

На правах рукописи

ШИТИКОВ ЕГОР АЛЕКСАНДРОВИЧ

**ГЕНОМНАЯ ВАРИАБЕЛЬНОСТЬ ВОЗБУДИТЕЛЕЙ
ЛЕКАРСТВЕННО-УСТОЙЧИВОГО ТУБЕРКУЛЕЗА,
РАСПРОСТРАНЕННЫХ НА ТЕРРИТОРИИ
РОССИЙСКОЙ ФЕДЕРАЦИИ**

03.01.04 – биохимия

АВТОРЕФЕРАТ

диссертации на соискание ученой степени
кандидата биологических наук

Москва – 2014

Работа выполнена в Федеральном государственном бюджетном учреждении науки «Научно-исследовательский институт физико-химической медицины Федерального медико-биологического агентства», в лаборатории молекулярной генетики микроорганизмов

Научные руководители: доктор биологических наук, доцент
Ильина Елена Николаевна

Официальные оппоненты: **Ажикина Татьяна Леодоровна**
доктор биологических наук,
Федеральное государственное бюджетное
учреждение науки Институт
биоорганической химии им. академиков
М.М. Шемякина и Ю.А. Овчинникова
Российской академии наук, лаборатория
структуры и функций генов человека,
ведущий научный сотрудник

Сергиев Петр Владимирович
доктор химических наук,
МГУ имени М.В. Ломоносова,
химический факультет, кафедра химии
природных соединений, профессор

Ведущая организация: Федеральное государственное бюджетное
учреждение науки Институт биохимии им.
А.Н. Баха Российской академии наук

Защита состоится «23» октября 2014 г. в 11 часов на заседании диссертационного совета Д 001.010.01 при Федеральном государственном бюджетном учреждении «Научно-исследовательский институт биомедицинской химии имени В.Н. Ореховича» РАМН по адресу: 119121, Москва, ул. Погодинская, д. 10, стр. 8.

С диссертацией можно ознакомиться в библиотеке и на сайте ФГБУ «ИБМХ» РАМН

Автореферат разослан _____ 2014 г.

Ученый секретарь диссертационного совета,
кандидат химических наук

Карпова Е.А.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность проблемы

В настоящее время туберкулез остается одной из наиболее значимых проблем здравоохранения во всем мире. По оценке Всемирной организации здравоохранения Россия является одной из 22 стран мира с наибольшим бременем данного заболевания. В стране регистрируется более трети всех новых случаев туберкулеза, выявленных в Европейском регионе, причем смертность превышает показатели стран Европы в 5 – 8 раз (WHO, 2012).

Инфекционным агентом, вызывающим данное заболевание, является микобактерия туберкулеза (*Mycobacterium tuberculosis*). Использование современных молекулярно-генетических методов типирования, таких как IS6110 RFLP-анализ (от англ. Restriction Fragment Length Polymorphisms IS6110, анализ полиморфизма длины рестрикционных фрагментов IS6110) (van Embden *et al.*, 1993), сполиготипирование (Kamerbeek *et al.*, 1997) и VNTR-анализ (от англ. Variable Number of Tandem Repeats, анализ числа tandemных повторов в различных локусах генома) (Supply *et al.*, 2006) позволило определить структуру популяции возбудителя туберкулеза в России. Согласно многочисленным исследованиям на территории страны преобладают изоляты генотипов Beijing, Ural и LAM (Mokrousov *et al.*, 2003; Kovalev *et al.*, 2005; Dymova *et al.*, 2011). При этом доля генотипа Beijing в зависимости от региона может достигать 80 %, а относящиеся к нему представители демонстрируют повышенную вирулентность, способность размножаться в макрофагах и быструю адаптацию к иммунной системе макроорганизма (Mokrousov, 2013). В настоящее время так же показана ассоциация изолятов генотипа Beijing с повышенной лекарственной устойчивостью, обусловленной точечными мутациями в геноме (Hanekom *et al.*, 2011). При этом следует отметить, что в России все чаще выявляются изоляты различных семейств *M. tuberculosis* с так называемой множественной лекарственной устойчивостью (устойчивые, как минимум, к рифампицину и изониазиду) и широкой лекарственной устойчивостью (устойчивые к изониазиду и рифампицину, одному из фторхинолонов, и, по крайней мере, к одному из трех инъекционных препаратов второго ряда – канамицину, капреомицину или амикацину). В связи с этим большое количество работ посвящено созданию и внедрению молекулярно-генетических подходов для выявления устойчивых форм *M. tuberculosis*, так как применяемые на сегодняшний день микробиологические методы весьма трудоемки, длительны, плохо стандартизуемы и дороги.

Стоит отметить, что по причине разнообразия молекулярных механизмов развития устойчивости не во всех случаях генетическое тестирование отражает реальный фенотип патогена, регистрируемый микробиологическими тестами (Zhang, 2005). Это обуславливает необходимость дальнейших исследований, направленных на уточнение

известных и поиск новых генетических детерминант лекарственной устойчивости возбудителя туберкулеза.

В настоящий момент развитие методов полногеномного секвенирования и сравнительной геномики способствует активизации усилий в решении сформулированных задач (Farhat *et al.*, 2013; Zhang *et al.*, 2013). Данные методы позволяют оценить как микроэволюционные изменения в геноме, приводящие, к примеру, к развитию лекарственной устойчивости, так и изучить макроэволюцию патогена, что является весьма актуальным в связи с появлением на территории страны штаммов *M. tuberculosis*, характеризующихся популяционной “успешностью”.

Цель исследования

Охарактеризовать микро- и макроэволюционные изменения в геномах эндемичных для России штаммов *M. tuberculosis*, обладающих разным профилем лекарственной устойчивости

Задачи работы

1. Формирование коллекции образцов геномной ДНК эндемичных для России штаммов *M. tuberculosis*, охарактеризованных по профилю лекарственной чувствительности

2. Разработка метода определения спוליгопрофиля *M. tuberculosis* с использованием реакции удлинения зонда и последующим MALDI-ToF масс-спектрометрическим анализом. Сполиготипирование коллекции образцов *M. tuberculosis*

3. Проведение полногеномного секвенирования отобранной группы образцов *M. tuberculosis* с последующей сборкой и аннотацией полученных данных

4. Проведение филогенетического анализа и определение генотип-специфических мутаций включенных в исследование генетических семейств

5. Изучение геномной организации штаммов генотипа Beijing, как наиболее распространенных на территории России

6. Изучение молекулярных основ развития лекарственной устойчивости как формы микроэволюции *M. tuberculosis*

7. Сравнительный анализ мутационного профиля секвенированных геномов для поиска кандидатных маркеров устойчивости

Личное участие автора в получении результатов

Основной вклад автора состоит в теоретическом обосновании выдвигаемых гипотез, детальном планировании и реализации экспериментальной части, анализе результатов, написании диссертационной работы. Автором осуществлена адаптация метода удлинения зонда с последующим масс-спектрометрическим анализом для проведения спוליготипирования возбудителя туберкулеза. Проведен сбор коллекции биологических образцов. Проанализированы результаты полногеномного

секвенирования. Обнаружены и экспериментально подтверждены неизвестные ранее рекомбинационные события в геноме *M. tuberculosis*. Предложена модель реконструкции наблюдаемой рекомбинации. Произведен поиск генетических маркеров устойчивости *M. tuberculosis* к противотуберкулезным препаратам, определен список новых геномных полиморфизмов, потенциально ассоциированных с лекарственной устойчивостью.

Научная новизна и практическая значимость работы

В работе использованы современные биохимические и молекулярно-генетические методы для изучения микро- и макроэволюционных изменений в геномах эндемичных для России изолятов *M. tuberculosis*.

На основе реакции удлинения зонда с последующим масс-спектрометрическим анализом разработан лабораторный метод для быстрого сполиготипирования *M. tuberculosis*. Метод показал точность выдаваемых результатов, высокую производительность и низкую себестоимость тестирования, что может быть использовано для проведения масштабного эпидемиологического исследования патогена.

На основании проведенного полногеномного секвенирования впервые описаны генотип-специфические полиморфизмы циркулирующих на территории России изолятов *M. tuberculosis*. Полученные результаты могут быть использованы как для быстрой дифференциации эндемичных для России возбудителей туберкулеза, так и для функционального анализа с дальнейшим соотнесением фенотипических особенностей с генетическим контекстом.

Впервые получена полная геномная последовательность эндемичного для России штамма генетического семейства LAM. Результаты могут быть использованы для дальнейшего анализа характерных особенностей представителей данного семейства.

Впервые в мире показана крупная перестройка сегментов хромосомы в *M. tuberculosis* кластера Beijing B0/W148. Данное открытие может привести к переосмыслению ряда представлений о крайней степени мономорфности генома *M. tuberculosis*. Дополнительно описаны кластер-специфические однонуклеотидные полиморфизмы, которые, совместно с описанными рекомбинациями, могут отчасти объяснить “успешность” представителей описываемого кластера, а также могут быть использованы для мониторинга.

Впервые на основании результатов полногеномного секвенирования эндемичных для России образцов ДНК *M. tuberculosis* проведена комплексная оценка маркеров устойчивости к противотуберкулезным препаратам. Проведен поиск новых детерминант устойчивости. Выявленные мутации могут служить в дальнейшем основой для дополнительных исследований в области антибиотикорезистентности.

В целом результаты диссертационной работы представляют большую практическую значимость и могут быть использованы для решения прикладных задач клинической микробиологии и эпидемиологии возбудителей туберкулеза.

Основные положения диссертации, выносимые на защиту

1) Разработан лабораторный метод сполиготипирования на основе реакции удлинения зонда с последующим масс-спектрометрическим анализом, показавший полную сходимость с результатами классического типирования и превосходящий его по скорости получения данных.

2) В ходе полногеномного секвенирования и сравнительного анализа выявлены специфические для генетических семейств полиморфизмы, которые могут быть использованы как для надежной идентификации генотипов Beijing, LAM и Ural, так и для функционального анализа с дальнейшим соотношением фенотипа с генотипом.

3) На основании сравнительной геномики описаны полиморфизмы и перестройки сегментов хромосомы, отчасти объясняющие “успешность” представителей кластера Beijing B0/W148, и которые могут служить основой для дальнейших прицельных исследований, а также создания систем генетического мониторинга указанного кластера.

4) Проведен поиск новых детерминант устойчивости к противотуберкулезным препаратам среди эндемичных для России штаммов *M. tuberculosis*. Показана ступенчатость развития антибиотикорезистентности, выявлены новые полиморфизмы, ассоциированные с лекарственной устойчивостью.

Апробация работы

Основные положения диссертационной работы доложены и обсуждены на расширенном межлабораторном заседании Отдела молекулярной биологии и генетики ФГБУН НИИ ФХМ ФМБА России (Москва, 22 мая 2014 г.), а также в ходе ряда международных конференций (41-я Всемирная конференция по легочным заболеваниям (Берлин, Германия, 2010), 5-я Европейская Конференция по Геномике Прокариот и Грибов (Геттинген, Германия, 2011), 5-я Международная школа молодых учёных по молекулярной генетике на тему «Непостоянство генома» (Звенигород, Россия, 2012).

Публикации

По теме диссертации опубликовано шесть работ в рецензируемых научных журналах.

Объем и структура диссертации

Диссертация изложена на 176 страницах машинописного текста, содержит 6 приложений, 22 таблицы и 18 рисунков. Состоит из следующих

разделов: «Введение», «Обзор литературы», «Материалы и методы», «Результаты», «Обсуждение», «Заключение и выводы», «Список литературы», который включает 242 источника.

МАТЕРИАЛЫ И МЕТОДЫ

Бактериальные штаммы и геномные последовательности

В исследование включены 500 препаратов ДНК штаммов *M. tuberculosis*, отобранных из коллекций трех центров по борьбе с туберкулезом: ЦНИИ Туберкулеза РАМН, Государственного казенного учреждения здравоохранения города Москвы «Московский городской научно-практический Центр борьбы с туберкулезом Департамента здравоохранения г. Москвы» и Санкт-Петербургского НИИ Фтизиопульмонологии. Для 150 образцов был известен уровень лекарственной устойчивости к препаратам первого и второго ряда. Для 310 образцов были известны сполиготипы, определенные по классической методике (Kamerbeek *et al.*, 1997).

Дополнительно использовали выборку геномных последовательностей 373 образцов *M. tuberculosis* и одного образца *M. canettii*, находящихся в открытом доступе в базе данных NCBI.

Генетический анализ образцов *M. tuberculosis*

Для амплификации фрагментов генома, ассоциированных с лекарственной устойчивостью, использовали праймеры, описанные ранее (Икрянникова *et al.*, 2007; Zimenkov *et al.*, 2013). Амплификацию DR-региона для сполиготипирования проводили согласно классическому протоколу с использованием небийотинилированных праймеров (Kamerbeek *et al.*, 1997). Амплификацию tandemных повторов в 24 локусах генома осуществляли по методике, описанной ранее (Supply *et al.*, 2006). Для подтверждения рекомбинационных событий были использованы праймеры, приведенные в таблице 1.

Таблица 1. 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 |

Амплификацию осуществляли на приборе DNA Engine Tetrad 2 (MJ Research, США). В реакционную смесь входили следующие компоненты: 66 мМ Tris-HCl pH 9.0; 16,6 мМ (NH₄)₂SO₄; 2,5 мМ MgCl₂, по 250 мкМ каждого дНТФ, 1 ед Taq-полимеразы (ЛИТЕХ, Россия) и по 10 пмоль соответствующих праймеров. Конечный объем реакции составлял 25 мкл. Общий вид профиля реакции: 94°C – 5 мин. 30–45 циклов 94°C – 15 сек. Та °С – 15 сек. 72°C – от 15 сек. до 3 мин. 72°C – 10 мин. 4 °С – хранение. Та – температура отжига праймера.

В случае необходимости проведения реакции удлинения зонда или секвенирования проводили подготовку продуктов амплификации. Для этого добавляли по 5 мкл смеси, содержащей 66 мМ Tris-HCl pH 9.0; 16.6 мМ (NH₄)₂SO₄; 2.5 мМ MgCl₂; 0.5 Ед щелочной фосфатазы арктических креветок (Shrimp Alkaline Phosphatase (Fermentas, Литва)) и 5 Ед экзонуклеазы I *E. coli* (ExoI, Fermentas, Литва). Полученную смесь инкубировали 20 минут при 37 °С с последующей инактивацией ферментов прогреванием в течение 10 минут при

85°C. В образцы, предназначенные для проведения реакции удлинения зонда, экзонуклеазу I не добавляли.

Для определения структуры DR-региона использовали реакцию удлинения зонда с последующим масс-спектрометрическим анализом продуктов. Для каждого из 43 спейсерных участков были разработаны высокоспецифичные праймеры, в дальнейшем сгруппированные в 8 систем. Реакцию удлинения зонда осуществляли в 10 мкл реакционной смеси, содержащей 66 мМ Tris-HCl (pH 9.0); 16.6 мМ (NH₄)₂SO₄; 2.5 мМ MgCl₂; по 0.2 мМ дидезоксирибонуклеозидтрифосфата (ддНТФ), необходимого для реакции; по 10 пмоль каждого праймера, который использовался в системе и 2 Ед TermiPol ДНК-полимеразы (Solis Biodyne, Эстония), используя в качестве матрицы предварительно очищенные амплифицированные фрагменты генома. Нарработку продуктов реакции осуществляли по универсальному профилю: 95 °С – 5 мин. 70 циклов: 94 °С – 20 сек, 60 °С – 20 сек, 72 °С – 15 сек.

Очистку продуктов реакции удлинения зонда осуществляли на катионно-обменной смоле (SPS-SAC-50, Синтез полимерных сорбентов, Россия). Масс-спектрометрический анализ продуктов реакции проводили на приборе Autoflex III smartbeam vertical MALDI ToF (Bruker Daltonics, Германия) согласно инструкциям производителя. 3-гидроксипиколиновая кислота (Fluka, Германия) служила главным компонентом матрицы.

Обработку масс-спектров осуществляли с использованием программного обеспечения фирмы Bruker Daltonics (Германия). Масс-спектры каждого образца сравнивали с контрольным, и по наличию изменения молекулярной массы, соответствующей массе используемого ддНТФ, определяли наличие того или иного спейсера.

Нуклеотидные последовательности фрагментов генома определяли по методу Сенгера с модификациями (Sanger *et al.*, 1997) с использованием прибора ABI Prism 3730xl Genetic Analyzer (“Applied Biosystems”, США; “Hitachi”, Япония). Реакцию секвенирования проводили с использованием тех же праймеров, что для реакции амплификации. Восстановление полноразмерных последовательностей, выравнивание и сравнительный анализ последовательностей осуществляли с использованием ContigExpress и AlignX программного пакета «Vector NTI 11.0» (Informax Inc, США).

Определение полногеномных последовательностей 54 штаммов проводили методом секвенирования с использованием высокопроизводительного секвенатора GS FLX + (Roche, США). Фрагментные библиотеки получали в соответствии с протоколом производителя (Rapid Library Preparation Method Manual GSFLX+ Series – XL+, Roche, США) с использованием коммерческого набора GS FLX Titanium Rapid Library Preparation Kit (Roche, США). Для образца SP21 дополнительно проводили секвенирование с использованием прибора IonTorrent PGM (LifeTechnologies, США) и двух библиотек парных фрагментов (с размером библиотеки 2000–3000 п.о. и 5000–6000 п.о.). Библиотеки готовили с использованием коммерческого набора 5500 SOLiDMate-Paired Library Construction Kit (Life Technologies, США) в соответствии с протоколом Ion Mate-Paired Library Preparation (Life Technologies Demonstrated Protocol).

Анализ данных полногеномного секвенирования, статистический анализ данных

Анализ данных полногеномного секвенирования проводили на вычислительном комплексе ФГБУН НИИ ФХМ ФМБА России с использованием открытого программного обеспечения, а также разработанных в лаборатории скриптов на языке программирования R.

Картирование чтений с прибора на референтный геном *M. tuberculosis* штамма H37Rv производили с помощью инструмента bowtie2 (Langmead and Salzberg, 2012) с использованием настроек по умолчанию. Анализ данных выравнивания проводили в

приложении samtools 0.1.19 (Li et al., 2009). Поиск однонуклеотидных полиморфизмов осуществляли с помощью программы VarScan v2.3.1 (Koboldt et al., 2012).

Выравнивание и поиск замен для полногеномных последовательностей производили с помощью программного пакета MUMMER 3 и входящих в него утилит nucmer, show-snps, show-coords (Delcher et al., 2002)

Для расчета dN/dS (dN – отношение несинонимичных замен на несинонимичные сайты; dS – отношение синонимичных замен на синонимичные сайты) замен использовали KaKs_Calculator (<https://code.google.com/p/kaks-calculator/>).

Для оценки разнообразия *M. tuberculosis* вычисляли матрицу расстояний Хэмминга на основании однонуклеотидных замен. Построение филогенетических деревьев осуществляли исходя из полученной матрицы с применением алгоритма Neighbor-joining. В качестве внешней группы использовали образец *M. canettii* (NC_015848.1). Для уменьшения размерности и двумерного представления разнообразия применяли принцип главных компонент.

Для определения значимости отличий dN/dS между группами использовали anova тест при значении p-value ≤ 0.05 . Определение достоверности отличий количества замен в COG категориях и оценка значимости различий в частотах встречаемости замен проводили с помощью биномиального теста (p-value ≤ 0.05). Для сравнения матриц расстояний, полученных на основании нуклеотидных замен и данных VNTR-анализа, использовали корреляции Спирмена и Пирсона.

Для определения достоверности ассоциации нуклеотидной замены с резистентностью использовали точный тест Фишера (p-value ≤ 0.05). Для исключения ложно-положительных ассоциаций использовали поправку Бенджамини-Гохберга. Для корректировки полученных результатов использовали логистическую регрессию, основанную на ступенчатом механизме развития лекарственной устойчивости. Последовательность возникновения устойчивости выглядела следующим образом: {Xn} {INH, RMP, STR, EMB, OFX, ETH, KAN, CPM}. Для каждой из ассоциированных с резистентностью к антибиотику {Xn} мутации проводили сравнение двух логистических моделей вида $X_n \sim X_1 + X_2 + \dots + X_{n-1}$ и $X_n \sim X_1 + X_2 + \dots + X_{n-1} + \text{SNP}$ с помощью хи-квадрат теста (p-value ≤ 0.05). В качестве нулевой гипотезы рассматривали равнозначность моделей.

РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ И ИХ ОБСУЖДЕНИЕ

Формирование экспериментальных групп образцов

Сформирована лабораторная коллекция препаратов ДНК, экстрагированных из 500 клинических изолятов *M. tuberculosis*, выделенных от пациентов трех центров по борьбе с туберкулезом. Для 150 образцов в бактериологических лабораториях соответствующих учреждений была установлена лекарственная чувствительность к противотуберкулезным препаратам первой и второй линии терапии. Группа образцов (n = 310) с определенным по классической схеме сполиготипом была использована для разработки инновационного метода сполиготипирования, базирующегося на применении реакции удлинения зонда с последующей MALDI-ToF масс-спектрометрической детекцией. Разработанный метод позволил установить сполигопаттерны оставшихся 190 образцов ДНК.

Из группы образцов, прошедших сполиготипирование и охарактеризованных по чувствительности к противотуберкулезным

препаратам первого и второго ряда, были отобраны 54 образца для осуществления полногеномного секвенирования. Представленные образцы относились к эндемичным для России генотипам и имели разный уровень лекарственной устойчивости, что позволяет изучить как генетическую вариабельность на уровне семейств, так и произвести поиск новых детерминант резистентности.

Сформирована коллекция геномных последовательностей 373 образцов *M. tuberculosis* и одного образца *M. canettii*, находящихся в открытом доступе в базе данных NCBI. Из них геномные последовательности 8 образцов (7 *M. tuberculosis* (KZN-4207, KZN-2475, 94_M4241A, CCDC5079, CCDC5180, W-148 и OSDD493) и 1 *M. canettii*), представленные в виде протяженных секвенсов, использованы для построения филогенетического дерева. Данные о 366 геномах МБТ, представленные в виде чтений с приборов высокопроизводительного секвенирования (Illumina), использованы для корректировки генотип-специфических полиморфизмов. Данные о 46 *M. tuberculosis* с известным профилем устойчивости к противотуберкулезным препаратам использованы для построения гипотез формирования резистентности.

Сполиготипирование клинических изолятов *M. tuberculosis*

Для молекулярно-генетического типирования микобактерий наиболее часто используются следующие методы: IS6110 RFLP-анализ, VNTR-типирование и сполиготипирование. Выбор сполиготипирования для описания образцов коллекции был обусловлен возможностью сравнения полученных результатов с уже известными в литературе и находящимися в международных базах данных SpolDB4 (Brudey *et al.*, 2006) и SITVIT (Demay *et al.*, 2012). Метод обладает меньшей дискриминирующей способностью по сравнению с IS6110 RFLP- и VNTR-типированием, однако, успешность его применения определялась быстротой и надежностью выявления эндемичных семейств популяции.

В настоящем исследовании разработана методика сполиготипирования, базирующаяся на реакции удлинения зонда с последующим масс-спектрометрическим анализом. В основе метода лежит ферментативное достраивание олигонуклеотидного зонда на определенный дидезоксирибонуклеозидтрифосфат с последующей масс-спектрометрической детекцией продуктов реакции. Подобранные зонды были сгруппированы в 8 мультиплексных реакций для определения всех 43 спейсеров DR-региона. При наличии места посадки детектировалось удлинение зонда, что говорило о наличии спейсера. Метод был апробирован на выборке из 310 образцов и показал полную сходимость с результатами, полученными по классической методике. В дальнейшем он был применен для тестирования 190 образцах с неизвестными сполиготипами.

Сравнительный анализ полученных для 500 образцов паттернов с интернациональными базами данных SpolDB4 и SITVIT_WEB разделил исследованную выборку на 45 генотипов (база SITVIT_WEB, номер SIT), входящих в состав 16 генетических семейств (база SpolDB4). Наибольшее количество образцов (88,2 %) относилось к семействам Beijing, Ural и LAM.

Таким образом, структура DR-региона была определена методом сполиготипирования для 500 образцов ДНК *M. tuberculosis*, полученных от ЦНИИ Туберкулеза РАМН, МНПЦБТ и СПбНИИФ.

Считывание и анализ геномных последовательностей отобранных изолятов *M. tuberculosis*

На основании суммарно накопленной информации была сформирована выборка из 150 образцов, наиболее полно охарактеризованных с клинико-эпидемиологической точки зрения (таблица 2).

Таблица 2. Характеристика 150 образцов коллекции с учетом уровня устойчивости и проведенного сполиготипирования

| Уровень устойчивости | Генетические семейства | | | | Всего |
|----------------------|------------------------|-----|------|-----------|-------|
| | Beijing | LAM | Ural | остальные | |
| Чувствительные | 8 | 2 | 3 | 7 | 20 |
| Монорезистентные | 1 | 0 | 0 | 1 | 2 |
| Полирезистентные | 11 | 0 | 1 | 4 | 16 |
| МЛУ | 70 | 4 | 7 | 2 | 83 |
| ШЛУ | 18 | 4 | 3 | 4 | 29 |
| Всего | 108 | 10 | 14 | 18 | 150 |

Из соответствующей коллекции были отобраны 54 образца для проведения полногеномного секвенирования. Выбранные образцы относились к наиболее часто встречаемым генотипам на территории России: Beijing (35 образцов), Ural (10) и LAM (9). В каждую группу были включены как лекарственно-чувствительные, так и лекарственно-устойчивые бактерии.

Полногеномное секвенирование было проведено на высокопроизводительном геномном секвенаторе GS FLX+ Series — XL+ (Roche, США) с использованием фрагментных бар-кодированных библиотек. Полученные чтения для каждого образца депонированы в Sequence read archive (SRA) (проект № PRJNA181180).

Для каждого образца была проведена как сборка *de novo*, так и выравнивание на референтный геном штамма *M. tuberculosis* H37Rv для идентификации однонуклеотидных замен. Сравнительный анализ данных показал низкую частоту встречаемости полиморфизмов в геномах (3 SNPs на 10 т.п.о.), что согласуется с данными мировой литературы и характеризует *M. tuberculosis* как генетически мономорфный микроорганизм. Для оценки силы и направленности естественного отбора, действующего на геном патогена, использовали отношение dN/dS. По результатам исследования значение этого показателя составило 0.573 в среднем для всех геномов (стандартное

отклонение 0.03), что согласуется с данными мировой литературы и свидетельствует об отрицательном естественном отборе.

Дополнительно в результате проведенных исследований впервые в России была определена полная геномная последовательность *M. tuberculosis* чувствительного штамма CTRI-2, относящегося к семейству LAM. Представленный геном был аннотирован и депонирован в международную базу данных NCBI (номер в GenBank: CP002992).

Сравнительный филогенетический анализ

На основании полученных данных полногеномного секвенирования построено филогенетическое дерево с учетом всех однонуклеотидных замен геномов. В контексте данного исследования филогенетический анализ использовался для оценки разнообразия секвенированных образцов и дальнейшего поиска специфических для изучаемых генетических семейств полиморфизмов. Для построения филогенетического дерева были использованы все полиморфизмы представленных геномов. В качестве внешней группы был выбран геном *M. canettii*. Такой способ плохо отражает филогенетические взаимоотношения в случае быстро эволюционирующих организмов, подверженных частым рекомбинациям, однако может быть успешно применен в случае генетически мономорфных микроорганизмов, к которым относится *M. tuberculosis* (Achtman, 2008).

В ходе анализа филогенетического дерева было установлено, что штаммы различных филогенетических линий сгруппированы в отдельные кластеры (рисунок 1). В кластер изолятов, соответствующих генотипу Beijing (Восточно-Азиатская линия), вошли штаммы MOS10, MOS11, MOS12, MOS14, MOS15, MOS16, MOS17, MOS18, SP1, SP2, SP3, SP4, SP5, SP6, SP7, SP8, SP9, SP10, SP11, SP12, SP13, SP15, SP16, SP17, SP18, SP21, SP22, SP23, SP26, SP27, SP29, SP30, SP34, SP45, CTRI-4, W-148, 94_M4241A, CCDC5180 и CCDC5179; генотипу Ural: MOS9, SP24, SP25, SP28, SP31, SP32, SP33, SP41, SP42, CTRI-3 и OSDD493; генотипу LAM: MOS2, MOS7, SP35, SP36, SP37, SP38, SP39, SP46, CTRI-2, KZN4207 и KZN2475. При этом образцы генетических семейств Ural и LAM расположились на одной внутренней ветви, что свидетельствует об их эволюционном родстве. Данные семейства относятся к распространенной Евро-Американской линии и несут схожие филогенетические маркеры.

Таким образом, результаты филогенетического анализа секвенированных штаммов показали сходимость с результатами первичного генотипирования собранных образцов.

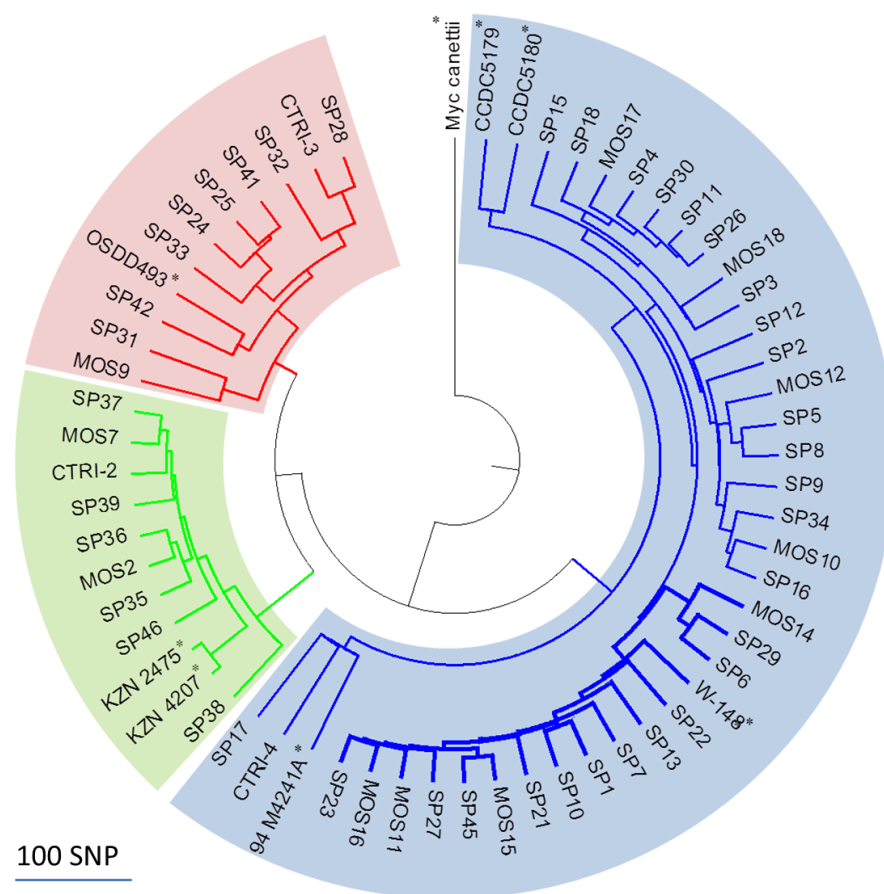


Рисунок 1. Сравнительный филогенетический анализ секвенированных штаммов. Алгоритм построения Neighbor-Joining. Синим цветом выделены образцы семейства Beijing, зеленым – LAM, красным – Ural. Звездочкой отмечены геномы из GenBank.

Определение и валидация кластер-специфических мутаций

В ходе сравнительного анализа данных полногеномного секвенирования для филогенетических семейств Beijing, Ural и LAM были отобраны генотип-специфические мутации. Представленные семейства относятся к двум из шести основных филогенетических линий, и в связи с этим для корректировки полученных значений в работу было включено 366 геномов *M. tuberculosis*, представляющих мировое филогенетическое разнообразие патогена.

По результатам исследования для представителей семейства Beijing было идентифицировано 422 генотип-специфических полиморфизмов; для Ural – 143 SNPs; для LAM – 309 SNPs. Полученные генотип-специфические nsSNPs были использованы для анализа вызываемых аминокислотных замен. Распределение белков, несущих эти замены, по COGs категориям (от англ. Clusters of Orthologous Groups, кластеры ортологичных групп) представлено на рисунке 2. Нормализация количества nsSNPs была проведена с учетом длин генов в каждом представленном COG и с учетом количества несинонимичных полиморфизмов, найденных в конкретном семействе.

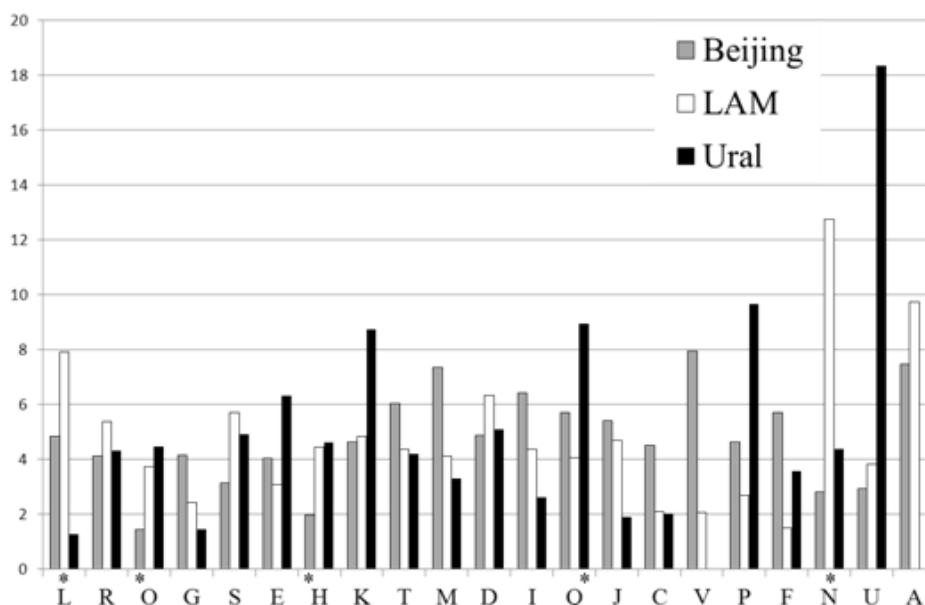


Рисунок 2. Гистограмма распределения генотип-специфических nsSNPs по категориям COG. Звездочкой отмечены статистические значимые различия в количестве nsSNPs в представленном COG определенного семейства ($p < 0.05$). А – процессинг и модификация РНК; С – получение и преобразование энергии; D – контроль клеточного цикла, клеточное деление, деление хромосом; E – транспорт и метаболизм аминокислот; F – транспорт и метаболизм нуклеиновых кислот; G – транспорт и метаболизм углеводов; H – транспорт и метаболизм коферментов, I – транспорт и метаболизм липидов, J – трансляция, рибосомные структуры и биогенез; K – транскрипция; L – репликация, рекомбинация и репарация; M – биосинтез клеточной стенки и мембраны, N – клеточная подвижность; O – посттрансляционные модификации, шапероны, пептидазы; P – транспорт и метаболизм неорганических ионов; Q – биосинтез вторичных метаболитов, транспорт и катаболизм; R – предсказана только основная функция; S – функция не известна; T – механизмы передачи сигналов, регуляция; U – внутриклеточный и везикулярный транспорт, секреция; V – защитные механизмы.

Согласно проведенному анализу штаммы семейства Beijing достоверно отличались от Ural и LAM накоплением замен в белках, относящихся к кластерам O и H. Количество замен в белках этих кластеров было ниже, чем у образцов других линий, что свидетельствует о возможном эволюционном давлении. В свою очередь представители генотипа LAM показали достоверно большее количество замен в кластерах L и N. Кластер L ассоциирован с системой репликации, репарации и рекомбинации и представляется весьма интересным с точки зрения формирования мутаторного генотипа. Штаммы семейства Ural достоверно отличались от LAM и Beijing накоплением замен в белках, относящихся к COG кластеру Q.

Характеристика кластера Beijing B0/W148

Сравнительный анализ VNTR-паттернов представителей семейства Beijing разделил исследуемую выборку на 14 VNTR-профилей. Согласно Мокроусову с соавторами наиболее обширные группы были отнесены к типам M2 (17 образцов) и M11 (15 образцов). При этом 12 образцов из типа

M11 имели профили, ассоциированные с кластером Beijing B0/W148 (223325173533424672454433 (9 образцов), 223325173533424572454433 (2 образца), 223325173533424672444433 (1 образец)) (Mokrousov, 2013). К кластеру Beijing B0/W148 относились образцы SP1, SP7, SP10, SP13, SP21, SP22, SP23, SP27, SP45, MOS11, MOS15 и MOS16. К этой группе образцов также относился геном штамма W-148, загруженный из базы данных GenBank.

Методами сравнительной геномики для образцов кластера Beijing B0/W148 был определен характерный мутационный профиль, включающий 63 полиморфизма. Наиболее значимые полиморфизмы в межгенных областях и в открытых рамках считывания представлены в таблице 3.

Таблица 3. Наиболее значимые замены в образцах кластера Beijing B0/W148

| Позиция | Ген | Ам-ная замена | Продукт гена и его функция |
|---------|-------------------|---------------|--|
| 404130 | <i>Rv0338c</i> | Glu571Gly | Предположительно, железосвязывающая редуктаза. Участвует в метаболизме и дыхании |
| 886661 | межгенная область | | Промоторная область оперона <i>Rv0789c-Rv0792c</i> . <i>gntR</i> (<i>Rv0792c</i>) является белком, относящимся к HutC подсемейству регуляторов транскрипции |
| 1035426 | <i>pstS3</i> | Phe175Cys | Фосфат-связывающий липопротеин PstS3 (PBP-3) (<i>PstS3</i>). Участвует в транспорте через мембрану соединений фосфора |
| 2078967 | <i>Rv1833c</i> | Asp275Tyr | Возможно, галогеноалкан дегалогеназа. Может действовать на широкий спектр 1-галогеналканы, галоген спирты, галогеналканы и некоторые галоид замещенным ароматические соединения |
| 2127011 | <i>Rv1877</i> | Trp370Gly | Предположительно, консервативный интегральный белок мембраны. Возможно связан с транспортом лекарств через мембрану |
| 3031090 | <i>chiZ</i> | Gln150Ter | Возможно, консервативный мембранный белок. Белок-регулятор клеточного деления <i>M. tuberculosis</i> |
| 3109512 | <i>Rv2800</i> | Trp366Ter | Возможно, гидролаза. Участвует в клеточном метаболизме |
| 4175847 | <i>Rv3728</i> | Trp325Ter | Предположительно, консервативный двух доменный мембранный белок. Белок относится к представителям эффлюксной системы MFS. Возможно участвует в транспорте углеводов и антибиотиков |
| 4338371 | <i>whiB6</i> | Thr51Pro | Возможно, регулятор транскрипции WhiB6 |

Для определения геномной организации представителей кластера использовался геном W-148, представленный в NCBI в виде одного скаффолда. Выравнивание геномов W-148 и H37Rv показало наличие двух больших инверсий в геноме W-148 (рисунок 3), которые впоследствии были подтверждены методом ПЦР для всех представителей кластера в имеющейся коллекции. Дополнительно наличие инверсии было проверено в ходе

полногеномного секвенирования библиотеки парных фрагментов образца SP21, относящегося к кластеру Beijing B0/W148.

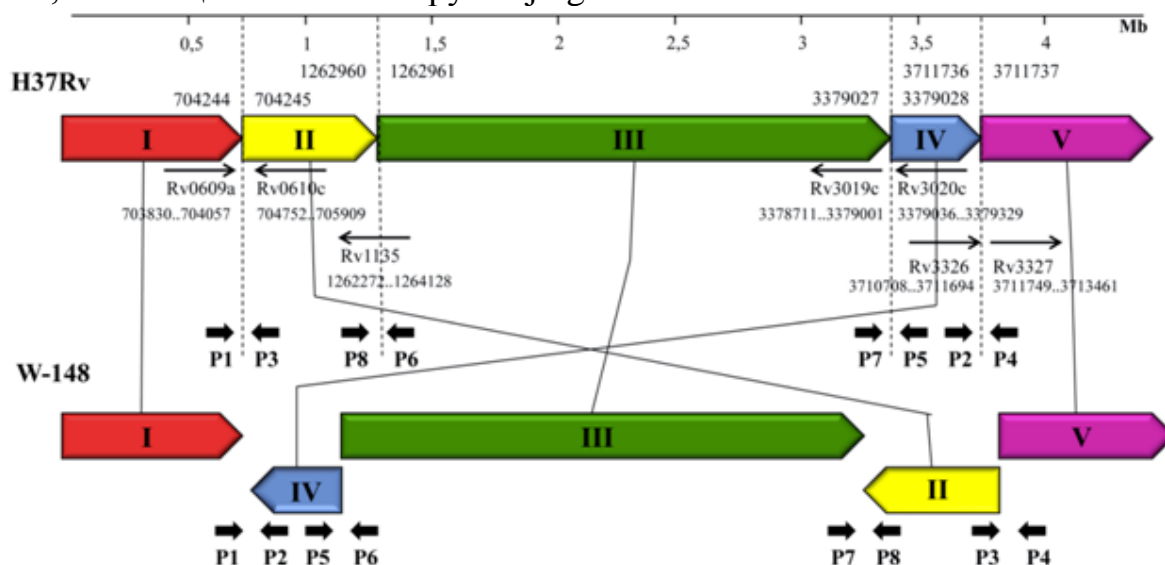


Рисунок 3. Схематическое изображение перестроек в геноме W-148 относительно H37Rv. Римскими цифрами I, III и IV отмечены консервативные блоки. Блоки II и IV в геноме W-148 инвертированы и переставлены. Тонкими стрелками отмечены гены и их позиции в области инверсии. Жирные стрелки соответствуют праймерам (таблица 1)

Методом *in silico* реконструированы возможные рекомбинационные события, произошедшие в геноме предкового штамма W-148. Согласно модели, порядок и ориентация коллинеарных блоков в геноме предшественника 1 (P1) соответствовали геному H37Rv. В ходе первого рекомбинационного события произошла инверсия участка генома длиной 3 м.п.о., затронувшая блоки II, III и IV. Относительно генома H37Rv рекомбинация хромосомной ДНК произошла между генами Rv0609a-Rv0610c с одной стороны и Rv3326-Rv3327 с другой.

Следующий этап рекомбинационных событий произошел в геноме предшественника 2 (P2), затронув блок III между генами Rv3019 и Rv3020c с одной стороны и разбив ген Rv1135c с другой. В результате данной инверсии было восстановлено исходное положение этого сегмента как в P1 и H37Rv, что привело к образованию W-148.

Описанные на основании сравнительной геномики полиморфизмы и геномные перестройки могут служить основой для дальнейших прицельных исследований, объясняющих успешность кластера Beijing B0/W148, а также использованы при создании систем генетического мониторинга указанного семейства.

Анализ генетических маркеров, ассоциированных с лекарственной устойчивостью *M. tuberculosis*

Для 54 секвенированных образцов коллекции был произведен поиск полиморфизмов, ассоциированных с резистентностью.

В исследуемой выборке 48 образцов были устойчивы как к изониазиду, так и к рифампицину, что было подтверждено на генетическом уровне.

Аминокислотная замена Ser315Thr в KatG была выявлена у подавляющего количества образцов (47 из 48) в независимости от генетического семейства. Следует отметить, что мутация С(-15)Т в промоторной области оперона *tabA(fabG) – inhA* была выявлена у 9.3 % и 62.5 % ИИ-устойчивых штаммов семейства Beijing и штаммов Евро-Американской линии, соответственно (OR 6.6; 95 % CI 1.6–27.7).

Мутации в гене *rpoB* были обнаружены среди всех рифампицин-устойчивых образцов, причем мутация в кодоне 450 была выявлена существенно чаще у штаммов семейства Beijing, нежели чем у штаммов Евро-Американской линии: 90.6 % против 56.3 %.

Анализ распределения мутаций в гене *embB* 35 этамбутол-устойчивых штаммов не показал существенных различий в изучаемых семействах. Большинство замен было выявлено в кодоне 306 гена *embB* (Met306Val и Met306Ile). Данные замены встретились в 27 секвенированных штаммах (23 устойчивых и 4 чувствительных), относящихся к разным генотипам, и были статистически достоверно ассоциированы с устойчивостью к этамбутолу.

Мутации в генах *rpsL* и *rrs*, ассоциированных с устойчивостью к стрептомицину, были выявлены у 44 образцов коллекции. Найденные мутации в большей степени были сопряжены с генетическими семействами: замена Lys88Arg была выявлена преимущественно в образцах семейства Ural, в то время как для образцов генотипа Beijing такого рода замена была определена лишь единожды. В свою очередь все устойчивые образцы семейства LAM несли мутацию A514C в гене *rrs*.

Согласно исследованию наиболее часто встречаемыми мутациями, приводящими к устойчивости к инъекционным препаратам, была мутация A1401G в гене *rrs* (52 % штаммов) и замены в промоторной области гена *eis* (51.5 %). Пять образцов из секвенированной коллекции, чувствительных к представленным препаратам, также содержали мутации в промоторной области гена *eis*.

Мутации в QRDR-регионе генов *gyrA* и *gyrB* были выявлены у 88 % штаммов, устойчивых к фторхинолонам. При этом наиболее часто замены встречались в 94 кодоне гена *gyrA* (16 из 27 устойчивых штаммов), ассоциированном с высоким МИК к офлоксацину.

Таким образом, в ходе исследования был проведен анализ канонических маркеров устойчивости к наиболее часто используемым противотуберкулезным препаратам как на уровне всей выборки секвенированных образцов, так и на уровне отдельных генотипов. Так, например, для большей части препаратов первого ряда (рифампицин, изониазид и стрептомицин) можно отметить не случайность распределения маркеров устойчивости в зависимости от генетической линии штамма. В свою очередь для этамбутола, а также препаратов второго ряда, таких закономерностей выявить не удалось. Данные результаты могут свидетельствовать как о возможном отборе на уровне генетических семейств,

так и о влиянии геномной организации патогена на закрепление тех или иных мутаций. Дополнительно следует отметить, что в целом полученные результаты согласуются с данными аналогичных исследований других авторов, однако, не для всех препаратов удалось обнаружить известные маркеры резистентности в изучаемых образцах.

Для поиска новых генетических маркеров, ассоциированных с лекарственной устойчивостью, в нашем исследовании был проведен анализ геномных последовательностей 100 образцов МБТ (54 секвенированных образца и 46 образцов из NCBI). Отличительная особенность представленной работы заключалась в выборе эндемичных для России штаммов. Дополнительно задачей исследования было определение наиболее вероятного пути развития лекарственной устойчивости.

Согласно изначальной предпосылке, мутации, ассоциированные с лекарственной устойчивостью, чаще встречаются в устойчивых, чем в чувствительных к противотуберкулезным препаратам штаммах. Для формирования списка таких мутаций был проведен анализ распределения однонуклеотидных полиморфизмов в геномах исследуемой выборки образцов. Для увеличения точности анализа были выбраны следующие критерии отбора SNPs: 1) покрытие однонуклеотидного полиморфизма не менее тремя чтениями с прибора; 2) встречаемость SNP менее чем у 90 % образцов; 3) анализируемые SNPs не должны встречаться одновременно только у двух образцов наименьшего кластера согласно филогенетическому анализу; 4) анализируемый SNP не должен находиться в генах, кодирующих семейство белков PE-PPE. В ходе применения выбранных критериев из общего массива 6403 SNPs было отобрано 4135 полиморфизмов. Для вычисления вероятности ассоциации мутаций с устойчивостью был использован точный тест Фишера. В ходе исследования было выявлено 104 SNPs в генах и 12 SNPs в межгенных участках ($p < 0.01$) (данные не приведены). Для дальнейшей проверки значимости полученных полиморфизмов была использована логистическая регрессия. Согласно анализу было получено 22 и 2 значимые мутации в генах (несинонимичные полиморфизмы) и межгенных участках, соответственно ($p < 0.01$) (таблица 4).

Согласно анализу полученных полиморфизмов 10 SNPs относились к уже изученным маркерам устойчивости к антибиотикам. Сопоставление оставшихся 14 SNPs с результатами двух описанных ранее исследований выявило один общий полиморфизм в гене *rpoC*. Замена Val483Gly в соответствующем белке является компенсаторной и описана ранее (Comas *et al.*, 2011).

Из специфических мутаций, характерных для образцов семейства Beijing типа M2, следует отметить замены Glu761Asp в RpoB и C517T в *rrs*. Данные мутации были выявлены при анализе генетических маркеров, ассоциированных с лекарственной устойчивостью и в представленном анализе. При этом замена Glu761Asp в RpoB была выявлена только у

штаммов, несущих мутации в RRDR регионе гена *rpoB*, что может свидетельствовать о ее возможном компенсаторном влиянии. Здесь также следует отметить, что других компенсаторных мутаций в данных образцах обнаружено не было. В свою очередь, замена C517T в *rrs* встретила только среди стрептомицин-устойчивых штаммах, не несущих других известных детерминант устойчивости к представленному препарату. Оставшиеся шесть аминокислотных замен были выявлены в белках, участвующих в посттрансляционных модификациях белков, в клеточном метаболизме и в транспорте веществ через клеточную стенку и представляют существенный интерес для дальнейших исследований.

Таблица 4. Характеристика 100 исследуемых образцов с учетом профиля лекарственной устойчивости в зависимости от генотипа

| Позиция | Ген | Замена | Противотуберкулезный препарат; | | | | | | | | | | | | количество | | | |
|---------|-----------------------|-----------|--------------------------------------|---|-----|---|-----|---|-----|---|-----|---|-----|---|------------|----|--|--|
| | | | чувствительных и устойчивых образцов | | | | | | | | | | | | | | | |
| | | | INH | | RIF | | STR | | EMB | | OFX | | KAN | | CAP | | | |
| R | S | R | S | R | S | R | S | R | S | R | S | R | S | R | S | | | |
| 7570 | <i>gyrA*</i> | Ala90Val | - | - | - | - | - | - | - | - | 5 | 0 | - | - | - | - | | |
| 7582 | <i>gyrA*</i> | Asp94Ala | 16 | 0 | - | - | - | - | 16 | 0 | 15 | 1 | - | - | - | - | | |
| 692217 | <i>lprL</i> | Asp65Gly | 12 | 0 | - | - | - | - | - | - | - | - | 11 | 1 | 0 | 12 | | |
| 761155 | <i>rpoB*</i> | Ser450Leu | 62 | 0 | 62 | 0 | - | - | - | - | - | - | - | - | - | - | | |
| 762089 | <i>rpoB</i> | Glu761Asp | 13 | 0 | - | - | - | - | - | - | - | - | 12 | 1 | 0 | 13 | | |
| 764817 | <i>rpoC</i> | Val483Gly | 12 | 0 | - | - | - | - | - | - | - | - | - | - | - | - | | |
| 781687 | <i>rpsL*</i> | Lys43Arg | 41 | 1 | - | - | 41 | 1 | - | - | - | - | - | - | - | - | | |
| 799029 | <i>Rv0698</i> | Ser66Ter | 13 | 0 | - | - | - | - | - | - | - | - | 12 | 1 | 0 | 13 | | |
| 1112625 | <i>Rv0996</i> | Pro81Leu | 13 | 0 | - | - | - | - | - | - | - | - | 12 | 1 | 0 | 13 | | |
| 1245775 | <i>gnd2</i> | Val216Gly | 13 | 0 | - | - | - | - | - | - | - | - | 12 | 1 | 0 | 13 | | |
| 1472362 | <i>rrs</i> | C517T | 19 | 0 | - | - | 18 | 1 | - | - | - | - | - | - | 1 | 18 | | |
| 1473246 | <i>rrs*</i> | A1401G | 24 | 0 | - | - | - | - | - | - | - | - | 24 | 0 | 20 | 4 | | |
| 1927734 | <i>Rv1702c</i> | Leu281Pro | 12 | 0 | - | - | - | - | - | - | - | - | 11 | 1 | 0 | 12 | | |
| 2101193 | <i>ureD</i> | Ser58Gly | - | - | - | - | - | - | - | - | - | - | 8 | 0 | - | - | | |
| 2155168 | <i>katG*</i> | Ser315Thr | 75 | 1 | - | - | - | - | - | - | - | - | - | - | - | - | | |
| 2335638 | <i>Rv2079</i> | Leu95Ser | 21 | 0 | - | - | - | - | - | - | - | - | - | - | - | - | | |
| 3149251 | <i>nusA</i> | Arg60Cys | 13 | 0 | - | - | - | - | - | - | - | - | 12 | 1 | 0 | 13 | | |
| 3297704 | <i>pks15</i> | Thr46Ile | 13 | 0 | - | - | - | - | - | - | - | - | 12 | 1 | 0 | 13 | | |
| 3521253 | <i>nuoJ</i> | Val39Ile | - | - | - | - | - | - | 10 | 0 | - | - | - | - | - | - | | |
| 4137219 | <i>Rv3695</i> | Val5Gly | - | - | - | - | - | - | - | - | - | - | 9 | 1 | - | - | | |
| 4247429 | <i>embB*</i> | Met306Val | 25 | 0 | - | - | - | - | 23 | 2 | - | - | - | - | - | - | | |
| 4247574 | <i>embB*</i> | Asp354Ala | 14 | 0 | - | - | - | - | - | - | - | - | 13 | 1 | 0 | 14 | | |
| 2715369 | <i>Rv2417c-eis*</i> | | 13 | 0 | - | - | - | - | - | - | - | - | 12 | 1 | 0 | 13 | | |
| 4327480 | <i>ethR-ethA*</i> | | 13 | 0 | - | - | - | - | - | - | - | - | 12 | 1 | 0 | 13 | | |

ЗАКЛЮЧЕНИЕ

В настоящее время туберкулез остается серьезным инфекционным заболеванием. Высокий уровень заболеваемости и смертности в России связан с рядом генетических и фенотипических особенностей патогена. Одной из особенностей является все больше распространение изолятов *M. tuberculosis* с множественной и широкой лекарственной устойчивостью. Использование методов молекулярного типирования позволило определить структуру популяции возбудителя туберкулеза и при этом выявить наиболее распространенные семейства. Так, на территории России преобладают штаммы семейств Beijing, Ural и LAM. Особый научный интерес представляет семейство Beijing, ассоциированное с повышенной вирулентностью и лекарственной устойчивостью. Таким образом, актуальным являются исследования направленные как на изучение популяционной структуры изолятов, преобладающих на территории Российской Федерации, так и их генетических особенностей, ассоциированных с лекарственной устойчивостью.

В данной работе нами предложен новый метод для типирования *M. tuberculosis*, сполиготипирование, основанный на реакции удлинения зонда с последующим масс-спектрометрическим анализом. Метод апробирован на большой выборке образцов и показал полную сходимость с результатами классического типирования и превосходящий его по скорости получения данных.

Так же, на основании проведенного полногеномного секвенирования описаны семейство-специфические полиморфизмы циркулирующих на территории России изолятов *M. tuberculosis*. Получена полная геномная последовательность эндемичного для России штамма семейства LAM. Проведена комплексная оценка маркеров устойчивости к противотуберкулезным препаратам, образцов ДНК *M. tuberculosis* эндемичных для России. Проведен поиск новых детерминант устойчивости.

В ходе исследования показана крупная перестройка сегментов хромосомы в *M. tuberculosis* кластера Beijing B0/W148. Дополнительно описаны кластер-специфические однонуклеотидные полиморфизмы, которые, совместно с описанными рекомбинациями, могут отчасти объяснить “успешность” представителей описываемого кластера, а также могут быть использованы для мониторинга кластера.

Полученные результаты могут быть использованы как для быстрой дифференциации эндемичных для России возбудителей туберкулеза, так и для функционального анализа с дальнейшим соотношением фенотипических особенностей с генетическим контекстом. Результаты работы представляют большую практическую значимость и могут быть использованы для решения прикладных задач клинической микробиологии и эпидемиологии возбудителей туберкулеза.

ВЫВОДЫ

1. Сформирована охарактеризованная коллекция из 500 образцов геномной ДНК штаммов *M. tuberculosis*.
2. Разработанный на основе MALDI-ToF масс-спектрометрического анализа способ определения споллигопрофиля *M. tuberculosis* реализован для типирования отобранных образцов ДНК.
3. Проведено полногеномное секвенирование 54 образцов геномной ДНК *M. tuberculosis*, относящихся к эндемичным для России генетическим семействам – Beijing (35), LAM (9) и Ural (10). Геномные данные депонированы в международную базу данных NCBI.
4. В ходе сравнительного филогенетического анализа определены полиморфизмы, обеспечивающие надежную дифференциацию штаммов *M. tuberculosis* на уровне семейств: Beijing (422 SNPs), Ural (143) и LAM (309).
5. Охарактеризован мутационный профиль “успешного” кластера Beijing B0/W148, включающий 63 полиморфизма. Описаны крупные перестройки сегментов хромосомной ДНК этих штаммов, предложена вероятная последовательность рекомбинационных событий.
6. Определены генетические детерминанты резистентности, характерные для лекарственно-устойчивых штаммов генотипов Beijing, Ural и LAM. В частности, показана ассоциация между принадлежностью к генетическому семейству Beijing и наличием замены Ser450Leu в RpoB.
7. Показана ступенчатость формирования лекарственной устойчивости для эндемичных для России штаммов *M. tuberculosis*. Выявлено 24 полиморфизма, показавших ассоциацию с лекарственной устойчивостью.

Список работ, опубликованных по теме диссертации

- 1) Shitikov E, Ilina E, Chernousova L, Borovskaya A, Rukin I, Afanas'ev M, Smirnova T, Vorobyeva A, Larionova E, Andreevskaya S, Kostrzewa M, Govorun V. Mass spectrometry based methods for the discrimination and typing of mycobacteria // Infect Genet Evol. 2012. Vol. 12. №4. P. 838-845.
- 2) Ilina EN, Shitikov EA, Ikryannikova LN, Alekseev DG, Kamashev DE, Malakhova MV, Parfenova TV, Afanas'ev MV, Ischenko DS, Bazaleev NA, Smirnova TG, Larionova EE, Chernousova LN, Beletsky AV, Mardanov AV, Ravin NV, Skryabin KG, Govorun VM. Comparative genomic analysis of *Mycobacterium tuberculosis* drug resistant strains from Russia // PLoS One. 2013. Vol. 8. №2. P.e56577.
- 3) Shitikov EA, Bespyatykh JA, Ischenko DS, Alexeev DG, Karpova IY, Kostryukova ES, Isaeva YD, Nosova EY, Mokrousov IV, Vyazovaya AA, Narvskaya OV, Vishnevsky BI, Otten TF, Zhuravlev VY, Yablonsky PK, Ilina EN, Govorun VM. Unusual large-scale chromosomal rearrangements in *Mycobacterium tuberculosis* Beijing B0/W148 cluster isolates // PLoS One. 2014. Vol. 9 № 1. P. e84971.
- 4) Chernyaeva EN, Shulgina MV, Rotkevich MS, Dobrynin PV, Simonov SA, Shitikov EA, Ischenko DS, Karpova IY, Kostryukova ES, Ilina EN, Govorun VM, Zhuravlev VY, Manicheva OA, Yablonsky PK, Isaeva YD, Nosova EY, Mokrousov IV, Vyazovaya AA, Narvskaya OV, Lapidus AL, O'Brien SJ. Genome-wide *Mycobacterium tuberculosis* variation (GMTV) database: a new tool for integrating sequence variations and epidemiology // BMC Genomics. 2014. Vol. 15. № 1:308.
- 5) Bespyatykh JA, Zimenkov DV, Shitikov EA, Kulagina EV, Lapa SA, Gryadunov DA, Ilina EN, Govorun VM. Spoligotyping of *Mycobacterium tuberculosis* complex isolates using hydrogel oligonucleotide microarrays // Infect Genet Evol. 2014. Vol. 26 P.41-46.
- 6) Зименков Д.В., Кулагина Е.В., Антонова О.В., Суржигов С.А., Беспятых Ю.А., Шитиков Е.А., Ильина Е.Н., Михайлович В.М., Заседателев А.С., Грядунов Д.А. Анализ генетических детерминант множественной и широкой лекарственной устойчивости возбудителя туберкулеза с использованием олигонуклеотидного микрочипа // Молекулярная биология. 2014. Т. 48. № 2. С. 251-264.
- 7) Е.А. Шитиков, Д.С. Ищенко, И.В. Мокроусов, О.В. Нарвская, Е.Н. Ильин. Определение кластер-специфических мутаций “успешного” клона B0 *M. tuberculosis*. Пятая Международная школа молодых учёных по молекулярной генетике на тему «Непостоянство генома». 2012, стр 70.
- 8) E.A. Shitikov, E.N. Ilina, A.D. Borovskaya, L.N. Chernousova, T.G. Smirnova, A.V. Vorobyeva, E.E. Larionova, M Kostrzewa. Mass spectrometry based methods for the discrimination and typing of mycobacteria. 41st World Conference on Lung Health. 2010. S242.
- 9) E. Shitikov, D. Ischenko, E. Ilina, L. Ikryannikova, M. Malakhova, T. Parfenova, I. Rukin, L. Chernousova, E. Larionova, V. Govorun. Genome sequencing of XDR *M. tuberculosis* strain endemic for Russian Federation. 5th European Conference on Prokaryotic and Fungal Genomics. 2011. p. 170.

Список сокращений

ДНК – дезоксирибонуклеиновая кислота

ддНТФ – дидезоксирибонуклеозидтрифосфат

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

МБТ – *Mycobacterium tuberculosis*, микобактерия туберкулеза

DR – Direct Repeat (англ. прямой повтор)

MALDI-ToF MS – Matrix-Assisted Laser Desorption/Ionization Time Of Flight Mass-Spectrometry (англ. времяпролётная масс-спектрометрия с матричной лазерной десорбционной ионизацией)

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

RD – Region of Difference (англ. регион различия)

RRDR – Rifampicin Resistance Determining Region (англ. участок, определяющий устойчивость к рифампицину)

SNP – Single Nucleotide Polymorphism (англ. однонуклеотидный полиморфизм)

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