, detoxification, adaptation	COG0791M
Rv1479	moxR1	40762.9	Cytoplasmic	transcriptional regulator MoxR1	Probable transcriptional regulatory protein MoxR1	Regulatory proteins	COG0714R

Rv1480	-	34301.3	Cytoplasmic	hypothetical protein Rv1480	Uncharacterized conserved protein (some members contain a von Willebrand factor type A (vWA) domain)	Regulatory proteins	COG1721R
Rv1481	-	36029.9	Unknown	membrane protein	Probable membrane protein	cell wall and cell processes	COG1240H
Rv1483	fabG1	25665.2	Cytoplasmic	3-oxoacyl-ACP reductase FabG1	3-oxoacyl-[acyl-carrier protein] reductase FabG1 (3-ketoacyl-acyl carrier protein reductase) (mycolic acid biosynthesis a protein)	lipid metabolism	COG1028IQR
Rv1484	inhA	28527.8	Cytoplasmic Membrane	NADH-dependent enoyl-[ACP] reductase	NADH-dependent enoyl-[acyl-carrier-protein] reductase InhA (NADH-dependent enoyl-ACP reductase)	lipid metabolism	COG0623I
Rv1488	-	41250.2	Cytoplasmic Membrane	hypothetical protein Rv1488	Possible exported conserved protein	cell wall and cell processes	COG0330O
Rv1489	-	12305.7	Cytoplasmic Membrane	hypothetical protein Rv1489	Conserved protein	conserved hypotheticals	-
Rv1492	mutA	64712.3	Cytoplasmic	methylmalonyl-CoA mutase small subunit	Probable methylmalonyl-CoA mutase small subunit MutA (MCM)	lipid metabolism	COG1884I
Rv1493	mutB	80604.2	Cytoplasmic	methylmalonyl-CoA mutase large subunit	Probable methylmalonyl-CoA mutase large subunit MutB (MCM)	lipid metabolism	COG1884I
Rv1496	-	36256.4	Cytoplasmic Membrane	transport system kinase	Possible transport system kinase	cell wall and cell processes	COG1703E
Rv1498A	-	7628.64	Cytoplasmic	hypothetical protein Rv1498A	Conserved protein	conserved hypotheticals	COG3360S
Rv1498c	-	23110	Cytoplasmic	methyltransferase	Probable methyltransferase	intermediary metabolism and respiration	COG2226H
Rv1502	-	33472.6	Cytoplasmic	hypothetical protein Rv1502	Hypothetical protein	unknown	-
Rv1507c	-	27072.6	Unknown	hypothetical protein Rv1507c	cell wall biosynthesis	cell wall and cell processes	-
Rv1508c	-	63856.9	Cytoplasmic Membrane	membrane protein	Probable membrane protein	cell wall and cell processes	-
Rv1509	-	33023.4	Cytoplasmic	hypothetical protein Rv1509	Hypothetical protein	unknown	COG2227H
Rv1511	gmdA	38310.4	Cytoplasmic	GDP-D-mannose dehydratase GmdA	GDP-D-mannose dehydratase GmdA (GDP-mannose 4,6 dehydratase) (GMD)	intermediary metabolism and respiration	COG1089M
Rv1512	epiA	34537.2	Cytoplasmic	nucleotide-sugar epimerase EpiA	Probable nucleotide-sugar epimerase EpiA	intermediary metabolism and respiration	COG0451MG
Rv1513	-	26684.3	Cytoplasmic	hypothetical protein Rv1513	SAM-dependent methyltransferases	Intermediary metabolism and respiration	-
Rv1514c	-	28932.8	Cytoplasmic	glycosyltransferase	Glycosyltransferases involved in cell wall biogenesis	cell wall and cell processes	COG0463M
Rv1516c	-	37932.4	Cytoplasmic	sugar transferase	Probable sugar transferase	intermediary metabolism and respiration	COG0463M

Rv1521	fadD25	63141.6	Cytoplasmic	fatty-acid--CoA ligase FadD25	Probable fatty-acid-AMP ligase FadD25 (fatty-acid-AMP synthetase) (fatty-acid-AMP synthase)	lipid metabolism	COG0318IQ
Rv1523	-	37915	Unknown	methyltransferase	Probable methyltransferase	intermediary metabolism and respiration	COG2226H
Rv1527c	pks5	223857	Unknown	polyketide synthase	Probable polyketide synthase Pks5	lipid metabolism	COG3321Q
Rv1529	fadD24	62797.4	Cytoplasmic	fatty-acid--CoA ligase FadD24	Probable fatty-acid-AMP ligase FadD24 (fatty-acid-AMP synthetase) (fatty-acid-AMP synthase)	lipid metabolism	COG0318IQ
Rv1530	adh	37996.5	Cytoplasmic	alcohol dehydrogenase	Probable alcohol dehydrogenase Adh	intermediary metabolism and respiration	COG1063ER
Rv1531	-	20795.7	Cytoplasmic Membrane	hypothetical protein Rv1531	Conserved protein	conserved hypotheticals	COG2128S
Rv1533	-	39620	Cytoplasmic	monooxygenase	Dioxygenases related to 2-nitropropane dioxygenase	Intermediary metabolism and respiration	COG2070R
Rv1534	-	24468.8	Cytoplasmic	transcriptional regulator	Probable transcriptional regulator	Regulatory proteins	COG1309K
Rv1536	ileS	117308	Cytoplasmic	isoleucine--tRNA ligase	Isoleucyl-tRNA synthetase IleS	information pathways	COG0060J
Rv1537	dinX	49075.6	Cytoplasmic	DNA polymerase IV	Probable DNA polymerase IV DinX (pol IV 1) (DNA nucleotidyltransferase (DNA-directed))	information pathways	-
Rv1538c	ansA	32946.5	Cytoplasmic	L-aporaginase	Probable L-aporaginase AnsA	intermediary metabolism and respiration	COG0252EJ
Rv1540	-	33016.7	Unknown	RNA pseudouridine synthase	Pseudouridylate synthases, 23S RNA-specific	Information pathways	COG0564J
Rv1542c	glbN	14448.6	Cytoplasmic	hemoglobin GlbN	Hemoglobin GlbN	intermediary metabolism and respiration	COG2346R
Rv1543	-	36820.9	Cytoplasmic	oxidoreductase	Possible fatty acyl-CoA reductase	lipid metabolism	COG0300R
Rv1544	-	28199.6	Cytoplasmic	ketoacyl reductase	Possible ketoacyl reductase	lipid metabolism	COG0300R
Rv1546	-	15324.5	Unknown	hypothetical protein Rv1546	Conserved protein	conserved hypotheticals	-
Rv1547	dnaE1	129323	Cytoplasmic	DNA polymerase III subunit alpha	Probable DNA polymerase III (alpha chain) DnaE1 (DNA nucleotidyltransferase)	information pathways	COG0587L
Rv1552	frdA	63763.1	Cytoplasmic Membrane	fumarate reductase flavoprotein subunit	Probable fumarate reductase [flavoprotein subunit] FrdA (fumarate dehydrogenase) (fumaric hydrogenase)	intermediary metabolism and respiration	COG1053C
Rv1556	-	22329.5	Cytoplasmic	HTH-type transcriptional regulator	Possible regulatory protein	Regulatory proteins	COG1309K
Rv1558	-	16347.5	Unknown	hypothetical protein Rv1558	Conserved protein	conserved hypotheticals	-
Rv1559	ilvA	45008.9	Cytoplasmic	threonine dehydratase IlvA	Probable threonine dehydratase IlvA	intermediary metabolism and respiration	COG1171E
Rv1562c	treZ	64076.7	Cytoplasmic	malto-oligosyltrehalose	Maltooligosyltrehalose trehalohydrolase TreZ	virulence,	COG0296G

				trehalohydrolase		detoxification, adaptation	
Rv1563c	treY	82706.8	Cytoplasmic	maltooligosyl trehalose synthase	Maltooligosyltrehalose synthase TreY	virulence, detoxification, adaptation	COG3280G
Rv1564c	treX	81081.1	Cytoplasmic	maltooligosyl trehalose synthase	Probable maltooligosyltrehalose synthase TreX	virulence, detoxification, adaptation	COG1523G
Rv1566c	-	23964.6	Cytoplasmic Membrane	hypothetical protein Rv1566c	Possible Inv protein	virulence, detoxification, adaptation	COG0791M
Rv1568	bioA	46319.4	Cytoplasmic	adenosylmethionine-8-amino-7-oxononoate aminotransferase	Adenosylmethionine-8-amino-7-oxononoate aminotransferase BioA	intermediary metabolism and respiration	COG0161H
Rv1592c	-	48043.1	Cytoplasmic	hypothetical protein Rv1592c	Secretory lipase	cell wall and cell processes	COG0657I
Rv1593c	-	25907.8	Cytoplasmic	hypothetical protein Rv1593c	ADP-ribose pyrophosphatase	Intermediary metabolism and respiration	COG1051F
Rv1594	nadA	37408.4	Cytoplasmic	quinolinate synthetase A	Probable quinolinate synthetase NadA	intermediary metabolism and respiration	COG0379H
Rv1595	nadB	53784.8	Cytoplasmic	L-aspartate oxidase	Probable L-aspartate oxidase NadB	intermediary metabolism and respiration	COG0029H
Rv1596	nadC	29951.1	Cytoplasmic	nicotinate-nucleotide pyrophosphatase	Probable nicotinate-nucleotide pyrophosphatase NadC	intermediary metabolism and respiration	COG0157H
Rv1597	-	27930.6	Cytoplasmic	hypothetical protein Rv1597	SAM-dependent methyltransferases	Intermediary metabolism and respiration	COG2226H
Rv1598c	-	14892.4	Cytoplasmic	hypothetical protein Rv1598c	Conserved protein	conserved hypotheticals	-
Rv1599	hisD	45346.2	Unknown	histidinol dehydrogenase	Probable histidinol dehydrogenase HisD (HDH)	intermediary metabolism and respiration	COG0141E
Rv1600	hisC1	40581.1	Cytoplasmic	histidinol-phosphate aminotransferase	Probable histidinol-phosphate aminotransferase HisC1	intermediary metabolism and respiration	COG0079E
Rv1602	hisH	21386.4	Cytoplasmic	imidazole glycerol phosphate synthase subunit HisH	Probable amidotransferase HisH	intermediary metabolism and respiration	COG0118E
Rv1603	hisA	25730.9	Cytoplasmic	phosphoribosyl isomerase A	Probable phosphoribosylformimino-5-aminoimidazole carboxamide ribotide isomerase HisA	intermediary metabolism and respiration	COG0106E

Rv1605	hisF	27220.9	Cytoplasmic	imidazole glycerol phosphate synthase subunit HisF	Probable cyclase HisF	intermediary metabolism and respiration	COG0107E
Rv1608c	bcpB	16894.5	Cytoplasmic	peroxidoxin	Probable peroxidoxin BcpB	virulence, detoxification, adaptation	COG1225O
Rv1609	trpE	55816.9	Cytoplasmic	anthranilate synthase component I	Anthranilate synthase component I TrpE (glutamine amidotransferase)	intermediary metabolism and respiration	COG0147EH
Rv1611	trpC	28023.1	Cytoplasmic	indole-3-glycerol phosphate synthase	Probable indole-3-glycerol phosphate synthase TrpC	intermediary metabolism and respiration	COG0134E
Rv1612	trpB	43286.7	Cytoplasmic	tryptophan synthase subunit beta	Tryptophan synthase, beta subunit TrpB	intermediary metabolism and respiration	COG0133E
Rv1613	trpA	27727.7	Unknown	tryptophan synthase subunit alpha	Probable tryptophan synthase, alpha subunit TrpA	intermediary metabolism and respiration	COG0159E
Rv1614	lgt	50392.2	Cytoplasmic Membrane	prolipoprotein diacylglyceryl transferase	Possible prolipoprotein diacylglyceryl transferases Lgt	cell wall and cell processes	COG0682M
Rv1617	pykA	50667.1	Cytoplasmic	pyruvate kinase	Probable pyruvate kinase PykA	intermediary metabolism and respiration	COG0469G
Rv1618	tesB1	33140.2	Unknown	acyl-CoA thioesterase TesB	Probable acyl-CoA thioesterase II TesB1	lipid metabolism	COG1946I
Rv1620c	cydC	59496.9	Cytoplasmic Membrane	cytochrome biosynthesis ATP-binding protein/permease CydC	Probable 'component linked with the assembly of cytochrome' transport transmembrane ATP-binding protein ABC transporter CydC	intermediary metabolism and respiration	COG4987CO
Rv1621c	cydD	54766.1	Cytoplasmic Membrane	cytochrome biosynthesis ABC transporter ATP-binding protein/permease CydD	Probable 'component linked with the assembly of cytochrome' transport transmembrane ATP-binding protein ABC transporter CydD	intermediary metabolism and respiration	COG4988CO
Rv1623c	cydA	53826.5	Cytoplasmic Membrane	cytochrome D ubiquinol oxidase subunit I CydA	Probable integral membrane cytochrome D ubiquinol oxidase (subunit I) CydA (cytochrome BD-I oxidase subunit I)	intermediary metabolism and respiration	COG1271C
Rv1626	-	22669.2	Cytoplasmic	two-component system transcriptional regulator	Probable two-component system transcriptional regulator	Regulatory proteins	COG3707T
Rv1627c	-	42386	Cytoplasmic	nonspecific lipid-transfer protein	Probable nonspecific lipid-transfer protein	lipid metabolism	COG0183I
Rv1628c	-	17721.9	Unknown	hypothetical protein Rv1628c	Predicted nucleic-acid-binding protein containing a Zn-ribbon	Regulatory proteins	COG1545R
Rv1629	polA	98440	Cytoplasmic	DNA polymerase I	Probable DNA polymerase I PolA	information pathways	COG0749L
Rv1630	rpsA	53201.8	Cytoplasmic	30S ribosomal protein S1	30S ribosomal protein S1 RpsA	information pathways	COG0539J
Rv1631	coaE	44668.6	Cytoplasmic	dephospho-CoA kinase CoaE	Probable dephospho-CoA kinase CoaE (dephosphocoenzyme a kinase)	intermediary metabolism and respiration	COG2320S
Rv1632c	-	16598.5	Cytoplasmic	hypothetical protein Rv1632c	Hypothetical protein	conserved hypotheticals	COG2306S
Rv1633	uvrB	78038.3	Cytoplasmic	excinuclease ABC subunit UvrB	Probable excinuclease ABC (subunit B - helicase) UvrB	information pathways	-

Rv1636	TB15.3	15312.4	Unknown	iron-regulated universal stress protein	Iron-regulated universal stress protein family protein TB15.3	virulence, detoxification, adaptation	COG0589T
Rv1637c	-	27936.3	Unknown	hypothetical protein Rv1637c	Zn-dependent hydrolases, including glyoxylases	Intermediary metabolism and respiration	COG0491R
Rv1638	uvrA	106099	Cytoplasmic	excinuclease ABC subunit UvrA	Probable excinuclease ABC (subunit A - DNA-binding ATPase) UvrA	information pathways	COG0178L
Rv1640c	lysX	128208	Cytoplasmic	bifunctional lysine--tRNA ligase/phosphatidylglycerol lysyltransferase	Lysyl-tRNA synthetase 2 LysX	information pathways	COG1190J
Rv1641	infC	22315.4	Cytoplasmic	initiation factor IF-3	Probable initiation factor if-3 InfC	information pathways	COG0290J
Rv1643	rplT	14526.7	Cytoplasmic	50S ribosomal protein L20	50S ribosomal protein L20 RplT	information pathways	COG0292J
Rv1647	-	33938.7	Cytoplasmic	adenylate cyclase	Adenylate cyclase (ATP pyrophosphate-lyase) (adenyl cyclase)	intermediary metabolism and respiration	COG2114T
Rv1649	pheS	37371.4	Cytoplasmic	phenylalanine--tRNA ligase subunit alpha	Probable phenylalanyl-tRNA synthetase, alpha chain PheS	information pathways	-
Rv1650	pheT	88374	Cytoplasmic	phenylalanine--tRNA ligase subunit beta	Probable phenylalanyl-tRNA synthetase, beta chain PheT	information pathways	COG0072J
Rv1652	argC	36304.3	Unknown	N-acetyl-gamma-glutamyl-phosphate reductase	Probable N-acetyl-gamma-glutamyl-phosphate reductase ArgC	intermediary metabolism and respiration	COG0002E
Rv1653	argJ	41115.3	Cytoplasmic	bifunctional glutamate N-acetyltransferase/amino-acid acetyltransferase	Probable glutamate N-acetyltransferase ArgJ	intermediary metabolism and respiration	COG1364E
Rv1654	argB	30936.8	Cytoplasmic	acetylglutamate kinase	Probable acetylglutamate kinase ArgB	intermediary metabolism and respiration	COG0548E
Rv1655	argD	40909.9	Cytoplasmic	acetylornithine aminotransferase	Probable acetylornithine aminotransferase ArgD	intermediary metabolism and respiration	COG4992E
Rv1656	argF	33025.2	Cytoplasmic	ornithine carbamoyltransferase	Probable ornithine carbamoyltransferase, anabolic ArgF	intermediary metabolism and respiration	COG0078E
Rv1658	argG	43682.4	Cytoplasmic	argininosuccinate synthase	Probable argininosuccinate synthase ArgG	intermediary metabolism and respiration	COG0137E
Rv1659	argH	49742.2	Unknown	argininosuccinate lyase	Probable argininosuccinate lyase ArgH	intermediary metabolism and respiration	COG0165E
Rv1660	pks10	37113.5	Cytoplasmic	chalcone synthase	Chalcone synthase Pks10	lipid metabolism	COG3424Q
Rv1661	pks7	221046	Cytoplasmic Membrane	polyketide synthase	Probable polyketide synthase Pks7	lipid metabolism	COG3321Q
Rv1662	pks8	167096	Cytoplasmic	polyketide synthase	Probable polyketide synthase Pks8	lipid metabolism	COG3321Q

			Membrane				
Rv1664	pks9	106371	Cytoplasmic Membrane	polyketide synthase	Probable polyketide synthase Pks9	lipid metabolism	COG3321Q
Rv1665	pks11	37639.2	Cytoplasmic	chalcone synthase	Chalcone synthase Pks11	lipid metabolism	COG3424Q
Rv1666c	cyp13 ₉	47832.7	Cytoplasmic	cytochrome P450 Cyp139	Probable cytochrome P450 139 Cyp139	intermediary metabolism and respiration	COG2124Q
Rv1676	-	25859.2	Cytoplasmic	hypothetical protein Rv1676	Antioxidant protein	virulence, detoxification, adaptation	COG0450O
Rv1679	fadE1 ₆	37876.2	Cytoplasmic	acyl-CoA dehydrogenase FadE16	Possible acyl-CoA dehydrogenase FadE16	lipid metabolism	COG1960I
Rv1680	-	30214	Cytoplasmic	hypothetical protein Rv1680	ABC-type phosphate/phosphonate transport system, periplasmic component	cell wall and cell processes	COG3221P
Rv1683	-	107362	Cytoplasmic	bifunctional long-chain acyl-CoA synthase/lipase	Possible bifunctional enzyme; long-chain acyl-CoA synthase and lipase.	lipid metabolism	COG3243I
Rv1689	tyrS	46330.4	Cytoplasmic	tyrosine--tRNA ligase	Probable tyrosyl-tRNA synthase TyrS (TYRRS)	information pathways	COG0162J
Rv1691	-	26650	Cytoplasmic	hypothetical protein Rv1691	membrane associated transport process	cell wall and cell processes	-
Rv1692	-	37010.9	Cytoplasmic	phosphatase	Probable phosphatase	intermediary metabolism and respiration	COG0647G
Rv1695	ppnK	32871.6	Cytoplasmic	inorganic polyphosphate/ATP-NAD kinase	Inorganic polyphosphate/ATP-NAD kinase PpnK (poly(P)/ATP NAD kinase)	intermediary metabolism and respiration	COG0061G
Rv1696	recN	62196.9	Cytoplasmic	DNA repair protein RecN	Probable DNA repair protein RecN (recombination protein N)	information pathways	COG0497L
Rv1697	-	42422.8	Cytoplasmic	hypothetical protein Rv1697	Predicted membrane protein	cell wall and cell processes	COG4825S
Rv1699	pyrG	63603	Cytoplasmic	CTP synthase	Probable CTP synthase PyrG	intermediary metabolism and respiration	COG0504F
Rv1701	-	33517.3	Cytoplasmic	tyrosine recombinase XerD	Probable integrase/recombinase	Insertion seqs and phages	COG4974L
Rv1703c	-	21537.5		methyltransferase	Probable catechol-O-methyltransferase	intermediary metabolism and respiration	-
Rv1707	-	51171.1	Cytoplasmic Membrane	transmembrane protein	Probable conserved transmembrane protein	cell wall and cell processes	COG0659P
Rv1708	-	34362.4	Cytoplasmic Membrane	initiation inhibitor protein	Putative initiation inhibitor protein	cell wall and cell processes	COG1192D
Rv1710	scpB	25244.6	Cytoplasmic	segregation and condensation protein ScpB	Possible segregation and condensation protein ScpB	cell wall and cell processes	COG1386K
Rv1711	-	27578.8	Cytoplasmic	RNA pseudouridine synthase	16S rRNA uridine-516 pseudouridylylate synthase and related pseudouridylylate synthases	Information pathways	COG1187J

Rv1712	cmk	24145.2	Cytoplasmic	cytidylate kinase	Cytidylate kinase Cmk (CMP kinase) (cytidine monophosphate kinase) (ck)	intermediary metabolism and respiration	COG0283F
Rv1713	engA	49925.7	Cytoplasmic Membrane	GTPase Der	Probable GTP-binding protein EngA	intermediary metabolism and respiration	COG1160R
Rv1719	-	27872.7	Cytoplasmic	transcriptional regulator	Probable transcriptional regulatory protein	Regulatory proteins	COG1414K
Rv1722	-	55428.8	Cytoplasmic	carboxylase	Possible carboxylase	lipid metabolism	COG0439I
Rv1731	gabD2	55324.5	Cytoplasmic	succinate-semialdehyde dehydrogenase	Possible succinate-semialdehyde dehydrogenase [NADP+] dependent (SSDH) GabD2	intermediary metabolism and respiration	COG1012C
Rv1732c	-	19398.9	Unknown	hypothetical protein Rv1732c	Thiol-disulfide isomerase and thioredoxins	Intermediary metabolism and respiration	COG1225O
Rv1736c	narX	72793.3	Cytoplasmic Membrane	nitrate reductase-like protein NarX	Probable nitrate reductase NarX	intermediary metabolism and respiration	COG5013C
Rv1738	-	10606.1	Cytoplasmic	hypothetical protein Rv1738	Conserved protein	conserved hypotheticals	-
Rv1742	-	27702.3	Cytoplasmic	hypothetical protein Rv1742	Unknown protein	unknown	-
Rv1743	pknE	60512.2	Cytoplasmic Membrane	serine/threonine-protein kinase PknE	Probable transmembrane serine/threonine-protein kinase E PknE (protein kinase E) (STPK E)	Regulatory proteins	COG0515RTKL
Rv1746	pknF	50668.1	Cytoplasmic Membrane	serine/threonine-protein kinase PknF	Anchored-membrane serine/threonine-protein kinase PknF (protein kinase F) (STPK F)	Regulatory proteins	COG0515RTKL
Rv1747	-	92121.2	Cytoplasmic Membrane	ABC transporter ATP-binding protein/permease	Probable conserved transmembrane ATP-binding protein ABC transporter	cell wall and cell processes	COG1131V
Rv1750c	fadD1	57413.5	Cytoplasmic Membrane	fatty-acid--CoA ligase FadD1	Possible fatty-acid-CoA ligase FadD1 (fatty-acid-CoA synthetase) (fatty-acid-CoA synthase)	lipid metabolism	COG0318IQ
Rv1751	-	50412.1	Cytoplasmic	oxidoreductase	Probable oxidoreductase	intermediary metabolism and respiration	COG0654HC
Rv1754c	-	60608.4	Cytoplasmic Membrane	hypothetical protein Rv1754c	Conserved protein	conserved hypotheticals	COG5373S
Rv1769	-	45380.9	Cytoplasmic	hypothetical protein Rv1769	Predicted amino acid aldolase or racemase	Intermediary metabolism and respiration	COG3616E
Rv1770	-	45972.6	Cytoplasmic	hypothetical protein Rv1770	Predicted aminopeptidases	Intermediary metabolism and respiration	COG2234R
Rv1771	-	48045.4	Cytoplasmic	L-gulono-1,4-lactone dehydrogenase	L-gulono-1,4-lactone dehydrogenase	intermediary metabolism and respiration	COG0277C
Rv1774	-	48340.6	Cytoplasmic	oxidoreductase	Probable oxidoreductase	intermediary metabolism and respiration	COG0277C

Rv1778c	-	15828	Unknown	hypothetical protein Rv1778c	Unknown protein	conserved hypotheticals	-
Rv1779c	-	63511.7	Cytoplasmic Membrane	integral membrane protein	Possible integral membrane protein	cell wall and cell processes	-
Rv1780	-	20456.5	Cytoplasmic	hypothetical protein Rv1780	Conserved protein	conserved hypotheticals	-
Rv1781c	malQ	79745	Cytoplasmic	4-alpha-glucanotransferase	Probable 4-alpha-glucanotransferase MalQ (amylomaltase) (disproportionating enzyme) (D-enzyme)	intermediary metabolism and respiration	COG1640G
Rv1782	eccB5	53689.2	Unknown	ESX-5 type VII secretion system protein EccB5	ESX conserved component EccB5. ESX-5 type VII secretion system protein. Probable membrane protein.	cell wall and cell processes	-
Rv1783	eccC5	152740	Cytoplasmic Membrane	ESX-5 type VII secretion system protein EccC5	ESX conserved component EccC5. ESX-5 type VII secretion system protein.	cell wall and cell processes	-
Rv1789	PPE26	38588.7	Cytoplasmic Membrane	PPE family protein PPE26	PPE family protein PPE26	PE/PPE	COG5651N
Rv1793	esxN	9941.95	Cytoplasmic	ESAT-6 like protein EsxN	Putative ESAT-6 like protein EsxN (ESAT-6 like protein 5)	cell wall and cell processes	-
Rv1794	-	32399.6	Cytoplasmic	hypothetical protein Rv1794	cell wall processes	cell wall and cell processes	-
Rv1795	eccD5	53435.2	Cytoplasmic Membrane	ESX-5 type VII secretion system protein EccD	ESX conserved component EccD5. ESX-5 type VII secretion system protein. Probable membrane protein.	cell wall and cell processes	-
Rv1796	mycP5	60027.6	Cytoplasmic Membrane	membrane-anchored mycosin MycP	Probable proline rich membrane-anchored mycosin MycP5 (serine protease) (subtilisin-like protease) (subtilase-like) (mycosin-5)	intermediary metabolism and respiration	COG1404O
Rv1797	eccE5	44178.1	Cytoplasmic Membrane	ESX-5 type VII secretion system protein EccE	ESX conserved component EccE5. ESX-5 type VII secretion system protein. Probable membrane protein.	cell wall and cell processes	-
Rv1798	eccA5	67756.1	Cytoplasmic	ESX-5 type VII secretion system protein EccA	ESX conserved component EccA5. ESX-5 type VII secretion system protein.	cell wall and cell processes	COG1223R
Rv1808	PPE32	39899.3	Cytoplasmic Membrane	PPE family protein PPE32	PPE family protein PPE32	PE/PPE	COG5651N
Rv1809	PPE33	47156.3	Cytoplasmic Membrane	PPE family protein PPE33	PPE family protein PPE33	PE/PPE	COG5651N
Rv1812c	-	41938	Cytoplasmic Membrane	dehydrogenase	Probable dehydrogenase	intermediary metabolism and respiration	COG1252C
Rv1817	-	51034.9	Cytoplasmic	flavoprotein	Possible flavoprotein	intermediary metabolism and respiration	COG1053C
Rv1819c	bacA	71299.1	Cytoplasmic Membrane	vitamin B12 transport ATP-binding protein BacA	Probable drug-transport transmembrane ATP-binding protein ABC transporter BacA	cell wall and cell processes	COG4178R
Rv1820	ilvG	57520.5	Cytoplasmic	acetolactate synthase large subunit IlvG	Probable acetolactate synthase IlvG (acetohydroxy-acid synthase)(ALS)	intermediary metabolism and respiration	COG0028EH
Rv1821	secA2	88919.7	Cytoplasmic	protein translocase subunit SecA	Possible preprotein translocase ATPase SecA2	cell wall and cell processes	COG0653U

Rv1825	-	30612.8	Unknown	hypothetical protein Rv1825	membrane associated process	cell wall and cell processes	COG3879S
Rv1827	garA	17218.8	Cytoplasmic	glycogen accumulation regulator GarA	FOG- FHA domain	Regulatory proteins	COG1716T
Rv1828	-	26411.3	Cytoplasmic	HTH-type transcriptional regulator	Predicted transcriptional regulators	Regulatory proteins	COG0789K
Rv1829	-	18114.5	Cytoplasmic	hypothetical protein Rv1829	phospholipid biosynthetic process	Lipid metabolism	COG1259S
Rv1830	-	23987.8	Cytoplasmic	HTH-type transcriptional regulator	Predicted transcriptional regulators	Regulatory proteins	COG0789K
Rv1832	gcvB	99478.8	Cytoplasmic	glycine dehydrogenase	Probable glycine dehydrogenase GcvB (glycine decarboxylase) (glycine cleavage system P-protein)	intermediary metabolism and respiration	COG1003E
Rv1833c	-	32151.1	Cytoplasmic	haloalkane dehalogenase	Possible haloalkane dehalogenase	intermediary metabolism and respiration	COG0596R
Rv1836c	-	69676.8	Cytoplasmic Membrane	hypothetical protein Rv1836c	Von Willebrand factor (vWF) type A domain-containing protein	Regulatory proteins	-
Rv1837c	glcB	80403	Cytoplasmic	malate synthase	Malate synthase G GlcB	intermediary metabolism and respiration	COG2225C
Rv1841c	-	36685	Cytoplasmic Membrane	hypothetical protein Rv1841c	Conserved hypothetical membrane protein	cell wall and cell processes	COG1253R
Rv1842c	-	48092	Cytoplasmic Membrane	hypothetical protein Rv1842c	Conserved hypothetical membrane protein	cell wall and cell processes	COG1253R
Rv1843c	guaB1	49932.8	Cytoplasmic	inosine-5'-monophosphate dehydrogenase	Probable inosine-5'-monophosphate dehydrogenase GuaB1 (imp dehydrogenase) (IMPDH) (IMPD)	intermediary metabolism and respiration	COG0516F
Rv1844c	gnd1	51547.1	Cytoplasmic	6-phosphogluconate dehydrogenase	Probable 6-phosphogluconate dehydrogenase Gnd1	intermediary metabolism and respiration	COG0362G
Rv1850	ureC	60825.1	Cytoplasmic	urease subunit alpha	Urease alpha subunit UreC (urea amidohydrolase)	intermediary metabolism and respiration	COG0804E
Rv1852	ureG	23348.4	Unknown	urease accessory protein UreG	Urease accessory protein UreG	intermediary metabolism and respiration	COG0378OK
Rv1854c	ndh	49619	Cytoplasmic Membrane	NADH dehydrogenase	Probable NADH dehydrogenase Ndh	intermediary metabolism and respiration	COG1252C
Rv1855c	-	33188.2	Cytoplasmic	oxidoreductase	Possible oxidoreductase	intermediary metabolism and respiration	COG2141C
Rv1856c	-	23228.1	Cytoplasmic	oxidoreductase	Possible oxidoreductase	intermediary metabolism and respiration	COG1028IQR

Rv1860	apa	32720.8	Extracellular	alanine/proline-rich secreted protein Apa	Alanine and proline rich secreted protein Apa (fibronectin attachment protein) (immunogenic protein MPT32) (antigen MPT-32) (45-kDa glycoprotein) (45/47 kDa antigen)	cell wall and cell processes	COG5373S
Rv1862	adhA	36631.7	Cytoplasmic	alcohol dehydrogenase A	Probable alcohol dehydrogenase AdhA	intermediary metabolism and respiration	COG1064R
Rv1864c	-	27207.8	Cytoplasmic	hypothetical protein Rv1864c	Conserved protein	conserved hypotheticals	COG2258S
Rv1865c	-	30372.9	Cytoplasmic	short-chain type dehydrogenase	Probable short-chain type dehydrogenase	intermediary metabolism and respiration	COG4221R
Rv1866	-	83182.5	Cytoplasmic	hypothetical protein Rv1866	Conserved protein	lipid metabolism	-
Rv1867	-	53560.5	Unknown	hypothetical protein Rv1867	Conserved protein	lipid metabolism	COG0183I
Rv1868	-	74365.1	Cytoplasmic	hypothetical protein Rv1868	Nucleoside-diphosphate-sugar epimerases	Intermediary metabolism and respiration	COG0451MG
Rv1869c	-	43629.3	Cytoplasmic	reductase	Probable reductase	intermediary metabolism and respiration	COG1251C
Rv1871c	-	14662.9	Cytoplasmic	hypothetical protein Rv1871c	DNA-repair	Information pathways	-
Rv1872c	lldD2	45341.6	Cytoplasmic Membrane	L-lactate dehydrogenase	Possible L-lactate dehydrogenase (cytochrome) LldD2	intermediary metabolism and respiration	COG1304C
Rv1874	-	24899.2	Cytoplasmic Membrane	hypothetical protein Rv1874	link to sulfotransferase activity	Intermediary metabolism and respiration	-
Rv1875	-	15979.1	Unknown	hypothetical protein Rv1875	Pyridoxamine protein	Intermediary metabolism and respiration	-
Rv1876	bfrA	18340.7	Cytoplasmic	bacterioferritin BfrA	Probable bacterioferritin BfrA	intermediary metabolism and respiration	COG2193P
Rv1878	glnA3	46720.1	Cytoplasmic	glutamine synthetase GlnA	Probable glutamine synthetase GlnA3 (glutamine synthase) (GS-I)	intermediary metabolism and respiration	COG0174E
Rv1880c	cyp140	48839.9	Cytoplasmic Membrane	cytochrome P450 Cyp140	Probable cytochrome P450 140 Cyp140	intermediary metabolism and respiration	-
Rv1882c	-	30010.4	Cytoplasmic	short-chain type dehydrogenase/reductase	Probable short-chain type dehydrogenase/reductase	intermediary metabolism and respiration	COG4221R
Rv1883c	-	17280.7	Unknown	hypothetical protein Rv1883c	Conserved hypothetical protein	conserved hypotheticals	-
Rv1886c	fbpB	34580.9	Extracellular	diacylglycerol acyltransferase/mycolyltransferase	Secreted antigen 85-B FbpB (85B) (antigen 85 complex B) (mycolyl transferase 85B) (fibronectin-binding protein B)	lipid metabolism	COG0627R

				e Ag85B	(extracellular alpha-antigen)		
Rv1894c	-	40185.7	Cytoplasmic	hypothetical protein Rv1894c	Dioxygenases related to 2-nitropropane dioxygenase	Intermediary metabolism and respiration	COG2070R
Rv1896c	-	33267.4	Cytoplasmic	S-adenosyl-L-methionine-dependent methyltransferase	O-Methyltransferase involved in polyketide biosynthesis	Lipid metabolism	COG3315Q
Rv1899c	lppD	35071.8	Cytoplasmic	lipoprotein LppD	Possible lipoprotein LppD	cell wall and cell processes	COG2110R
Rv1900c	lipJ	49684.4	Cytoplasmic	lignin peroxidase LipJ	Probable lignin peroxidase LipJ	intermediary metabolism and respiration	COG2114T
Rv1901	cinA	45512	Cytoplasmic	protein CinA	Probable CinA-like protein CinA	virulence, detoxification, adaptation	COG1058R
Rv1908c	katG	80572.8	Cytoplasmic	catalase-peroxidase	Catalase-peroxidase-peroxynitritase T KatG	virulence, detoxification, adaptation	COG0376P
Rv1911c	lppC	19798.3	Unknown	lipoprotein LppC	Probable lipoprotein LppC	cell wall and cell processes	COG1881R
Rv1912c	fadB5	35813.2	Cytoplasmic	oxidoreductase FadB	Possible oxidoreductase FadB5	lipid metabolism	COG0604CR
Rv1914c	-	14719.3	Cytoplasmic Membrane	hypothetical protein Rv1914c	Unknown protein	conserved hypotheticals	-
Rv1915	aceAa	40489.8	Cytoplasmic	isocitrate lyase AceAa	Probable isocitrate lyase AceAa [first part] (isocitrase) (isocitratase) (Icl)	intermediary metabolism and respiration	COG2224C
Rv1916	aceAb	44581.8	Cytoplasmic	isocitrate lyase AceAb	Probable isocitrate lyase AceAb [second part] (isocitrase) (isocitratase) (Icl)	intermediary metabolism and respiration	COG2224C
Rv1919c	-	16803.1	Unknown	hypothetical protein Rv1919c	Conserved protein	conserved hypotheticals	-
Rv1920	-	31948.8	Cytoplasmic	membrane protein	Probable membrane protein	cell wall and cell processes	COG0204I
Rv1922	-	39062	Extracellular	lipoprotein	Probable conserved lipoprotein	cell wall and cell processes	COG1680V
Rv1923	lipD	47151.8	Unknown	lipase LipD	Probable lipase LipD	intermediary metabolism and respiration	COG1680V
Rv1925	fadD31	66310.1	Cytoplasmic	fatty-acid--CoA ligase FadD31	Probable acyl-CoA ligase FadD31 (acyl-CoA synthetase) (acyl-CoA synthase)	lipid metabolism	-
Rv1926c	mpt63	16514	Extracellular	immunogenic protein Mpt63	Immunogenic protein Mpt63 (antigen Mpt63/MPB63) (16 kDa immunoprotective extracellular protein)	cell wall and cell processes	-
Rv1927	-	28554	Cytoplasmic Membrane	hypothetical protein Rv1927	Conserved hypothetical protein	conserved hypotheticals	COG3361S
Rv1929c	-	23977.7	Cytoplasmic Membrane	hypothetical protein Rv1929c	Conserved hypothetical protein	conserved hypotheticals	-

Rv1932	tpx	16896.1	Unknown	thiol peroxidase	Probable thiol peroxidase Tpx	virulence, detoxification, adaptation	COG2077O
Rv1944c	-	20795.2	Cytoplasmic	hypothetical protein Rv1944c	Conserved protein	conserved hypotheticals	COG3318R
Rv1957	-	20106.4	Cytoplasmic	SecB-like chaperone	Hypothetical protein	conserved hypotheticals	-
Rv1963c	mce3R	44341.7	Cytoplasmic	transcriptional repressor Mce3R	Probable transcriptional repressor (probably TetR-family) Mce3R	Regulatory proteins	COG1309K
Rv1977	-	38307.9	Cytoplasmic	hypothetical protein Rv1977	Zn-dependent protease with chaperone function	Information pathways	COG0501O
Rv1980c	mpt64	24823.1	Extracellular	immunogenic protein Mpt64	Immunogenic protein Mpt64 (antigen Mpt64/MPB64)	cell wall and cell processes	-
Rv1981c	nrdF1	36589.9	Unknown	ribonucleoside-diphosphate reductase subunit beta NrdF1	Ribonucleoside-diphosphate reductase (beta chain) NrdF1 (ribonucleotide reductase small subunit) (R2F protein)	information pathways	COG0208F
Rv1984c	cfp21	21782.4	Unknown	cutinase	Probable cutinase precursor CFP21	cell wall and cell processes	COG3946U
Rv1987	-	14919	Cytoplasmic Membrane	chitinase	Possible chitinase	cell wall and cell processes	COG5297G
Rv1992c	ctpG	79277.4	Cytoplasmic Membrane	cation transporter ATPase G	Probable metal cation transporter P-type ATPase G CtpG	cell wall and cell processes	COG2217P
Rv1993c	-	9333.73	Unknown	hypothetical protein Rv1993c	Conserved protein	conserved hypotheticals	-
Rv1996	-	33879.7	Cytoplasmic Membrane	universal stress protein	Universal stress protein family protein	virulence, detoxification, adaptation	COG0589T
Rv2000	-	59894.8	Cytoplasmic	hypothetical protein Rv2000	Unknown protein	conserved hypotheticals	COG1249C
Rv2001	-	28701.5	Cytoplasmic	hypothetical protein Rv2001	Acyl-ACP thioesterase	Intermediary metabolism and respiration	COG3884I
Rv2002	fabG3	27029.7	Cytoplasmic	20-beta-hydroxysteroid dehydrogenase	Possible 20-beta-hydroxysteroid dehydrogenase FabG3 (cortisone reductase) ((R)-20-hydroxysteroid dehydrogenase)	lipid metabolism	COG1028IQR
Rv2004c	-	54422.5	Cytoplasmic	hypothetical protein Rv2004c	Predicted kinase	Regulatory proteins	COG2187S
Rv2005c	-	30985.4	Cytoplasmic	universal stress protein	Universal stress protein family protein	virulence, detoxification, adaptation	COG0589T
Rv2006	otsB1	145752	Cytoplasmic	trehalose-6-phosphate phosphatase OtsB	Probable trehalose-6-phosphate phosphatase OtsB1 (trehalose-phosphatase) (TPP)	virulence, detoxification, adaptation	COG1554G
Rv2007c	fdxA	12031.7	Unknown	ferredoxin	Ferredoxin FdxA	intermediary metabolism and respiration	COG1146C
Rv2008c	-	48392.8	Cytoplasmic Membrane	hypothetical protein Rv2008c	Predicted ATPase (AAA+ superfamily)	Intermediary metabolism and respiration	COG1373R

Rv2024c	-	57273.9	Cytoplasmic	hypothetical protein Rv2024c	Predicted helicase	Information pathways	COG4889R
Rv2027c	dosT	62136.5	Cytoplasmic Membrane	two component sensor histidine kinase DosT	Two component sensor histidine kinase DosT	Regulatory proteins	COG4585T
Rv2028c	-	29441.6	Cytoplasmic Membrane	universal stress protein	Universal stress protein family protein	virulence, detoxification, adaptation	COG0589T
Rv2029c	pfkB	35401.4	Cytoplasmic Membrane	6-phosphofructokinase PfkB	6-phosphofructokinase PfkB (phosphohexokinase) (phosphofructokinase)	intermediary metabolism and respiration	COG1105G
Rv2030c	-	74898.2	Cytoplasmic	hypothetical protein Rv2030c	Erythromycin esterase homolog	Intermediary metabolism and respiration	COG2312R
Rv2031c	hspX	16227.3	Cellwall	alpha-crystallin	Heat shock protein HspX (alpha-crystallin homolog) (14 kDa antigen) (HSP16.3)	virulence, detoxification, adaptation	COG0071O
Rv2032	acg	36559.3	Cytoplasmic	NAD(P)H nitroreductase acg	nitroreductase	Lipid metabolism	-
Rv2035	-	18514.9	Cytoplasmic	hypothetical protein Rv2035	Conserved hypothetical protein	conserved hypotheticals	COG3832S
Rv2037c	-	34656.5	Cytoplasmic Membrane	transmembrane protein	Conserved transmembrane protein	cell wall and cell processes	COG1752R
Rv2038c	-	38886.7	Cytoplasmic Membrane	sugar ABC transporter ATP-binding protein	Probable sugar-transport ATP-binding protein ABC transporter	cell wall and cell processes	COG3839G
Rv2041c	-	47876.2	Unknown	sugar ABC transporter substrate-binding lipoprotein	Probable sugar-binding lipoprotein	cell wall and cell processes	COG1653G
Rv2042c	-	29091.2	Unknown	hypothetical protein Rv2042c	Nuclear transport factor 2	Regulatory proteins	-
Rv2045c	lipT	56055.8	Cytoplasmic	carboxylesterase LipT	Carboxylesterase LipT	intermediary metabolism and respiration	COG2272I
Rv2047c	-	91352.8	Cytoplasmic Membrane	hypothetical protein Rv2047c	Nucleoside-diphosphate-sugar epimerases	Intermediary metabolism and respiration	COG0451MG
Rv2048c	pks12	431607	Unknown	polyketide synthase	Polyketide synthase Pks12	lipid metabolism	-
Rv2051c	ppm1	93792.4	Cytoplasmic Membrane	polyprenol-monophosphomannose synthase	Polyprenol-monophosphomannose synthase Ppm1	cell wall and cell processes	COG0815M
Rv2052c	-	55821.8	Unknown	hypothetical protein Rv2052c	Predicted metal-dependent hydrolase with the TIM-barrel fold	Intermediary metabolism and respiration	COG1574R
Rv2054	-	25183.1	Cytoplasmic Membrane	hypothetical protein Rv2054	Dienelactone hydrolase and related enzymes	Intermediary metabolism and respiration	COG0412Q
Rv2061c	-	14782	Unknown	hypothetical protein Rv2061c	Predicted flavin-nucleotide-binding protein structurally related to pyridoxine 5-phosphate oxidase	Information pathways	COG3576R
Rv2062c	cobN	129178	Cytoplasmic	cobalamin biosynthesis protein CobN	Cobalamin biosynthesis protein CobN	intermediary metabolism and respiration	COG1429H

Rv2065	cobH	21581.7	Cytoplasmic	precorrin-8X methylmutase	Precorrin-8X methylmutase CobH (aka precorrin isomerase)	intermediary metabolism and respiration	COG2082H
Rv2066	cobIJ	53911	Unknown	bifunctional S-adenosyl-L-methionine-precorrin-2 methyl transferase/precorrin-3 methylase	Probable bifunctional protein, CobI-COBI fusion protein: S-adenosyl-L-methionine-precorrin-2 methyl transferase + precorrin-3 methylase	intermediary metabolism and respiration	COG1010H
Rv2067c	-	45898.4	Cytoplasmic	hypothetical protein Rv2067c	SAM-dependent methyltransferases	Intermediary metabolism and respiration	COG2227H
Rv2068c	blaC	32567.9	Extracellular	beta-lactamase	Class a beta-lactamase BlaC	intermediary metabolism and respiration	COG2367V
Rv2072c	cobL	41854.4	Cytoplasmic Membrane	precorrin-6Y C(5,15)-methyltransferase	Precorrin-6Y C(5,15)-methyltransferase (decarboxylating) CobL	intermediary metabolism and respiration	COG2242H
Rv2073c	-	25890	Cytoplasmic Membrane	oxidoreductase	Probable shortchain dehydrogenase	intermediary metabolism and respiration	COG0300R
Rv2074	-	14992	Cytoplasmic	pyridoxamine 5'-phosphate oxidase	Possible pyridoxamine 5'-phosphate oxidase (PNP/PMP oxidase) (pyridoxinephosphate oxidase) (PNPOX) (pyridoxine 5'-phosphate oxidase)	intermediary metabolism and respiration	-
Rv2088	pknJ	61563.3	Cytoplasmic Membrane	transmembrane serine/threonine-protein kinase PknJ	Transmembrane serine/threonine-protein kinase J PknJ (protein kinase J) (STPK J)	Regulatory proteins	COG0515RTKL
Rv2090	-	41939	Cytoplasmic	5'-3' exonuclease	Probable 5'-3' exonuclease	information pathways	COG0258L
Rv2091c	-	26019.8	Cytoplasmic Membrane	membrane protein	Probable membrane protein	cell wall and cell processes	-
Rv2092c	helY	99574	Cytoplasmic	ATP-dependent DNA helicase HelY	ATP-dependent DNA helicase HelY	information pathways	COG4581L
Rv2095c	pafC	33764.4	Cytoplasmic	proteasome accessory factor C	Proteasome accessory factor C PafC	intermediary metabolism and respiration	COG2378K
Rv2096c	pafB	35300	Cytoplasmic	proteasome accessory factor B	Proteasome accessory factor B PafB	intermediary metabolism and respiration	COG2378K
Rv2097c	pafA	51352	Cytoplasmic	proteasome accessory factor PafA	Proteasome accessory factor a PafA	intermediary metabolism and respiration	-
Rv2101	helZ	111612	Cytoplasmic	helicase HelZ	Probable helicase HelZ	information pathways	-
Rv2109c	prcA	26849	Cytoplasmic Membrane	proteasome subunit alpha	Proteasome alpha subunit PrcA; assembles with beta subunit PrcB.	intermediary metabolism and respiration	COG0638O
Rv2110c	prcB	30273	Unknown	proteasome subunit beta	Proteasome beta subunit PrcB; assembles with alpha subunit PrcA.	intermediary metabolism and respiration	COG0638O
Rv2111c	pup	6944.31	Cytoplasmic	ubiquitin-like protein Pup	Prokaryotic ubiquitin-like protein Pup	intermediary	-

						metabolism and respiration	
Rv2112c	dop	60129.2	Cytoplasmic	pup deamidase/depupylase	Deamidase of pup Dop	intermediary metabolism and respiration	-
Rv2113	-	42394.2	Cytoplasmic Membrane	integral membrane protein	Probable integral membrane protein	cell wall and cell processes	-
Rv2114	-	22774.7	Unknown	hypothetical protein Rv2114	membrane associated process	cell wall and cell processes	-
Rv2115c	mpa	67401.1	Cytoplasmic Membrane	proteasome-associated ATPase	Mycobacterial proteasome ATPase Mpa	cell wall and cell processes	COG0464O
Rv2118c	-	30090.4	Cytoplasmic	tRNA (adenine(58)-N(1))-methyltransferase	RNA methyltransferase	intermediary metabolism and respiration	COG2519J
Rv2120c	-	16371.5	Cytoplasmic Membrane	integral membrane protein	Probable conserved integral membrane protein	cell wall and cell processes	COG3918S
Rv2121c	hisG	30480.8	Cytoplasmic	ATP phosphoribosyltransferase	ATP phosphoribosyltransferase HisG	intermediary metabolism and respiration	COG0040E
Rv2124c	metH	130291	Cytoplasmic	methionine synthase	5-methyltetrahydrofolate--homocystein methyltransferase MetH (methionine synthase, vitamin-B12 dependent isozyme) (ms)	intermediary metabolism and respiration	COG1410E
Rv2129c	-	31331.3	Cytoplasmic	oxidoreductase	Probable oxidoreductase	intermediary metabolism and respiration	COG4221R
Rv2130c	mshC	45594.5	Cytoplasmic	cysteine:1D-myo-inosityl 2-amino-2-deoxy--D-glucopyranoside ligase	Cysteine:1D-myo-inosityl 2-amino-2-deoxy--D-glucopyranoside ligase MshC	intermediary metabolism and respiration	COG0215J
Rv2131c	cysQ	28415.1	Cytoplasmic	3'-phosphoadenosine 5'-phosphate phosphatase	Monophosphatase CysQ	intermediary metabolism and respiration	COG1218P
Rv2134c	-	21216.9	Cytoplasmic	hypothetical protein Rv2134c	Conserved protein	conserved hypotheticals	-
Rv2135c	-	25122.4	Cytoplasmic	hypothetical protein Rv2135c	Fructose-2,6-bisphosphatase	Intermediary metabolism and respiration	COG0406G
Rv2138	lppL	36818.6	Unknown	lipoprotein LppL	Probable conserved lipoprotein LppL	cell wall and cell processes	-
Rv2139	pyrD	37998.6	Cytoplasmic Membrane	dihydroorotate dehydrogenase	Probable dihydroorotate dehydrogenase PyrD	intermediary metabolism and respiration	COG0167F
Rv2140c	TB18.6	18634	Cytoplasmic	hypothetical protein Rv2140c	Phospholipid-binding protein	Intermediary metabolism and respiration	COG1881R
Rv2141c	-	48110.3	Unknown	hypothetical protein Rv2141c	Conserved protein	intermediary	COG0624E

						metabolism and respiration	
Rv2145c	wag31	28277.2	Cytoplasmic	cell wall synthesis protein Wag31	Diviva family protein Wag31	cell wall and cell processes	COG3599D
Rv2148c	-	27693.4	Cytoplasmic	hypothetical protein Rv2148c	Predicted enzyme with a TIM-barrel fold	Intermediary metabolism and respiration	COG0325R
Rv2149c	yfiH	25946.5	Unknown	laccase domain-containing protein	cell division	cell wall and cell processes	COG1496S
Rv2150c	ftsZ	38755.9	Cytoplasmic	cell division protein FtsZ	Cell division protein FtsZ	cell wall and cell processes	COG0206D
Rv2151c	ftsQ	33933.5	Unknown	cell division protein FtsQ	Possible cell division protein FtsQ	cell wall and cell processes	COG1589M
Rv2152c	murC	51144.5	Cytoplasmic	UDP-N-acetylmuramate--alanine ligase	Probable UDP-N-acetylmuramate-alanine ligase MurC	cell wall and cell processes	COG0773M
Rv2153c	murG	41828.1	Cytoplasmic Membrane	UDP-N-acetylglucosamine--N-acetylmuramyl-(pentapeptide) pyrophosphoryl-undecaprenol-N-acetylglucosamine transferase	Probable UDP-N-acetylglucosamine-N-acetylmuramyl-(pentapeptide) pyrophosphoryl-undecaprenol-N-acetylglucosamine transferase MurG	cell wall and cell processes	COG0707M
Rv2155c	murD	49313.4	Cytoplasmic	UDP-N-acetylmuramoylalanine--D-glutamate ligase	Probable UDP-N-acetylmuramoylalanine-D-glutamate ligase MurD	cell wall and cell processes	-
Rv2157c	murF	51632.6	Cytoplasmic	UDP-N-acetylmuramoyl-tripeptide--D-alanyl-D-alanine ligase	Probable UDP-N-acetylmuramoylalanyl-D-glutamyl-2,6-diaminopimelate-D-alanyl-D-alanyl ligase MurF	cell wall and cell processes	COG0770M
Rv2158c	murE	55308.7	Cytoplasmic	UDP-N-acetylmuramoylalanyl-D-glutamate--2,6-diaminopimelate ligase	Probable UDP-N-acetylmuramoylalanyl-D-glutamate-2,6-diaminopimelate ligase MurE	cell wall and cell processes	COG0769M
Rv2159c	-	36376.5	Cytoplasmic Membrane	hypothetical protein Rv2159c	Peroxidase	Intermediary metabolism and respiration	COG0599S
Rv2161c	-	30768.3	Cytoplasmic	hypothetical protein Rv2161c	Conserved protein	intermediary metabolism and respiration	COG2141C
Rv2165c	-	42497.2	Cytoplasmic	rRNA small subunit methyltransferase H	Predicted S-adenosylmethionine-dependent methyltransferase involved in cell envelope biogenesis	cell wall and cell processes	-
Rv2166c	-	15911.8	Unknown	protein MraZ	cell division	cell wall and cell processes	COG2001S
Rv2171	lppM	23814	Cytoplasmic Membrane	lipoprotein LppM	Probable conserved lipoprotein LppM	cell wall and cell processes	-
Rv2172c	-	33006.9	Cytoplasmic	hypothetical protein Rv2172c	membrane associated process	cell wall and cell processes	-
Rv2178c	aroG	50641.4	Cytoplasmic	phospho-2-dehydro-3-deoxyheptonate aldolase AroG	3-deoxy-D-arabino-heptulosonate 7-phosphate synthase AroG (DAHP synthetase, phenylalanine-repressible)	intermediary metabolism and respiration	COG3200E
Rv2182c	-	26955.1	Unknown	1-acylglycerol-3-phosphate O-	1-acyl-sn-glycerol-3-phosphate acyltransferase	Intermediary	COG0204I

				acyltransferase		metabolism and respiration	
Rv2184c	-	40838.9	Cytoplasmic	hypothetical protein Rv2184c	Oxyanion-translocating ATPase	Intermediary metabolism and respiration	COG0003P
Rv2185c	TB16.3	16292.7	Cytoplasmic	hypothetical protein Rv2185c	Cyclase/Dehydrase	Intermediary metabolism and respiration	COG2867I
Rv2187	fadD15	64002.9	Cytoplasmic	long-chain-fatty-acid--CoA ligase FadD15	Long-chain-fatty-acid-CoA ligase FadD15 (fatty-acid-CoA synthetase) (fatty-acid-CoA synthase)	lipid metabolism	COG1022I
Rv2192c	trpD	37707.8	Cytoplasmic Membrane	anthranilate phosphoribosyltransferase	Probable anthranilate phosphoribosyltransferase TrpD	intermediary metabolism and respiration	COG0547E
Rv2194	qcrC	29120.4	Cytoplasmic Membrane	ubiquinol-cytochrome C reductase cytochrome subunit C	Probable ubiquinol-cytochrome C reductase QcrC (cytochrome C subunit)	intermediary metabolism and respiration	COG2010C
Rv2195	qcrA	46922.6	Cytoplasmic Membrane	ubiquinol-cytochrome C reductase rieske iron-sulfur subunit	Probable rieske iron-sulfur protein QcrA	intermediary metabolism and respiration	COG0723C
Rv2196	qcrB	61014.4	Cytoplasmic Membrane	ubiquinol-cytochrome C reductase cytochrome subunit B	Probable ubiquinol-cytochrome C reductase QcrB (cytochrome B subunit)	intermediary metabolism and respiration	COG1290C
Rv2198c	mmpS3	30955	Unknown	membrane protein MmpS3	Probable conserved membrane protein MmpS3	cell wall and cell processes	COG3889R
Rv2200c	ctaC	40447.6	Cytoplasmic Membrane	cytochrome C oxidase subunit II	Probable transmembrane cytochrome C oxidase (subunit II) CtaC	intermediary metabolism and respiration	COG1622C
Rv2201	asnB	72117.5	Cytoplasmic	asparagine synthetase	Probable asparagine synthetase AsnB	intermediary metabolism and respiration	COG0367E
Rv2202c	adoK	34440	Unknown	adenosine kinase	Adenosine kinase	intermediary metabolism and respiration	COG0524G
Rv2204c	-	12543.9	Cytoplasmic	hypothetical protein Rv2204c	Conserved protein	conserved hypotheticals	COG0316S
Rv2207	cobT	36411.3	Cytoplasmic	nicotinate-nucleotide-dimethylbenzimidazol phosphoribosyltransferase	Probable nicotinate-nucleotide-dimethylbenzimidazol phosphoribosyltransferase CobT	intermediary metabolism and respiration	COG2038H
Rv2210c	ilvE	39666.1	Cytoplasmic	branched-chain amino acid aminotransferase	Branched-chain amino acid transaminase IlvE	intermediary metabolism and respiration	COG0115EH
Rv2211c	gcvT	39597	Cytoplasmic	aminomethyltransferase	Probable aminomethyltransferase GcvT (glycine cleavage system T protein)	intermediary metabolism and respiration	COG0404E
Rv2212	-	39743	Cytoplasmic	adenylyl cyclase	Adenylyl cyclase (ATP pyrophosphate-lyase) (adenylate cyclase)	intermediary	COG2114T

						metabolism and respiration	
Rv2213	pepB	53449.3	Cytoplasmic	cytosol aminopeptidase	Probable aminopeptidase PepB	intermediary metabolism and respiration	COG0260E
Rv2214c	ephD	64029.8	Cytoplasmic	oxidoreductase EphD	Possible short-chain dehydrogenase EphD	virulence, detoxification, adaptation	COG0300R
Rv2215	dlaT	57087.7	Cytoplasmic	pyruvate dehydrogenase E2 component dihydrolipoamide acyltransferase	DlaT, dihydrolipoamide acyltransferase, E2 component of pyruvate dehydrogenase	intermediary metabolism and respiration	COG0508C
Rv2216	-	31671.9	Cytoplasmic Membrane	epimerase family protein	Predicted nucleoside-diphosphate sugar epimerase	Intermediary metabolism and respiration	COG1090R
Rv2219	-	26863.6	Cytoplasmic Membrane	transmembrane protein	Probable conserved transmembrane protein	cell wall and cell processes	-
Rv2220	glnA1	53537.9	Cytoplasmic	glutamine synthetase	Glutamine synthetase GlnA1 (glutamine synthase) (GS-I)	intermediary metabolism and respiration	COG0174E
Rv2221c	glnE	109106	Cytoplasmic	glutamate-ammonia-ligase adenyltransferase	Glutamate-ammonia-ligase adenyltransferase GlnE (glutamine-synthetase adenyltransferase)	intermediary metabolism and respiration	COG1391OT
Rv2222c	glnA2	49607.8	Cytoplasmic	glutamine synthetase	Probable glutamine synthetase GlnA2 (glutamine synthase) (GS-II)	intermediary metabolism and respiration	COG0174E
Rv2223c	-	55077.8	Unknown	carboxylesterase B	Probable exported protease	cell wall and cell processes	COG0596R
Rv2224c	caeA	55892.4	Cytoplasmic	carboxylesterase A	Probable carboxylesterase CaeA	cell wall and cell processes	COG0596R
Rv2225	panB	29336.4	Unknown	3-methyl-2-oxobutanoate hydroxymethyltransferase	3-methyl-2-oxobutanoate hydroxymethyltransferase PanB	intermediary metabolism and respiration	COG0413H
Rv2226	-	56333.2	Cytoplasmic	hypothetical protein Rv2226	Protein involved in cAMP biosynthetic process	Intermediary metabolism and respiration	COG5607S
Rv2228c	-	39113.1	Cellwall	multifunctional RNASE H/alpha-ribazole phosphatase/acid phosphatase	Multifunctional protein. Has RNASE H, alpha-ribazole phosphatase, and acid phosphatase activities.	information pathways	COG0406G
Rv2229c	-	26850.9	Cytoplasmic	hypothetical protein Rv2229c	Zn-ribbon protein, possibly nucleic acid-binding	Regulatory proteins	COG1579R
Rv2230c	-	39597.8	Unknown	GTP cyclohydrolase	DNA-associated processes	Regulatory proteins	COG3323S
Rv2231c	cobC	38904.7	Cytoplasmic	aminotransferase	Possible aminotransferase CobC	intermediary metabolism and respiration	COG0079E
Rv2235	-	29762.4	Unknown	transmembrane protein	Probable conserved transmembrane protein	cell wall and cell processes	COG3346S

Rv2237	-	29067.5	Cytoplasmic Membrane	hypothetical protein Rv2237	Conserved protein	conserved hypotheticals	COG3662S
Rv2238c	ahpE	16819	Cytoplasmic	peroxiredoxin	Probable peroxiredoxin AhpE	virulence, detoxification, adaptation	COG1225O
Rv2241	aceE	100185	Cytoplasmic	pyruvate dehydrogenase E1 component	Pyruvate dehydrogenase E1 component AceE (pyruvate decarboxylase) (pyruvate dehydrogenase) (pyruvic dehydrogenase)	intermediary metabolism and respiration	-
Rv2242	-	44605.6	Cytoplasmic	hypothetical protein Rv2242	Regulator of polyketide synthase expression	Lipid metabolism	COG2508TQ
Rv2243	fabD	30756.1	Cytoplasmic	malonyl CoA-acyl carrier protein transacylase	Malonyl CoA-acyl carrier protein transacylase FabD (malonyl CoA:ACPM acyltransferase) (MCT)	lipid metabolism	COG0331I
Rv2244	acpM	12491.9	Cytoplasmic	meromycolate extension acyl carrier protein	Meromycolate extension acyl carrier protein AcpM	lipid metabolism	COG0236IQ
Rv2245	kasA	43284.1	Cytoplasmic Membrane	3-oxoacyl-ACP synthase 1	3-oxoacyl-[acyl-carrier protein] synthase 1 KasA (beta-ketoacyl-ACP synthase) (KAS I)	lipid metabolism	COG0304IQ
Rv2246	kasB	46388.7	Cytoplasmic	3-oxoacyl-ACP synthase 2	3-oxoacyl-[acyl-carrier protein] synthase 2 KasB (beta-ketoacyl-ACP synthase) (KAS I)	lipid metabolism	COG0304IQ
Rv2247	accD6	50136	Cytoplasmic	acetyl-/propionyl-CoA carboxylase subunit beta	Acetyl/propionyl-CoA carboxylase (beta subunit) AccD6	lipid metabolism	COG4799I
Rv2249c	glpD1	54143.9	Cytoplasmic	glycerol-3-phosphate dehydrogenase	Probable glycerol-3-phosphate dehydrogenase GlpD1	intermediary metabolism and respiration	COG0578C
Rv2251	-	49782.5	Cytoplasmic	flavoprotein	Possible flavoprotein	intermediary metabolism and respiration	COG0277C
Rv2252	-	32805.4	Cytoplasmic	diacylglycerol kinase	Diacylglycerol kinase	lipid metabolism	COG1597IR
Rv2257c	-	28385.1	Unknown	hypothetical protein Rv2257c	Beta-lactamase class C and other penicillin binding proteins	cell wall and cell processes	COG1680V
Rv2258c	-	37513.6	Cytoplasmic	transcriptional regulator	Possible transcriptional regulatory protein	Regulatory proteins	COG2226H
Rv2259	mScR	37888.2	Cytoplasmic	S-nitrosomycotohiol reductase MscR	S-nitrosomycotohiol reductase MscR	intermediary metabolism and respiration	COG1062C
Rv2277c	-	32163.3	Unknown	glycerolphosphodiesterase	Possible glycerolphosphodiesterase	intermediary metabolism and respiration	COG0584C
Rv2281	pitB	58789.4	Cytoplasmic Membrane	phosphate permease	Putative phosphate-transport permease PitB	cell wall and cell processes	COG0306P
Rv2282c	-	32989.1	Cytoplasmic	LysR family HTH-type transcriptional regulator	Probable transcription regulator (LysR family)	Regulatory proteins	COG0583K
Rv2284	lipM	46681.4	Cytoplasmic Membrane	esterase LipM	Probable esterase LipM	intermediary metabolism and respiration	COG0657I
Rv2285	-	47675	Cytoplasmic Membrane	diacylglycerol acyltransferase	Possible triacylglycerol synthase (diacylglycerol acyltransferase)	lipid metabolism	COG4908R

Rv2289	cdh	28576.4	Cytoplasmic Membrane	CDP-diacylglycerol pyrophosphatase	Probable CDP-diacylglycerol pyrophosphatase Cdh (CDP-diacylglycerol diphosphatase) (CDP-diacylglycerol phosphatidylhydrolase)	lipid metabolism	COG2134I
Rv2291	sseB	29367.9	Cytoplasmic	thiosulfate sulfurtransferase SseB	Probable thiosulfate sulfurtransferase SseB	intermediary metabolism and respiration	COG2897P
Rv2294	-	44251.1	Cytoplasmic	cystathionine beta-lyase	Probable aminotransferase	intermediary metabolism and respiration	COG1168E
Rv2296	-	33357.9	Cytoplasmic	haloalkane dehalogenase	Probable haloalkane dehalogenase	intermediary metabolism and respiration	COG0596R
Rv2297	-	16467.9	Unknown	hypothetical protein Rv2297	Unknown protein	conserved hypotheticals	-
Rv2298	-	34986.1	Cytoplasmic	oxidoreductase	Predicted oxidoreductases (related to aryl-alcohol dehydrogenases)	Intermediary metabolism and respiration	COG0667C
Rv2299c	htpG	72961.1	Cytoplasmic	chaperone protein HtpG	Probable chaperone protein HtpG (heat shock protein) (HSP90 family protein) (high temperature protein G)	virulence, detoxification, adaptation	COG0326O
Rv2301	cut2	23894.1	Unknown	cutinase	Probable cutinase Cut2	cell wall and cell processes	-
Rv2303c	-	34550.5	Cytoplasmic	antibiotic-resistance protein	Probable antibiotic-resistance protein	virulence, detoxification, adaptation	COG2159R
Rv2305	-	46124	Cytoplasmic	hypothetical protein Rv2305	Uncharacterized protein containing a NRPS condensation (elongation) domain	cell wall and cell processes	COG4908R
Rv2314c	-	48712	Cytoplasmic Membrane	hypothetical protein Rv2314c	Predicted Zn-dependent proteases and their inactivated homologs	Intermediary metabolism and respiration	COG0312R
Rv2315c	-	54546.1	Unknown	hypothetical protein Rv2315c	Predicted Zn-dependent proteases and their inactivated homologs	Intermediary metabolism and respiration	COG0312R
Rv2319c	-	31834.6	Unknown	universal stress protein	Universal stress protein family protein	virulence, detoxification, adaptation	COG0589T
Rv2326c	-	72835.5	Cytoplasmic Membrane	ABC transporter ATP-binding protein	Possible transmembrane ATP-binding protein ABC transporter	cell wall and cell processes	COG1123R
Rv2332	mez	59422.6	Cytoplasmic Membrane	malate oxidoreductase	Probable [NAD] dependent malate oxidoreductase Mez (malic enzyme) (NAD-malic enzyme) (malate dehydrogenase (oxaloacetate decarboxylating)) (pyruvic-malic carboxylase) (NAD-me)	intermediary metabolism and respiration	COG0281C
Rv2334	cysK1	32752.7	Cytoplasmic	O-acetylserine sulfhydrylase	Cysteine synthase a CysK1 (O-acetylserine sulfhydrylase A) (O-acetylserine (thiol)-lyase A) (CSASE A)	intermediary metabolism and respiration	COG0031E

Rv2335	cysE	23770.4	Cytoplasmic	serine acetyltransferase	Probable serine acetyltransferase CysE (sat)	intermediary metabolism and respiration	COG1045E
Rv2337c	-	41272.4	Cytoplasmic	hypothetical protein Rv2337c	Hypothetical protein	conserved hypotheticals	-
Rv2338c	moeW	35271.3	Cytoplasmic	molybdopterin biosynthesis protein MoeW	Possible molybdopterin biosynthesis protein MoeW	intermediary metabolism and respiration	COG0476H
Rv2343c	dnaG	69593	Cytoplasmic	DNA primase	Probable DNA primase DnaG	information pathways	COG0358L
Rv2344c	dgt	46879.6	Cytoplasmic	deoxyguanosine triphosphate triphosphohydrolase	Probable deoxyguanosine triphosphate triphosphohydrolase Dgt (dGTPase) (dGTP triphosphohydrolase)	intermediary metabolism and respiration	COG0232F
Rv2345	-	70029.7	Cytoplasmic Membrane	transmembrane protein	Possible conserved transmembrane protein	cell wall and cell processes	COG1512R
Rv2346c	esxO	9954.01	Cytoplasmic	ESAT-6 like protein EsxO	Putative ESAT-6 like protein EsxO (ESAT-6 like protein 6)	cell wall and cell processes	-
Rv2350c	plcB	55623.3	Cytoplasmic Membrane	membrane-associated phospholipase B	Membrane-associated phospholipase C 2 PlcB	intermediary metabolism and respiration	COG3511M
Rv2351c	plcA	55381.2	Cytoplasmic Membrane	membrane-associated phospholipase A	Membrane-associated phospholipase C 1 PlcA (MTP40 antigen)	intermediary metabolism and respiration	COG3511M
Rv2357c	glyS	52905.7	Cytoplasmic	glycine--tRNA ligase	Probable glycyl-tRNA synthetase GlyS (glycine--tRNA ligase) (GLYRS)	information pathways	COG0423J
Rv2361c	-	33759.3	Cytoplasmic	decaprenyl diphosphate synthase	Long (C50) chain Z-isoprenyl diphosphate synthase (Z-decaprenyl diphosphate synthase)	cell wall and cell processes	COG0020I
Rv2363	amiA2	50851.7	Cytoplasmic	amidase	Probable amidase AmiA2 (aminohydrolase)	intermediary metabolism and respiration	COG0154J
Rv2368c	phoH1	37767.3	Cytoplasmic	protein PhoH	Probable PHOH-like protein PhoH1 (phosphate starvation-inducible protein PSIH)	intermediary metabolism and respiration	COG1702T
Rv2373c	dnaJ2	40456.8	Cytoplasmic	chaperone protein DnaJ	Probable chaperone protein DnaJ2	virulence, detoxification, adaptation	COG0484O
Rv2374c	hrcA	36474.5	Cytoplasmic	heat-inducible transcription repressor HrcA	Probable heat shock protein transcriptional repressor HrcA	virulence, detoxification, adaptation	COG1420K
Rv2378c	mbtG	46943.9	Unknown	L-lysine N6-monoxygenase	Lysine-N-oxygenase MbtG (L-lysine 6-monoxygenase) (lysine N6-hydroxylase)	lipid metabolism	COG3486Q
Rv2379c	mbtF	156748	Cytoplasmic	peptide synthetase	Peptide synthetase MbtF (peptide synthase)	lipid metabolism	COG1020Q
Rv2383c	mbtB	151602	Cytoplasmic	phenyloxazoline synthase	Phenyloxazoline synthase MbtB (phenyloxazoline synthetase)	lipid metabolism	COG1020Q
Rv2387	-	45008.5	Cytoplasmic Membrane	hypothetical protein Rv2387	Predicted permease	cell wall and cell processes	COG3329R
Rv2388c	hemN	40333.6	Cytoplasmic	oxygen-independent	Probable oxygen-independent coproporphyrinogen III oxidase	intermediary	COG0635H

				coproporphyrinogen III oxidase	HemN (coproporphyrinogenase) (coprogen oxidase)	metabolism and respiration	
Rv2391	sirA	62997.5	Cytoplasmic	sulfite reductase	Ferredoxin-dependent sulfite reductase SirA	intermediary metabolism and respiration	COG0155P
Rv2392	cysH	27422.9	Cytoplasmic	phosphoadenosine phosphosulfate reductase	Probable 3'-phosphoadenosine 5'-phosphosulfate reductase CysH (PAPS reductase, thioredoxin DEP.) (padops reductase) (3'-phosphoadenylylsulfate reductase) (PAPS sulfotransferase)	intermediary metabolism and respiration	COG0175EH
Rv2394	ggtB	66527.5	Extracellular	gamma-glutamyltranspeptidase precursor GgtB	Probable gamma-glutamyltranspeptidase precursor GgtB (gamma-glutamyltransferase) (glutamyl transpeptidase)	intermediary metabolism and respiration	COG0405E
Rv2397c	cysA1	37443.8	Cytoplasmic Membrane	sulfate ABC transporter ATP-binding protein CysA	Sulfate-transport ATP-binding protein ABC transporter CysA1	cell wall and cell processes	COG1118P
Rv2400c	subI	37418.3	Unknown	sulfate ABC transporter substrate-binding lipoprotein SubI	Probable sulfate-binding lipoprotein SubI	cell wall and cell processes	COG1613P
Rv2402	-	72734	Cytoplasmic Membrane	trehalase	Glucoamylase and related glycosyl hydrolases	Intermediary metabolism and respiration	-
Rv2403c	lppR	26306.8	Unknown	lipoprotein LppR	Probable conserved lipoprotein LppR	cell wall and cell processes	-
Rv2404c	lepA	72363.7	Cytoplasmic Membrane	GTP-binding protein LepA	Probable GTP-binding protein LepA (GTP-binding elongation factor)	intermediary metabolism and respiration	COG0481M
Rv2405	-	21241.1	Cytoplasmic Membrane	hypothetical protein Rv2405	lipid metabolism	Lipid metabolism	-
Rv2406c	-	15053.5	Cytoplasmic	hypothetical protein Rv2406c	FOG- CBS domain	Regulatory proteins	COG2905T
Rv2409c	-	31028.8	Cytoplasmic	hypothetical protein Rv2409c	Transglutaminase-like enzymes, putative cysteine proteases	Intermediary metabolism and respiration	COG1305E
Rv2410c	-	36020.6	Cytoplasmic Membrane	hypothetical protein Rv2410c	proteasome	Intermediary metabolism and respiration	COG2307S
Rv2411c	-	61383.4	Cytoplasmic	hypothetical protein Rv2411c	proteasome	Intermediary metabolism and respiration	COG2308S
Rv2412	rpsT	9372.81	Cytoplasmic	30S ribosomal protein S20	30S ribosomal protein S20 RpsT	information pathways	COG0268J
Rv2413c	-	33113.7	Cellwall	hypothetical protein Rv2413c	DNA polymerase III, delta subunit	Information pathways	-
Rv2416c	eis	43771.8	Unknown	enhanced intracellular survival protein	Enhanced intracellular survival protein Eis, GCN5-related N-acetyltransferase	virulence, detoxification, adaptation	COG4552R
Rv2417c	-	28455.7	Unknown	DegV domain-containing protein	Conserved protein	conserved hypotheticals	COG1307S
Rv2418c	-	27225.2	Cytoplasmic	hypothetical protein Rv2418c	Unknown protein	conserved hypotheticals	-
Rv2419c	gpgP	24174.3	Unknown	glucosyl-3-phosphoglycerate	Glucosyl-3-phosphoglycerate phosphatase GpgP	intermediary	COG0406G

				phosphatase		metabolism and respiration	
Rv2420c	-	13826.5	Cytoplasmic	hypothetical protein Rv2420c	Uncharacterized homolog of plant Iojap protein	Regulatory proteins	COG0799S
Rv2421c	nadD	23053	Cytoplasmic	nicotinate-nucleotide adenyltransferase	Probable nicotinate-nucleotide adenyltransferase NadD (deamido-NAD(+) pyrophosphorylase) (deamido-NAD(+) diphosphorylase) (nicotinate mononucleotide adenyltransferase) (NAMN adenyltransferase)	intermediary metabolism and respiration	COG1057H
Rv2423	-	38958.2	Cytoplasmic	hypothetical protein Rv2423	SAM-dependent methyltransferases	Intermediary metabolism and respiration	COG0030J
Rv2426c	-	31918.7	Cytoplasmic	hypothetical protein Rv2426c	MoxR-like ATPases	Intermediary metabolism and respiration	COG0714R
Rv2427c	proA	43744.7	Cytoplasmic	gamma-glutamyl phosphate reductase	Probable gamma-glutamyl phosphate reductase protein ProA (GPR) (glutamate-5-semialdehyde dehydrogenase) (glutamyl-gamma-semialdehyde dehydrogenase)	intermediary metabolism and respiration	COG0014E
Rv2428	ahpC	21566.3	Cytoplasmic	alkyl hydroperoxide reductase subunit C	Alkyl hydroperoxide reductase C protein AhpC (alkyl hydroperoxidase C)	virulence, detoxification, adaptation	COG0450O
Rv2429	ahpD	18780.6	Unknown	alkyl hydroperoxide reductase AphD	Alkyl hydroperoxide reductase D protein AhpD (alkyl hydroperoxidase D)	virulence, detoxification, adaptation	COG2128S
Rv2438c	nadE	74682.5	Unknown	glutamine-dependent NAD(+) synthetase	Glutamine-dependent NAD(+) synthetase NadE (NAD(+) synthase [glutamine-hydrolysing])	intermediary metabolism and respiration	COG0171H
Rv2439c	proB	38788.3	Unknown	glutamate 5-kinase protein	Probable glutamate 5-kinase protein ProB (gamma-glutamyl kinase) (GK)	intermediary metabolism and respiration	COG0263E
Rv2440c	obg	50429	Unknown	GTPase Obg	Probable GTP1/Obg-family GTP-binding protein Obg	intermediary metabolism and respiration	COG0536R
Rv2441c	rpmA	8968.23	Cytoplasmic	50S ribosomal protein L27	50S ribosomal protein L27 RpmA	information pathways	COG0211J
Rv2442c	rplU	11151.1	Unknown	50S ribosomal protein L21	50S ribosomal protein L21 RplU	information pathways	COG0261J
Rv2444c	rne	103358	Cytoplasmic	ribonuclease E	Possible ribonuclease E Rne	information pathways	COG1530J
Rv2445c	ndkA	14475.5	Cytoplasmic	nucleoside diphosphate kinase	Probable nucleoside diphosphate kinase NdkA (NDK) (NDP kinase) (nucleoside-2-P kinase)	intermediary metabolism and respiration	COG0105F
Rv2447c	folC	50779.3	Cytoplasmic	folylpolyglutamate synthase FolC	Probable folylpolyglutamate synthase protein FolC (folylpoly-gamma-glutamate synthetase) (FPGS)	intermediary metabolism and respiration	COG0285H
Rv2448c	valS	97820.5	Cytoplasmic	valine--tRNA ligase	Probable valyl-tRNA synthase protein ValS (valyl-tRNA synthetase) (valine--tRNA ligase) (valine transase)	information pathways	COG0525J
Rv2449c	-	44346.5	Cytoplasmic	trans-acting enoyl reductase	Conserved protein	conserved hypotheticals	COG3268S
Rv2453c	mobA	21052.4	Cytoplasmic	molybdenum cofactor	Probable molybdopterin-guanine dinucleotide biosynthesis	intermediary	COG0746H

				guanylyltransferase	protein A MobA	metabolism and respiration	
Rv2454c	-	40143.2	Cytoplasmic	2-oxoglutarate oxidoreductase subunit KorB	Probable oxidoreductase (beta subunit)	intermediary metabolism and respiration	COG1013C
Rv2455c	-	69150.4	Cytoplasmic	2-oxoglutarate oxidoreductase subunit KorA	Probable oxidoreductase (alpha subunit)	intermediary metabolism and respiration	COG0674C
Rv2457c	clpX	46782.7	Cytoplasmic	ATP-dependent CLP protease ATP-binding subunit ClpX	Probable ATP-dependent CLP protease ATP-binding subunit ClpX	intermediary metabolism and respiration	COG1219O
Rv2460c	clpP2	23507.7	Cytoplasmic	ATP-dependent CLP protease proteolytic subunit 2	Probable ATP-dependent CLP protease proteolytic subunit 2 ClpP2 (endopeptidase CLP 2)	intermediary metabolism and respiration	COG0740OU
Rv2461c	clpP1	21675.6	Cytoplasmic	ATP-dependent CLP protease proteolytic subunit 1	Probable ATP-dependent CLP protease proteolytic subunit 1 ClpP1 (endopeptidase CLP)	intermediary metabolism and respiration	COG0740OU
Rv2462c	tig	50584	Cytoplasmic	trigger factor	Probable trigger factor (TF) protein Tig	cell wall and cell processes	COG0544O
Rv2465c	rpiB	17277.5	Cytoplasmic	ribose-5-phosphate isomerase B	Ribose-5-phosphate isomerase	intermediary metabolism and respiration	COG0698G
Rv2466c	-	23035	Cytoplasmic	hypothetical protein Rv2466c	Protein disulfide oxidoreductase	Intermediary metabolism and respiration	COG2761Q
Rv2467	pepN	94252.7	Cytoplasmic	aminopeptidase PepN	Probable aminopeptidase N PepN (Lysyl aminopeptidase) (LYS-AP) (alanine aminopeptidase)	intermediary metabolism and respiration	COG0308E
Rv2468c	-	17288.6	Cytoplasmic	hypothetical protein Rv2468c	Conserved protein	conserved hypotheticals	-
Rv2473	-	25101.3	Unknown	alanine/proline-rich membrane protein	Possible alanine and proline rich membrane protein	cell wall and cell processes	-
Rv2475c	-	15733	Cytoplasmic	hypothetical protein Rv2475c	Predicted thioesterase	Intermediary metabolism and respiration	COG0824R
Rv2476c	gdh	176900	Cytoplasmic Membrane	NAD-dependent glutamate dehydrogenase	Probable NAD-dependent glutamate dehydrogenase Gdh (NAD-Gdh) (NAD-dependent glutamic dehydrogenase)	intermediary metabolism and respiration	COG2902E
Rv2477c	-	61893.1	Cytoplasmic	macrolide ABC transporter ATP-binding protein	Probable macrolide-transport ATP-binding protein ABC transporter	cell wall and cell processes	COG0488R
Rv2482c	plsB2	88283	Cytoplasmic Membrane	glycerol-3-phosphate acyltransferase	Probable glycerol-3-phosphate acyltransferase PlsB2 (GPAT)	lipid metabolism	COG2937I
Rv2483c	plsC	61788.1	Cytoplasmic Membrane	bifunctional L-3-phosphoserine phosphatase/1-acyl-sn-glycerol-3-phosphate acyltransferase	Possible transmembrane phospholipid biosynthesis bifunctional enzyme PlsC: putative L-3-phosphoserine phosphatase (O-phosphoserine phosphohydrolase) (PSP) (pspase) + 1-acyl-SN-	lipid metabolism	COG0560E

					glycerol-3-phosphate acyltransferase (1-AGP acyltransferase) (1-AGPAT) (lysophosphatidic acid acyltransferase) (LPAAT)		
Rv2484c	-	52309.1	Cytoplasmic	diacylglycerol O-acyltransferase	Possible triacylglycerol synthase (diacylglycerol acyltransferase)	lipid metabolism	-
Rv2486	echA14	26280.1	Cytoplasmic	enoyl-CoA hydratase EchA14	Probable enoyl-CoA hydratase EchA14 (enoyl hydratase) (unsaturated acyl-CoA hydratase) (crotonase)	lipid metabolism	COG1024I
Rv2488c	-	120818	Cytoplasmic	LuxR family transcriptional regulator	Probable transcriptional regulatory protein (LuxR-family)	Regulatory proteins	COG3903R
Rv2494	vapC38	15392.9	Unknown	ribonuclease VapC38	Possible toxin VapC38. Contains PIN domain.	virulence, detoxification, adaptation	COG1848R
Rv2495c	bkdC	41061	Cytoplasmic	branched-chain keto acid dehydrogenase E2 component	Probable branched-chain keto acid dehydrogenase E2 component BkdC	intermediary metabolism and respiration	COG0508C
Rv2496c	bkdB	38064.3	Cytoplasmic	branched-chain keto acid dehydrogenase E1 component subunit beta	Probable branched-chain keto acid dehydrogenase E1 component, beta subunit BkdB	intermediary metabolism and respiration	COG0022C
Rv2497c	bkdA	40615.8	Cytoplasmic	branched-chain keto acid dehydrogenase E1 component subunit alpha	Probable branched-chain keto acid dehydrogenase E1 component, alpha subunit BkdA	intermediary metabolism and respiration	COG1071C
Rv2498c	citE	28885.8	Cytoplasmic	citrate (pro-3S)-lyase subunit beta	Probable citrate (pro-3S)-lyase (beta subunit) CitE (citrase) (citratase) (citritase) (citridesmolase) (citrase aldolase)	intermediary metabolism and respiration	COG2301G
Rv2500c	fadE19	41916.4	Cytoplasmic	acyl-CoA dehydrogenase FadE19	Possible acyl-CoA dehydrogenase FadE19 (MMGC)	lipid metabolism	COG1960I
Rv2501c	accA1	70592.2	Cytoplasmic	acetyl-/propionyl-CoA carboxylase subunit alpha	Probable acetyl-/propionyl-coenzyme A carboxylase alpha chain (alpha subunit) AccA1: biotin carboxylase + biotin carboxyl carrier protein (BCCP)	lipid metabolism	COG4770I
Rv2502c	accD1	56745.7	Cytoplasmic	acetyl-/propionyl-CoA carboxylase subunit beta	Probable acetyl-/propionyl-CoA carboxylase (beta subunit) AccD1	lipid metabolism	COG4799I
Rv2503c	scoB	22898.4	Cytoplasmic	succinyl-CoA:3-ketoacid-CoA transferase subunit B	Probable succinyl-CoA:3-ketoacid-coenzyme A transferase (beta subunit) ScoB (3-oxo-acid:CoA transferase) (OXCT B) (succinyl CoA:3-oxoacid CoA-transferase)	lipid metabolism	COG2057I
Rv2504c	scoA	26275.7	Cytoplasmic	succinyl-CoA:3-ketoacid-CoA transferase subunit A	Probable succinyl-CoA:3-ketoacid-coenzyme A transferase (alpha subunit) ScoA (3-oxo acid:CoA transferase) (OXCT A) (succinyl-CoA:3-oxoacid-coenzyme A transferase)	lipid metabolism	COG1788I
Rv2505c	fadD35	59746.4	Cytoplasmic	fatty-acid--CoA ligase FadD35	Probable fatty-acid-CoA ligase FadD35 (fatty-acid-CoA synthetase) (fatty-acid-CoA synthase)	lipid metabolism	COG0318IQ
Rv2509	-	28013.2	Cytoplasmic	short-chain type dehydrogenase/reductase	Probable short-chain type dehydrogenase/reductase	intermediary metabolism and respiration	COG0300R
Rv2510c	-	56546.7	Unknown	hypothetical protein Rv2510c	Predicted ATPase	Intermediary metabolism and respiration	COG0433R
Rv2511	orn	23410.4	Cytoplasmic	oligoribonuclease	Oligoribonuclease Orn	intermediary metabolism and	COG1949A

						respiration	
Rv2516c	-	29355.3	Cytoplasmic	hypothetical protein Rv2516c	Hypothetical protein	conserved hypotheticals	-
Rv2518c	ldtB	43365.9	Unknown	L,D-transpeptidase LdtB	Probable L,D-transpeptidase LdtB	cell wall and cell processes	COG1376S
Rv2520c	-	8341.7	Cytoplasmic	membrane protein	Possible conserved membrane protein	cell wall and cell processes	-
Rv2521	bcp	17045.3	Cellwall	peroxiredoxin	Probable bacterioferritin comigratory protein Bcp	virulence, detoxification, adaptation	COG1225O
Rv2524c	fas	326222	Cytoplasmic Membrane	fatty acid synthase	Probable fatty acid synthase Fas (fatty acid synthetase)	lipid metabolism	COG4982I
Rv2528c	mrr	33648.2	Cytoplasmic	restriction system protein	Probable restriction system protein Mrr	information pathways	COG1715V
Rv2529	-	50160.1	Cytoplasmic	hypothetical protein Rv2529	Nuclease	Information pathways	COG1948L
Rv2531c	-	106029	Cytoplasmic	amino acid decarboxylase	Probable amino acid decarboxylase	intermediary metabolism and respiration	COG1982E
Rv2533c	nusB	16740.2	Cytoplasmic	N utilization substance protein B	N utilization substance protein NusB (NusB protein)	information pathways	COG0781K
Rv2534c	efp	20375.1	Cytoplasmic	elongation factor P	Probable elongation factor P Efp	information pathways	COG0231J
Rv2535c	pepQ	38758.9	Cytoplasmic	cytoplasmic peptidase PepQ	Probable cytoplasmic peptidase PepQ	intermediary metabolism and respiration	COG0006E
Rv2536	-	24626.3	Cytoplasmic Membrane	transmembrane protein	Probable conserved transmembrane protein	cell wall and cell processes	-
Rv2537c	aroD	15790	Cytoplasmic	3-dehydroquinate dehydratase	3-dehydroquinate dehydratase AroD (AROQ) (3- dehydroquinase) (type II dhqase)	intermediary metabolism and respiration	COG0757E
Rv2538c	aroB	38119.1	Cytoplasmic	3-dehydroquinate synthase	3-dehydroquinate synthase AroB	intermediary metabolism and respiration	COG0337E
Rv2540c	aroF	41760.4	Cytoplasmic	chorismate synthase	Probable chorismate synthase AroF (5-enolpyruvylshikimate-3- phosphate phospholyase)	intermediary metabolism and respiration	COG0082E
Rv2553c	-	45473.8	Cytoplasmic	membrane protein	Probable conserved membrane protein	cell wall and cell processes	COG1559R
Rv2554c	-	18033.4	Cytoplasmic	Holliday junction resolvase	Predicted endonuclease involved in recombination (possible Holliday junction resolvase in Mycoplasmas and B. subtilis)	Information pathways	COG0816L
Rv2555c	alaS	97324.8	Cytoplasmic	alanine--tRNA ligase	Probable alanyl-tRNA synthetase AlaS (alanine--tRNA ligase) (alanine transase) (ALARS)	information pathways	COG0013J
Rv2557	-	24277.3	Cytoplasmic	hypothetical protein Rv2557	Conserved protein	conserved hypotheticals	-
Rv2558	-	25700.4	Cytoplasmic	hypothetical protein Rv2558	Conserved protein	conserved hypotheticals	-
Rv2559c	-	47499.2	Cytoplasmic	alanine/leucine/valine-rich protein	ATPase related to the helicase subunit of the Holliday junction	Information pathways	COG2256L

					resolvase		
Rv2563	-	36025.4	Cytoplasmic Membrane	glutamine ABC transporter permease	Probable glutamine-transport transmembrane protein ABC transporter	cell wall and cell processes	COG4591M
Rv2564	glnQ	35428.7	Cytoplasmic Membrane	glutamine ABC transporter ATP-binding protein	Probable glutamine-transport ATP-binding protein ABC transporter GlnQ	cell wall and cell processes	COG1136V
Rv2565	-	62123.8	Cytoplasmic	NTE family protein	cAMP-binding proteins - catabolite gene activator and regulatory subunit of cAMP-dependent protein kinases	Regulatory proteins	COG1752R
Rv2566	-	124504	Cytoplasmic	hypothetical protein Rv2566	Transglutaminase-like enzymes, putative cysteine proteases	Intermediary metabolism and respiration	COG4196S
Rv2567	-	95448.6	Cytoplasmic Membrane	alanine/leucine-rich protein	proteasome	Intermediary metabolism and respiration	COG2308S
Rv2569c	-	34378.5	Unknown	hypothetical protein Rv2569c	Transglutaminase-like enzymes, putative cysteine proteases	Intermediary metabolism and respiration	COG1305E
Rv2571c	-	36789.9	Cytoplasmic Membrane	transmembrane protein	Probable transmembrane alanine and valine and leucine rich protein	cell wall and cell processes	COG4129S
Rv2572c	aspS	65194.8	Cytoplasmic	aspartate--tRNA ligase	Probable aspartyl-tRNA synthetase AspS (aspartate--tRNA ligase) (ASPRS) (aspartic acid translase)	information pathways	COG0173J
Rv2574	-	19001.9	Cytoplasmic Membrane	hypothetical protein Rv2574	Conserved protein	conserved hypotheticals	COG3427S
Rv2575	-	30806.2	Unknown	membrane protein	Possible conserved membrane glycine rich protein	cell wall and cell processes	COG2321R
Rv2577	-	57239.5	Unknown	hypothetical protein Rv2577	Predicted phosphohydrolases	Intermediary metabolism and respiration	COG1409R
Rv2579	dhaA	33728.2	Cytoplasmic	haloalkane dehalogenase	Possible haloalkane dehalogenase DhaA (1-chlorohexane halidohydrolase)	intermediary metabolism and respiration	COG0596R
Rv2580c	hisS	45117.1	Cytoplasmic	histidine--tRNA ligase	Probable histidyl-tRNA synthetase HisS (histidine--tRNA ligase) (HISRS) (histidine--translase)	information pathways	COG0124J
Rv2581c	-	24165.6	Cytoplasmic	glyoxalase II	Possible glyoxalase II (hydroxyacylglutathione hydrolase) (GLX II)	virulence, detoxification, adaptation	COG0491R
Rv2582	ppiB	32370.9	Cytoplasmic	peptidyl-prolyl cis-trans isomerase B	Probable peptidyl-prolyl cis-trans isomerase B PpiB (cyclophilin) (PPIase) (rotamase) (peptidylprolyl isomerase)	information pathways	COG0652O
Rv2583c	relA	87321.8	Cytoplasmic	bifunctional (p)ppGpp synthase/hydrolase RelA	Probable GTP pyrophosphokinase RelA (ATP:GTP 3'-pyrophosphotransferase) (PPGPP synthetase I) ((P)PPGPP synthetase) (GTP diphosphokinase)	intermediary metabolism and respiration	COG0317TK
Rv2584c	apt	23155.7	Cytoplasmic	adenine phosphoribosyltransferase	Adenine phosphoribosyltransferase Apt (APRT) (AMP diphosphorylase) (AMP pyrophosphorylase) (transphosphoribosidase)	intermediary metabolism and respiration	COG0503F
Rv2585c	-	58602.6	Unknown	lipoprotein	Possible conserved lipoprotein	cell wall and cell processes	COG0747E

Rv2586c	secF	47013.6	Cytoplasmic Membrane	protein translocase subunit SecF	Probable protein-export membrane protein SecF	cell wall and cell processes	COG0341U
Rv2587c	secD	60235.1	Cytoplasmic Membrane	protein translocase subunit SecD	Probable protein-export membrane protein SecD	cell wall and cell processes	COG0342U
Rv2588c	yajC	12966.6	Cytoplasmic	membrane protein secretion factor YajC	Probable conserved membrane protein secretion factor YajC	cell wall and cell processes	COG1862U
Rv2589	gabT	46780.6	Cytoplasmic	4-aminobutyrate aminotransferase	4-aminobutyrate aminotransferase GabT (gamma-amino-N-butyrate transaminase) (GABA transaminase) (glutamate:succinic semialdehyde transaminase) (GABA aminotransferase) (GABA-at)	intermediary metabolism and respiration	COG0160E
Rv2590	fadD9	127328	Cytoplasmic	fatty-acid--CoA ligase FadD9	Probable fatty-acid-CoA ligase FadD9 (fatty-acid-CoA synthetase) (fatty-acid-CoA synthase)	lipid metabolism	COG1022I
Rv2597	-	21910.9	Unknown	membrane protein	Probable membrane protein	cell wall and cell processes	-
Rv2603c	-	26798.8	Cytoplasmic	transcriptional regulator	Highly conserved protein	conserved hypotheticals	COG0217S
Rv2605c	tesB2	31536.7	Cytoplasmic	acyl-CoA thioesterase TesB	Probable acyl-CoA thioesterase II TesB2 (TEII)	lipid metabolism	COG1946I
Rv2606c	snzP	31365.1	Cytoplasmic	pyridoxine biosynthesis protein	Possible pyridoxine biosynthesis protein SnzP	intermediary metabolism and respiration	COG0214H
Rv2607	pdxH	25186.3	Cytoplasmic	pyridoxine/pyridoxamine 5'-phosphate oxidase	Probable pyridoxamine 5'-phosphate oxidase PdxH (PNP/PMP oxidase) (pyridoxinephosphate oxidase) (PNPOX) (pyridoxine 5'-phosphate oxidase)	intermediary metabolism and respiration	COG0259H
Rv2609c	-	38096.4	Cytoplasmic	membrane protein	Probable conserved membrane protein	cell wall and cell processes	COG0494LR
Rv2610c	pimA	40445.5	Cytoplasmic	alpha-(1-2)-phosphatidylinositol mannosyltransferase	Alpha-mannosyltransferase PimA	lipid metabolism	COG0438M
Rv2611c	-	35094.3	Unknown	phosphatidylinositol mannoside acyltransferase	Probable acyltransferase	lipid metabolism	COG1560M
Rv2613c	-	21836.7	Cytoplasmic Membrane	AP-4-A phosphorylase	Conserved protein	lipid metabolism	COG0537FGR
Rv2614c	thrS	77121.9	Cytoplasmic	threonine--tRNA ligase	Probable threonyl-tRNA synthetase ThrS (threonine-tRNA synthetase)(ThrRS) (threonine-tRNA ligase)	information pathways	COG0441J
Rv2622	-	29542	Cytoplasmic	methyltransferase	Possible methyltransferase (methylase)	intermediary metabolism and respiration	COG2226H
Rv2623	TB31.7	31652.1	Cytoplasmic	universal stress protein	Universal stress protein family protein TB31.7	virulence, detoxification, adaptation	COG0589T
Rv2624c	-	29400.5	Cytoplasmic	universal stress protein	Universal stress protein family protein	virulence, detoxification, adaptation	COG0589T
Rv2625c	-	41480.2	Cytoplasmic Membrane	zinc metalloprotease Rip3	Probable conserved transmembrane alanine and leucine rich protein	cell wall and cell processes	COG1994R
Rv2626c	hrp1	15517.7	Unknown	hypoxic response protein	FOG- CBS domain	Regulatory proteins	COG3620K

Rv2627c	-	46252.3	Cytoplasmic	hypothetical protein Rv2627c	Conserved protein	conserved hypotheticals	-
Rv2629	-	40839.9	Cytoplasmic	hypothetical protein Rv2629	Conserved protein	conserved hypotheticals	COG1503J
Rv2632c	-	10083.3	Cytoplasmic	hypothetical protein Rv2632c	Conserved protein	conserved hypotheticals	-
Rv2668	-	18345.8	Unknown	alanine/valine-rich protein	Possible exported alanine and valine rich protein	cell wall and cell processes	-
Rv2671	ribD	27693.5	Cytoplasmic	bifunctional diaminohydroxyphosphoribosylaminopyrimidine deaminase/5-amino-6-(5-phosphoribosylamino)uracil reductase	Possible bifunctional enzyme riboflavin biosynthesis protein RibD: diaminohydroxyphosphoribosylaminopyrimidine deaminase (riboflavin-specific deaminase) + 5-amino-6-(5-phosphoribosylamino)uracil reductase (HTP reductase)	intermediary metabolism and respiration	COG1985H
Rv2672	-	54015.7	Unknown	protease	Possible secreted protease	intermediary metabolism and respiration	COG0596R
Rv2673	aftC	48882.7	Cytoplasmic Membrane	alpha-(1-3)-arabinofuranosyltransferase	Possible arabinofuranosyltransferase AftC	cell wall and cell processes	-
Rv2675c	-	27545.1	Cytoplasmic	hypothetical protein Rv2675c	SAM-dependent methyltransferases	Intermediary metabolism and respiration	COG2226H
Rv2676c	-	26249.8	Cytoplasmic	hypothetical protein Rv2676c	porphyrin biosynthetic process	Regulatory proteins	COG3253S
Rv2677c	hemY	46846.4	Cytoplasmic	protoporphyrinogen oxidase	Probable protoporphyrinogen oxidase HemY (protoporphyrinogen-IX oxidase) (protoporphyrinogenase) (PPO)	intermediary metabolism and respiration	COG1232H
Rv2678c	hemE	37606.3	Cytoplasmic	uroporphyrinogen decarboxylase	Probable uroporphyrinogen decarboxylase HemE (uroporphyrinogen III decarboxylase) (URO-D) (UPD)	intermediary metabolism and respiration	COG0407H
Rv2679	echA15	29634.9	Cytoplasmic	enoyl-CoA hydratase EchA15	Probable enoyl-CoA hydratase EchA15 (enoyl hydratase) (unsaturated acyl-CoA hydratase) (crotonase)	lipid metabolism	COG1024I
Rv2680	-	22555.1	Cytoplasmic	hypothetical protein Rv2680	DNA/RNA-associated process	Regulatory proteins	-
Rv2681	-	47224.5	Cytoplasmic	alanine-rich protein	Ribonuclease D	Information pathways	COG0349J
Rv2682c	dxs1	67885.4	Cytoplasmic	1-deoxy-D-xylulose 5-phosphate synthase	Probable 1-deoxy-D-xylulose 5-phosphate synthase Dxs1 (1-deoxyxylulose-5-phosphate synthase) (DXP synthase) (DXPS)	intermediary metabolism and respiration	COG1154HI
Rv2683	-	17729.5	Unknown	hypothetical protein Rv2683	FOG- CBS domain	Regulatory proteins	-
Rv2688c	-	33099.7	Cytoplasmic Membrane	antibiotic ABC transporter ATP-binding protein	Antibiotic-transport ATP-binding protein ABC transporter	cell wall and cell processes	COG1131V
Rv2691	ceoB	24239.5	Cytoplasmic Membrane	TRK system potassium uptake protein CeoB	TRK system potassium uptake protein CeoB	cell wall and cell processes	COG0569P
Rv2692	ceoC	23947.4	Cytoplasmic Membrane	TRK system potassium uptake protein CeoC	TRK system potassium uptake protein CeoC	cell wall and cell processes	COG0569P
Rv2694c	-	13570.3	Cytoplasmic	hypothetical protein Rv2694c	Nucleic acid binding, OB-fold,tRNA/helicase-type	Information pathways	COG1200LK

Rv2696c	-	27216.7	Cytoplasmic	hypothetical protein Rv2696c	nucleotide metabolism	Regulatory proteins	-
Rv2697c	dut	15770.9	Cytoplasmic Membrane	deoxyuridine 5'-triphosphate nucleotidohydrolase	Probable deoxyuridine 5'-triphosphate nucleotidohydrolase Dut (dUTPase) (dUTP pyrophosphatase) (deoxyuridine 5'-triphosphatase) (dUTP diphosphatase) (deoxyuridine-triphosphatase)	intermediary metabolism and respiration	COG0756F
Rv2701c	suhB	29995	Cytoplasmic	inositol-1-monophosphatase SuhB	Inositol-1-monophosphatase SuhB	intermediary metabolism and respiration	COG0483G
Rv2702	ppgK	27429.1	Cytoplasmic	polyphosphate glucokinase	Polyphosphate glucokinase PpgK (polyphosphate-glucose phosphotransferase)	intermediary metabolism and respiration	COG1940KG
Rv2703	sigA	57768.2	Cytoplasmic	RNA polymerase sigma factor SigA	RNA polymerase sigma factor SigA (sigma-A)	information pathways	COG0568K
Rv2710	sigB	36343.4	Cytoplasmic	RNA polymerase sigma factor SigB	RNA polymerase sigma factor SigB	information pathways	COG0568K
Rv2711	ideR	25232.9	Cytoplasmic	iron-dependent repressor and activator IdeR	Iron-dependent repressor and activator IdeR	Regulatory proteins	COG1321K
Rv2713	sthA	50754	Cytoplasmic	pyridine nucleotide transhydrogenase	Probable soluble pyridine nucleotide transhydrogenase SthA (STH) (NAD(P)(+) transhydrogenase [B-specific]) (nicotinamide nucleotide transhydrogenase)	intermediary metabolism and respiration	COG1249C
Rv2714	-	35519.8	Cytoplasmic	alanine/leucine-rich protein	Conserved alanine and leucine rich protein	conserved hypotheticals	-
Rv2715	-	36932.6	Cytoplasmic	hydrolase	Possible hydrolase	intermediary metabolism and respiration	COG0596R
Rv2716	-	24574.3	Cytoplasmic	hypothetical protein Rv2716	Predicted epimerase, PhzC/PhzF homolog	Intermediary metabolism and respiration	COG0384R
Rv2717c	-	17846.3	Cytoplasmic	hypothetical protein Rv2717c	UPF0678 fatty acid-binding protein-like protein	Lipid metabolism	-
Rv2718c	nrdR	17177.3	Cytoplasmic	transcriptional regulator NrdR	Probable transcriptional regulatory protein NrdR	Regulatory proteins	COG1327K
Rv2720	lexA	24841.2	Unknown	repressor LexA	Repressor LexA	Regulatory proteins	COG1974KT
Rv2721c	-	72339.7	Cellwall	transmembrane alanine/glycine-rich protein	Possible conserved transmembrane alanine and glycine rich protein	cell wall and cell processes	COG5479M
Rv2724c	fadE20	42844.7	Cytoplasmic	acyl-CoA dehydrogenase FadE20	Probable acyl-CoA dehydrogenase FadE20	lipid metabolism	COG1960I
Rv2725c	hflX	53327.7	Cytoplasmic	GTP-binding protein HflX	Probable GTP-binding protein HflX	intermediary metabolism and respiration	COG2262R
Rv2726c	dapF	29697.6	Cytoplasmic	diaminopimelate epimerase	Probable diaminopimelate epimerase DapF (DAP epimerase)	intermediary metabolism and respiration	COG0253E
Rv2727c	miaA	34414.3	Cytoplasmic	tRNA delta(2)-isopentenylpyrophosphate transferase	Probable tRNA delta(2)-isopentenylpyrophosphate transferase MiaA (IPP transferase) (isopentenyl-diphosphate:tRNA isopentenyltransferase) (iptase) (IPPT)	intermediary metabolism and respiration	COG0324J
Rv2731	-	49796.1	Cytoplasmic	alanine/arginine-rich protein	Conserved alanine and arginine rich protein	conserved	COG4717S

						hypotheticals	
Rv2732c	-	21989.4	Cytoplasmic Membrane	transmembrane protein	Probable conserved transmembrane protein	cell wall and cell processes	-
Rv2737c	recA	85389.1	Cytoplasmic	recombinase A	RecA protein (recombinase A) [contains: endonuclease PI-MTUI (MTU RecA intein)].	information pathways	COG0468L
Rv2739c	-	39962.3	Unknown	alanine-rich transferase	Possible alanine rich transferase	intermediary metabolism and respiration	COG1819GC
Rv2740	ephG	16593	Unknown	epoxide hydrolase	Epoxide hydrolase	virulence, detoxification, adaptation	COG4308Q
Rv2743c	-	28961.6	Cytoplasmic Membrane	transmembrane alanine-rich protein	Possible conserved transmembrane alanine rich protein	cell wall and cell processes	-
Rv2744c	35kd_ag	29257.9	Cytoplasmic	alanine-rich protein	Phage shock protein A (IM30), suppresses sigma54-dependent transcription	Insertion seqs and phages	COG1842KT
Rv2748c	ftsK	94406.5	Cytoplasmic Membrane	DNA translocase FtsK	Possible cell division transmembrane protein FtsK	cell wall and cell processes	COG1674D
Rv2751	-	33610.4	Cytoplasmic Membrane	hypothetical protein Rv2751	O-Methyltransferase involved in polyketide biosynthesis	Lipid metabolism	COG3315Q
Rv2752c	-	59500.4	Cytoplasmic	ribonuclease J	Predicted hydrolase of the metallo-beta-lactamase superfamily	Intermediary metabolism and respiration	COG0595R
Rv2753c	dapA	30826.2	Cytoplasmic	4-hydroxy-tetrahydrodipicolinate synthase	Probable dihydrodipicolinate synthase DapA (DHDPS) (dihydrodipicolinate synthetase)	intermediary metabolism and respiration	COG0329EM
Rv2754c	thyX	27559.2	Cytoplasmic	thymidylate synthase ThyX	Probable thymidylate synthase ThyX (ts) (TSase)	intermediary metabolism and respiration	COG1351F
Rv2756c	hsdM	60084.1	Cytoplasmic	type I restriction/modification system DNA methylase HsdM	Possible type I restriction/modification system DNA methylase HsdM (M protein) (DNA methyltransferase)	information pathways	COG0286V
Rv2761c	hsdS	39211.9	Cytoplasmic Membrane	type I restriction/modification system specificity determinant HsdS	Possible type I restriction/modification system specificity determinant HsdS (S protein)	information pathways	COG0732V
Rv2763c	dfrA	17640	Cytoplasmic	dihydrofolate reductase	Dihydrofolate reductase DfrA (DHFR) (tetrahydrofolate dehydrogenase)	intermediary metabolism and respiration	COG0262H
Rv2764c	thyA	29820.8	Cytoplasmic	thymidylate synthase ThyA	Probable thymidylate synthase ThyA (ts) (TSASE)	intermediary metabolism and respiration	COG0207F
Rv2765	-	26159.5	Unknown	alanine-rich hydrolase	Probable alanine rich hydrolase	intermediary metabolism and respiration	COG0412Q
Rv2766c	-	27140.1	Cytoplasmic	short-chain type dehydrogenase/reductase	Probable short-chain type dehydrogenase/reductase	intermediary metabolism and respiration	COG1028IQR

Rv2771c	-	16000.3	Cytoplasmic	hypothetical protein Rv2771c	Multimeric flavodoxin WrbA	Intermediary metabolism and respiration	COG0655R
Rv2773c	dapB	25733.4	Cytoplasmic	4-hydroxy-tetrahydrodipicolinate reductase	Dihydrodipicolinate reductase DapB (DHPR)	intermediary metabolism and respiration	COG0289E
Rv2775	-	17405.9	Cytoplasmic	GCN5-like N-acetyltransferase	GCN5-related N-acetyltransferase	intermediary metabolism and respiration	-
Rv2778c	-	17268.5	Cytoplasmic	hypothetical protein Rv2778c	Conserved protein	conserved hypotheticals	-
Rv2780	ald	38713.2	Extracellular	L-alanine dehydrogenase	Secreted L-alanine dehydrogenase Ald (40 kDa antigen) (TB43)	intermediary metabolism and respiration	COG0686E
Rv2781c	-	35392.8	Unknown	alanine-rich oxidoreductase	Possible alanine rich oxidoreductase	intermediary metabolism and respiration	COG2070R
Rv2782c	pepR	47072.5	Cytoplasmic	zinc protease	Probable zinc protease PepR	intermediary metabolism and respiration	COG0612R
Rv2783c	gpsI	79734.7	Cytoplasmic	bifunctional guanosine pentaphosphate synthetase/polyribonucleotide nucleotidyltransferase	Bifunctional protein polyribonucleotide nucleotidyltransferase GpsI: guanosine pentaphosphate synthetase + polyribonucleotide nucleotidyltransferase (polynucleotide phosphorylase) (pnpase)	information pathways	COG1185J
Rv2785c	rpsO	10443.2	Cytoplasmic	30S ribosomal protein S15	30S ribosomal protein S15 RpsO	information pathways	COG0184J
Rv2786c	ribF	36058.1	Cytoplasmic	bifunctional riboflavin kinase /FMN adenyltransferase	Probable bifunctional FAD synthetase/riboflavin biosynthesis protein RibF: riboflavin kinase (flavokinase) + FMN adenyltransferase (FAD pyrophosphorylase) (FAD synthetase)(FAD diphosphorylase) (flavin adenine dinucleotide synthetase)	intermediary metabolism and respiration	COG0196H
Rv2788	sirR	24919.1	Cytoplasmic	transcriptional repressor SirR	Probable transcriptional repressor SirR	Regulatory proteins	COG1321K
Rv2789c	fadE21	44743.4	Cytoplasmic	acyl-CoA dehydrogenase FadE21	Probable acyl-CoA dehydrogenase FadE21	lipid metabolism	COG1960I
Rv2790c	ltp1	42904.1	Cytoplasmic	lipid-transfer protein	Probable lipid-transfer protein Ltp1	cell wall and cell processes	-
Rv2793c	truB	31819.5	Cytoplasmic	tRNA pseudouridine synthase B	Probable tRNA pseudouridine synthase B TruB (tRNA pseudouridine 55 synthase) (PSI55 synthase) (pseudouridylate synthase) (uracil hydrolyase)	information pathways	COG0130J
Rv2794c	pptT	24708.5	Cytoplasmic	phosphopantetheinyl transferase	Phosphopantetheinyl transferase PptT (CoA:APO-[ACP]pantetheinophosphotransferase) (CoA:APO-[acyl-carrier protein]pantetheinophosphotransferase)	lipid metabolism	COG2977Q
Rv2795c	-	37568.5	Cytoplasmic	hypothetical protein Rv2795c	Predicted phosphohydrolases	Intermediary metabolism and respiration	COG1409R
Rv2799	-	22870	Cytoplasmic	membrane protein	Probable membrane protein	cell wall and cell	-

			Membrane			processes	
Rv2802c	-	38793.9	Cytoplasmic	arginine/alanine-rich protein	Hypothetical arginine and alanine rich protein	conserved hypotheticals	COG5586S
Rv2826c	-	33056.7	Cytoplasmic	hypothetical protein Rv2826c	Hypothetical protein	conserved hypotheticals	COG2253S
Rv2831	echA16	26630.4	Unknown	enoyl-CoA hydratase EchA16	Probable enoyl-CoA hydratase EchA16 (enoyl hydratase) (unsaturated acyl-CoA hydratase) (crotonase)	lipid metabolism	COG1024I
Rv2837c	-	35383.2	Unknown	bifunctional oligoribonuclease/PAP phosphatase NrnA	Exopolyphosphatase-related proteins	Intermediary metabolism and respiration	COG0618R
Rv2838c	rbfA	18998.3	Cytoplasmic	ribosome-binding factor RbfA	Probable ribosome-binding factor a RbfA (P15B protein)	information pathways	COG0858J
Rv2839c	infB	94008.9	Cytoplasmic	translation initiation factor IF-2	Probable translation initiation factor if-2 InfB	information pathways	COG0532J
Rv2841c	nusA	37641.6	Cytoplasmic	transcription termination/antitermination protein NusA	Probable N utilization substance protein A NusA	information pathways	COG0195K
Rv2842c	-	19533.3	Cytoplasmic	ribosome maturation factor RimP	transcription/translation	Regulatory proteins	COG0779S
Rv2845c	proS	63269.6	Cytoplasmic	proline--tRNA ligase	Probable prolyl-tRNA synthetase ProS (proline--tRNA ligase) (PRORS) (global RNA synthesis factor) (proline transase)	information pathways	COG0442J
Rv2847c	cysG	41938.4	Cytoplasmic	multifunctional uroporphyrin-III C-methyltransferase/precorrin-2 oxidase/ferrochelatase	Possible multifunctional enzyme siroheme synthase CysG: uroporphyrin-III C-methyltransferase (urogen III methylase) (SUMT) (uroporphyrinogen III methylase) (UROM) + precorrin-2 oxidase + ferrochelatase	intermediary metabolism and respiration	COG0007H
Rv2850c	-	66955	Cytoplasmic	magnesium chelatase	Possible magnesium chelatase	intermediary metabolism and respiration	COG1239H
Rv2852c	mqo	53563.2	Cellwall	malate:quinone oxidoreductase	Probable malate:quinone oxidoreductase Mqo (malate dehydrogenase [acceptor])	intermediary metabolism and respiration	COG0579R
Rv2854	-	38191.2	Cytoplasmic	hypothetical protein Rv2854	Lysophospholipase	Lipid metabolism	COG2267I
Rv2855	mtr	49945.7	Cytoplasmic	mycothione reductase	NADPH-dependent mycothiol reductase Mtr	intermediary metabolism and respiration	COG1249C
Rv2857c	-	26772.4	Cytoplasmic	short-chain type dehydrogenase/reductase	Probable short-chain type dehydrogenase/reductase	intermediary metabolism and respiration	COG1028IQR
Rv2858c	aldC	48258.7	Cytoplasmic	aldehyde dehydrogenase AldC	Probable aldehyde dehydrogenase AldC	intermediary metabolism and respiration	COG1012C
Rv2859c	-	32409.6	Cytoplasmic	glutamine amidotransferase	Possible amidotransferase	intermediary metabolism and respiration	COG2071R
Rv2860c	glnA4	49715.4	Cytoplasmic	glutamine synthetase	Probable glutamine synthetase GlnA4 (glutamine synthase) (GS-II)	intermediary metabolism and respiration	COG0174E
Rv2861c	mapB	30891	Cytoplasmic	methionine aminopeptidase	Methionine aminopeptidase MapB (map) (peptidase M)	intermediary	COG0024J

						metabolism and respiration	
Rv2868c	gcpE	40450.3	Cytoplasmic	4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase	Enzyme involved in the deoxyxylulose pathway of isoprenoid biosynthesis	Intermediary metabolism and respiration	COG0821I
Rv2869c	rip	42834.2	Cytoplasmic Membrane	zinc metalloprotease	Membrane bound metalloprotease	cell wall and cell processes	COG0750M
Rv2870c	dxr	42821.4	Cytoplasmic	1-deoxy-D-xylulose 5-phosphate reductoisomerase	Probable 1-deoxy-D-xylulose 5-phosphate reductoisomerase Dxr (DXP reductoisomerase) (1-deoxyxylulose-5-phosphate reductoisomerase)	intermediary metabolism and respiration	COG0743I
Rv2872	vapC43	16595.7	Cytoplasmic	ribonuclease VapC43	Possible toxin VapC43. Contains PIN domain.	virulence, detoxification, adaptation	COG1848R
Rv2873	mpt83	22070.1	Cytoplasmic Membrane	cell surface lipoprotein	Cell surface lipoprotein Mpt83 (lipoprotein P23)	cell wall and cell processes	COG2335M
Rv2878c	mpt53	18383	Extracellular	soluble secreted antigen Mpt53	Soluble secreted antigen Mpt53 precursor	cell wall and cell processes	COG1225O
Rv2882c	frr	20827.6	Cytoplasmic	ribosome recycling factor	Ribosome recycling factor Frr (ribosome releasing factor) (RRF)	information pathways	COG0233J
Rv2883c	pyrH	27429.6	Cytoplasmic	uridylate kinase	Probable uridylate kinase PyrH (UK) (uridine monophosphate kinase) (UMP kinase)	intermediary metabolism and respiration	COG0528F
Rv2888c	amiC	50916.1	Cytoplasmic	amidase AmiC	Probable amidase AmiC (aminohydrolase)	intermediary metabolism and respiration	COG0154J
Rv2889c	tsf	28754.7	Cytoplasmic	elongation factor EF-Ts	Probable elongation factor Tsf (EF-ts)	information pathways	COG0264J
Rv2890c	rpsB	31089.3	Cytoplasmic	30S ribosomal protein S2	30S ribosomal protein S2 RpsB	information pathways	COG0052J
Rv2893	-	34617.2	Cytoplasmic	oxidoreductase	Possible oxidoreductase	intermediary metabolism and respiration	COG2141C
Rv2895c	viuB	30570.6	Cytoplasmic	mycobactin utilization protein ViuB	Possible mycobactin utilization protein ViuB	intermediary metabolism and respiration	COG2375P
Rv2900c	fdhF	84566.1	Cytoplasmic	formate dehydrogenase subunit alpha FdhF	Possible formate dehydrogenase H FdhF (formate-hydrogen-lyase-linked, selenocysteine-containing polypeptide) (formate dehydrogenase-H alpha subunit) (FDH-H)	intermediary metabolism and respiration	COG0243C
Rv2901c	-	12225.8	Cytoplasmic	hypothetical protein Rv2901c	Conserved protein	conserved hypotheticals	-
Rv2903c	lepB	31848.2	Cytoplasmic Membrane	signal peptidase	Probable signal peptidase I LepB (SPASE I) (leader peptidase I).	cell wall and cell processes	COG0681U
Rv2904c	rplS	13013.1	Cytoplasmic	50S ribosomal protein L19	50S ribosomal protein L19 RplS	information pathways	COG0335J
Rv2906c	trmD	25134.4	Unknown	tRNA (guanine-N1)-methyltransferase	Probable tRNA (guanine-N1)-methyltransferase TrmD (M1G-methyltransferase) (tRNA [GM37] methyltransferase)	information pathways	COG0336J
Rv2907c	rimM	18616	Cytoplasmic	16S rRNA processing protein RimM	Probable 16S rRNA processing protein RimM	information pathways	COG0806J

Rv2908c	-	8520.68	Cytoplasmic	hypothetical protein Rv2908c	Predicted RNA-binding protein (contains KH domain)	Regulatory proteins	COG1837R
Rv2909c	rpsP	17437.1	Cytoplasmic	30S ribosomal protein S16	30S ribosomal protein S16 RpsP	information pathways	COG0228J
Rv2911	dacB2	29746.7	Cytoplasmic Membrane	penicillin-binding protein DacB2	Probable penicillin-binding protein DacB2 (D-alanyl-D-alanine carboxypeptidase) (DD-peptidase) (DD-carboxypeptidase) (PBP) (DD-transpeptidase) (serine-type D-ala-D-ala carboxypeptidase) (D-amino acid hydrolase)	cell wall and cell processes	COG1686M
Rv2914c	pknI	61804.8	Cytoplasmic Membrane	serine/threonine-protein kinase PknI	Probable transmembrane serine/threonine-protein kinase I PknI (protein kinase I) (STPK I) (phosphorylase B kinase kinase) (hydroxyalkyl-protein kinase)	Regulatory proteins	COG0515RTKL
Rv2916c	ffh	54969.5	Cytoplasmic Membrane	signal recognition particle protein	Probable signal recognition particle protein Ffh (fifty-four homolog) (SRP protein)	cell wall and cell processes	COG0541U
Rv2918c	glnD	86437.8	Cytoplasmic	bifunctional uridylyltransferase/uridylyl-removing enzyme	Probable [protein-PII] uridylyltransferase GlnD (PII uridylyltransferase) (uridylyl removing enzyme) (UTASE)	intermediary metabolism and respiration	COG2844O
Rv2919c	glnB	12227	Cytoplasmic Membrane	nitrogen regulatory protein P-II	Probable nitrogen regulatory protein P-II GlnB	Regulatory proteins	COG0347E
Rv2921c	ftsY	43999.2	Cytoplasmic Membrane	signal recognition particle receptor FtsY	Probable cell division protein FtsY (SRP receptor) (signal recognition particle receptor)	cell wall and cell processes	COG0552U
Rv2922c	smc	130605	Cytoplasmic	chromosome partition protein Smc	Probable chromosome partition protein Smc	cell wall and cell processes	COG1196D
Rv2925c	rnc	25398.9	Cytoplasmic	ribonuclease III	Probable ribonuclease III Rnc (RNase III)	information pathways	COG0571K
Rv2927c	-	26953.8	Cytoplasmic	hypothetical protein Rv2927c	Cell division initiation protein	Regulatory proteins	COG3599D
Rv2928	tesA	29151.8	Cytoplasmic	thioesterase TesA	Probable thioesterase TesA	lipid metabolism	COG3208Q
Rv2930	fadD26	63044	Cytoplasmic	fatty-acid--CoA ligase FadD26	Fatty-acid-AMP ligase FadD26 (fatty-acid-AMP synthetase) (fatty-acid-AMP synthase)	lipid metabolism	COG0318IQ
Rv2931	ppsA	198835	Unknown	phthiocerol synthesis polyketide synthase type I PpsA	Phenolphthiocerol synthesis type-I polyketide synthase PpsA	lipid metabolism	COG3321Q
Rv2932	ppsB	162496	Cytoplasmic Membrane	phthiocerol synthesis polyketide synthase type I PpsB	Phenolphthiocerol synthesis type-I polyketide synthase PpsB	lipid metabolism	COG3321Q
Rv2933	ppsC	230622	Cytoplasmic Membrane	phthiocerol synthesis polyketide synthase type I PpsC	Phenolphthiocerol synthesis type-I polyketide synthase PpsC	lipid metabolism	COG3321Q
Rv2934	ppsD	193315	Cytoplasmic Membrane	phthiocerol synthesis polyketide synthase type I PpsD	Phenolphthiocerol synthesis type-I polyketide synthase PpsD	lipid metabolism	COG3321Q
Rv2935	ppsE	158713	Cytoplasmic Membrane	phthiocerol synthesis polyketide synthase type I PpsE	Phenolphthiocerol synthesis type-I polyketide synthase PpsE	lipid metabolism	COG3321Q
Rv2936	drrA	35818.2	Cytoplasmic Membrane	daunorubicin ABC transporter ATP-binding protein DrrA	Daunorubicin-dim-transport ATP-binding protein ABC transporter DrrA	cell wall and cell processes	COG1131V
Rv2939	papA5	45428.8	Cytoplasmic	phthiocerol/phthiodiolone dimycocerosyl transferase	Possible conserved polyketide synthase associated protein PapA5	lipid metabolism	-
Rv2940c	mas	224380	Cytoplasmic Membrane	multifunctional mycocerosic acid synthase	Probable multifunctional mycocerosic acid synthase membrane-associated Mas	lipid metabolism	COG3321Q
Rv2941	fadD28	62640.7	Cytoplasmic	long-chain-fatty-acid--AMP ligase FadD28	Fatty-acid-AMP ligase FadD28 (fatty-acid-AMP synthetase) (fatty-acid-AMP synthase)	lipid metabolism	COG0318IQ
Rv2942	mmpL	95121.9	Cytoplasmic	transmembrane transport protein	Conserved transmembrane transport protein MmpL7	cell wall and cell	COG2409R

	7		Membrane	MmpL7		processes	
Rv2945c	lppX	24140.2	Cytoplasmic Membrane	lipoprotein LppX	Probable conserved lipoprotein LppX	cell wall and cell processes	-
Rv2946c	pks1	166640	Cytoplasmic Membrane	polyketide synthase	Probable polyketide synthase Pks1	lipid metabolism	COG3321Q
Rv2947c	pks15	51669.8	Cytoplasmic Membrane	polyketide synthase	Probable polyketide synthase Pks15	lipid metabolism	COG3321Q
Rv2948c	fadD22	75198.2	Cytoplasmic Membrane	p-hydroxybenzoyl--AMP ligase	P-hydroxybenzoyl-AMP ligase FadD22	lipid metabolism	COG0318IQ
Rv2949c	-	22587	Cytoplasmic	chorismate pyruvate-lyase	Chorismate pyruvate lyase	intermediary metabolism and respiration	COG3161H
Rv2950c	fadD29	67446.5	Cytoplasmic	long-chain-fatty-acid--AMP ligase FadD29	Fatty-acid-AMP ligase FadD29 (fatty-acid-AMP synthetase) (fatty-acid-AMP synthase)	lipid metabolism	-
Rv2951c	-	41315.6	Cytoplasmic	phthiodiolone/phenolphthiodiolone dimycocerosates ketoreductase	Possible oxidoreductase	intermediary metabolism and respiration	COG2141C
Rv2953	-	45104.2	Unknown	trans-acting enoyl reductase	Enoyl reductase	lipid metabolism	COG3268S
Rv2954c	-	27144.1	Cytoplasmic	hypothetical protein Rv2954c	SAM-dependent methyltransferases	Intermediary metabolism and respiration	COG0030J
Rv2955c	-	35918.3	Cytoplasmic	hypothetical protein Rv2955c	SAM-dependent methyltransferases	Intermediary metabolism and respiration	COG4123R
Rv2956	-	26534.2	Cytoplasmic	hypothetical protein Rv2956	SAM-dependent methyltransferases	Intermediary metabolism and respiration	-
Rv2957	-	31046.1	Cytoplasmic Membrane	PGL/p-HBAD biosynthesis glycosyltransferase	Possible glycosyl transferase	intermediary metabolism and respiration	COG0463M
Rv2958c	-	46795.9	Cytoplasmic Membrane	PGL/p-HBAD biosynthesis glycosyltransferase	Possible glycosyl transferase	intermediary metabolism and respiration	COG1819GC
Rv2959c	-	27844.5	Cytoplasmic	rhamnosyl O-methyltransferase	Possible methyltransferase (methylase)	intermediary metabolism and respiration	COG3510V
Rv2962c	-	49401	Cytoplasmic Membrane	PGL/p-HBAD biosynthesis rhamnosyltransferase	Possible glycosyl transferase	intermediary metabolism and respiration	COG1819GC
Rv2967c	pca	120392	Cytoplasmic	pyruvate carboxylase	Probable pyruvate carboxylase Pca (pyruvic carboxylase)	intermediary metabolism and respiration	COG1038C
Rv2969c	-	26829.5	Cytoplasmic	hypothetical protein Rv2969c	Possible conserved membrane or secreted protein	cell wall and cell processes	COG1651O
Rv2970c	lipN	40134.7	Cytoplasmic	lipase/esterase LipN	Probable lipase/esterase LipN	intermediary	COG0657I

						metabolism and respiration	
Rv2971	-	30332	Cytoplasmic	oxidoreductase	Probable oxidoreductase	intermediary metabolism and respiration	COG0656R
Rv2973c	recG	80329	Cytoplasmic	ATP-dependent DNA helicase RecG	Probable ATP-dependent DNA helicase RecG	information pathways	COG1200LK
Rv2974c	-	46685	Cytoplasmic	alanine-rich protein	Predicted kinase related to dihydroxyacetone kinase	Intermediary metabolism and respiration	COG1461R
Rv2975c	-	8835.89	Unknown	hypothetical protein Rv2975c	Predicted kinase related to dihydroxyacetone kinase	Intermediary metabolism and respiration	COG1461R
Rv2976c	ung	24449.1	Cytoplasmic	uracil-DNA glycosylase	Probable uracil-DNA glycosylase Ung (UDG)	information pathways	COG0692L
Rv2977c	thiL	34298.5	Cytoplasmic	thiamine-monophosphate kinase	Probable thiamine-monophosphate kinase ThiL (thiamine-phosphate kinase)	intermediary metabolism and respiration	COG0611H
Rv2981c	ddlA	39678.2	Cytoplasmic	D-alanine--D-alanine ligase	Probable D-alanine--D-alanine ligase DdlA (D-alanylalanine synthetase) (D-ala-D-ala ligase)	cell wall and cell processes	COG1181M
Rv2982c	gpdA2	33992	Cytoplasmic Membrane	glycerol-3-phosphate dehydrogenase	Probable glycerol-3-phosphate dehydrogenase [NAD(P)+] GpdA2 (NAD(P)H- dependent glycerol-3-phosphate dehydrogenase)	lipid metabolism	COG0240C
Rv2984	ppk1	83039.3	Cytoplasmic Membrane	polyphosphate kinase	Polyphosphate kinase PPK (polyphosphoric acid kinase) (ATP-polyphosphate phosphotransferase)	intermediary metabolism and respiration	COG0855P
Rv2985	mutT1	34748.3	Cytoplasmic	8-oxo-dGTP diphosphatase	Possible hydrolase MutT1	information pathways	COG1051F
Rv2986c	hupB	22187.3	Cytoplasmic	DNA-binding protein HU	DNA-binding protein HU homolog HupB (histone-like protein) (HLP) (21-kDa laminin-2-binding protein)	information pathways	COG0776L
Rv2987c	leuD	21779.8	Cytoplasmic	3-isopropylmalate dehydratase small subunit	Probable 3-isopropylmalate dehydratase (small subunit) LeuD (isopropylmalate isomerase) (alpha-IPM isomerase) (IPMI)	intermediary metabolism and respiration	COG0066E
Rv2988c	leuC	50185	Cytoplasmic	3-isopropylmalate dehydratase large subunit	Probable 3-isopropylmalate dehydratase (large subunit) LeuC (isopropylmalate isomerase) (alpha-IPM isomerase) (IPMI)	intermediary metabolism and respiration	COG0065E
Rv2989	-	24283.9	Cytoplasmic Membrane	transcriptional regulator	Probable transcriptional regulatory protein	Regulatory proteins	COG1414K
Rv2991	-	18203.7	Cytoplasmic	hypothetical protein Rv2991	Pyridoxamine 5-phosphate oxidase	Intermediary metabolism and respiration	COG3871R
Rv2992c	gltS	53832	Cytoplasmic	glutamate--tRNA ligase	Glutamyl-tRNA synthetase GltS (glutamate--tRNA ligase) (glutamyl-tRNA synthase) (GLURS)	information pathways	COG0008J
Rv2993c	-	25341.1	Unknown	2-hydroxyhepta-2,4-diene-1,7-dioate isomerase	Possible 2-hydroxyhepta-2,4-diene-1,7-dioate isomerase (HHDD isomerase)	intermediary metabolism and respiration	-
Rv2995c	leuB	35306.1	Cytoplasmic	3-isopropylmalate dehydrogenase	Probable 3-isopropylmalate dehydrogenase LeuB (beta-IPM	intermediary	COG0473CE

					dehydrogenase) (IMDH) (3-IPM-DH)	metabolism and respiration	
Rv2996c	serA1	54522.2	Cytoplasmic	D-3-phosphoglycerate dehydrogenase	Probable D-3-phosphoglycerate dehydrogenase SerA1 (PGDH)	intermediary metabolism and respiration	COG0111HE
Rv2999	lppY	33949.6	Unknown	lipoprotein LppY	Probable conserved lipoprotein LppY	cell wall and cell processes	-
Rv3001c	ilvC	36090.1	Cytoplasmic	ketol-acid reductoisomerase	Probable KETOL-acid reductoisomerase IlvC (acetohydroxy-acid isomeroreductase) (alpha-keto-beta-hydroxylacil reductoisomerase)	intermediary metabolism and respiration	-
Rv3002c	ilvN	18186.9	Cytoplasmic	acetolactate synthase small subunit	Probable acetolactate synthase (small subunit) IlvN (acetohydroxy-acid synthase) (AHAS) (ALS)	intermediary metabolism and respiration	COG0440E
Rv3003c	ilvB1	66090.7	Cytoplasmic	acetolactate synthase large subunit IlvB	Acetolactate synthase (large subunit) IlvB1 (acetohydroxy-acid synthase)	intermediary metabolism and respiration	COG0028EH
Rv3006	lppZ	38752.1	Unknown	lipoprotein LppZ	Probable conserved lipoprotein LppZ	cell wall and cell processes	COG2133G
Rv3009c	gatB	54632.9	Cytoplasmic	aspartyl/glutamyl-tRNA(Asn/Gln) amidotransferase subunit B	Probable glutamyl-tRNA(GLN) amidotransferase (subunit B) GatB (Glu-ADT subunit B)	information pathways	COG0064J
Rv3010c	pfkA	36880	Cytoplasmic	6-phosphofructokinase	Probable 6-phosphofructokinase PfkA (phosphohexokinase) (phosphofructokinase)	intermediary metabolism and respiration	COG0205G
Rv3011c	gatA	51388.2	Cytoplasmic	glutamyl-tRNA(GLN) amidotransferase subunit A	Probable glutamyl-tRNA(GLN) amidotransferase (subunit A) GatA (Glu-ADT subunit A)	information pathways	-
Rv3013	-	22967.2	Cytoplasmic	hypothetical protein Rv3013	Amino acid-binding ACT protein	Regulatory proteins	COG0317TK
Rv3014c	ligA	75257.1	Cytoplasmic	DNA ligase A	DNA ligase [NAD dependent] LigA (polydeoxyribonucleotide synthase [NAD+])	information pathways	COG0272L
Rv3015c	-	34211.9	Unknown	hypothetical protein Rv3015c	B12 independent protein	Regulatory proteins	COG0620E
Rv3024c	trmU	38120.1	Cytoplasmic	tRNA-specific 2-thiouridylase	Probable tRNA (5-methylaminomethyl-2-thiouridylate)-methyltransferase TrmU	information pathways	COG0482J
Rv3027c	-	30513.7	Cytoplasmic	GCN5-like N-acetyltransferase	GCN5-related N-acetyltransferase	intermediary metabolism and respiration	-
Rv3028c		31691	Cytoplasmic	electron transfer flavoprotein subunit alpha	Probable electron transfer flavoprotein (alpha-subunit) FixB (alpha-ETF) (electron transfer flavoprotein large subunit) (ETFLS)	intermediary metabolism and respiration	COG2025C
Rv3029c	fixA	28080.8	Cytoplasmic	electron transfer flavoprotein subunit beta	Probable electron transfer flavoprotein (beta-subunit) FixA (beta-ETF) (electron transfer flavoprotein small subunit) (ETFSS)	intermediary metabolism and respiration	COG2086C
Rv3031	-	57797.3	Cytoplasmic	1,4-alpha-glucan branching protein	linked to sugar transferase activity	Intermediary metabolism and respiration	COG1543S
Rv3034c	-	33032.3	Unknown	acetyltransferase	Possible transferase	intermediary	COG0110R

						metabolism and respiration	
Rv3035	-	37305	Cytoplasmic Membrane	hypothetical protein Rv3035	FOG- WD40-like repeat	Regulatory proteins	COG1520S
Rv3036c	TB22.2	24406.7	Extracellular	secreted protein	Probable conserved secreted protein TB22.2	cell wall and cell processes	-
Rv3038c	-	36049.7	Cytoplasmic	hypothetical protein Rv3038c	SAM-dependent methyltransferases	Intermediary metabolism and respiration	COG2226H
Rv3039c	echA17	26505.9	Unknown	enoyl-CoA hydratase EchA17	Probable enoyl-CoA hydratase EchA17 (crotonase) (unsaturated acyl-CoA hydratase) (enoyl hydratase)	lipid metabolism	COG1024I
Rv3040c	-	31485.2	Cytoplasmic	hypothetical protein Rv3040c	NTP pyrophosphohydrolases including oxidative damage repair enzymes	Information pathways	-
Rv3041c	-	31249.4	Cytoplasmic Membrane	ABC transporter ATP-binding protein	Probable conserved ATP-binding protein ABC transporter	cell wall and cell processes	COG1119P
Rv3042c	serB2	43058.7	Cytoplasmic	phosphoserine phosphatase SerB	Probable phosphoserine phosphatase SerB2 (PSP) (O-phosphoserine phosphohydrolase) (pspase)	intermediary metabolism and respiration	COG0560E
Rv3043c	ctaD	63654.7	Cytoplasmic Membrane	cytochrome C oxidase cytochrome 1	Probable cytochrome C oxidase polypeptide I CtaD (cytochrome AA3 subunit 1)	intermediary metabolism and respiration	COG0843C
Rv3044	fecB	36910.4	Unknown	FeIII-dicitrate-binding periplasmic lipoprotein	Probable FEIII-dicitrate-binding periplasmic lipoprotein FecB	cell wall and cell processes	COG0614P
Rv3045	adhC	37075.3	Cytoplasmic	NADP-dependent alcohol dehydrogenase	Probable NADP-dependent alcohol dehydrogenase AdhC	intermediary metabolism and respiration	COG1064R
Rv3046c	-	13350.8	Cytoplasmic Membrane	hypothetical protein Rv3046c	Conserved protein	conserved hypotheticals	-
Rv3048c	nrdF2	36959.5	Unknown	ribonucleoside-diphosphate reductase subunit beta NrdF2	Ribonucleoside-diphosphate reductase (beta chain) NrdF2 (ribonucleotide reductase small subunit) (R2F protein)	information pathways	COG0208F
Rv3049c	-	58744.6	Cytoplasmic Membrane	monooxygenase	Probable monooxygenase	intermediary metabolism and respiration	COG2072P
Rv3050c	-	27586.6	Cytoplasmic	AsnC family transcriptional regulator	Probable transcriptional regulatory protein (probably AsnC-family)	Regulatory proteins	COG1309K
Rv3051c	nrdE	79290.8	Cytoplasmic	ribonucleoside-diphosphate reductase subunit alpha	Ribonucleoside-diphosphate reductase (alpha chain) NrdE (ribonucleotide reductase small subunit) (R1F protein)	information pathways	COG0209F
Rv3057c	-	31153.8	Cytoplasmic	short chain alcohol dehydrogenase/reductase	Probable short chain alcohol dehydrogenase/reductase	intermediary metabolism and respiration	COG4221R
Rv3058c	-	23803	Cytoplasmic	TetR family transcriptional regulator	Possible transcriptional regulatory protein (probably TetR-family)	Regulatory proteins	COG1309K
Rv3059	cyp136	56227.7	Cytoplasmic	cytochrome P450 Cyp136	Probable cytochrome P450 136 Cyp136	intermediary metabolism and respiration	COG2124Q

Rv3061c	fadE2 2	76741.1	Cytoplasmic	acyl-CoA dehydrogenase FadE22	Probable acyl-CoA dehydrogenase FadE22	lipid metabolism	COG1960I
Rv3062	ligB	53704.6	Unknown	DNA ligase	Probable ATP-dependent DNA ligase LigB (polydeoxyribonucleotide synthase [ATP]) (polynucleotide ligase [ATP]) (sealase) (DNA repair protein) (DNA joinase)	information pathways	COG1793L
Rv3068c	pgmA	58265.7	Cytoplasmic	phosphoglucomutase PgmA	Probable phosphoglucomutase PgmA (glucose phosphomutase) (PGM)	intermediary metabolism and respiration	COG0033G
Rv3071	-	40466.2	Cytoplasmic	hypothetical protein Rv3071	membrane associated process	cell wall and cell processes	COG1993S
Rv3075c	-	33100.2	Cytoplasmic	hypothetical protein Rv3075c	Citrate lyase beta subunit	Intermediary metabolism and respiration	COG2301G
Rv3076	-	17121.6	Cytoplasmic	hypothetical protein Rv3076	Cyclase/Dehydrase	Intermediary metabolism and respiration	-
Rv3077	-	66740.9	Cytoplasmic	hydrolase	Possible hydrolase	intermediary metabolism and respiration	COG3119P
Rv3079c	-	30064.4	Cytoplasmic	hypothetical protein Rv3079c	Coenzyme F420-dependent N5,N10-methylene tetrahydromethanopterin reductase and related flavin-dependent oxidoreductases	Intermediary metabolism and respiration	COG2141C
Rv3080c	pknK	119418	Cytoplasmic Membrane	serine/threonine-protein kinase PknK	Serine/threonine-protein kinase transcriptional regulatory protein PknK (protein kinase K) (STPK K)	Regulatory proteins	COG2909K
Rv3081	-	45788.1	Unknown	hypothetical protein Rv3081	Conserved hypothetical protein	conserved hypotheticals	COG4914R
Rv3087	-	52597.1	Cytoplasmic	diacylglycerol O-acyltransferase	Possible triacylglycerol synthase (diacylglycerol acyltransferase)	lipid metabolism	COG4908R
Rv3088	tgs4	50886.5	Unknown	diacylglycerol O-acyltransferase	Putative triacylglycerol synthase (diacylglycerol acyltransferase) Tgs4	lipid metabolism	-
Rv3089	fadD1 3	54459.5	Cytoplasmic	long chain-fatty-acid--CoA ligase FadD13	Probable chain -fatty-acid-CoA ligase FadD13 (fatty-acyl-CoA synthetase)	lipid metabolism	COG0318IQ
Rv3090	-	32270.9	Unknown	alanine/valine-rich protein	Membrane protease subunits, stomatin/prohibitin homologs	cell wall and cell processes	COG0330O
Rv3091	-	61645.5	Cytoplasmic	hypothetical protein Rv3091	Conserved protein	conserved hypotheticals	COG1752R
Rv3094c	-	40296.8	Cytoplasmic	hypothetical protein Rv3094c	Acyl-CoA dehydrogenases	Intermediary metabolism and respiration	COG1960I
Rv3096	-	43129.8	Cytoplasmic Membrane	hypothetical protein Rv3096	Conserved hypothetical protein	conserved hypotheticals	COG3934G
Rv3099c	-	30475.7	Cytoplasmic	hypothetical protein Rv3099c	Conserved protein	conserved hypotheticals	-
Rv3101c	ftsX	32787.4	Cytoplasmic Membrane	cell division protein FtsX	Putative cell division protein FtsX (septation component-transport integral membrane protein ABC transporter)	cell wall and cell processes	COG2177D
Rv3102c	ftsE	25596.5	Cytoplasmic	cell division ATP-binding protein	Putative cell division ATP-binding protein FtsE (septation	cell wall and cell	COG2884D

			Membrane	FtsE	component-transport ATP-binding protein ABC transporter)	processes	
Rv3104c	-	33380.6	Cytoplasmic Membrane	transmembrane protein	Possible conserved transmembrane protein	cell wall and cell processes	COG0668M
Rv3105c	prfB	42125	Cytoplasmic	peptide chain release factor PrfB	Probable peptide chain release factor 2 PrfB (RF-2)	information pathways	COG1186J
Rv3106	fprA	49341.1	Cytoplasmic	NADPH-ferredoxin reductase FprA	NADPH:adrenodoxin oxidoreductase FprA (NADPH-ferredoxin reductase)	intermediary metabolism and respiration	COG0493ER
Rv3107c	agpS	56520.5	Cytoplasmic	alkyldihydroxyacetonephosphate synthase	Possible alkyldihydroxyacetonephosphate synthase AgpS (alkyl-DHAP synthase) (alkylglycerone-phosphate synthase)	lipid metabolism	COG0277C
Rv3117	cysA3	31014.6	Cytoplasmic	thiosulfate sulfurtransferase	Probable thiosulfate sulfurtransferase CysA3 (rhodanese-like protein) (thiosulfate cyanide transsulfurase) (thiosulfate thiotransferase)	intermediary metabolism and respiration	COG2897P
Rv3118	-	10165.3	Unknown	hypothetical protein Rv3118	Conserved hypothetical protein SseC1	intermediary metabolism and respiration	-
Rv3119	moaE1	15920.2	Cytoplasmic	molybdopterin synthase catalytic subunit 1	Probable molybdenum cofactor biosynthesis protein E MoaE1 (molybdopterin converting factor large subunit) (molybdopterin [MPT] converting factor, subunit 2)	intermediary metabolism and respiration	COG0314H
Rv3127	-	38488.6	Cytoplasmic	hypothetical protein Rv3127	Conserved protein	conserved hypotheticals	-
Rv3130c	tgs1	50720.6	Cytoplasmic	diacylglycerol O-acyltransferase	Triacylglycerol synthase (diacylglycerol acyltransferase) Tgs1	lipid metabolism	-
Rv3131	-	35978.4	Cytoplasmic Membrane	NAD(P)H nitroreductase	diacylglycerol O-acyltransferase activity	Intermediary metabolism and respiration	-
Rv3132c	devS	62240.6	Cytoplasmic Membrane	two component sensor histidine kinase DevS	Two component sensor histidine kinase DevS	Regulatory proteins	COG4585T
Rv3133c	devR	23261.9	Cytoplasmic	two component transcriptional regulator DevR	Two component transcriptional regulatory protein DevR (probably LuxR/UhpA-family)	Regulatory proteins	COG2197TK
Rv3134c	-	28007.9	Cytoplasmic	universal stress protein	Universal stress protein family protein	virulence, detoxification, adaptation	COG0589T
Rv3136	PPE51	37979.9	Cytoplasmic Membrane	PPE family protein PPE51	PPE family protein PPE51	PE/PPE	COG5651N
Rv3137	-	27661.1	Unknown	histidinol-phosphatase	Probable monophosphatase	intermediary metabolism and respiration	COG0483G
Rv3139	fadE24	49646.4	Cytoplasmic	acyl-CoA dehydrogenase FadE24	Probable acyl-CoA dehydrogenase FadE24	lipid metabolism	COG1960I
Rv3140	fadE23	43344.2	Cytoplasmic	acyl-CoA dehydrogenase FadE23	Probable acyl-CoA dehydrogenase FadE23	lipid metabolism	COG1960I
Rv3141	fadB4	33852.8	Cytoplasmic	NADPH quinone oxidoreductase FadB	Probable NADPH quinone oxidoreductase FadB4 (NADPH:quinone reductase) (zeta-crystallin)	lipid metabolism	COG0604CR
Rv3143	-	14353.6	Unknown	response regulator	Probable response regulator	Regulatory proteins	COG0745TK
Rv3146	nuoB	20186.8	Cytoplasmic	NADH-quinone oxidoreductase	Probable NADH dehydrogenase I (chain B) NuoB (NADH-	intermediary	COG0377C

			Membrane	subunit B	ubiquinone oxidoreductase chain B)	metabolism and respiration	
Rv3147	nuoC	26932.1	Cytoplasmic	NADH-quinone oxidoreductase subunit C	Probable NADH dehydrogenase I (chain C) NuoC (NADH-ubiquinone oxidoreductase chain C)	intermediary metabolism and respiration	COG0852C
Rv3148	nuoD	48164.5	Cytoplasmic	NADH-quinone oxidoreductase subunit D	Probable NADH dehydrogenase I (chain D) NuoD (NADH-ubiquinone oxidoreductase chain D)	intermediary metabolism and respiration	COG0649C
Rv3149	nuoE	27165.6	Cytoplasmic	NADH-quinone oxidoreductase subunit E	Probable NADH dehydrogenase I (chain E) NuoE (NADH-ubiquinone oxidoreductase chain E)	intermediary metabolism and respiration	COG1905C
Rv3150	nuoF	48101.6	Cytoplasmic	NADH-quinone oxidoreductase subunit F	Probable NADH dehydrogenase I (chain F) NuoF (NADH-ubiquinone oxidoreductase chain F)	intermediary metabolism and respiration	COG1894C
Rv3151	nuoG	85392.1	Cytoplasmic	NADH-quinone oxidoreductase subunit G	Probable NADH dehydrogenase I (chain G) NuoG (NADH-ubiquinone oxidoreductase chain G)	intermediary metabolism and respiration	COG1034C
Rv3153	nuoI	23434.4	Unknown	NADH-quinone oxidoreductase subunit I	Probable NADH dehydrogenase I (chain I) NuoI (NADH-ubiquinone oxidoreductase chain I)	intermediary metabolism and respiration	COG1143C
Rv3154	nuoJ	27589.3	Cytoplasmic Membrane	NADH-quinone oxidoreductase subunit J	Probable NADH dehydrogenase I (chain J) NuoJ (NADH-ubiquinone oxidoreductase chain J)	intermediary metabolism and respiration	COG0839C
Rv3157	nuoM	59175.1	Cytoplasmic Membrane	NADH-quinone oxidoreductase subunit M	Probable NADH dehydrogenase I (chain M) NUOK (NADH-ubiquinone oxidoreductase chain M)	intermediary metabolism and respiration	COG1008C
Rv3160c	-	23120.7	Cytoplasmic	TetR family transcriptional regulator	Possible transcriptional regulatory protein (probably TetR-family)	Regulatory proteins	COG1309K
Rv3161c	-	42504.2	Unknown	dioxygenase	Possible dioxygenase	intermediary metabolism and respiration	COG4638PR
Rv3164c	moxR ₃	35300.8	Cytoplasmic	methanol dehydrogenase transcriptional regulator MoxR	Probable methanol dehydrogenase transcriptional regulatory protein MoxR3	Regulatory proteins	COG0714R
Rv3168	-	42529.2	Cytoplasmic	aminoglycoside phosphotransferase	Putative aminoglycoside phosphotransferase	virulence, detoxification, adaptation	COG3173R
Rv3169	-	41759.8	Cytoplasmic	hypothetical protein Rv3169	Conserved protein	conserved hypotheticals	-
Rv3170	aofH	48407.6	Cytoplasmic	flavin-containing monoamine oxidase	Probable flavin-containing monoamine oxidase AofH (amine oxidase) (MAO)	intermediary metabolism and respiration	COG1231E
Rv3173c	-	21566.6	Unknown	TetR/Acr family transcriptional regulator	Probable transcriptional regulatory protein (probably TetR/AcrR-family)	Regulatory proteins	COG1309K
Rv3179	-	48021	Cytoplasmic	hypothetical protein Rv3179	Predicted ATPase (AAA+ superfamily)	Intermediary metabolism and	COG1373R

						respiration	
Rv3193c	-	107470	Cytoplasmic Membrane	transmembrane protein	Probable conserved transmembrane protein	cell wall and cell processes	COG1615S
Rv3194c	-	35440.6	Unknown	secreted protein	Possible conserved secreted protein	cell wall and cell processes	COG3480T
Rv3197	-	49121.9	Cytoplasmic Membrane	ABC transporter ATP-binding protein	Probable conserved ATP-binding protein ABC transporter	cell wall and cell processes	COG0661R
Rv3198A	-	8963.21	Unknown	glutaredoxin protein	Possible glutaredoxin protein	intermediary metabolism and respiration	COG0695O
Rv3198c	uvrD2	75603.7	Cytoplasmic	ATP-dependent DNA helicase UvrD	Probable ATP-dependent DNA helicase II UvrD2	information pathways	COG0210L
Rv3200c	-	38099.9	Cytoplasmic Membrane	transmembrane cation transporter	Possible transmembrane cation transporter	cell wall and cell processes	COG0569P
Rv3203	lipV	23624.1	Cytoplasmic	lipase LipV	Possible lipase LipV	intermediary metabolism and respiration	COG0596R
Rv3205c	-	31351.9	Cytoplasmic	hypothetical protein Rv3205c	membrane associated process	cell wall and cell processes	-
Rv3206c	moeB1	42141.4	Cytoplasmic	adenylyltransferase/sulfurtransferase	Probable molybdenum cofactor biosynthesis protein MoeB1 (MPT-synthase sulfurylase) (molybdopterin synthase sulphurylase)	intermediary metabolism and respiration	COG0476H
Rv3208	-	25166.3	Cytoplasmic	TetR family transcriptional regulator	Probable transcriptional regulatory protein (probably TetR-family)	Regulatory proteins	COG1309K
Rv3208A	TB9.4	9400.53	Cytoplasmic	hypothetical protein Rv3208A	Conserved protein TB9.4	conserved hypotheticals	-
Rv3210c	-	25121.4	Cytoplasmic	hypothetical protein Rv3210c	Conserved protein	conserved hypotheticals	COG4445FJ
Rv3211	rhIE	56703.1	Cytoplasmic	ATP-dependent RNA helicase RhIE	Probable ATP-dependent RNA helicase RhIE	information pathways	COG0513LKJ
Rv3212	-	42506.1	Unknown	alanine/valine-rich protein	intermediary metabolism and respiration , Valine rich protein	Intermediary metabolism and respiration	-
Rv3213c	-	28575.3	Cytoplasmic Membrane	SOJ/ParA-like protein	Possible SOJ/para-related protein	cell wall and cell processes	COG1192D
Rv3214	gpm2	21948.8	Cytoplasmic	phosphoglycerate mutase	Possible phosphoglycerate mutase Gpm2 (phosphoglyceromutase) (PGAM) (BPG-dependent PGAM)	intermediary metabolism and respiration	COG0406G
Rv3215	entC	38905.7	Cytoplasmic	isochorismate synthase	Probable isochorismate synthase EntC (isochorismate hydroxymutase) (enterochelin biosynthesis)	intermediary metabolism and respiration	COG1169HQ
Rv3218	-	34474.5	Unknown	hypothetical protein Rv3218	Sphingosine kinase and enzymes related to eukaryotic diacylglycerol kinase	cell wall and cell processes	COG1597IR
Rv3220c	-	54012.5	Cytoplasmic	two component sensor kinase	Probable two component sensor kinase	Regulatory proteins	COG3920T

Rv3223c	sigH	24225.1	Cytoplasmic	ECF RNA polymerase sigma factor SigH	Alternative RNA polymerase sigma-E factor (sigma-24) SigH (RPOE)	information pathways	COG1595K
Rv3224	-	29814.4	Cytoplasmic	iron-regulated short-chain dehydrogenase/reductase	Possible iron-regulated short-chain dehydrogenase/reductase	intermediary metabolism and respiration	COG1028IQR
Rv3227	aroA	46393.8	Unknown	3-phosphoshikimate 1-carboxyvinyltransferase	3-phosphoshikimate 1-carboxyvinyltransferase AroA (5-enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS)	intermediary metabolism and respiration	COG0128E
Rv3228	-	34857.8	Cytoplasmic	hypothetical protein Rv3228	Predicted GTPases	Intermediary metabolism and respiration	COG1162R
Rv3229c	desA3	48443	Cytoplasmic	stearoyl-CoA 9-desaturase	Possible linoleoyl-CoA desaturase (delta(6)-desaturase)	lipid metabolism	COG3239I
Rv3230c	-	40794.7	Cytoplasmic	stearoyl-CoA 9-desaturase electron transfer partner	Hypothetical oxidoreductase	intermediary metabolism and respiration	COG1018C
Rv3232c	ppk2	34092	Cytoplasmic	polyphosphate kinase	Polyphosphate kinase Ppk2 (polyphosphoric acid kinase)	intermediary metabolism and respiration	COG2326S
Rv3234c	tgs3	30346.8	Cytoplasmic	diacylglycerol O-acyltransferase	Putative triacylglycerol synthase (diacylglycerol acyltransferase) Tgs3	lipid metabolism	-
Rv3237c	-	17233.6	Cytoplasmic	hypothetical protein Rv3237c	Putative regulatory, ligand-binding protein related to C-terminal domains of K ⁺ channels	cell wall and cell processes	COG0490P
Rv3240c	secA1	105990	Cytoplasmic	protein translocase subunit SecA	Probable preprotein translocase SecA1 1 subunit	cell wall and cell processes	COG0653U
Rv3241c	-	24534.7	Cytoplasmic	hypothetical protein Rv3241c	Conserved protein	information pathways	-
Rv3243c	-	30689.3	Cytoplasmic Membrane	hypothetical protein Rv3243c	Unknown protein	conserved hypotheticals	-
Rv3244c	lpqB	61151.1	Unknown	lipoprotein LpqB	Probable conserved lipoprotein LpqB	cell wall and cell processes	-
Rv3245c	mtrB	61636.6	Cytoplasmic Membrane	two component sensory histidine kinase MtrB	Two component sensory transduction histidine kinase MtrB	Regulatory proteins	COG0642T
Rv3246c	mtrA	25279.3	Cytoplasmic	two component DNA-binding response regulator MtrA	Two component sensory transduction transcriptional regulatory protein MtrA	Regulatory proteins	COG0745TK
Rv3248c	sahH	54323.6	Cytoplasmic	adenosylhomocysteinase	Probable adenosylhomocysteinase SahH (S-adenosyl-L-homocysteine hydrolase) (adohcayase)	intermediary metabolism and respiration	COG0499H
Rv3254	-	51518.1	Cytoplasmic Membrane	hypothetical protein Rv3254	Monoxygenase	Intermediary metabolism and respiration	COG0654HC
Rv3255c	manA	43339.2	Cytoplasmic	mannose-6-phosphate isomerase	Probable mannose-6-phosphate isomerase ManA (phosphomannose isomerase) (phosphomannoisomerase) (PMI) (phosphohexoisomerase) (phosphohexomutase)	intermediary metabolism and respiration	COG1482G
Rv3256c	-	35277.4	Unknown	hypothetical protein Rv3256c	Sugar binding protein	Intermediary metabolism and respiration	-

Rv3257c	pmmA	49041	Cytoplasmic	phosphomannomutase PmmA	Probable phosphomannomutase PmmA (PMM) (phosphomannose mutase)	intermediary metabolism and respiration	COG1109G
Rv3259	-	15649	Cytoplasmic	hypothetical protein Rv3259	Conserved hypothetical protein	conserved hypotheticals	-
Rv3261	fbiA	35303.1	Cytoplasmic	2-phospho-L-lactate transferase	Probable F420 biosynthesis protein FbiA	intermediary metabolism and respiration	COG0391S
Rv3262	fbiB	47560.3	Cytoplasmic	coenzyme F420:L-glutamate ligase	Probable F420 biosynthesis protein FbiB	intermediary metabolism and respiration	COG1478S
Rv3264c	manB	37839	Cytoplasmic	D-alpha-D-mannose-1-phosphate guanylyltransferase ManB	D-alpha-D-mannose-1-phosphate guanylyltransferase ManB (D-alpha-D-heptose-1-phosphate guanylyltransferase)	cell wall and cell processes	COG1208MJ
Rv3265c	wbbL1	33262.8	Unknown	N-acetylglucosaminyl-diphosphodecaprenol L-rhamnosyltransferase	dTDP-RHA:a-D-GlcNAc-diphosphoryl polyprenol, a-3-L-rhamnosyl transferase WbbL1 (alpha-L-rhamnose-(1->3)-alpha-D-GlcNAc(1->P)-P-decaprenyl)	cell wall and cell processes	-
Rv3269	-	9750.17	Cytoplasmic	hypothetical protein Rv3269	Conserved protein	virulence, detoxification, adaptation	-
Rv3270	ctpC	76495.1	Cytoplasmic Membrane	manganese/zinc-exporting P-type ATPase	Probable metal cation-transporting P-type ATPase C CtpC	cell wall and cell processes	COG2217P
Rv3273	-	80533.6	Cytoplasmic Membrane	transmembrane carbonic anhydrase	Probable transmembrane carbonic anhydrase (carbonate dehydratase) (carbonic dehydratase)	intermediary metabolism and respiration	COG0659P
Rv3274c	fadE25	41723.4	Cytoplasmic	acyl-CoA dehydrogenase FadE25	Probable acyl-CoA dehydrogenase FadE25	lipid metabolism	COG1960I
Rv3276c	purK	45695.1	Cytoplasmic Membrane	N5-carboxyaminoimidazole ribonucleotide synthase	Probable phosphoribosylaminoimidazole carboxylase ATPase subunit PurK (air carboxylase) (AIRC)	intermediary metabolism and respiration	COG0026F
Rv3278c	-	19820	Cytoplasmic Membrane	transmembrane protein	Probable conserved transmembrane protein	cell wall and cell processes	COG3428S
Rv3279c	birA	28089	Cytoplasmic	bifunctional biotin operon repressor/biotin--[acetyl-CoA-carboxylase] synthetase	Possible bifunctional protein BirA: biotin operon repressor + biotin--[acetyl-CoA-carboxylase] synthetase (biotin--protein ligase)	intermediary metabolism and respiration	COG0340H
Rv3280	accD5	59354.2	Cytoplasmic	propionyl-CoA carboxylase subunit beta	Probable propionyl-CoA carboxylase beta chain 5 AccD5 (pccase) (propanoyl-CoA:carbon dioxide ligase)	lipid metabolism	COG4799I
Rv3283	sseA	33288.4	Cytoplasmic	thiosulfate sulfurtransferase SseA	Probable thiosulfate sulfurtransferase SseA (rhodanese) (thiosulfate cyanide transsulfurase) (thiosulfate thiotransferase)	intermediary metabolism and respiration	COG2897P
Rv3285	accA3	63751	Cytoplasmic	bifunctional protein acetyl-/propionyl-CoA carboxylase subunit alpha AccA	Probable bifunctional protein acetyl-/propionyl-coenzyme A carboxylase (alpha chain) AccA3: biotin carboxylase + biotin carboxyl carrier protein (BCCP)	lipid metabolism	COG4770I
Rv3287c	rsbW	15550.6	Unknown	anti-sigma factor RsbW	Anti-sigma factor RsbW (sigma negative effector)	information pathways	-
Rv3290c	lat	49012.1	Cytoplasmic	L-lysine-epsilon aminotransferase	Probable L-lysine-epsilon aminotransferase Lat (L-lysine aminotransferase) (lysine 6-aminotransferase)	intermediary metabolism and	COG0160E

						respiration	
Rv3291c	lrpA	16510.6	Unknown	transcriptional regulator LrpA	Probable transcriptional regulatory protein LrpA (Lrp/AsnC-family)	Regulatory proteins	COG1522K
Rv3292	-	45332	Cytoplasmic	hypothetical protein Rv3292	Conserved hypothetical protein	conserved hypotheticals	COG5383S
Rv3293	pcd	51312.2	Cytoplasmic	piperideine-6-carboxylic acid dehydrogenase	Probable piperideine-6-carboxylic acid dehydrogenase Pcd (piperideine-6-carboxylate dehydrogenase)	intermediary metabolism and respiration	COG1012C
Rv3295	-	24805.3	Cytoplasmic Membrane	TetR family transcriptional regulator	Probable transcriptional regulatory protein (probably TetR-family)	Regulatory proteins	COG1309K
Rv3296	lhr	161348	Cytoplasmic Membrane	ATP-dependent helicase	Probable ATP-dependent helicase Lhr (large helicase-related protein)	information pathways	COG1201R
Rv3299c	atsB	105648	Cytoplasmic	arylsulfatase AtsB	Probable arylsulfatase AtsB (aryl-sulfate sulphohydrolase) (sulfatase)	intermediary metabolism and respiration	COG3119P
Rv3301c	phoY1	24826.5	Cytoplasmic	phosphate transport system transcriptional regulator PhoY	Probable phosphate-transport system transcriptional regulatory protein PhoU homolog 1 PhoY1	cell wall and cell processes	COG0704P
Rv3302c	glpD2	62745.8	Cytoplasmic	glycerol-3-phosphate dehydrogenase	Probable glycerol-3-phosphate dehydrogenase GlpD2	intermediary metabolism and respiration	COG0578C
Rv3303c	lpdA	51456.9	Cytoplasmic	NAD(P)H quinone reductase LpdA	NAD(P)H quinone reductase LpdA	intermediary metabolism and respiration	COG1249C
Rv3304	-	17456.6	Cytoplasmic	hypothetical protein Rv3304	Conserved protein	conserved hypotheticals	-
Rv3305c	amiA1	41227.1	Cytoplasmic	N-acyl-L-amino acid amidohydrolase AmiA	Possible N-acyl-L-amino acid amidohydrolase AmiA1 (N-acyl-L-amino acid aminohydrolase)	intermediary metabolism and respiration	COG1473R
Rv3306c	amiB1	40740.3	Cytoplasmic	amidase AmiB	Probable amidohydrolase AmiB1 (aminohydrolase)	intermediary metabolism and respiration	COG1473R
Rv3308	pmmB	56163.8	Cytoplasmic	phosphomannomutase PmmB	Probable phosphomannomutase PmmB (phosphomannose mutase)	intermediary metabolism and respiration	COG1109G
Rv3311	-	45732.3	Cytoplasmic	hypothetical protein Rv3311	Conserved protein	conserved hypotheticals	-
Rv3313c	add	39743.2	Cytoplasmic	adenosine deaminase	Probable adenosine deaminase Add (adenosine aminohydrolase)	intermediary metabolism and respiration	COG1816F
Rv3314c	deoA	44454	Cytoplasmic	thymidine phosphorylase	Probable thymidine phosphorylase DeoA (tdrpase) (pyrimidine phosphorylase)	intermediary metabolism and respiration	COG0213F
Rv3315c	cdd	14072.1	Unknown	cytidine deaminase	Probable cytidine deaminase Cdd (cytidine aminohydrolase) (cytidine nucleoside deaminase)	intermediary metabolism and respiration	COG0295F

Rv3318	sdhA	64807.5	Cytoplasmic Membrane	succinate dehydrogenase flavoprotein subunit	Probable succinate dehydrogenase (flavoprotein subunit) SdhA (succinic dehydrogenase) (fumarate reductase) (fumarate dehydrogenase) (fumaric hydrogenase)	intermediary metabolism and respiration	COG1053C
Rv3319	sdhB	29334.9	Cytoplasmic Membrane	succinate dehydrogenase iron-sulphur protein subunit	Probable succinate dehydrogenase (iron-sulphur protein subunit) SdhB (succinic dehydrogenase) (fumarate reductase) (fumarate dehydrogenase) (fumaric hydrogenase)	intermediary metabolism and respiration	COG0479C
Rv3322c	-	21614.6	Cytoplasmic	methyltransferase	Possible methyltransferase	intermediary metabolism and respiration	COG2890J
Rv3323c	moaX	24050.4	Cytoplasmic	MoaD-MoaE fusion protein MoaX	Probable MoaD-MoaE fusion protein MoaX	intermediary metabolism and respiration	COG0314H
Rv3324c	moaC ₃	18523.2	Cytoplasmic	cyclic pyranopterin monophosphate synthase accessory protein	Probable molybdenum cofactor biosynthesis protein C 3 MoaC3	intermediary metabolism and respiration	-
Rv3327	-	62961	Cytoplasmic	transposase fusion protein	Probable transposase fusion protein	Insertion seqs and phages	COG2128S
Rv3328c	sigJ	33530.2	Cytoplasmic	ECF RNA polymerase sigma factor SigJ	Probable alternative RNA polymerase sigma factor (fragment) SigJ	information pathways	COG1595K
Rv3329	-	47699.2	Cytoplasmic	aminotransferase	Probable aminotransferase	intermediary metabolism and respiration	-
Rv3332	nagA	38773	Cytoplasmic	N-acetylglucosamine-6-phosphate deacetylase NagA	Probable N-acetylglucosamine-6-phosphate deacetylase NagA (GlcNAc 6-P deacetylase)	cell wall and cell processes	COG1820G
Rv3336c	trpS	36272	Cytoplasmic	tryptophan--tRNA ligase	Probable tryptophanyl-tRNA synthetase TrpS (tryptophan--tRNA ligase) (TRPRS) (tryptophan translase)	information pathways	COG0180J
Rv3339c	icd1	45513.8	Unknown	isocitrate dehydrogenase	Probable isocitrate dehydrogenase [NADP] Icd1 (oxalosuccinate decarboxylase) (IDH) (NADP+-specific ICDH) (IDP)	intermediary metabolism and respiration	COG0538C
Rv3340	metC	47371.2	Cytoplasmic	O-acetylhomoserine sulfhydrylase	Probable O-acetylhomoserine sulfhydrylase MetC (homocysteine synthase) (O-acetylhomoserine (thiol)-lyase) (OAH sulfhydrylase) (O-acetyl-L-homoserine sulfhydrylase)	intermediary metabolism and respiration	COG2873E
Rv3342	-	26608.6	Unknown	methyltransferase	Possible methyltransferase (methylase)	intermediary metabolism and respiration	COG2226H
Rv3356c	folD	29451.7	Cytoplasmic	bifunctional methylenetetrahydrofolate dehydrogenase/methenyltetrahydrofolate cyclohydrolase	Probable bifunctional protein FolD: methylenetetrahydrofolate dehydrogenase + methenyltetrahydrofolate cyclohydrolase	intermediary metabolism and respiration	COG0190H
Rv3362c	-	20704.7	Cytoplasmic Membrane	ATP/GTP-binding protein	Predicted GTPase	Intermediary metabolism and respiration	COG2229R
Rv3364c	-	13440.5	Unknown	hypothetical protein Rv3364c	Uncharacterized distant relative of homeotic protein bithoraxoid	cell wall and cell processes	COG2018R
Rv3365c	-	93409.6	Cytoplasmic	hypothetical protein Rv3365c	Signal transduction histidine kinase	Regulatory proteins	COG0642T

			Membrane				
Rv3368c	-	23733	Cytoplasmic	oxidoreductase	Possible oxidoreductase	intermediary metabolism and respiration	COG0778C
Rv3369	-	15717.8	Cytoplasmic	hypothetical protein Rv3369	FMN binding protein	Intermediary metabolism and respiration	COG3871R
Rv3372	otsB2	41688.2	Cytoplasmic	trehalose 6-phosphate phosphatase	Trehalose 6-phosphate phosphatase OtsB2 (trehalose-phosphatase) (TPP)	virulence, detoxification, adaptation	COG1877G
Rv3373	echA18	22662.2	Cytoplasmic	enoyl-CoA hydratase	Probable enoyl-CoA hydratase EchA18 (enoyl hydratase) (unsaturated acyl-CoA hydratase) (crotonase)	lipid metabolism	COG1024I
Rv3376	-	23645.8	Cytoplasmic	phosphatase	Predicted hydrolase (HAD superfamily)	Intermediary metabolism and respiration	COG1011R
Rv3377c	-	55262.4	Cytoplasmic	type B diterpene cyclase	Halimadienyl diphosphate synthase	intermediary metabolism and respiration	COG1657I
Rv3384c	vapC46	14316.7	Cytoplasmic	ribonuclease VapC46	Possible toxin VapC46. Contains PIN domain.	virulence, detoxification, adaptation	COG1848R
Rv3389c	htdY	30296.3	Cytoplasmic	3-hydroxyacyl-thioester dehydratase HtdY	Probable 3-hydroxyacyl-thioester dehydratase HtdY	intermediary metabolism and respiration	COG2030I
Rv3390	lpqD	25056.4	Cytoplasmic Membrane	lipoprotein LpqD	Probable conserved lipoprotein LpqD	cell wall and cell processes	COG0406G
Rv3391	acrA1	70940.2	Cytoplasmic	acyl-CoA-reductase AcrA	Possible multi-functional enzyme with acyl-CoA-reductase activity AcrA1	lipid metabolism	COG3320Q
Rv3392c	cmaA1	32461.2	Cytoplasmic	cyclopropane mycolic acid synthase CmaA	Cyclopropane-fatty-acyl-phospholipid synthase 1 CmaA1 (cyclopropane fatty acid synthase) (CFA synthase) (cyclopropane mycolic acid synthase 1)	lipid metabolism	COG2230M
Rv3396c	guaA	56027.8	Cytoplasmic	GMP synthase	Probable GMP synthase [glutamine-hydrolyzing] GuaA (glutamine amidotransferase) (GMP synthetase)	intermediary metabolism and respiration	COG0519F
Rv3399	-	38140.2	Cytoplasmic	S-adenosylmethionine-dependent methyltransferase	Possible S-adenosylmethionine-dependent methyltransferase	lipid metabolism	COG3315Q
Rv3400	-	28227.7	Cytoplasmic	hydrolase	Probable hydrolase	intermediary metabolism and respiration	COG0637R
Rv3401	-	87315.5	Cytoplasmic	glycosyl hydrolase	Conserved protein	intermediary metabolism and respiration	COG1554G
Rv3404c	-	26483	Cytoplasmic	hypothetical protein Rv3404c	Methionyl-tRNA formyltransferase	Information pathways	COG0223J
Rv3409c	choD	63023.8	Unknown	cholesterol oxidase	Cholesterol oxidase ChoD (cholesterol-O2 oxidoreductase)	lipid metabolism	COG2303E
Rv3410c	guaB3	38990.6	Cytoplasmic	oxidoreductase	Probable inosine-5'-monophosphate dehydrogenase GuaB3 (imp	intermediary	COG0516F

					dehydrogenase) (inosinic acid dehydrogenase) (inosinate dehydrogenase) (imp oxidoreductase) (inosine-5'-monophosphate oxidoreductase) (IMPDH) (IMPD)	metabolism and respiration	
Rv3411c	guaB2	54867	Unknown	inosine-5'-monophosphate dehydrogenase	Probable inosine-5'-monophosphate dehydrogenase GuaB2 (imp dehydrogenase) (inosinic acid dehydrogenase) (inosinate dehydrogenase) (imp oxidoreductase) (inosine-5'-monophosphate oxidoreductase) (IMPDH) (IMPD)	intermediary metabolism and respiration	COG0516F
Rv3417c	groEL1	55877.5	Cytoplasmic	chaperonin GroEL	60 kDa chaperonin 1 GroEL1 (protein CPN60-1) (GroEL protein 1)	virulence, detoxification, adaptation	COG0459O
Rv3418c	groES	10772.2	Cytoplasmic	chaperonin GroES	10 kDa chaperonin GroES (protein CPN10) (protein GroES) (BCG-a heat shock protein) (10 kDa antigen)	virulence, detoxification, adaptation	COG0234O
Rv3422c	-	18195.7	Cytoplasmic	tRNA threonylcarbamoyladenosine biosynthesis protein	Predicted ATPase or kinase	Intermediary metabolism and respiration	COG0802R
Rv3423c	alr	43355.6	Cytoplasmic	alanine racemase	Alanine racemase Alr	intermediary metabolism and respiration	COG0787M
Rv3432c	gadB	50747.7	Cytoplasmic	glutamate decarboxylase GadB	Probable glutamate decarboxylase GadB	intermediary metabolism and respiration	COG0076E
Rv3433c	-	46857.3	Unknown	bifunctional ADP-dependent (S)-NAD(P)H-hydrate dehydratase/NAD(P)H-hydrate epimerase	Predicted sugar kinase	Intermediary metabolism and respiration	COG0063G
Rv3436c	glmS	67539.7	Cytoplasmic	glucosamine--fructose-6-phosphate aminotransferase	Probable glucosamine--fructose-6-phosphate aminotransferase [isomerizing] GlmS (hexosephosphate aminotransferase) (D-fructose-6-phosphate amidotransferase) (GFAT) (L-glutamine-D-fructose-6-phosphate amidotransferase) (glucosamine-6-phosphate synthase)	intermediary metabolism and respiration	COG0449M
Rv3438	-	29208.9	Unknown	hypothetical protein Rv3438	membrane associated process	cell wall and cell processes	COG0412Q
Rv3441c	mrsA	45873.1	Cytoplasmic	phosphoglucosamine mutase	Probable phospho-sugar mutase / MrsA protein homolog	intermediary metabolism and respiration	COG1109G
Rv3442c	rpsI	16436	Cytoplasmic	30S ribosomal protein S9	30S ribosomal protein S9 RpsI	information pathways	COG0103J
Rv3443c	rplM	16304.9	Cytoplasmic	50S ribosomal protein L13	50S ribosomal protein L13 RplM	information pathways	COG0102J
Rv3455c	truA	28059.9	Cytoplasmic	tRNA pseudouridine synthase A	Probable tRNA pseudouridine synthase a TruA (pseudouridylylate synthase I) (pseudouridine synthase I) (uracil hydrolyase)	information pathways	-
Rv3456c	rplQ	19475.2	Cytoplasmic	50S ribosomal protein L17	50S ribosomal protein L17 RplQ	information pathways	COG0203J
Rv3457c	rpoA	37706.5	Cytoplasmic	DNA-directed RNA polymerase subunit alpha	Probable DNA-directed RNA polymerase (alpha chain) RpoA (transcriptase alpha chain) (RNA polymerase alpha subunit) (DNA-directed RNA nucleotidyltransferase)	information pathways	COG0202K
Rv3458c	rpsD	23475.8	Cytoplasmic	30S ribosomal protein S4	30S ribosomal protein S4 RpsD	information pathways	COG0522J

Rv3459c	rpsK	14784.9	Cytoplasmic	30S ribosomal protein S11	30S ribosomal protein S11 RpsK	information pathways	COG0100J
Rv3460c	rpsM	14350.7	Cytoplasmic	30S ribosomal protein S13	30S ribosomal protein S13 RpsM	information pathways	COG0099J
Rv3462c	infA	8488.93	Cytoplasmic	translation initiation factor IF-1	Probable translation initiation factor if-1 InfA	information pathways	COG0361J
Rv3463	-	30650.8	Cytoplasmic	hypothetical protein Rv3463	Coenzyme F420-dependent N5,N10-methylene tetrahydromethanopterin reductase and related flavin-dependent oxidoreductases	Intermediary metabolism and respiration	COG2141C
Rv3464	rmlB	37557.7	Cytoplasmic	dTDP-glucose 4,6-dehydratase	dTDP-glucose 4,6-dehydratase RmlB	intermediary metabolism and respiration	COG1088M
Rv3465	rmlC	22314	Unknown	dTDP-4-dehydrorhamnose 3,5-epimerase	dTDP-4-dehydrorhamnose 3,5-epimerase RmlC (dTDP-4-keto-6-deoxyglucose 3,5-epimerase) (dTDP-L-rhamnose synthetase) (thymidine diphospho-4-keto-rhamnose 3,5-epimerase)	intermediary metabolism and respiration	COG1898M
Rv3478	PPE60	39381.6	Cytoplasmic Membrane	PE family protein PPE60	PE family protein PPE60	PE/PPE	COG5651N
Rv3479	-	109260	Cytoplasmic Membrane	transmembrane protein	Possible transmembrane protein	cell wall and cell processes	COG1752R
Rv3480c	-	53282.5	Cytoplasmic Membrane	diacylglycerol O-acyltransferase	Possible triacylglycerol synthase (diacylglycerol acyltransferase)	lipid metabolism	COG4908R
Rv3485c	-	33193.8	Cytoplasmic	short-chain type dehydrogenase/reductase	Probable short-chain type dehydrogenase/reductase	intermediary metabolism and respiration	COG1028IQR
Rv3490	otsA	55862.5	Cytoplasmic	trehalose-phosphate synthase	Alpha, alpha-trehalose-phosphate synthase [UDP-forming] OtsA (trehalose-6-phosphate synthase) (UDP-glucose-glucosephosphate glucosyltransferase) (trehalosephosphate-UDP glucosyltransferase) (trehalose-6-phosphate synthetase) (trehalose-phosphate synthase) (trehalose-phosphate synthetase) (transglucosylase) (trehalosephosphate-UDP glucosyl transferase)	virulence, detoxification, adaptation	COG0380G
Rv3492c	-	18006.9	Cytoplasmic Membrane	Mce associated protein	Conserved hypothetical Mce associated protein	conserved hypotheticals	-
Rv3494c	mce4F	59636.8	Unknown	Mce family protein Mce4	Mce-family protein Mce4F	virulence, detoxification, adaptation	COG1463Q
Rv3495c	lprN	41293	Unknown	Mce family lipoprotein LprN	Possible Mce-family lipoprotein LprN (Mce-family lipoprotein Mce4E)	cell wall and cell processes	COG1463Q
Rv3496c	mce4D	47201.4	Unknown	Mce family protein Mce4D	Mce-family protein Mce4D	virulence, detoxification, adaptation	COG1463Q
Rv3497c	mce4C	38132.4	Cytoplasmic Membrane	Mce family protein Mce4C	Mce-family protein Mce4C	virulence, detoxification, adaptation	COG1463Q
Rv3498c	mce4B	36880.3	Unknown	Mce family protein Mce4B	Mce-family protein Mce4B	virulence, detoxification, adaptation	COG1463Q
Rv3499c	mce4	42418.3	Unknown	Mce family protein Mce4A	Mce-family protein Mce4A	virulence,	COG1463Q

	A					detoxification, adaptation	
Rv3501c	yrbE4A	26621.7	Cytoplasmic Membrane	integral membrane protein	Conserved integral membrane protein YrbE4A. Possible ABC transporter.	virulence, detoxification, adaptation	COG0767Q
Rv3502c	-	32734.8	Cytoplasmic	short-chain type dehydrogenase/reductase	Probable short-chain type dehydrogenase/reductase. Possible 17-beta-hydroxysteroid dehydrogenase.	intermediary metabolism and respiration	COG1028IQR
Rv3504	fadE26	43784.6	Cytoplasmic	acyl-CoA dehydrogenase FadE26	Probable acyl-CoA dehydrogenase FadE26	lipid metabolism	COG1960I
Rv3505	fadE27	39007.9	Cytoplasmic	acyl-CoA dehydrogenase FadE27	Probable acyl-CoA dehydrogenase FadE27	lipid metabolism	COG1960I
Rv3509c	ilvX	52040.3	Cytoplasmic Membrane	acetohydroxyacid synthase large subunit	Probable acetohydroxyacid synthase IlvX (acetolactate synthase)	intermediary metabolism and respiration	COG0028EH
Rv3510c	-	31067.7	Cytoplasmic Membrane	hypothetical protein Rv3510c	Predicted metal-dependent hydrolase of the TIM-barrel fold	Intermediary metabolism and respiration	COG2159R
Rv3515c	fadD19	59708.5	Cytoplasmic	long-chain-fatty-acid--CoA ligase FadD19	Fatty-acid-CoA ligase FadD19 (fatty-acid-CoA synthetase) (fatty-acid-CoA synthase)	lipid metabolism	COG0318IQ
Rv3516	echA19	28216.5	Cytoplasmic	enoyl-CoA hydratase EchA19	Possible enoyl-CoA hydratase EchA19 (enoyl hydratase) (unsaturated acyl-CoA hydratase) (crotonase)	lipid metabolism	COG1024I
Rv3518c	cyp142	44398.5	Cytoplasmic Membrane	cytochrome P450 monooxygenase Cyp142	Probable cytochrome P450 monooxygenase 142 Cyp142	intermediary metabolism and respiration	COG2124Q
Rv3519	-	25960	Cytoplasmic	hypothetical protein Rv3519	Carboxy-Lyase	Intermediary metabolism and respiration	-
Rv3520c	-	37584.9	Cytoplasmic	coenzyme F420-dependent oxidoreductase	Possible coenzyme F420-dependent oxidoreductase	intermediary metabolism and respiration	COG2141C
Rv3523	ltp3	41421.3	Cytoplasmic	lipid carrier protein	Probable lipid carrier protein or keto acyl-CoA thiolase Ltp3	lipid metabolism	COG0183I
Rv3526	kshA	44235.5	Cytoplasmic	3-ketosteroid-9-alpha-hydroxylase oxygenase subunit	Oxygenase component of 3-ketosteroid-9-alpha-hydroxylase KshA	intermediary metabolism and respiration	COG4638PR
Rv3528c	-	27440.4	Cytoplasmic	hypothetical protein Rv3528c	Unknown protein	conserved hypotheticals	-
Rv3534c	hsaF	36442.4	Cytoplasmic	4-hydroxy-2-oxovalerate aldolase	Probable 4-hydroxy-2-oxovalerate aldolase (HOA)	intermediary metabolism and respiration	COG0119E
Rv3535c	hsaG	32008.5	Cytoplasmic	acetaldehyde dehydrogenase	Probable acetaldehyde dehydrogenase (acetaldehyde dehydrogenase [acetylating])	intermediary metabolism and respiration	COG4569Q
Rv3536c	hsaE	27891.6	Cytoplasmic	hydratase	Probable hydratase	intermediary metabolism and	COG3971Q

						respiration	
Rv3537	kstD	60636	Cytoplasmic Membrane	3-oxosteroid 1-dehydrogenase	Probable dehydrogenase	intermediary metabolism and respiration	COG1053C
Rv3544c	fadE28	35468.5	Unknown	acyl-CoA dehydrogenase FadE28	Probable acyl-CoA dehydrogenase FadE28	lipid metabolism	COG1960I
Rv3545c	cyp125	48400.6	Cytoplasmic Membrane	steroid C26-monooxygenase	Probable cytochrome P450 125 Cyp125	intermediary metabolism and respiration	COG2124Q
Rv3546	fadA5	41328.8	Cytoplasmic	acetyl-CoA acetyltransferase FadA	Probable acetyl-CoA acetyltransferase FadA5 (acetoacetyl-CoA thiolase)	lipid metabolism	COG0183I
Rv3547	ddn	17371	Unknown	deazaflavin-dependent nitroreductase	Deazaflavin-dependent nitroreductase Ddn	intermediary metabolism and respiration	-
Rv3548c	-	31070.1	Unknown	short-chain type dehydrogenase/reductase	Probable short-chain type dehydrogenase/reductase	intermediary metabolism and respiration	COG1028IQR
Rv3550	echA20	26332.9	Cytoplasmic	enoyl-CoA hydratase EchA20	Probable enoyl-CoA hydratase EchA20 (enoyl hydratase) (unsaturated acyl-CoA hydratase) (crotonase)	lipid metabolism	COG1024I
Rv3554	fdxB	76533.7	Cytoplasmic Membrane	electron transfer protein FdxB	Possible electron transfer protein FdxB	intermediary metabolism and respiration	COG3239I
Rv3556c	fadA6	40902.3	Cytoplasmic	acetyl-CoA acetyltransferase FadA	Probable acetyl-CoA acetyltransferase FadA6 (acetoacetyl-CoA thiolase)	lipid metabolism	COG0183I
Rv3559c	-	27429.9	Cytoplasmic	oxidoreductase	Probable oxidoreductase	intermediary metabolism and respiration	COG1028IQR
Rv3560c	fadE30	42732.1	Cytoplasmic	acyl-CoA dehydrogenase FadE30	Probable acyl-CoA dehydrogenase FadE30	lipid metabolism	COG1960I
Rv3561	fadD3	54076.5	Cytoplasmic	fatty-acid--CoA ligase FadD3	Probable fatty-acid-CoA ligase FadD3 (fatty-acid-CoA synthetase) (fatty-acid-CoA synthase)	lipid metabolism	COG0318IQ
Rv3562	fadE31	41332.5	Cytoplasmic	acyl-CoA dehydrogenase FadE31	Probable acyl-CoA dehydrogenase FadE31	lipid metabolism	COG1960I
Rv3563	fadE32	33455.1	Cytoplasmic	acyl-CoA dehydrogenase FadE32	Probable acyl-CoA dehydrogenase FadE32	lipid metabolism	COG1960I
Rv3565	aspB	41048.7	Unknown	aspartate aminotransferase AspB	Possible aspartate aminotransferase AspB (transaminase A) (ASPAT) (glutamic--oxaloacetic transaminase) (glutamic--aspartic transaminase)	intermediary metabolism and respiration	COG0436E
Rv3566c	nat	31029.2	Cytoplasmic	arylamine N-acetyltransferase	Arylamine N-acetyltransferase Nat (arylamine acetylase)	intermediary metabolism and respiration	COG2162Q
Rv3568c	hsaC	33582.3	Cytoplasmic	extradiol dioxygenase	3,4-DHSA dioxygenase	intermediary metabolism and respiration	COG2514R
Rv3569c	hsaD	31875.4	Cytoplasmic	4,5:9,10-diseco-3-hydroxy-	4,9-DHSA hydrolase	intermediary	COG0596R

				5,9,17-trioxoandrosta-1(10),2-diene-4-oate hydrolase		metabolism and respiration	
Rv3570c	hsaA	43110.5	Cytoplasmic Membrane	flavin-dependent monooxygenase oxygenase subunit HsaA	Possible oxidoreductase. Possible 3-hydroxy-9,10-seconandrost-1,3,5(10)-triene-9,17-dione hydroxylase.	intermediary metabolism and respiration	COG1960I
Rv3573c	fadE34	74546.2	Cytoplasmic	acyl-CoA dehydrogenase FadE34	Probable acyl-CoA dehydrogenase FadE34	lipid metabolism	COG1960I
Rv3575c	-	38177.7	Cytoplasmic	LacI family transcriptional regulator	Transcriptional regulatory protein (probably LacI-family)	Regulatory proteins	COG1609K
Rv3579c	-	34009.8	Unknown	tRNA/rRNA methyltransferase	Possible tRNA/rRNA methyltransferase	intermediary metabolism and respiration	COG0566J
Rv3580c	cysS1	51822.5	Cytoplasmic	cysteine--tRNA ligase	Cysteinyl-tRNA synthetase 1 CysS1 (cysteine--tRNA ligase 1) (CYSRS 1) (cysteine transase)	information pathways	COG0215J
Rv3581c	ispF	16425.8	Cytoplasmic Membrane	2C-methyl-D-erythritol 2,4-cyclodiphosphate synthase	Probable 2C-methyl-D-erythritol 2,4-cyclodiphosphate synthase IspF (MECPS)	intermediary metabolism and respiration	COG0245I
Rv3582c	ispD	24041.5	Cytoplasmic	2-C-methyl-D-erythritol 4-phosphate cytidyltransferase	4-diphosphocytidyl-2C-methyl-D-erythritol synthase IspD (MEP cytidyltransferase) (MCT)	intermediary metabolism and respiration	COG1211I
Rv3583c	-	17907.4	Cytoplasmic	RNA polymerase-binding transcription factor CarD	Possible transcription factor	Regulatory proteins	COG1329K
Rv3584	lpqE	18786.6	Unknown	lipoprotein LpqE	Possible conserved lipoprotein LpqE	cell wall and cell processes	-
Rv3586	-	39031.1	Cytoplasmic	DNA integrity scanning protein DisA	Predicted nucleic-acid-binding protein (contains the HHH domain)	Regulatory proteins	COG1623R
Rv3587c	-	27066.9	Cytoplasmic Membrane	membrane protein	Probable conserved membrane protein	cell wall and cell processes	-
Rv3591c	-	28510.4	Cytoplasmic	hydrolase	Possible hydrolase	intermediary metabolism and respiration	COG0596R
Rv3592	-	11184.7	Cytoplasmic	heme-degrading monooxygenase	Possible heme degrading protein MhuD	intermediary metabolism and respiration	COG2329R
Rv3596c	clpC1	93552.4	Cytoplasmic	ATP-dependent protease ATP-binding subunit ClpC	Probable ATP-dependent protease ATP-binding subunit ClpC1	intermediary metabolism and respiration	COG0542O
Rv3597c	lsr2	12098.4	Cytoplasmic	iron-regulated H-NS-like protein	Iron-regulated H-NS-like protein Lsr2	information pathways	-
Rv3598c	lysS	55677.3	Cytoplasmic	lysine--tRNA ligase	Lysyl-tRNA synthetase 1 LysS (lysine--tRNA ligase 1) (LysRS 1) (lysine transase)	information pathways	COG1190J
Rv3600c	-	29272.6	Cytoplasmic	type III pantothenate kinase	Putative transcriptional regulator, homolog of Bvg accessory factor	Regulatory proteins	COG1521K
Rv3602c	panC	32677.5	Cytoplasmic	pantothenate synthetase	Pantoate--beta-alanine ligase PanC (pantothenate synthetase) (pantoate activating enzyme)	intermediary metabolism and respiration	COG0414H

Rv3603c	-	31104.6	Cytoplasmic	alanine/leucine-rich protein	folic acid-containing compound biosynthetic process	Intermediary metabolism and respiration	COG5495S
Rv3604c	-	43058.4	Cytoplasmic Membrane	transmembrane protein	Probable conserved transmembrane protein rich in alanine and arginine and proline	cell wall and cell processes	COG5178A
Rv3608c	folP1	28810.9	Unknown	dihydropteroate synthase	Dihydropteroate synthase 1 FolP (DHPS 1) (dihydropteroate pyrophosphorylase 1) (dihydropteroate diphosphorylase 1)	intermediary metabolism and respiration	COG0294H
Rv3609c	folE	22394.7	Unknown	GTP cyclohydrolase I	GTP cyclohydrolase I FolE (GTP-ch-I)	intermediary metabolism and respiration	COG0302H
Rv3610c	ftsH	81985.5	Cytoplasmic Membrane	zinc metalloprotease FtsH	Membrane-bound protease FtsH (cell division protein)	cell wall and cell processes	COG0465O
Rv3614c	espD	19802.5	Unknown	ESX-1 secretion-associated protein EspD	ESX-1 secretion-associated protein EspD	cell wall and cell processes	-
Rv3616c	espA	39888.3	Cytoplasmic	ESX-1 secretion-associated protein EspA	ESX-1 secretion-associated protein A, EspA	cell wall and cell processes	-
Rv3617	ephA	35096.6	Cytoplasmic	epoxide hydrolase EphA	Probable epoxide hydrolase EphA (epoxide hydratase) (arene-oxide hydratase)	virulence, detoxification, adaptation	COG0596R
Rv3618	-	43277.1	Cytoplasmic	monooxygenase	Possible monooxygenase	intermediary metabolism and respiration	COG2141C
Rv3619c	esxV	9832.78	Unknown	ESAT-6 like protein EsxV	Putative ESAT-6 like protein EsxV (ESAT-6 like protein 1)	cell wall and cell processes	-
Rv3623	lpqG	24836.8	Unknown	lipoprotein LpqG	Probable conserved lipoprotein LpqG	cell wall and cell processes	COG2968S
Rv3624c	hpt	23603.8	Cytoplasmic	hypoxanthine-guanine phosphoribosyltransferase	Hypoxanthine-guanine phosphoribosyltransferase Hpt (HGPR) (HGPRase) (hypoxanthine phosphoribosyltransferase) (imp pyrophosphorylase) (imp diphosphorylase) (transphosphoribosyltransferase) (guanine phosphoribosyltransferase)	intermediary metabolism and respiration	COG0634F
Rv3625c	mesJ	34129.9	Cytoplasmic	tRNA(Ile)-lysine synthase	Possible cell cycle protein MesJ	cell wall and cell processes	COG0037D
Rv3627c	-	46834.3	Unknown	hypothetical protein Rv3627c	D-alanyl-D-alanine carboxypeptidase (penicillin-binding protein 4)	cell wall and cell processes	COG2027M
Rv3628	ppa	18296.6	Cytoplasmic	inorganic pyrophosphatase	Inorganic pyrophosphatase Ppa (pyrophosphate phosphohydrolase) (PPASE) (inorganic diphosphatase) (diphosphate phosphohydrolase)	intermediary metabolism and respiration	COG0221C
Rv3631	-	26212.9	Cytoplasmic	transferase	Possible transferase (possibly glycosyltransferase)	intermediary metabolism and respiration	COG0463M
Rv3633	-	32447	Cytoplasmic	hypothetical protein Rv3633	Protein involved in biosynthesis of mitomycin antibiotics/polyketide fumonisin	Lipid metabolism	COG5285Q
Rv3634c	galE1	33549.6	Cytoplasmic	UDP-glucose 4-epimerase	UDP-glucose 4-epimerase GalE1 (galactowaldenase) (UDP-	intermediary	COG0451MG

					galactose 4-epimerase) (uridine diphosphate galactose 4-epimerase) (uridine diphospho-galactose 4-epimerase)	metabolism and respiration	
Rv3644c	-	41784.5	Cytoplasmic	DNA polymerase	Possible DNA polymerase	information pathways	COG0470L
Rv3645	-	58759.3	Cytoplasmic Membrane	transmembrane protein	Probable conserved transmembrane protein	cell wall and cell processes	COG2114T
Rv3646c	topA	102318	Cytoplasmic	DNA topoisomerase I	DNA topoisomerase I TopA (omega-protein) (relaxing enzyme) (untwisting enzyme) (swivelase) (type I DNA topoisomerase) (nicking-closing enzyme) (TOPO I)	information pathways	COG0550L
Rv3648c	cspA	7370.2	Cytoplasmic	cold shock protein A	Probable cold shock protein A CspA	virulence, detoxification, adaptation	COG1278K
Rv3649	-	81410.4	Cytoplasmic	helicase	Probable helicase	information pathways	COG1205R
Rv3651	-	37967.3	Unknown	hypothetical protein Rv3651	Conserved hypothetical protein	conserved hypotheticals	-
Rv3661	-	30679.1	Unknown	hypothetical protein Rv3661	Conserved hypothetical protein	virulence, detoxification, adaptation	COG0560E
Rv3663c	dppD	58943.6	Cytoplasmic Membrane	dipeptide ABC transporter ATP-binding protein DppD	Probable dipeptide-transport ATP-binding protein ABC transporter DppD	cell wall and cell processes	COG1123R
Rv3666c	dppA	58381.5	Cellwall	dipeptide ABC transporter substrate-binding lipoprotein DppA	Probable periplasmic dipeptide-binding lipoprotein DppA	cell wall and cell processes	COG4166E
Rv3667	acs	71443.9	Cytoplasmic	acetyl-CoAsynthetase	Acetyl-coenzyme A synthetase Acs (acetate--CoA ligase) (acetyl-CoA synthetase) (acetyl-CoA synthase) (acyl-activating enzyme) (acetate thiokinase) (acetyl-activating enzyme) (acetate-coenzyme A ligase) (acetyl-coenzyme A synthase)	lipid metabolism	COG0365I
Rv3669	-	18887.1	Cytoplasmic Membrane	transmembrane protein	Probable conserved transmembrane protein	cell wall and cell processes	-
Rv3670	ephE	36333.6	Cytoplasmic	epoxide hydrolase EphE	Possible epoxide hydrolase EphE (epoxide hydratase) (arene-oxide hydratase)	virulence, detoxification, adaptation	COG0596R
Rv3671c	-	40721.4	Cytoplasmic Membrane	serine protease	Membrane-associated serine protease	intermediary metabolism and respiration	COG0265O
Rv3672c	-	28890.8	Cytoplasmic Membrane	hypothetical protein Rv3672c	NTP pyrophosphohydrolases including oxidative damage repair enzymes	Information pathways	COG1051F
Rv3674c	nth	26998.3	Cytoplasmic Membrane	endonuclease III	Probable endonuclease III Nth (DNA-(apurinic or apyrimidinic site)lyase) (AP lyase) (AP endonuclease class I) (endodeoxyribonuclease (apurinic or apyrimidinic)) (deoxyribonuclease (apurinic or apyrimidinic))	information pathways	COG0177L
Rv3676	crp	24759.3	Cytoplasmic	cAMP receptor protein	Transcriptional regulatory protein Crp (Crp/Fnr-family)	Regulatory proteins	COG0664T
Rv3677c	-	28490	Cytoplasmic	beta lactamase	Possible hydrolase	intermediary metabolism and respiration	COG0491R
Rv3678c	-	15245.6	Cytoplasmic	hypothetical protein Rv3678c	Putative translation initiation inhibitor, yjgF family	Information pathways	COG0251J

Rv3679	-	35856.3	Cytoplasmic	anion transporter ATPase	Probable anion transporter ATPase	cell wall and cell processes	COG0003P
Rv3680	-	41404.6	Cytoplasmic Membrane	anion transporter ATPase	Probable anion transporter ATPase	cell wall and cell processes	COG0003P
Rv3682	ponA2	84635.8	Extracellular	bifunctional penicillin-insensitive transglycosylase/penicillin-sensitive transpeptidase	Probable bifunctional membrane-associated penicillin-binding protein 1A/1B PonA2 (murein polymerase) [includes: penicillin-insensitive transglycosylase (peptidoglycan TGASE) + penicillin-sensitive transpeptidase (DD-transpeptidase)]	cell wall and cell processes	COG0744M
Rv3683	-	34394.4	Unknown	hypothetical protein Rv3683	Predicted phosphohydrolases	Intermediary metabolism and respiration	COG1408R
Rv3684	-	37623.3	Unknown	lyase	Probable lyase	intermediary metabolism and respiration	-
Rv3685c	cyp13 ₇	52265.6	Cytoplasmic	cytochrome P450 Cyp137	Probable cytochrome P450 137 Cyp137	intermediary metabolism and respiration	COG2124Q
Rv3688c	-	16701.1	Cytoplasmic	hypothetical protein Rv3688c	Conserved protein	conserved hypotheticals	COG1610S
Rv3691	-	35492.5	Cytoplasmic Membrane	hypothetical protein Rv3691	membrane associated process	cell wall and cell processes	-
Rv3692	moxR ₂	37904.4	Cytoplasmic	methanol dehydrogenase transcriptional regulator MoxR	Probable methanol dehydrogenase transcriptional regulatory protein MoxR2	Regulatory proteins	COG0714R
Rv3693	-	47687.2	Cytoplasmic Membrane	membrane protein	Possible conserved membrane protein	cell wall and cell processes	COG1721R
Rv3695	-	33167.9	Cytoplasmic Membrane	membrane protein	Possible conserved membrane protein	cell wall and cell processes	COG1714S
Rv3696c	glpK	55827.8	Cytoplasmic	glycerol kinase	Probable glycerol kinase GlpK (ATP:glycerol 3-phosphotransferase) (glycerokinase) (GK)	intermediary metabolism and respiration	COG0554C
Rv3698	-	56929.5	Cytoplasmic	hypothetical protein Rv3698	Glutamate-Cysteine ligase	Intermediary metabolism and respiration	-
Rv3699	-	24998.1	Cytoplasmic	hypothetical protein Rv3699	SAM-dependent methyltransferases	Intermediary metabolism and respiration	COG2227H
Rv3700c	egtE	40446.1	Cytoplasmic	pyridoxal-phosphate-dependent protein EgtE	Conserved hypothetical protein	intermediary metabolism and respiration	COG0520E
Rv3701c	egtD	35404.1	Cytoplasmic	histidine-specific methyltransferase EtgD	related to cytosine deaminase activity	Intermediary metabolism and respiration	COG4301S
Rv3702c	egtC	24627	Unknown	amidohydrolase EgtC	Predicted glutamine amidotransferase	Intermediary metabolism and respiration	COG0121R

Rv3708c	asd	36230.2	Cytoplasmic	aspartate-semialdehyde dehydrogenase	Aspartate-semialdehyde dehydrogenase Asd (ASA dehydrogenase) (ASADH) (aspartic semialdehyde dehydrogenase) (L-aspartate-beta-semialdehyde dehydrogenase)	intermediary metabolism and respiration	COG0136E
Rv3709c	ask	44429.6	Cytoplasmic	aspartokinase	Aspartokinase Ask (aspartate kinase) [contains: aspartokinase alpha subunit (Ask-alpha); and aspartokinase beta subunit (Ask-beta)]	intermediary metabolism and respiration	COG0527E
Rv3710	leuA	70081.6	Cytoplasmic	2-isopropylmalate synthase	2-isopropylmalate synthase LeuA (alpha-isopropylmalate synthase) (alpha-IPM synthetase) (IPMS)	intermediary metabolism and respiration	COG0119E
Rv3712	-	43389.4	Cytoplasmic	ligase	Possible ligase	intermediary metabolism and respiration	COG0769M
Rv3713	cobQ2	24125.7	Cytoplasmic	cobryric acid synthase CobQ	Possible cobryric acid synthase CobQ2	intermediary metabolism and respiration	COG3442R
Rv3715c	recR	22119.3	Cytoplasmic	recombination protein RecR	Probable recombination protein RecR	information pathways	COG0353L
Rv3716c	-	13357.5	Unknown	hypothetical protein Rv3716c	DNA-repair	Information pathways	COG0718S
Rv3718c	-	15661.8	Unknown	hypothetical protein Rv3718c	Conserved protein	conserved hypotheticals	-
Rv3719	-	53007	Unknown	hypothetical protein Rv3719	FAD/FMN-containing dehydrogenases	Intermediary metabolism and respiration	COG0277C
Rv3720	-	46852.8	Cytoplasmic	fatty acid synthase	Possible fatty acid synthase	lipid metabolism	-
Rv3721c	dnaZ X	61891.3	Cytoplasmic	DNA polymerase III subunit gamma/tau	DNA polymerase III (subunit gamma/tau) DnaZ/X	information pathways	COG2812L
Rv3722c	-	47309	Unknown	hypothetical protein Rv3722c	Transcriptional regulators containing a DNA-binding HTH domain and an aminotransferase domain (MocR family) and their eukaryotic orthologs	Regulatory proteins	COG1167KE
Rv3723	-	27367.4	Cytoplasmic Membrane	transmembrane protein	Probable conserved transmembrane protein	cell wall and cell processes	COG1178P
Rv3725	-	32725.7	Unknown	oxidoreductase	Possible oxidoreductase	intermediary metabolism and respiration	COG0451MG
Rv3726	-	42035.8	Cytoplasmic	dehydrogenase	Possible dehydrogenase	intermediary metabolism and respiration	COG1063ER
Rv3729	-	84129	Cytoplasmic	transferase	Possible transferase	intermediary metabolism and respiration	COG1964R
Rv3730c	-	38543.7	Unknown	hypothetical protein Rv3730c	Predicted eukaryotic-type DNA primase	Information pathways	COG3285L
Rv3731	ligC	40159.7	Cytoplasmic	DNA ligase C	Possible ATP-dependent DNA ligase LigC (polydeoxyribonucleotide synthase [ATP]) (polynucleotide ligase [ATP]) (sealase) (DNA repair protein) (DNA joinase)	information pathways	COG1793L
Rv3732	-	37532.8	Unknown	hypothetical protein Rv3732	Conserved protein	conserved hypotheticals	COG4402S

Rv3734c	tgs2	49303.8	Cytoplasmic	diacylglycerol O-acyltransferase	Putative triacylglycerol synthase (diacylglycerol acyltransferase) Tgs2	lipid metabolism	-
Rv3747	-	13507.5	Unknown	hypothetical protein Rv3747	link to PE family	Regulatory proteins	-
Rv3753c	-	17917	Cytoplasmic	hypothetical protein Rv3753c	Conserved protein	conserved hypotheticals	-
Rv3754	tyrA	30991.1	Unknown	prephenate dehydrogenase TyrA	Prephenate dehydrogenase TyrA (PDH) (hydroxyphenylpyruvate synthase)	intermediary metabolism and respiration	-
Rv3755c	-	21969.8	Cytoplasmic	hypothetical protein Rv3755c	Conserved protein	conserved hypotheticals	COG3554S
Rv3758c	proV	39552.3	Cytoplasmic Membrane	glycine betaine/carnitine/choline/L-proline ABC transporter ATP-binding protein ProV	Possible osmoprotectant (glycine betaine/carnitine/choline/L-proline) transport ATP-binding protein ABC transporter ProV	virulence, detoxification, adaptation	COG1125E
Rv3759c	proX	33357.9	Unknown	glycine betaine/carnitine/choline/L-proline ABC transporter substrate-binding lipoprotein ProX	Possible osmoprotectant (glycine betaine/carnitine/choline/L-proline) binding lipoprotein ProX	virulence, detoxification, adaptation	COG1732M
Rv3761c	fadE36	38532.7	Cytoplasmic	acyl-CoA dehydrogenase FadE36	Possible acyl-CoA dehydrogenase FadE36	lipid metabolism	COG3173R
Rv3762c	-	68451	Cytoplasmic	hydrolase	Possible hydrolase	intermediary metabolism and respiration	COG2015Q
Rv3763	lpqH	15114.8	Unknown	lipoprotein LpqH	19 kDa lipoprotein antigen precursor LpqH	cell wall and cell processes	-
Rv3767c	-	34723.2	Unknown	S-adenosylmethionine-dependent methyltransferase	Possible S-adenosylmethionine-dependent methyltransferase	lipid metabolism	COG3315Q
Rv3768	-	13166.8	Unknown	hypothetical protein Rv3768	Unknown protein	conserved hypotheticals	COG4319S
Rv3772	hisC2	38007.5	Unknown	histidinol-phosphate aminotransferase	Probable histidinol-phosphate aminotransferase HisC2 (imidazole acetol-phosphate transaminase) (imidazolylacetolphosphate aminotransferase)	intermediary metabolism and respiration	COG0079E
Rv3774	echA21	29101.2	Unknown	enoyl-CoA hydratase EchA21	Possible enoyl-CoA hydratase EchA21 (enoyl hydratase) (unsaturated acyl-CoA hydratase) (crotonase)	lipid metabolism	COG1024I
Rv3775	lipE	45273.7	Unknown	lipase LipE	Probable lipase LipE	intermediary metabolism and respiration	COG1680V
Rv3777	-	33931.3	Cytoplasmic	oxidoreductase	Probable oxidoreductase	intermediary metabolism and respiration	COG0604CR
Rv3778c	-	41851.7	Cytoplasmic	aminotransferase	Possible aminotransferase	intermediary metabolism and respiration	COG0520E
Rv3779	-	71762.8	Cytoplasmic	transmembrane protein	Probable conserved transmembrane protein alanine and leucine	cell wall and cell	COG2194R

			Membrane		rich	processes	
Rv3780	-	19484	Cytoplasmic	hypothetical protein Rv3780	Conserved protein	conserved hypotheticals	-
Rv3781	rfbE	29960.3	Cytoplasmic Membrane	O-antigen/lipopolysaccharide ABC transporter ATP-binding protein RfbE	Probable O-antigen/lipopolysaccharide transport ATP-binding protein ABC transporter RfbE	cell wall and cell processes	COG1134GM
Rv3782	glfT1	33864	Cytoplasmic Membrane	galactofuranosyl transferase GlfT	UDP-galactofuranosyl transferase GlfT1	cell wall and cell processes	COG1216R
Rv3786c	-	44944.1	Cytoplasmic	hypothetical protein Rv3786c	Membrane proteins related to metalloendopeptidases	Intermediary metabolism and respiration	COG0739M
Rv3787c	-	33358.4	Cytoplasmic	S-adenosyl-L-methionine- dependent methyltransferase	O-Methyltransferase involved in polyketide biosynthesis	Lipid metabolism	COG3315Q
Rv3788	-	17374.4	Unknown	hypothetical protein Rv3788	Transcription elongation factor	Information pathways	COG0782K
Rv3790	dprE1	50163.2	Unknown	decaprenylphosphoryl-beta-D- ribose oxidase	Decaprenylphosphoryl-beta-D-ribose 2'-oxidase	lipid metabolism	COG0277C
Rv3791	dprE2	27468.8	Cytoplasmic	decaprenylphosphoryl-D-2-keto erythropentose reductase	Decaprenylphosphoryl-D-2-keto erythro pentose reductase	lipid metabolism	COG0300R
Rv3793	embC	117490	Cytoplasmic Membrane	arabinoxyltransferase C	Integral membrane indolylacetylaminositol arabinoxyltransferase EmbC (arabinoxylindolylacetylaminositol synthase)	cell wall and cell processes	-
Rv3794	embA	115692	Cytoplasmic Membrane	arabinoxyltransferase A	Integral membrane indolylacetylaminositol arabinoxyltransferase EmbA (arabinoxylindolylacetylaminositol synthase)	cell wall and cell processes	COG1807M
Rv3795	embB	118021	Cytoplasmic Membrane	arabinoxyltransferase B	Integral membrane indolylacetylaminositol arabinoxyltransferase EmbB (arabinoxylindolylacetylaminositol synthase)	cell wall and cell processes	-
Rv3796	-	39443.6	Cytoplasmic Membrane	hypothetical protein Rv3796	Metal-dependent hydrolases of the beta-lactamase superfamily III	Intermediary metabolism and respiration	COG1234R
Rv3799c	accD4	56647.7	Cytoplasmic	propionyl-CoA carboxylase subunit beta AccD	Probable propionyl-CoA carboxylase beta chain 4 AccD4 (pccase) (propanoyl-CoA:carbon dioxide ligase)	lipid metabolism	COG4799I
Rv3800c	pks13	186446	Cytoplasmic Membrane	polyketide synthase	Polyketide synthase Pks13	lipid metabolism	COG3321Q
Rv3801c	fadD3 2	69232	Cytoplasmic	long-chain-fatty-acid--AMP ligase FadD32	Fatty-acid-AMP ligase FadD32 (fatty-acid-AMP synthetase) (fatty-acid-AMP synthase). Also shown to have acyl-ACP ligase activity.	lipid metabolism	COG0318IQ
Rv3803c	fbpD	31088.9	Extracellular	MPT51/MPB51 antigen	Secreted MPT51/MPB51 antigen protein FbpD (MPT51/MPB51 antigen 85 complex C) (AG58C) (mycolyl transferase 85C) (fibronectin-binding protein C) (85C)	lipid metabolism	COG0627R
Rv3804c	fbpA	35686.2	Extracellular	diacylglycerol acyltransferase/mycolyltransferase Ag85A	Secreted antigen 85-a FbpA (mycolyl transferase 85A) (fibronectin-binding protein A) (antigen 85 complex A)	lipid metabolism	COG0627R
Rv3806c	ubiA	32654.5	Cytoplasmic Membrane	decaprenyl-phosphate phosphoribosyltransferase	Decaprenylphosphoryl-5-phosphoribose (DPPR) synthase (decaprenyl-phosphate 5-phosphoribosyltransferase)	cell wall and cell processes	COG0382H
Rv3807c	-	17218.4	Cytoplasmic Membrane	decaprenylphosphoryl-5- phosphoribose phosphatase	Possible conserved transmembrane protein	cell wall and cell processes	COG0671I

Rv3808c	glfT2	71506.5	Cytoplasmic	galactofuranosyl transferase GlfT	Bifunctional UDP-galactofuranosyl transferase GlfT2	cell wall and cell processes	COG1216R
Rv3809c	glf	45814.4	Cytoplasmic	UDP-galactopyranose mutase	UDP-galactopyranose mutase Glf (UDP-GALP mutase) (NAD ⁺ -flavin adenine dinucleotide-requiring enzyme)	cell wall and cell processes	COG0562M
Rv3813c	-	28515.4	Cytoplasmic	hypothetical protein Rv3813c	Predicted hydrolases of the HAD superfamily	Intermediary metabolism and respiration	COG0561R
Rv3814c	-	29290.9	Cytoplasmic	acyltransferase	Possible acyltransferase	intermediary metabolism and respiration	COG0204I
Rv3816c	-	28421.8	Cytoplasmic	acyltransferase	Possible acyltransferase	intermediary metabolism and respiration	COG0204I
Rv3818	-	57615.2	Cytoplasmic	hypothetical protein Rv3818	Predicted Zn-dependent hydrolases of the beta-lactamase fold	Intermediary metabolism and respiration	COG2220R
Rv3819	-	12825.1	Cytoplasmic	hypothetical protein Rv3819	Unknown protein	conserved hypotheticals	-
Rv3820c	papA2	52118	Cytoplasmic	trehalose-2-sulfate acyltransferase	Possible conserved polyketide synthase associated protein PapA2	lipid metabolism	-
Rv3825c	pks2	225758	Cytoplasmic Membrane	phthioceranic/hydroxyphthioceranic acid synthase	Polyketide synthase Pks2	lipid metabolism	COG3321Q
Rv3826	fadD23	62840.8	Cytoplasmic	long-chain-fatty-acid--CoA ligase FadD23	Probable fatty-acid-AMP ligase FadD23 (fatty-acid-AMP synthetase) (fatty-acid-AMP synthase)	lipid metabolism	COG0318IQ
Rv3829c	-	57073.2	Cytoplasmic	dehydrogenase	Probable dehydrogenase	intermediary metabolism and respiration	COG1233Q
Rv3831	-	17867.6	Cytoplasmic Membrane	hypothetical protein Rv3831	Hypothetical protein	conserved hypotheticals	-
Rv3834c	serS	45292.3	Cytoplasmic	serine--tRNA ligase	SERYL-tRNA synthetase SerS (serine--tRNA ligase) (SERRS) (serine transase)	information pathways	COG0172J
Rv3838c	pheA	33601.2	Cytoplasmic	prephenate dehydratase	Prephenate dehydratase PheA	intermediary metabolism and respiration	COG0077E
Rv3841	bfrB	20442	Cytoplasmic	bacterioferritin BfrB	Bacterioferritin BfrB	intermediary metabolism and respiration	COG1528P
Rv3842c	glpQ1	29917	Cytoplasmic	glycerophosphoryl diester phosphodiesterase	Probable glycerophosphoryl diester phosphodiesterase GlpQ1 (glycerophosphodiester phosphodiesterase)	intermediary metabolism and respiration	COG0584C
Rv3843c	-	37352.8	Cytoplasmic Membrane	transmembrane protein	Probable conserved transmembrane protein	cell wall and cell processes	-
Rv3846	sodA	23001.9	Extracellular	superoxide dismutase	Superoxide dismutase [FE] SodA	virulence, detoxification, adaptation	COG0605P
Rv3849	espR	14708.7	Cytoplasmic	ESX-1 transcriptional regulator	ESX-1 transcriptional regulatory protein EspR	Regulatory proteins	-

				EspR			
Rv3850	-	23810.9	Cytoplasmic	hypothetical protein Rv3850	Conserved protein	conserved hypotheticals	-
Rv3852	hns	13823.1	Unknown	histone-like protein Hns	Possible histone-like protein Hns	information pathways	-
Rv3853	rraA	16203.4	Unknown	RNase E regulator RraA	Regulator of RNase E activity a RraA	Regulatory proteins	COG0684H
Rv3854c	ethA	55326.1	Cytoplasmic	monooxygenase EthA	Monooxygenase EthA	intermediary metabolism and respiration	COG2072P
Rv3855	ethR	23724.7	Cytoplasmic	HTH-type transcriptional repressor EthR	Transcriptional regulatory repressor protein (TetR-family) EthR	Regulatory proteins	COG1309K
Rv3858c	gltD	53451.8	Cytoplasmic	glutamate synthase small subunit	Probable NADH-dependent glutamate synthase (small subunit) GltD (L-glutamate synthase) (L-glutamate synthetase) (NADH-glutamate synthase) (glutamate synthase (NADH)) (GLTS beta chain) (NADPH-GOGAT)	intermediary metabolism and respiration	COG0493ER
Rv3859c	gltB	165943	Cytoplasmic Membrane	glutamate synthase large subunit	Probable ferredoxin-dependent glutamate synthase [NADPH] (large subunit) GltB (L-glutamate synthase) (L-glutamate synthetase) (NADH-glutamate synthase) (glutamate synthase (NADH))(NADPH-GOGAT)	intermediary metabolism and respiration	COG0069E
Rv3860	-	42133.9	Unknown	hypothetical protein Rv3860	ATPases involved in chromosome partitioning	cell wall and cell processes	COG0455D
Rv3864	espE	42068.8	Cytoplasmic	ESX-1 secretion-associated protein EspE	ESX-1 secretion-associated protein EspE	cell wall and cell processes	-
Rv3865	espF	10618	Unknown	ESX-1 secretion-associated protein EspF	ESX-1 secretion-associated protein EspF	cell wall and cell processes	-
Rv3866	espG1	30064.1	Cytoplasmic	ESX-1 secretion-associated protein EspG	ESX-1 secretion-associated protein EspG1	cell wall and cell processes	COG4170V
Rv3867	espH	19945.6	Cytoplasmic	ESX-1 secretion-associated protein EspH	ESX-1 secretion-associated protein EspH	cell wall and cell processes	-
Rv3868	eccA1	62425.6	Cytoplasmic	ESX-1 secretion system protein EccA1	ESX conserved component EccA1. ESX-1 type VII secretion system protein.	cell wall and cell processes	COG0464O
Rv3869	eccB1	51092.7	Cytoplasmic Membrane	ESX-1 secretion system protein EccB	ESX conserved component EccB1. ESX-1 type VII secretion system protein. Possible membrane protein.	cell wall and cell processes	-
Rv3870	eccCa 1	80913	Cytoplasmic Membrane	ESX-1 secretion system protein EccCa	ESX conserved component EccCa1. ESX-1 type VII secretion system protein. Possible transmembrane protein.	cell wall and cell processes	COG1674D
Rv3871	eccCb 1	64560.8	Cytoplasmic	ESX-1 secretion system protein EccCb	ESX conserved component EccCb1. ESX-1 type VII secretion system protein.	cell wall and cell processes	COG1674D
Rv3873	PPE68	37330	Cytoplasmic Membrane	PPE family protein PPE68	PPE family protein PPE68	PE/PPE	COG5651N
Rv3874	esxB	10793.8	Extracellular	ESAT-6-like protein EsxB	10 kDa culture filtrate antigen EsxB (LHP) (CFP10)	cell wall and cell processes	COG4842S
Rv3875	esxA	9903.87	Extracellular	ESAT-6 protein EsxA	6 kDa early secretory antigenic target EsxA (ESAT-6)	cell wall and cell processes	COG4842S
Rv3876	espI	70645.1	Cytoplasmic	ESX-1 secretion-associated protein EspI	ESX-1 secretion-associated protein EspI. Conserved proline and alanine rich protein.	cell wall and cell processes	COG0455D

Rv3877	eccD1	53981.3	Cytoplasmic Membrane	ESX-1 secretion system protein EccD1	ESX conserved component EccD1. ESX-1 type VII secretion system protein. Probable transmembrane protein.	cell wall and cell processes	-
Rv3879c	espK	74492.3	Unknown	ESX-1 secretion-associated protein EspK	ESX-1 secretion-associated protein EspK. Alanine and proline rich protein.	cell wall and cell processes	COG0810M
Rv3880c	espL	12167.5	Unknown	ESX-1 secretion-associated protein EspL	ESX-1 secretion-associated protein EspL	cell wall and cell processes	-
Rv3881c	espB	47593.8	Cytoplasmic	ESX-1 secretion-associated protein EspB	Secreted ESX-1 substrate protein B, EspB. Conserved alanine and glycine rich protein	cell wall and cell processes	COG0038P
Rv3882c	eccE1	50396.7	Cytoplasmic	ESX-1 secretion system protein EccE1	ESX conserved component EccE1. ESX-1 type VII secretion system protein. Possible membrane protein.	cell wall and cell processes	-
Rv3883c	mycP1	45086	Extracellular	membrane-anchored mycosin	Membrane-anchored mycosin MycP1 (serine protease) (subtilisin-like protease) (subtilase-like) (mycosin-1)	intermediary metabolism and respiration	COG1404O
Rv3888c	-	37691.8	Cytoplasmic	membrane protein	Probable conserved membrane protein	cell wall and cell processes	COG0455D
Rv3894c	eccC2	153700	Cytoplasmic Membrane	ESX-2 type VII secretion system protein EccC	ESX conserved component EccC2. ESX-2 type VII secretion system protein. Possible membrane protein.	cell wall and cell processes	COG1674D
Rv3902c	-	19790.3	Unknown	hypothetical protein Rv3902c	Hypothetical protein	conserved hypotheticals	-
Rv3907c	pcnA	53055.4	Cytoplasmic	poly(A) polymerase PcnA	Probable poly(A) polymerase PcnA (polynucleotide adenylyltransferase) (NTP polymerase) (RNA adenylyating enzyme) (poly(A) polymerase)	information pathways	COG0617J
Rv3909	-	83877.4	Cytoplasmic Membrane	hypothetical protein Rv3909	cell wall	cell wall and cell processes	-
Rv3910	-	123564	Cytoplasmic Membrane	peptidoglycan biosynthesis protein	Probable conserved transmembrane protein	cell wall and cell processes	COG0728R
Rv3913	trxB2	35643	Cytoplasmic	thioredoxin reductase	Probable thioredoxin reductase TrxB2 (TRXR) (TR)	intermediary metabolism and respiration	COG0492O
Rv3914	trxC	12544.5	Cytoplasmic	thioredoxin TrxC	Thioredoxin TrxC (TRX) (MPT46)	intermediary metabolism and respiration	COG3118O
Rv3915	-	43911.5	Cellwall	peptidoglycan hydrolase	Probable peptidoglycan hydrolase	intermediary metabolism and respiration	COG0860M
Rv3917c	parB	37018.3	Cytoplasmic	chromosome partitioning protein ParB	Probable chromosome partitioning protein ParB	cell wall and cell processes	COG1475K
Rv3918c	parA	37516.8	Cytoplasmic Membrane	chromosome partitioning protein ParA	Probable chromosome partitioning protein ParA	cell wall and cell processes	COG1192D
Rv3920c	-	20558.6	Cytoplasmic	hypothetical protein Rv3920c	Predicted RNA-binding protein	Regulatory proteins	COG1847R
Rv3921c	-	40884.4	Cytoplasmic Membrane	membrane protein insertase YidC	Probable conserved transmembrane protein	cell wall and cell processes	COG0706U

ПРИЛОЖЕНИЕ 5. Дифференциальные белки для *M. tuberculosis* кластера Beijing B0/W148.

Таблица 5.1.

Высокопредставленные белки в штаммах кластера Beijing B0/W148 по сравнению с H37Rv

PID	Operon	Synonym	Gene	Accession	Annotation TubercuList	Functional category	Final localization	COG	p-value	log2 fold-change	Genomic differences in Beijing
15607527		Rv0386	-	gi 15607527	Probable transcriptional regulatory protein (probably LuxR/UhpA-family)	Regulatory proteins	Cytoplasmic	COG3903R	-	-	sSNP
15608085	184	Rv0945	-	gi 15608085	Probable short-chain type dehydrogenase/reductase	intermediary metabolism and respiration	Cytoplasmic	COG0300R	-	-	Arg180Gly
15608098	188	Rv0958	-	gi 15608098	Possible magnesium chelatase	intermediary metabolism and respiration	Cytoplasmic	COG1239H	0,015747	3,03484622	Ser274Pro
15608116	190	Rv0976c	-	gi 15608116	related to synthesis and oxidation of fatty acids	Lipid metabolism	Cytoplasmic	-	-	-	Upstream region (1090189insG)
15608459	252	Rv1319c	-	gi 15608459	Possible adenylate cyclase (ATP pyrophosphate-lyase) (adenyl cyclase)	intermediary metabolism and respiration	Cytoplasmic Membrane	COG2114T	-	-	Gly106Asp
15608569	278	Rv1431	-	gi 15608569	Conserved membrane protein	cell wall and cell processes	Cytoplasmic Membrane	-	0,029048	3,71365928	Asn65Thr
15609161		Rv2024c	-	gi 15609161	Predicted helicase	Information pathways	Cytoplasmic	COG4889R	0,04164	4,37188561	Trp47Arg; Asp154G
15609414		Rv2277c	-	gi 15609414	Possible glycerolphosphodiesterase	intermediary metabolism and respiration	Unknown	COG0584C	-	-	IS6110 is near the start of the gene
15609612	473	Rv2475c	-	gi 15609612	Predicted thioesterase	Intermediary metabolism and respiration	Cytoplasmic	COG0824R	-	-	IS6110 is near the start of the gene
15609963		Rv2826c	-	gi 15609963	Hypothetical protein	conserved hypotheticals	Cytoplasmic	COG2253S	-	-	without changes
15610093		Rv2956	-	gi 15610093	SAM-dependent methyltransferases	Intermediary metabolism and respiration	Cytoplasmic	-	0,035516	2,47739948	without changes
15610108	571	Rv2971	-	gi 15610108	Probable oxidoreductase	intermediary metabolism and respiration	Cytoplasmic	COG0656R	0,041065	0,49555832	Asn152His
15610111	572	Rv2974c	-	gi 15610111	Predicted kinase related to dihydroxyacetone kinase	Intermediary metabolism and respiration	Cytoplasmic	COG1461R	-	-	Upstream region (3101512delCG)

15610231	592	Rv3094c	-	gi 15610231	Acyl-CoA dehydrogenases	Intermediary metabolism and respiration	Cytoplasmic	COG1960I	0,038694	9,61066369	without changes
15610336		Rv3200c	-	gi 15610336	Possible transmembrane cation transporter	cell wall and cell processes	Cytoplasmic Membrane	COG0569P	0,038591	1,42894962	without changes
15610349		Rv3213c	-	gi 15610349	Possible SOJ/para-related protein	cell wall and cell processes	Cytoplasmic Membrane	COG1192D	0,046739	0,78724527	Lys144Glu
15610364	616	Rv3228	-	gi 15610364	Predicted GTPases	Intermediary metabolism and respiration	Cytoplasmic	COG1162R	-	-	sSNP
448824809		Rv3329	-	gi 448824809	Probable aminotransferase	intermediary metabolism and respiration	Cytoplasmic	-	0,006929	1,0515286	Gln122His
57117112		Rv3479	-	gi 57117112	Possible transmembrane protein	cell wall and cell processes	Cytoplasmic Membrane	COG1752R	0,005988	3,42337554	Leu174Arg
15610728	701	Rv3592	-	gi 15610728	Possible heme degrading protein MhuD	intermediary metabolism and respiration	Cytoplasmic	COG2329R	-	-	without changes
448824819		Rv3753c	-	gi 448824819	Conserved protein	conserved hypotheticals	Cytoplasmic	-	-	-	without changes
15610965	741	Rv3829c	-	gi 15610965	Probable dehydrogenase	intermediary metabolism and respiration	Cytoplasmic	COG1233Q	0,007094	2,05688191	another annotation in Beijing strains
15610244		Rv3107c	agpS	gi 15610244	Possible alkylidihydroxyacetonephosphate synthase AgpS (alkyl-DHAP synthase) (alkylglycerone-phosphate synthase)	lipid metabolism	Cytoplasmic	COG0277C	0,022586	3,50212982	without changes
15608850	335	Rv1712	cmk	gi 15608850	Cytidylate kinase Cmk (CMP kinase) (cytidine monophosphate kinase) (ck)	intermediary metabolism and respiration	Cytoplasmic	COG0283F	-	-	without changes
15608602	283	Rv1464	csd	gi 15608602	Probable cysteine desulfurase Csd	intermediary metabolism and respiration	Cytoplasmic	COG0520E	0,048907	1,55571668	without changes
15608935	351	Rv1798	eccA5	gi 15608935	ESX conserved component EccA5. ESX-5 type VII secretion system protein.	cell wall and cell processes	Cytoplasmic	COG1223R	0,04164	1,76461225	without changes
15609324		Rv2187	fadD15	gi 15609324	Long-chain-fatty-acid-CoA ligase FadD15 (fatty-acid-CoA synthetase) (fatty-acid-CoA synthase)	lipid metabolism	Cytoplasmic	COG1022I	0,049359	0,63987833	Thr100Ile
15610243		Rv3106	fprA	gi 15610243	NADPH:adrenodoxin oxidoreductase FprA (NADPH-ferredoxin reductase)	intermediary metabolism and respiration	Cytoplasmic	COG0493ER	0,021691	1,0945977	without changes
57116925		Rv1844c	gnd1	gi 57116925	Probable 6-phosphogluconate dehydrogenase Gnd1	intermediary metabolism and	Cytoplasmic	COG0362G	0,02211	1,61085849	Rv1844: 2094915 del CGTCAG

						respiration					
15607650	93	Rv0509	hemA	gi 15607650	Probable glutamyl-tRNA reductase HemA (GLUTR)	intermediary metabolism and respiration	Cytoplasmic	COG0373H	-	-	without changes
15610706	696	Rv3570c	hsaA	gi 15610706	Possible oxidoreductase. Possible 3-hydroxy-9,10-seconandrost-1,3,5(10)-triene-9,17-dione hydroxylase.	intermediary metabolism and respiration	Cytoplasmic Membrane	COG1960I	0,043815	1,00243225	without changes
15610630	679	Rv3494c	mce4F	gi 15610630	Mce-family protein Mce4F	virulence, detoxification, adaptation	Unknown	COG1463Q	0,025993	1,46143783	without changes
57117090	640	Rv3323c	moaX	gi 57117090	Probable MoaD-MoaE fusion protein MoaX	intermediary metabolism and respiration	Cytoplasmic	COG0314H	-	-	IS6110 is near the start of the gene
15607267		Rv0125	pepA	gi 15607267	Probable serine protease PepA (serine proteinase) (MTB32A)	intermediary metabolism and respiration	Unknown	COG0265O	-	-	without changes
15607551	73	Rv0410c	pknG	gi 15607551	Serine/threonine-protein kinase PknG (protein kinase G) (STPK G)	Regulatory proteins	Cytoplasmic Membrane	COG0515RTKL	0,022586	0,93897666	without changes
15609488		Rv2351c	plcA	gi 15609488	Membrane-associated phospholipase C 1 PlcA (MTP40 antigen)	intermediary metabolism and respiration	Cytoplasmic Membrane	COG3511M	-	-	Thr446Ala
15610393	623	Rv3257c	pmmA	gi 15610393	Probable phosphomannomutase PmmA (PMM) (phosphomannose mutase)	intermediary metabolism and respiration	Cytoplasmic	COG1109G	0,028727	1,71695929	Ser206Leu
15608520	267	Rv1380	pyrB	gi 15608520	Probable aspartate carbamoyltransferase PyrB (ATCase) (aspartate transcarbamylase)	intermediary metabolism and respiration	Cytoplasmic	COG0540F	0,034757	1,68716099	without changes
15610370	619	Rv3234c	tgs3	gi 15610370	Putative triacylglycerol synthase (diacylglycerol acyltransferase) Tgs3	lipid metabolism	Cytoplasmic	-	0,007094	2,5875231	Rv3234c: 3610391 insC
15610026	555	Rv2889c	tsf	gi 15610026	Probable elongation factor Tsf (EF-ts)	information pathways	Cytoplasmic	COG0264J	0,01944	1,10514952	without changes
15608535	270	Rv1397c	vapC10	gi 15608535	Possible toxin VapC10	virulence, detoxification, adaptation	Unknown	COG1487R	-	-	without changes

Низкопредставленные белки в штаммах кластера Beijing B0/W148 по сравнению с H37Rv

PID	Operon	Synonym	Gene	Accession	Annotation TubercuList	Functional category	Final localization	COG	p-value	log2 fold-change	Genomic differences in Beijing
2E+07	7	Rv0034	-	gi 15607176	Ketosteroid isomerase-related protein	Intermediary metabolism and respiration	Cytoplasmic	COG3631R	-	-	sSNP
6E+07	7	Rv0035	fadD34	gi 57116683	Probable fatty-acid-CoA ligase FadD34 (fatty-acid-CoA synthetase) (fatty-acid-CoA synthase)	lipid metabolism	Cytoplasmic Membrane	COG0318IQ	-	-	Ser16Trp; Thr26Ser
2E+07		Rv0045c	-	gi 15607187	Possible hydrolase	lipid metabolism	Cytoplasmic	COG0596R	-	-	Val194Ile
2E+07	10	Rv0046c	ino1	gi 15607188	myo-inositol-1-phosphate synthase Ino1 (inositol 1-phosphate synthetase) (D-glucose 6-phosphate cycloaldolase) (glucose 6-phosphate cyclase) (glucocycloaldolase)	intermediary metabolism and respiration	Cytoplasmic	COG1260I	0,026787	-0,86811	Arg190Gly
2E+07	16	Rv0072	-	gi 15607214	Probable glutamine-transport transmembrane protein ABC transporter	cell wall and cell processes	Cytoplasmic Membrane	COG4591M	-	-	part of RD105
2E+07	16	Rv0073	-	gi 15607215	Probable glutamine-transport ATP-binding protein ABC transporter	cell wall and cell processes	Cytoplasmic Membrane	COG1136V	0,000362	0,75926	without changes
2E+07	18	Rv0081	-	gi 15607223	Probable transcriptional regulatory protein	Regulatory proteins	Cytoplasmic	COG0640K	-	-	without changes
2E+07	18	Rv0087	hycE	gi 15607229	Possible formate hydrogenase HycE (FHL)	intermediary metabolism and respiration	Cytoplasmic	COG3261C	-	-	without changes
2E+07		Rv0088	-	gi 15607230	Possible polyketide cyclase/dehydrase	lipid metabolism	Unknown	COG2867I	0,037668	-0,86647	sSNP
2E+07	19	Rv0101	nrp	gi 15607243	Probable peptide synthetase Nrp (peptide synthase)	lipid metabolism	Cytoplasmic	COG1020Q	0,00007	0,98446	sSNP
2E+07		Rv0119	fadD7	gi 15607261	Probable fatty-acid-CoA ligase FadD7 (fatty-acid-CoA synthetase) (fatty-acid-CoA synthase)	lipid metabolism	Cytoplasmic	COG0318IQ	4,62E-10	-1,95912	without changes
2E+07		Rv0143c	-	gi 15607285	Probable conserved transmembrane protein	cell wall and cell processes	Cytoplasmic Membrane	COG0038P	-	-	without changes
2E+07	27	Rv0154c	fadE2	gi 15607296	Probable acyl-CoA dehydrogenase FadE2	lipid metabolism	Cytoplasmic	COG1960I	0,000806	-0,59373	sSNP
6E+07		Rv0164	TB18.5	gi 57116699	Cyclase/Dehydrase	Intermediary metabolism and	Cytoplasmic	COG2867I	0,02162	-1,1725	without changes

						respiration					
6E+07	29	Rv0169	mce1A	gi 57116701	Mce-family protein Mce1A	virulence, detoxification, adaptation	Cytoplasmic Membrane	COG1463Q	5,76E-06	-1,44113	Ser313Ala, TΦ, TSS
2E+07	29	Rv0170	mce1B	gi 15607311	Mce-family protein Mce1B	virulence, detoxification, adaptation	Unknown	COG1463Q	0,001715	-1,37691	without changes
2E+07	29	Rv0172	mce1D	gi 15607313	Mce-family protein Mce1D	virulence, detoxification, adaptation	Unknown	COG1463Q	0,000154	-1,33027	Ile188Thr
2E+07	29	Rv0173	lprK	gi 15607314	Possible Mce-family lipoprotein LprK (Mce-family lipoprotein Mce1E)	cell wall and cell processes	Cytoplasmic Membrane	COG1463Q	0,000101	-1,36906	sSNP
2E+07	29	Rv0174	mce1F	gi 15607315	Mce-family protein Mce1F	virulence, detoxification, adaptation	Unknown	COG1463Q	7,20E-05	-0,84091	Leu370Pro
2E+07	29	Rv0176	-	gi 15607317	Probable conserved Mce associated transmembrane protein	cell wall and cell processes	Cytoplasmic Membrane	COG1714S	0,017555	-1,80599	without changes
2E+07		Rv0180c	-	gi 15607321	Probable conserved transmembrane protein	cell wall and cell processes	Cytoplasmic Membrane	COG1511S	0,001535	-1,10742	without changes
2E+07		Rv0182c	sigG	gi 15607323	Probable alternative RNA polymerase sigma factor SigG (RNA polymerase ECF type sigma factor)	information pathways	Cytoplasmic	COG1595K	-	-	Asp332Tyr; sSNP
2E+07	43	Rv0241c	htdX	gi 15607382	Probable 3-hydroxyacyl-thioester dehydratase HtdX	intermediary metabolism and respiration	Cytoplasmic	COG2030I	0,038694	-0,86377	without changes
2E+07		Rv0243	fadA2	gi 15607384	Probable acetyl-CoA acyltransferase FadA2 (3-ketoacyl-CoA thiolase) (beta-ketothiolase)	lipid metabolism	Cytoplasmic	COG0183I	0,016661	-0,91371	without changes
2E+07	44	Rv0248c	-	gi 15607389	Probable succinate dehydrogenase [iron-sulfur subunit] (succinic dehydrogenase)	intermediary metabolism and respiration	Cytoplasmic Membrane	COG1053C	0,001715	-0,63154	without changes
2E+07	51	Rv0282	eccA3	gi 15607423	ESX conserved component EccA3. ESX-3 type VII secretion system protein.	cell wall and cell processes	Cytoplasmic	COG0464O	1,67E-06	-1,46087	Glu6Ala
2E+07	51	Rv0283	eccB3	gi 15607424	ESX conserved component EccB3. ESX-3 type VII secretion system protein. Possible membrane protein.	cell wall and cell processes	Cytoplasmic Membrane	-	0,000183	-0,96747	without changes
2E+07	51	Rv0284	eccC3	gi 15607425	ESX conserved component EccC3. ESX-3 type VII secretion system protein. Possible membrane protein.	cell wall and cell processes	Cytoplasmic Membrane	COG1674D	0,000782	-1,05623	Pro214Arg
2E+07		Rv0311	-	gi 15607452	Unknown protein	conserved hypotheticals	Cytoplasmic	-	-	-	Glu119Asp

2E+07		Rv0315	-	gi 15607456	Possible beta-1,3-glucanase precursor	intermediary metabolism and respiration	Unknown	COG2273G	0,002776	-1,17799	Lys260Thr
2E+07	58	Rv0334	rmlA	gi 15607475	Alpha-D-glucose-1-phosphate thymidyltransferase RmlA (dTDP-glucose synthase) (dTDP-glucose pyrophosphorylase)	intermediary metabolism and respiration	Cytoplasmic	COG1209M	7,30E-05	-1,41106	without changes
2E+07		Rv0338c	-	gi 15607479	Probable iron-sulfur-binding reductase	intermediary metabolism and respiration	Cytoplasmic Membrane	COG0247C	0,001404	-1,00554	Arg506Gly; Glu571Gly; Ala621Val
2E+07	61	Rv0350	dnaK	gi 15607491	Probable chaperone protein DnaK (heat shock protein 70) (heat shock 70 kDa protein) (HSP70)	virulence, detoxification, adaptation	Cytoplasmic	COG0443O	0,001624	-0,44943	sSNP
2E+07	71	Rv0405	pks6	gi 15607546	Probable membrane bound polyketide synthase Pks6	lipid metabolism	Cytoplasmic Membrane	COG3321Q	-	-	Arg1402Pro
2E+07	74	Rv0417	thiG	gi 15607558	Probable thiamin biosynthesis protein ThiG (thiazole biosynthesis protein)	intermediary metabolism and respiration	Cytoplasmic	COG2022H	-	-	Ser75Cys
2E+07		Rv0419	lpqM	gi 15607560	Possible lipoprotein peptidase LpqM	cell wall and cell processes	Cytoplasmic Membrane	COG2321R	0,04164	-1,14388	Ala297Thr
2E+07	78	Rv0433	-	gi 15607574	Conserved hypothetical protein	conserved hypotheticals	Cytoplasmic	COG2170S	0,024943	-0,66706	without changes
2E+07		Rv0439c	-	gi 15607580	Probable dehydrogenase/reductase	intermediary metabolism and respiration	Unknown	COG1028IQR	7,30E-05	-2,00839	sSNP
2E+07		Rv0445c	sigK	gi 15607586	Alternative RNA polymerase sigma factor SigK	information pathways	Cytoplasmic Membrane	COG1595K	0,002695	-1,92304	without changes
2E+07		Rv0468	fadB2	gi 15607609	3-hydroxybutyryl-CoA dehydrogenase FadB2 (beta-hydroxybutyryl-CoA dehydrogenase) (BHBD)	lipid metabolism	Cytoplasmic	COG1250I	2,15E-07	-1,33024	without changes
6E+07		Rv0469	umaA	gi 57116735	Possible mycolic acid synthase UmaA	lipid metabolism	Cytoplasmic	COG2230M	0,007094	-0,53196	without changes
2E+07	86	Rv0479c	-	gi 15607620	Probable conserved membrane protein	cell wall and cell processes	Unknown	-	0,007095	-0,70188	without changes
2E+07		Rv0495c	-	gi 15607636	DNA-associated process	Regulatory proteins	Cytoplasmic	-	-	-	without changes
2E+07		Rv0500	proC	gi 15607641	Probable pyrroline-5-carboxylate reductase ProC (P5CR) (P5C reductase)	intermediary metabolism and respiration	Cytoplasmic	COG0345E	0,009979	-1,2638	sSNP
2E+07		Rv0562	grcC1	gi 15607702	Probable polyprenyl-diphosphate synthase GrcC1 (polyprenyl pyrophosphate synthetase)	intermediary metabolism and respiration	Cytoplasmic	COG0142H	0,022246	-0,89802	without changes
2E+07		Rv0604	lpqO	gi 15607744	Probable conserved lipoprotein LpqO	cell wall and cell	Cytoplasmic	-	-	-	without

						processes	Membrane				changes
2E+07	120	Rv0651	rplJ	gi 15607791	50S ribosomal protein L10 RplJ	information pathways	Cytoplasmic	COG0244J	0,005988	-1,32699	without changes
2E+07	124	Rv0667	rpoB	gi 15607807	DNA-directed RNA polymerase (beta chain) RpoB (transcriptase beta chain) (RNA polymerase beta subunit)	information pathways	Cytoplasmic	COG0085K	0,001404	-0,6019	Ser450Leu
6E+07	125	Rv0675	echA5	gi 57116767	Probable enoyl-CoA hydratase EchA5 (enoyl hydratase) (unsaturated acyl-CoA hydratase) (crotonase)	lipid metabolism	Cytoplasmic	COG1024I	0,00076	-1,0963	without changes
2E+07	128	Rv0683	rpsG	gi 15607823	30S ribosomal protein S7 RpsG	information pathways	Cytoplasmic	COG0049J	0,017784	-1,29281	without changes
2E+07	133	Rv0702	rplD	gi 15607842	50S ribosomal protein L4 RplD	information pathways	Cytoplasmic	COG0088J	0,003053	-0,93935	without changes
2E+07		Rv0730	-	gi 15607870	GCN5-related N-acetyltransferase	intermediary metabolism and respiration	Cytoplasmic	COG0188L	0,033976	-0,49701	without changes
2E+07		Rv0731c	-	gi 15607871	Possible S-adenosylmethionine-dependent methyltransferase	lipid metabolism	Cytoplasmic	COG3315Q	0,002576	-0,90112	without changes
2E+07	145	Rv0768	aldA	gi 15607908	Probable aldehyde dehydrogenase NAD dependent AldA (aldehyde dehydrogenase [NAD+])	intermediary metabolism and respiration	Cytoplasmic	COG1012C	-	-	Leu102Arg
6E+07		Rv0824c	desA1	gi 57116786	Probable acyl-[acyl-carrier protein] desaturase DesA1 (acyl-[ACP] desaturase) (stearoyl-ACP desaturase) (protein Des)	lipid metabolism	Cytoplasmic	-	5,76E-06	-1,97695	without changes
2E+07		Rv0825c	-	gi 15607965	Transcriptional regulator	Regulatory proteins	Unknown	COG1309K	-	-	Phe68Cys
2E+07		Rv0844c	narL	gi 15607984	Possible nitrate/nitrite response transcriptional regulatory protein NarL	Regulatory proteins	Cytoplasmic	COG2197TK	-	-	Gly169Arg
2E+07	163	Rv0859	-	gi 15607999	Possible acyl-CoA thiolase FadA	lipid metabolism	Cytoplasmic	COG0183I	0,017555	-0,55642	Ser150Gly
2E+07	163	Rv0860	fadB	gi 15608000	Probable fatty oxidation protein FadB	lipid metabolism	Cytoplasmic	COG1250I	0,008655	-0,47085	sSNP
2E+07		Rv0884c	serC	gi 15608024	Possible phosphoserine aminotransferase SerC (PSAT)	intermediary metabolism and respiration	Unknown	COG1932HE	0,007831	-0,91074	without changes
2E+07		Rv0888	-	gi 15608028	Probable exported protein	cell wall and cell processes	Unknown	COG3021S	-	-	Rv0888: 987586 insGG
2E+07	172	Rv0905	echA6	gi 15608045	Possible enoyl-CoA hydratase EchA6 (enoyl hydratase) (unsaturated acyl-CoA hydratase) (crotonase)	lipid metabolism	Cytoplasmic	COG1024I	0,002615	-0,96051	without changes
2E+07	173	Rv0908	ctpE	gi 15608048	Probable metal cation transporter ATPase P-type CtpE	cell wall and cell processes	Cytoplasmic Membrane	COG0474P	-	-	without changes

2E+07	179	Rv0927c	-	gi 15608067	Probable short-chain type dehydrogenase/reductase	intermediary metabolism and respiration	Cytoplasmic	COG1028IQR	3,69E-05	-2,86967	Rv0927c: 1034211 del TGC
2E+07	182	Rv0933	pstB	gi 15608073	Phosphate-transport ATP-binding protein ABC transporter PstB	cell wall and cell processes	Cytoplasmic Membrane	COG1117P	4,65E-07	-2,40935	without changes
6E+07	182	Rv0934	pstS1	gi 57116801	Periplasmic phosphate-binding lipoprotein PstS1 (PBP-1) (PstS1)	cell wall and cell processes	Unknown	COG0226P	0,001559	-1,11713	without changes
4E+08	191	Rv0981	mprA	gi 448824752	Mycobacterial persistence regulator MRPA (two component response transcriptional regulatory protein)	Regulatory proteins	Cytoplasmic	-	0,001451	-1,34097	without changes
2E+07	191	Rv0984	moaB2	gi 15608124	Possible pterin-4-alpha-carbinolamine dehydratase MoaB2 (PHS) (4-alpha-hydroxy-tetrahydropterin dehydratase) (pterin-4-a-carbinolamine dehydratase) (phenylalanine hydroxylase-stimulating protein) (PHS) (pterin carbinolamine dehydratase) (PCD)	intermediary metabolism and respiration	Cytoplasmic	COG0521H	0,005587	-1,16	without changes
2E+07	192	Rv0987	-	gi 15608127	Probable adhesion component transport transmembrane protein ABC transporter	cell wall and cell processes	Cytoplasmic Membrane	COG4591M	-	-	Ala372Val; Phe717Val
2E+07	192	Rv0988	-	gi 15608128	Possible conserved exported protein	cell wall and cell processes	Cytoplasmic Membrane	COG5621R	-	-	Leu191Ala
2E+07		Rv0989c	grcC2	gi 15608129	Probable polyprenyl-diphosphate synthase GrcC2 (polyprenyl pyrophosphate synthetase)	intermediary metabolism and respiration	Cytoplasmic	COG0142H	-	-	Ile321Val
2E+07	195	Rv1015c	rplY	gi 15608155	50S ribosomal protein L25 RplY	information pathways	Cytoplasmic	COG1825J	0,001693	-0,98111	without changes
2E+07	197	Rv1023	eno	gi 15608163	Probable enolase Eno	intermediary metabolism and respiration	Cytoplasmic	COG0148G	1,08E-06	-1,41623	Arg8Gly; Lys429Gln
2E+07	206	Rv1071c	echA9	gi 15608211	Possible enoyl-CoA hydratase EchA9 (enoyl hydrase) (unsaturated acyl-CoA hydratase) (crotonase)	lipid metabolism	Cytoplasmic	COG1024I	0,002788	-1,26081	without changes
2E+07		Rv1078	-	gi 15608218	cell wall and cell processes , Predicted membrane protein/domain, proline-rich antigen homolog	cell wall and cell processes	Cytoplasmic Membrane	COG5178A	0,015898	-0,95889	sSNP
2E+07		Rv1094	desA2	gi 15608234	Possible acyl-[acyl-carrier protein] desaturase DesA2 (acyl-[ACP] desaturase) (stearoyl-ACP desaturase)	lipid metabolism	Cytoplasmic	COG0208F	5,20E-09	-1,91831	without changes
2E+07		Rv1095	phoH2	gi 15608235	Probable PHOH-like protein PhoH2 (phosphate starvation-inducible protein PSIH)	intermediary metabolism and respiration	Cytoplasmic	COG1875T	0,000101	-3,39899	without changes

4E+08		Rv1099c	glpX	gi 448824754	Fructose 1,6-bisphosphatase GlpX	intermediary metabolism and respiration	Cytoplasmic	-	0,000782	-0,66491	sSNP
2E+07		Rv1129c	-	gi 15608269	Probable transcriptional regulator protein	Regulatory proteins	Cytoplasmic	COG3800R	0,007142	-1,35114	TΦ
2E+07	220	Rv1144	-	gi 15608284	Probable short-chain type dehydrogenase/reductase	intermediary metabolism and respiration	Cytoplasmic	COG1028IQR	0,046988	-1,08539	without changes
2E+07		Rv1165	typA	gi 15608305	Possible GTP-binding translation elongation factor TypA (tyrosine phosphorylated protein A) (GTP-binding protein)	information pathways	Cytoplasmic	COG1217T	0,049764	-0,72774	without changes
2E+07		Rv1179c	-	gi 15608319	DNA or RNA helicases of superfamily II	Information pathways	Cytoplasmic Membrane	COG1061KL	0,018225	-0,87501	without changes
2E+07	226	Rv1181	pks4	gi 15608321	Probable polyketide beta-ketoacyl synthase Pks4	lipid metabolism	Cytoplasmic Membrane	COG3321Q	0,006874	-1,28729	sSNP
2E+07	226	Rv1182	papA3	gi 15608322	Probable conserved polyketide synthase associated protein PapA3	lipid metabolism	Cytoplasmic	-	-	-	without changes
2E+07	226	Rv1183	mmpL10	gi 15608323	Probable conserved transmembrane transport protein MmpL10	cell wall and cell processes	Cytoplasmic Membrane	COG2409R	-	-	Thr408Arg
2E+07		Rv1185c	fadD21	gi 15608325	Probable fatty-acid-AMP ligase FadD21 (fatty-acid-AMP synthetase) (fatty-acid-AMP synthase)	lipid metabolism	Cytoplasmic	COG0318IQ	0,004873	-1,34882	without changes
2E+07		Rv1192	-	gi 15608332	Protein involved in GPI anchor metabolic process	cell wall and cell processes	Cytoplasmic	COG0596R	0,000633	-1,91451	without changes
2E+07		Rv1194c	-	gi 15608334	Regulator of polyketide synthase expression	Lipid metabolism	Cytoplasmic	COG2508TQ	0,000217	-1,5581	without changes
6E+07		Rv1196	PPE18	gi 57116841	PPE family protein PPE18	PE/PPE	Cytoplasmic Membrane	COG5651N	1,27E-11	-6,47296	Arg287Gln, TSS
2E+07		Rv1220c	-	gi 15608360	Probable methyltransferase	intermediary metabolism and respiration	Cytoplasmic	COG4122R	0,048818	-1,1874	without changes
2E+07		Rv1240	mdh	gi 15608380	Probable malate dehydrogenase Mdh	intermediary metabolism and respiration	Unknown	COG0039C	8,85E-06	-1,36533	without changes
2E+07	246	Rv1281c	oppD	gi 15608421	Probable oligopeptide-transport ATP-binding protein ABC transporter OppD	cell wall and cell processes	Cytoplasmic Membrane	COG0444EP	0,045018	-0,59734	without changes
2E+07	248	Rv1292	argS	gi 15608432	Probable arginyl-tRNA synthetase ArgS (ARGRS) (arginine--tRNA ligase)	information pathways	Cytoplasmic	COG0018J	0,043815	-0,64607	Ala119Thr

2E+07	250	Rv1306	atpF	gi 15608446	Probable ATP synthase B chain AtpF	intermediary metabolism and respiration	Unknown	COG0711C	0,001044	-1,8072	without changes
2E+07	250	Rv1307	atpH	gi 15608447	Probable ATP synthase delta chain AtpH	intermediary metabolism and respiration	Cytoplasmic	COG0712C	0,000445	-0,88756	without changes
2E+07	250	Rv1308	atpA	gi 15608448	Probable ATP synthase alpha chain AtpA	intermediary metabolism and respiration	Cytoplasmic	COG0056C	0,00133	-0,46146	without changes
2E+07	250	Rv1309	atpG	gi 15608449	Probable ATP synthase gamma chain AtpG	intermediary metabolism and respiration	Cytoplasmic	COG0224C	0,015747	-0,73473	without changes
2E+07	250	Rv1310	atpD	gi 15608450	Probable ATP synthase beta chain AtpD	intermediary metabolism and respiration	Cytoplasmic Membrane	COG0055C	0,000609	-0,66026	without changes
2E+07	251	Rv1317c	alkA	gi 15608457	Probable bifunctional regulatory protein and DNA repair enzyme AlkA (regulatory protein of adaptative response) (methylphosphotriester-DNA--protein-cysteine S-methyltransferase)	information pathways	Cytoplasmic	COG0122L	-	-	Ile12Val
2E+07	263	Rv1363c	-	gi 15608503	Possible membrane protein	cell wall and cell processes	Cytoplasmic Membrane	-	0,000212	-3,30026	without changes
6E+07	267	Rv1384	carB	gi 57116856	Probable carbamoyl-phosphate synthase large chain CarB (carbamoyl-phosphate synthetase ammonia chain)	intermediary metabolism and respiration	Unknown	COG0458EF	0,027524	-0,81598	without changes
6E+07	268	Rv1386	PE15	gi 57116857	PE family protein PE15	PE/PPE	Cytoplasmic Membrane	-	-	-	without changes
2E+07		Rv1392	metK	gi 15608531	Probable S-adenosylmethionine synthetase MetK (mat) (AdoMet synthetase) (methionine adenosyltransferase)	intermediary metabolism and respiration	Cytoplasmic	COG0192H	0,039752	-0,33106	without changes
2E+07	271	Rv1400c	lipI	gi 15608538	Probable lipase LipH	intermediary metabolism and respiration	Cytoplasmic	COG0657I	-	-	Met42Val
2E+07	274	Rv1410c	-	gi 15608548	Aminoglycosides/tetracycline-transport integral membrane protein	cell wall and cell processes	Cytoplasmic Membrane	COG2814G	0,003882	-2,993	without changes
2E+07	280	Rv1437	pgk	gi 15608575	Probable phosphoglycerate kinase Pkg	intermediary metabolism and respiration	Cytoplasmic	COG0126G	0,002498	-0,92706	without changes
6E+07		Rv1489	-	gi 57116875	Conserved protein	conserved hypotheticals	Cytoplasmic Membrane	-	0,000437	-4,21534	Lys52Thr
2E+07		Rv1508c	-	gi 15608646	Probable membrane protein	cell wall and cell processes	Cytoplasmic Membrane	-	0,01028	-1,16187	Pro16Ser, TSS

2E+07		Rv1536	ileS	gi 15608674	Isoleucyl-tRNA synthetase IleS	information pathways	Cytoplasmic	COG0060J	0,000732	-1,45409	without changes
2E+07	301	Rv1543	-	gi 15608681	Possible fatty acyl-CoA reductase	lipid metabolism	Cytoplasmic	COG0300R	0,008639	-0,83522	without changes
2E+07	301	Rv1544	-	gi 15608682	Possible ketoacyl reductase	lipid metabolism	Cytoplasmic	COG0300R	0,002695	-1,23163	without changes
2E+07		Rv1576c	-	gi 15608714	Probable PhiRv1 phage protein	Insertion seqs and phages	Cytoplasmic Membrane	COG4653R	-	-	part of RD149
2E+07	307	Rv1586c	-	gi 15608724	Probable PhiRv1 integrase	Insertion seqs and phages	Cytoplasmic	COG1961L	-	-	part of RD149
2E+07	308	Rv1589	bioB	gi 15608727	Probable biotin synthetase BioB	intermediary metabolism and respiration	Cytoplasmic	COG0502H	-	-	without changes
2E+07	308	Rv1591	-	gi 15608729	Probable transmembrane protein	cell wall and cell processes	Cytoplasmic Membrane	-	-	-	without changes
2E+07		Rv1593c	-	gi 15608731	ADP-ribose pyrophosphatase	Intermediary metabolism and respiration	Cytoplasmic	COG1051F	0,020621	-1,01074	without changes
2E+07	309	Rv1596	nadC	gi 15608734	Probable nicotinate-nucleotide pyrophosphatase NadC	intermediary metabolism and respiration	Cytoplasmic	COG0157H	3,88E-05	-1,57016	without changes
2E+07	312	Rv1614	lgt	gi 15608752	Possible prolipoprotein diacylglycerol transferases Lgt	cell wall and cell processes	Cytoplasmic Membrane	COG0682M	0,000654	-1,53451	without changes
2E+07	314	Rv1617	pykA	gi 15608755	Probable pyruvate kinase PykA	intermediary metabolism and respiration	Cytoplasmic	COG0469G	0,013802	-0,59584	without changes
6E+07	315	Rv1623c	cydA	gi 57116893	Probable integral membrane cytochrome D ubiquinol oxidase (subunit I) CydA (cytochrome BD-I oxidase subunit I)	intermediary metabolism and respiration	Cytoplasmic Membrane	COG1271C	0,000595	-2,48477	without changes
2E+07		Rv1689	tyrS	gi 15608827	Probable tyrosyl-tRNA synthase TyrS (TYRRS)	information pathways	Cytoplasmic	COG0162J	0,028324	-2,34012	without changes
2E+07	340	Rv1729c	-	gi 15608867	Possible S-adenosylmethionine-dependent methyltransferase	lipid metabolism	Cytoplasmic	COG3315Q	-	-	without changes
2E+07	344	Rv1762c	-	gi 15608900	Unknown protein	conserved hypotheticals	Cytoplasmic	COG0393S	-	-	part of RD152
6E+07	352	Rv1808	PPE32	gi 57116921	PPE family protein PPE32	PE/PPE	Cytoplasmic Membrane	COG5651N	0,000954	-1,03628	Gly301Ala
2E+07	353	Rv1822	pgsA2	gi 15608959	Probable CDP-diacylglycerol--glycerol-3-phosphate 3-phosphatidyltransferase PgsA2 (PGP synthase) (phosphatidylglycerophosphate synthase) (3-phosphatidyl-1'-glycerol-3'phosphate	lipid metabolism	Cytoplasmic Membrane	COG0558I	-	-	without changes

					synthase)						
2E+07		Rv1832	gcvB	gi 15608969	Probable glycine dehydrogenase GcvB (glycine decarboxylase) (glycine cleavage system P-protein)	intermediary metabolism and respiration	Cytoplasmic	COG1003E	0,003864	-0,74702	sSNP
2E+07		Rv1835c	-	gi 15608972	Predicted acyl esterases	Intermediary metabolism and respiration	Cytoplasmic	COG2936R	-	-	Cys179Arg
2E+07		Rv1836c	-	gi 15608973	Von Willebrand factor (vWF) type A domain-containing protein	Regulatory proteins	Cytoplasmic Membrane	-	0,046988	-0,87507	without changes
2E+07		Rv1843c	guaB1	gi 15608980	Probable inosine-5'-monophosphate dehydrogenase GuaB1 (imp dehydrogenase) (IMPDH) (IMPD)	intermediary metabolism and respiration	Cytoplasmic	COG0516F	0,048987	-0,63838	without changes
2E+07	358	Rv1852	ureG	gi 15608989	Urease accessory protein UreG	intermediary metabolism and respiration	Unknown	COG0378OK	0,00133	-1,22355	without changes
2E+07	359	Rv1856c	-	gi 15608993	Possible oxidoreductase	intermediary metabolism and respiration	Cytoplasmic	COG1028IQR	0,037951	-1,05784	without changes
2E+07		Rv1872c	lldD2	gi 15609009	Possible L-lactate dehydrogenase (cytochrome) LldD2	intermediary metabolism and respiration	Cytoplasmic Membrane	COG1304C	9,56E-06	-1,10998	Gly59Ala; Val253Met, TΦ, TSS
2E+07		Rv1932	tpx	gi 15609069	Probable thiol peroxidase Tpx	virulence, detoxification, adaptation	Unknown	COG2077O	0,015898	-1,23196	without changes
2E+07		Rv1978	-	gi 15609115	SAM-dependent methyltransferases	Intermediary metabolism and respiration	Cytoplasmic Membrane	COG2226H	-	-	Ser226Thr
2E+07		Rv1997	ctpF	gi 15609134	Probable metal cation transporter P-type ATPase A CtpF	cell wall and cell processes	Cytoplasmic Membrane	COG0474P	-	-	Rv1997: 2241032 delG
2E+07		Rv2003c	-	gi 15609140	SAM-dependent methyltransferases	Intermediary metabolism and respiration	Unknown	COG2226H	-	-	without changes
2E+07		Rv2004c	-	gi 15609141	Predicted kinase	Regulatory proteins	Cytoplasmic	COG2187S	0,000437	-1,45085	without changes
2E+07		Rv2005c	-	gi 15609142	Universal stress protein family protein	virulence, detoxification, adaptation	Cytoplasmic	COG0589T	0,023446	-1,23836	without changes
2E+07	385	Rv2029c	pfkB	gi 15609166	6-phosphofructokinase PfkB (phosphohexokinase) (phosphofructokinase)	intermediary metabolism and respiration	Cytoplasmic Membrane	COG1105G	0,000633	-3,00266	without changes
4E+08	388	Rv2048c	pks12	gi 448824772	Polyketide synthase Pks12	lipid metabolism	Unknown	-	0,039341	-1,37853	Arg1652Cys;

											Pro3649Ala; Val3768Leu
2E+07		Rv2050	-	gi 15609187	Conserved protein	conserved hypotheticals	Cytoplasmic	-	-	-	without changes
2E+07	400	Rv2103c	vapC37	gi 15609240	Possible toxin VapC37. Contains PIN domain.	virulence, detoxification, adaptation	Unknown	COG1848R	-	-	Thr16Ala
2E+07		Rv2129c	-	gi 15609266	Probable oxidoreductase	intermediary metabolism and respiration	Cytoplasmic	COG4221R	0,032315	-0,79239	without changes
2E+07	411	Rv2159c	-	gi 15609296	Peroxidase	Intermediary metabolism and respiration	Cytoplasmic Membrane	COG0599S	7,02E-15	-4,28086	without changes
2E+07		Rv2161c	-	gi 15609298	Conserved protein	intermediary metabolism and respiration	Cytoplasmic	COG2141C	0,000408	-2,9799	without changes
2E+07		Rv2172c	-	gi 15609309	membrane associated process	cell wall and cell processes	Cytoplasmic	-	0,017555	-1,23852	without changes
2E+07		Rv2192c	trpD	gi 15609329	Probable anthranilate phosphoribosyltransferase TrpD	intermediary metabolism and respiration	Cytoplasmic Membrane	COG0547E	0,034757	-0,97865	without changes
2E+07		Rv2203	-	gi 15609340	Possible conserved membrane protein	cell wall and cell processes	Unknown	-	-	-	Upstream region (G2468180A)
2E+07	425	Rv2224c	caeA	gi 15609361	Probable carboxylesterase CaeA	cell wall and cell processes	Cytoplasmic	COG0596R	0,019823	-0,62883	without changes
4E+08		Rv2240c	-	gi 448824780	Unknown protein	conserved hypotheticals	Unknown	-	-	-	without changes
2E+07		Rv2263	-	gi 15609400	Possible oxidoreductase	intermediary metabolism and respiration	Cytoplasmic	COG1028IQR	-	-	part of RD181
2E+07		Rv2289	cdh	gi 15609426	Probable CDP-diacylglycerol pyrophosphatase Cdh (CDP-diacylglycerol diphosphatase) (CDP-diacylglycerol phosphatidylhydrolase)	lipid metabolism	Cytoplasmic Membrane	COG2134I	7,27E-05	-1,86557	without changes
4E+08	454	Rv2380c	mbtE	gi 448824784	Peptide synthetase MbtE (peptide synthase)	lipid metabolism	Cytoplasmic	-	-	-	sSNP
2E+07	457	Rv2394	ggtB	gi 15609531	Probable gamma-glutamyltranspeptidase precursor GgtB (gamma-glutamyltransferase) (glutamyl transpeptidase)	intermediary metabolism and respiration	Extracellular	COG0405E	0,000189	-2,69709	without changes
2E+07	460	Rv2410c	-	gi 15609547	proteasome	Intermediary metabolism and respiration	Cytoplasmic Membrane	COG2307S	0,0083	-1,42021	without changes

2E+07	468	Rv2454c	-	gi 15609591	Probable oxidoreductase (beta subunit)	intermediary metabolism and respiration	Cytoplasmic	COG1013C	3,59E-05	-1,03335	without changes
2E+07	468	Rv2455c	-	gi 15609592	Probable oxidoreductase (alpha subunit)	intermediary metabolism and respiration	Cytoplasmic	COG0674C	0,026498	-0,65363	without changes
2E+07		Rv2457c	clpX	gi 15609594	Probable ATP-dependent CLP protease ATP-binding subunit ClpX	intermediary metabolism and respiration	Cytoplasmic	COG1219O	0,017555	-2,15591	without changes
2E+07	469	Rv2460c	clpP2	gi 15609597	Probable ATP-dependent CLP protease proteolytic subunit 2 ClpP2 (endopeptidase CLP 2)	intermediary metabolism and respiration	Cytoplasmic	COG0740OU	0,049839	-0,68961	without changes
2E+07		Rv2528c	mrr	gi 15609665	Probable restriction system protein Mrr	information pathways	Cytoplasmic	COG1715V	0,001785	-1,07	without changes
2E+07		Rv2563	-	gi 15609700	Probable glutamine-transport transmembrane protein ABC transporter	cell wall and cell processes	Cytoplasmic Membrane	COG4591M	0,003882	-0,89077	without changes
2E+07		Rv2564	glnQ	gi 15609701	Probable glutamine-transport ATP-binding protein ABC transporter GlnQ	cell wall and cell processes	Cytoplasmic Membrane	COG1136V	0,000189	-0,81912	Met243Leu
2E+07		Rv2590	fadD9	gi 15609727	Probable fatty-acid-CoA ligase FadD9 (fatty-acid-CoA synthetase) (fatty-acid-CoA synthase)	lipid metabolism	Cytoplasmic	COG1022I	6,78E-05	-0,99946	without changes
2E+07		Rv2616	-	gi 15609753	Conserved protein	conserved hypotheticals	Unknown	COG4762S	-	-	without changes
2E+07		Rv2623	TB31.7	gi 15609760	Universal stress protein family protein TB31.7	virulence, detoxification, adaptation	Cytoplasmic	COG0589T	0,015747	-1,25589	without changes
2E+07		Rv2626c	hrp1	gi 15609763	FOG- CBS domain	Regulatory proteins	Unknown	COG3620K	0,021691	-1,56171	without changes
2E+07		Rv2627c	-	gi 15609764	Conserved protein	conserved hypotheticals	Cytoplasmic	-	0,005833	-1,40046	Arg104Gly, TSS
2E+07	510	Rv2673	aftC	gi 15609810	Possible arabinofuranosyltransferase AftC	cell wall and cell processes	Cytoplasmic Membrane	-	6,80E-06	-2,33565	without changes
2E+07		Rv2690c	-	gi 15609827	Probable conserved integral membrane alanine and valine and leucine rich protein	cell wall and cell processes	Cytoplasmic Membrane	COG0531E	-	-	without changes
2E+07	517	Rv2711	ideR	gi 15609848	Iron-dependent repressor and activator IdeR	Regulatory proteins	Cytoplasmic	COG1321K	0,001589	-1,44989	sSNP, TSS
2E+07	518	Rv2714	-	gi 15609851	Conserved alanine and leucine rich protein	conserved hypotheticals	Cytoplasmic	-	0,032451	-1,12288	Val245Alla
2E+07		Rv2778c	-	gi 15609915	Conserved protein	conserved hypotheticals	Cytoplasmic	-	0,001176	-2,78773	without changes
2E+07	540	Rv2818c	-	gi 15609955	Hypothetical protein	unknown	Cytoplasmic	-	-	-	part of RD207

2E+07		Rv2831	echA16	gi 15609968	Probable enoyl-CoA hydratase EchA16 (enoyl hydratase) (unsaturated acyl-CoA hydratase) (crotonase)	lipid metabolism	Unknown	COG1024I	0,000782	-1,29362	Asp156Tyr
2E+07		Rv2868c	gcpE	gi 15610005	Enzyme involved in the deoxyxylulose pathway of isoprenoid biosynthesis	Intermediary metabolism and respiration	Cytoplasmic	COG0821I	0,000212	-0,95576	without changes
2E+07		Rv2882c	frf	gi 15610019	Ribosome recycling factor Frr (ribosome releasing factor) (RRF)	information pathways	Cytoplasmic	COG0233J	0,007831	-1,40184	without changes
2E+07		Rv2883c	pyrH	gi 15610020	Probable uridylate kinase PyrH (UK) (uridine monophosphate kinase) (UMP kinase)	intermediary metabolism and respiration	Cytoplasmic	COG0528F	0,015747	-0,92656	without changes
2E+07	559	Rv2908c	-	gi 15610045	Predicted RNA-binding protein (contains KH domain)	Regulatory proteins	Cytoplasmic	COG1837R	0,000842	-1,66553	without changes
2E+07	562	Rv2921c	ftsY	gi 15610058	Probable cell division protein FtsY (SRP receptor) (signal recognition particle receptor)	cell wall and cell processes	Cytoplasmic Membrane	COG0552U	0,014793	-0,96432	without changes
2E+07	564	Rv2925c	rnc	gi 15610062	Probable ribonuclease III Rnc (RNase III)	information pathways	Cytoplasmic	COG0571K	0,028727	-2,12524	without changes
2E+07	565	Rv2928	tesA	gi 15610065	Probable thioesterase TesA	lipid metabolism	Cytoplasmic	COG3208Q	0,026787	-0,98984	without changes
2E+07		Rv2951c	-	gi 15610088	Possible oxidoreductase	intermediary metabolism and respiration	Cytoplasmic	COG2141C	0,04164	-0,8752	without changes
2E+07		Rv2952	-	gi 15610089	Possible methyltransferase (methylase)	intermediary metabolism and respiration	Cytoplasmic	COG2226H	-	-	Gly176Arg
2E+07		Rv2967c	pca	gi 15610104	Probable pyruvate carboxylase Pca (pyruvic carboxylase)	intermediary metabolism and respiration	Cytoplasmic	COG1038C	0,037123	-0,44832	without changes
2E+07	570	Rv2969c	-	gi 15610106	Possible conserved membrane or secreted protein	cell wall and cell processes	Cytoplasmic	COG1651O	0,005795	-0,97065	without changes
2E+07		Rv2982c	gpdA2	gi 15610119	Probable glycerol-3-phosphate dehydrogenase [NAD(P)+] GpdA2 (NAD(P)H- dependent glycerol-3-phosphate dehydrogenase)	lipid metabolism	Cytoplasmic Membrane	COG0240C	0,000819	-1,5091	Pro133Ala
2E+07	574	Rv2987c	leuD	gi 15610124	Probable 3-isopropylmalate dehydratase (small subunit) LeuD (isopropylmalate isomerase) (alpha-IPM isomerase) (IPMI)	intermediary metabolism and respiration	Cytoplasmic	COG0066E	0,007831	-1,16094	without changes
2E+07		Rv3006	lppZ	gi 15610143	Probable conserved lipoprotein LppZ	cell wall and cell processes	Unknown	COG2133G	0,000975	-0,78255	without changes

2E+07		Rv3010c	pfkA	gi 15610147	Probable 6-phosphofructokinase PfkA (phosphohexokinase) (phosphofructokinase)	intermediary metabolism and respiration	Cytoplasmic	COG0205G	0,009545	-1,26766	without changes
2E+07		Rv3044	fecB	gi 15610181	Probable FEIII-dictrate-binding periplasmic lipoprotein FecB	cell wall and cell processes	Unknown	COG0614P	0,014793	-0,92762	without changes
6E+07		Rv3048c	nrdF2	gi 57117051	Ribonucleoside-diphosphate reductase (beta chain) NrdF2 (ribonucleotide reductase small subunit) (R2F protein)	information pathways	Unknown	COG0208F	0,038694	-0,64318	without changes
2E+07		Rv3099c	-	gi 15610236	Conserved protein	conserved hypotheticals	Cytoplasmic	-	0,015898	-0,79309	without changes
2E+07		Rv3130c	tgsl	gi 15610266	Triacylglycerol synthase (diacylglycerol acyltransferase) Tgs1	lipid metabolism	Cytoplasmic	-	0,000116	-2,28598	without changes
4E+08		Rv3131	-	gi 448824799	diacylglycerol O-acyltransferase activity	Intermediary metabolism and respiration	Cytoplasmic Membrane	-	0,002641	-1,69236	without changes
6E+07		Rv3136	PPE51	gi 57117065	PPE family protein PPE51	PE/PPE	Cytoplasmic Membrane	COG5651N	0,000837	-1,64032	without changes
2E+07	600	Rv3158	nuoN	gi 15610294	Probable NADH dehydrogenase I (chain N) NuoN (NADH-ubiquinone oxidoreductase chain N)	intermediary metabolism and respiration	Cytoplasmic Membrane	COG1007C	-	-	without changes
2E+07	613	Rv3215	entC	gi 15610351	Probable isochorismate synthase EntC (isochorismate hydroxymutase) (enterochelin biosynthesis)	intermediary metabolism and respiration	Cytoplasmic	COG1169HQ	0,015791	-1,75397	without changes
2E+07	616	Rv3227	aroA	gi 15610363	3-phosphoshikimate 1-carboxyvinyltransferase AroA (5-enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS)	intermediary metabolism and respiration	Unknown	COG0128E	0,001892	-0,87917	without changes
2E+07	617	Rv3230c	-	gi 15610366	Hypothetical oxidoreductase	intermediary metabolism and respiration	Cytoplasmic	COG1018C	0,032334	-1,37623	without changes
2E+07		Rv3272	-	gi 15610408	Predicted acyl-CoA transferases/carnitine dehydratase	Intermediary metabolism and respiration	Cytoplasmic	COG1804C	-	-	Val181Met
2E+07		Rv3273	-	gi 15610409	Probable transmembrane carbonic anhydrase (carbonate dehydratase) (carbonic dehydratase)	intermediary metabolism and respiration	Cytoplasmic Membrane	COG0659P	3,18E-06	-1,20821	without changes
2E+07	630	Rv3283	sseA	gi 15610419	Probable thiosulfate sulfurtransferase SseA (rhodanese) (thiosulfate cyanide transsulfurase) (thiosulfate thiotransferase)	intermediary metabolism and respiration	Cytoplasmic	COG2897P	8,49E-18	-4,63167	Glu276Lys
2E+07	638	Rv3318	sdhA	gi 15610454	Probable succinate dehydrogenase (flavoprotein subunit) SdhA (succinic dehydrogenase) (fumarate reductase) (fumarate dehydrogenase) (fumaric	intermediary metabolism and respiration	Cytoplasmic Membrane	COG1053C	0,015159	-0,3754	without changes

					hydrogenase)						
2E+07		Rv3389c	htdY	gi 15610525	Probable 3-hydroxyacyl-thioester dehydratase HtdY	intermediary metabolism and respiration	Cytoplasmic	COG2030I	0,001097	-0,77623	Pro165Ser
2E+07	655	Rv3400	-	gi 15610536	Probable hydrolase	intermediary metabolism and respiration	Cytoplasmic	COG0637R	0,045163	-0,92577	without changes
2E+07		Rv3409c	choD	gi 15610545	Cholesterol oxidase ChoD (cholesterol-O2 oxidoreductase)	lipid metabolism	Unknown	COG2303E	3,59E-05	-1,80358	without changes
2E+07	669	Rv3456c	rplQ	gi 15610592	50S ribosomal protein L17 RplQ	information pathways	Cytoplasmic	COG0203J	0,043168	-1,30221	without changes
2E+07	669	Rv3457c	rpoA	gi 15610593	Probable DNA-directed RNA polymerase (alpha chain) RpoA (transcriptase alpha chain) (RNA polymerase alpha subunit) (DNA-directed RNA nucleotidyltransferase)	information pathways	Cytoplasmic	COG0202K	0,001589	-0,75917	without changes
2E+07	677	Rv3482c	-	gi 15610618	Probable conserved membrane protein	cell wall and cell processes	Unknown	-	-	-	without changes
2E+07	682	Rv3509c	ilvX	gi 15610645	Probable acetohydroxyacid synthase IlvX (acetolactate synthase)	intermediary metabolism and respiration	Cytoplasmic Membrane	COG0028EH	0,008018	-0,66135	without changes
2E+07		Rv3520c	-	gi 15610656	Possible coenzyme F420-dependent oxidoreductase	intermediary metabolism and respiration	Cytoplasmic	COG2141C	0,026787	-0,75499	without changes
2E+07		Rv3597c	lsr2	gi 15610733	Iron-regulated H-NS-like protein Lsr2	information pathways	Cytoplasmic	-	0,002169	-1,05164	without changes
2E+07	702	Rv3602c	panC	gi 15610738	Pantoate--beta-alanine ligase PanC (pantothenate synthetase) (pantoate activating enzyme)	intermediary metabolism and respiration	Cytoplasmic	COG0414H	0,001444	-1,25369	without changes
2E+07	707	Rv3626c	-	gi 15610762	proteolysis/nucleotide biosynthesis	Intermediary metabolism and respiration	Cytoplasmic	COG5282S	-	-	without changes
2E+07		Rv3694c	-	gi 15610830	Possible conserved transmembrane protein	cell wall and cell processes	Cytoplasmic Membrane	COG1300S	-	-	Upstream region (T4137219G)
4E+08	726	Rv3720	-	gi 448824818	Possible fatty acid synthase	lipid metabolism	Cytoplasmic	-	0,000954	-1,01974	without changes
2E+07	736	Rv3790	dprE1	gi 15610926	Decaprenylphosphoryl-beta-D-ribose 2'-oxidase	lipid metabolism	Unknown	COG0277C	0,00014	-0,8145	without changes

2E+07	736	Rv3791	dprE2	gi 15610927	Decaprenylphosphoryl-D-2-keto erythro pentose reductase	lipid metabolism	Cytoplasmic	COG0300R	0,003882	-0,84582	without changes
2E+07		Rv3797	fadE35	gi 15610933	Probable acyl-CoA dehydrogenase FadE35	lipid metabolism	Cytoplasmic	COG1960I	-	-	Ser328Pro; sSNP
2E+07		Rv3802c	-	gi 15610938	Probable conserved membrane protein	cell wall and cell processes	Cytoplasmic	-	-	-	Upstream region (4264382insC)
2E+07		Rv3805c	aftB	gi 15610941	Possible arabinofuranosyltransferase AftB	cell wall and cell processes	Cytoplasmic Membrane	-	-	-	Asp397Gly
2E+07		Rv3825c	pks2	gi 15610961	Polyketide synthase Pks2	lipid metabolism	Cytoplasmic Membrane	COG3321Q	0,031607	-1,00127	without changes
2E+07	756	Rv3885c	eccE2	gi 15611021	ESX conserved component EccE2. ESX- 2 type VII secretion system protein. Possible membrane protein.	cell wall and cell processes	Unknown	-	-	-	sSNP
2E+07	758	Rv3895c	eccB2	gi 15611031	ESX conserved component EccB2. ESX- 2 type VII secretion system protein. Probable membrane protein.	cell wall and cell processes	Unknown	-	-	-	without changes
2E+07	764	Rv3921c	-	gi 15611057	Probable conserved transmembrane protein	cell wall and cell processes	Cytoplasmic Membrane	COG0706U	0,008743	-1,6701	without changes