

**ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ
УЧРЕЖДЕНИЕ «НАУЧНО-ИССЛЕДОВАТЕЛЬСКИЙ ИНСТИТУТ
ФИЗИКО-ХИМИЧЕСКОЙ МЕДИЦИНЫ ФЕДЕРАЛЬНОГО МЕДИКО-
БИОЛОГИЧЕСКОГО АГЕНТСТВА»**

на правах рукописи

ШИТИКОВ ЕГОР АЛЕКСАНДРОВИЧ

**Геномная вариабельность возбудителей
лекарственно-устойчивого туберкулеза,
распространенных на территории
Российской Федерации**

03.01.04 – биохимия

Диссертация на соискание ученой степени
кандидата биологических наук

Научный руководитель
д.б.н., доцент, Ильина Е.Н.

Москва 2014

Оглавление

ВВЕДЕНИЕ.....	4
1 ОБЗОР ЛИТЕРАТУРЫ.....	10
1.1 Общая характеристика <i>Mycobacterium tuberculosis</i>	10
1.2 Особенности геномной организации <i>M. tuberculosis</i>	12
1.3 Эволюция микобактерий туберкулезного комплекса	14
1.4 Популяционная структура <i>M. tuberculosis</i>	19
1.5 Молекулярная эпидемиология туберкулеза в России	27
1.6 Генетическое семейство Beijing <i>M. tuberculosis</i>	32
1.7 Молекулярные основы возникновения устойчивости к противотуберкулезным препаратам.....	35
1.7.1 Механизм возникновения лекарственно-устойчивых штаммов.....	37
1.7.2 Формирование устойчивости к различным противотуберкулезным препаратам	41
2. МАТЕРИАЛЫ И МЕТОДЫ.....	47
2.1 Бактериальные штаммы	47
2.2 Геномные последовательности.....	48
2.3 Амплификация фрагментов генома <i>M. tuberculosis</i>	50
2.4 Подготовка продуктов амплификации для дальнейшего анализа.....	52
2.5 Реакция удлинения зонда и масс-спектрометрический анализ продуктов реакции	52
2.6 Секвенирование фрагментов генома модифицированным методом Сенгера	56
2.7 VNTR-типирование.....	56
2.8 Полногеномное секвенирование	57
2.9 Анализ данных полногеномного секвенирования.....	58
2.10 Статистическая обработка данных.....	59
3. РЕЗУЛЬТАТЫ	61
3.1 Формирование экспериментальных групп образцов ДНК штаммов <i>M.</i> <i>tuberculosis</i>	61
3.2 Формирование коллекции геномных последовательностей	61
3.3 Сполиготипирование штаммов <i>M. tuberculosis</i>	62
3.4 Считывание и первичный анализ геномных последовательностей отобранных эндемичных для России клинических изолятов <i>M. tuberculosis</i>	67

3.5 Сборка полных геномных последовательностей.....	71
3.5 Сравнительный филогенетический анализ секвенированных штаммов	73
3.6 Определение генотип-специфических мутаций	75
3.7 Валидация генотип-специфических SNPs.....	76
3.8 Характеристика кластера Beijing B0/W148.....	78
3.8.1 Beijing B0/W148 кластер-специфические SNPs.....	79
3.8.2 Структурная организация генома Beijing B0/W148	82
3.9 Анализ генетических маркеров, ассоциированных с лекарственной устойчивостью <i>M. tuberculosis</i>	87
3.10 Построение гипотез формирования устойчивости к противотуберкулезным препаратам.....	92
4. ОБСУЖДЕНИЕ	95
4.1. Формирование экспериментальных групп образцов ДНК штаммов <i>M. tuberculosis</i>	95
4.2 Сполиготипирование штаммов <i>M. tuberculosis</i>	95
4.3 Считывание и анализ геномных последовательностей отобранных изолятов <i>M. tuberculosis</i>	97
4.4 Сравнительный филогенетический анализ секвенированных штаммов	102
4.5 Определение и валидация генотип-специфических мутаций	105
4.6 Характеристика кластера Beijing B0/W148.....	108
4.7 Анализ генетических маркеров, ассоциированных с лекарственной устойчивостью <i>M. tuberculosis</i>	114
4.8 Построение гипотез формирования устойчивости к противотуберкулезным препаратам.....	118
ЗАКЛЮЧЕНИЕ	122
ВЫВОДЫ.....	124
Список используемых сокращений.....	125
Список литературы	127
ПРИЛОЖЕНИЯ.....	147

ВВЕДЕНИЕ

Актуальность темы исследования

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 сполиготипов Beijing, Ural и LAM (Mokrousov *et al.*, 2003; Kovalev *et al.*, 2005; Думова *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 сполиготипа 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 (WHO, 2006). 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

Научная новизна и практическая значимость работы

В работе использованы современные 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 спוליготипирования *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 семейство-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 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.

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, 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 кандидатные полиморфизмы, ассоциированные с лекарственной устойчивостью.

Апробация работы

По теме диссертации опубликовано шесть работ в рецензируемых научных журналах.

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).

1 ОБЗОР ЛИТЕРАТУРЫ

1.1 Общая характеристика *Mycobacterium tuberculosis*

Mycobacterium tuberculosis (МБТ) – инфекционный агент, вызывающий туберкулез. Заболевание характеризуется различной локализацией и клиническими проявлениями с тенденцией к хроническому и рецидивирующему течению.

Заболевание известно с давних времен: описание симптомов туберкулеза можно найти в трудах Гиппократ, Аристотеля, Клариссимуса Галена и др. Греки называли его «фтизис» – слово, совмещающее два значения: «кашлять кровью» и «чахнуть, быстро терять вес». Первые описания болезни и инфекционной природы возбудителя были представлены в итальянской медицинской литературе еще в XVII веке, однако крупнейшее научное событие в истории изучения туберкулеза произошло в 1882 году (Die Aetiologie der Tuberculose, 24 марта 1882), когда немецкий бактериолог Роберт Кох (1843-1910) после 17 лет лабораторных исследований выявил в мокроте больного возбудителя заболевания. В окуляре Кох увидел нечто, напоминающее «палочку» (отсюда и название – «палочка Коха») («Under the microscope the structures of the animal tissues, such as the nucleus and its breakdown products are brown, while the tubercle bacteria are a beautiful blue») (Koch, 1982). В дальнейшем, в ходе улучшения разрешающей способности микроскопов, бактериологам удалось установить, что возбудителем туберкулеза являются микобактерии. Изучение органов и тканей, пораженных микобактериями, установило наличие «бугорков» (от лат. «tuberculum») в очаге поражения, и поэтому болезнь стали называть «бугорчаткой» или «туберкулезом». В настоящее время используется только один термин – туберкулез.

Mycobacterium tuberculosis являются грамположительными палочками, длиной 1-10 мкм и диаметром 0.2-0.6 мкм. Морфологически выделяют, как прямые, так и слегка изогнутые формы. По типу дыхания микроорганизм относится к аэробам. Однако следует отметить, что в процессе жизнедеятельности

в неблагоприятных условиях метаболизм может изменяться, и бактерии могут трансформироваться в микроаэрофилы или становиться анаэробами. Температурные границы роста находятся между 29 и 42 °С (оптимальная – 37–38 °С). Размножается патоген поперечным делением. Процесс происходит крайне медленно – одно деление за 14–20 ч. При посеве патологического материала *M. tuberculosis* образуют первичный рост через 3-4 недели. Пассажируемые культуры растут быстрее – на 10-21 сутки. При культивировании на плотной яичной среде, содержащей глицерин; колонии шероховатые (R-колонии), имеют кремовый цвет, но так же могут быть гладкие, сливающиеся между собой. На жидкой питательной среде микобактерии туберкулеза образуют морщинистую грубую пленку, а иногда даже придонный крошковатый рост (Мишин, 2005).

В качестве стандартной среды для культивирования микобактерий туберкулеза ВОЗ рекомендована плотная яичная среда Левенштейна — Йенсена. В России и некоторых других странах широкое распространение получила рекомендованная в качестве второй стандартной яичная среда Финн-II (Кольчев, 2003; Мишин, 2005). Для повышения вероятности роста микобактерий в настоящее время рекомендуется засеивание патологического материала на 2—3 среды одновременно.

Биохимической дифференциальной особенностью МБТ является способность синтезировать большие количества ниацина. Так же для биохимической идентификации МБТ используют их способность к пигментообразованию, уреазную, никотинамидазную и пиразинамидазную активности.

Важнейшим элементом МБТ является клеточная стенка, состоящая из 3–4 связанных слоев толщиной до 200–250 нанометров. Главным компонентом стенки является пептидогликан, связанный с арабиногалактаном, который, в свою очередь, образует сложные эфиры с миколевой кислотой. Дополнительно клеточная стенка содержит специфичные воска (микозиды) и полисахариды (Kaig *et al.*, 2009; Niederweis *et al.*, 2010).

1.2 Особенности геномной организации *M. tuberculosis*

Последовательность генома *M. tuberculosis* штамма H37Rv была полностью расшифрована в Сенгеровском институте в 1998 году (Cole *et al.*, 1998) (Рисунок 1). Это был третий опубликованный бактериальный геном после *Haemophilus influenzae* (Fleischmann *et al.*, 1995) и *Mycoplasma genitalium* (Fraser *et al.*, 1995). На данный момент в GenBank представлено 23 полностью прочитанных и аннотированных генома МБТ, а проекты секвенирования открыты еще для более 2000 штаммов.

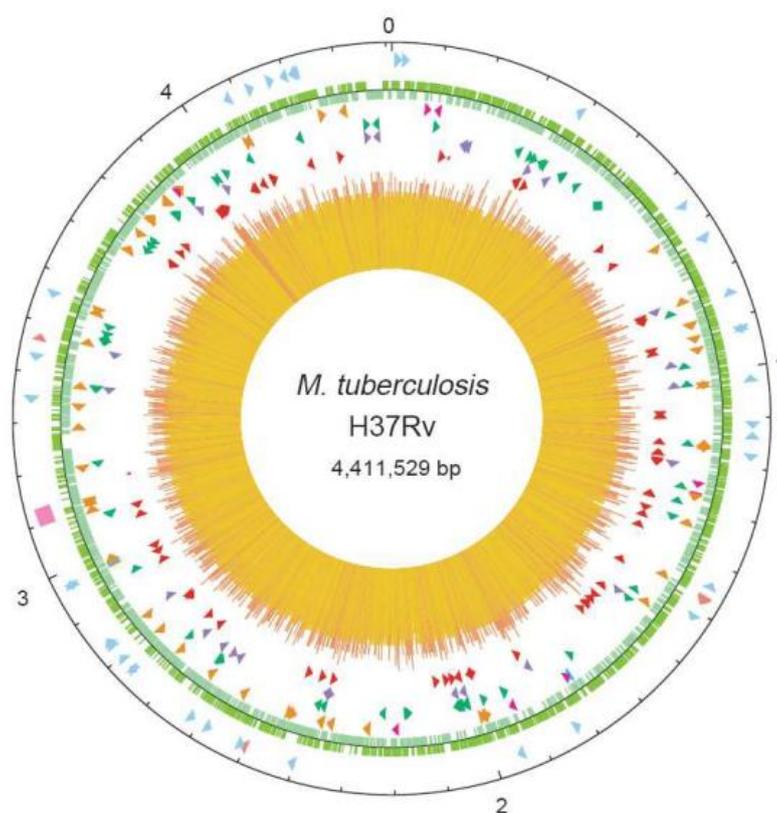


Рисунок 1. Круговая карта генома *M. tuberculosis* штамма H37Rv. По направлению изнутри наружу: 1) гистограмма G+C состава (красным >65%, желтым <65%); 2) первое (темно-красное), второе (лиловое) и третье (зеленое) кольца отражают положение генов семейств PPE, PE и PE-PGRS, соответственно; 3) четвертое 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)

Следует отметить, что ключевые особенности организации генома одинаковы для всех штаммов патогена. Геномы представлены кольцевой молекулой ДНК протяженностью около 4400 тысяч пар оснований (т.п.о.) и характеризуются высоким содержанием G+C пар (~ 65.5 %). При этом существует несколько регионов, отличающихся по G+C составу. К участкам с высоким G+C составом относятся крупные генные семейства PE и PPE, названные так в соответствии с N-терминальными мотивами ProGlu (PE), или ProProGlu (PPE), и состоящие из 100 и 67 членов в геноме штамма H37Rv, соответственно. При этом часть представителей семейства PE имеет домен с высоким содержанием глицина и соответственно G+C пар (PGRS, polymorphic G+C-rich sequence). Роль представленных семейств остается не до конца изученной, однако высказываются предположения об их значении в патогенезе и антигенной вариабельности (Mukhopadhyay and Balaji, 2011). К участкам с низким G+C составом (менее 50 %) относятся некоторые гены, кодирующие трансмембранные белки. Эволюционно это вызвано тем, что гидрофобные аминокислоты, входящие в состав трансмембранных доменов, кодируются кодонами с низким содержанием гуанина и цитозина.

Углубленный анализ генома *M. tuberculosis* штамма H37Rv выявил около 4000 генов, кодирующих белки. При этом следует отметить, что альтернативный старт трансляции GTG встретился в 35 % случаев, что существенно чаще, чем 14 % и 9 % в геномах *Bacillus subtilis* и *Escherichia coli*, соответственно. Был найден один набор рибосомных генов и 45 транспортных РНК.

Дальнейший анализ показал, что 3.4 % генома H37Rv занимают инсерционные элементы (от англ. Insertion Sequence elements; IS elements) и профаги (phiRv1 и phiRv2). Профаг phiRv1 интегрирован в область повторяющихся последовательностей семейства 13E12. В других штаммах МБТ представленный профаг может отсутствовать, либо находиться в других участках генома (Fleischmann *et al.*, 2002). На данный момент описано 7 потенциальных сайтов интеграции профага (Cole, 1999). Профаг phiRv2 более стабилен и показывает крайне малую вариабельность среди штаммов. Среди 56 локусов IS

элементов (семейства IS3, IS5, IS21, IS30, IS110, IS256 и ISL3), описанных в геноме H37Rv, наибольший интерес представляет инсерционный элемент IS6110, относящийся к семейству IS3. В связи с частыми транспозициями данный элемент широко используется в молекулярной эпидемиологии для штаммовой дифференциации (van Embden *et al.*, 1993). Геном H37Rv содержит шестнадцать повторов IS6110.

Гены белков, участвующих в липидном метаболизме, занимают около 8 % генома, что говорит об их весомом значении для жизненного цикла МБТ. Представленные данные согласуются с наличием широкого спектра липидов, липогликанов, гликолипидов и поликетидов в клеточной стенке патогена, а также указывают на то, что МБТ может использовать липиды и стеролы макроорганизма в качестве источника энергии.

Другой интересной особенностью *M. tuberculosis* является наличие очень эффективной и точной системы репарации. Mizrahi с соавторами было показано отсутствие в геноме микобактерий белков семейства MutHLS (MutS и MutL), ответственных за репарацию неспаренных оснований, что в свою очередь, возможно, компенсируется наличием 45 генов, в том числе трех копий гена *mutT*, вовлеченных так же в процессы репарации (Mizrahi and Andersen, 1998). Продукт гена *mutT*, специфичная пирофосфатаза, гидролизует дГТФ до дГМФ и пирофосфата, тем самым обеспечивая эксцизионную репарацию (base excision repair). Данный тип репарации является наиболее важным для микобактерий, так как высокий процентный G+C состав делает их восприимчивее к гуанин-специфическому стрессу. Так же у микобактерий обнаружены все гены, вовлеченные в систему SOS-ответа, за исключением *polB* и *umuD* (Cole *et al.*, 1998; Mizrahi and Andersen, 1998).

1.3 Эволюция микобактерий туберкулезного комплекса

Микобактерии туберкулезного комплекса (от англ. *Mycobacterium tuberculosis* complex) – это группа тесно взаимосвязанных видов и подвидов кислотоустойчивых бактерий, способных вызывать туберкулез (Smith *et al.*, 2006). К представителям комплекса относятся следующие виды: *M. africanum*

(Vasconcellos *et al.*, 2010), *M. bovis* (Garnier *et al.*, 2003), *M. canettii* (van Soolingen *et al.*, 1997), *M. caprae* (Niemann *et al.*, 2002), *M. microti* (Frota *et al.*, 2004), *M. mungi* (Alexander *et al.*, 2010), *M. orygis* (van Ingen *et al.*, 2012), *M. pinnipedii* (Cousins *et al.*, 2003) и *M. tuberculosis* (Cole *et al.*, 1998). Данные микроорганизмы характеризуются крайне низкой вероятностью горизонтального переноса генов между штаммами (Gutacker *et al.*, 2002; Smith *et al.*, 2003; Supply *et al.*, 2003; Hirsh *et al.*, 2004), и, что более существенно, являются одним из наиболее крайних примеров генетической гомогенности на уровне значений 0.01–0.03 % однонуклеотидных полиморфизмов (от англ. Single Nucleotide Polymorphisms; SNP). Исключением является *M. canettii* и другие «гладкие» микобактерии (образуют гладкие колонии при культивировании) (Sreevatsan *et al.*, 1997; Cole *et al.*, 1998; Fleischmann *et al.*, 2002; Gutacker *et al.*, 2002). Следует отметить, что для данных патогенов характерна выраженная, хоть и не абсолютная, специфичность в выборе макроорганизма-хозяина. Так, например, *M. tuberculosis*, *M. canettii* и *M. africanum* наиболее часто являются возбудителями туберкулеза человека, но известны случаи передачи *M. tuberculosis* приматам и крупному рогатому скоту (Vervenne *et al.*, 2004; Осепек *et al.*, 2005). *M. microti* и *M. pinnipedii* вызывают заболевание у грызунов и морских львов, соответственно, но также в редких случаях могут быть причиной туберкулеза у людей (Kiers *et al.*, 2008; Panteix *et al.*, 2010). *M. bovis* и *M. caprae* обладают более широким кругом хозяев и способны инфицировать как крупный рогатый скот, так и людей (Kubica *et al.*, 2003). Тем самым представленные виды могут быть рассмотрены как экотипы одного генетического вида, эволюционировавшие вследствие адаптации к разным макроорганизмам-хозяевам (Smith *et al.*, 2006; Djelouadji *et al.*, 2011).

В ходе реконструкции эволюционных событий, произошедших с микобактериями туберкулезного комплекса, было выдвинуто предположение, что его члены являются клональными потомками единого успешного предка, образовавшегося в результате эволюционного эффекта «бутылочного горлышка» 35000–20000 лет назад (Sreevatsan *et al.*, 1997; Gutacker *et al.*, 2002; Hughes *et al.*, 2002). При этом природа и географические рамки предшествовавшего

бактериального пула долго оставались невыясненными. В дальнейшем было определено, что родственные виды микобактерий претерпели эволюцию путем делеций крупных фрагментов генома (от англ. Large Sequence Polymorphisms; LSP) в так называемых регионах различия (от англ. Region of Difference; RD), что привело к возникновению представителей комплекса из единого гипотетического вида-предшественника, позднее названного *M. prototuberculosis* (Brosch *et al.*, 2002; Gutierrez *et al.*, 2005). Представленные информативные маркеры, делеции, были выявлены на основе анализа результатов сравнительной гибридизации полных геномов и являются однонаправленными в случае микобактерий туберкулезного комплекса (Behr *et al.*, 1999; Mostowy *et al.*, 2004; Tsolaki *et al.*, 2004; Tsolaki *et al.*, 2005; Azhikina *et al.*, 2006; Gutacker *et al.*, 2006). Обобщенные данные по исследованию необратимых хромосомных делеций и анализ однонуклеотидных полиморфизмов позволили исследователям определить наиболее вероятную схему эволюции представителей комплекса (Рисунок 2). В частности, была показана ошибочность теории о происхождении *M. tuberculosis* от *M. bovis* в ходе одомашнивания крупного рогатого скота (Stead, 1997). Напротив, было показано, что в геноме *M. bovis* произошла серия однонаправленных крупных делеций независимо от *M. tuberculosis*, обособивших его от общего предка. Дополнительно было выявлено существование «древних» и «современных» штаммов *M. tuberculosis*, а также определено, что *M. canettii*, редкий вид с необычным фенотипом (van Soolingen *et al.*, 1997), может представлять наиболее древнюю линию внутри микобактерий туберкулезного комплекса (Fabre *et al.*, 2004), не подвергшуюся эффекту «бутылочного горлышка».

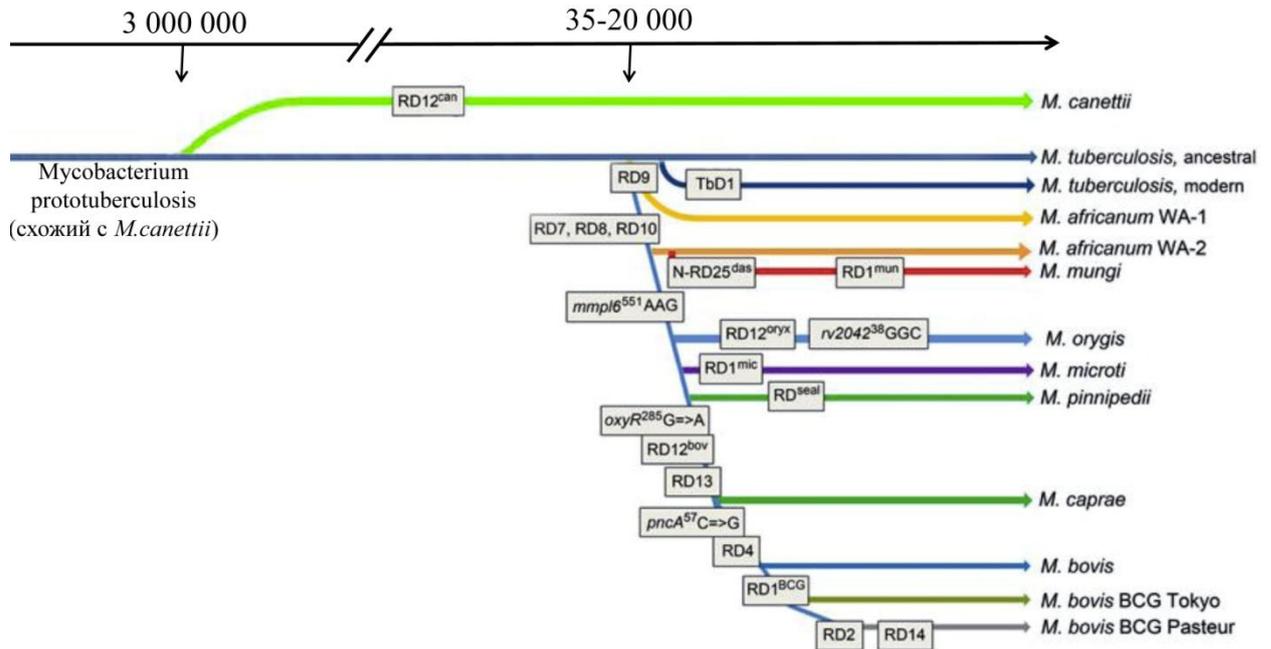


Рисунок 2. Схема эволюционных событий для микобактерий туберкулезного комплекса на основе различных информативных маркеров (делеции, SNP). Маркеры в прямоугольниках включают делетированные регионы RD и SNP (адаптировано из (Brosch *et al.*, 2002; Ernst *et al.*, 2007; Djelouadji *et al.*, 2011; van Ingen *et al.*, 2012)).

Для более глубокого изучения эволюционных событий, а также роли горизонтального переноса генов в формировании генетического разнообразия *M. tuberculosis*, Gutierrez с соавторами (Gutierrez *et al.*, 2005) исследовали ряд генов «домашнего хозяйства» (от англ. Housekeeping genes) среди представителей комплекса. Авторы показали, что, несмотря на гомогенность, геном *M. tuberculosis* представляет собой результат множественных событий генетического переноса, предшествовавших клональной экспансии вида. Также было установлено, что популяционная структура МТБ является малой частью более разнообразного предкового вида, современные представители которого состоят из атипичных изолятов возбудителя туберкулеза человека в Восточной Африке (*M. canettii* и другие «гладкие» микобактерии). Проведенный филогенетический анализ показал, что члены комплекса образуют компактную ветвь внутри разветвленной сети, сформированной различными представителями «гладких» микобактерий (Рисунок 3). Причем коллекция из нескольких десятков «гладких»

микобактерий характеризуется большим генетическим разнообразием, чем мировая популяция микобактерий туберкулезного комплекса.

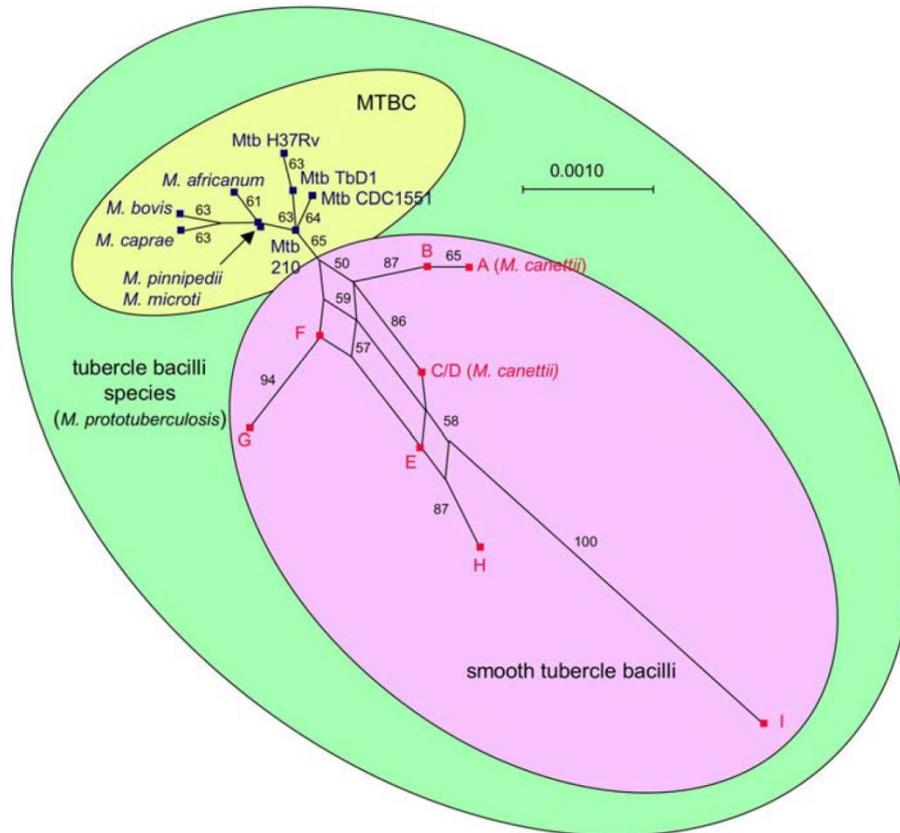


Рисунок 3. Расщепляемый-граф (от англ. Split-graph) на основе анализа 17 участков 6 генов «домашнего хозяйства» (housekeeping genes). Узлы представляют отдельные штаммы микобактерий туберкулезного комплекса (красные точки – штаммы с «гладкими» колониями, синие – остальные штаммы микобактерий туберкулезного комплекса). Цифры на ветвях показывают значения бутстрэп-анализа (от англ. bootstrap analysis) в %, на основе 1000 повторностей (Gutierrez *et al.*, 2005).

Следует отметить, что сетевая структура филогении для «гладких» микобактерий свидетельствует о возможности рекомбинации среди штаммов. Ярким примером внутривидового горизонтального переноса генов послужило выявление мозаичности в последовательности генов *gyrA* и *gyrB* «гладких» микобактерий. При этом анализ этих же генов среди других членов комплекса не выявил рекомбинации, что соответствует и предыдущим публикациям (12-14). В ходе исследования была выдвинута гипотеза, что микобактерии туберкулезного комплекса являются успешной клональной субпопуляцией, эволюционировавшей

из гораздо более древнего и крупного бактериального вида (*M. prototuberculosis*), включающего в частности *M. canettii* и другие «гладкие» варианты. Таким образом, представленные исследования привели к отказу от гипотезы о «недавнем» происхождении патогена (Sreevatsan *et al.*, 1997) и определили его возраст в 3 млн. лет. Неким подтверждением выдвинутого сценария является то, что почти все «гладкие» штаммы были выявлены в восточной Африке (Джибути), в регионе присутствия ранних гоминидов 3 млн. лет назад (Semaw *et al.*, 2005).

Подтверждением гипотезы о древнем происхождении возбудителя туберкулеза послужило определение в 2013 году полных геномных последовательностей 5 штаммов «гладких» микобактерий (Supply *et al.*, 2013). Анализ секвенированных геномов выявил множественные рекомбинационные события, происходящие внутри штаммов *M. canettii*. Около 10 % белок-кодирующих последовательностей изучаемых геномов имели мозаичное строение. Также было определено, что геномы секвенированных образцов на 10-115 т.п.о. больше по сравнению с геномами других микобактерий туберкулезного комплекса, что согласуется с данными об эволюции патогена путем крупных делеций. Анализ однонуклеотидных полиморфизмов выявил большое разнообразие «гладких» микобактерий. Количество SNP среди 5 секвенированных образцов было в среднем в 25 раз больше, чем количество полиморфизмов среди всех оставшихся членов микобактерий туберкулезного комплекса. Таким образом, P. Supply с соавторами (Supply *et al.*, 2013) было показано, что представители вида *M. canettii* отделились от последнего общего предка задолго до клональной экспансии штаммов туберкулезного комплекса.

1.4 Популяционная структура *M. tuberculosis*

В последние годы молекулярно-генетические методы типирования патогенов становятся рутинной практикой. Исследования, использующие данные методы, можно разделить на несколько областей: 1) «классическая» молекулярная эпидемиология, 2) изучение филогении и эволюции, 3) классификация штаммов. Следует отметить, что существующие методы генотипирования далеко не всегда применимы для решения всех поставленных задач одновременно. Так, например,

для молекулярно-эпидемиологических исследований, направленных на изучение вопросов и закономерностей развития эпидемий туберкулезной инфекции, нахождение отличий между эндогенной реактивацией и суперинфекцией при рецидиве туберкулеза, выявление смешанных инфекций или лабораторных кросс-контаминацией, необходимо использовать методы с высокой дискриминирующей способностью. Наиболее часто для этого применяются анализ полиморфизма длины рестрикционных фрагментов *IS6110* (от англ. Restriction Fragment Length Polymorphisms *IS6110*; *IS6110* RFLP) и анализ числа тандемных повторов в различных локусах генома (от англ. Variable Number of Tandem Repeats; VNTR). Подход *IS6110* RFLP-типирования был предложен van Embden с соавторами в 1993 году (van Embden *et al.*, 1993). В основе метода лежит определение количества и локализации в геноме *M. tuberculosis* мобильного генетического элемента – инсерционной последовательности *IS6110* (около 1350 п.о.), имеющей сайт рестрикции *PvuII* (Thierry *et al.*, 1990). Данный элемент распределен в геноме *M. tuberculosis* случайным образом, а его число может варьировать от 0 до 26 копий (McHugh and Gillespie, 1998; McEvoy *et al.*, 2007). Следует отметить, что представленный метод, до недавнего времени считавшийся «золотым стандартом» молекулярно-генетического типирования *M. tuberculosis*, обладает рядом серьезных недостатков. Методика постановки достаточно сложная и длительная, требует большого количества исходной геномной ДНК (и, следовательно, бактериальной массы), процесс учета результатов тяжело поддается стандартизации. Кроме того, разрешающая способность метода недостаточна для типирования штаммов с малым числом фрагментов *IS6110* в геноме (McEvoy *et al.*, 2007). Перечисленные выше недостатки затрудняют широкое применение представленного метода в лабораторной практике. Другой способ типирования, VNTR-анализ, основан на амплификационной технологии и базируется на оценке длины продукта полимеразной цепной реакции (ПЦР) (Supply *et al.*, 2001; Supply *et al.*, 2006). Данный метод в последние годы все чаще используется в качестве основного метода генотипирования. В межгенных участках генома *M. tuberculosis* на сегодняшний день идентифицировано

несколько десятков локусов, содержащих различное количество тандемных повторов (VNTR-локусов) – структур, представляющих собой идентичные последовательности геномной ДНК, длиной от 40 до 200 п.о. (Frothingham and Meeker-O'Connell, 1998; Supply *et al.*, 2000; Roring *et al.*, 2002; Skuce *et al.*, 2002). Первоначально метод был основан на анализе 6 локусов ETR (так называемых точных тандемных повторов, от англ. Exact Tandem Repeats) (Frothingham and Meeker-O'Connell, 1998), дальнейшее его развитие включило 12 локусов MIRU (микобактериальные рассеянные повторяющиеся единицы, от англ. Mycobacterial Interspersed Repeat Units) (Supply *et al.*, 2001), которых все же было недостаточно для полноценной дискриминации штаммов (Supply *et al.*, 2001; Kwara *et al.*, 2003; Surikova *et al.*, 2005). Последний предложенный формат для максимальной дифференциации штаммов включает 24 VNTR-локуса, 15 из которых составляют дискриминирующий набор (Supply *et al.*, 2006; Oelemann *et al.*, 2007).

Представленные методики, основанные на анализе мобильных (IS6110 RFLP) или повторяющихся элементов генома (VNTR-анализ), обладают высокой дискриминирующей силой для дифференциации штаммов и могут быть успешно использованы для генотипирования взаимосвязанных образцов. Однако слишком быстрая эволюция этих маркеров в редких случаях может приводить к возникновению схожих профилей у неродственных штаммов в результате гомоплазии (Filliol *et al.*, 2006; Gutacker *et al.*, 2006; Schurch and van Soolingen, 2012), что затрудняет их использование для филогенетического анализа и штаммовой классификации (Hirsh *et al.*, 2004; Monot *et al.*, 2005; Comas *et al.*, 2009).

Другой широко применяемой методикой, используемой для молекулярной эпидемиологии и исследования эволюционных взаимосвязей, является сполиготипирование (spoligotyping, от англ. Spacer Oligonucleotide Typing). Метод показывает меньшую дискриминирующую способность по сравнению с IS6110 RFLP-типированием и VNTR-анализом, и в связи с этим используется чаще как вспомогательный для эпидемиологических исследований. Сполиготипирование основано на определении структуры DR (от англ. Direct Repeat, прямые повторы)

региона и представляет собой эффективную, воспроизводимую и стандартизированную технологию молекулярно-генетического типирования МБТ (Kamerbeek *et al.*, 1997). Структурно DR регион, относящийся к семейству CRISPR (от англ. Clustered Regularly Interspaced Palindromic Repeats, кластеризованные короткие палиндромные повторы, разделенные спейсерными участками), представляет собой уникальные последовательности (спейсеры) (от 35 до 41 п.о.), разграниченные повторяющимися участками (36 п.о.). Впервые повторы такого рода были найдены в 1987 г. в ходе изучения гена *iap* *E. coli* (Ishino *et al.*, 1987). В последующем CRISPR система была найдена у *M. tuberculosis* (Hermans *et al.*, 1991), *Haloferax mediterranei* (Mojica *et al.*, 1995), *Methanocaldococcus jannaschii* (Bult *et al.*, 1996) и в других бактериях и археях. При этом сам термин CRISPR был введен в 2002 году Jansen с соавторами для обозначения специфической структуры этого локуса (Jansen *et al.*, 2002). Данные компьютерного анализа показывают, что CRISPR встречается у 40 % бактерий и 90 % архей (Kunin *et al.*, 2007). Функция этой системы оставалась не ясной до недавнего времени. В 2005 году 3 независимые исследовательские группы обнаружили, что спейсерные последовательности содержат участки ДНК плазмид или фагов, и тем самым, возможно, играют роль в CRISPR опосредованном иммунитете против внехромосомных агентов (Bolotin *et al.*, 2005; Mojica *et al.*, 2005; Pourcel *et al.*, 2005). В процессе перемещения IS-элементов, а также в ходе гомологической рекомбинации, структура DR региона может изменяться, что в случае МБТ может быть зафиксировано путем гибридизации специфических олигонуклеотидов, нанесенных на мембрану, с уникальными спейсерными последовательностями (Kamerbeek *et al.*, 1997). Представленный метод позволяет получать специфические для штаммов профили, которые могут быть представлены в числовом формате и использоваться для сравнения между лабораториями. На сегодняшний день наиболее крупными международными информационными базами являются SpolDB4 (Brudey *et al.*, 2006) и SITVIT (Demay *et al.*, 2012), которые содержат результаты сполиготипирования нескольких десятков тысяч штаммов *M. tuberculosis*. Таким образом, результаты,

полученные в разных лабораториях, могут быть соотнесены друг с другом на основании общей номенклатуры (Gori *et al.*, 2005). Основываясь на встречаемости определенных паттернов и их характерном профиле, на основании представленной методики удалось определить многие линии и генетические семейства МБТ. Так, к примеру, представители генетического семейства Beijing показывают типичный споллигопаттерн, характеризующийся отсутствием сигналов между 1 и 34 спейсерной последовательностью (Bifani *et al.*, 2002). Однако и в случае споллиготипирования, использование DR региона в качестве маркера в редких случаях может быть затруднено из-за гомоплазии. В недавнем исследовании Fenner с соавторами было показано существование «псевдо-Beijing» клинических изолятов микобактерий туберкулеза, имеющих характерный для Beijing споллигопаттерн, но относящихся к другим филогенетическим линиям (Fenner *et al.*, 2011).

Двумя другими маркерами, используемыми непосредственно для филогенетического анализа и штаммовой классификации, являются частично описанные ранее крупные делеции и однонуклеотидные полиморфизмы. Представленные маркеры в случае генетически мономорфного микроорганизма МБТ представляются уникальными и однонаправленными. SNP выступают в качестве филогенетически информативных мутаций, поскольку низкий уровень ДНК полиморфизма в целом у *M. tuberculosis* делает независимую обратную мутацию практически невероятной. В свою очередь отсутствие горизонтального переноса генов, в случае LSP, исключает восстановление крупных участков генома, утраченных в процессе эволюции. В ходе многочисленных исследований было выявлено большое количество геномных регионов, отсутствующих в единственном к тому времени референтном геноме H37Rv, что послужило основой для определения эволюционной истории микобактерий туберкулезного комплекса и набора дискретных филогенетических линий (Behr *et al.*, 1999; Brosch *et al.*, 2002; Hirsh *et al.*, 2004; Mostowy *et al.*, 2004; Mostowy *et al.*, 2004; Tsolaki *et al.*, 2005; Gagneux *et al.*, 2006). Однако следует отметить, что представленный маркер используется скорее для построения кладограмм (от англ.

cladogram), а не для определения эволюционных дистанций, так как геномные делеции носят случайный характер и для них не разработано эволюционных моделей. Более того, одним из ограничений использования геномных делеций для филогенетических исследований является тот факт, что данные маркеры обычно определяются на основании одностороннего сравнения с референтным геномом H37Rv.

Типирование на основе однонуклеотидных полиморфизмов в последние годы все чаще используется для генотипирования микобактерий туберкулезного комплекса. Данный маркер является практически идеальным для классификации штаммов и отнесения их к тем или иным филогенетическим линиям. При этом полиморфизмы могут быть идентифицированы как путем сравнения *in silico* опубликованных полных геномов штаммов *M. tuberculosis* (Fleischmann *et al.*, 2002; Gutacker *et al.*, 2002; Alland *et al.*, 2003; Garnier *et al.*, 2003; Baker *et al.*, 2004), так и *de novo* анализом (Dos Vultos *et al.*, 2008; Hershberg *et al.*, 2008), когда первоначально выбираются гены, а потом в них проводится поиск SNP. Первый вариант является менее предпочтительным, так как использование полиморфизмов, подобранных на основе сравнения полных геномов только нескольких штаммов, причем, в основном, одной линии, содержит риск получения непредставительной коллекции маркеров, и как результат коллапсирование филогенетических ветвей (Alland *et al.*, 2003; Pearson *et al.*, 2004; Achtman, 2008; Smith *et al.*, 2009). Использование SNP маркеров, идентифицированных *de novo*, представляется наиболее удачным вариантом для определения филогенетических взаимоотношений между штаммами, а в некоторых случаях и для молекулярной эпидемиологии. Здесь следует отметить, что в последние годы для построения углубленных филогений все чаще применяется полногеномное секвенирование с последующим анализом однонуклеотидных полиморфизмов. Публикации последних лет также подчеркивают, что секвенирование геномов становится своего рода новым «золотым стандартом» для исследования молекулярной эпидемиологии. Так, например, показано, что оно обладает значительно большей дискриминирующей

способностью, чем любые из стандартных методов генотипирования, и что образцы, имеющие одинаковые генетические профили, могут иметь существенные отличия на геномном уровне, что в дальнейшем оказывается важным для интерпретации паттернов трансмиссии (Niemann *et al.*, 2009; Gardy *et al.*, 2011; Casali *et al.*, 2012) или для изучения смешанных инфекций (Comas *et al.*, 2011; Saunders *et al.*, 2011). В случае филогенетического анализа сравнительная геномика позволяет установить степень неоднородности популяции и определить генетические дистанции. Основываясь на недавно опубликованных исследованиях геномных последовательностей было выделено шесть основных филогенетических линий для микобактерий туберкулезного комплекса, ассоциированных с туберкулезом человека (Comas *et al.*, 2010; Bentley *et al.*, 2012) (Рисунок 4). Четыре линии относятся к *M. tuberculosis*, а две к *M. africanum*. Как упоминалось ранее, анализ геномных данных позволяет вычислить эволюционные дистанции. Было установлено, что на уровне геномов отличия между линиями составляют в среднем 2000 SNPs, что эквивалентно, к примеру, эволюционной дистанции между *M. tuberculosis* и *M. bovis* (Garnier *et al.*, 2003), а разнообразие мировой популяции МБТ в целом выше, чем разнообразие всех остальных представителей комплекса при сравнении друг с другом (без учета *M. canettii* и других «гладких» микобактерий).

В заключении следует отметить, что использование филогенетически информативных и надежных маркеров (большое количество SNPs и LSPs) при выраженной клональности *M. tuberculosis* ведет к выявлению схожих филогений (Рисунок 4).

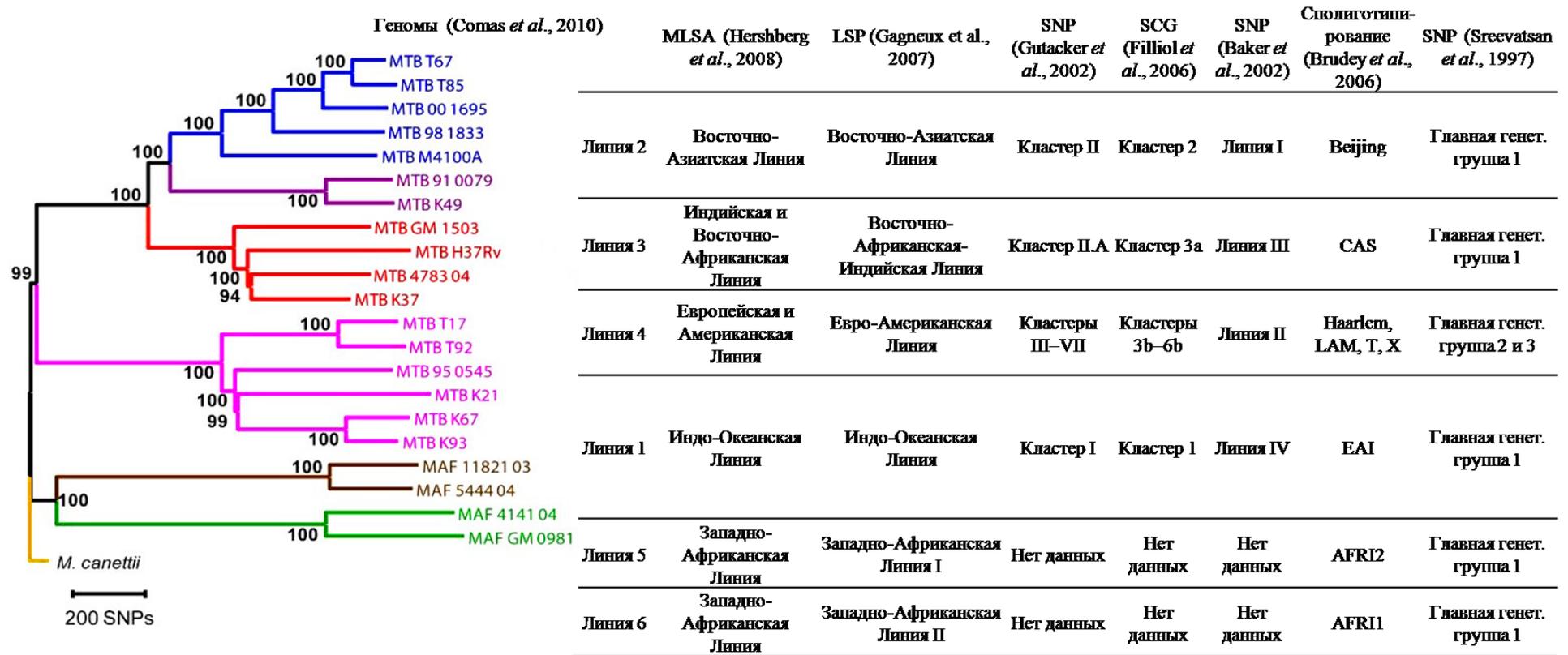


Рисунок 4. Сравнение терминологии и молекулярных маркеров для определения основных линий внутри *M. tuberculosis* и *M. africanum*.

1.5 Молекулярная эпидемиология туберкулеза в России

В Российской Федерации (РФ) туберкулез остается одной из основных глобальных проблем здравоохранения. Несмотря на общую тенденцию к снижению заболеваемости впервые выявленными активными формами туберкулеза, ситуация остается чрезвычайно напряженной. В 2012 году заболеваемость туберкулезом составила 68 человек на 100 тыс. населения, а всего за тот год в РФ было зарегистрировано около 100 тыс. новых случаев заболевания (Нечаева, 2013).

По оценкам Всемирной Организации Здравоохранения (ВОЗ) Россия является одной из 22 стран с наибольшим бременем туберкулеза. В стране регистрируется более трети всех новых случаев туберкулеза, зарегистрированных в Европейском регионе, причем смертности от туберкулеза превышает показатели в странах Европы в 5–8 раз (Туберкулез в Российской Федерации, 2013).

Для решения задач эпидемиологического мониторинга заболевания в последние годы все чаще используются методы молекулярно-генетического типирования патогена. Полученные данные позволяют изучать динамику и закономерности географического распространения туберкулеза, как в популяции в целом, так и среди определенных групп населения.

Показано, что в структуре популяции возбудителя туберкулеза в России преобладают штаммы генетического семейства Beijing (от 50 % до 80 %) (Mokrousov *et al.*, 2003; Норкина, 2003). На основании VNTR-анализа представители данного генотипа могут быть разделены на несколько типов (Surikova *et al.*, 2005; Mokrousov *et al.*, 2008). Наиболее представленными являются типы M2 и M11, составляя около 88 % всех изолятов генетического семейства Beijing, циркулирующих на территории азиатской части РФ (Drobniewski *et al.*, 2005), и 77 % — в европейской части РФ (Surikova *et al.*, 2005), при этом частота встречаемости данных типов варьируется в зависимости от региона. В Центральном, Северо-Западном, Сибирском и Приволжском федеральных округах (ФО) преобладает тип M2, составляя от

50 % до 60 % популяции Beijing (Медведева, 2004; Nikolayevskyy *et al.*, 2006; Mokrousov *et al.*, 2012) В Уральском ФО, по результатам исследования Ковалева с соавторами, наиболее представленным типом является M11 (56 %), в то время как вариант M2 выявляется чуть более чем в 20 % случаев (Kovalev *et al.*, 2005). По данным Мокроусова с соавторами, тип M11 является предковым по отношению к M2, и высокая частота встречаемости данного варианта подразумевает исторически отдаленное время его первоначального проникновения в Россию. На основании одной из последних гипотез, тип был занесен на территорию России в XIII-XV веках во времена экспансии Монгольской империи Чингисхана. При этом считается, что он зародился в северных областях Китая более 2000 лет назад, что, вероятно, является минимальной оценкой возраста представителей генотипа Beijing (Mokrousov, 2008). Обобщенные данные по встречаемости типов генетического семейства Beijing и генотипа в целом на территории России представлены в таблицах 1 и 2.

Другим часто встречаемым на территории России генотипом является LAM (от англ. Latin American and Mediterranean, латино-американско-средиземноморское семейство). Частота встречаемости этого семейства, так же как и для описанного выше генетического семейства Beijing, варьируется в зависимости от региона. Наиболее часто (от 10 % до 45 % случаев) представителей генотипа выявляют в Центральном регионе РФ (Шемякин, 2003). В Сибирском ФО этот показатель варьирует от 8 % до 17 % (Думова *et al.*, 2011), в Северо-Западном – около 9 % (Mokrousov *et al.*, 2009). Brudey с соавторами отмечают, что данное семейство демонстрирует большую способность к трансмиссии и к широкому распространению по всему миру, после семейства Beijing (Brudey *et al.*, 2004). По данным литературы семейство Ural является еще одним распространенным семейством на территории России. Данный генотип, также как и генотип LAM, относится к Евро-Американской линии. В России представители этого генетического семейства были впервые выявлены на основании 12-локусного VNTR-анализа образцов из Екатеринбургской области с частотой встречаемости

около 15 % (Kovalev *et al.*, 2005). Для этого генотипа характерно отсутствие сигналов по спейсерам 29-31 и 33-36, а также наличие одной копии в локусе MIRU26. В европейской части России данное генетическое семейство встречается с частотой 4-5 % (Нарвская с соавт., 2002; Mokrousov *et al.*, 2009; Afanas'ev *et al.*, 2011). Немного чаще генотип выявляется в Восточной Сибири (8–9 %) (Ogarkov *et al.*, 2012), и на среднем Урале (7 %) (Umpeleva, 2010). Возможной областью зарождения представителей генотипа считаются страны северной и северо-восточной частей Черного моря, где данный генотип выявляется с частотой до 15 % (Niemann *et al.*, 2010). Обобщенные данные по встречаемости представителей генетических семейств LAM и Ural представлены в таблице 2. Среди прочих генотипов на территории России встречаются изоляты генотипов T (превалирует T1), Haarlem, T1_RUS и MANU2. В среднем частота встречаемости последних не превышает 9 % от общей популяции патогена.

В заключение, следует отметить, что на территории России преобладают представители генетического семейства Beijing, характеризующиеся высокой степенью вирулентности. В связи с этим мы считаем необходимым дать развернутое описание данного генотипа в следующем разделе.

Таблица 1.
Географическое распространение основных типов генотипа Beijing на территории России

Регион, количество штаммов	Тип, количество и (%) штаммов от общей популяции Beijing в данном регионе									источник
	M1	M2	M8	M11	M12	M16	M33	M44	M122	
Псков (Северо- Западный ФО), 40		13 (32.5)	2 (5)	11 (27.5)	1 (2.5)					Mokrousov <i>et al.</i> , 2012
Центральный ФО, 129		75 (58.1)		31 (24.0)	2 (1.6)				3 (2.3)	Nikolayevskyy <i>et al.</i> , 2006
Уральский ФО, 50		11 (22.0)	1 (2.0)	28 (56)	3 (6.0)		2 (4.0)			Kovalev <i>et al.</i> , 2005
Восточная Сибирь, 113		57 (50.4)		29 (25.7)		1 (0.9)	1 (0.9)	2 (1.8)		Медведева с соавт., 2004
Сибирский ФО, 50		15 (30)		9 (17)			4 (8)			Dymova <i>et al.</i> , 2011
Приволжский ФО, 113		68 (60)		31 (27)						Drobniowski <i>et al.</i> , 2005

Таблица 2. Географическое распространение основных генетических семейств МБТ на территории России

Регион	Годы	Число штаммов	Генетическое семейство, встречаемость в выборке				Источник
			Beijing	LAM	Ural	Другие	
Сибирский ФО							
Иркутск	-	72	70 %				Медведева с соавт., 2004
Иркутск	2007–2009	101	51 %				Ogarkov <i>et al.</i> , 2012
Новосибирск	-	99	51 %				Surikova <i>et al.</i> , 2005
Новосибирская обл.	-	106	47 %	17 %	9 %		Dumova <i>et al.</i> , 2011
Респ. Тува	-	32	50 %				Матрашкин с соавт., 2004
Уральский ФО							
Екатеринбургская обл.	2001–2002	92	54 %		15 %	Haarlem – 2 %	Kovalev <i>et al.</i> , 2005

-	2009–2011	178	55.7 %	13 %	12 %	Haarlem – 8 %	Умпелева, 2013
Приволжский ФО							
Самарская обл.	2001–2002	880	66.6 %				Drobniewski <i>et al.</i> , 2005
Самарская обл.	2008–2010	2348	72 %			T, Haarlem, LAM – 22 %	Casali <i>et al.</i> , 2012
Центральный ФО							
Тула, Серпухов	1998–2001	217 (МЛЮ)	42 %	45 %	10 %	T – 1.8 %	Lipin <i>et al.</i> , 2007
Тульская обл. (Озерск)	2001–2002	87	44 %	45 %	6 %	T1 – 1 %, T1_RUS – 1 %	Ignatova <i>et al.</i> , 2006
Москва	2005–2006	115	66 %	10 %	2 %	T – 8.7 %	Afanas'ev <i>et al.</i> , 2011
-	-	353	34 %	49.6 %			Иванов с соавт., 2004
Северо-Западный ФО							
Архангельская обл.	1998–1999	119	44.5 %				Toungoussova <i>et al.</i> , 2002
Мурманская обл., Архангельская обл., респ. Карелия, респ. Коми	2004–2006	176	47 %	6 %	11 %	T – 24 %, MANU2 – 5 %	Baranov <i>et al.</i> , 2009
Калининград	2006	90	45.6 %				Mokrousov <i>et al.</i> , 2009
Мурманск	2003–2004	1226	50 %	4–9 %	14 %	T2 – 6.7 %	Makinen <i>et al.</i> , 2011
Псков	2008–2009	90	44 %	20 %	6 %	T – 14 %, Haarlem – 11 %	Mokrousov <i>et al.</i> , 2012

1.6 Генетическое семейство Beijing *M. tuberculosis*

Изоляты генетического семейства Beijing впервые были обнаружены в 1992–1994 годах, в одноименном городе Пекин (Китай) (van Soolingen *et al.*, 1995). В ходе дальнейших исследований штаммы данного генотипа были выявлены на территории США, Восточной Азии, России и стран бывшего СССР (Bifani *et al.*, 2002; Toungousova *et al.*, 2002; Glynn *et al.*, 2005; Garcia de Viedma *et al.*, 2006; Ignatova *et al.*, 2006; Millet *et al.*, 2007). Согласно данным последних лет штаммы генотипа Beijing составляют 13 % от глобальной популяции изолятов *M. tuberculosis*, являясь единственным семейством, столь широко распространенными по всему миру (Parwati *et al.*, 2010).

Генотип Beijing имеет ряд характерных молекулярно-генетических признаков, отличающих его от других генетических семейств (Bifani *et al.*, 2002; Ferdinand *et al.*, 2004; Kremer *et al.*, 2004; Mokrousov, 2008). Изоляты семейства Beijing в соответствии с интернациональной базой данных SITVIT имеют сполиготип SIT 1 (от англ. Spoligotype International Type, международный сполиготип) определяющийся наличием 9 из 43 спейсеров (с 35 по 43) в DR локусе. Данный паттерн легко выявляется при стандартном сполиготипировании в виде характерного профиля гибридизации [8]. Согласно классификации Gagneux с соавторами семейство Beijing принадлежит к Восточно-Азиатской линии и характеризуется делециями определенных участков генома (Gagneux and Small, 2007). Внутри генотипа выделяют атипичные/древние и типичные/современные штаммы (Tsolaki *et al.*, 2004; Mokrousov *et al.*, 2005; Tsolaki *et al.*, 2005). На основании участков различия все штаммы генотипа характеризуются делецией региона RD207, в свою очередь для древних Beijing маркерными являются RD150 и RD142, а для современных RD105 и RD181 (Tsolaki *et al.*, 2004; Tsolaki *et al.*, 2005). Основные молекулярно-генетические особенности описываемого семейства приведены в таблице 3.

Таблица 3.

Молекулярно-генетические характеристики семейства Beijing (Hanekom *et al.*, 2011)

Общее с рядом других генотипов	Уникальное для генотипа Beijing
Главная генетическая группа 1	Сполиготип SIT 1
Интактный ген <i>pks15/1</i>	IS6110 интегрирован между генами <i>dnaA</i> и <i>dnaN</i>
линия современных МБТ (делетирован регион TbD1)	Наличие фрагмента длиной 3.5 т.п.о. после IS6110 RFLP-типирования, несущего повторяющийся элемент IS1081
<i>RvD2</i> (Mb1785–1787) и <i>RvD3</i> (MT1812–1813)	Отсутствие или 1–2 копии IS6110 в NTF регионе в древних и современных штаммах, соответственно
SNP в 507 кодоне гена <i>fadD28</i>	Делеция региона RD207 и регионов RD105, RD181, RD150, RD142 (в зависимости от типа штамма)
	Делеция в гене <i>Rv0927c</i> и SNP в межгенном регионе <i>Rv0927c</i>
	Увеличенная экспрессия гена <i>Rv3130c</i> (DosR-контролируемый регулон МБТ)
	A191C мутация в гене <i>Rv2629</i> (ассоциирована с резистентностью к рифампицину)
	Делеция в гене <i>Rv0279c</i>
	Полиморфизмы в 3 генах <i>mut</i> : <i>mutT2</i> , <i>mutT4</i> и <i>ogt</i> (в зависимости от типа штамма)

В настоящее время уделяется все большее внимания изучению данного генетического семейства с клинической точки зрения, т.к. изоляты семейства Beijing демонстрируют важные свойства патогенности. В ходе многочисленных исследований доказана ассоциация генотипа Beijing с лекарственной устойчивостью (Bifani *et al.*, 2002; Mokrousov *et al.*, 2002; Toungousova *et al.*, 2002; Cox *et al.*, 2005; Hillemann *et al.*, 2005; Park *et al.*, 2005; Lipin *et al.*, 2007). По одной из гипотез данное обстоятельство объясняется наличием у изолятов семейства мутаций в так называемых генах мутаторов, связанных с системой репарации ДНК (Parwati *et al.*, 2010).

Согласно другой гипотезе появления устойчивых штаммов связано со специфичностью строения клеточной стенки (Reed *et al.*, 2007). В свою очередь при лечении туберкулеза, вызванного штаммами с повышенной вирулентностью, увеличивается доза и продолжительность приема лекарственных препаратов, что так же может вести к развитию лекарственной устойчивости (Parwati *et al.*, 2010). Сравнение свойств патогенности различных генетических семейств *M. tuberculosis*, в экспериментах *in vivo*, показало, что бактерии генотипа Beijing характеризуются большей вирулентностью (Dormans *et al.*, 2004). В работе Lopez с соавторами была продемонстрирована повышенная вирулентность штаммов Beijing при инфицировании мышей, вакцинированных БЦЖ (Бацилла Кальметта-Герена, от франц. Bacillus Calmette—Guérin) (Lopez *et al.*, 2003). В связи с этим сделано предположение, что длительная массовая вакцинация БЦЖ может быть одним из факторов, обеспечивающих отбор и распространение штаммов Beijing (Lasunskaja *et al.*, 2010; Parwati *et al.*, 2010). Изоляты семейства так же демонстрируют способность размножаться в макрофагах (Черноусова с соавт. 2008). В исследованиях Barczak с соавторами на мышинных моделях показано, что штаммы генотипа Beijing являлись доминирующими в смешанных инфекциях, даже если изначально при инфицировании их концентрация была меньше (Barczak *et al.*, 2005). Обобщенные данные об особенностях представителей данного генотипа, способствующих его глобальной диссеминации, представлены в таблице 4.

Таблица 4.

Клинические особенности изолятов *M tuberculosis* генотипа Beijing (Parwati *et al.*, 2010)

Свойство	Свидетельство	Последствие
«Ускользание» от БЦЖ вакцины	Вакцина в меньшей степени защищает от инфицирования современными штаммами Beijing (на животных моделях)	Более латентная инфекция; активное течение заболевания
Повышенный	Ассоциация штаммов	Замедленный или

уровень устойчивости к антибиотикам	генотипа Beijing с множественной лекарственной устойчивостью	отсутствие ответа на противотуберкулезную терапию
Повышенная вирулентность	Быстрое распространение патогена, уменьшенная выживаемость (на животных моделях)	Быстрая прогрессия заболевания, высокая бактериальная нагрузка
Высокая иммуногенность	Повышенный ответ провоспалительных цитокинов; некроз зараженных макрофагов	Образование большого количества каверн, большое количество бактерий в мокроте
Адаптация (ко-эволюция) к иммунной системе человека	Ассоциация туберкулеза, вызванного штаммами генотипа Beijing, с полиморфизмами в TLR2 и SLC11A1 во Вьетнаме и Индонезии, соответственно	Повышенная активность течения болезни

1.7 Молекулярные основы возникновения устойчивости к противотуберкулезным препаратам

В последние годы во многих странах мира наблюдается повсеместное распространение лекарственно-устойчивых штаммов *M. tuberculosis*. Согласно классификации ВОЗ резистентные штаммы делятся на четыре группы: монорезистентные (устойчивые к одному из противотуберкулезных препаратов), полирезистентные (устойчивые к двум и более противотуберкулезным препаратам, но не к рифампицину и изониазиду одновременно), штаммы с множественной лекарственной устойчивостью (МЛУ – устойчивые, как минимум, к рифампицину и изониазиду) и штаммы с широкой лекарственной устойчивостью (ШЛУ – устойчивые к изониазиду и рифампицину, одному из фторхинолонов, и, по крайней мере, к одному из

трех инъекционных препаратов второго ряда – канамицину, капреомицину или амикацину) (WHO, 2006). Также в последние годы выделяют тотально устойчивые штаммы, т.е. устойчивые ко всем известным противотуберкулезным препаратам (Migliori *et al.*, 2007).

Феномен лекарственной устойчивости МБТ был обнаружен вскоре после начала применения стрептомицина, первого антибиотика, используемого в виде монотерапии заболевания (Crofton and Mitchison, 1948). Препарат был открыт в 1944 году Зельманом Ваксманом и Альбертом Шацем (Schatz *et al.*, 2005), а в дальнейшем, в «золотые годы антибиотиков» (1940–1960 г), были разработаны и остальные наиболее часто используемые противотуберкулезные препараты (ПТП). Применяемые на сегодняшний день ПТП можно разделить на две группы. К препаратам первой группы относят рифампицин (RIF), изониазид (INH), этамбутол (EMB), пиразинамид (PZA) и стрептомицин (STR). К препаратам второй группы относят канамицин (KAN), амикацин (AMK), капреомицин (CAP), этионамид (ETH), циклосерин (CS), пара-аминосалицилат натрия (PAS) и фторхинолоны (FQ). Основные исторические этапы в открытии противотуберкулезных препаратов и изменения в схемах лечения представлены на рисунке 5.

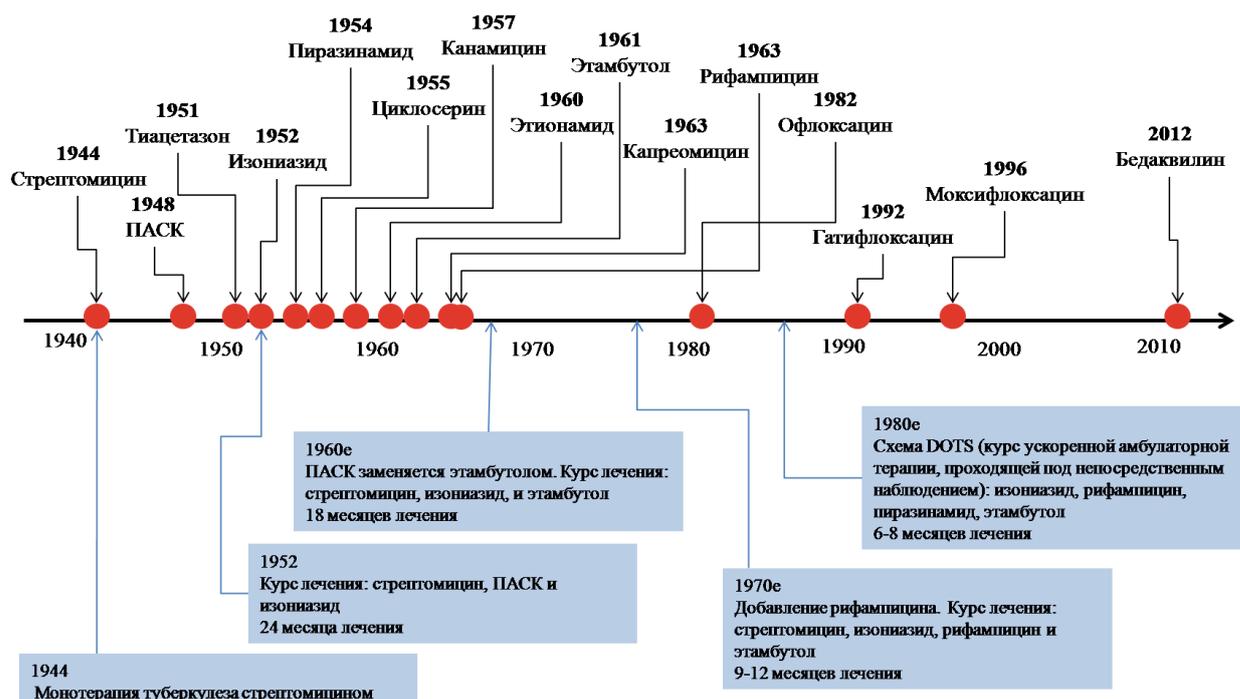


Рисунок 5. История открытия антибактериальных препаратов и развитие схем лечения туберкулеза (адаптировано из (Ma *et al.*, 2010)).

В общем, механизмы устойчивости к антибактериальным препаратам могут носить как естественный, не связанный с генетическими изменениями, так и приобретенный характер, связанный с появлением мутаций или поступлением нового генетического материала. Стоит отметить, что у *M. tuberculosis* реализуются оба вышеперечисленных механизма. К естественным механизмам устойчивости у МБТ можно отнести мощную гидрофобную клеточную стенку, затрудняющую проникновение внутрь клетки различных веществ, в том числе и антибактериальных, наличие систем активного транспорта, систем ферментативной деградации и модификации антибиотиков (Chambers *et al.*, 1995; Danilchanka *et al.*, 2008). Однако наибольшую клиническую значимость имеет приобретенная устойчивость, которая у МБТ наиболее часто связана с точечными мутациями, приводящими к модификации мишеней действия лекарственного препарата или нарушению процесса активации препарата (Zhang, 2005; Almeida Da Silva and Palomino, 2011).

1.7.1 Механизм возникновения лекарственно-устойчивых штаммов

Механизм возникновения лекарственно-устойчивых штаммов условно можно разделить на несколько стадий (zur Wiesch *et al.*, 2011). На первой стадии появление генетических мутантов, обладающих устойчивостью к лекарственным препаратам, происходит спонтанно, и отдельные лекарственно-устойчивые микобактерии обычно присутствуют в составе нормальной популяции бактерий, которые никогда не подвергались воздействию противотуберкулезных препаратов. При этом абсолютное количество лекарственно-устойчивых мутантных клеток в популяции дикого типа зависит от происхождения штамма и в значительной мере от общего количества бактерий. Зависимости количества мутантных клеток, устойчивых к тому или иному противотуберкулезному препарату, от общего

количества клеток в популяции были достаточно хорошо изучены в 60–70х годах прошлого века. Так, например, было показано, что количество клеток микобактерий туберкулеза, устойчивых к изониазиду, равно 1 на 10^6 , а к рифампицину – 1 на 10^8 . Последнее обстоятельство позволяет использовать определение устойчивости к рифампицину как суррогатный маркер множественной лекарственной устойчивости. При этом предполагается, что при устойчивости к рифампицину вероятность устойчивости к изониазиду очень высока (Gillespie, 2002). Как упоминалось ранее, количество мутантных клеток может напрямую зависеть от происхождения штамма, т.е. от той филогенетической линии, к которой он принадлежит. Работы последних лет показывают, что частоты мутаций, и тем самым вероятность возникновения лекарственно-устойчивых штаммов различаются среди представителей разных филогенетических семейств. Так в работе Ford с соавторами было проведено сравнение частот возникновения устойчивости к рифампицину среди микобактерий филогенетической линии 2 (включает изоляты Beijing семейства) и линии 4 (Ford *et al.*, 2013). Было показано, что устойчивость к рифампицину возникает на порядок чаще у штаммов линии 2 и составляет 10^{-7} – 10^{-8} против 10^{-8} – 10^{-9} для штаммов четвертой линии. Авторами было установлено, что данное обстоятельство связано с более высокой частотой мутаций для штаммов линии 2, что в свою очередь может быть частично объяснено наличием относительно большого количества мутаций в системе репликации, репарации и рекомбинации (Dos Vultos *et al.*, 2008). Также следует отметить, что происхождение штамма, а вместе с ним и генетический контекст на геномном уровне, могут влиять на тип возникающих мутаций, ассоциированных с лекарственной устойчивостью. Так, например, было показано, что для представителей филогенетической линии 1 мутации, ассоциированные с устойчивостью к изониазиду, статистически наиболее часто встречаются в промоторной области гена *inhA*. При этом мутации в этой области связаны с высоким уровнем устойчивости. Напротив, для представителей других линий таких ассоциаций установлено

не было, а мутации в описанной области приводили к устойчивости низкого уровня (Fenner *et al.*, 2012).

На следующей стадии количество лекарственно-устойчивых микобактерий увеличивается до такого уровня, что вероятность их спонтанного элиминирования из популяции становится крайне низкой. При этом закрепление мутации в популяции будет зависеть от размера популяции и действия направленного отбора. На данной стадии преимущество нарастающей популяции обеспечивается единственным признаком – устойчивостью к лекарственному препарату, то есть могут отобраться штаммы со сниженной способностью к выживанию в других условиях или обладающие крайне низкой вирулентностью (Andersson and Levin, 1999; Andersson, 2006)[17, 18].

Третья стадия характеризуется активным увеличением численности устойчивых бактерий до размера популяции и носит детерминированный характер, который в большей степени зависит от влияния антибактериального препарата. Образовавшаяся популяция может вытеснить или сосуществовать с чувствительными штаммами. При этом полному вытеснению чувствительных штаммов может препятствовать обратная мутация в устойчивом штамме или если приобретение устойчивости сопровождается снижением приспособленности (англ. *fitness*) резистентных штаммов (Hastings *et al.*, 2002).

Как было отмечено ранее, приобретение детерминант резистентности довольно часто ассоциировано с уменьшением приспособленности в случае отсутствия антибактериального препарата. Степень проявления этого эффекта зависит от особенностей генетической организации конкретного штамма или филогенетической линии, а также может модулироваться путем приобретения дополнительных детерминант устойчивости или компенсаторных мутаций. Наиболее интересным положительным эпистатическим эффектом обладают компенсаторные мутации. Данные мутации увеличивают приспособленность организма, несущего

детерминанту устойчивости, но при этом оказывают негативный (или в лучшем случае нейтральный) эффект на организмы дикого типа. Такого рода компенсация описана для многих бактериальных видов и антибиотиков (Andersson and Hughes, 2010). Для МБТ наиболее известным примером является мутация в промоторной области гена *ahpC*, приводящая к увеличению транскрипции данного гена и тем самым компенсирующая потерю каталазно-пероксидазной (KatG) активности в изониазид-устойчивых штаммах. Белок KatG активирует изониазид, а также защищает бактерию от оксидативного стресса, вызванного клетками макроорганизмом. В случае мутаций в гене *katG* происходит уменьшение активности фермента. Увеличение выработки белка AhpC, кодирующего алкил-гидропероксид редуктазу, как считается, частично защищает такие бактерии от повреждений, вызванных активными формами кислорода (Sherman *et al.*, 1996). Однако данная мутация встречается довольно редко в клинических изолятах, что может отражать слабое влияние на распространение лекарственно-устойчивых штаммов (Gagneux *et al.*, 2006). Схожий компенсаторный механизм также был выявлен в гене 16S рибосомной РНК для микобактерий туберкулезного комплекса в случае развития устойчивости к аминогликозидам (Shcherbakov *et al.*, 2010). Однако описанные мутации также практически не встречаются в клинической практике (Georghiou *et al.*, 2012). Наиболее интересные результаты были получены в случае резистентности к рифампицину. Comas с соавторами описали набор компенсаторных мутаций в генах *rpoA* и *rpoC*, кодирующих альфа- и бета-субъединицы РНК-полимеразы (Comas *et al.*, 2011). В изученной выборке описанные мутации встречались только у рифампицин-устойчивых штаммов, несущих мутации в гене *rpoB*, и были ассоциированы с увеличением приспособленности *in vitro*. Более того, мутации были выявлены более чем у 30 % клинических изолятов с множественной лекарственной устойчивостью с тенденцией к увеличению встречаемости в странах с высоким бременем МЛУ форм туберкулеза. Последнее обстоятельство свидетельствует о том,

что компенсаторные мутации способствуют распространению лекарственно-устойчивых штаммов.

1.7.2 Формирование устойчивости к различным противотуберкулезным препаратам

Как было упомянуто выше, устойчивость к ПТП у микобактерий туберкулеза наиболее часто связана с точечными мутациями, приводящими к модификации мишеней действия лекарственного препарата или нарушению процесса его активации. В данном разделе кратко будут описаны наиболее изученные детерминанты устойчивости к основным препаратам (таблица 5).

Таблица 5.

Генетические локусы, связанные с возникновением устойчивости микобактерий к противотуберкулезным препаратам

Антибиотик	Механизм действия	Ген	Функция гена	Роль
препараты первой линии				
Изониазид	ингибирует биосинтез миколиевых кислот и другие метаболические процессы	<i>katG</i>	каталаза-пероксидаза	активатор пролекарства
		<i>oxyR-ahpC</i>	алкилгидроредуктаза	маркер резистентности
		<i>inhA</i>	еноил-АСР-редуктаза	мишень лекарства
		<i>kasA</i>	кетоацил-АСР-синтаза	
		<i>ndh</i>	NADH-гидролаза	модуляция активности
Рифампицин	ингибирует транскрипцию	<i>rpoB</i>	β -субъединица РНК полимеразы	мишень лекарства
Пиразинамид	ингибирует транс-транслокацию	<i>pncA</i>	пиразинамидаза	активатор пролекарства
		<i>rpsA</i>	рибосомный белок S1	мишень лекарства
Этамбутол	ингибирует синтез арабиногалактанов	<i>embCAB</i>	арабинозилтрансферазы	мишень лекарства
Стрептомицин	ингибирует трансляцию	<i>rpsL</i>	рибосомный белок S12	мишень лекарства
		<i>rrs</i>	16SpРНК	мишень лекарства
		<i>gidB</i>	метилтрансфераза 16SpРНК	модификация мишени
препараты второй линии				
Амикацин/ Канамицин/ Капреомицин	ингибирует трансляцию	<i>rrs</i>	16S pРНК	мишень лекарства
		<i>eis</i>	ацетилтрансфераза	модификация лекарства
Этионамид	ингибирует биосинтез миколиевых кислот	<i>inhA</i>	еноил-АСР-редуктаза	мишень лекарства
		<i>ethA</i>	флавопротеин-монооксигеназа	активация пролекарства

		<i>ethR</i>	репрессор транскрипции <i>ethA</i>	экспрессия активатора пролекарства
		<i>ndh</i>		
		<i>mshA</i>	NADH дегидрогеназа II	модуляция активности
			гликозилтрансфераза	активация пролекарства
Фторхинолоны	ингибирует ДНК-гиразы	<i>gyrA</i>	субъединица А ДНК-гиразы	мишень лекарства
		<i>gyrB</i>	субъединица В ДНК-гиразы	связывание лекарства

Ингибиторы синтеза клеточной стенки

Изониазид является одним из важнейших ПТП первой линии и широко применяется для лечения различных форм туберкулеза. Препарат является пролекарством. Активация происходит под действием фермента каталазы-пероксидазы (KatG) микобактерии и сопровождается образованием свободнорадикальных форм изоникотиновой кислоты и кислорода (Chouchane *et al.*, 2000). Мишенью действия препарата является NADH-зависимая-АСР-редуктаза (InhA) – основной фермент синтеза миколовой кислоты, основного компонента клеточной стенки МБТ. Формирование устойчивости к препарату связывают с мутациями в нескольких генетических локусах: *katG* (ген каталазы-пероксидазы KatG), *kasA* (ген β-кето-ацил-АСР-синтетазы), *ndh* (ген NADH-дегидрогеназы), регион *inhA* (ген NADH-зависимой-АСР-редуктазы InhA) и регион *ahpC* (ген алкил-гидропероксидредуктазы). Следует отметить, что наиболее часто мутации встречаются в гене *katG*. При этом замена в 315 кодоне белка (Ser315Thr) выявляется у подавляющего количества изониазид-устойчивых штаммов (от 60 % до 90 %, в зависимости от выборки) (Ramaswamy and Musser, 1998; van Soolingen *et al.*, 2000; Mokrousov *et al.*, 2002; Zhang *et al.*, 2005). Кроме мутаций в *katG*, к формированию устойчивости часто могут приводить изменения в регионе гена *inhA* (Banerjee *et al.*, 1994). Точечные замены в позициях -15 и -8 промоторного региона *mabA – inhA* оперона ведут к гиперпродукции белка – мишени.

Этионамид – препарат второй линии, сходный по структуре с изониазидом. Оба препарата являются пролекарствами, однако в случае с этионамидом активация происходит под действием фермента монооксигеназы EthA (Baulard *et al.*, 2000; DeBarber *et al.*, 2000). Основной мишенью действия препарата является фермент InhA (Banerjee *et al.*, 1994). Формирование резистентности к этионамиду связано с мутациями в генах *ethA* и *inhA* (Morlock *et al.*, 2003).

Этамбутол – препарат первой линии, оказывающий бактериостатическое действие на активно делящиеся микобактерии. Механизм действия основан на ингибировании полимеризации арабинана в арабиногалактан и липоарабиноманнан, что ведет к нарушению синтеза клеточной стенки (Mikusova *et al.*, 1995). Мишенью действия препарата является фермент арабинозилтрансфераза B (EmbB) (Telenti *et al.*, 1997). Совместно с белками EmbA и EmbC, представленный фермент входит в состав *embCAB* оперона, кодирующего трансмембранный белок. Около 65 % этамбутол-устойчивых клинических изолятов содержат мутации в *embCAB* опероне, причем наиболее часто мутации встречаются в 306 кодоне гена *embB* (Mokrousov *et al.*, 2002; Ahmad *et al.*, 2004; Isola *et al.*, 2005).

Ингибиторы синтеза нуклеиновых кислот

Рифампицин – препарат первой линии, оказывающий бактерицидное действие на активно делящиеся микобактерии, локализованные как внутриклеточно, так и внеклеточно (Jindani *et al.*, 1980). Механизм действия препарата основан на связывании с β -субъединицей бактериальной ДНК-зависимой РНК-, что приводит к блокированию процесса транскрипции. Установлено, что до 95–98 % рифампицин-устойчивых штаммов содержат мутации в гене *rpoB*, причем 90–95 % таких мутаций расположены в так называемом RRDR (от англ. Rifampin Resistance Determining Region, участок, определяющий устойчивость к рифампицину) регионе, между 426 и 452 кодонами гена (507-533 кодона по *E. coli*) (Telenti *et al.*, 1993; Ramaswamy and Musser, 1998).

Фторхинолоны – группа препаратов второго ряда, обладающих бактерицидной активностью. Основной мишенью действия препаратов данной группы является фермент ДНК-гираза (GyrA-GyrB). Попадая в клетку, препарат связывается с ферментом, тем самым нарушая негативную сверхспирализацию ДНК, что в конечном итоге приводит к нарушению репликации (Drlica and Malik, 2003). Формирование устойчивости к фторхинолонам связано в основном с модификациями мишени действия, обусловленными мутациями в небольшом локусе гена *gyrA* (кодоны Gly88, Asp89, Ala90, Ser91 и Asp94), так же называемым как «регион, определяющий устойчивость к хинолонам» (от англ. quinolone resistance-determining region, QRDR). Мутации в гене *gyrB* встречаются реже и, как правило, сопровождают мутации в *gyrA* (Guillemin *et al.*, 1998).

Ингибиторы синтеза белка

Стрептомицин – препарат первого ряда, оказывающий бактерицидное действие на активно делящиеся микобактерии, расположенные вне клеток организма-хозяина. Препарат нарушает синтез белков, связываясь с 16S субъединицей рибосомной РНК и белком S12 бактериальной рибосомы (Garvin *et al.*, 1974), что приводит к появлению ошибок в полипептидной цепи в процессе трансляции. Согласно литературным данным, до 75 % стрептомицин-устойчивых штаммов имеют мутации в генах *rpsL* (кодирует рибосомный белок S12), *rrs* (кодирует 16S рРНК) и *gidB* (Finken *et al.*, 1993; Cooksey *et al.*, 1996).

Канамицин/Амикацин/Капреомицин – препараты второго ряда. По механизму действия схожи со стрептомицином. Устойчивость связывают с точечными мутациями в гене 16S рРНК (*rrs*) в позициях 1400, 1401, 1402 и 1484, а также в гене *gidB*. Предполагается так же, что устойчивость к капреомицину может быть обусловлена точечными мутациями в гене рРНК-метилтрансферазы (*tlyA*). С мутациями в промоторной области гена *eis* ассоциируют развитие устойчивости к канамицину/амикацину (Georghiou *et al.*, 2012).

Нарушение мембранного транспорта

Пиразинамид – препарат первой линии, обладающий бактерицидной активностью на покоящиеся клетки. Под действием фермента пиразиназы, кодируемого геном *pncA*, пролекарство пиразинамид превращается в активную форму – пиразиновую кислоту. Протонированная форма кислоты вызывает закисление цитоплазмы, нарушает протонный потенциал и функции мембранного транспорта (Zhang *et al.*, 2003). Возникновение устойчивости к пиразинамиду связывают с дефектами фермента пиразиназы, обусловленными точечными мутациями, делециями или инсерциями (до 1355 п.о.) в гене *pncA*. Мутации не имеют определенной локализации и выявляются в различных кодонах *pncA* у 97 % пиразинамид-устойчивых штаммов (Scorpio and Zhang, 1996).

В заключении следует отметить, что 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. Так, например, в случае этамбутол-устойчивых штаммов мутации в *embCAB* опероне выявляются в среднем с частотой 70 % (Sreevatsan *et al.*, 1997). Для препаратов второго ряда такого рода ассоциации зачастую еще ниже. В связи с этим актуальны исследования, направленные на выявление и описание новых детерминант устойчивости. Для решения этих задач наиболее перспективными представляются технологии полногеномного секвенирования с последующим сравнительным анализом геномов (Warner and Mizrahi, 2013). В 2013 году были опубликованы результаты двух крупных исследований по поиску генетических детерминант устойчивости методами сравнительной геномики большого количества образцов. Так в исследовании Farhat с соавторами был проведен анализ 123 геномов МБТ, представляющих основные филогенетические линии (Farhat *et al.*, 2013). Применяя новый разработанный филогенетический тест, авторы обнаружили

как известные детерминанты устойчивости (11 мутаций), так и не описанные ранее мутации (39 мутаций), локализованные в 30 генах и межгенных участках. Большинство выявленных мутаций располагались в генах семейства PE/PPE, роль которых остается до конца не ясна. Пять других мутаций были найдены в генах, вовлеченных в биосинтез клеточной стенки. Что не маловажно, исследователи экспериментально показали влияние мутаций в одном из генов, *ponA1*, на устойчивость к ПТП. В другом исследовании Zhang с соавторами проанализировали выборку из Китая, включающую 161 штамм МБТ (Zhang *et al.*, 2013). Наибольшее количество образцов при этом относилось к филогенетической линии два. 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. Дополнительно авторами было показано, что ПТП оказывают слабый положительный отбор на геном МБТ.

2. МАТЕРИАЛЫ И МЕТОДЫ

2.1 Бактериальные штаммы

В исследование было включено 500 препаратов ДНК штаммов *M. tuberculosis*, отобранных из коллекций трех центров по борьбе с туберкулезом: ЦНИИ Туберкулеза РАМН – 349 образцов, Государственного казенного учреждения здравоохранения города Москвы «Московский городской научно-практический Центр борьбы с туберкулезом Департамента здравоохранения г. Москвы» (МНПЦБТ) – 54 образца и Санкт-Петербургского НИИ Фтизиопульмонологии (СПбНИИФ) – 97 образцов. Перечисленные образцы были переданы в ФГБУН НИИ физико-химической медицины ФМБА России в рамках проведения опытно-конструкторских работ по Госконтракту № 16.522.11.2003 от 19 марта 2012 года на тему: «Разработка и выпуск опытных образцов тест-системы для обнаружения лекарственно-устойчивого туберкулеза с учетом особенностей геномной организации эндемичных для России штаммов».

Сформированные коллекции включали в себя штаммы *M. tuberculosis*, полученные от пациентов, проходивших лечение с 2005 по 2012 год в поликлиниках научно-исследовательских учреждений. Идентификация и выделение чистой культуры патогена проводились в специализированных лабораториях соответствующих учреждений согласно разработанным протоколам (Министерство здравоохранения РФ. О совершенствовании противотуберкулезных мероприятий в Российской Федерации.// Приказ № 109 от 21 марта 2003 г.)

Для 150 штаммов коллекции на базах соответствующих учреждений методом абсолютных концентраций (Министерство здравоохранения РФ. О совершенствовании противотуберкулезных мероприятий в Российской Федерации.// Приказ № 109 от 21 марта 2003 г.) ранее был определен уровень лекарственной устойчивости к препаратам первого (рифампицин, изониазид,

стрептомицин и этамбутол) и второго ряда (офлоксацин, амикацин, капреомицин и канамицин). По данным тестирования из 150 образцов *M. tuberculosis* 19 (12.7 %) были идентифицированы как чувствительные, 2 (1.3 %) – монорезистентные, 13 (8.7 %) – полирезистентные, 83 (55.3 %) – МЛУ, 33 (22 %) – ШЛУ. Профили антибиотикорезистентности для каждого тестируемого образца представлены в приложении 1.

Выделение тотальной геномной ДНК образцов коллекции было проведено в соответствующих учреждениях по методу van Embden с соавторами (van Embden *et al.*, 1993) либо с использованием набора «ПРОБА-НК» производства ООО «ДНК-Технология».

Для 310 образцов коллекции в лабораториях ЦНИИ Туберкулеза РАМН и НИИ эпидемиологии и микробиологии им. Пастера были определены сполиготипы. Сполиготипирование было проведено по классической методике с использованием иммобилизованных на нейлоновой мембране уникальных зондов, комплементарных фрагментам DR-региона (Kamerbeek *et al.*, 1997). Для определения семейства и сполиготипа по международной классификации полученные профили сравнивали с таковыми, представленными в базе сполигопрофилей SpolDB4.0 (Brudey *et al.*, 2006) и в новой версии постоянно обновляемой компьютерной базы данных SITVIT (http://www.pasteur-guadeloupe.fr:8081/SITVIT_ONLINE/) (Demay *et al.*, 2012). Наибольшее количество образцов относилось к семействам Beijing (226/310; 73 %), LAM (31/310; 10 %), Ural (16/310; 5 %) и T1 (9/310; 3 %). На долю семейств H2, H3, MANU2, T1_RUS2, T2-T3, T4, T5_RUS1 и U пришлось менее 10 % образцов.

2.2 Геномные последовательности

В исследование была включена выборка геномных последовательностей 373 образцов *M. tuberculosis* и одного образца *M. canettii*, находящихся в открытом доступе в базе данных NCBI. Геномные

последовательности 8 образцов представлены в виде протяженных секвенсов – полные геномы, скаффолды, контиги (таблица 6)

Таблица 6.

Коллекция геномных секвенсов образцов

Образец	Генетическое семейство	Номер в GenBank
<i>M. tuberculosis</i>		
KZN-4207	LAM	CP001662
KZN-2475	LAM	ACVT00000000
94_M4241A	Beijing	ABL00000000
CCDC5079	Beijing	CP001641
CCDC5180	Beijing	CP001642
W-148	Beijing	GL877853.1
OSDD493	Ural	AVQJ00000000.1
<i>M. canettii</i>		
CIPT 140010059		NC_015848.1

Геномы 366 образцов представлены в виде чтений с приборов высокопроизводительного секвенирования (Illumina) и входят в состав четырех проектов: «*Mycobacterium tuberculosis* pilot» (92 образца, проект № PRJEB2221), «Discovery of sequence diversity in *Mycobacterium tuberculosis* (key strains)» (84 образца, проект № PRJEB2070) и «Discovery of sequence diversity in *Mycobacterium tuberculosis* (Russia collection)» (160 образцов, проект № PRJEB2138) и «*Mycobacterium tuberculosis* Genome sequencing» (30 образцов, проект № PRJNA218508).

Согласно исследованиям Casali с соавторами профиль лекарственной устойчивости был известен для 22 образцов МБТ из коллекции «Discovery of sequence diversity in *Mycobacterium tuberculosis* (Russia collection)» (Casali *et al.*, 2012). Данные о лекарственной устойчивости 24 образцов из коллекции «*Mycobacterium tuberculosis* Genome sequencing» были любезно предоставлены сотрудниками Центра геномной биоинформатики им. Ф. Г. Добржанского Санкт-Петербургского государственного университета. Из 46 образцов МБТ девять (19.6 %) были идентифицированы как чувствительные,

2 (4.3 %) – монорезистентные, 4 (8.7 %) – полирезистентные, 22 (47.8 %) – МЛУ и 9 (19.6 %) – ШЛУ.

2.3 Амплификация фрагментов генома *M. tuberculosis*

Амплификацию осуществляли в реакционной смеси: 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.

Подбор праймеров для амплификации проводили с помощью программы Oligo 6.71 (Molecular Biology Insights, Inc., Cascade, США). Праймеры для амплификации соответствующих локусов геномов STR1-2 и SP28 представлены в приложении 6. Для амплификации фрагментов генома, участвующих в рекомбинационных событиях, праймеры приведены в таблице 7. Праймеры для амплификации фрагментов генома, ассоциированных с лекарственной устойчивостью, приведены в таблице 8.

Таблица 7.

Праймеры для амплификации фрагментов генома, участвующих в рекомбинационных событиях

Название праймера	Последовательность (5' → 3')
P1	gtgtgtacattgggcatcg
P2	ggtgtacatatcgaagctcg
P3	gctcgacgaagtgagattgc
P4	ttggcgatccgatacagtgc
P5	ctgccaagcactggacagc
P6	caagtctccggtattcaagg
P7	agccttggctcgtcctacc
P8	cacggctctcccaacgtgg

Таблица 8.

Праймеры для амплификации фрагментов генома, ассоциированных с лекарственной устойчивостью

Генетический локус	Название праймера	Последовательность (5'→3')	Ожидаемый размер ПЦР-продукта (п.о.)	Источник
<i>rpoB</i>	MtrpoBf MtrpoBr	gagcgatcacaccgcagac ggtacggcggttcgatgaac	321	Ikryannikova <i>et al.</i> , 2007
<i>katG</i>	MtkatGf MtkatGr	acccgaggetgctccgctgg cagctcccactcgtagccgt	168	Ikryannikova <i>et al.</i> , 2007
<i>fabG-inhA</i> промоторная область	MtfabGf MtfabGr	gcctcgctgccagaaagg ctccggatccacgggggt	320	Ikryannikova <i>et al.</i> , 2007
<i>embB</i>	EMB306A EMB306B EMB 406F EMB 406R	ccgacgccgtggtgatattcgct	167	Ahmad <i>et al.</i> , 2004
		gtaataccagccgaaggatcctc ccatggtcttgctgacc cacaccagtgtgaatgc	170	
<i>gyrA-gyrB</i>	gyrA-F gyrA-R gyrB-F gyrB-R	ggaggtgcgcgacgggctcaagc	203	Zimenkov <i>et al.</i> , 2013
		ccagcgggtagcgcagcgacca caaggctgtgtcctcgccgcaagcc cggtgccagcgcctgatgat	311	
<i>rrs</i>	rrs-F rrs-R	tctcagttcggatcggggctgcaactcg ccagttggggcgtttctggtgctcc	280	Zimenkov <i>et al.</i> , 2013
<i>eis</i> промоторная область	eis-F eis-R	gcgaaattcgtcgtgattctgcagtgcc cccggccagtcgtctcggtcg	176	Zimenkov <i>et al.</i> , 2013

Амплификацию проводили на приборе DNA Engine Tetrad 2 (MJ Research, США). В зависимости от ожидаемой длины ПЦР-продукта были использованы различные профили амплификации исследуемых локусов генома. Общий вид профиля представлен ниже:

94°C – 5 мин.

30–45 циклов 94°C – 15 сек.

 Ta°C – 15 сек. где Ta – температура отжига праймера

 72°C – от 15 сек. до 3 мин.

72°C – 10 мин.

4°C – хранение.

Детекцию продуктов амплификации осуществляли методом электрофореза в 1 % агарозном геле. Окраску гелей проводили бромистым этидием.

Амплификацию тандемных повторов в 24 локусах генома для образцов *M. tuberculosis* осуществляли по методике, описанной ранее (Supply *et al.*, 2006).

Амплификацию DR-региона для сполиготипирования проводили согласно классическому протоколу с использованием неботионилированных праймеров (Kamerbeek *et al.*, 1997).

2.4 Подготовка продуктов амплификации для дальнейшего анализа

Перед реакцией удлинения зонда и секвенированием проводили дефосфорилирование концевых фосфатных групп дезоксирибонуклеозидтрифосфатов и элиминацию оставшихся праймеров в пост-амплификационной смеси. К продуктам реакции амплификации добавляли по 5 мкл смеси, содержащей 66 мМ Tris-HCl pH 9.0; 16.6 мМ $(\text{NH}_4)_2\text{SO}_4$; 2.5 мМ MgCl_2 ; 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°C. Инактивацию фермента осуществляли прогреванием в течение 10 минут при 85°C. В образцы, предназначенные для проведения реакции удлинения зонда, экзонуклеазу I не добавляли.

2.5 Реакция удлинения зонда и масс-спектрометрический анализ продуктов реакции

Проведение реакции удлинения зонда

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 (зонды). Зонды подбирали с использованием программы Oligo 6.71 (Molecular Biology Insights, Inc., Cascade, США) и

были сгруппированы в 8 систем. Олигонуклеотидные праймеры для реакции удлинения зонда приведены в таблице 9.

Таблица 9.

Олигонуклеотидные зонды, используемые для определения структуры DR-региона

Спейсер	Назв пр-ра	5'-3' последовательность	Масса продукта реакции при отсутств спейсера	Масса продукта реакции при присутств спейсера (+дд сист)	ддНТФ сист
система I					
1	SP_1	taggggagcgtgatccag	5605	5902	ддАТФ
2	SP_2	gtcgccaagcatgaggca	5823	6120	ддАТФ
3	SP_3	gctccagtgatcgcttct	6035	6332	ддАТФ
4	SP_4	gacctgatgattggtcgcgct	6494	6791	ддАТФ
5	SP_5	taatcccgcacaaagtggcag	6416	6713	ддАТФ
6	SP_6	tgcgtcgtcattccggcttca	6669	6966	ддАТФ
система II					
7	SP_7	gagtactcggggctgcc	5459	5772	ддГТФ
8	SP_8	cgcgtgaaaccgccccca	5375	5688	ддГТФ
9	SP_9	tgacgacatgcgaatccgct	6102	6415	ддГТФ
10	SP_10	actagagcgggtgtcgaac	5878	6191	ддГТФ
11	SP_11	gacggaaaccaggtgagcaac	6640	6953	ддГТФ
система III					
12	SP_12	tgacgacctcggacagc	5181	5478	ддАТФ
13	SP_13	gacacgtggggagagagg	5366	5663	ддАТФ
14	SP_14	cgccaacccttccggtgtg	5756	6053	ддАТФ
15	SP_15	cttgaataacgcgcagtga	5837	6134	ддАТФ
16	SP_16	cgatttacgacgctgacggga	6472	6769	ддАТФ
система IV					
17	SP_17	accggcaacaatcgcg	5179	5452	ддЦТФ
18	SP_18	acctcccggaccatctg	5092	5365	ддЦТФ
19	SP_19	ctggattgcgctaactgg	5531	5804	ддЦТФ
20	SP_20	ccacatcgatttcctgac	5699	5972	ддЦТФ
21	SP_21	tgacgactcgcgattggga	6182	6455	ддЦТФ
система V					
22	SP_22	gacggcacgattgagacaa	5871	6184	ддГТФ
23	SP_23	cgctgatcggtccagctc	5781	6094	ддГТФ
24	SP_24	gcacgtctcaccagca	5101	5414	ддГТФ
25	SP_25	actcggggagccgatca	5221	5534	ддГТФ
26	SP_26	cggatgatggtggtgctgaag	6583	6896	ддГТФ
27	SP_27	tcacgaatccggcgtcgaagg	6457	6770	ддГТФ
система VI					
28	SP_28	ttagcgatcacacaacca	6054	6327	ддЦТФ
29	SP_29	ccagcgaaatacaggctc	5478	5751	ддЦТФ
30	SP_30	acgacggctggaaaagg	5623	5896	ддЦТФ
31	SP_31	tgacgaccggacttgatcga	6142	6415	ддЦТФ

32	SP_32	tcactcccgaccaaataggtat	6664	6937	ддЦТФ
система VII					
33	SP_33	acggcgggtgccgcactt	5188	5501	ддГТФ
34	SP_34	gcttcgacggtgtgggc	5259	5572	ддГТФ
35	SP_35	cgctcgaaacgctccaac	5718	6031	ддГТФ
36	SP_36	gtcgaaatccagcaccatcc	6634	6947	ддГТФ
37	SP_37	tctgacgacctgaaagggg	5878	6191	ддГТФ
38	SP_38	ccacaacgcatccgccatccac	6555	6868	ддГТФ
система VIII					
39	SP_39	ggcgtgggtaaccgatcc	5541	5854	ддГТФ
40	SP_40	gacctgtatgaggcacagat	6166	6479	ддГТФ
41	SP_41	gtctgacgactacctgatagaa	6744	7057	ддГТФ
42	SP_42	tcgtccatgtcccaccatga	6013	6326	ддГТФ
43	SP_43	ccgggtgcgggaggtgca	5622	5935	ддГТФ

Для каждой системы реакцию удлинения зонда осуществляли в 10 мкл реакционной смеси, содержащей 66 мМ Tris-HCl (pH 9.0); 16.6 мМ $(\text{NH}_4)_2\text{SO}_4$; 2.5 мМ MgCl_2 ; 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. Реакцию проводили в термоциклере DNA Engine Tetrad 2 (MJ Research, США). Для всех систем праймеров была подобрана единая оптимальная температура отжига. Нарботку продуктов реакции осуществляли по универсальному профилю:

95 °С – 5 мин

70 циклов: 94 °С – 20 сек

60 °С – 20 сек

72 °С – 15 сек

Очистка продуктов реакции удлинения зонда

Очистку продуктов реакции удлинения зонда осуществляли на катионно-обменной смоле (SPS-SAC-50, Синтез полимерных сорбентов, Россия). Полученную смесь выдерживали при комнатной температуре в течение часа. В дальнейшем сорбент осаждали центрифугированием в

течение 5 мин при 16100 g. Супернатант использовали для проведения масс-спектрометрического анализа.

MALDI-ToF масс-спектрометрический анализ продуктов реакций удлинения зонда

Масс-спектрометрический анализ продуктов реакции проводили на приборе Autoflex III smartbeam vertical MALDI-ToF (Bruker Daltonics, Германия) согласно инструкциям производителя. Для этого аликвоту образца (0.3 мкл) наносили на матрицу, предварительно высушенную на планшете AnchorChip (600 мкм, Bruker Daltonics, Германия), и давали ей высохнуть. Матрицу готовили из насыщенного раствора 3-гидроксипиколиновой кислоты (Fluka, Германия) в 50 % ацетонитриле (Merck, Германия) с добавлением 10 г/л цитрата аммония двухосновного (Fluka, Германия). Полученные планшеты загружали в прибор.

Анализ масс-спектров и определение сполиготипа

Обработку масс-спектров осуществляли с использованием программного обеспечения фирмы Bruker Daltonics (Германия): flexControl 2.4 (Build 38) и flexAnalysis 2.4 (Build 11).

При получении масс спектров фиксировали пики, соответствующие определенной ожидаемой молекулярной массе. Масс-спектры каждого образца сравнивали с контрольным, и по наличию изменения молекулярной массы, соответствующей массе используемого ддНТФ, определяли наличие того или иного спейсера. Сполиготип для каждого образца записывали в двоичном формате, где за 1 принималось наличие спейсера, а за 0 – его отсутствие.

Для определения семейства и сполиготипа по международной классификации полученные профили сравнивали с таковыми, представленными в базе сполигопрофилей SpolDB4.0 (Brudey *et al.*, 2006) и в новой версии постоянно обновляемой компьютерной базы данных SITVIT (http://www.pasteur-guadeloupe.fr:8081/SITVIT_ONLINE/) (Demay *et al.*, 2012).

2.6 Секвенирование фрагментов генома модифицированным методом Сенгера

Нуклеотидные последовательности фрагментов генома определяли по методу Сенгера с модификациями (Sanger *et al.*, 1992) 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 (таблицы 7 и 8, приложение 6).

Восстановление полноразмерных последовательностей, выравнивание и сравнительный анализ последовательностей осуществляли с использованием ContigExpress и AlignX программного пакета «Vector NTI 11.0» (Informax Inc, США).

Продукты реакции амплификации анализировали методом электрофореза в 1 % агарозном геле в 1×TBE-буфере. Окраску проводили бромистым этидием. Длину ПЦР-продукта определяли путем сравнения размера полученного фрагмента с маркером GeneRuler™ 100 bp DNA Ladders (Fermentas, США) в программном пакете Quantity One v.4.4.0 (BioRad, США).

Число tandemных повторов в соответствующих 24 локусах вычисляли путем сравнения табличных данных с размерами ПЦР-продукта.

2.7 VNTR-типирование

Продукты реакции амплификации анализировали методом электрофореза в 1 % агарозном геле в 1×TBE-буфере. Окраску проводили бромистым этидием. Длину ПЦР-продукта определяли путем сравнения размера полученного фрагмента с маркером GeneRuler™ 100 bp DNA Ladders (Fermentas, США) в программном пакете Quantity One v.4.4.0 (BioRad, США).

Число tandemных повторов в соответствующих 24 локусах вычисляли путем сравнения табличных данных с размерами ПЦР-продукта.

2.8 Полногеномное секвенирование

Определение полногеномных последовательностей 54 штаммов было проведено методом секвенирования с использованием высокопроизводительного секвенатора GS FLX + (Roche, США). Фрагментные библиотеки были получены в соответствии с протоколом производителя (Rapid Library Preparation Method Manual GS FLX+ Series – XL+, Roche, США) с использованием коммерческого набора GS FLX Titanium Rapid Library Preparation Kit (Roche, США). Средний размер библиотек соответствовал 1400–1800 п.о. Средняя длина чтений с прибора (ридов) для секвенируемых образцов составила 780 нуклеотидов.

Для образца SP21 было дополнительно проведено секвенирование с использованием прибора Ion Torrent PGM (Life Technologies, США). Для этого были получены две библиотеки парных фрагментов (с размером библиотеки 2000–3000 п.о. и 5000–6000 п.о.) с помощью коммерческого набора 5500 SOLiD Mate-Paired Library Construction Kit (Life Technologies, США) в соответствии с протоколом Ion Mate-Paired Library Preparation (Life Technologies Demonstrated Protocol). Средняя длина чтений с прибора составила 100 нуклеотидов.

Для первичной оценки характеристик полногеномного секвенирования использовали сборку (ассемблинг) *de novo* с помощью программы GS *de novo* assembler версии 2.5 (Roche, США). Для оценки перекрытия референтного генома *M. tuberculosis* штамма H37Rv и получения файлов выравнивания использовали программу 454 GS Reference Mapper (Roche, США). Результаты выравнивания визуализировали с использованием программы Mauve версии 2.3.1. (<http://gel.ahabs.wisc.edu/mauve/>) (Darling *et al.*, 2004).

Для определения последовательностей между контигами и, тем самым, установления полной геномной последовательности дополнительно использовали секвенирование по Сенгеру как описано ранее (пункт 2.6).

2.9 Анализ данных полногеномного секвенирования

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

Картирование чтений с прибора на референтный геном *M. tuberculosis* штамма H37Rv производили с помощью инструмента bowtie2 (Langmead and Salzberg, 2012) (<http://bowtie-bio.sourceforge.net/bowtie2/index.shtml>) с использованием настроек по умолчанию. Анализ данных выравнивания проводили в приложении samtools 0.1.19 (Li *et al.*, 2009) (<http://samtools.sourceforge.net/>). Поиск однонуклеотидных полиморфизмов осуществлялся с помощью программы VarScan v2.3.1 (Koboldt *et al.*, 2012) (<http://varscan.sourceforge.net>). Пример запрос на определение однонуклеотидных полиморфизмов представлен ниже:

```
# align reads with bowtie2
bowtie2 -t -p 1 -x $BWTREF -1 $WF\_1.fastq -2 $WF\_2.fastq -S $WF.sam

# converting and sorting bowtie2 results
samtools view -bS $WF.sam > $WF.bam
samtools sort $WF.bam $WF.sort

# get snps
samtools mpileup -f $FREF $WF.sort.bam | java -jar $VARSCAN mpileup2cns --
variants 1 --min-coverage 1 --min-reads2 1 --p-value 0.05 > $WF.snps
```

Выравнивание и поиск замен для полногеномных последовательностей производилось 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 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/>). Если отношение < 1 , то считается, что последовательности находятся под стабилизирующим отбором, если значение числа примерно равно 1, то последовательность эволюционирует нейтрально, если же значение > 1 , то считается, что последовательность находится под положительным (позитивным) отбором.

Для оценки разнообразия *M. tuberculosis* была вычислена матрица расстояний Хэмминга на основании однонуклеотидных замен. Построение филогенетических деревьев осуществлялось исходя из полученной матрицы с применением алгоритма Neighbor-joining. В качестве внешней группы был добавлен образец *M. canettii* (NC_015848.1). Для уменьшения размерности и двухмерного представления разнообразия был применен принцип главных компонент.

2.10 Статистическая обработка данных

Дискриминирующую способность методов типирования определяли с помощью индекса Хантера-Гастона (*HGDI*):

$$HGDI = 1 - \frac{1}{N(N-1)} \sum_{j=1}^S n_j (n_j - 1),$$

где S – число групп, на которое метод разделяет всю выборку штаммов, n_j – число штаммов в j -й группе и N – общее число штаммов в исследованной выборке.

Согласно общепринятой практике, методы типирования (или их комбинации), имеющие значение индекса Хантера-Гастона 0,9 и выше, считаются приемлемыми.

Для определения значимости отличий dN/dS между группами использовался апоста тест при значении p -value ≤ 0.05 . Определение достоверности отличий количества замен в COG категориях и оценка значимости различий в частотах встречаемости замен проводились с

помощью биномиального теста ($p\text{-value} \leq 0.05$). Для сравнения матриц расстояний, полученных на основании нуклеотидных замен и данных VNTR-анализа, использовали корреляции Спирмэна и Пирсона.

Для определения достоверности ассоциации нуклеотидной замены с резистентностью использовался точный тест Фишера ($p\text{-value} \leq 0.05$, в качестве нулевой гипотезы было принято отсутствие ассоциации, в качестве альтернативной гипотезы — ее наличие). Для исключения ложноположительных ассоциаций использовалась поправка Бенджамини-Гохберга. Для корректировки полученных результатов была использована логистическая регрессия, основанная на ступенчатом механизме развития лекарственной устойчивости. Последовательность возникновения устойчивости выглядела следующим образом: $\{X_n\}$ $\{INH, RMP, STR, EMB, OFX, ETH, KAN, CPM\}$. Для каждой из ассоциированных с резистентностью к антибиотику $\{X_n\}$ мутации, было проведено сравнение двух логистических моделей вида $X_n \sim X_1 + X_2 + \dots + X_{n-1}$ и $X_n \sim X_1 + X_2 + \dots + X_{n-1} + SNP$ с помощью хи-квадрат теста ($p\text{-value} \leq 0.05$). В качестве нулевой гипотезы было выбрана равнозначность моделей.

3. РЕЗУЛЬТАТЫ

3.1 Формирование экспериментальных групп образцов ДНК штаммов

M. tuberculosis

Сформирована лабораторная коллекция препаратов ДНК, экстрагированных из 500 штаммов *M. tuberculosis*, выделенных от пациентов трех центров по борьбе с туберкулезом: ЦНИИ Туберкулеза РАМН, МНПЦБТ и СПбНИИФ.

Для 150 образцов был установлен профиль лекарственной устойчивости к препаратам первого (изониазид, рифампицин, этамбутол, стрептомицин) и второго (офлоксацин, канамицин, амикацин, капреомицин) ряда терапии туберкулеза (приложение 1). Определение чувствительности проводили методом абсолютных концентраций в бактериологических лабораториях соответствующих учреждений.

Группа образцов ($n = 310$) с определенным по классической схеме (Kamerbeek *et al.*, 1997) сполиготипом была использована для разработки собственного лабораторного метода сполиготипирования с применением MALDI-ToF масс-спектрометрии (Matrix-Assisted Laser Desorption/Ionization Time Of Flight Mass-Spectrometry, времяпролетная масс-спектрометрия с матричной лазерной десорбционной ионизацией). Разработанный метод позволил установить сполигопаттерны оставшихся 190 образцов ДНК.

Из группы штаммов, охарактеризованных по степени чувствительности к противотуберкулезным препаратам первого и второго ряда, были отобраны 54 образца для осуществления полногеномного секвенирования.

3.2 Формирование коллекции геномных последовательностей

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* и 1 *M. canettii*), представленные в виде протяженных секвенсов, использованы для построения филогенетического дерева (таблица 6).

Данные о 366 геномах МБТ, представленные в виде чтений с приборов высокопроизводительного секвенирования (Illumina), использованы для валидации специфических для генотипов полиморфизмов. В настоящее исследование вошли данные четырех проектов: «*Mycobacterium tuberculosis* pilot» (92 образца), «Discovery of sequence diversity in *Mycobacterium tuberculosis* (key strains)» (84 образца), «Discovery of sequence diversity in *Mycobacterium tuberculosis* (Russia collection)» (160 образцов) и «*Mycobacterium tuberculosis* Genome sequencing» (30 образцов).

Для построения гипотез формирования резистентности дополнительно были использованы геномные секвенсы 46 штаммов *M. tuberculosis* с известным профилем устойчивости к противотуберкулезным препаратам. Данные геномы также представлены в виде чтений с приборов высокопроизводительного секвенирования и входят в состав двух проектов: «Discovery of sequence diversity in *Mycobacterium tuberculosis* (Russia collection)» (22 образца) и «*Mycobacterium tuberculosis* Genome sequencing» (24 образца).

3.3 Сполиготипирование штаммов *M. tuberculosis*

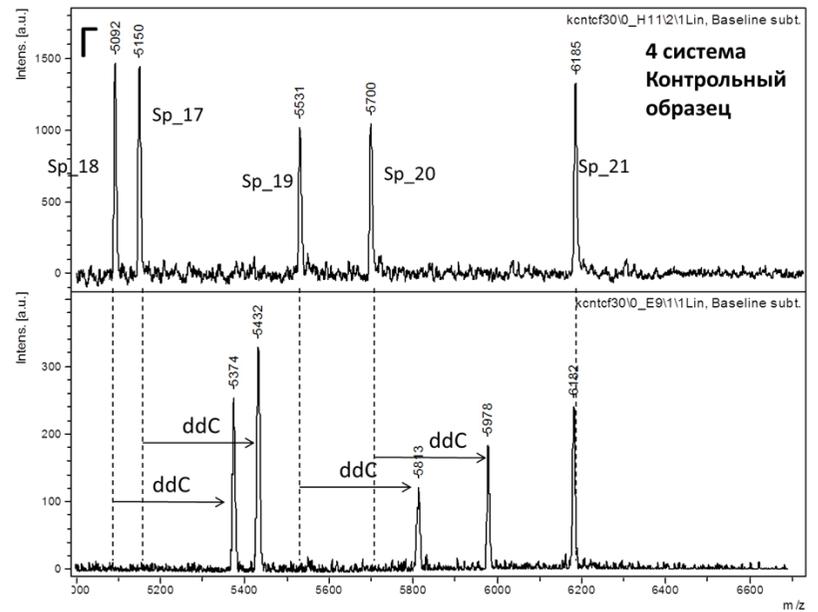
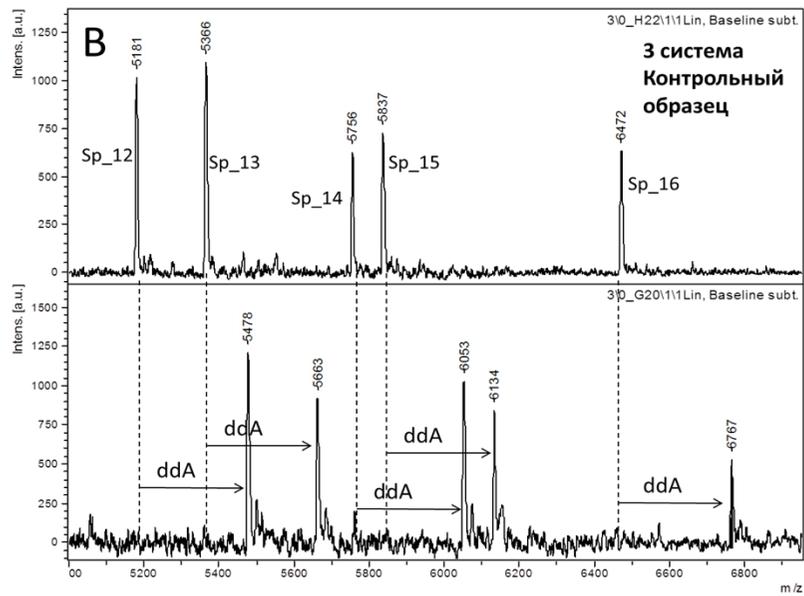
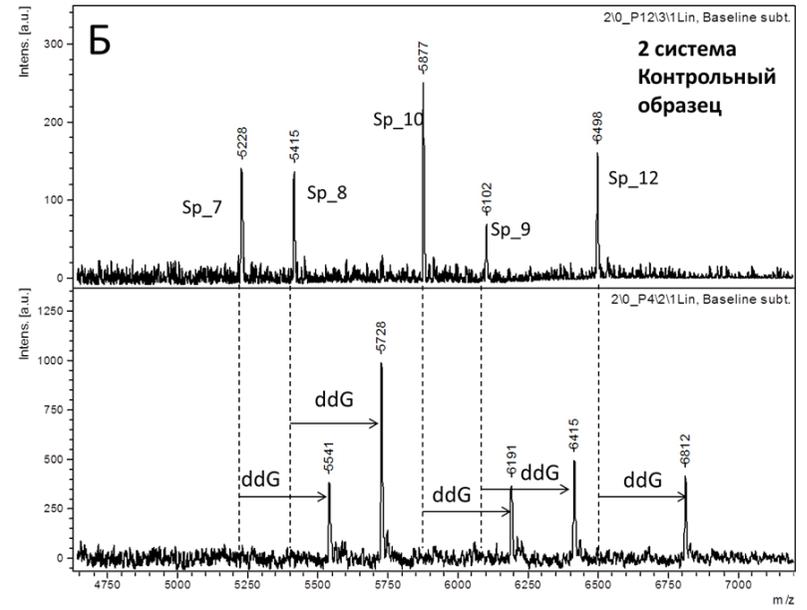
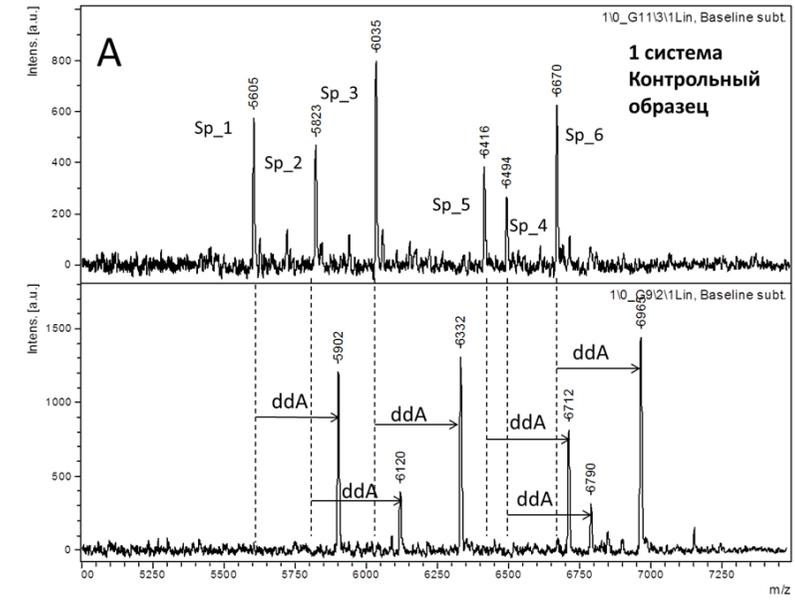
Для группы из 310 образцов был известен гибридизационный профиль, полученный в ходе проведения сполиготипирования по классической методике с использованием иммобилизованных на нейлоновой мембране уникальных зондов, комплементарных фрагментам DR-региона.

Для определения структуры DR-региона оставшихся 190 образцов был разработан лабораторный метод сполиготипирования, основанный на применении реакции удлинения зонда с последующей MALDI-ToF масс-спектрометрической детекцией.

Метод основан на реакции ферментативного достраивания олигонуклеотидного зонда с участием дидезоксирибонуклеотидтрифосфатов, так называемых терминаторов, которые, в свою очередь, ответственны за специфическую остановку ферментативной реакции. Таким образом, в ходе реакции образуются нуклеотидные цепочки, отличающиеся по длине на 1 нуклеотид. Матрицей при этом служит предварительно очищенный ПЦР-фрагмент DR-региона. На завершающем этапе проводится MALDI-ToF масс-спектрометрическая детекция олигонуклеотидов, отличающихся на массу одного основания от исходного зонда.

Для каждого из 43 спейсерных участков региона были разработаны высокоспецифичные праймеры (зонды) для реакции удлинения зонда (таблица 9). Зонды были сгруппированы в 8 систем для повышения общей производительности анализа. Пример результатов типирования образца семейства LAM9 (SIT42) представлен на рисунке 6.

Разработанный метод был апробирован на 310 образцах с известными паттернами сполиготипирования. Метод показал полную сходимость результатов с данными классического сполиготипирования.



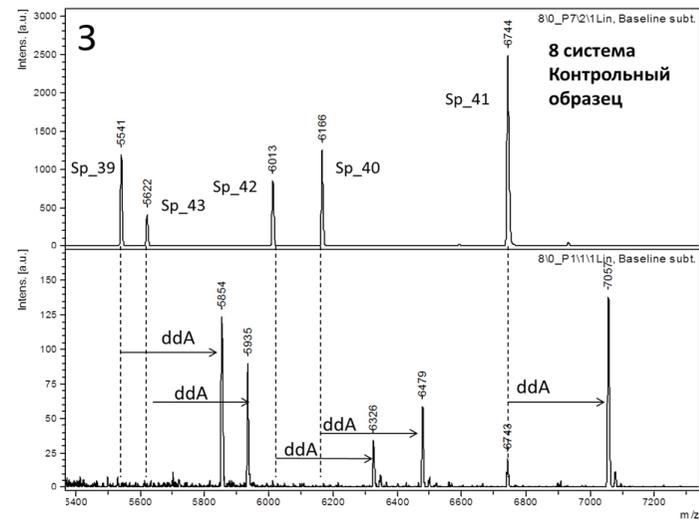
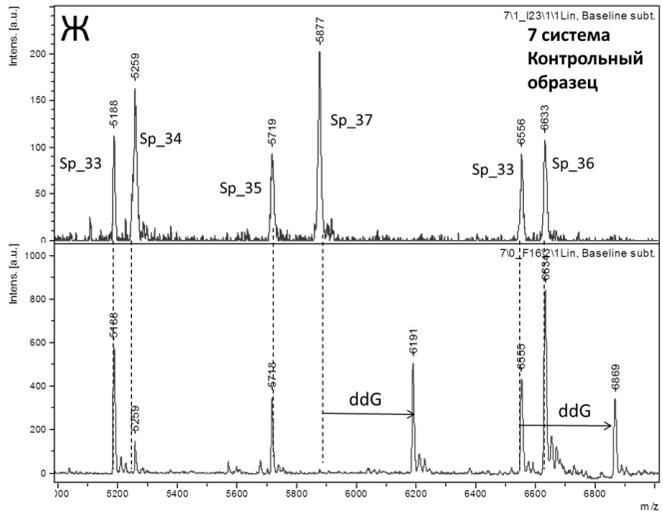
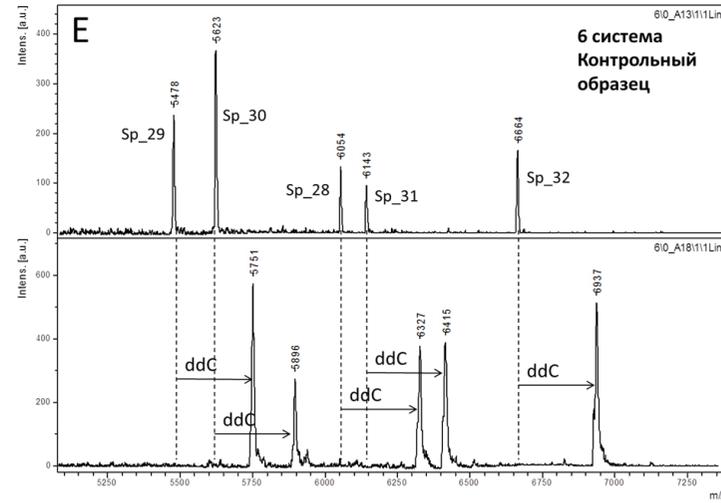
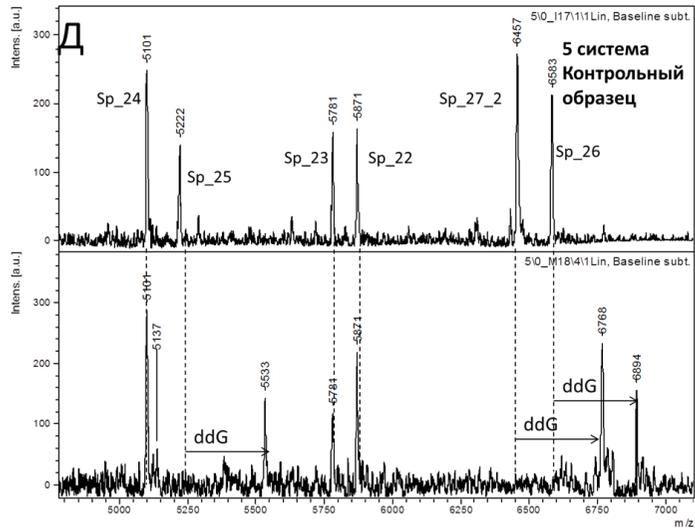


Рисунок 6. Масс-спектры образца, принадлежащего к LAM9 семейству. А – первая система, Б – 2 система, В – 3 система, Г – 4 система, Д – 5 система, Е – 6 система, Ж – 7 система, З – 8 система. На каждом из рисунков верхний спектр соответствует контрольному образцу, нижний – исследуемому.

Полногеномное секвенирование проведено с помощью высокопроизводительного геномного секвенатора нового поколения GS FLX+ Series — XL+ (Roche, США) с первичным приготовлением фрагментных баркодированных библиотек (Секвенирование проведено в лаборатории постгеномных методов исследования ФГБУН НИИ ФХМ ФМБА России Карповой Ириной Юрьевной и Беспярых Юлией Андреевной). Полученные чтения для каждого образца депонированы в Sequence read archive (SRA) (проект № PRJNA181180).

По результатам сборки *de novo* для каждого образца был выведен отчет с указанием количества чтений и оснований, вошедших в сборку; количества контигов, т.е. последовательностей, собранных программой из ридов с применением специального алгоритма обработки; N50, показателя качества сборки; покрытие представленных последовательностей. Суммарная отчетная информация по секвенированным геномам представлена в таблице 11.

Таблица 11.

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

Образец	Количество ридов	Количество оснований	Количество больших контигов при сборке <i>de novo</i> (>500 п.о.)	N50*	Покрывание
SP1	81909	30438396	816	8127	6.0
SP2	134659	50189911	211	44811	10.0
SP3	104076	39000963	370	22197	8.0
SP4	175471	72894347	133	93844	16.0
SP5	116786	39778839	398	21472	8.0
SP6	129883	48136569	231	41663	10.0
SP7	175956	77764794	136	74958	16.0
SP8	111927	43808801	273	32356	9.0
SP9	80637	32005337	619	11169	7.0
SP10	116700	46517045	224	37833	10.0
SP11	148437	58889195	162	61099	12.0
SP12	89186	34876051	453	16540	7.0
SP13	62826	24729962	1232	4703	5.0
SP15	16393	6570335	1468	966	2.9
SP16	56107	29868228	709	9004	6.0
SP17	28001	15390577	1474	2212	3.1

SP18	42120	22734779	1068	5072	4.9
SP21	67767	36221974	385	17366	7.0
SP22	38731	20765922	1254	3909	4.0
SP23	157250	65500638	150	64630	14.0
SP24	137671	78772357	121	82557	16.0
SP25	176266	110952555	109	114358	25.0
SP26	227636	92737403	131	73111	20.0
SP27	89027	35873657	480	15797	8.0
SP28	196465	123227872	108	103541	27.0
SP29	63509	40332940	225	39252	9.0
SP30	162613	65519164	141	70707	14.0
SP31	27545	15066753	1479	2204	3.1
SP32	31019	17101387	1461	2655	3.9
SP33	33255	18008741	1527	2617	3.9
SP34	90843	49551661	195	46205	10.0
SP35	41582	22694086	1167	4472	4.1
SP36	81444	44363395	212	39823	9.0
SP37	62303	34894013	402	18137	7.0
SP38	78925	44565858	212	37554	10.0
SP39	52401	27928246	902	6522	5
SP41	129553	84824150	132	77185	18.0
SP42	135916	90642224	134	77520	19.0
SP45	133845	89650837	126	71074	19.0
SP46	146010	95197972	111	77233	21.0
MOS2	236155	121692160	95	128825	26.0
MOS7	128313	55325788	153	64051	12.0
MOS9	343923	147036759	112	98092	31.0
MOS10	107965	34328273	824	8692	7.0
MOS11	111548	35370500	570	11956	7.0
MOS12	91856	29354590	964	6794	6.0
MOS14	108024	35894213	653	10946	7.0
MOS15	183859	109501832	134	73836	24.0
MOS16	76989	46115919	306	26698	9.0
MOS17	118374	66553110	199	45728	13.0
MOS18	53863	29622932	938	6443	5.0
CTRI-2	223321	49931803	501	15974	11.0
CTRI-3	216795	47309214	612	11445	10.0
CTRI-4	217033	45937341	617	11042	10.0

N50 – показатель качества сборки. Рассчитывается следующим образом: все контиги сортируются по длине, определяется суммарная длина контигов с учетом покрытия и ее половинное значение, далее проводится сложение длин контигов до того момента, когда длина станет соответствовать показателю половинного значения, длина соответствующего контига при суммировании и будет называться N50

В ходе анализа секвенированных геномов коллекции для каждого образца было определено местонахождение однонуклеотидных полиморфизмов в кодирующей области или в межгенном пространстве относительно референтного генома *M. tuberculosis* штамма H37Rv. В случае выявления SNP в кодирующей области дополнительно была определена

синонимичность произошедшей замены. Полиморфизм обозначался как синонимичный (sSNP), если нуклеотидная замена не приводила к изменению аминокислоты в белке, или несинонимичный (nsSNP), если замена приводила к изменению аминокислоты. Дополнительно для каждого генома было определено соотношение транзиций (замена пурин - пурин, пиримидин - пиримидин) и трансверсий (замена пурин ↔ пиримидин), а также отношение dN/dS (таблице 12).

Таблица 12.

Сравнительный анализ секвенированных геномов

Образец	sSNP	nsSNP	Межгенный участок	% транзиций	% трансверсий	dN/dS
SP1	527	820	175	60.7	39.3	0.568
SP2	542	886	184	61.0	39.0	0.597
SP3	552	872	216	61.0	39.0	0.577
SP4	591	933	211	60.6	39.4	0.576
SP5	523	875	199	61.5	38.5	0.611
SP6	548	886	212	61.2	38.8	0.59
SP7	610	920	220	61.0	39.0	0.551
SP8	552	988	227	58.1	41.9	0.653
SP9	505	840	191	61.4	38.6	0.607
SP10	574	893	206	60.0	40.0	0.568
SP11	598	940	213	60.6	39.4	0.574
SP12	516	840	190	61.2	38.8	0.594
SP13	582	902	244	60.1	39.9	0.566
SP15	325	502	133	62.1	37.9	0.564
SP16	533	828	196	61.4	38.6	0.567
SP17	340	540	115	61.2	38.8	0.58
SP18	463	739	195	61.8	38.2	0.583
SP21	545	871	220	60.1	39.9	0.583
SP22	450	738	160	60.0	40.0	0.599
SP23	622	931	216	60.6	39.4	0.546
SP24	454	645	127	59.2	40.8	0.519
SP25	500	715	173	58.8	41.2	0.522
SP26	612	957	219	60.2	39.8	0.571
SP27	565	874	219	60.5	39.5	0.565
SP28	424	616	124	59.5	40.5	0.53
SP29	566	912	210	61.1	38.9	0.588
SP30	591	927	212	60.6	39.4	0.573
SP31	255	408	97	62.0	38.0	0.584
SP32	259	397	85	58.8	41.2	0.56
SP33	302	416	92	61.1	38.9	0.503
SP34	570	893	215	60.9	39.1	0.572
SP35	297	489	90	59.2	40.8	0.601

SP36	381	593	103	58.5	41.5	0.568
SP37	366	578	109	58.8	41.2	0.576
SP38	390	588	103	58.9	41.1	0.55
SP39	324	517	94	60.1	39.9	0.582
SP41	452	634	122	59.9	40.1	0.512
SP42	421	655	143	59.4	40.6	0.568
SP45	633	951	218	59.7	40.3	0.548
SP46	379	585	119	59.2	40.8	0.563
MOS2	439	649	120	58.3	41.7	0.54
MOS7	376	597	114	58.0	42.0	0.58
MOS9	501	746	183	58.3	41.7	0.544
MOS10	489	842	172	61.9	38.1	0.629
MOS11	543	873	208	61.0	39.0	0.587
MOS12	515	847	218	61.0	39.0	0.6
MOS14	518	874	194	61.6	38.4	0.616
MOS15	665	990	216	60.2	39.8	0.543
MOS16	580	903	221	60.6	39.4	0.568
MOS17	575	908	205	60.8	39.2	0.576
MOS18	536	846	187	61.1	38.9	0.576
CTRI-2	286	479	101	61.9	38.1	0.611
CTRI-3	344	506	114	60.5	39.5	0.537
CTRI-4	431	754	164	60.9	39.1	0.639

3.5 Сборка полных геномных последовательностей

Для образцов CTRI-2 (LAM), SP28 (Ural) и SP21 (Beijing) было проведено дополнительное секвенирование, что позволило представить их геномные последовательности меньшим числом контигов, чем было получено при первичном анализе. Геном SP28 представлен в виде 5 контигов, геном SP21 в виде 12 скаффолдов, депонированных в базу данных NCBI (AOUF00000000.1). Для образца CTRI-2 была определена полная кольцевая последовательность геномной ДНК (GenBank accession number CP002992). Последовательность генома составила 4398525 пар оснований с GC-составом порядка 65.6 %. Был определен один набор рРНК, 45 тРНК и 3946 белок-кодирующих генов, составивших 90.3 % генома. Более 97 % генов соответствовали генам H37Rv (Pina *et al.*, 2013).

В геноме было выявлено 100 вставок и 104 делеции относительно генома H37Rv. Наиболее крупные вставки и делеции представлены в таблице 13.

Таблица 13.

Крупные вставки и делеции в геноме CTRI-2 (более 1000 п.о.), не связанные с повторяющимся элементом IS6110. Вставки/делеции, идентичные геномам CDC1551 и F11, выделены серым цветом. Вставки/делеции, идентичные геному F11, выделены черным. Вставки/делеции, идентичные геному CDC1551, выделены жирным.

Делеции (n = 6)			Вставки (n = 4)			
Позиция в H37Rv	ORF	Продукт гена	Позиция в CTRI-2 ^{SENS}	Позиция в F11	ORF	Продукт гена (согласно аннотации генома F11)
453367 – 455972	Rv0376c	hypothetical protein	334409 – 337250	334971 – 337812	TBFG_10283	PE-PGRS family protein
	Rv0377	LysR family transcriptional regulator			TBFG_10284	PE-PGRS family protein
	Rv0378	glycine rich protein			TBFG_10285	PE-PGRS family protein
1779243 - 1788513	Rv1573-Rv1585	phiRV1 phage proteins	1987509 - 1988461	2002082 - 2003034	TBFG_11789	hypothetical protein
	Rv1586c	phiRv1 integrase			TBFG_11790	hypothetical protein
1986626 - 1987702	Rv1754c	hypothetical protein	2262290 - 2268647	2278338 - 2283337	TBFG_12059	hypothetical protein
	Rv1755c	phospholipase C 4 PLCD			TBFG_12060	hypothetical protein
2635577 - 2638997	Rv2353c	PPE family protein	3704771 - 3710210	3705270 - 3709351		Product (according to CDC1551 annotation)
	Rv2354	transposase			MT3426	pterin-4-alpha-carbinolamine dehydratase
	Rv2355	transposase			MT3427	molybdopterin cofactor biosynthesis protein A
	Rv2356c	PPE family protein			MT3427.1	hypothetical protein
3842767 - 3847692	Rv3426	PPE family protein			MT3428	AfsR/DnrI/RedD family transcriptional regulator
	Rv3427c	transposase			MT3429	hypothetical protein
	Rv3428c	transposase			MT3427.1	hypothetical protein
3945959 - 3950163	Rv3514	PE-PGRS family protein				

3.5 Сравнительный филогенетический анализ секвенированных штаммов

На основании полученных данных полногеномного секвенирования построено филогенетическое дерево с учетом всех однонуклеотидных замен геномов (исследование проведено совместно с Ищенко Д.С). Дополнительно в анализ было включено 7 геномов МБТ из международной базы данных NCBI. В качестве внешней группы был выбран геном *M. canettii* (таблица 6). Для построения дерева использовался алгоритм Neighbor-Joining. Полученное филогенетическое дерево представлено на рисунке 7.

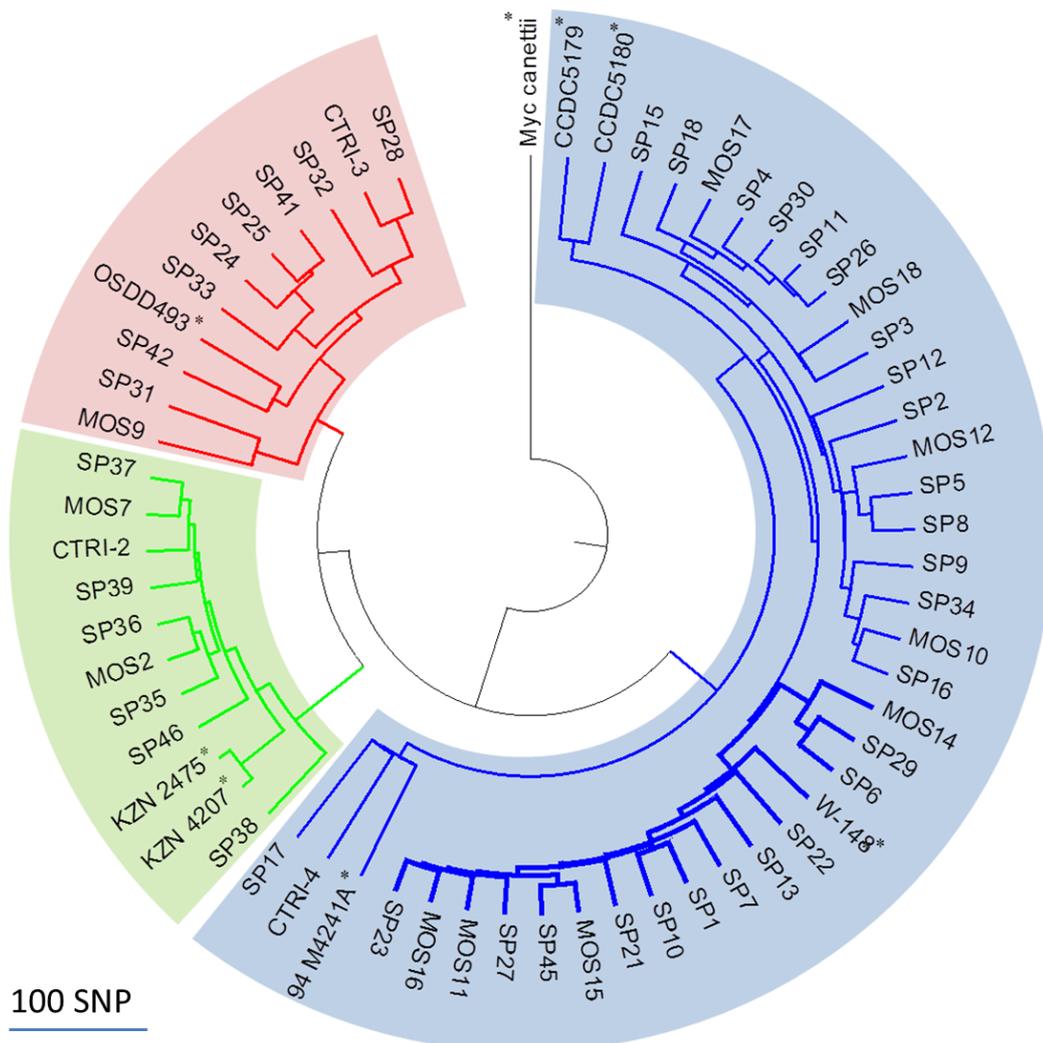


Рисунок 7. Сравнительный филогенетический анализ секвенированных штаммов. Алгоритм построения Neighbor-Joining. Синим цветом выделены образцы семейства Beijing, зеленым – LAM, красным – Ural. Звездочкой отмечены геномы из GenBank.

В ходе анализа филогенетического дерева было установлено, что штаммы различных филогенетических линий сгруппированы в отдельные кластеры. В кластер группирующихся изолятов, соответствующих генотипу Beijing, 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. Таким образом, результаты филогенетического анализа секвенированных штаммов показали сходимость с результатами первичного генотипирования собранных образцов.

Для оценки сходимости результатов секвенирования с методом VNTR-типирования был проведен анализ числа tandemных повторов по 24 локусам для всех секвенированных штаммов (приложение 2) (проведено совместно с Беспятыми Ю.А.). По результатам исследования было обнаружено 29 различных VNTR-профилей. Двадцать один из них были уникальными, а 33 оставшихся образца образовывали 8 кластеров, включающих в себя от 2 до 9 штаммов. Значение показателя индекса Хантера-Гастона составило 0.944. Корреляции Спирмана и Пирсона между филогенетическими расстояниями, полученными на основании VNTR-анализа и анализа единичных нуклеотидных полиморфизмов, составили 0.82 и 0.96, соответственно.

Дополнительно результаты VNTR-анализа были сопоставлены с базой данных MIRU-VNTR-plus (<http://www.miru-vntrplus.org/>). Построенные филогенетические деревья показали полную сходимость на уровне генетических семейств.

На основании литературных данных был проведен поиск молекулярных маркеров для определения основных линий (семейств) внутри *M. tuberculosis*. Результаты анализа филогенетических маркеров представлены в таблице 14.

Маркеры филогенетических линий *M. tuberculosis*

Образец	Spoligotype (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)
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	Beijing	Главная генетическая группа 1	Кластер 2	Линия 1	Восточно-Азиатская линия
MOS2, MOS7, SP35, SP36, SP37, SP38, SP39, SP46, CTRI-2	LAM	Главная генетическая группа 2	Кластер 5	Линия 2	Евро-Американская линия
MOS9, SP24, SP25, SP28, SP31, SP32, SP33, SP41, SP42, CTRI-3	Ural	Главная генетическая группа 2	Кластер 3b	Линия 2	Евро-Американская линия

3.6 Определение генотип-специфических мутаций

В ходе сравнительного анализа данных полногеномного секвенирования для филогенетических семейств Beijing, Ural и LAM были отобраны генотип-специфические мутации (рисунок 8).

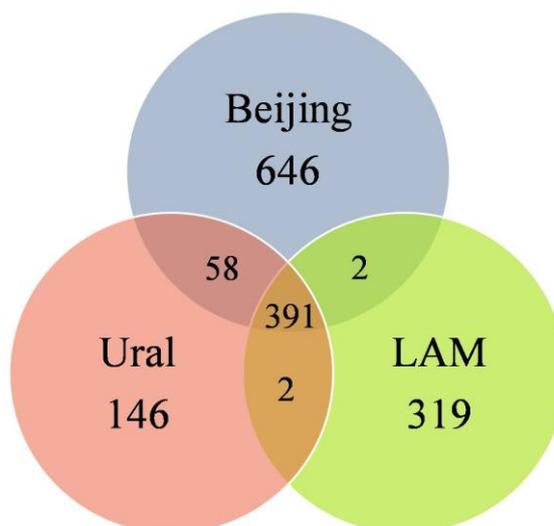


Рисунок 8. Диаграмма Венна, иллюстрирующая пересечение однонуклеотидных полиморфизмов для трех изучаемых семейств: Beijing, Ural и LAM.

По результатам анализа для семейства Beijing было идентифицировано 646 специфических однонуклеотидных полиморфизмов в генах и межгенных участках, не встретившихся среди представителей семейств LAM и Ural (приложение 3). Для семейства LAM таких полиморфизмов обнаружено 319, для Ural – 146 (приложения 4 и 5, соответственно). В ходе анализа установлено, что семейства Beijing и Ural имеют 58 общих полиморфизмов, тогда как попарное сравнение семейств Beijing - LAM и Ural - LAM выявило по 2 общих полиморфизма. Общими среди представителей трех описываемых семейств были определены 391 SNPs.

3.7 Валидация генотип-специфических SNPs

Валидация генотип-специфических полиморфизмов была проведена с использованием 366 геномов *M. tuberculosis*, представленных в базе данных NCBI. В ходе анализа геномов для каждого образца было определено местонахождение однонуклеотидных полиморфизмов в кодирующей области или в межгенном пространстве относительно референтного генома *M. tuberculosis* штамма H37Rv.

На основании исследований Homolka с соавторами (Homolka *et al.*, 2012) 366 представленных геномов были отнесены к шести основным филогенетическим линиям (совместно с Ищенко Д. С.). Десять образцов не были дифференцированы. Дополнительно было построено филогенетическое дерево с учетом всех однонуклеотидных замен геномов (данные не приведены). Результаты исследований показали полную сходимость в кластеризации образцов. Не дифференцированные образцы по первому методу согласно филогенетическому анализу были отнесены к семейству Beijing (таблица 15).

Таблица 15.

Распределение 366 образцов по филогенетическим линиям

Филогенетическая линия	Семейство	количество образцов
1	EAI	30
2	Beijing	134+10*

3	CAS	38
4	Haarlem	22
	LAM	38
	Ural	28
	Ghana	6
	Cameroon	4
	X-type	38
5	AFRI2	8
6	AFRI1	10

* 10 образцов добавлено в ходе филогенетического анализа

Полученные однонуклеотидные полиморфизмы были использованы для корректировки генотип-специфических паттернов для описанных ранее семейств Beijing, Ural и LAM. По результатам исследования для представителей семейства Beijing было идентифицировано 422 генотип-специфических полиморфизмов; для Ural – 143 SNPs; для LAM – 309 SNPs (приложения 3, 4 и 5). 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) [10] представлено на рисунке 9 (совместно с Ищенко Д.С.). Нормализация количества nsSNPs была проведена с учетом длин генов в каждом представленном COG и с учетом количества несинонимичных полиморфизмов, найденных в конкретном семействе.

Таким образом, отобрано необходимое и достаточное количество полиморфизмов, обеспечивающих надежную дифференциацию штаммов *M. tuberculosis* на уровне семейств.

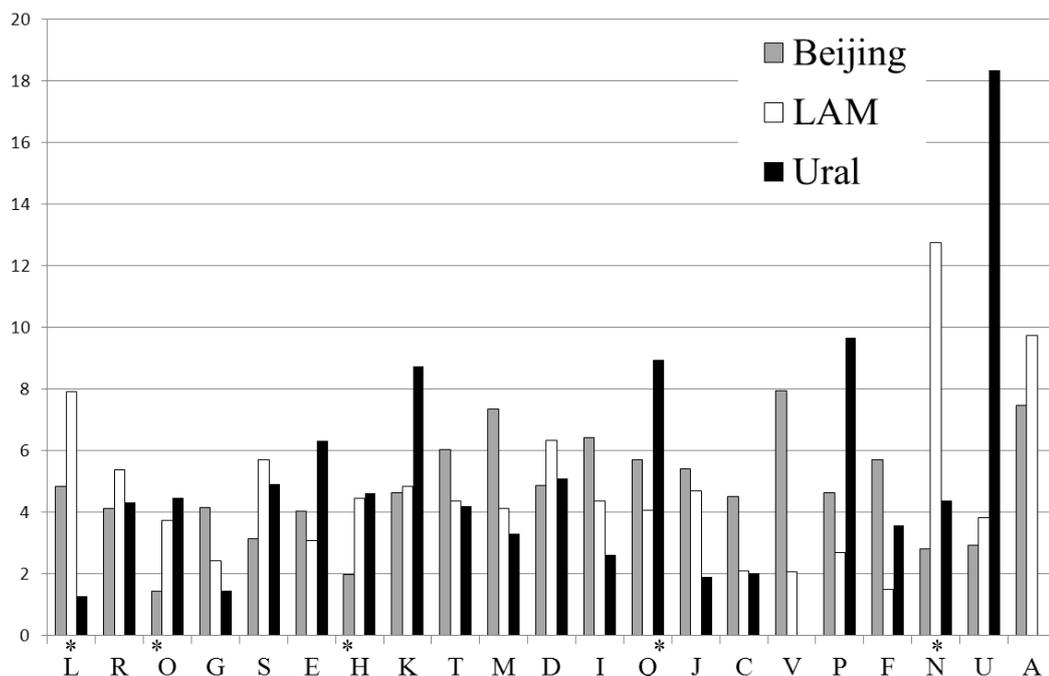


Рисунок 9. Гистограмма распределения генотип-специфических nsSNPs по COG категориям. Звездочкой отмечены статистически значимые различия в количестве nsSNPs в представленном COG определенно семейства ($p < 0.05$). А – процессинг и модификация РНК; С – получение и преобразование энергии; D – контроль клеточного цикла, клеточное деление, деление хромосом; Е – транспорт и метаболизм аминокислот; F – транспорт и метаболизм нуклеиновых кислот; G – транспорт и метаболизм углеводов; Н – 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.

3.8 Характеристика кластера Beijing B0/W148

Сравнительный анализ VNTR-паттернов представителей семейства Beijing разделил исследуемую выборку на 14 VNTR-профилей. Обширная группа из 15 образцов относилась к типу M11 согласно Мокроусову с соавторами [3], причем 12 из них имели профили, ассоциированные с кластером Beijing B0/W148 (223325173533424672454433 (9 образцов), 223325173533424572454433 (2 образца), 223325173533424672444433 (1

образец)) (Mokrousov, 2013). Наличие кластера B0/W148 было так же подтверждено результатами построения филогенетического дерева (рисунок 7, образцы с W-148 по SP23). Дополнительно принадлежность образцов к кластеру была проверена методом ПЦР согласно публикации Мокроусова с соавторами (Mokrousov *et al.*, 2012). 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.

3.8.1 Beijing B0/W148 кластер-специфические SNPs

По результатам сравнительного анализа и дополнительной валидации для образцов кластера Beijing B0/W148 был определен характерный мутационный профиль. В итоге было идентифицировано 63 полиморфизма. Среди них 7 полиморфизмов находилось в межгенной области. Из мутаций, найденных в кодирующей области генома, 29 были синонимичными, а 27 приводили к несинонимичным заменам. Полиморфизмы в межгенных областях и nsSNPs представлены в таблице 16. Три полиморфизма в кодирующие области относились к нонсенс-мутациям и приводили к образованию стоп-кодона. Это мутации в гене Rv2719с, участвующем в регуляции клеточного деления *M. tuberculosis*, в гене Rv2800, участвующем в клеточном метаболизме, и в гене Rv3728, являющемся элементом эффлюксной системы патогена.

Таким образом, на основании сравнительной геномики секвенированных образцов удалось выделить специфические полиморфизмы, характерные для представителей кластера Beijing B0/W148.

Таблица 16.

Beijing B0/W148 кластер-специфические однонуклеотидные полиморфизмы

Позиция	Ген	№ локуса	Ам-та реф	Ам-та мут	Вес аминокислотной замены по BLOSUM62	Синонимичность	Нукл. реф	Нукл. мут.	Продукт гена	Функция белка
143120	охсА	Rv0118c	Ala	Ser	1	nsSNP	С	А	Probable oxalyl-CoA decarboxylase ОхсА	участвует в метаболизме щавелевой кислоты, в дыхании
404130	-	Rv0338c	Glu	Gly	-2	nsSNP	Т	С	Probable iron-sulfur-binding reductase	участвует в метаболизме и дыхании
886661	межген.						С	Т		
1023883	межген.						Г	А		
1035426	pstS3	Rv0928	Phe	Cys	-2	nsSNP	Т	Г	Periplasmic phosphate-binding lipoprotein PstS3 (PBP-3) (PstS3)	Участвует в транспорте через мембрану соединений фосфора
1155884	kdpC	Rv1031	Met	Thr	-1	nsSNP	Т	С	Probable potassium-transporting ATPase C chain KdpC	Участвует в транспорте калия
1482185	-	Rv1319c	Gly	Asp	-1	nsSNP	С	Т	Possible adenylate cyclase (ATP pyrophosphate-lyase) (adenylate cyclase)	Аденилат циклаза. Участвует в регуляции клеточного метаболизма.
1619513	PE_PGRS26	Rv1441c	Tyr	Asp	-3	nsSNP	А	С	PE-PGRS family protein PE_PGRS26	функция не известна
1710601	-	Rv1518	Ter	Lys		nsSNP	Т	А	hypothetical protein	Возможно гликозил трансфераза. Участвует в синтезе экзополисахарида
1947325	vapC12	Rv1720c	His	Arg	0	nsSNP	Т	С	Possible toxin VapC12	Токсин VAPC12, участник системы токсин-антитоксин. Участвует в росте бактерии
2078967	-	Rv1833c	Asp	Tyr	-3	nsSNP	С	А	Possible haloalkane dehalogenase	Может действовать на широкий спектр 1-галогеналканы, галоген спирты, галогеналканы и некоторые галоидзамещенным ароматические соединения
2099129	ureC	Rv1850	Ala	Val	0	nsSNP	С	Т	Urease alpha subunit UreC (urea amidohydrolase)	Фермент уреазы. Участвует в клеточном ответе на нехватку азота
2127011	-	Rv1877	Trp	Gly	-2	nsSNP	Т	Г	Probable conserved integral membrane protein	Возможно связан с транспортом лекарств через мембрану
2205550	межген.						А	Г		

2210745	mce3B	Rv1967	Ser	Ala	1	nsSNP	T	G	Mce-family protein Mce3B	Белок эффлюксной системы MCE. Вовлечен в инвазию клеток макроорганизма.
2335638	-	Rv2079	Leu	Ser	-2	nsSNP	T	C	hypothetical protein	функция не известна
2468180	межген.						G	A		
2517129	fabD	Rv2243	Thr	Ala	0	nsSNP	A	G	Malonyl CoA-acyl carrier protein transacylase FabD (malonyl CoA:ACPM acyltransferase) (MCT)	Катализирует малонил-КоА-АСП трансацилазы (MCAT) активно использует голо-АСПМ в качестве субстрата для трансацилирования
2544135	cyp128	Rv2268c	Ser	Ala	1	nsSNP	A	C	Probable cytochrome P450 128 Cyp128	Цитохром P450 128 CYP128
2604740	-	Rv2331A	Met	Val	1	nsSNP	A	G	Hypothetical protein	Белок участвует в транспорте ионов
2964454	-	Rv2638	Pro	Leu	-3	nsSNP	C	T	hypothetical protein	функция не известна
3031090	-	Rv2719c	Gln	Ter		nsSNP	G	A	Possible conserved membrane protein	Белок-регулятор клеточного деления <i>M. tuberculosis</i>
3109512	-	Rv2800	Trp	Ter		nsSNP	G	A	Possible hydrolase	Фермент гидролаза, участвует в клеточном метаболизме
3557244	межген.						C	G		
3566107	-	Rv3196	Ala	Asp	-2	nsSNP	C	A	hypothetical protein	функция не известна
3610335	-	Rv3233c	Arg	Ser	-1	nsSNP	G	T	Possible triacylglycerol synthase (diacylglycerol acyltransferase)	Триаилглицерол синтаза
4175847	-	Rv3728	Trp	Ter		nsSNP	G	A	Probable conserved two-domain membrane protein	Белок относится к представителям эффлюксной системы MFS. Возможно участвует в транспорте углеводов и антибиотиков
4216783	межген.						A	C		
4221591	-	Rv3776	Thr	Ser	1	nsSNP	C	G	hypothetical protein	функция не известна
4221609	-	Rv3776	Ser	Thr	1	nsSNP	G	C	hypothetical protein	функция не известна
4321479	межген.						G	A		
4338371	whiB6	Rv3862c	Thr	Pro	-1	nsSNP	T	G	Possible transcriptional regulatory protein WhiB-like WhiB6	Белок WhiB6. Участвует в транскрипции
4349982	eccCb1	Rv3871	Ala	Ser	1	nsSNP	G	T	ESX conserved component EccCb1 ESX-1 type VII secretion system protein	функция не известна
4367633	eccE2	Rv3885c	Thr	Ala	0	nsSNP	T	C	ESX conserved component EccE2 ESX-2 type VII secretion system protein Possible membrane protein	функция не известна

3.8.2 Структурная организация генома Beijing B0/W148

В ходе филогенетического анализа было установлено сходство секвенированных образцов кластера Beijing B0/W148 со штаммом W-148, представленным в базе NCBI в виде одного скаффолда. Выравнивание геномов W-148 и H37Rv в программе Mauve 2.3.1 показало наличие двух больших инверсий в геноме W-148 (Shitikov *et al.*, 2014) (проведено совместно с Беспятых Ю.А.). При этом геном W-148 был разделен на пять коллинеарных блоков (таблица 17). Первый, третий и пятый блоки сохранили свою ориентацию, в то время как второй и четвертый были инвертированы и переставлены в W-148 по отношению к геному H37Rv (рисунок 10).

Таблица 17.

Распределение коллинеарных блоков между геномами H37Rv и W-148

Коллинеарный блок	Тип	Позиция в H37Rv	Позиция в W-148
I	Консервативен	1–704243	1–769418
II	Перестроен	704246–1262963	3247025–3814187
III	Консервативен	1262961–3379027	1108934–3245669
IV	Перестроен	3379025–3710381	773801–1107578
V	Консервативен	3711737–4411532	3814961–4539434

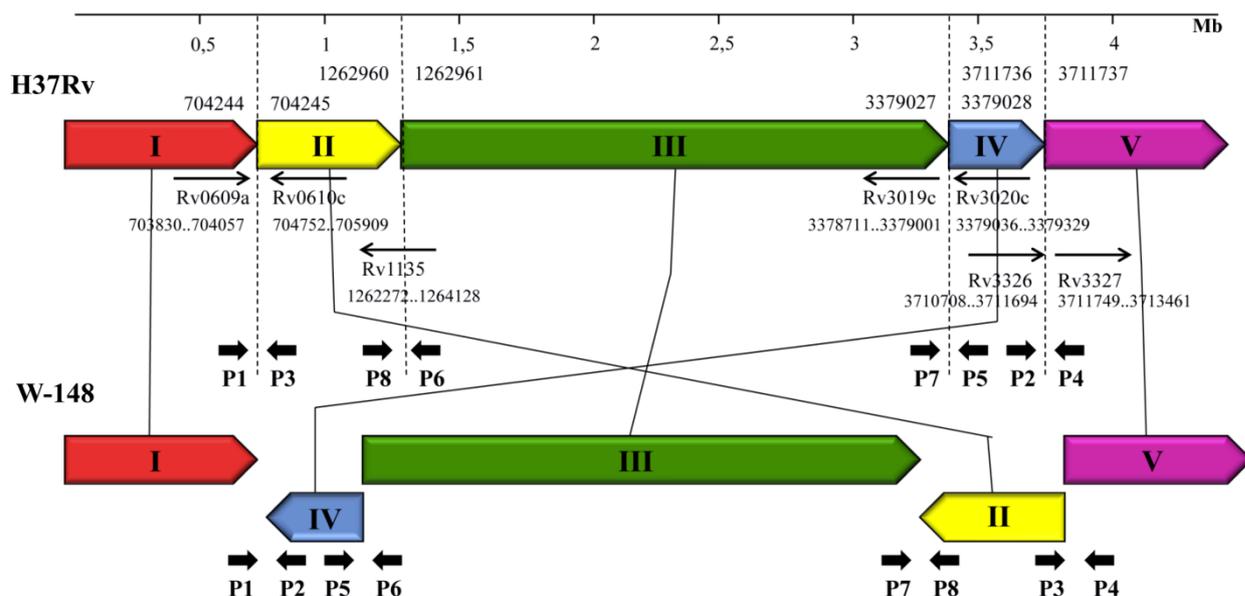


Рисунок 10. Схематическое изображение перестроек в геноме W-148 относительно H37Rv. Коллинеарные блоки соответствуют таковым в таблице

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).

Подтверждение хромосомных перестроек в исследуемых образцах кластера Beijing B0/W148

Для подтверждения наличия инверсий в исследуемых штаммах кластера Beijing B0/W148 было проведено полногеномное секвенирование штамма SP21 с использованием библиотек парных фрагментов. Дополнительно была разработана ПЦР-система для детекции инвертированных участков.

Полногеномное секвенирование штамма SP21

Считывание последовательности генома штамма SP21 было проведено на секвенаторе нового поколения IonTorrent с использованием библиотек парных фрагментов. Сборка последовательности генома SP21 была выполнена путем объединения чтений с прибора GS FLX+ Series — XL+ (Roche) (70 тыс. секвенсов, 540 п.о.) и IonTorrent (650 К ридов, средняя длина прочтения 180 п.о.). Начальная сборка была выполнена с помощью GS De Novo Assembler с получением 391 контига с длинами от 500 до 69788 пар оснований. По результатам анализа парных чтений из 391 контига было получено 12 скаффолдов общей длиной 4.45 м.п.о. Выравнивание геномов H37Rv, W-148 и SP21 показало наличие одинаковых крупномасштабных инверсии в SP21 и в W-148 (рисунок 11). Скаффолд 5 (392333 п.о.) включал полную последовательность блока IV и часть блоков I и III (29 т.п.о. и 16 т.п.о., соответственно). Скаффолд 3 (738393 п.о.) содержал большие части блоков III и II (16 т.п.о. и 132 т.п.о., соответственно). Скаффолд 9 (162927 п.о.) включал части блоков II и IV (132 т.п.о. и 31 т.п.о., соответственно).

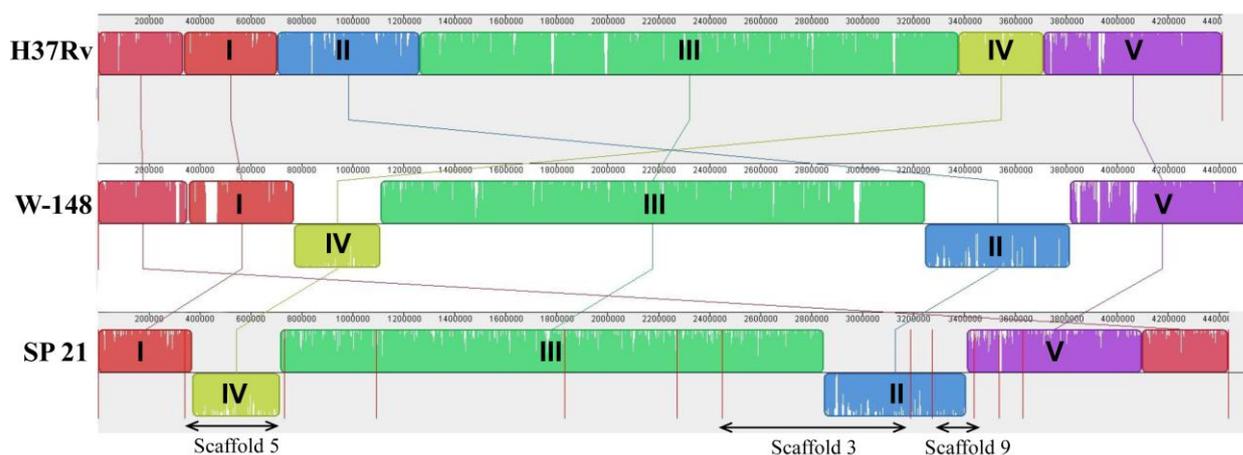


Рисунок 11. Выравнивание геномов H37Rv, W-148 и SP21 с использованием программы Mauve 2.3.1. Коллинеарные блоки соответствуют таблице 17 и рисунку 10. Вертикальные красные линии на геноме SP21 соответствуют границам скаффолдов. Скаффолды 5, 3 и 9 отмечены двусторонними стрелками и содержат последовательности инвертированных регионов.

Подтверждение инверсий с использованием ПЦР

Для определения наличия или отсутствия инверсий в секвенированных штаммах *M. tuberculosis* были подобраны 8 праймеров для амплификации границ инвертированных участков (таблица 7 и рисунок 10).

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 10). Для анализа геномной организации штаммов, не относящихся к кластеру Beijing B0/W148, использованы пары праймеров: P1-P3, P2-P4, P6-P8 и P5-P7. Полученные ПЦР-продукты представлены в таблице 18.

Таблица 18.

Пары праймеров, используемые для детекции рекомбинационных событий

Пары праймеров	Праймеры	Размер ПЦР-продукта		
		B0/W148 Beijing	Другие Beijing	He 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 была выполнена для образцов семейств Ural (16 образцов), LAM (16) и Beijing (32), среди которых 12 образцов относились к кластеру Beijing B0/W148.

В ходе проведенного анализа было выявлено, что ПЦР-продукты для образцов кластера Beijing B0/W148 образуются только при использовании пар праймеров 1, 2, 5 и 6 (таблица 18 и рисунок 12А, дорожки 1, 2, 5, 6), тогда как для образцов других семейств ПЦР-продукты получаются только при использовании пар 3, 4, 7 и 8 (таблица 18 и рисунок 12В, дорожки 3, 4, 7 и 8). Различия в длинах ампликонов, получаемых при использовании пар праймеров 4, 7 и 8 для штаммов различных семейств связаны с наличием или отсутствием копии повторяющегося элемента *IS6110* в анализируемом регионе. Все полученные специфичные ПЦР-продукты были подтверждены секвенированием по Сенгеру.

Таким образом, показано наличие инверсий в изученных штаммах Beijing B0/W148 ($n = 12$), которые, возможно, являются уникальными и специфичными для данного кластера.

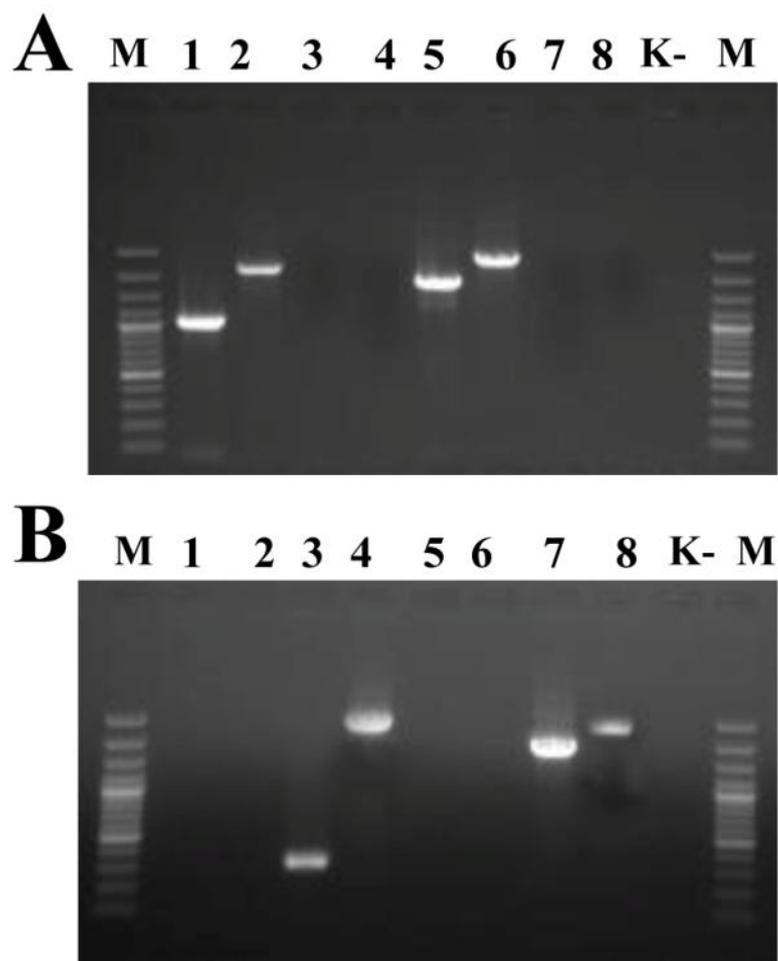


Рисунок 12. Электрофореграммы ПЦР-продуктов, полученных в ходе амплификации с праймерами из таблицы 7. А – амплификация ДНК образца SP21 кластера B0/W148 Beijing. В – амплификация ДНК образца SP5, не относящегося к кластеру B0/W148 Beijing. Бэнды 1 – 8 соответствуют парам праймеров из таблицы 18. М – маркер молекулярного веса GeneRuler 100 bp Plus DNA Ladder (SM0324, Fermentas, США). К – отрицательный контроль.

Реконструкция вероятной последовательности рекомбинационных событий в штаммах Beijing B0/W148

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 (рисунок 13). 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 это имеет место в P1 и H37Rv, что привело к образованию W-148.

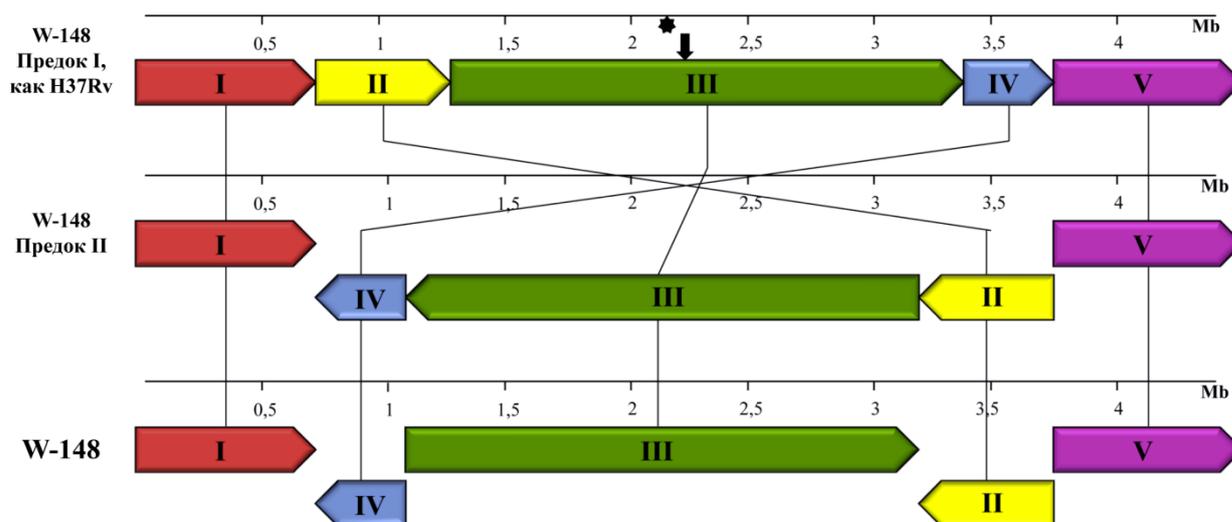


Рисунок 13. Реконструкция рекомбинационных событий в геноме W-148. Геномные перестройки показаны для геномов W-148 Предок I (схож с H37Rv), W-148 Предок II и W-148. Звездочкой обозначен сайт терминации репликации, рассчитанный с использованием программы GraphDNA [19]

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

Для 54 секвенированных образцов коллекции был произведен поиск полиморфизмов, ассоциированных с лекарственной резистентностью. Дополнительно для всех образцов было проведено секвенирование по Сенгеру генов и межгенных участков, ассоциированных с лекарственной устойчивостью. Результаты исследования представлены в таблице 19.

Анализ мутаций ассоциированных с устойчивостью к изониазиду

Среди 48 изониазид-устойчивых штаммов мутации в кодоне Ser315 гена *katG* и/или позиции -15 промоторной области *mabA(fabG)* – *inhA* оперона были обнаружены у всех образцов. У большей части образцов

(47/48) была выявлена нуклеотидная замена AGC→ACC (Ser→Thr) в 315 кодоне гена *katG*.

Тринадцать образцов имели замену в -15 позиции промотора *mabA(fabG) – inhA*. У 12 образцов были выявлены мутации одновременно в 315 кодоне гена *katG* и позиции -15 промоторной области оперона.

Анализ мутаций ассоциированных с устойчивостью к рифампицину

Мутации в RRDR-регионе гена *rpoB* были обнаружены среди всех 48 рифампицин-устойчивых штаммов. Замены Ser450Leu и Asp435Val встретились у 38 и 7 образцов, соответственно. Замены Ser431Arg, Asp435Tyr, His445Tyr и His445Asp встретились по одному разу. Один образец имел аминокислотные замены одновременно в двух разных позициях белка: Asp435Val/Ser431Arg.

Анализ мутаций ассоциированных с устойчивостью к этамбутолу

Среди 35 этамбутол-устойчивых штаммов мутации в гене *embB* были обнаружены у 34 образцов. Было выявлено девять вариантов замен: Met306Val (14 образцов), Met306Ile (9), Asp354Ala (4), Asn296His (2), Gln497Arg (2), Ser297Ala (1), Tyr319Ser (1), Asn399Thr (1) и Gly406Asp (1). Один образец содержал замены Met306Ile и Gly406Asp одновременно.

У 19 этамбутол-чувствительных штаммов обнаружены следующие аминокислотные замены: Met306Val (2), Met306Ile (2), Tyr319Ser (2), Ser297Ala (1), Asp354Ala (1), Gly406Ala (1) и Gln497Arg (1).

Анализ мутаций ассоциированных с устойчивостью к стрептомицину

Среди 47 стрептомицин-устойчивых штаммов мутации в генах *rpsL* и *rrs* были обнаружены у 44 образцов. В гене *rpsL* было выявлено два типа замен: Lys43Arg (22 образца) и Lys88Arg (6). В гене *rrs* были обнаружены мутации A514C (8) и C517T (8). Мутация C517T была обнаружена у одного стрептомицин-чувствительного образца.

Анализ мутаций ассоциированных с устойчивостью к канамицину/амикацину и капреомицину

Среди 33 штаммов, устойчивых хотя бы к одному из инъекционных антибиотиков, мутации в гене *rrs* и/или промоторной области гена *eis* были выявлены у 32 образцов. Наиболее часто встречаемой мутацией в гене *rrs* была замена A1401G (17 образцов). Также для одного образца была выявлена мутация G1484T. В промоторной области гена *eis* были выявлены мутации G(-10)A (5 образцов), G(-37)T (5 образцов), C(-14)T (4 образца) и C(-12)T (3 образца). У трех образцов были выявлены мутации в гене *rrs* и промоторной области гена *eis* одновременно: A1401G / C(-12)T (2 образца) и G1484T / C(-12)T (1 образец).

Среди 21 штамма, чувствительного ко всем представленным инъекционным препаратам, пять образцов содержали замены в промоторной области гена *eis*: G(-10)A (3 образца), C(-12)T (1 образец) и C(-14)T (1 образец).

Анализ мутаций ассоциированных с устойчивостью к фторхинолонам

Мутации в QRDR-регионе генов *gyrA* и *gyrB* были обнаружены у 25 из 27 штаммов, устойчивых к фторхинолонам. Наиболее часто были обнаружены замены в гене *gyrA*: Asp94Gly (11 образцов), Ala90Val (4), Ser91Pro (3), Asp94Ala (2), Asp94Tyr (2), Ala74Ser (1), Ala90Gly (1) и Asp94Asn (1). В гене *gyrB* были выявлены следующие замены: Ala504Val (2), Thr500Ala (1) и Thr500Ile (1). Четыре образца содержали одновременно несколько аминокислотных замен: Ala90Val/Asp94Gly, Thr500Ala/Ala90Val, Thr500Ile/Asp94Asn, Ala504Val/Ser91Pro.

Один фторхинолон-чувствительный образец содержал замену Ala504Val в гене *gyrB*.

Таблица 19.

Варианты нуклеотидных и аминокислотных замен, ассоциированных с лекарственной устойчивостью, среди изучаемых штаммов *M. tuberculosis*

Образец	Генотип	INH	KatG	Пром_inh	RIF	RpoB	EMB	EmbB	STR	RpsL	rrs	CAP	AM	KAN	rrs	Пром_eis	OFX	GyrB	GyrA
SP1	Beijing	R	S315T		R	S450L	R	M306I	R	K43R		S	S	R		G(-10)A	R		A90V
SP2	Beijing	R	S315T		R	S450L	S	M306V	R	K43R		S	S	R		G(-10)A	S	A504V	
SP3	Beijing	R	S315T		R	S450L	R	M306V	R		C517T	S	S	S		C(-14)T	R	A504V	
SP4	Beijing	R	S315T		R	S450L	R	D354A	R		C517T	S	S	R		G(-37)T	R		D94G
SP5	Beijing	S			S		S		S			S	S	S			S		
SP6	Beijing	R	S315T		R	S450L	R	M306V	R	K43R		S	S	R		G(-10)A	S		
SP7	Beijing	R	S315T		R	S450L	R	Q497R	R	K43R		R	R	R	A1401G		R		D94G
SP8	Beijing	R	S315T		R	S450L	S		R	K43R		S	S	S			S		
SP9	Beijing	S			S		S		S			S	S	S			S		
SP10	Beijing	R	S315T		R	S450L	R	M306I	R	K43R		S	S	S		G(-10)A	R		A74S
SP11	Beijing	R	S315T		R	S450L	R	D354A	R		C517T	S	S	R		G(-37)T	R		D94G
SP12	Beijing	R	S315T		R	S450L	R	M306V	R	K43R		R	S	R	A1401G		S		
SP13	Beijing	R	S315T		R	S450L	R	N296H	R	K43R		S	S	R		G(-10)A	S		
SP15	Beijing	R	S315T		R	S450L	S		R		C517T	S	S	S			S		
SP16	Beijing	R	S315T		R	S450L	R	M306I G406D	R	K43R		S	S	S			S		
SP17	Beijing	R	S315T		R	S450L	R	Q497R	R			R	R	S		C(-14)T	S		
SP18	Beijing	R	S315T		R	S450L	R	D354A	R		C517T	S	S	R		G(-37)T	S		
SP21	Beijing	R	S315T		R	S450L	R	N296H	R	K43R		R	S	R	A1401G		R	T500I	D94N
SP22	Beijing	R	S315T		R	S450L	S	M306I	R	K43R		R	S	R	A1401G		R		
SP23	Beijing	R	S315T	C(-15)T	R	S450L	R	M306V	R	K43R		R	S	R	A1401G		R		D94G
SP24	Ural	R	S315T	C(-15)T	R	S450L	S	S297A	R	K88R		S	S	S		C(-12)T	R	T500A	A90V
SP25	Ural	R	S315T	C(-15)T	R	S450L	R	S297A	R	K88R		R	S	R	A1401G	C(-12)T	R		D94Y
SP26	Beijing	R	S315T		R	S450L	R	M306V	R		C517T	R	R	S		C(-14)T	R		D94A
SP27	Beijing	R	S315T		R	S450L	R	M306V	R	K43R		R	R	S		C(-14)T	R		D94G
SP28	Ural	R	S315T		R	S450L	R	M306I	R			R	R	R	A1401G		R		D94Y
SP29	Beijing	R	S315T		R	D435Y	R		R	K43R		R	S	R	G1484T	C(-12)T	R		S91P
SP30	Beijing	R	S315T		R	S450L	R	M306V	R		C517T	S	R	S		C(-14)T	R	A504V	S91P
SP31	Ural	R	S315T		R	D435V	S		R	K88R		S	S	S			S		
SP32	Ural	R	S315T		R	S450L	R	M306V	R	K43R		S	S	S			R		S91P
SP33	Ural	R	S315T		R	S450L	R	N399T	R	K88R		S	S	S			R		D94G

SP34	Beijing	R	S315T		R	S450L	R	M306V	R	K43R		S	S	S			S		
SP35	LAM	R	S315T	C(-15)T	R	S450L	S	Y319S	R		A514C	S	S	S			S		
SP36	LAM	R	S315T	C(-15)T	R	S450L	S	Y319S	R		A514C	S	S	S			S		
SP37	LAM	R	S315T	C(-15)T	R	D435V	R	M306I	R		A514C	R	R	R	A1401G		S		
SP38	LAM	R	S315T	C(-15)T	R	D435V	R	M306I	R		A514C	R	R	R	A1401G		R		D94A
SP39	LAM	R	S315T	C(-15)T	R	S450L	S	Q497R	R		A514C	S	R	R	A1401G	C(-12)T	R		A90G
SP41	Ural	R	S315T		R	S450L	S	M306I	R	K88R		S	S	R		G(-10)A	S		
SP42	Ural	S			S		S		S			S	S	S			S		
SP45	Beijing	S			S		S		S			S	S	S			S		
SP46	LAM	S			S		S		S			S	S	S			S		
MOS2	LAM	R	S315T	C(-15)T	R	D435V	R	Y319S	R		A514C	S	-	R	A1401G		R		D94G
MOS7	LAM	R	S315T	C(-15)T	R	D435V S431R	R	M306I	R		A514C	R	R	R	A1401G		R		D94G
MOS9	Ural	R	S315T	G(-17)T	R	H445D	S	G406A	R		A514C	S	S	S		G(-10)A	S		
MOS10	Beijing	R	S315T	C(-15)T	R	H445Y	R	M306V	R	K43R		R	-	R	A1401G		R		A90V
MOS11	Beijing	R	S315T	G(-17)T	R	S450L	R	M306V	R	K43R		R	-	R	A1401G		R		A90V D94G
MOS12	Beijing	R	A109T	C(-15)T	R	S450L	R	M306V	R	K43R		R	R	R	A1401G		R		D94G
MOS14	Beijing	R	S315T		R	S450L	R	M306V	R	K43R		R	R	R	A1401G		R		D94G
MOS15	Beijing	R	S315T		R	S450L	S	M306V	R	K43R		S	S	S			S		
MOS16	Beijing	R	S315T		R	S450L	R	M306I	R	K43R		S	S	S		G(-10)A	S		
MOS17	Beijing	R	S315T		R	S450L	S	D354A	S		C517T	S	S	R		G(-37)T	S		
MOS18	Beijing	R	S315T		R	S450L	R	D354A	R		C517T	S	S	R		G(-37)T	S		
CTRI-2	LAM	S			S		S		S			S	S	S			S		
CTRI-3	Ural	R	S315T	C(-15)T	R	D435V	R	M306I	R			R	R	-	A1401G		S		
CTRI-4	Beijing	R	S315T		R	D435V	R	M306V	R	K88R		R	R	-			R		

Примечание: изониазид (INH), рифампицин (RIF), этамбутол (EMB), стрептомицин (STR), канамицин (KAN), амикацин (AMK), капреомицин (CAP) и офлоксацин (OFX).

3.10 Построение гипотез формирования устойчивости к противотуберкулезным препаратам

Для построения гипотез формирования лекарственной устойчивости к ПТП был проведен анализ геномов 100 образцов МБТ (проведено совместно с Ищенко Д.С.). Представленная выборка включала 54 образца, секвенированных в ходе исследования, а также 46 образцов, загруженных из базы данных NCBI. Для всех образцов были известны как генотипы, так и профили лекарственной устойчивости к семи препаратам (таблица 20).

Таблица 20.

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

Лекарственная устойчивость		Генетическое семейство, N			Всего
		Beijing, 72	LAM, 13	Ural, 15	
INH	S	9	4	3	16
	R	63	9	12	84
RMP	S	12	5	4	21
	R	60	8	11	79
STR	S	11	5	4	20
	R	61	8	11	80
EMB	S	23	9	8	40
	R	49	4	7	60
OFX	S	44	9	10	63
	R	28	4	5	37
KAN	S	32	8	11	51
	R	40	5	4	49
CAP	S	52	10	12	74
	R	20	3	3	26

Согласно нашему предположению мутации, ассоциированные с лекарственной устойчивостью, должны встречаться чаще в устойчивых, чем в чувствительных к ПТП штаммах. Для проверки выдвинутого предположения был проведен анализ распределения однонуклеотидных полиморфизмов в геномах изучаемой выборки образцов. Для увеличения точности анализа были выбраны следующие критерии отбора SNPs: 1) покрытие однонуклеотидного полиморфизма не менее тремя чтениями с

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

* гены и межгенные участки, для которых ассоциация с лекарственной устойчивостью показана ранее

4. ОБСУЖДЕНИЕ

4.1. Формирование экспериментальных групп образцов ДНК штаммов *M. tuberculosis*

Исследование проводилось на препаратах ДНК клинических изолятов *M. tuberculosis*, предоставленных на разных этапах выполнения работы соответствующими противотуберкулезными центрами.

Большой объем проанализированных образцов ($n = 500$), качественная микробиологическая, клиническая и генетическая характеристика исходных штаммов позволили сформировать основные экспериментальные группы. В том числе, 310 образцов были задействованы в разработке инновационного метода сполиготипирования, базирующегося на применении реакции удлинения зонда с последующей MALDI-ToF масс-спектрометрической детекцией (Shitikov *et al.*, 2012), и 54 образца составили группу для проведения полногеномного секвенирования. В последнюю группу вошли образцы, относящиеся к эндемичным для России генотипам, что позволяет изучить генетическую вариабельность на уровне семейств. Кроме этого, каждое семейство было представлено как лекарственно-чувствительными, так и лекарственно-устойчивыми микобактериями (приложение 1). Последнее, в свою очередь, дало возможность оценить вариабельность внутри семейства и провести поиск новых детерминант устойчивости.

4.2 Сполиготипирование штаммов *M. tuberculosis*

Как было упомянуто ранее, для молекулярно-генетического типирования микобактерий наиболее часто используются следующие методы: IS6110 RFLP-анализ (van Embden *et al.*, 1993), VNTR-типирование (Supply *et al.*, 2006) и сполиготипирование (Kamerbeek *et al.*, 1997). Выбор сполиготипирования для описания образцов коллекции был обусловлен возможностью сравнения полученных паттернов 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, в нашем случае успешность его применения определяется быстротой и надежностью выявления эндемичных семейств популяции.

В настоящем исследовании разработана эффективная и быстрая методика определения структуры DR-региона, базирующаяся на реакции удлинения зонда с последующим MALDI-ToF масс-спектрометрическим анализом. Данный подход ранее был успешно применен на практике коллегами из Германии. Honish с соавторами (Honisch *et al.*, 2010) для определения структуры DR-региона применяли закрытую коммерческую систему компании Sequenom, которая позволяла определить наличие/отсутствие 43 спейсеров с использованием двух реакций удлинения зонда. Для упрощения использования метода в практике нами было увеличено количество реакций для одного образца, что снижало сложность интерпретации результата. В предложенном методе олигонуклеотидные зонды, соответствующие спейсерам, были сгруппированы в восемь систем, содержащих по 5–6 зондов. Согласно исследованиям Sauer (Sauer, 2006) данное количество зондов в составе одной реакционной смеси позволяет повысить производительность системы без снижения ее стабильности, что часто используется в практике.

Разработанный метод может быть реализован в формате 96- и 384-луночного планшета, позволяя анализировать до 48 образцов за 8 часов. Для сравнения, сполиготипирование 45 образцов по классической методике занимает около 24 часов. Другим достоинством метода является высокая чувствительность масс-спектрометрического анализа. Современные масс-спектрометры позволяют детектировать фемтомоль вещества, что дает возможность анализировать образцы с исходно малым количеством ДНК.

Представленная платформа с использованием масс-спектрометрии позволяет проводить дополнительные исследования нуклеиновых кислот и белков. Помимо сполиготипирования на уровне нуклеиновых кислот

возможен анализ генов, ассоциированных с устойчивостью (Ikryannikova *et al.*, 2007; Afanas'ev *et al.*, 2011). В свою очередь прямое белковое профилирование позволяет осуществлять видовую идентификацию микобактерий (Hettick *et al.*, 2006; Pignone *et al.*, 2006; Shitikov *et al.*, 2012). Таким образом, данная технология может служить одним из компонентов комплексного анализа микробного сообщества, объединенного единой инструментальной базой (MALDI-ToF масс-спектрометр), и потому быть востребованной для использования в практике клинко-диагностических лабораторий.

В представленном исследовании разработанный метод был апробирован на 310 образцах с известными паттернами сполиготипирования и показал полную сходимость результатов с данными классического сполиготипирования. В связи с этим мы использовали его для определения сполиготипов оставшихся 190 образцов *M. tuberculosis*. Как и ожидалось, разрешающая способность метода оказалась низкой (HGDI=0.519). В свою очередь для образцов семейства Beijing значение индекса HGDI составило всего 0.105, а для группы образцов, не относящихся к этому семейству, разрешающая способность метода показала высокое значение индекса (HGDI=0.906). К сожалению, полученные данные не могут быть использованы для описания молекулярно-генетических особенностей региональной микробной популяции из-за не случайности выбора образцов. Однако следует отметить, что к преобладающим генетическим семействам коллекций профильных противотуберкулезных учреждений относились Beijing, Ural и LAM, что соответствует исследованиям других авторов (Mokrousov *et al.*, 2009; Mokrousov, 2012; Mokrousov *et al.*, 2012).

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

На основании суммарно накопленной информации была сформирована выборка из 150 образцов, наиболее полно охарактеризованных с клинко-эпидемиологической точки зрения, информация о которых представлена в

приложении 1 и таблице 22. Представители семейства Beijing в совокупной выборке составили 72 %. Следует отметить, что в России частота встречаемости данного семейства варьирует от 50 % до 80 % (Норкина с соавт., 2003; Mokrousov *et al.*, 2003). Семейства LAM и Ural были выявлены среди 8.7 % и 6.7 % образцов, соответственно. Согласно литературным данным частота встречаемости данных семейств в России в среднем составляет 15 % и 10 %, соответственно (Шемякин, 2003; Mokrousov, 2012).

Таблица 22.

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), Ural (10) и LAM (9). В каждую группу были включены как лекарственно-чувствительные, так и лекарственно-устойчивые бактерии (приложение 1).

Считывание геномных последовательностей было проведено на секвенаторе GS FLX+ Series — XL+ (Roche) с использованием фрагментных баркодированных библиотек. Представленный прибор относится к секвенаторам нового поколения и позволяет в короткие сроки определить геномную последовательность изучаемого образца. По сравнению с другими приборами полногеномного секвенирования особенность данной платформы заключается в том, что получаемые чтения обладают достаточно большой длиной (около 800 нуклеотидных оснований), при этом точность

определения последовательности одного прочитанного фрагмента составляет не менее 99 %. Таким образом, полученные данные представляются достоверными, а чтения, в силу большей длины, могут быть использованы не только для определения однонуклеотидных полиморфизмов, но и рекомбинационных событий в геноме.

Полученные чтения с прибора были использованы как для сборки *de novo*, так и для выравнивания на референтный геном *M. tuberculosis* штамма H37Rv. Сборка *de novo* использовалась для оценки качества секвенирования и корректировки дальнейших запусков прибора. Зависимость количества собранных контигов от покрытия представлена на рисунке 14. В ходе анализа было установлено, что начиная с покрытия 14 кол-во контигов составляет 100+/-15. При данном покрытии количество собранных контигов в большей степени зависит от присутствия повторяющихся последовательностей генома, а не от участков с нулевым покрытием.

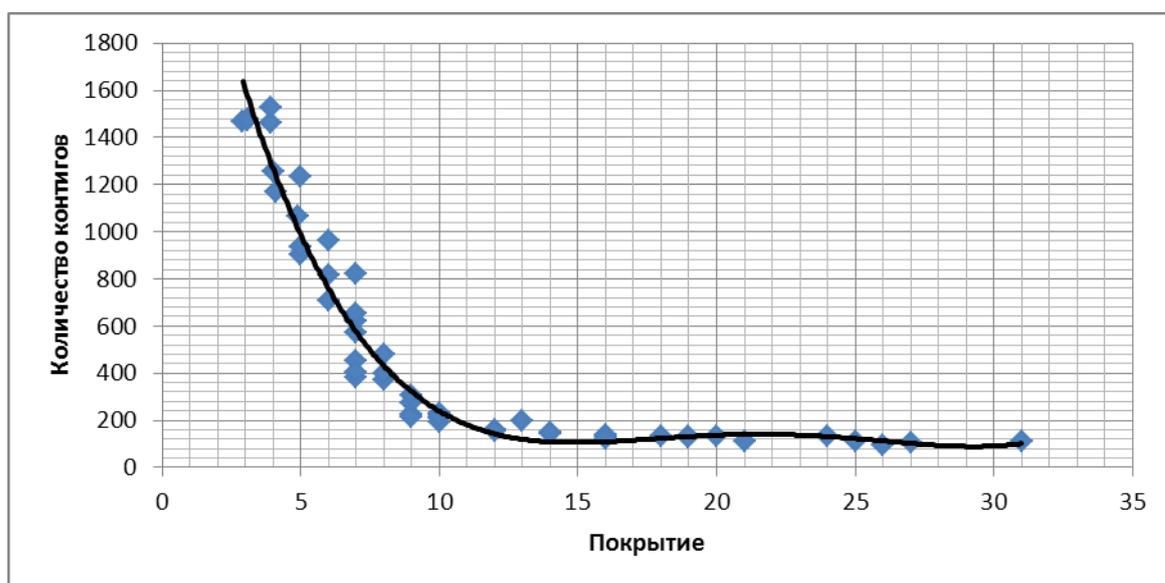


Рисунок 14. Зависимость количества контигов от покрытия генома для секвенированных образцов

В ходе выравнивания чтений с секвенатора были идентифицированы однонуклеотидные замены в изучаемых образцах относительно генома референтного штамма H37Rv. 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 (Achtman, 2008). Для оценки силы и направленности естественного отбора, действующего на геном патогена, использовали отношение dN/dS (таблица 12). 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 (Fleischmann *et al.*, 2002; Zhang *et al.*, 2013). В свою очередь следует отметить, что дисперсионный анализ значений dN/dS показал достоверные отличия между исследуемыми семействами ($p = 3.1 \times 10^{-5}$) (рисунок 15). Полученные данные свидетельствуют о разной степени влияния естественного отбора.

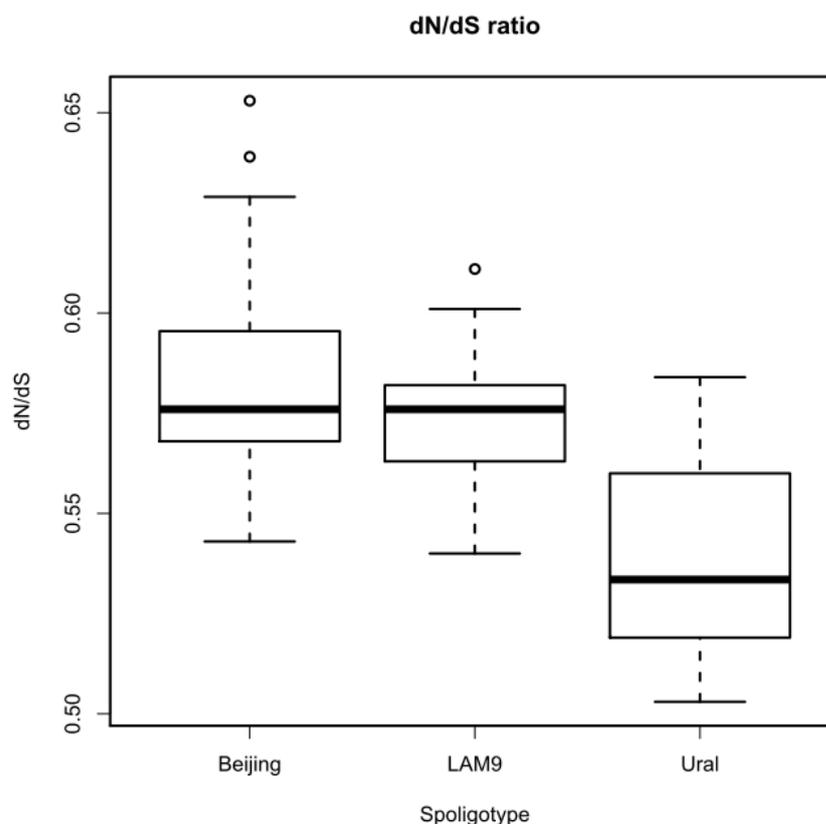


Рисунок 15. Дисперсионный анализ значений dN/dS для семейств Beijing, LAM и Ural. Жирной горизонтальной линией отмечено медианное значение dN/dS, границами прямоугольной области служат 25-й и 75-й процентиля.

По результатам первичного анализа геномных последовательностей образцы CTRI-2, SP21 и SP28 были отобраны для определения полных геномов. Данные образцы относились к семействам LAM, Beijing и Ural, соответственно. При первичной сборке генома *de novo* для образца CTRI-2 было получено более 500 контигов. Полученные последовательности были выравнены на референтный геном H37Rv и использовались для конструирования праймеров для дальнейшего соединения контигов секвенированием по Сенгеру. Разрывы контигов были обусловлены наличием повторяющихся последовательностей и областей с нулевым покрытием. Участки с нулевым покрытием наиболее часто были связаны с высоким G+C составом генома патогена и обусловлены необходимостью проведения амплификации перед полногеномным секвенированием [24]. Пример отсутствия чтений из-за высокого GC-состава представлен на рисунке 16.

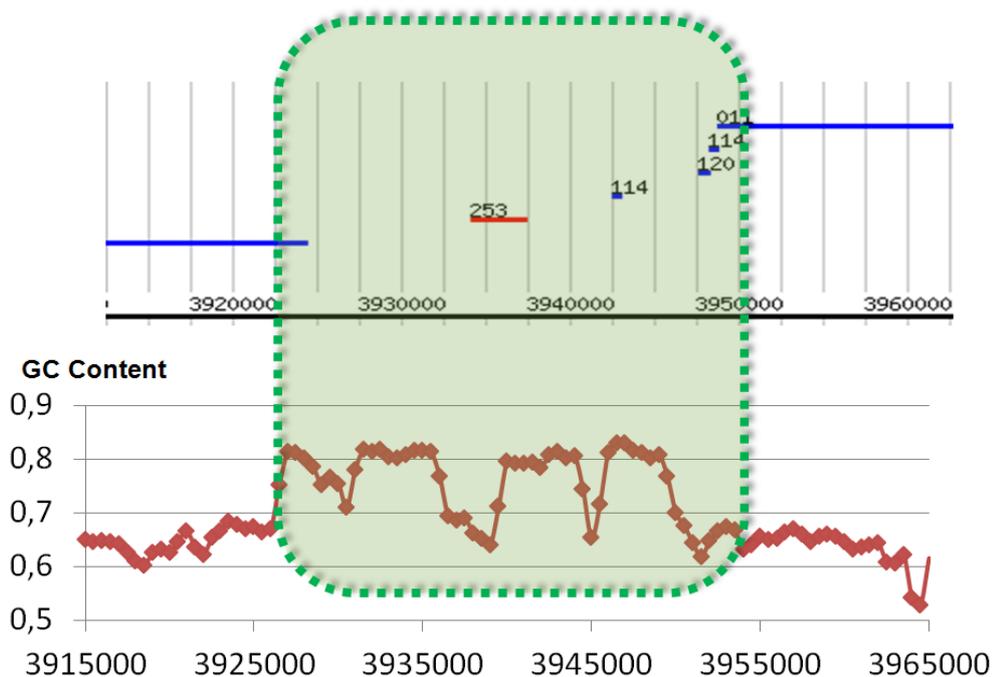


Рисунок 16. Наложение результатов выравнивания контигов генома CTRI-2 на референтный геном H37Rv в зависимости от GC-состава. Контиги обозначены горизонтальными линиями (синий цвет – обратная ориентация относительно H37Rv; красный – прямая). На горизонтальной шкале отмечена область генома H37Rv.

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) (Pina *et al.*, 2013).

4.4 Сравнительный филогенетический анализ секвенированных штаммов

В контексте данного исследования филогенетический анализ использовался для оценки разнообразия секвенированных образцов и дальнейшего поиска семейство-специфических полиморфизмов. Дополнительно для сравнения полученных результатов с мировыми данными в анализ было включено 7 образцов *M. tuberculosis*, представленных в GenBank в виде протяженных секвенсов (таблица 6).

Для построения филогенетического дерева были использованы все полиморфизмы представленных геномов. 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).

Анализ филогенетического дерева выявил три больших генетических кластера, которые сформировали образцы семейств Beijing, Ural и LAM (рисунок 7). Следует отметить, что образцы генетических семейств Ural и LAM располагаются на одной внутренней ветви, что свидетельствует об их эволюционном родстве. Данные семейства относятся к распространенной Евро-Американской линии и несут схожие филогенетические маркеры (таблица 14). Согласно Sreevatsan с соавторами (Sreevatsan *et al.*, 1997) и Baker с соавторами (Baker *et al.*, 2004) данные семейства относятся к главной генетической группе 2 и линии 2, соответственно. Однако на основании однонуклеотидных полиморфизмов по Filliol с соавторами (Filliol *et al.*, 2006) они соответствуют разным генетическим кластерам: LAM – кластер 5, Ural –

кластер 3b. Следует отметить, что полученные данные по классификации семейства Ural являются новыми и ранее не описаны в литературе.

Наиболее обширной группой на филогенетическом дереве было семейство Beijing (35 образцов) (рисунок 7). Представители данного семейства относятся к Восточно-Азиатской линии и характеризуются набором специфических молекулярных маркеров (таблица 3). Следует отметить, что ветвь Beijing была разделена на две клады. Одна из них состояла из образцов SP17, CTRI-4 и 94_M4241A, в свою очередь другая клада включала все оставшиеся образцы генотипа Beijing. Согласно литературным данным семейство Beijing делится на атипичные/древние и типичные/современные штаммы (Tsolaki *et al.*, 2004; Mokrousov *et al.*, 2005; Tsolaki *et al.*, 2005). На сегодняшний день существуют различные методики определения древних и современных Beijing: на основании RD (Tsolaki *et al.*, 2005; Hanekom *et al.*, 2007; Reed *et al.*, 2007); по наличию IS6110 в NTF регионе генома (Kurepina *et al.*, 1998; Mokrousov *et al.*, 2002; Mokrousov *et al.*, 2006); на основании SNP анализа (Ebrahimi-Rad *et al.*, 2003; Schurch and van Soolingen, 2012). Согласно SNP анализу по Rad с соавторами (Ebrahimi-Rad *et al.*, 2003) и Shurch с соавторами (Schurch and van Soolingen, 2012) выявленные в представленном филогенетическом анализе клады соответствовали древним и современным Beijing. К наиболее обширной ветви относились образцы современных Beijing, что соответствует литературным данным о высокой представленности данной группы в мире (Hanekom *et al.*, 2007; Iwamoto *et al.*, 2008; Schurch *et al.*, 2011).

С целью получения наиболее полной информации о генетическом разнообразии коллекции секвенированных образцов дополнительно было проведено VNTR-типирование по 24 локусам генома. В ходе анализа было выявлено 29 уникальных профилей, что оказалось существенно больше 8 SIT паттернов, полученных при использовании метода сполиготипирования. Как и ожидалось, разрешающая способность 24-локусного VNTR-анализа

значительно превосходила таковую для сполиготипирования – 0.944 и 0.519, соответственно.

Применение VNTR-анализа позволило разделить современные штаммы Beijing на 12 паттернов (приложение 2). Семь из них были уникальными, в то время как 5 паттернов включали в себя от 2 до 9 кластеризующихся образцов. Согласно исследованиям Мокроусова с соавторами (Mokrousov *et al.*, 2008) представленные паттерны можно отнести к 3 типам: M2 (17 образцов), M11 (15 образцов) и M33 (1 образец). Вариант M33 представлен единственным образцом SP8 и отличается от образцов типа M2 по локусу MIRU26. Следует отметить, что образцы типа M11 на филогенетическом дереве образуют монофилетическую группу (рисунок 7, образцы MOS14-SP23). Образцы типов M2 и M33 также входят в состав единой ветви. Полученные результаты в полной мере соответствуют эпидемиологическим данным по России, согласно которым типы M2 и M11 составляют около 80 % изолятов генотипа Beijing (Норкина с соавт., 2003; Mokrousov *et al.*, 2003).

С целью дополнительной классификации представленных семейств на основании однонуклеотидных полиморфизмов был проведен анализ SNPs в представленных геномах. Согласно работам Schurch с соавторами (Schurch and van Soolingen, 2012) и Mestre с соавторами (Mestre *et al.*, 2011) изучаемые семейства попадают в одну группу, представители которой распространены по всему миру. Однако по результатам другой работы Schurch с соавторами (Schurch *et al.*, 2011) на основании анализа RD-регионов и SNPs семейства удалось причислить к разным генетическим группам. Образцы типа M2 относились к группе BST18, в то время как образцы типа M11 – к BST19. Следует отметить, что согласно исследованию изоляты этих групп также были собраны по всему миру.

В связи с выше сказанным особый интерес представляет более детальный анализ однонуклеотидных полиморфизмов в геномах различных семейств *M. tuberculosis*. Развернутое описание по поиску генотип-специфических SNPs приведено в следующем разделе.

4.5 Определение и валидация генотип-специфических мутаций

Сравнительный анализ данных полногеномного секвенирования образцов семейств Beijing, Ural и LAM позволил идентифицировать генотип-специфические мутации (рисунок 8). Как описано выше, представленные семейства относятся к двум из шести основных филогенетических линий: Евро-Американской (LAM и Ural) и Восточно-Азиатской (Beijing). Из рисунка 8 видно, что образцы семейства Beijing характеризуются большим количеством полиморфизмов по сравнению с представителями генотипов Ural и LAM. Данное обстоятельство связано с тем, что образцы семейств Ural, LAM и референтный геном *M. tuberculosis* штамм H37Rv, относительно которого был проведен поиск однонуклеотидных полиморфизмов, принадлежат к одной филогенетической линии. Большое количество общих SNPs между всеми генотипами связано как с ошибками референтного генома H37Rv (Koser *et al.*, 2012), так и с его полиморфизмами относительно изучаемых образцов.

С целью верификации найденных полиморфизмов в работу было включено 366 геномов *M. tuberculosis*, представленных в базе данных NCBI. Согласно исследованиям Nomolka с соавторами (Nomolka *et al.*, 2012), а также построенным филогенетическим деревьям указанные геномы были отнесены к шести основным филогенетическим линиям (таблица 15). На основании полученных полиморфизмов была проведена корректировка генотип-специфических SNPs (приложения 3, 4 и 5). Таким образом, в результате проведенного исследования были получены паттерны семейство-специфических полиморфизмов, которые могут быть использованы как для достоверной дифференциации штаммов *M. tuberculosis*, так и для изучения влияния генетических изменений на фенотип патогена на уровне семейств.

В частности, было проанализировано распределение по COG категориям белков, содержащих аминокислотные замены, характерные для изучаемых семейств (рисунок 9). Согласно проведенному анализу штаммы семейства Beijing достоверно отличались от Ural и LAM накоплением замен

в белках, относящихся к категориям O (посттрансляционные модификации, сборка белков, шапероны) и H (транспорт ко-ферментов и метаболизм). Количество замен в белках этих категорий 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 COG категориях L (репликация, репарация и рекомбинация) и N (клеточная подвижность). При этом наибольший интерес представляют замены в белках категории L по сравнению с N, к которому относится большое количество генов семейства PE/PPE. Белки категории L ассоциированы с системой репликации, репарации и рекомбинации и представляются весьма интересным с точки зрения формирования мутаторного генотипа. Полученные результаты для генетического семейства LAM являются новыми для российской популяции МБТ и не в полной мере согласуются с данными мировой литературы по данному семейству (Das *et al.*, 2009). Дополнительно интересным оказался тот факт, что специфических замен в данной категории белков для генотипа LAM было больше, чем для генотипа Beijing, для которого по одной из гипотез ассоциация с лекарственной устойчивостью объясняется наличием мутаций в так называемых генах мутаторов, связанных с системой репарации ДНК (Dos Vultos *et al.*, 2008). Штаммы семейства Ural достоверно отличались от LAM и Beijing накоплением замен в белках, относящихся к COG категории Q (биосинтез вторичных метаболитов транспорт и катаболизм).

Оценить генетическую вариабельность образцов внутри семейства помогло построение матрицы расстояний (матрица попарных расстояний между образцами) на основании расстояний Хэмминга (по количеству различных нуклеотидов). Для уменьшения размерности матрицы был применен метод главных компонент. На рисунке 17 представлены результаты анализа, а также в виде боксплотов отображены внутригрупповые попарные расстояния между образцами. Проведенный дисперсионный анализ выявил существенные различия между генотипами ($p < 2.2e-16$). Согласно

исследованию наибольшим генетическим разнообразием обладали представители генотипа Ural. В свою очередь члены семейства LAM характеризовались наименьшим разнообразием, несмотря на относительно большое количество полиморфизмов в системе репликации, репарации и рекомбинации (рисунок 17Б). Для образцов семейства Beijing можно отметить довольно плотную кластеризацию. Исключение составляют два образца (на рисунке 17А обведены в пунктирный круг), отстоящие на значительном расстоянии от остальных Beijing. Согласно ранее проведенному филогенетическому анализу данные штаммы принадлежат к древним Beijing.

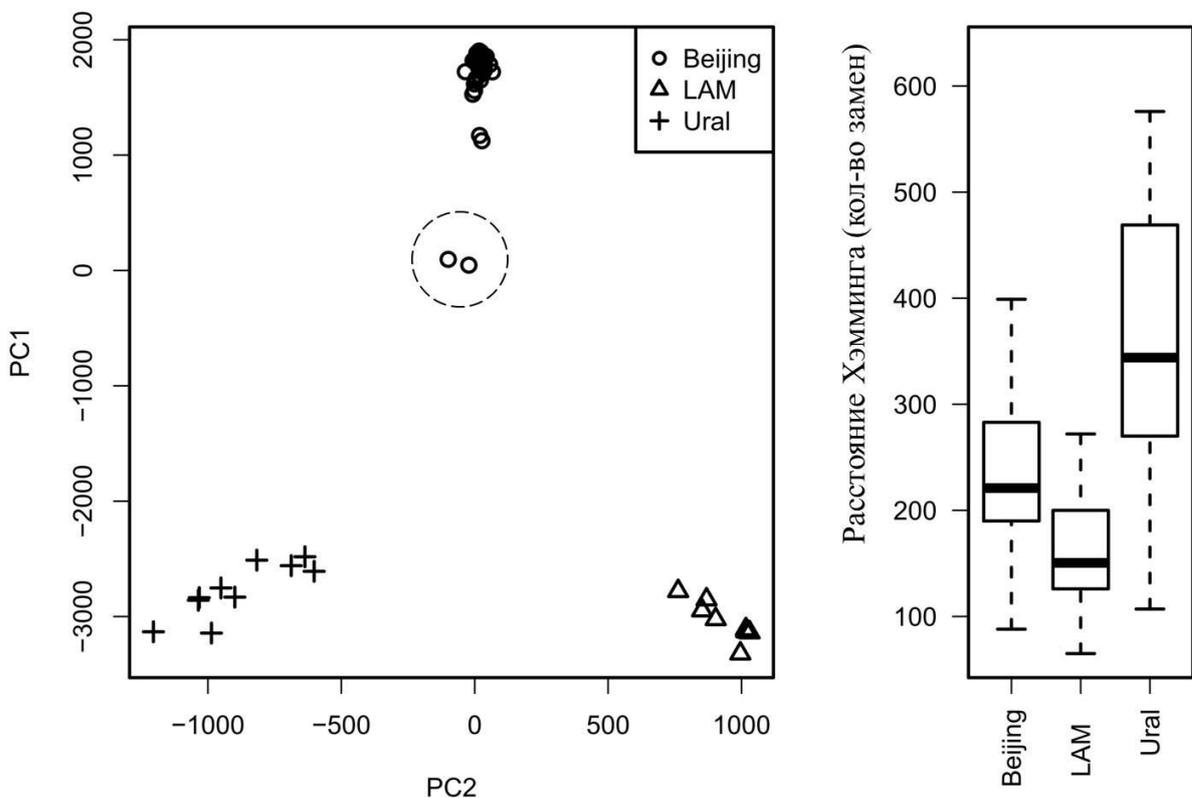


Рисунок 17. Оценка генетической вариабельности семейств. А — метод главных компонент. Окружностью выделены два образца, относящиеся к древним Beijing. Б — боксплот, показывающий разбросы внутрисемейственных расстояний между образцами. Жирной горизонтальной линией отмечено медианное значение, границами прямоугольной области служат 25-й и 75-й процентиля.

Применение данного метода не позволило выявить отличия между описанными ранее типами M2 и M11 генетического семейства Beijing. Тип

M11 занимает более предковое положение по сравнению с типом M2 и превалирует в Евразии (Mokrousov, 2008). Согласно Мокроусову с соавторами данный тип включает образцы кластера B0/W148, широко и недавно распространившегося в России. 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 и увеличенной трансмиссивностью (Mokrousov, 2013). В связи с этим мы решили сконцентрировать внимание на генетических особенностях представителей данного кластера с целью исследования механизмов их адаптации и выживания в организме хозяина.

4.6 Характеристика кластера Beijing B0/W148

Согласно проведенному филогенетическому анализу и VNTR-типированию 12 из 15 секвенированных образцов типа M11 относились к кластеру Beijing B0/W148 (рисунок 7, образцы с W-148 по SP23). Данные образцы располагались на единой внутренней ветви филогенетического дерева и были удалены от трех других представителей типа M11 (рисунок 7, образцы MOS14-SP6). Следует отметить, что согласно VNTR-анализу группы отличались по единственному локусу QUB26. В образцах кластера Beijing B0/W148 в данном локусе было выявлено 7 тандемных повторов, тогда как в других геномах типа M11 таких повторов было шесть. Дополнительно принадлежность образцов к кластеру была проверена методом ПЦР согласно публикации Мокроусова с соавторами (Mokrousov *et al.*, 2012). Образцы, не принадлежащие к кластеру Beijing B0/W148, показали отрицательные результаты.

Методами сравнительной геномики для образцов кластера Beijing B0/W148 был определен характерный мутационный профиль, включающий 63 полиморфизма. Для дальнейшего анализа были использованы полиморфизмы в межгенных областях и полиморфизмы, приводящие к

несинонимичным заменам (таблица 16). Из семи мутаций, найденных в межгенных областях, только одна (С(-15)Т) находилась в промоторной области оперона Rv0789c-Rv0792c, в то время как другие замены находились на значительных расстояниях от сайтов инициации транскрипции. При этом генам Rv0789c-Rv0791c соответствуют белки с неизвестной функцией. Продуктом гена *gntR* (Rv0792c) является белок, относящийся к HutC подсемейству регуляторов транскрипции, однако его функции не изучены в полной мере.

В ходе анализа 27 несинонимичных замен было установлено, что три мутации приводят к терминации трансляции. Так, например, для гена *chiZ* (Rv2719c), состоящего из 165 кодонов, нонсенс-мутация произошла в 150 кодоне. Данный ген интересен тем, что он кодирует гидролазу клеточной стенки, индуцируемую повреждением ДНК или во время роста бактерии в макрофаге, и является регулятором клеточного деления (Vadrevu *et al.*, 2011). Гиперпродукция данного белка приводит к дестабилизации септального кольца и, тем самым, нарушению нормального деления клетки. Следует отметить, что белок состоит из трех доменов: N-концевого (обладает гидролитической активностью), трансмембранного и С-концевого (содержит пептидогликан-связывающий участок). Найденная в ходе исследования нонсенс-мутация приводит к нарушению синтеза пептидогликан-связывающего домена, функция которого, к сожалению, не до конца выяснена. Две другие кластер-специфические нонсенс-мутации произошли в генах Rv3728 и Rv2800. Ген Rv3728 кодирует двух-доменный консервативный мембранный белок эффлюксной системы, относящийся к суперсемейству Major facilitator и ассоциированный с лекарственной устойчивостью (Gupta *et al.*, 2010). В свою очередь ген Rv2800 кодирует белок гидролазу, участвующий в клеточном метаболизме. Однако в контексте представленного исследования мы не можем судить о значимости описанных мутаций и степени повреждения белков. Дополнительно в ходе анализа удалось выявить существенные аминокислотные замены в белках,

ответственных за метаболизм патогена (Rv0338c и Rv1833c) и транспорт веществ через мембрану (Rv1877 и Rv0928).

Таким образом, описанные на основании сравнительной геномики полиморфизмы отчасти объясняют успешность кластера Beijing B0/W148, и могут служить основой для дальнейших прицельных исследований, а также создания систем генетического мониторинга указанного семейства.

Рассмотрев особенности представителей кластера Beijing B0/W148 на уровне SNPs, следующей задачей было определение структурной организации генома. Как написано ранее, в данной работе использовался секвенатор нового поколения GS FLX+ Series — XL+ (Roche), позволяющий получать протяженные чтения (более 800 п.о.). Анализ данных с прибора указал на вероятные различия в распределении повторяющегося элемента IS6110 относительно референтного генома H37Rv, однако в связи с использованием фрагментных библиотек оценить геномные перестройки в полной мере не представлялось возможным.

В ходе филогенетического анализа было установлено сходство секвенированных штаммов кластера Beijing B0/W148 со штаммом W-148, представленным в GenBank в виде одного скаффолда. Дальнейшее выравнивание геномов W-148 и H37Rv показало наличие двух больших инверсий в геноме W-148 (рисунок 10), которые впоследствии были подтверждены для всех представителей кластера в имеющейся коллекции (Shitikov, 2014).

Следует отметить, что крупные хромосомные перестройки в геномах представителей рода *Mycobacterium* были описаны совсем недавно в ряде статей (Garcia-Betancur *et al.*, 2012). Так, например, было показано наличие дубликации длиной 56 т.п.о. в геноме *Mycobacterium smegmatis* штамма mc(2)155 относительно предкового генома ATCC 607. Описанная дубликация располагалась между двумя копиями повторяющегося элемента IS1096 (Wang *et al.*, 2008). В свою очередь методами сравнительной геномики штаммов *Mycobacterium bovis* BCG было установлено наличие двух крупных

хромосомных дупликаций: DU1 и DU2. При этом дупликация DU1 была найдена только в геномах штаммов BCG Pasteur, а четыре различные формы дупликации DU2 были найдены в разных штаммах BCG (Brosch *et al.*, 2000). В литературе также описаны два случая возникновения крупных дупликаций в геноме штаммов МБТ, относящихся к генетическим линиям 2 и 4 (Domenech *et al.*, 2010; Weiner *et al.*, 2012). Некоторые из описанных дупликаций так же были заключены между элементами IS6110, подтверждая тот факт, что большинство крупных геномных перестроек связано с мобильными генетическими элементами или повторами (Смирнов, 2010).

Рекомбинационные события другого рода, инверсии, были обнаружены в геномах штаммов *Mycobacterium avium* и только единожды упоминаются для *M. tuberculosis*. Так для трех штаммов МБТ из Южной Африки (провинция KwaZulu-Natal), секвенированных в Broad Institute, было показано наличие крупной инверсии сегмента хромосомы длиной около 2.5 м.п.о. относительно области начала репликации. Однако повторное секвенирование одного из штаммов в другой лаборатории не подтвердило наличие данной перестройки (Joerger *et al.*, 2009). В геноме *M. avium* инверсия была обнаружена между подвидами *hominissuis* и *paratuberculosis* (Hsu *et al.*, 2011).

В данной работе были обнаружены крупные хромосомные инверсии в геноме *M. tuberculosis* штаммов кластера Beijing B0/W148 (рисунок 10). Следует отметить, что рекомбинационные события такого масштаба показаны для МБТ впервые. Инверсии были подтверждены для всех штаммов коллекции, относящихся к кластеру Beijing B0/W148, методом ПЦР с последующим секвенированием по Сенгеру, а также проведением полногеномного секвенирования одного образца с применением библиотек парных фрагментов (рисунки 11 и 12). Дополнительно были реконструированы возможные рекомбинационные события, произошедшие в геноме предкового штамма W-148 (рисунок 13). Согласно предложенной теории, эволюция штаммов проходила в два этапа. На первом этапе

произошла крупномасштабная инверсия сегмента генома длиной около 3 м.п.о. Данное предположение основано на том, что в данном случае границы инвертированного региона равноудалены от сайта терминации репликации. В литературе имеется много данных о геномных перестройках у других бактерий, происходящих вокруг *ter* региона (Eisen *et al.*, 2000) и возможно данный механизм так же был использован микобактериями. Между тем, остается не ясным, почему на границах инверсии были обнаружены только части элемента *IS6110*. На втором этапе эволюционных событий произошла инверсия блока III, частично восставившая ориентацию крупного сегмента генома. Как было отмечено ранее, блок III находится между повторами *IS6110* в геномах всех образцов генотипа Beijing, представленных в GenBank. Одной из характеристик повторов данного типа является дупликация 3–4 пар оснований непосредственно по краю сайта встраивания в момент интеграции (Dale, 1995). В работе было проверено наличие таких дупликаций в геномах образцов кластера Beijing B0/W148, а также в образцах других Beijing. В образцах не-B0/W148 Beijing была найдена дупликация оснований AGC в сайте интеграции *IS6110* между блоками II и III, в то время как основания CAG были дублированы между блоками III и IV (рисунок 18). В образцах B0/W148 последовательность дублированных триплетов в блоке III оказалась в той же ориентации, в то время как последовательность триплетов в блоках II и IV была инвертирована и перегруппирована, что соответствует происхождению W-148 от Предка 2 (Рисунки 15 и 18). В данном случае гомологичная рекомбинация между элементами *IS6110* является наиболее вероятным механизмом инверсии.

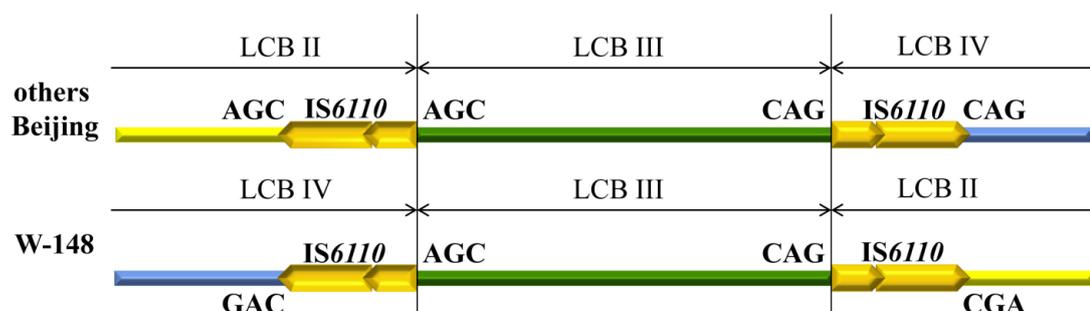


Рисунок 18. Схематическое представление дублированных оснований, расположенных по краям сайтов встраивания *IS6110* между блоками II–III и III–IV

Учитывая тот факт, что представленные рекомбинационные события являются уникальными для образцов кластера Beijing B0/W148 и в целом впервые описаны для *M. tuberculosis*, мы предприняли попытку проанализировать кластер-специфические полиморфизмы в генах системы репарации, рекомбинации и репликации (глава 3.8.1) (Muttucumaru and Parish, 2004; Dos Vultos *et al.*, 2008). К нашему сожалению, в ходе анализа не было обнаружено несинонимичных замен в изучаемых генах. Единственная синонимичная замена GGG – GGT была обнаружена в 269 кодоне гена *recF*, что в свою очередь вряд ли могло быть причиной столь крупных геномных перестроек. Исходя из этого, нами было сделано предположение, что данные события могут быть обусловлены предрасположенностью к рекомбинациям на генетическом уровне для генотипа Beijing в целом.

Для определения генетической сублинии секвенированных штаммов на основании литературных данных был проведен анализ наличия в геноме пяти специфических LSPs (RD105, RD207, RD181, RD150 и RD142) (Tsolaki *et al.*, 2005; Gagneux *et al.*, 2006). Согласно полученным данным все секвенированные образцы относились к Beijing сублинии 3 (делеция RD105, RD207 и RD181), так же как и штаммы с недавно описанными дупликациями (Domenech *et al.*, 2010; Weiner *et al.*, 2012). Соответствующие исследования описывают крупную дупликацию, произошедшую в образцах Beijing сублиний 3, 4 и 5, которая затрагивает 350 т.п.о. хромосомы между генами Rv3128с и Rv3427с. Стоит отметить, что дупликация произошла между двумя копиями IS6110, находящимися в одинаковой ориентации. Дополнительно, согласно работе Weiner с соавторами (Weiner *et al.*, 2012) граница дупликации для одного из образцов находилась в области гена Rv3327с, что соответствует границе коллинеарного блока V в геноме W-148. Интересно отметить, что данный регион так же соответствует RvD5 в геноме H37Rv (Brosch *et al.*, 1999).

Суммируя представленную 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.

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

Для поиска и анализа генетических изменений, приводящих к возникновению лекарственной устойчивости, в настоящем исследовании было проведено полногеномное секвенирование 54 эндемичных для России штаммов МБТ.

Согласно литературным данным подавляющее большинство изменений представлено точечными мутациями в генах, ассоциированных с устойчивостью к основным ПТП (Almeida Da Silva and Palomino, 2011). Эти мутации могут рассматриваться в качестве маркеров устойчивости, и в настоящее время все больше внимание уделяется разработке и адаптации методов их обнаружения в клинико-лабораторной практике.

В ходе исследования был проведен анализ генетических маркеров резистентности к основным наиболее распространенным ПТП первого и второго ряда: рифампицину, изониазиду, этамбутолу, стрептомицину, канамицину, амикацину, капреомицину и офлоксацину.

Согласно анализу микробиологических данных в исследуемой выборке 48 образцов были устойчивы как к изониазиду, так и к рифампицину, что было подтверждено на генетическом уровне в ходе полногеномного секвенирования. Аминокислотная замена Ser315Thr в KatG была выявлена у подавляющего количества образцов (47 из 48) независимо от генетического семейства. При этом один образец семейства Beijing имел мутацию в 109 кодоне гена *katG*, однако ее роль в формировании лекарственной устойчивости не доказана. Следует отметить, что мутация С(-15)Т в промоторной области *mabA(fabG) – inhA* оперона была выявлена у 9.3 % и 62.5 % ИН-устойчивых штаммов семейства Beijing и штаммов Евро-Американской линии, соответственно (OR 6.6; 95 % CI 1.6–27.7). Согласно исследованиям Casalli с соавторами (Casali *et al.*, 2014) настоящая мутация в

большей степени ассоциирована с устойчивостью к этионамиду, и из полученных результатов можно предположить, что у штаммов семейства Beijing реализуется другой механизм устойчивости к данному антибиотику.

Мутации в гене *rpoB* были обнаружены среди всех рифампицин-устойчивых образцов, причем мутация в кодоне 450 была выявлена несколько чаще у штаммов семейства Beijing, нежели чем у штаммов Евро-Американской линии: 90.6 % против 56.3 %. Следует отметить, что мутации в описанном кодоне являются наиболее часто встречаемыми среди RIF-устойчивых штаммов, распространенных по всему миру (Ramaswamy and Musser, 1998; Qian *et al.*, 2002). Дополнительно для секвенированных образцов был проведен поиск компенсаторных мутаций в генах *rpoA* и *rpoC* (Comas, 2011). В ходе исследования было выявлено, что анализируемые мутации встречаются в среднем у 31 % RIF-устойчивых образцов независимо от генетического семейства. При этом следует отметить, что наиболее часто встречалась мутация в 483 кодоне гена *rpoC*, приводящая к замене V483G. Для выявленных типов M2 и M11 генотипа Beijing также не удалось найти существенных различий в распределении компенсаторных мутаций. При этом в большом количестве рифампицин-устойчивых образцов типа M2 Beijing была обнаружена замена Glu761Asp в RpoB, роль которой предстоит выяснить в будущем. В заключении стоит отметить, что согласно исследованиям Comas с соавторами (Comas, 2011) доля компенсаторных мутаций среди МЛУ форм туберкулеза в странах с высоким бременем заболевания достигает 21 %. Полученные нами данные свидетельствуют о большей частоте встречаемости данных мутаций в российской популяции патогена, что может говорить о более легком распространении лекарственно-устойчивых форм в будущем.

Анализ распределения мутаций в *embB* гене 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 ($p < 0.05$). Дополнительно семь разных аминокислотных замен было выявлено среди этамбутол-устойчивых штаммов, однако их ассоциацию с лекарственной устойчивостью подтвердить не удалось. Необходимо подчеркнуть, что в настоящее время многие авторы ставят под сомнение связь мутаций в гене *embB* с устойчивостью к этамбутолу (Mokrousov *et al.*, 2002; Tracevska *et al.*, 2004). В нашем исследовании у 4 из 19 (21 %) EMB-чувствительных образцов были обнаружены мутации в кодоне 306.

Мутации в генах *rpsL* и *rrs*, ассоциированных с устойчивостью к стрептомицину, были выявлены у 44 образцов коллекции. Как упоминалось ранее, данный препарат был впервые применен в 1944 г. и в настоящий момент практически не используется в связи с большим процентом устойчивых к нему штаммов. Однако следует подчеркнуть, что найденные мутации в большей степени были сопряжены с генетическими семействами. Так, например, замена Lys88Arg, связанная с устойчивостью высокого уровня, была выявлена преимущественно в образцах семейства Ural (5 штаммов), в то время как для образцов генотипа Beijing такого рода замена была определена лишь единожды (OR 22; 95 % CI 2.3 207.0). В свою очередь все устойчивые образцы семейства LAM (7 штаммов) несли мутацию A514C в гене *rrs*. Данная мутация также была выявлена в одном образце семейства Ural. Мутация C517T в гене *rrs* была детектирована только у образцов, относящихся к типу M2 генотипа Beijing (8 штаммов), причем один образец данной группы был стрептомицин-чувствительным. Дополнительно следует отметить, что для трех стрептомицин-устойчивых штаммов нам так и не удалось найти мутаций в описанных генах.

Анализ мутаций, ассоциированных с устойчивостью к инъекционным препаратам, показал общую сходимость с результатами системного обзора, опубликованного Georghiou с соавторами в 2012 г. (Georghiou *et al.*, 2012). Коллеги из США проанализировали информацию по 22 независимым

публикациям, объединив таким образом 1585 штаммов, изолированных в 18 странах. Согласно исследованию наиболее часто встречаемой мутацией, приводящей к устойчивости к инъекционным препаратам, была мутация A1401G в гене *rrs* (56–78 % штаммов). В нашем исследовании представленная мутация встречалась несколько реже и была выявлена у 52 % устойчивых штаммов. В свою очередь доля мутаций в промоторной области гена *eis* у нас была несколько выше (51.5 % образцов), чем в исследовании коллег (17–44 %). Пять образцов из секвенированной коллекции, чувствительных к представленным препаратам, также содержали мутации в промоторной области гена *eis*, что, однако, также согласуется с исследованием Georghiou с соавторами. Дополнительно следует отметить, что закономерностей в распределении мутаций по генотипам нами обнаружено не было.

Мутации в QRDR-регионе генов *gyrA* и *gyrB* были выявлены у 88 % штаммов, устойчивых к фторхинолонам. Это несколько больше, чем в целом по миру (Maruri *et al.*, 2012), однако в полной мере согласуется с данными по России (Mokrousov *et al.*, 2008). При этом следует отметить, что наиболее часто замены встречались в 94 кодоне гена *gyrA* (16 из 27 устойчивых штаммов), ассоциированном с высокой МИК к офлоксацину. Мутации в *gyrB* были найдены за пределами QRDR-региона и их роль в формировании устойчивости пока не определена. Дополнительно, как и в случае с этамбутолом и инъекционными препаратами, нам не удалось выявить различий в распределении описанных мутаций по генотипам.

В заключении следует отметить, что в ходе исследования был проведен анализ канонических маркеров устойчивости к наиболее часто используемым ПТП 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 ПТП удалось обнаружить известные маркеры резистентности в изучаемых образцах. Так в нашем исследовании для 7 из 54 образцов не удалось обнаружить маркеров устойчивости к тем или иным препаратам, что свидетельствует о реализации иных, не описанных ранее, механизмов устойчивости к ПТП.

4.8 Построение гипотез формирования устойчивости к противотуберкулезным препаратам

Развитие алгоритмов обработки данных, компьютерной техники и методов полногеномного секвенирования в последние годы привело к широкому распространению геномики. В области изучения *M. tuberculosis* данные технологии уже с успехом применяются для изучения макро- и микро-эволюционных событий в геноме, установления путей передачи патогена (Roetzer *et al.*, 2013), поиска молекулярных маркеров устойчивости к противотуберкулезным препаратам (Farhat *et al.*, 2013; Zhang *et al.*, 2013). Последняя задача может решаться методами сравнительной геномики с привлечением большого количества геномов. Как упоминалось в пункте 1.7.2 литературного обзора на сегодняшний день опубликовано два крупных исследования такого рода, включающих анализ 123 и 161 генома, соответственно. Для поиска генетических маркеров, ассоциированных с лекарственной устойчивостью, в нашем исследовании был проведен анализ геномных последовательностей 100 образцов МБГ (таблица 20). Отличительная особенность представленной работы заключалась в выборе эндемичных для России штаммов. Дополнительно задачей исследования было определение наиболее вероятного пути развития лекарственной устойчивости. По одной из существующих гипотез лекарственная

устойчивость может развиваться ступенчато, путем приобретения мутаций в генах, кодирующих ферменты активации лекарства или мишени действия препаратов. Предполагается, что такой механизм наиболее вероятен для МБТ. В свою очередь другой механизм проявляется скачкообразным нарастанием устойчивости одновременно к нескольким препаратам и связан с мутациями в одном или небольшом количестве генов. К примеру, это могут быть мутации, приводящие к уменьшению проницаемости клеточной стенки или к увеличению активности эффлюксных насосов (Farhat *et al.*, 2013).

Как говорилось ранее, наше предположение основывалось на том, что мутации, ассоциированные с лекарственной устойчивостью, должны встречаться чаще в устойчивых, чем в чувствительных к ПТП штаммах. На первом этапе работы были разработаны основные критерии отбора SNPs. Одним из них было наличие SNP менее чем у 90 % образцов, что может быть связано как с ошибками референтного генома H37Rv, так и с его полиморфизмами относительно изучаемых образцов. Другим критерием служило исключение полиморфизмов в генах, кодирующих семейство белков PE-PPE. Данное обстоятельство связано с большой схожестью представленных генов и высоким составом G+C пар в них, что затрудняет как секвенирование этих участков, так и их дальнейший биоинформатический анализ. Следующий критерий заключался в исключении филогенетических полиморфизмов. Данный критерий позволил отсеять кластер-специфические SNPs, которые могли бы дать ложноположительные результаты на следующем этапе анализа. В дальнейшем для 4135 отобранных полиморфизмов был применен тест на основе гипергеометрического распределения. Данный тест позволяет оценить вероятность ассоциации мутации с резистентностью к каждому тестируемому антибиотику. В ходе анализа было выявлено 116 мутаций, достоверно ассоциированных с лекарственной устойчивостью. Причем многие из найденных SNPs находились в хорошо изученных и описанных генах, ассоциированных с резистентностью. Следует отметить, что данный

тест не позволил выявить ассоциацию найденных SNPs с конкретным препаратом. Так, например, мутация в кодоне Ser315 гена *katG* была достоверно ассоциирована с устойчивостью ко всем анализируемым ПТП. Данное обстоятельство скорее связано со ступенчатым механизмом формирования устойчивости, происходящим параллельно с поэтапным использованием ПТП. Для проверки нашего предположения и оценки значимости для того или иного антибиотика отобранных после первичного анализа SNPs была использована логистическая регрессия. На основании схемы лечения туберкулеза и разницы в важности ПТП наиболее вероятная цепочка препаратов, и соответствующая ей последовательность развития лекарственной устойчивости, выглядела следующим образом: изониазид, рифампицин, стрептомицин, этамбутол, офлоксацин, канамицин и капреомицин. По результатам исследования 22 и 2 SNPs в генах и межгенных участках, соответственно, показали значимую ассоциацию с проявлением фенотипа резистентности. При этом большое количество полиморфизмов удалось соотнести с устойчивостью к определенным ПТП (таблица 21). Данное обстоятельство подтверждает гипотезу ступенчатого приобретения лекарственной устойчивости.

Согласно анализу полученных полиморфизмов 10 SNPs относились к уже изученным маркерам устойчивости к ПТП. 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). Следует отметить, что это была единственная общая замена, выявленная при попарном сравнении исследований. Пять из оставшихся 13, не описанных ранее полиморфизмов, находились в генах с гипотетической функцией, что затрудняет их анализ. Однако следует отметить, что замена Leu95Ser в белке Rv2079, относящегося к COG M (биосинтез клеточной стенки), встретилась только в устойчивых к ПТП образцах кластера Beijing B0/W148, что вероятно может говорить о значимости данной мутации. 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 (пункт 3.9) 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* встретилась 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.

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 соотнесением 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 спוליгопрофиля *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 строгую ассоциацию с лекарственной устойчивостью.

Список используемых сокращений

ВОЗ – Всемирная Организация Здравоохранения

ДНК – дезоксирибонуклеиновая кислота

дНТФ – дезоксирибонуклеозидтрифосфат

МБТ – *Mycobacterium tuberculosis*, микобактерия туберкулеза

МЛУ – множественная лекарственная устойчивость

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

ПТП – противотуберкулезный препарат

ПЦР – полимеразная цепная реакция

РФ – Российская Федерация

ФО – Федеральный округ

ШЛУ – широкая лекарственная устойчивость

АМК – амикацин

САР – капреомицин

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

CS - циклосерин

DR – Direct Repeat (англ. прямой повтор)

EMB – этамбутол

ETH – этионамид

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

FQ – фторхинолоны

INH – изониазид

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

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

KAN – канамицин

- LAM – Latin American Mediterranean (генетическое семейство вида *M. tuberculosis*)
- LSP – Large Sequence Polymorphism (англ. крупный нуклеотидный полиморфизм)
- MALDI-ToF MS – Matrix-Assisted Laser Desorption/Ionization Time Of Flight Mass-Spectrometry (англ. времяпролётная масс-спектрометрия с матричной лазерной десорбционной ионизацией)
- MIRU – Mycobacterial Interspersed Repeat Units (англ. микобактериальные рассеянные повторяющиеся единицы)
- OFX – офлоксацин
- QRDR – Quinolone Resistance-Determining Region (англ. участок, определяющий устойчивость к хинолонам)
- PAS – пара-аминосалицилат натрия
- PE/PPE – белки с характерными N-терминальными мотивами ProGlu (PE) / ProProGlu (PPE)
- RD – Region of Difference (англ. регион различия)
- RIF – рифампицин
- RRDR – Rifampicin Resistance Determining Region (англ. участок, определяющий устойчивость к рифампицину)
- PZA – пипразинамид
- SIT – Spoligotype International Type (англ. международный сполиготип)
- SNP – Single Nucleotide Polymorphism (англ. однонуклеотидный полиморфизм)
- SRA – sequence read archive (архив секвенированных последовательностей)
- STR – стрептомицин
- VNTR – Variable Number of Tandem Repeats (англ. переменное количество tandemных повторов)

Список литературы

1. Иванов, И. Ю., Степаншина, В. Н., Липин, М. Ю., Коробова, О. В., Шемякин И. Г. Спוליгогипы клинических штаммов *Mycobacterium tuberculosis*, выделенных у больных туберкулезом в Центральном регионе России // Проблемы туберкулеза и болезней легких. 2004. №4. С. 23-27.
2. Колычев, Н. М., Госманов, Р. Г., Ветеринарная микробиология и иммунология // М.: Колос, 2003. 432 с.
3. Матрашкин, А. Г., Месько, Е. М., Белякова, Н.К., Андреевская, С.Н., Смирнова, с соавт. Генотипическая характеристика штаммов *Mycobacterium tuberculosis* из Республики Тыва // Проблемы туберкулеза и болезней легких. 2004. №3. С. 37-40.
4. Медведева, Т. В., Огарков, О. Б., Некипелов, О. М., Ушаков, И. В., Козьякова, Е. С., Скворцова, Р. Г. MIRU-VNTR-генотипирование штаммов *Mycobacterium tuberculosis* в Восточной Сибири: семейство Beijing против Kilimanjaro // Молекулярная генетика микробиология и вирусология. 2004. №4. С. 33-38.
5. Министерство здравоохранения РФ. О совершенствовании противотуберкулезных мероприятий в Российской Федерации.// Приказ № 109 от 21 марта 2003 г.
6. Мишин, В. Ю. Лекарственно-устойчивый туберкулез легких: Учебное пособие для врачей // М.: ГОУ ВПО МГМСУ. 2005. 142 с.
7. Нарвская, О. В., Мокроусов, И. В., Лимещенко, Е. В., Стеклова, Л. Н., Оттен, Т. Ф., Вишневский, Б. И. Характеристика циркулирующих на Северо-Западе России штаммов *Mycobacterium tuberculosis* с использованием спוליгогипирования // Проблемы туберкулеза. 2002. №4. С. 44-48.
8. Нечаева, О.Б. Ситуация по туберкулезу в российской федерации // Отчет ФГБУ "ЦНИИОИЗ" Минздравсоцразвития России. Москва. 2013 г.
9. 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.
10. Смирнов, Г. Б. Повторы в геномах бактерий и жизни // Молекулярная генетика, микробиология и вирусология. 2010. №2. С. 10-20.
11. Туберкулез в Российской Федерации 2011 г. Аналитический обзор статистических показателей, используемых в Российской Федерации и в мире // М. 2013 г. 280 с.
12. Умпелева, Т. В., Кравченко, М. А., Еремеева, Н.И., Вязовая, А. А., Нарвская, О. В. Молекулярно-генетическая характеристика штаммов *Mycobacterium tuberculosis*, циркулирующих на территории Уральского региона России // Инфекция и иммунитет. 2013. Т. 3 № 1. С. 21–28
13. Черноусова, Л. Н., Андреевская, С. Н., Смирнова, Т. Г., Земскова, З. С., Ларионова, Е. Е. Биологические свойства штаммов *M. tuberculosis*

- кластера W // Проблемы туберкулеза и болезней легких. 2008. №10. С. 45-50.
14. Шемякин, И. Г., Степаншина, В. Н., Иванов, И. Ю., Липин, М. Ю., Коробова, О. В., Анисимова, В. А. Характеристика клинических изолятов *Mycobacterium tuberculosis* с использованием молекулярно-биологических методов // Молекулярная генетика микробиология и вирусология. 2003. №1. С. 32-40.
 15. Achtman, M. Evolution, population structure, and phylogeography of genetically monomorphic bacterial pathogens // Annu Rev Microbiol. 2008. Vol. 62. P.53-70.
 16. Afanas'ev, M. V., Ikryannikova, L. N., Il'ina, E. N., Kuz'min, A. V., Larionova, E. E., *et al.* Molecular typing of *Mycobacterium tuberculosis* circulated in Moscow, Russian Federation // Eur J Clin Microbiol Infect Dis. 2011. Vol. 30. № 2. P. 181-191.
 17. 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.
 18. Alexander, K. A., Laver, P. N., Michel, A. L., Williams, M., van Helden, P. D., *et al.* Novel *Mycobacterium tuberculosis* complex pathogen, *M. mungi* // Emerg Infect Dis. 2010. Vol. 16. № 8. P. 1296-1299.
 19. Alland, D., Whittam, T. S., Murray, M. B., Cave, M. D., Hazbon, M. H., *et al.* Modeling bacterial evolution with comparative-genome-based marker systems: application to *Mycobacterium tuberculosis* evolution and pathogenesis // J Bacteriol. 2003. Vol. 185. № 11. P. 3392-3399.
 20. 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.
 21. Andersson, D. I. The biological cost of mutational antibiotic resistance: any practical conclusions? // Curr Opin Microbiol. 2006. Vol. 9. № 5. P. 461-465.
 22. Andersson, D. I. and Hughes, D. Antibiotic resistance and its cost: is it possible to reverse resistance? // Nat Rev Microbiol. 2010. Vol. 8. № 4. P. 260-271.
 23. Andersson, D. I. and Levin, B. R. The biological cost of antibiotic resistance // Curr Opin Microbiol. 1999. Vol. 2. № 5. P. 489-493.
 24. Azhikina, T., Gvozdevsky, N., Botvinnik, A., Fushan, A., Shemyakin, I., *et al.* A genome-wide sequence-independent comparative analysis of insertion-deletion polymorphisms in multiple *Mycobacterium tuberculosis* strains // Res Microbiol. 2006. Vol. 157. № 3. P. 282-290.
 25. 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.
 26. Banerjee, A., Dubnau, E., Quemard, A., Balasubramanian, V., Um, K. S., *et al.* inhA, a gene encoding a target for isoniazid and ethionamide in *Mycobacterium tuberculosis* // Science. 1994. Vol. 263. № 5144. P. 227-230.

27. Baranov, A. A., Mariandyshev, A. O., Mannsaker, T., Dahle, U. R. and Bjune, G. A. Molecular epidemiology and drug resistance of widespread genotypes of *Mycobacterium tuberculosis* in northwestern Russia // Int J Tuberc Lung Dis. 2009. Vol. 13. № 10. P. 1288-1293.
28. Barczak, A. K., Domenech, P., Boshoff, H. I., Reed, M. B., Manca, C., *et al.* In vivo phenotypic dominance in mouse mixed infections with *Mycobacterium tuberculosis* clinical isolates // J Infect Dis. 2005. Vol. 192. № 4. P. 600-606.
29. Baulard, A. R., Betts, J. C., Engohang-Ndong, J., Quan, S., McAdam, R. A., *et al.* Activation of the pro-drug ethionamide is regulated in mycobacteria // J Biol Chem. 2000. Vol. 275. № 36. P. 28326-28331.
30. 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.
31. Bentley, S. D., Comas, I., Bryant, J. M., Walker, D., Smith, N. H., *et al.* The genome of *Mycobacterium africanum* West African 2 reveals a lineage-specific locus and genome erosion common to the *M. tuberculosis* complex // PLoS Negl Trop Dis. 2012. Vol. 6. № 2. P. e1552.
32. 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.
33. Bolotin, A., Quinquis, B., Sorokin, A. and Ehrlich, S. D. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin // Microbiology. 2005. Vol. 151. № 8. P. 2551-2561.
34. Brosch, R., Gordon, S. V., Buchrieser, C., Pym, A. S., Garnier, T., *et al.* Comparative genomics uncovers large tandem chromosomal duplications in *Mycobacterium bovis* BCG Pasteur // Yeast. 2000. Vol. 17. № 2. P. 111-123.
35. Brosch, R., Gordon, S. V., Marmiesse, M., Brodin, P., Buchrieser, C., *et al.* A new evolutionary scenario for the *Mycobacterium tuberculosis* complex // Proc Natl Acad Sci U S A. 2002. Vol. 99. № 6. P. 3684-3689.
36. 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.
37. 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.
38. Brudey, K., Gordon, M., Mostrom, P., Svensson, L., Jonsson, B., *et al.* Molecular epidemiology of *Mycobacterium tuberculosis* in western Sweden // J Clin Microbiol. 2004. Vol. 42. № 7. P. 3046-3051.
39. Bult, C. J., White, O., Olsen, G. J., Zhou, L., Fleischmann, R. D., *et al.* Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii* // Science. 1996. Vol. 273. № 5278. P. 1058-1073.

40. 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.
41. Casali, N., Nikolayevskyy, V., Balabanova, Y., Ignatyeva, O., Kontsevaya, I., *et al.* Microevolution of extensively drug-resistant tuberculosis in Russia // *Genome Res.* 2012. Vol. 22. № 4. P. 735-745.
42. Chambers, H. F., Moreau, D., Yajko, D., Miick, C., Wagner, C., *et al.* Can penicillins and other beta-lactam antibiotics be used to treat tuberculosis? // *Antimicrob Agents Chemother.* 1995. Vol. 39. № 12. P. 2620-2624.
43. Chouchane, S., Lippai, I. and Magliozzo, R. S. Catalase-peroxidase (*Mycobacterium tuberculosis* KatG) catalysis and isoniazid activation // *Biochemistry.* 2000. Vol. 39. № 32. P. 9975-9983.
44. Cole, S. T. Learning from the genome sequence of *Mycobacterium tuberculosis* H37Rv // *FEBS Lett.* 1999. Vol. 452. № 1-2. P. 7-10.
45. 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.
46. Comas, I., Borrell, S., Roetzer, A., Rose, G., Malla, B., *et al.* Whole-genome sequencing of rifampicin-resistant *Mycobacterium tuberculosis* strains identifies compensatory mutations in RNA polymerase genes // *Nat Genet.* 2011. Vol. 44. № 1. P. 106-110.
47. 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.
48. Comas, I., Homolka, S., Niemann, S. and Gagneux, S. Genotyping of genetically monomorphic bacteria: DNA sequencing in *Mycobacterium tuberculosis* highlights the limitations of current methodologies // *PLoS One.* 2009. Vol. 4. № 11. P. e7815.
49. Cooksey, R. C., Morlock, G. P., McQueen, A., Glickman, S. E. and Crawford, J. T. Characterization of streptomycin resistance mechanisms among *Mycobacterium tuberculosis* isolates from patients in New York City // *Antimicrob Agents Chemother.* 1996. Vol. 40. № 5. P. 1186-1188.
50. Cousins, D. V., Bastida, R., Cataldi, A., Quse, V., Redrobe, S., *et al.* Tuberculosis in seals caused by a novel member of the *Mycobacterium tuberculosis* complex: *Mycobacterium pinnipedii* sp. nov // *Int J Syst Evol Microbiol.* 2003. Vol. 53. № Pt 5. P. 1305-1314.
51. 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.
52. Crofton, J. and Mitchison, D. A. Streptomycin resistance in pulmonary tuberculosis // *Br Med J.* 1948. Vol. 2. № 4588. P. 1009-1015.
53. Dale, J. W. Mobile genetic elements in mycobacteria // *Eur Respir J Suppl.* 1995. Vol. 20. № 633s-648s.

54. Danilchanka, O., Pavlenok, M. and Niederweis, M. Role of porins for uptake of antibiotics by *Mycobacterium smegmatis* // *Antimicrob Agents Chemother.* 2008. Vol. 52. № 9. P. 3127-3134.
55. 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.
56. Das, S., Yennamalli, R. M., Vishnoi, A., Gupta, P. and Bhattacharya, A. Single-nucleotide variations associated with *Mycobacterium tuberculosis* KwaZulu-Natal strains // *J Biosci.* 2009. Vol. 34. № 3. P. 397-404.
57. DeBarber, A. E., Mdluli, K., Bosman, M., Bekker, L. G. and Barry, C. E., 3rd. Ethionamide activation and sensitivity in multidrug-resistant *Mycobacterium tuberculosis* // *Proc Natl Acad Sci U S A.* 2000. Vol. 97. № 17. P. 9677-9682.
58. 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.
59. 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.
60. Djelouadji, Z., Raoult, D. and Drancourt, M. Palaeogenomics of *Mycobacterium tuberculosis*: epidemic bursts with a degrading genome // *Lancet Infect Dis.* 2011. Vol. 11. № 8. P. 641-650.
61. 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.
62. 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.
63. Dos Vultos, T., Mestre, O., Rauzier, J., Golec, M., Rastogi, N., *et al.* Evolution and diversity of clonal bacteria: the paradigm of *Mycobacterium tuberculosis* // *PLoS One.* 2008. Vol. 3. № 2. P. e1538.
64. Drlica, K. and Malik, M. Fluoroquinolones: action and resistance // *Curr Top Med Chem.* 2003. Vol. 3. № 3. P. 249-282.
65. Drobniowski, 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.
66. Dymova, M. A., Kinsht, V. N., Cherednichenko, A. G., Khrapov, E. A., Svistelnik, A. V., *et al.* Highest prevalence of the *Mycobacterium tuberculosis* Beijing genotype isolates in patients newly diagnosed with tuberculosis in the Novosibirsk oblast, Russian Federation // *J Med Microbiol.* 2011. Vol. 60. № Pt 7. P. 1003-1009.

67. 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.
68. Eisen, J. A., Heidelberg, J. F., White, O. and Salzberg, S. L. Evidence for symmetric chromosomal inversions around the replication origin in bacteria // *Genome Biol.* 2000. Vol. 1. № 6.
69. Ernst, J. D., Trevejo-Nunez, G. and Banaiee, N. Genomics and the evolution, pathogenesis, and diagnosis of tuberculosis // *J Clin Invest.* 2007. Vol. 117. № 7. P. 1738-1745.
70. Fabre, M., Koeck, J. L., Le Fleche, P., Simon, F., Herve, V., *et al.* High genetic diversity revealed by variable-number tandem repeat genotyping and analysis of hsp65 gene polymorphism in a large collection of "Mycobacterium canettii" strains indicates that the *M. tuberculosis* complex is a recently emerged clone of "M. canettii" // *J Clin Microbiol.* 2004. Vol. 42. № 7. P. 3248-3255.
71. 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.
72. Fenner, L., Egger, M., Bodmer, T., Altpeter, E., Zwahlen, M., *et al.* Effect of mutation and genetic background on drug resistance in *Mycobacterium tuberculosis* // *Antimicrob Agents Chemother.* 2012. Vol. 56. № 6. P. 3047-3053.
73. Fenner, L., Malla, B., Ninet, B., Dubuis, O., Stucki, D., *et al.* "Pseudo-Beijing": evidence for convergent evolution in the direct repeat region of *Mycobacterium tuberculosis* // *PLoS One.* 2011. Vol. 6. № 9. P. e24737.
74. Ferdinand, S., Valetudie, G., Sola, C. and Rastogi, N. Data mining of *Mycobacterium tuberculosis* complex genotyping results using mycobacterial interspersed repetitive units validates the clonal structure of spoligotyping-defined families // *Res Microbiol.* 2004. Vol. 155. № 8. P. 647-654.
75. 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.
76. Finken, M., Kirschner, P., Meier, A., Wrede, A. and Bottger, E. C. Molecular basis of streptomycin resistance in *Mycobacterium tuberculosis*: alterations of the ribosomal protein S12 gene and point mutations within a functional 16S ribosomal RNA pseudoknot // *Mol Microbiol.* 1993. Vol. 9. № 6. P. 1239-1246.
77. Fleischmann, R. D., Adams, M. D., White, O., Clayton, R. A., Kirkness, E. F., *et al.* Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd // *Science.* 1995. Vol. 269. № 5223. P. 496-512.

78. 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.
79. 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.
80. Fraser, C. M., Gocayne, J. D., White, O., Adams, M. D., Clayton, R. A., *et al.* The minimal gene complement of *Mycoplasma genitalium* // Science. 1995. Vol. 270. № 5235. P. 397-403.
81. Frota, C. C., Hunt, D. M., Buxton, R. S., Rickman, L., Hinds, J., *et al.* Genome structure in the vole bacillus, *Mycobacterium microti*, a member of the *Mycobacterium tuberculosis* complex with a low virulence for humans // Microbiology. 2004. Vol. 150. № Pt 5. P. 1519-1527.
82. 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.
83. Gagneux, S., Burgos, M. V., DeRiemer, K., Encisco, A., Munoz, S., *et al.* Impact of bacterial genetics on the transmission of isoniazid-resistant *Mycobacterium tuberculosis* // PLoS Pathog. 2006. Vol. 2. № 6. P. e61.
84. 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.
85. 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.
86. Garcia-Betancur, J. C., Menendez, M. C., Del Portillo, P. and Garcia, M. J. Alignment of multiple complete genomes suggests that gene rearrangements may contribute towards the speciation of Mycobacteria // Infect Genet Evol. 2012. Vol. 12. № 4. P. 819-826.
87. Garcia de Viedma, D., Chaves, F. and Inigo, J. New route of importation of *Mycobacterium tuberculosis* Beijing genotype // Emerg Infect Dis. 2006. Vol. 12. № 1. P. 169-170.
88. Gardy, J. L., Johnston, J. C., Ho Sui, S. J., Cook, V. J., Shah, L., *et al.* Whole-genome sequencing and social-network analysis of a tuberculosis outbreak // N Engl J Med. 2011. Vol. 364. № 8. P. 730-739.
89. Garnier, T., Eiglmeier, K., Camus, J. C., Medina, N., Mansoor, H., *et al.* The complete genome sequence of *Mycobacterium bovis* // Proc Natl Acad Sci U S A. 2003. Vol. 100. № 13. P. 7877-7882.
90. Garvin, R. T., Biswas, D. K. and Gorini, L. The effects of streptomycin or dihydrostreptomycin binding to 16S RNA or to 30S ribosomal subunits // Proc Natl Acad Sci U S A. 1974. Vol. 71. № 10. P. 3814-3818.
91. Georghiou, S. B., Magana, M., Garfein, R. S., Catanzaro, D. G., Catanzaro, A., *et al.* Evaluation of genetic mutations associated with *Mycobacterium*

- tuberculosis* resistance to amikacin, kanamycin and capreomycin: a systematic review // PLoS One. 2012. Vol. 7. № 3. P. e33275.
92. Gillespie, S. H. Evolution of drug resistance in *Mycobacterium tuberculosis*: clinical and molecular perspective // Antimicrob Agents Chemother. 2002. Vol. 46. № 2. P. 267-274.
 93. Glynn, J. R., Crampin, A. C., Traore, H., Yates, M. D., Mwaungulu, F. D., *et al.* *Mycobacterium tuberculosis* Beijing genotype, northern Malawi // Emerg Infect Dis. 2005. Vol. 11. № 1. P. 150-153.
 94. Gori, A., Bandera, A., Marchetti, G., Degli Esposti, A., Catozzi, L., *et al.* Spoligotyping and *Mycobacterium tuberculosis* // Emerg Infect Dis. 2005. Vol. 11. № 8. P. 1242-1248.
 95. Guillemain, I., Jarlier, V. and Cambau, E. Correlation between quinolone susceptibility patterns and sequences in the A and B subunits of DNA gyrase in mycobacteria // Antimicrob Agents Chemother. 1998. Vol. 42. № 8. P. 2084-2088.
 96. Gupta, A. K., Katoch, V. M., Chauhan, D. S., Sharma, R., Singh, M., *et al.* Microarray analysis of efflux pump genes in multidrug-resistant *Mycobacterium tuberculosis* during stress induced by common anti-tuberculous drugs // Microb Drug Resist. 2010. Vol. 16. № 1. P. 21-28.
 97. 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.
 98. Gutacker, M. M., Smoot, J. C., Migliaccio, C. A., Ricklefs, S. M., Hua, S., *et al.* Genome-wide analysis of synonymous single nucleotide polymorphisms in *Mycobacterium tuberculosis* complex organisms: resolution of genetic relationships among closely related microbial strains // Genetics. 2002. Vol. 162. № 4. P. 1533-1543.
 99. 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.
 100. 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.
 101. Hanekom, M., van der Spuy, G. D., Streicher, E., Ndabambi, S. L., McEvoy, C. R., *et al.* A recently evolved sublineage of the *Mycobacterium tuberculosis* Beijing strain family is associated with an increased ability to spread and cause disease // J Clin Microbiol. 2007. Vol. 45. № 5. P. 1483-1490.
 102. Hastings, I. M., Watkins, W. M. and White, N. J. The evolution of drug-resistant malaria: the role of drug elimination half-life // Philos Trans R Soc Lond B Biol Sci. 2002. Vol. 357. № 1420. P. 505-519.
 103. Hermans, P. W., van Soolingen, D., Bik, E. M., de Haas, P. E., Dale, J. W., *et al.* Insertion element IS987 from *Mycobacterium bovis* BCG is located in a hot-spot integration region for insertion elements in *Mycobacterium*

- tuberculosis* complex strains // Infect Immun. 1991. Vol. 59. № 8. P. 2695-2705.
104. Hershberg, R., Lipatov, M., Small, P. M., Sheffer, H., Niemann, S., *et al.* High functional diversity in *Mycobacterium tuberculosis* driven by genetic drift and human demography // PLoS Biol. 2008. Vol. 6. № 12. P. e311.
 105. Hettick, J. M., Kashon, M. L., Slaven, J. E., Ma, Y., Simpson, J. P., *et al.* Discrimination of intact mycobacteria at the strain level: a combined MALDI-TOF MS and biostatistical analysis // Proteomics. 2006. Vol. 6. № 24. P. 6416-6425.
 106. Hillemann, D., Kubica, T., Rusch-Gerdes, S. and Niemann, S. Disequilibrium in distribution of resistance mutations among *Mycobacterium tuberculosis* Beijing and non-Beijing strains isolated from patients in Germany // Antimicrob Agents Chemother. 2005. Vol. 49. № 3. P. 1229-1231.
 107. 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.
 108. 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.
 109. 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.
 110. 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.
 111. Hughes, A. L., Friedman, R. and Murray, M. Genomewide pattern of synonymous nucleotide substitution in two complete genomes of *Mycobacterium tuberculosis* // Emerg Infect Dis. 2002. Vol. 8. № 11. P. 1342-1346.
 112. Hunter, P. R., Gaston, M. A. Numerical index of the discriminatory ability of typing systems: an application of Simpson's index of diversity // J. Clin. Microbiol. 1988. Vol. 26. № 11. P. 2465-2466.
 113. Ignatova, A., Dubiley, S., Stepanshina, V. and Shemyakin, I. Predominance of multi-drug-resistant LAM and Beijing family strains among *Mycobacterium tuberculosis* isolates recovered from prison inmates in Tula Region, Russia // J Med Microbiol. 2006. Vol. 55. № Pt 10. P. 1413-1418.
 114. Ikryannikova, L. N., Afanas'ev, M. V., Akopian, T. A., Il'ina, E. N., Kuz'min, A. V., *et al.* Mass-spectrometry based minisequencing method for the rapid detection of drug resistance in *Mycobacterium tuberculosis* // J Microbiol Methods. 2007. Vol. 70. № 3. P. 395-405.

115. 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.
116. 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.
117. Ishino, Y., Shinagawa, H., Makino, K., Amemura, M. and Nakata, A. Nucleotide sequence of the *iap* gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product // J Bacteriol. 1987. Vol. 169. № 12. P. 5429-5433.
118. Isola, D., Pardini, M., Varaine, F., Niemann, S., Rusch-Gerdes, S., *et al.* A Pyrosequencing assay for rapid recognition of SNPs in *Mycobacterium tuberculosis* embB306 region // J Microbiol Methods. 2005. Vol. 62. № 1. P. 113-120.
119. Iwamoto, T., Yoshida, S., Suzuki, K. and Wada, T. Population structure analysis of the *Mycobacterium tuberculosis* Beijing family indicates an association between certain sublineages and multidrug resistance // Antimicrob Agents Chemother. 2008. Vol. 52. № 10. P. 3805-3809.
120. Jansen, R., Embden, J. D., Gaastra, W. and Schouls, L. M. Identification of genes that are associated with DNA repeats in prokaryotes // Mol Microbiol. 2002. Vol. 43. № 6. P. 1565-1575.
121. Jindani, A., Aber, V. R., Edwards, E. A. and Mitchison, D. A. The early bactericidal activity of drugs in patients with pulmonary tuberculosis // Am Rev Respir Dis. 1980. Vol. 121. № 6. P. 939-949.
122. 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.
123. Kaur, D., Guerin, M. E., Skovierova, H., Brennan, P. J. and Jackson, M. Chapter 2: Biogenesis of the cell wall and other glycoconjugates of *Mycobacterium tuberculosis* // Adv Appl Microbiol. 2009. Vol. 69. P. 23-78.
124. Kiers, A., Klarenbeek, A., Mendelts, B., Van Soolingen, D. and Koeter, G. Transmission of *Mycobacterium pinnipedii* to humans in a zoo with marine mammals // Int J Tuberc Lung Dis. 2008. Vol. 12. № 12. P. 1469-1473.
125. 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.
126. Koch, R. Classics in infectious diseases. The etiology of tuberculosis: Robert Koch. Berlin, Germany 1882 // Rev Infect Dis. 1982. Vol. 4. № 6. P. 1270-1274.
127. Koser, C. U., Niemann, S., Summers, D. K. and Archer, J. A. Overview of errors in the reference sequence and annotation of *Mycobacterium tuberculosis*

- H37Rv, and variation amongst its isolates // *Infect Genet Evol.* 2012. Vol. 12. № 4. P. 807-810.
128. Kovalev, S. Y., Kamaev, E. Y., Kravchenko, M. A., Kurepina, N. E. and Skorniakov, S. N. Genetic analysis of *Mycobacterium tuberculosis* strains isolated in Ural region, Russian Federation, by MIRU-VNTR genotyping // *Int J Tuberc Lung Dis.* 2005. Vol. 9. № 7. P. 746-752.
 129. 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.
 130. Kubica, T., Rusch-Gerdes, S. and Niemann, S. *Mycobacterium bovis* subsp. *caprae* caused one-third of human *M. bovis*-associated tuberculosis cases reported in Germany between 1999 and 2001 // *J Clin Microbiol.* 2003. Vol. 41. № 7. P. 3070-3077.
 131. Kunin, V., Sorek, R. and Hugenholtz, P. Evolutionary conservation of sequence and secondary structures in CRISPR repeats // *Genome Biol.* 2007. Vol. 8. № 4. P. R61.
 132. 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.
 133. Kwara, A., Schiro, R., Cowan, L. S., Hyslop, N. E., Wiser, M. F., *et al.* Evaluation of the epidemiologic utility of secondary typing methods for differentiation of *Mycobacterium tuberculosis* isolates // *J Clin Microbiol.* 2003. Vol. 41. № 6. P. 2683-2685.
 134. Langmead, B. and Salzberg, S. L. Fast gapped-read alignment with Bowtie 2 // *Nat Methods.* 2012. Vol. 9. № 4. P. 357-359.
 135. 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.
 136. 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.
 137. 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.
 138. 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.

139. Ma, Z., Lienhardt, C., McIlleron, H., Nunn, A. J. and Wang, X. Global tuberculosis drug development pipeline: the need and the reality // *Lancet*. 2010. Vol. 375. № 9731. P. 2100-2109.
140. Makinen, J., Marjamaki, M., Haanpera-Heikkinen, M., Marttila, H., Endourova, L. B., *et al.* Extremely high prevalence of multidrug resistant tuberculosis in Murmansk, Russia: a population-based study // *Eur J Clin Microbiol Infect Dis*. 2011. Vol. 30. № 9. P. 1119-1126.
141. Maruri, F., Sterling, T. R., Kaiga, A. W., Blackman, A., van der Heijden, Y. F., *et al.* A systematic review of gyrase mutations associated with fluoroquinolone-resistant *Mycobacterium tuberculosis* and a proposed gyrase numbering system // *J Antimicrob Chemother*. 2012. Vol. 67. № 4. P. 819-831.
142. 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.
143. McHugh, T. D. and Gillespie, S. H. Nonrandom association of IS6110 and *Mycobacterium tuberculosis*: implications for molecular epidemiological studies // *J Clin Microbiol*. 1998. Vol. 36. № 5. P. 1410-1413.
144. Mestre, O., Luo, T., Dos Vultos, T., Kremer, K., Murray, A., *et al.* Phylogeny of *Mycobacterium tuberculosis* Beijing strains constructed from polymorphisms in genes involved in DNA replication, recombination and repair // *PLoS One*. 2011. Vol. 6. № 1. P. e16020.
145. Migliori, G. B., Loddenkemper, R., Blasi, F. and Raviglione, M. C. 125 years after Robert Koch's discovery of the tubercle bacillus: the new XDR-TB threat. Is "science" enough to tackle the epidemic? // *Eur Respir J*. 2007. Vol. 29. № 3. P. 423-427.
146. Mikusova, K., Slayden, R. A., Besra, G. S. and Brennan, P. J. Biogenesis of the mycobacterial cell wall and the site of action of ethambutol // *Antimicrob Agents Chemother*. 1995. Vol. 39. № 11. P. 2484-2489.
147. Millet, J., Miyagi-Shiohira, C., Yamane, N., Sola, C. and Rastogi, N. Assessment of mycobacterial interspersed repetitive unit-QUB markers to further discriminate the Beijing genotype in a population-based study of the genetic diversity of *Mycobacterium tuberculosis* clinical isolates from Okinawa, Ryukyu Islands, Japan // *J Clin Microbiol*. 2007. Vol. 45. № 11. P. 3606-3615.
148. Mizrahi, V. and Andersen, S. J. DNA repair in *Mycobacterium tuberculosis*. What have we learnt from the genome sequence? // *Mol Microbiol*. 1998. Vol. 29. № 6. P. 1331-1339.
149. Mojica, F. J., Diez-Villasenor, C., Garcia-Martinez, J. and Soria, E. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements // *J Mol Evol*. 2005. Vol. 60. № 2. P. 174-182.
150. Mojica, F. J., Ferrer, C., Juez, G. and Rodriguez-Valera, F. Long stretches of short tandem repeats are present in the largest replicons of the Archaea *Haloferax mediterranei* and *Haloferax volcanii* and could be involved in replicon partitioning // *Mol Microbiol*. 1995. Vol. 17. № 1. P. 85-93.

151. 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.
152. Mokrousov, I. The quiet and controversial: Ural family of *Mycobacterium tuberculosis* // Infect Genet Evol. 2012. Vol. 12. № 4. P. 619-629.
153. 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.
154. Mokrousov, I., Jiao, W. W., Sun, G. Z., Liu, J. W., Valcheva, V., et al. Evolution of drug resistance in different sublineages of *Mycobacterium tuberculosis* Beijing genotype // Antimicrob Agents Chemother. 2006. Vol. 50. № 8. P. 2820-2823.
155. 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.
156. Mokrousov, I., Narvskaya, O., Otten, T., Limeschenko, E., Steklova, L., et al. High prevalence of KatG Ser315Thr substitution among isoniazid-resistant *Mycobacterium tuberculosis* clinical isolates from northwestern Russia, 1996 to 2001 // Antimicrob Agents Chemother. 2002. Vol. 46. № 5. P. 1417-1424.
157. 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.
158. 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.
159. 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.
160. Mokrousov, I., Otten, T., Manicheva, O., Potapova, Y., Vishnevsky, B., et al. Molecular characterization of ofloxacin-resistant *Mycobacterium tuberculosis* strains from Russia // Antimicrob Agents Chemother. 2008. Vol. 52. № 8. P. 2937-2939.
161. 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.
162. Mokrousov, I., Otten, T., Vyshnevskiy, B. and Narvskaya, O. Detection of embB306 mutations in ethambutol-susceptible clinical isolates of *Mycobacterium tuberculosis* from Northwestern Russia: implications for

- genotypic resistance testing // *J Clin Microbiol.* 2002. Vol. 40. № 10. P. 3810-3813.
163. Mokrousov, I., Otten, T., Zozio, T., Turkin, E., Nazemtseva, V., *et al.* At Baltic crossroads: a molecular snapshot of *Mycobacterium tuberculosis* population diversity in Kaliningrad, Russia // *FEMS Immunol Med Microbiol.* 2009. Vol. 55. № 1. P. 13-22.
164. Mokrousov, I., Vyazovaya, A., Otten, T., Zhuravlev, V., Pavlova, E., *et al.* *Mycobacterium tuberculosis* population in northwestern Russia: an update from Russian-EU/Latvian border region // *PLoS One.* 2012. Vol. 7. № 7. P. e41318.
165. Monot, M., Honore, N., Garnier, T., Araoz, R., Coppee, J. Y., *et al.* On the origin of leprosy // *Science.* 2005. Vol. 308. № 5724. P. 1040-1042.
166. Morlock, G. P., Metchock, B., Sikes, D., Crawford, J. T. and Cooksey, R. C. *ethA*, *inhA*, and *katG* loci of ethionamide-resistant clinical *Mycobacterium tuberculosis* isolates // *Antimicrob Agents Chemother.* 2003. Vol. 47. № 12. P. 3799-3805.
167. Mostowy, S., Cousins, D. and Behr, M. A. Genomic interrogation of the dassie bacillus reveals it as a unique RD1 mutant within the *Mycobacterium tuberculosis* complex // *J Bacteriol.* 2004. Vol. 186. № 1. P. 104-109.
168. Mostowy, S., Onipede, A., Gagneux, S., Niemann, S., Kremer, K., *et al.* Genomic analysis distinguishes *Mycobacterium africanum* // *J Clin Microbiol.* 2004. Vol. 42. № 8. P. 3594-3599.
169. Mukhopadhyay, S. and Balaji, K. N. The PE and PPE proteins of *Mycobacterium tuberculosis* // *Tuberculosis (Edinb).* 2011. Vol. 91. № 5. P. 441-447.
170. Muttucumar, D. G. and Parish, T. The molecular biology of recombination in *Mycobacteria*: what do we know and how can we use it? // *Curr Issues Mol Biol.* 2004. Vol. 6. № 2. P. 145-157.
171. Niederweis, M., Danilchanka, O., Huff, J., Hoffmann, C. and Engelhardt, H. *Mycobacterial* outer membranes: in search of proteins // *Trends Microbiol.* 2010. Vol. 18. № 3. P. 109-116.
172. 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.
173. 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.
174. Niemann, S., Richter, E. and Rusch-Gerdes, S. Biochemical and genetic evidence for the transfer of *Mycobacterium tuberculosis* subsp. *caprae* Aranaz *et al.* 1999 to the species *Mycobacterium bovis* Karlson and Lessel 1970 (approved lists 1980) as *Mycobacterium bovis* subsp. *caprae* comb. nov // *Int J Syst Evol Microbiol.* 2002. Vol. 52. № Pt 2. P. 433-436.

175. Nikolayevskyy, V., Gopaul, K., Balabanova, Y., Brown, T., Fedorin, I., *et al.* Differentiation of tuberculosis strains in a population with mainly Beijing-family strains // *Emerg Infect Dis.* 2006. Vol. 12. № 9. P. 1406-1413.
176. Ocepek, M., Pate, M., Zolnir-Dovc, M. and Poljak, M. Transmission of *Mycobacterium tuberculosis* from human to cattle // *J Clin Microbiol.* 2005. Vol. 43. № 7. P. 3555-3557.
177. 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.
178. Ogarkov, O., Mokrousov, I., Sinkov, V., Zhdanova, S., Antipina, S., *et al.* 'Lethal' combination of *Mycobacterium tuberculosis* Beijing genotype and human CD209 -336G allele in Russian male population // *Infect Genet Evol.* 2012. Vol. 12. № 4. P. 732-736.
179. Panteix, G., Gutierrez, M. C., Boschioli, M. L., Rouviere, M., Plaidy, A., *et al.* Pulmonary tuberculosis due to *Mycobacterium microti*: a study of six recent cases in France // *J Med Microbiol.* 2010. Vol. 59. № Pt 8. P. 984-989.
180. 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.
181. Parwati, I., van Crevel, R. and van Soolingen, D. Possible underlying mechanisms for successful emergence of the *Mycobacterium tuberculosis* Beijing genotype strains // *Lancet Infect Dis.* 2010. Vol. 10. № 2. P. 103-111.
182. Pearson, T., Busch, J. D., Ravel, J., Read, T. D., Rhoton, S. D., *et al.* Phylogenetic discovery bias in *Bacillus anthracis* using single-nucleotide polymorphisms from whole-genome sequencing // *Proc Natl Acad Sci U S A.* 2004. Vol. 101. № 37. P. 13536-13541.
183. Pignone, M., Greth, K. M., Cooper, J., Emerson, D. and Tang, J. Identification of mycobacteria by matrix-assisted laser desorption ionization-time-of-flight mass spectrometry // *J Clin Microbiol.* 2006. Vol. 44. № 6. P. 1963-1970.
184. Pourcel, C., Salvignol, G. and Vergnaud, G. CRISPR elements in *Yersinia pestis* acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies // *Microbiology.* 2005. Vol. 151. № Pt 3. P. 653-663.
185. Qian, L., Abe, C., Lin, T. P., Yu, M. C., Cho, S. N., *et al.* *rpoB* genotypes of *Mycobacterium tuberculosis* Beijing family isolates from East Asian countries // *J Clin Microbiol.* 2002. Vol. 40. № 3. P. 1091-1094.
186. Ramaswamy, S. and Musser, J. M. Molecular genetic basis of antimicrobial agent resistance in *Mycobacterium tuberculosis*: 1998 update // *Tuber Lung Dis.* 1998. Vol. 79. № 1. P. 3-29.

187. Reed, M. B., Gagneux, S., Deriemer, K., Small, P. M. and Barry, C. E., 3rd. The W-Beijing lineage of *Mycobacterium tuberculosis* overproduces triglycerides and has the DosR dormancy regulon constitutively upregulated // J Bacteriol. 2007. Vol. 189. № 7. P. 2583-2589.
188. 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.
189. 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.
190. Sanger, F., Nicklen, S. and Coulson, A. R. DNA sequencing with chain-terminating inhibitors. 1977 // Biotechnology. 1992. Vol. 24. № 104-108.
191. Sauer, S. Typing of single nucleotide polymorphisms by MALDI mass spectrometry: principles and diagnostic applications // Clin Chim Acta. 2006. Vol. 363. № 1-2. P. 95-105.
192. Saunders, N. J., Trivedi, U. H., Thomson, M. L., Doig, C., Laurenson, I. F., *et al.* Deep resequencing of serial sputum isolates of *Mycobacterium tuberculosis* during therapeutic failure due to poor compliance reveals stepwise mutation of key resistance genes on an otherwise stable genetic background // J Infect. 2011. Vol. 62. № 3. P. 212-217.
193. Schatz, A., Bugie, E. and Waksman, S. A. Streptomycin, a substance exhibiting antibiotic activity against gram-positive and gram-negative bacteria. 1944 // Clin Orthop Relat Res. 2005. Vol. № 437. P. 3-6.
194. 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.
195. Schurch, A. C., Kremer, K., Warren, R. M., Hung, N. V., Zhao, Y., *et al.* Mutations in the regulatory network underlie the recent clonal expansion of a dominant subclone of the *Mycobacterium tuberculosis* Beijing genotype // Infect Genet Evol. 2011. Vol. 11. № 3. P. 587-597.
196. Schurch, A. C. and van Soolingen, D. DNA fingerprinting of *Mycobacterium tuberculosis*: from phage typing to whole-genome sequencing // Infect Genet Evol. 2012. Vol. 12. № 4. P. 602-609.
197. Scorpio, A. and Zhang, Y. Mutations in *pncA*, a gene encoding pyrazinamidase/nicotinamidase, cause resistance to the antituberculous drug pyrazinamide in tubercle bacillus // Nat Med. 1996. Vol. 2. № 6. P. 662-667.
198. Semaw, S., Simpson, S. W., Quade, J., Renne, P. R., Butler, R. F., *et al.* Early Pliocene hominids from Gona, Ethiopia // Nature. 2005. Vol. 433. № 7023. P. 301-305.

199. Shcherbakov, D., Akbergenov, R., Matt, T., Sander, P., Andersson, D. I., *et al.* Directed mutagenesis of *Mycobacterium smegmatis* 16S rRNA to reconstruct the in-vivo evolution of aminoglycoside resistance in *Mycobacterium tuberculosis* // *Mol Microbiol.* 2010. Vol. 1. № 1
200. Sherman, D. R., Mdluli, K., Hickey, M. J., Arain, T. M., Morris, S. L., *et al.* Compensatory *ahpC* gene expression in isoniazid-resistant *Mycobacterium tuberculosis* // *Science.* 1996. Vol. 272. № 5268. P. 1641-1643.
201. 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.
202. 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.
203. Skuce, R. A., McCorry, T. P., McCarroll, J. F., Roring, S. M., Scott, A. N., *et al.* Discrimination of *Mycobacterium tuberculosis* complex bacteria using novel VNTR-PCR targets // *Microbiology.* 2002. Vol. 148. № Pt 2. P. 519-528.
204. Smith, N. H., Dale, J., Inwald, J., Palmer, S., Gordon, S. V., *et al.* The population structure of *Mycobacterium bovis* in Great Britain: clonal expansion // *Proc Natl Acad Sci U S A.* 2003. Vol. 100. № 25. P. 15271-15275.
205. Smith, N. H., Hewinson, R. G., Kremer, K., Brosch, R. and Gordon, S. V. Myths and misconceptions: the origin and evolution of *Mycobacterium tuberculosis* // *Nat Rev Microbiol.* 2009. Vol. 7. № 7. P. 537-544.
206. Smith, N. H., Kremer, K., Inwald, J., Dale, J., Driscoll, J. R., *et al.* Ecotypes of the *Mycobacterium tuberculosis* complex // *J Theor Biol.* 2006. Vol. 239. № 2. P. 220-225.
207. 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.
208. Sreevatsan, S., Stockbauer, K. E., Pan, X., Kreiswirth, B. N., Moghazeh, S. L., *et al.* Ethambutol resistance in *Mycobacterium tuberculosis*: critical role of *embB* mutations // *Antimicrob Agents Chemother.* 1997. Vol. 41. № 8. P. 1677-1681.
209. Stead, W. W. The origin and erratic global spread of tuberculosis. How the past explains the present and is the key to the future // *Clin Chest Med.* 1997. Vol. 18. № 1. P. 65-77.
210. 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.
211. Supply, P., Lesjean, S., Savine, E., Kremer, K., van Soolingen, D., *et al.* Automated high-throughput genotyping for study of global epidemiology of

- Mycobacterium tuberculosis* based on mycobacterial interspersed repetitive units // J Clin Microbiol. 2001. Vol. 39. № 10. P. 3563-3571.
212. 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.
213. 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.
214. Supply, P., Warren, R. M., Banuls, A. L., Lesjean, S., Van Der Spuy, G. D., *et al.* Linkage disequilibrium between minisatellite loci supports clonal evolution of *Mycobacterium tuberculosis* in a high tuberculosis incidence area // Mol Microbiol. 2003. Vol. 47. № 2. P. 529-538.
215. Surikova, O. V., Voitech, D. S., Kuzmicheva, G., Tatkov, S. I., Mokrousov, I. V., *et al.* Efficient differentiation of *Mycobacterium tuberculosis* strains of the W-Beijing family from Russia using highly polymorphic VNTR loci // Eur J Epidemiol. 2005. Vol. 20. № 11. P. 963-974.
216. Telenti, A., Imboden, P., Marchesi, F., Lowrie, D., Cole, S., *et al.* Detection of rifampicin-resistance mutations in *Mycobacterium tuberculosis* // Lancet. 1993. Vol. 341. № 8846. P. 647-650.
217. Telenti, A., Philipp, W. J., Sreevatsan, S., Bernasconi, C., Stockbauer, K. E., *et al.* The emb operon, a gene cluster of *Mycobacterium tuberculosis* involved in resistance to ethambutol // Nat Med. 1997. Vol. 3. № 5. P. 567-570.
218. Thierry, D., Cave, M. D., Eisenach, K. D., Crawford, J. T., Bates, J. H., *et al.* IS6110, an IS-like element of *Mycobacterium tuberculosis* complex // Nucleic Acids Res. 1990. Vol. 18. № 1. P. 188.
219. Tounghousova, 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.
220. Tracevska, T., Jansone, I., Nodieva, A., Marga, O., Skenders, G., *et al.* Characterisation of rpsL, rrs and embB mutations associated with streptomycin and ethambutol resistance in *Mycobacterium tuberculosis* // Res Microbiol. 2004. Vol. 155. № 10. P. 830-834.
221. 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.
222. 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.
223. Umpeleva, T. V., Kravchenko, M. A., Ereemeeva, N. I., Kamaev, E. Y., 2010. Genotyping of *M. tuberculosis* isolates recovered from residents of

- Yekaterinburg city. // In: Proceedings of VII All-Russian conference 'Molecular Diagnostics– 2010. Vol. 1 (<http://www.md2010.org/works.php>)'. pp. 188-190. In Russian.
224. Vadrevu, I. S., Lofton, H., Sarva, K., Blasczyk, E., Plocinska, R., *et al.* ChiZ levels modulate cell division process in mycobacteria // *Tuberculosis (Edinb)*. 2011. Vol. 91. № Suppl 1. P. S128-135.
225. 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.
226. van Ingen, J., Rahim, Z., Mulder, A., Boeree, M. J., Simeone, R., *et al.* Characterization of *Mycobacterium orygis* as *M. tuberculosis* complex subspecies // *Emerg Infect Dis*. 2012. Vol. 18. № 4. P. 653-655.
227. van Soolingen, D., de Haas, P. E., van Doorn, H. R., Kuijper, E., Rinder, H., *et al.* Mutations at amino acid position 315 of the katG gene are associated with high-level resistance to isoniazid, other drug resistance, and successful transmission of *Mycobacterium tuberculosis* in the Netherlands // *J Infect Dis*. 2000. Vol. 182. № 6. P. 1788-1790.
228. van Soolingen, D., Hoogenboezem, T., de Haas, P. E., Hermans, P. W., Koedam, M. A., *et al.* A novel pathogenic taxon of the *Mycobacterium tuberculosis* complex, Canetti: characterization of an exceptional isolate from Africa // *Int J Syst Bacteriol*. 1997. Vol. 47. № 4. P. 1236-1245.
229. 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.
230. Vasconcellos, S. E., Huard, R. C., Niemann, S., Kremer, K., Santos, A. R., *et al.* Distinct genotypic profiles of the two major clades of *Mycobacterium africanum* // *BMC Infect Dis*. 2010. Vol. 10. № 80.
231. Vervenne, R. A., Jones, S. L., van Soolingen, D., van der Laan, T., Andersen, P., *et al.* TB diagnosis in non-human primates: comparison of two interferon-gamma assays and the skin test for identification of *Mycobacterium tuberculosis* infection // *Vet Immunol Immunopathol*. 2004. Vol. 100. № 1-2. P. 61-71.
232. Wang, X. M., Galamba, A., Warner, D. F., Soetaert, K., Merkel, J. S., *et al.* IS1096-mediated DNA rearrangements play a key role in genome evolution of *Mycobacterium smegmatis* // *Tuberculosis (Edinb)*. 2008. Vol. 88. № 5. P. 399-409.
233. Warner, D. F. and Mizrahi, V. Complex genetics of drug resistance in *Mycobacterium tuberculosis* // *Nat Genet*. 2013. Vol. 45. № 10. P. 1107-1108.
234. 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.

235. WHO. Extensively drug-resistant tuberculosis (XDR-TB): recommendations for prevention and control // *Wkly Epidemiol Rec.* 2006. Vol. 81. № 45. P. 430-432.
236. WHO. Extensively drug-resistant tuberculosis (XDR-TB): recommendations for prevention and control // *Wkly Epidemiol Rec.* 2012. Vol. 81. № 45. P. 430-432.
237. 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.
238. 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.
239. Zhang, Y. The magic bullets and tuberculosis drug targets // *Annu Rev Pharmacol Toxicol.* 2005. Vol. 45. P. 529-564.
240. Zhang, Y., Wade, M. M., Scorpio, A., Zhang, H. and Sun, Z. Mode of action of pyrazinamide: disruption of *Mycobacterium tuberculosis* membrane transport and energetics by pyrazinoic acid // *J Antimicrob Chemother.* 2003. Vol. 52. № 5. P. 790-795.
241. Zimenkov, D. V., Antonova, O. V., Kuz'min, A. V., Isaeva, Y. D., Krylova, L. Y., *et al.* Detection of second-line drug resistance in *Mycobacterium tuberculosis* using oligonucleotide microarrays // *BMC Infect Dis.* 2013. Vol. 13. № 240.
242. zur Wiesch, P. A., Kouyos, R., Engelstadter, J., Regoes, R. R. and Bonhoeffer, S. Population biological principles of drug-resistance evolution in infectious diseases // *Lancet Infect Dis.* 2011. Vol. 11. №3. P. 236-247.

ПРИЛОЖЕНИЯ

Приложение 1. Молекулярно-генетическая и фенотипическая характеристика 150 образцов коллекции. Серым цветом отмечены образцы для полногеномного секвенирования.

№ п/п	№ образца	№ культуры	I N H	R I F	E M B	S T R	O F L	A M K	C A P	K A N	Характеристика	Spol DB4.0	SITVIT_WEB
1	MOS1	35876	s	s	s	s	s	s	s	s	чув	H4(Ural)	262
2	MOS2	42165	r	r	r	r	r	r	s	r	ШЛУ	LAM9	42
3	MOS3	41712	r	r	s	r	r	r	s	r	ШЛУ	Beijing	1
4	MOS4	39915	r	r	r	r	s	s	s	s	МЛУ	Beijing	1
5	MOS5	32322	s	s	s	r	s	s	s	s	моно	Beijing	1
6	MOS6	44694	r	r	s	r	r	s	s	s	МЛУ	Beijing	1
7	MOS7	41731	r	r	r	r	r	r	r	r	ШЛУ	LAM9	770
8	MOS8	41458	s	r	s	r	r	s	s	r	поли	Beijing	1
9	MOS9	36376	r	r	s	r	s	s	s	s	МЛУ	H4(Ural)	не определено
10	MOS10	37312	r	r	r	r	r	r	r	r	ШЛУ	Beijing	1
11	MOS11	38602	r	r	r	r	r	r	r	r	ШЛУ	Beijing	1
12	MOS12	49190	r	r	r	r	r	r	r	r	ШЛУ	Beijing	1
13	MOS13	37910	r	r	r	r	s	s	s	r	МЛУ	Beijing	190
14	MOS14	42710	r	r	r	r	r	r	r	r	ШЛУ	Beijing	1
15	MOS15	13217	r	r	s	r	s	s	s	s	МЛУ	Beijing	1
16	MOS16	26064	r	r	r	r	s	s	r	s	МЛУ	Beijing	1
17	MOS17	26767	r	r	s	s	s	s	s	r	МЛУ	Beijing	1
18	MOS18	30137	r	r	r	r	s	s	s	r	МЛУ	Beijing	1
19	MOS19	38040	s	s	s	s	s	s	s	s	чув	T1	159
20	MOS20	35990	r	r	r	r	r	s	r	r	ШЛУ	T1	154
21	MOS21	46078	r	r	r	r	r	s	s	s	МЛУ	Beijing	1
22	MOS22	40779	r	r	r	r	r	s	s	r	МЛУ	Beijing	1
23	MOS23	41240	r	r	s	r	s	s	s	s	МЛУ	Beijing	1
24	MOS24	42841	r	r	s	r	s	s	s	r	МЛУ	Beijing	1
25	MOS25	40242	r	r	r	r	s	s	s	s	МЛУ	Beijing	1
26	MOS26	42243	r	r	r	r	s	s	s	r	МЛУ	Beijing	1
27	MOS27	50112	r	r	r	r	r	s	s	r	МЛУ	Beijing	1
28	MOS28	56392	r	r	r	r	s	r	s	s	МЛУ	Beijing	1
29	MOS29	64550	r	r	r	r	r	s	s	s	МЛУ	Beijing	1
30	MOS30	64757	r	s	s	r	r	s	r	r	поли	Beijing	1
31	MOS31	66481	r	r	s	r	s	s	s	r	МЛУ	Beijing	1
32	MOS32	68101	r	r	r	r	r	s	s	r	МЛУ	Beijing	1
33	MOS33	76229	r	r	r	r	s	r	r	r	МЛУ	Beijing	1

34	MOS34	79230	r	r	r	r	s	r	r	r	МЛУ	Beijing	1
35	MOS35	70231	r	r	s	r	s	r	r	s	МЛУ	Beijing	1
36	MOS36	71282	r	r	s	r	r	s	s	s	МЛУ	Beijing	1
37	MOS37	69806	s	s	s	r	s	s	s	s	моно	T1	353
38	MOS38	69303	s	s	s	s	s	s	s	s	чув	T5_RUS1 (LAM)	254
39	MOS39	67539	s	s	s	s	r	r	r	s	поли	T5_RUS1 (LAM)	496
40	MOS40	59681	r	r	r	r	r	s	r	r	ШЛУ	T5_RUS1 (LAM)	496
41	MOS41	51417	r	s	s	r	s	s	s	s	поли	Beijing	190
42	MOS42	67040	s	s	s	s	r	r	r	s	поли	T5_RUS1 (LAM)	496
43	MOS43	68302	r	r	r	r	r	r	r	r	ШЛУ	T5_RUS1 (LAM)	496
44	MOS44	72586	s	s	s	s	s	s	s	s	чув	H3	50
45	MOS45	23663	r	r	r	r	r	s	s	r	МЛУ	Beijing	1
46	MOS46	29149	r	r	r	r	r	s	s	s	МЛУ	Beijing	1
47	MOS47	49229	r	r	r	r	r	r	s	s	МЛУ	Beijing	1
48	MOS48	50341	r	r	r	r	r	r	r	r	МЛУ	Beijing	1
49	MOS49	66568	r	r	s	r	r	s	s	s	МЛУ	Beijing	1
50	MOS50	72620	r	r	r	r	s	s	s	r	МЛУ	Beijing	1
51	SP1	7283	r	r	r	r	r	s	s	s	МЛУ	Beijing	1
52	SP2	3171	r	r	s	r	s	s	s	r	МЛУ	Beijing	1
53	SP3	3227	r	r	r	r	r	s	s	s	МЛУ	Beijing	1
54	SP4	2050	r	r	r	r	r	s	s	r	ШЛУ	Beijing	1
55	SP5	8692	s	s	s	s	s	s	s	s	чув	Beijing	1
56	SP6	6675	r	r	r	r	s	s	s	r	МЛУ	Beijing	1
57	SP7	6682	r	r	r	r	r	r	r	r	ШЛУ	Beijing	1
58	SP8	1059	r	r	s	r	s	s	s	s	МЛУ	Beijing	1
59	SP9	7941	s	s	s	s	s	s	s	s	чув	Beijing	1
60	SP10	3670	r	r	r	r	r	s	s	s	МЛУ	Beijing	1
61	SP11	8190	r	r	r	r	r	s	s	r	ШЛУ	Beijing	1
62	SP12	8340	r	r	r	r	s	s	r	r	МЛУ	Beijing	1
63	SP13	45	r	r	r	r	s	s	s	r	МЛУ	Beijing	1
64	SP14	7227	r	r	r	r	r	r	r	r	ШЛУ	Beijing	1
65	SP15	5582	r	r	s	r	s	s	s	s	МЛУ	Beijing	1
66	SP16	81	r	r	r	r	s	s	s	s	МЛУ	Beijing	1
67	SP17	5116	r	r	r	r	s	r	r	s	МЛУ	Beijing	1
68	SP18	975	r	r	r	r	s	s	s	r	МЛУ	Beijing	1
69	SP19	7344	r	r	r	r	r	r	r	r	ШЛУ	Beijing	1
70	SP20	6793	r	r	r	r	r	s	r	r	ШЛУ	Beijing	1
71	SP21	3851	r	r	r	r	r	s	r	r	ШЛУ	Beijing	1
72	SP22	2531	r	r	s	r	r	s	r	r	ШЛУ	Beijing	1
73	SP23	1317	r	r	r	r	r	s	r	r	ШЛУ	Beijing	1
74	SP24	7713	r	r	s	r	r	s	s	s	МЛУ	H4(Ural)	262
75	SP25	8968	r	r	r	r	r	s	r	r	ШЛУ	H4(Ural)	262

76	SP26	3128	r	r	r	r	r	r	r	r	ШЛЮ	Beijing	1
77	SP27	3083	r	r	r	r	r	r	r	s	ШЛЮ	Beijing	1
78	SP28	7861	r	r	r	r	r	r	r	s	ШЛЮ	H4(Ural)	262
79	SP29	7824	r	r	r	r	r	s	s	r	ШЛЮ	Beijing	1
80	SP30	4999	r	r	r	r	r	r	s	s	ШЛЮ	Beijing	1
81	SP31	6295	r	r	s	r	s	s	s	s	МЛЮ	H4(Ural)	1134
82	SP32	110	r	r	r	r	r	s	s	s	МЛЮ	H4(Ural)	262
83	SP33	1048	r	r	r	r	r	s	s	s	МЛЮ	H4(Ural)	262
84	SP34	7517	r	r	r	r	s	s	s	s	МЛЮ	Beijing	1
85	SP35	1290	r	r	s	r	s	s	s	s	МЛЮ	LAM9	42
86	SP36	1582	r	r	s	r	s	s	s	s	МЛЮ	LAM9	42
87	SP37	2269	r	r	r	r	s	r	r	r	МЛЮ	LAM9	770
88	SP38	1731	r	r	r	r	r	r	r	r	ШЛЮ	LAM9	770
89	SP39	7099	r	r	s	r	r	r	s	r	ШЛЮ	LAM9	42
90	SP40	7776	r	r	s	r	s	s	r	r	МЛЮ	Beijing	1
91	SP41	3210	r	r	s	r	s	s	s	r	МЛЮ	H4(Ural)	262
92	SP42	3963	s	s	s	s	s	s	s	s	чув	H4(Ural)	777
93	SP43	3336	r	r	r	r	r	r	r	r	ШЛЮ	Beijing	1
94	SP44	2343	r	r	r	r	r	r	r	r	ШЛЮ	Beijing	1
95	SP45	6679	s	s	s	s	s	s	s	s	чув	Beijing	1
96	SP46	1510	s	s	s	s	s	s	s	s	чув	LAM9	42
97	SP47	3694	s	s	s	s	s	s	s	s	чув	Beijing	1
98	SP48	1433	s	s	s	s	s	s	s	s	чув	Beijing	1
99	SP49	5149	s	s	s	s	s	s	s	s	чув	Beijing	1
100	SP50	3595	s	s	s	s	s	s	s	s	чув	Beijing	1
101	SP51	9424	r	r	r	r	r	r	s	r	ШЛЮ	Beijing	1
102	SP52	2757	r	r	r	r	r	s	s	s	МЛЮ	Beijing	1
103	SP53	7778	r	r	s	r	s	r	r	r	МЛЮ	Beijing	1
104	SP54	7779	r	r	s	r	s	r	r	s	МЛЮ	Beijing	1
105	SP55	7939	r	r	r	r	r	s	s	r	МЛЮ	Beijing	1
106	SP56	8334	r	r	s	r	s	s	s	r	МЛЮ	Beijing	1
107	SP57	8335	r	s	s	r	s	s	s	s	поли	Beijing	1
108	SP58	8545	r	r	s	r	s	s	s	r	МЛЮ	Beijing	1
109	SP59	8583	r	r	r	r	r	s	s	s	МЛЮ	Beijing	1
110	SP60	8856	r	r	r	r	r	s	s	s	МЛЮ	Beijing	1
111	SP61	8903	r	r	r	r	s	r	r	r	МЛЮ	Beijing	1
112	SP62	9080	r	r	r	r	r	s	s	r	МЛЮ	Beijing	1
113	SP63	9116	r	r	s	r	s	r	s	r	МЛЮ	Beijing	1
114	SP64	9622	r	r	r	r	r	s	s	r	МЛЮ	Beijing	1
115	SP65	9682	r	s	s	r	r	r	s	s	поли	Beijing	1
116	SP66	9683	r	s	s	r	s	s	s	s	поли	Beijing	1
117	SP67	9118	r	r	s	r	r	s	r	r	МЛЮ	T5_RUS1	496
118	SP68	8185	r	s	s	s	s	s	r	s	поли	H3	50

119	SP69	758	r	r	r	r	r	s	s	s	МЛУ	Beijing	1
120	SP70	1469	r	r	s	r	s	r	r	r	МЛУ	Beijing	1
121	SP71	5849	r	s	r	r	s	r	r	r	поли	Beijing	1
122	SP72	6141	r	r	r	r	s	s	s	r	МЛУ	Beijing	1
123	SP73	6591	r	r	r	r	s	r	r	s	МЛУ	Beijing	1
124	SP74	6676	r	r	s	r	s	s	s	r	МЛУ	Beijing	1
125	SP75	7106	r	r	r	r	s	r	r	s	МЛУ	Beijing	1
126	SP76	7829	r	r	r	r	s	s	s	r	МЛУ	Beijing	1
127	SP77	7910	r	r	s	r	s	r	s	s	МЛУ	Beijing	1
128	SP78	8864	r	r	s	r	s	s	s	r	МЛУ	Beijing	1
129	SP79	3050	s	s	s	s	s	s	s	s	чув	T1	53
130	SP80	1145	r	r	r	r	s	s	s	s	МЛУ	Beijing	1
131	SP81	2191	r	r	r	r	s	s	s	r	МЛУ	Beijing	1
132	SP82	2835	r	r	s	r	s	s	s	s	МЛУ	Beijing	1
133	SP83	7425	r	r	r	r	s	s	s	s	МЛУ	Beijing	1
134	SP84	7684	r	s	r	r	s	s	s	s	поли	Beijing	1
135	SP85	9660	r	s	r	r	s	s	s	r	МЛУ	Beijing	1
136	SP86	1056	s	s	s	s	s	s	s	s	чув	H4(Ural)	35
137	SP87	6081	s	s	s	s	s	s	s	s	чув	T2	52
138	SP88	1655	s	s	s	s	s	s	s	s	чув	H3	50
139	SP89	5691	r	s	r	r	r	r	r	r	МЛУ	Beijing	1
140	SP90	5966	r	r	r	r	r	s	s	r	МЛУ	Beijing	1
141	SP91	3892	r	r	s	r	s	s	s	r	МЛУ	LAM9	42
142	SP92	3163	r	s	s	r	s	s	s	s	поли	H4(Ural)	35
143	SP93	2747	r	r	s	r	r	s	s	s	МЛУ	H3-T3	36
144	SP94	5944	r	s	s	r	r	s	s	r	ШЛУ	H4(Ural)	262
145	SP95	5962	s	s	s	r	r	s	s	r	поли	T5_RUS1 (LAM)	254
146	SP96	5974	r	r	r	r	r	r	s	r	ШЛУ	T5_RUS1 (LAM)	496
147	SP97	2927	s	s	s	s	s	s	s	s	чув	H3	50
148	CTRI-2	R894	s	s	s	s	s	s	s	s	чув	LAM9	42
149	CTRI-3	R975	r	r	r	r	s	r	r	r	МЛУ	H4(Ural)	262
150	CTRI-4	R849	r	r	r	r	r	r	r	r	ШЛУ	Beijing	269

серым цветом отмечены образцы для
полногеномного секвенирования

Приложение 2. Результаты VNTR-типирования 54 штаммов *M. tuberculosis*. Образцы сгруппированы согласно результатам сполитоирования

№ образца	сполитоип	VNTR-локусы																				MIRU-тип согласно Мокроусову с соавторами				
		MIRU 2*	Mtub 04	ETR-C	MIRU 4*	MIRU 40*	MIRU 10*	MIRU 16*	Mtub 21	MIRU 20*	QUB 11b	ETR-A	Mtub 29	Mtub 30	ETR-B	MIRU 23*	MIRU 24*	MIRU 26*	MIRU 27*	Mtub 34	MIRU 31*		Mtub 39	QUB 26	QUB 4156	MIRU 39*
SP1	Beijing	2	4	4	2	3	3	3	5	2	6	4	4	4	2	5	1	7	3	3	5	3	7	2	3	M11
SP7	Beijing	2	4	4	2	3	3	3	5	2	6	4	4	4	2	5	1	7	3	3	5	3	7	2	3	M11
SP13	Beijing	2	4	4	2	3	3	3	5	2	6	4	4	4	2	5	1	7	3	3	5	3	7	2	3	M11
SP21	Beijing	2	4	4	2	3	3	3	5	2	6	4	4	4	2	5	1	7	3	3	5	3	7	2	3	M11
SP23	Beijing	2	4	4	2	3	3	3	5	2	6	4	4	4	2	5	1	7	3	3	5	3	7	2	3	M11
SP27	Beijing	2	4	4	2	3	3	3	5	2	6	4	4	4	2	5	1	7	3	3	5	3	7	2	3	M11
SP45	Beijing	2	4	4	2	3	3	3	5	2	6	4	4	4	2	5	1	7	3	3	5	3	7	2	3	M11
MOS11	Beijing	2	4	4	2	3	3	3	5	2	6	4	4	4	2	5	1	7	3	3	5	3	7	2	3	M11
MOS15	Beijing	2	4	4	2	3	3	3	5	2	6	4	4	4	2	5	1	7	3	3	5	3	7	2	3	M11
SP6	Beijing	2	4	4	2	3	3	3	5	2	6	4	4	4	2	5	1	7	3	3	5	3	6	2	3	M11
SP29	Beijing	2	4	4	2	3	3	3	5	2	6	4	4	4	2	5	1	7	3	3	5	3	6	2	3	M11
MOS14	Beijing	2	4	4	2	3	3	3	5	2	6	4	4	4	2	5	1	7	3	3	5	3	6	2	3	M11
SP22	Beijing	2	4	4	2	3	3	3	5	2	5	4	4	4	2	5	1	7	3	3	5	3	7	2	3	M11
MOS16	Beijing	2	4	4	2	3	3	3	5	2	5	4	4	4	2	5	1	7	3	3	5	3	7	2	3	M11
SP10	Beijing	2	4	4	2	3	3	3	4	2	6	4	4	4	2	5	1	7	3	3	5	3	7	2	3	M11
SP3	Beijing	2	4	4	2	3	3	3	5	2	6	4	4	4	2	5	1	5	3	3	5	3	6	2	3	M2
SP4	Beijing	2	4	4	2	3	3	3	5	2	6	4	4	4	2	5	1	5	3	3	5	3	6	2	3	M2
SP18	Beijing	2	4	4	2	3	3	3	5	2	6	4	4	4	2	5	1	5	3	3	5	3	6	2	3	M2
SP26	Beijing	2	4	4	2	3	3	3	5	2	6	4	4	4	2	5	1	5	3	3	5	3	6	2	3	M2

SP30	Beijing	2	4	4	2	3	3	3	5	2	6	4	4	4	2	5	1	5	3	3	5	3	6	2	3	M2	
SP34	Beijing	2	4	4	2	3	3	3	5	2	6	4	4	4	2	5	1	5	3	3	5	3	6	2	3	M2	
MOS17	Beijing	2	4	4	2	3	3	3	5	2	6	4	4	4	2	5	1	5	3	3	5	3	6	2	3	M2	
MOS18	Beijing	2	4	4	2	3	3	3	5	2	6	4	4	4	2	5	1	5	3	3	5	3	6	2	3	M2	
SP5	Beijing	2	4	4	2	3	3	3	5	2	6	4	4	4	2	5	1	5	3	3	5	3	8	2	3	M2	
SP12	Beijing	2	4	4	2	3	3	3	5	2	6	4	4	4	2	5	1	5	3	3	5	3	8	2	3	M2	
SP16	Beijing	2	4	4	2	3	3	3	5	2	6	4	4	4	2	5	1	5	3	3	5	3	8	2	3	M2	
MOS10	Beijing	2	4	4	2	3	3	3	5	2	6	4	4	4	2	5	1	5	3	3	5	3	8	2	3	M2	
SP9	Beijing	2	4	4	2	3	3	3	5	2	6	4	4	4	2	5	1	5	3	3	5	3	7	2	3	M2	
SP11	Beijing	2	2	4	2	3	3	3	5	2	6	4	4	4	2	5	1	5	3	3	5	3	6	2	3	M2	
SP15	Beijing	2	4	4	2	3	3	3	5	2	5	4	4	4	2	5	1	5	3	3	5	3	8	2	3	M2	
MOS12	Beijing	2	4	4	2	3	3	3	5	2	6	4	4	2	2	5	1	5	3	3	5	3	8	2	3	M2	
SP2	Beijing	2	4	4	2	3	3	3	5	2	6	3	4	4	2	5	1	5	3	3	5	3	8	2	3	M2	
SP8	Beijing	2	4	4	2	3	3	3	5	2	6	4	4	4	2	5	1	6	3	3	5	3	8	2	3	M33	
SP17	Beijing	2	4	4	2	3	1	3	3	2	1	4	4	4	2	5	1	7	3	3	5	3	9	2	3	M12	
CTRI-4	Beijing-like	2	2	5	2	3	3	3	4	2	6	4	4	4	2	5	1	7	3	3	4	3	7	3	2	Нет данных	
SP25	H4 (Ural)	2	3	5	2	2	7	2	3	2	2	4	4	4	2	5	1	1	3	3	2	3	6	3	2		
SP33	H4 (Ural)	2	3	5	2	2	7	2	3	2	2	4	4	4	2	5	1	1	3	3	2	3	6	3	2		
SP24	H4 (Ural)	2	3	5	2	4	7	2	3	2	2	4	4	4	2	5	1	1	3	3	2	3	6	3	2		
SP28	H4 (Ural)	2	3	5	2	3	7	1	3	2	2	6	4	4	2	5	1	1	3	3	2	3	6	3	2		
SP31	H4 (Ural)	2	3	4	2	2	7	2	3	2	1	6	4	4	2	5	1	1	3	3	3	3	8	3	2		
SP32	H4 (Ural)	2	3	5	2	3	7	2	3	2	2	4	4	4	2	5	1	1	3	3	2	3	8	3	2		
SP41	H4 (Ural)	2	3	5	2	3	6	2	3	2	2	4	4	4	2	5	1	1	3	3	2	3	6	3	2		
SP42	H4 (Ural)	2	3	4	2	3	6	2	2	2	1	3	4	4	2	5	1	1	3	3	2	3	7	3	2		
MOS9	H4 (Ural)	2	3	4	2	2	7	2	3	2	2	4	4	4	2	5	1	1	3	3	3	3	8	3	2		
CTRI-3	H4 (Ural)	2	3	5	2	4	7	1	3	2	2	4	4	4	2	5	1	1	3	3	2	4	6	3	2		
SP35	LAM9	1	3	2	2	5	4	3	3	2	2	2	4	1	2	5	1	5	3	3	2	2	6	2	2		
SP37	LAM9	1	3	2	2	5	4	3	3	2	2	2	4	1	2	5	1	5	3	3	2	2	6	2	2		

SP39	LAM9	1	3	2	2	5	4	3	3	2	2	2	4	1	2	5	1	5	3	3	2	2	6	2	2
MOS2	LAM9	1	3	2	2	5	3	3	3	2	2	2	4	1	2	5	1	5	3	3	2	2	6	2	2
MOS7	LAM9	1	3	2	2	5	3	3	3	2	2	2	4	1	2	5	1	5	3	3	2	2	6	2	2
SP36	LAM9	1	2	2	2	5	4	3	3	2	2	2	4	1	2	5	1	5	3	3	2	2	6	2	2
SP38	LAM9	2	3	2	2	5	4	3	3	2	2	2	4	1	2	5	1	5	3	3	2	2	6	2	2
SP46	LAM9	1	3	2	2	4	4	3	3	2	2	2	4	1	2	5	1	5	3	3	2	2	6	2	2
CTRI-2	LAM9	1	3	2	2	5	4	2	3	2	1	2	4	1	2	5	1	5	3	3	2	3	6	2	2

Приложение 3. Специфические полиморфизмы для представителей генетического семейства Beijing. Серым цветом отмечены полиморфизмы, исключенные из анализа в ходе процесса валидации.

Позиция	Ген	№ локуса	Ам-та реф	Ам-та мут	Нукл. реф	Нукл. мут.	Позиция	Ген	№ локуса	Ам-та реф	Ам-та мут	Нукл. реф	Нукл. мут.
1849	межген.				C	A	2288085	-	Rv2042c	Ala	Ala	G	C
11820	межген.				C	G	2294007	-	Rv2047c	Thr	Ala	T	C
14861	-	Rv0012	Gly	Val	G	T	2295685	pks12	Rv2048c	Val	Leu	C	A
15117	trpG	Rv0013	Ile	Met	C	G	2298194	pks12	Rv2048c	Ala	Ala	G	T
16119	rknB	Rv0014c	Arg	Leu	C	A	2314689	rpmG1	Rv2057c	Arg	Leu	C	A
25610	межген.				G	C	2315386	-	Rv2059	His	Gln	C	A
35608	bioF2	Rv0032	Ala	Ala	C	T	2321975	cobG	Rv2064	Asp	Asp	T	C
36008	bioF2	Rv0032	Asp	His	G	C	2328543	cobM	Rv2071c	Ile	Met	T	C
37305	fadD34	Rv0035	Ser	Trp	C	G	2331620	-	Rv2075c	Gly	Gly	A	C
37334	fadD34	Rv0035	Thr	Ser	A	T	2331789	-	Rv2075c	Pro	Gln	G	T
39158	-	Rv0036c	Arg	Arg	C	G	2337179	-	Rv2079	Gln	Ter	C	T
42281	-	Rv0039c	Cys	Phe	C	A	2339255	-	Rv2082	Thr	Ala	A	G
48503	-	Rv0044c	Arg	Arg	T	C	2339835	-	Rv2082	Pro	Leu	C	T
49360	-	Rv0045c	Val	Ile	C	T	2344246	межген.				G	C
63146	межген.				G	T	2344499	rknJ	Rv2088	Leu	Arg	T	G
65150	-	Rv0061c	Trp	Ter	C	T	2346318	pepE	Rv2089c	Pro	Leu	G	A
70267	-	Rv0064	Val	Phe	G	T	2354791	pafB	Rv2096c	Ser	Pro	A	G
75233	межген.				C	A	2364475	varC37	Rv2103c	Thr	Ala	T	C
87468	-	Rv0078A	Glu	Lys	C	T	2369186	prcA	Rv2109c	Arg	Gly	G	C
94388	hycQ	Rv0086	Gly	Gly	G	A	2388641	ansP1	Rv2127	Gly	Asp	G	A
96894	межген.				C	T	2411496	murC	Rv2152c	Phe	Phe	G	A
97350	-	Rv0088	Arg	Arg	A	C	2421816	-	Rv2160A	Cys	Arg	A	G
103756	-	Rv0094c	Pro	His	G	T	2421816	-	Rv2160c	Val	Ala	A	G
105139	-	Rv0095c	Gly	Val	C	A	2425471	pbpB	Rv2163c	Arg	Arg	T	C
105736	PPE1	Rv0096	Val	Ala	T	C	2442468	-	Rv2180c	Trp	Ter	C	T
123454	-	Rv0104	Gln	Ter	C	T	2448458	fadD15	Rv2187	Thr	Ile	C	T
123520	-	Rv0104	Tyr	His	T	C	2451700	-	Rv2189c	Ala	Ala	A	C
135329	-	Rv0111	Pro	Pro	C	T	2491597	glnE	Rv2221c	Arg	Arg	G	T
139297	hddA	Rv0115	Gly	Val	G	T	2505085	cobC	Rv2231c	Ala	Ala	G	A
143207	oxcA	Rv0118c	Ser	Gly	T	C	2510350	межген.				C	G
146236	fusA2	Rv0120c	Ala	Ala	C	G	2521428	accD6	Rv2247	Asp	Gly	A	G
147262	fusA2	Rv0120c	Asp	Glu	G	T	2522955	-	Rv2248	Val	Gly	T	G

163706	cyp138	Rv0136	Pro	Leu	C	T	2533377	-	Rv2260	Asp	Asp	T	C
166565	-	Rv0139	Ile	Val	A	G	2536599	-	Rv2264c	Gly	Gly	C	T
182824	fadE2	Rv0154c	Ile	Ile	G	A	2543395	cyp128	Rv2268c	Glu	Glu	T	C
190816	-	Rv0161	Ser	Ser	A	C	2546684	mazF8	Rv2274c	Gly	Val	C	A
195682	fadD5	Rv0166	Val	Val	C	G	2548700	cyp121	Rv2276	His	Tyr	C	T
199470	mce1A	Rv0169	Ser	Ala	T	G	2562644	lppO	Rv2290	Ala	Ser	G	T
203038	mce1D	Rv0172	Ile	Thr	T	C	2581052	-	Rv2308	Asp	Asn	G	A
206481	mce1F	Rv0174	Pro	Pro	C	G	2586076	-	Rv2314c	Arg	Arg	C	G
208590	-	Rv0177	Arg	Arg	A	C	2589036	uspA	Rv2316	Asp	His	G	C
213147	sigG	Rv0182c	Asp	Tyr	C	A	2600260	-	Rv2327	Arg	Arg	A	G
213281	sigG	Rv0182c	Gly	Asp	C	T	2601576	PE23	Rv2328	Ser	Arg	T	G
217201	bgIS	Rv0186	Asn	Asn	T	C	2601760	PE23	Rv2328	Ala	Thr	G	A
222925	-	Rv0191	Ala	Thr	G	A	2612256	-	Rv2337c	Leu	Leu	G	T
230170	-	Rv0194	Pro	Leu	C	T	2619271	межген.				T	C
249522	-	Rv0209	Val	Ala	T	C	2627618	plcC	Rv2349c	Gly	Cys	C	A
251575	-	Rv0210	Ala	Thr	G	A	2627946	plcC	Rv2349c	Arg	Arg	A	G
262268	-	Rv0219	Ala	Ala	A	T	2630740	plcA	Rv2351c	Thr	Ala	T	C
264129	-	Rv0221	Met	Ile	G	A	2634282	межген.				T	G
267751	-	Rv0223c	Ala	Gly	G	C	2637162	межген.				C	T
270430	-	Rv0226c	Pro	Ala	G	C	2647905	-	Rv2366c	Arg	Cys	G	A
278021	-	Rv0232	Arg	Arg	T	C	2663795	mbtE	Rv2380c	Ala	Ala	G	A
282892	aftD	Rv0236c	Thr	Thr	C	T	2673818	mbtB	Rv2383c	Val	Leu	C	G
283610	aftD	Rv0236c	Ala	Val	G	A	2697489	subI	Rv2400c	Ala	Thr	C	T
288260	lpqI	Rv0237	Leu	Leu	C	T	2706663	-	Rv2409c	Thr	Asn	G	T
296312	-	Rv0245	Ser	Phe	C	T	2706663	PE24	Rv2408	Gly	Val	G	T
305188	nirB	Rv0252	Val	Leu	G	T	2738274	obg	Rv2440c	Ser	Arg	A	C
311889	-	Rv0260c	Thr	Thr	C	T	2740693	межген.				T	C
325505	fadD2	Rv0270	Val	Val	T	C	2741209	dctA	Rv2443	Leu	Leu	G	A
342146	eccA3	Rv0282	Glu	Ala	A	C	2752122	rpfE	Rv2450c	Thr	Arg	G	C
353309	eccD3	Rv0290	Ser	Asn	G	A	2774555	aglA	Rv2471	Val	Ala	T	C
353365	eccD3	Rv0290	Ala	Thr	G	A	2775361	-	Rv2472	Arg	Arg	C	T
363464	-	Rv0298	Arg	Arg	G	A	2789798	plsC	Rv2483c	Arg	Arg	G	T
374353	PPE6	Rv0305c	Arg	Arg	A	G	2791098	-	Rv2484c	Gly	Asp	C	T
377014	-	Rv0307c	Leu	Leu	C	T	2793810	lipQ	Rv2485c	Arg	Leu	C	A
379528	-	Rv0311	Glu	Asp	G	T	2794071	межген.				G	A
383716	-	Rv0315	Pro	Ser	C	T	2807486	-	Rv2492	Ala	Asp	C	A
392261	-	Rv0325	Ter	Gln	T	C	2809633	bkdC	Rv2495c	Tyr	Asp	A	C
396199	межген.				T	C	2817502	accD1	Rv2502c	Phe	Val	A	C

424981	PPE8	Rv0355c	Pro	Pro	G	A	2825466	-	Rv2509	Lys	Lys	G	A
428698	PPE8	Rv0355c	Ala	Ala	C	T	2825581	-	Rv2510c	Ile	Ile	T	G
432536	PPE8	Rv0355c	Gly	Asp	C	T	2833329	-	Rv2516c	Ala	Val	G	A
441823	fba	Rv0363c	Glu	Glu	C	T	2835984	PE26	Rv2519	Val	Ala	T	C
445780	-	Rv0368c	Arg	His	C	T	2838897	-	Rv2522c	Met	Ile	C	G
454333	-	Rv0376c	Thr	Pro	T	G	2841022	fas	Rv2524c	Cys	Arg	A	G
456594	-	Rv0380c	Met	Val	T	C	2847281	fas	Rv2524c	Asp	Asp	A	G
460413	clpB	Rv0384c	Lys	Lys	C	T	2850631	межген.				T	C
465300	-	Rv0386	Phe	Phe	C	T	2866607	lppA	Rv2543	Gly	Asp	G	A
484005	fadD30	Rv0404	Arg	Leu	G	T	2866647	lppA	Rv2543	Lys	Lys	G	A
484596	fadD30	Rv0404	Pro	Leu	C	T	2866671	lppA	Rv2543	Glu	Glu	G	A
491742	fgd1	Rv0407	Phe	Phe	T	C	2866677	lppA	Rv2543	Leu	Leu	A	C
492150	pta	Rv0408	Gly	Ala	G	C	2866749	lppA	Rv2543	Asp	Glu	C	A
505974	lpqM	Rv0419	Ala	Thr	G	A	2866757	lppA	Rv2543	Asn	Thr	A	C
518987	-	Rv0430	Arg	Arg	C	T	2866765	lppA	Rv2543	Gln	Lys	C	A
533547	rskA	Rv0444c	Glu	Asp	C	G	2866770	lppA	Rv2543	Ser	Ser	T	C
542014	межген.				C	G	2866831	lppA	Rv2543	Asn	Asp	A	G
542514	-	Rv0452	His	Asp	C	G	2866839	lppA	Rv2543	Val	Val	A	T
546357	echA2	Rv0456c	Thr	Thr	A	G	2866842	lppA	Rv2543	Met	Ile	G	C
555991	-	Rv0465c	Cys	Arg	A	G	2866845	lppA	Rv2543	Phe	Phe	C	T
561776	межген.				C	T	2866863	lppA	Rv2543	Ala	Ala	C	G
561777	межген.				C	T	2866876	lppA	Rv2543	Ile	Val	A	G
572591	lprQ	Rv0483	Asp	Asp	C	T	2866880	lppA	Rv2543	Ala	Val	C	T
575907	mshA	Rv0486	Ala	Val	C	T	2866882	lppA	Rv2543	Ala	Thr	G	A
583171	-	Rv0492c	Ser	Ala	A	C	2867461	lppB	Rv2544	Ala	Gly	C	G
584171	-	Rv0493c	Ser	Gly	T	C	2867483	lppB	Rv2544	Met	Ile	G	C
599868	mmpL2	Rv0507	Arg	Arg	A	G	2867487	lppB	Rv2544	Lys	Glu	A	G
605150	hemB	Rv0512	Arg	Arg	G	A	2867492	lppB	Rv2544	Pro	Pro	G	A
611048	межген.				T	G	2867501	lppB	Rv2544	Phe	Phe	C	T
628864	-	Rv0537c	Thr	Ala	T	C	2867522	lppB	Rv2544	Glu	Glu	A	G
640954	межген.				A	G	2867532	lppB	Rv2544	Ile	Val	A	G
648856	mgtA	Rv0557	Gly	Gly	T	C	2867536	lppB	Rv2544	Ala	Val	C	T
655986	межген.				T	G	2867538	lppB	Rv2544	Ala	Thr	G	A
657081	-	Rv0565c	Val	Val	C	T	2867552	lppB	Rv2544	Arg	Arg	A	G
657142	-	Rv0565c	Arg	His	C	T	2867575	lppB	Rv2544	Val	Ala	T	C
659341	межген.				T	C	2867594	lppB	Rv2544	Leu	Leu	G	A
662911	nrdZ	Rv0570	Ala	Ala	T	C	2878550	межген.				G	A
669033	-	Rv0575c	Asp	Val	T	A	2881974	-	Rv2562	Ala	Thr	G	A

670545	-	Rv0576	Arg	His	G	A	2884068	glnQ	Rv2564	Met	Leu	A	C
686264	yrbE2B	Rv0588	Ala	Thr	G	A	2886570	-	Rv2566	Glu	Glu	A	G
690450	mce2C	Rv0591	Ala	Ala	A	C	2898522	-	Rv2573	Tyr	Tyr	T	C
694531	mce2F	Rv0594	Asn	Ser	A	G	2903439	-	Rv2578c	Ser	Ser	G	A
696917	межген.				G	T	2925962	varC40	Rv2596	Cys	Arg	T	C
699840	межген.				A	C	2948230	-	Rv2621c	Ala	Val	G	A
713310	galK	Rv0620	Cys	Arg	T	C	2955957	-	Rv2629	Asp	Ala	A	C
721373	recD	Rv0629c	Glu	Asp	T	G	2969197	межген.				A	G
738522	mmaA2	Rv0644c	Glu	Asp	T	G	2975227	-	Rv2650c	Asn	Thr	T	G
757139	atsD	Rv0663	Arg	Ser	C	A	2980861	межген.				A	G
757182	atsD	Rv0663	Asp	Gly	A	G	2983873	-	Rv2666	Ter	Ter	A	G
763031	ррoB	Rv0667	Ala	Ala	T	C	2988630	-	Rv2672	His	Asp	C	G
773497	echA4	Rv0673	Phe	Leu	C	A	2994187	hemE	Rv2678c	Leu	Leu	T	C
776100	mmpL5	Rv0676c	Thr	Ile	G	A	3003115	arsB1	Rv2685	Ala	Gly	C	G
776182	mmpL5	Rv0676c	Asp	Asn	C	T	3009834	ceoB	Rv2691	Asp	Gly	A	G
779615	-	Rv0679c	Asn	Lys	G	C	3010420	ceoC	Rv2692	Ile	Val	A	G
797736	-	Rv0697	Leu	Leu	C	T	3010993	-	Rv2693c	Gly	Arg	C	T
798355	-	Rv0697	Ala	Pro	G	C	3024021	ideR	Rv2711	Arg	Arg	C	A
799666	-	Rv0699	Asp	Gly	A	G	3024343	-	Rv2712c	Thr	Ile	G	A
811753	rplX	Rv0715	His	His	C	T	3027798	-	Rv2714	Val	Ala	T	C
820752	serA2	Rv0728c	Arg	His	C	T	3031168	-	Rv2719c	Tyr	His	A	G
835611	-	Rv0745	Thr	Met	C	T	3036826	fadE20	Rv2724c	Val	Met	C	T
847995	межген.				T	C	3056767	-	Rv2743c	Pro	Thr	G	T
861216	aldA	Rv0768	Leu	Arg	T	G	3079877	PPE44	Rv2770c	Phe	Ser	A	G
866051	ggtA	Rv0773c	Arg	Cys	G	A	3104189	-	Rv2795c	Cys	Cys	A	G
876141	ptrBb	Rv0782	Thr	Thr	G	T	3111476	-	Rv2802c	Ser	Ser	C	T
892416	-	Rv0799c	Val	Val	C	T	3112877	-	Rv2805	Gly	Asp	G	A
903537	purF	Rv0808	Arg	Leu	G	T	3112877	-	Rv2804c	Ala	Val	G	A
909166	cysA2	Rv0815c	Leu	Leu	C	T	3115384	межген.				G	C
919351	-	Rv0825c	Phe	Cys	A	C	3135606	межген.				A	G
931123	lpqQ	Rv0835	Tyr	Tyr	T	C	3137736	echA16	Rv2831	Asp	Tyr	G	T
932280	-	Rv0836c	Ter	Trp	T	C	3144623	-	Rv2837c	Ser	Asn	C	T
934230	межген.				C	G	3144623	rbfA	Rv2838c	Glu	Glu	C	T
934611	межген.				G	T	3174496	-	Rv2862c	Cys	Arg	A	G
940602	narL	Rv0844c	Gly	Arg	C	G	3177452	межген.				C	A
941722	-	Rv0845	Ile	Ser	T	G	3180988	rip	Rv2869c	Val	Phe	C	A
941845	-	Rv0845	Ala	Glu	C	A	3189242	mpt53	Rv2878c	Ala	Ala	A	G
951702	far	Rv0855	Ala	Val	C	T	3200282	-	Rv2891	Ala	Val	C	T

954101	dapC	Rv0858c	Ala	Thr	C	T	3214869	lppW	Rv2905	Gln	Arg	A	G
960367	-	Rv0862c	Leu	Ser	A	G	3249025	ppsA	Rv2931	Leu	Arg	T	G
972980	-	Rv0874c	Gly	Ser	C	T	3266030	ppsD	Rv2934	Ser	Ser	A	G
991896	-	Rv0890c	Glu	Gly	T	C	3284640	fadD28	Rv2941	Thr	Ala	A	G
996205	-	Rv0893c	Lys	Glu	T	C	3284855	fadD28	Rv2941	Ile	Ile	C	T
996346	межген.				G	A	3304966	-	Rv2952	Gly	Arg	G	A
1011511	-	Rv0907	Tyr	Ser	A	C	3310174	-	Rv2957	Lys	Lys	G	A
1022003	межген.				A	C	3310427	межген.				T	C
1024346	-	Rv0918	Ser	Gly	A	G	3310626	межген.				T	C
1028217	-	Rv0922	Gly	Asp	G	A	3314412	-	Rv2962c	Ala	Ala	A	G
1034758	межген.				C	T	3316955	purU	Rv2964	Met	Leu	A	C
1040251	pstS2	Rv0932c	Pro	Pro	C	A	3326554	-	Rv2971	Asn	His	A	C
1054784	-	Rv0945	Arg	Gly	C	G	3375165	-	Rv3015c	Phe	Cys	A	C
1071966	-	Rv0959	Asn	Asp	A	G	3379432	межген.				T	C
1076689	межген.				A	C	3380534	межген.				A	G
1076880	-	Rv0965c	Ala	Ala	C	T	3408150	-	Rv3047c	Thr	Ala	T	C
1080192	ctpV	Rv0969	Asp	Asn	G	A	3412611	nrdE	Rv3051c	Gln	Arg	T	C
1097220	mprA	Rv0981	Ser	Ser	C	T	3446699	-	Rv3081	Phe	Leu	C	G
1097442	mprA	Rv0981	Asp	Asp	C	T	3454263	-	Rv3087	Leu	Val	C	G
1098523	mprB	Rv0982	Leu	His	T	A	3467465	-	Rv3098c	Ala	Ala	C	G
1099058	межген.				G	A	3477942	moaA1	Rv3109	Thr	Thr	A	G
1102117	-	Rv0986	Val	Val	G	A	3480789	-	Rv3114	Ser	Pro	T	C
1102468	-	Rv0986	Gly	Gly	C	A	3484126	cysA3	Rv3117	Leu	Leu	G	A
1102646	-	Rv0987	Ala	Ala	G	A	3487108	cyp141	Rv3121	Ala	Ala	C	T
1103656	-	Rv0987	Ala	Val	C	T	3498198	devS	Rv3132c	Glu	Glu	C	T
1104690	-	Rv0987	Phe	Val	T	G	3500149	-	Rv3134c	Leu	Leu	G	A
1107940	-	Rv0990c	Ser	Ala	A	C	3500243	-	Rv3134c	Ala	Ala	G	C
1108521	межген.				G	A	3508970	fadB4	Rv3141	Glu	Glu	A	G
1111678	rimJ	Rv0995	Gly	Ser	G	A	3509626	межген.				C	A
1144585	eno	Rv1023	Arg	Gly	A	G	3510120	PPE52	Rv3144c	Lys	Gln	T	G
1145848	eno	Rv1023	Lys	Gln	A	C	3530955	-	Rv3161c	Val	Leu	C	G
1148259	межген.				A	G	3542049	межген.				G	A
1148930	kdpE	Rv1027c	Gly	Ser	C	T	3542219	-	Rv3174	Gly	Ser	G	A
1151054	kdpD	Rv1028c	Ser	Ser	C	G	3550347	межген.				C	T
1151304	kdpD	Rv1028c	Arg	Leu	C	A	3550789	-	Rv3183	Ala	Ala	C	G
1164336	межген.				G	A	3571828	nudC	Rv3199c	Pro	Arg	G	C
1165114	-	Rv1041c	Asp	Asn	C	T	3576231	-	Rv3201c	Gln	Arg	T	C
1165114	-	Rv1042c	Gly	Glu	C	T	3587446	-	Rv3210c	Leu	Leu	G	A

1168776	-	Rv1046c	Arg	Ser	T	G	3597249	-	Rv3220c	Gly	Arg	C	G
1172085	-	Rv1048c	Glu	Asp	C	A	3600576	-	Rv3224	Cys	Cys	C	T
1173750	-	Rv1050	Ala	Ala	G	T	3612813	-	Rv3236c	Thr	Ala	T	C
1175343	межген.				C	T	3628520	sahH	Rv3248c	Ser	Ser	G	A
1177157	-	Rv1054	Gly	Asp	G	A	3638093	pmmA	Rv3257c	Ser	Leu	G	A
1197157	межген.				T	G	3643985	-	Rv3263	Glu	Ala	A	C
1201581	lipU	Rv1076	Gln	Pro	A	C	3653988	-	Rv3272	Val	Met	G	A
1211369	-	Rv1086	Ser	Arg	A	C	3664457	-	Rv3282	Ala	Asp	C	A
1227830	glpX	Rv1099c	Thr	Thr	G	T	3665753	sseA	Rv3283	Glu	Lys	G	A
1230778	mazF3	Rv1102c	Thr	Ile	G	A	3670040	-	Rv3289c	Ala	Ala	C	T
1248382	-	Rv1125	Ser	Gly	A	G	3674157	pcd	Rv3293	Cys	Ser	T	A
1248936	-	Rv1125	Pro	Pro	G	C	3675504	-	Rv3294c	Ile	Ile	G	T
1250340	ppdK	Rv1127c	Ala	Ala	A	G	3678249	lhr	Rv3296	Lys	Thr	A	C
1254562	prpD	Rv1130	Asp	Gly	A	G	3680932	lhr	Rv3296	Pro	Pro	G	A
1265070	межген.				G	T	3690016	lpdA	Rv3303c	Leu	Ser	A	G
1278278	-	Rv1149	Gly	Glu	G	A	3693681	amiB1	Rv3306c	Ala	Ala	A	C
1281771	-	Rv1155	Ser	Pro	T	C	3704770	sdhD	Rv3317	Thr	Ala	A	G
1286766	mutT2	Rv1160	Gly	Arg	G	C	3735802	PPE54	Rv3343c	Thr	Thr	C	G
1288698	narG	Rv1161	Gly	Gly	G	A	3743656	межген.				G	A
1299305	PPE17	Rv1168c	Pro	Leu	G	A	3743899	PPE55	Rv3347c	Leu	Leu	G	A
1317655	pks4	Rv1181	Leu	Leu	C	T	3750185	PPE55	Rv3347c	Ser	Ser	C	G
1322741	mmpL10	Rv1183	Thr	Ala	A	G	3750193	PPE55	Rv3347c	Leu	Phe	G	A
1329234	-	Rv1186c	Asp	Asp	G	A	3750205	PPE55	Rv3347c	Asp	Asn	C	T
1336164	fadD36	Rv1193	Pro	Leu	C	T	3750407	PPE55	Rv3347c	Gly	Gly	G	C
1340208	PPE18	Rv1196	Arg	Gln	G	A	3750417	PPE55	Rv3347c	Phe	Tyr	A	T
1342581	-	Rv1199c	Thr	Ala	T	C	3750421	PPE55	Rv3347c	Ser	Gly	T	C
1349733	fadD6	Rv1206	Glu	Asp	G	C	3754508	-	Rv3349c	Leu	Val	G	C
1351172	межген.				A	G	3762013	PPE56	Rv3350c	Leu	Arg	A	C
1361190	-	Rv1217c	Ser	Ser	G	A	3775409	-	Rv3365c	Gln	Arg	T	C
1361285	-	Rv1217c	Ala	Thr	C	T	3778396	межген.				C	T
1390763	-	Rv1248c	Val	Met	C	T	3798596	межген.				A	G
1446733	argS	Rv1292	Ala	Thr	G	A	3804299	htdY	Rv3389c	Pro	Ser	G	A
1468492	-	Rv1313c	Leu	Leu	C	A	3811672	межген.				G	T
1477596	ogt	Rv1316c	Gly	Gly	C	T	3811675	межген.				A	G
1479085	alkA	Rv1317c	Ile	Val	T	C	3813266	guaA	Rv3396c	Val	Val	C	T
1490905	glgB	Rv1326c	Ser	Pro	A	G	3821845	-	Rv3402c	Ala	Glu	G	T
1498951	dinG	Rv1329c	Thr	Lys	G	T	3826501	vapB47	Rv3407	Arg	Cys	C	T
1501700	-	Rv1333	Val	Val	G	A	3830349	guaB2	Rv3411c	Ala	Thr	C	T

1523817	moeY	Rv1355c	Arg	Arg	C	T	3830695	guaB2	Rv3411c	Ala	Ala	A	G
1526311	межген.				G	T	3834486	-	Rv3415c	Gln	Gln	T	C
1540141	-	Rv1367c	Ile	Val	T	C	3841790	-	Rv3424c	Val	Ala	A	G
1541798	lprF	Rv1368	Ala	Val	C	T	3845695	-	Rv3428c	Arg	Arg	C	T
1544255	-	Rv1371	Arg	Arg	C	T	3847215	PPE59	Rv3429	Gly	Gly	T	C
1546703	-	Rv1373	Pro	Leu	C	T	3863138	esxU	Rv3445c	Pro	Ser	G	A
1556787	carA	Rv1383	Ser	Pro	T	C	3870010	mycP4	Rv3449	Thr	Ala	A	G
1561939	PPE20	Rv1387	Glu	Asp	G	C	3883178	rmlC	Rv3465	Glu	Asp	A	C
1576481	lipI	Rv1400c	Thr	Pro	T	G	3893480	kgtP	Rv3476c	Leu	Phe	G	A
1577241	-	Rv1401	Pro	Pro	G	A	3895585	PPE60	Rv3478	Pro	Leu	C	T
1594906	uvrC	Rv1420	Val	Ile	G	A	3895727	межген.				C	A
1595342	uvrC	Rv1420	Val	Ala	T	C	3906311	lipF	Rv3487c	Arg	Cys	G	A
1606673	PE16	Rv1430	Ala	Ala	G	T	3908465	otsA	Rv3490	Glu	Gly	A	G
1608276	-	Rv1431	Asn	Thr	A	C	3909235	otsA	Rv3490	Val	Leu	G	C
1616061	tpi	Rv1438	Ser	Ser	C	T	3916386	mce4C	Rv3497c	Arg	Ser	G	T
1617302	-	Rv1439c	Leu	Ile	G	T	3952308	fadD19	Rv3515c	Asp	Tyr	C	A
1620181	bisC	Rv1442	Leu	Val	C	G	3959730	ltp3	Rv3523	Phe	Leu	T	C
1638364	межген.				C	T	3962187	-	Rv3525c	Asp	Asp	G	A
1643864	-	Rv1458c	Thr	Ala	T	C	3964518	-	Rv3528c	Pro	Pro	G	A
1647436	-	Rv1461	Val	Ile	G	A	3968006	-	Rv3531c	Phe	Leu	A	C
1651308	-	Rv1463	Glu	Gly	A	G	3970594	межген.				C	G
1678706	-	Rv1489	Lys	Thr	A	C	3976595	kstD	Rv3537	Arg	Ser	G	T
1688064	lipL	Rv1497	Met	Val	A	G	3979990	ltp2	Rv3540c	Val	Leu	C	G
1688300	lipL	Rv1497	Phe	Phe	T	C	3984321	cyp125	Rv3545c	His	His	G	A
1695037	-	Rv1504c	Phe	Phe	G	A	3993058	fdxB	Rv3554	Gly	Glu	G	A
1695796	-	Rv1505c	Ala	Thr	C	T	4001622	межген.				T	C
1697896	межген.				C	T	4005114	fadE32	Rv3563	Trp	Ser	G	C
1699849	-	Rv1508c	Pro	Ser	G	A	4008747	hsaB	Rv3567c	Ile	Thr	A	G
1701215	межген.				C	T	4022388	-	Rv3579c	Pro	Pro	G	A
1706685	-	Rv1515c	Gly	Val	C	A	4041581	lysS	Rv3598c	Ala	Thr	C	T
1711670	-	Rv1520	Arg	Cys	C	T	4050811	ftsH	Rv3610c	Tyr	Tyr	G	A
1722200	pks5	Rv1527c	Val	Val	C	G	4056416	межген.				C	A
1733134	-	Rv1532c	Gln	Gln	C	T	4056693	межген.				G	T
1774659	межген.				C	T	4059143	-	Rv3618	Glu	Ala	A	C
1792777	-	Rv1592c	Ile	Val	T	C	4060588	esxW	Rv3620c	Thr	Ala	T	C
1792778	-	Rv1592c	Glu	Glu	T	C	4089058	-	Rv3649	Leu	Pro	T	C
1811375	trpB	Rv1612	Gly	Gly	G	T	4095295	-	Rv3655c	Glu	Glu	T	C
1831226	polA	Rv1629	Arg	Gly	A	G	4107074	dppA	Rv3666c	Gln	Arg	T	C

1834177	rpsA	Rv1630	Arg	Arg	A	C	4112429	-	Rv3671c	Ile	Val	T	C
1839759	-	Rv1634	Gly	Arg	G	C	4115890	nth	Rv3674c	Pro	Arg	G	C
1849051	lysX	Rv1640c	Pro	Pro	C	T	4120390	-	Rv3680	Gly	Ala	G	C
1849609	lysX	Rv1640c	Arg	Arg	T	C	4122287	ponA2	Rv3682	Ala	Ala	G	T
1849934	lysX	Rv1640c	Ile	Thr	A	G	4123770	ponA2	Rv3682	Ala	Thr	G	A
1859559	pheS	Rv1649	Ala	Asp	C	A	4133316	-	Rv3691	Thr	Ser	A	T
1877744	pks7	Rv1661	Glu	Ala	A	C	4137829	-	Rv3695	Ala	Ala	C	T
1892017	pks11	Rv1665	His	His	T	C	4145737	-	Rv3703c	Tyr	Tyr	A	G
1897608	-	Rv1672c	Leu	Leu	C	T	4156503	dnaQ	Rv3711c	Gly	Asp	C	T
1902337	межген.				A	C	4160371	recR	Rv3715c	Gly	Cys	C	A
1906078	moeX	Rv1681	Thr	Ala	A	G	4166441	dnaZX	Rv3721c	His	His	G	A
1924008	pyrG	Rv1699	Pro	Pro	G	A	4167656	-	Rv3722c	Met	Thr	A	G
1931718	PPE22	Rv1705c	Leu	Val	G	C	4169852	cut5b	Rv3724B	Ala	Ser	G	T
1955941	-	Rv1730c	Asp	Glu	G	C	4170739	-	Rv3725	Val	Leu	G	C
1961735	-	Rv1735c	Ala	Ala	A	G	4174131	-	Rv3727	Gly	Arg	G	C
1967026	-	Rv1739c	Ala	Ala	C	T	4179089	-	Rv3729	Pro	Ser	C	T
1967543	-	Rv1739c	Gly	Val	C	A	4179832	-	Rv3729	Gln	His	G	C
1971725	idi	Rv1745c	Arg	Arg	G	C	4182387	ligC	Rv3731	Val	Val	G	C
1980652	-	Rv1751	Pro	Pro	G	T	4186678	-	Rv3736	Leu	Leu	G	A
2003105	-	Rv1769	Phe	Phe	C	T	4189210	-	Rv3737	Pro	Pro	G	T
2003827	-	Rv1769	Asp	Ala	A	C	4194501	ctpJ	Rv3743c	Phe	Phe	G	A
2010614	межген.				G	A	4204168	proV	Rv3758c	Asn	Asp	T	C
2050822	PPE32	Rv1808	Gly	Ala	G	C	4205325	proX	Rv3759c	Trp	Arg	A	G
2067684	secA2	Rv1821	Val	Met	G	A	4215433	-	Rv3770c	Ala	Thr	C	T
2069771	-	Rv1823	Gln	Lys	C	A	4215484	-	Rv3770c	Pro	Ala	G	C
2072313	gcvH	Rv1826	Thr	Lys	C	A	4217557	hisC2	Rv3772	Thr	Ala	A	G
2078246	gcvB	Rv1832	Gly	Gly	C	G	4218941	echA21	Rv3774	Gly	Gly	G	A
2082053	-	Rv1835c	Cys	Arg	A	G	4221423	-	Rv3776	Ser	Leu	C	T
2086672	glcB	Rv1837c	Gly	Ser	C	T	4229167	glfT1	Rv3782	Val	Ala	T	C
2097270	-	Rv1847	Arg	Trp	C	T	4236237	dprE1	Rv3790	Thr	Thr	C	T
2108890	межген.				A	C	4243460	embA	Rv3794	Cys	Cys	C	T
2112832	-	Rv1865c	Ala	Ala	A	C	4251297	fadE35	Rv3797	Gly	Gly	G	C
2122976	lldD2	Rv1872c	Gly	Ala	C	G	4252066	fadE35	Rv3797	Ser	Pro	T	C
2138436	межген.				A	G	4254006	-	Rv3798	Leu	Leu	G	T
2138453	-	Rv1888A	Gly	Gly	C	T	4254431	accD4	Rv3799c	Asp	Asp	G	A
2142266	межген.				C	A	4256210	pks13	Rv3800c	Ala	Asp	G	T
2151222	-	Rv1904	Leu	Trp	T	G	4258106	pks13	Rv3800c	Thr	Met	G	A
2154724	katG	Rv1908c	Arg	Leu	C	A	4262388	fadD32	Rv3801c	Gly	Ser	C	T

2158109	fadB5	Rv1912c	Asp	Gly	T	C	4267647	aftB	Rv3805c	Asp	Gly	T	C
2167926	PPE35	Rv1918c	Leu	Ser	A	G	4284429	papA2	Rv3820c	Pro	Leu	G	A
2172380	-	Rv1920	Glu	Ala	A	C	4287164	-	Rv3822	Gly	Gly	A	G
2173860	межген.				A	C	4290135	mmpL8	Rv3823c	Leu	Leu	C	T
2177366	fadD31	Rv1925	Arg	Cys	C	T	4301075	fadD23	Rv3826	Glu	Gln	G	C
2179825	-	Rv1927	Glu	Asp	G	C	4308395	serS	Rv3834c	Leu	Leu	G	A
2187587	-	Rv1936	Gln	His	G	C	4311664	-	Rv3837c	Gly	Asp	C	T
2191498	ephB	Rv1938	Gly	Trp	G	T	4313128	-	Rv3839	Pro	Ser	C	T
2199052	-	Rv1948c	Gly	Arg	C	G	4314800	ncrMT3949	RVnc0044	Val	Val	G	A
2202500	higA	Rv1956	His	His	C	T	4314800	glpQ1	Rv3842c	His	Tyr	G	A
2209465	mce3A	Rv1966	Ala	Thr	G	A	4319985	межген.				G	C
2221584	-	Rv1978	Ser	Thr	G	C	4329782	межген.				G	A
2225365	межген.				T	C	4340028	межген.				G	C
2230045	-	Rv1986	Cys	Tyr	G	A	4350305	eccCb1	Rv3871	Pro	Pro	G	A
2235087	ctpG	Rv1992c	Phe	Phe	G	A	4359653	espK	Rv3879c	Asp	Asn	C	T
2244421	-	Rv1999c	Leu	Met	G	T	4359828	межген.				G	A
2245864	-	Rv2000	Pro	Gln	C	A	4368084	eccE2	Rv3885c	Val	Val	G	C
2251605	-	Rv2005c	Ser	Ser	G	A	4372353	-	Rv3888c	Arg	Arg	G	C
2260100	межген.				C	T	4378504	eccC2	Rv3894c	Asp	Gly	T	C
2262620	-	Rv2015c	Glu	Asp	T	A	4384007	межген.				C	G
2262857	-	Rv2015c	Ser	Ser	T	C	4393590	mutT4	Rv3908	Arg	Gly	C	G
2267015	-	Rv2021c	Ala	Ala	T	C	4397736	-	Rv3910	Ile	Ile	C	T
2268433	межген.				C	A	4407588	gid	Rv3919c	Ala	Ala	T	C
2268627	межген.				G	C	4407927	gid	Rv3919c	Glu	Asp	T	G
2269545	-	Rv2024c	Ser	Ser	G	A	4408923	межген.				C	T

Приложение 4. Специфические полиморфизмы для представителей генетического семейства LAM. Серым цветом отмечены полиморфизмы, исключенные из анализа в ходе процесса валидации.

Позиция	Ген	№ локуса	Ам-та реф	Ам-та мут	Нукл. реф	Нукл. мут.	Позиция	Ген	№ локуса	Ам-та реф	Ам-та мут	Нукл. реф	Нукл. мут.
8040	gyrA	Rv0006	Gly	Ser	G	A	2231132	-	Rv1987	Ser	Asn	G	A
14251	-	Rv0012	Asp	Asn	G	A	2245532	-	Rv2000	Cys	Cys	T	C
17608	pknA	Rv0015c	Ser	Arg	G	C	2267372	-	Rv2022c	Val	Ala	A	G
27463	межген.				C	G	2270546	межген.				G	A
33457	-	Rv0030	His	His	C	T	2271187	-	Rv2025c	Asp	Asn	C	T
33551	-	Rv0030	Ter	Gly	T	G	2283030	-	Rv2037c	Ile	Thr	A	G
34527	bioF2	Rv0032	Cys	Ser	G	C	2306306	pks12	Rv2048c	Ala	Ala	A	G
37763	fadD34	Rv0035	Asp	His	G	C	2306472	pks12	Rv2048c	Tyr	Phe	T	A
49080	-	Rv0045c	Leu	Pro	A	G	2306684	pks12	Rv2048c	Asp	Asp	A	G
63771	-	Rv0059	Pro	Leu	C	T	2306690	pks12	Rv2048c	Phe	Phe	G	A
74059	icd2	Rv0066c	Lys	Lys	C	T	2306696	pks12	Rv2048c	Gly	Gly	C	G
77058	sdaA	Rv0069c	Ile	Val	T	C	2306699	pks12	Rv2048c	Thr	Thr	A	G
133862	межген.				G	A	2306707	pks12	Rv2048c	Ala	Thr	C	T
137273	межген.				G	C	2311233	-	Rv2052c	Ala	Thr	C	T
157292	fbpC	Rv0129c	Glu	Glu	C	T	2315669	-	Rv2059	Val	Ile	G	A
162581	ephF	Rv0134	Gly	Ser	G	A	2315748	-	Rv2059	Asp	Ala	A	C
178941	PE1	Rv0151c	Asn	Lys	G	T	2347442	-	Rv2090	Asp	His	G	C
180025	PE2	Rv0152c	Gly	Glu	C	T	2437971	pknL	Rv2176	Glu	Lys	G	A
196522	fadD5	Rv0166	Asp	Asp	C	T	2439401	-	Rv2177c	Tyr	His	A	G
197289	yrbE1A	Rv0167	His	His	C	T	2439458	-	Rv2177c	Ala	Thr	C	T
203269	mce1D	Rv0172	Ala	Val	C	T	2439460	-	Rv2177c	Ile	Thr	A	G
207226	-	Rv0175	Met	Thr	T	C	2453879	-	Rv2191	Asp	Asn	G	A
212353	-	Rv0181c	Arg	His	C	T	2463094	межген.				G	A
239059	mmpL11	Rv0202c	Arg	Gln	C	T	2484255	-	Rv2216	Ala	Ala	T	G
272267	-	Rv0227c	Gly	Gly	A	G	2502073	-	Rv2228c	Arg	Arg	C	T
275009	php	Rv0230c	Phe	Val	A	C	2503625	-	Rv2230c	Leu	Leu	C	T
275634	php	Rv0230c	Leu	Leu	G	A	2518919	kasA	Rv2245	Gly	Ser	G	A
286415	aftD	Rv0236c	Met	Lys	A	T	2569934	-	Rv2298	Asp	Tyr	G	T
294841	fadE5	Rv0244c	Ala	Pro	C	G	2621058	dnaG	Rv2343c	Pro	Pro	G	A
327897	fadE6	Rv0271c	Val	Ile	C	T	2630946	plcA	Rv2351c	Gly	Ala	C	G
329826	-	Rv0273c	Cys	Tyr	C	T	2654371	hrcA	Rv2374c	Asn	Asn	G	A
348304	eccC3	Rv0284	Arg	Arg	C	T	2736434	nadE	Rv2438c	Arg	Leu	C	A

387353	pcp	Rv0319	Gly	Asp	G	A	2745889	-	Rv2446c	Ala	Thr	C	T
392479	-	Rv0326	Glu	Asp	G	C	2746340	folC	Rv2447c	Ala	Val	G	A
403364	-	Rv0338c	Pro	Pro	G	A	2775938	-	Rv2473	Gln	Pro	A	C
403920	-	Rv0338c	Arg	His	C	T	2814961	accA1	Rv2501c	Gly	Gly	G	A
408723	-	Rv0340	Ile	Ile	C	T	2871048	-	Rv2551c	Leu	Leu	C	T
412339	iniA	Rv0342	Ser	Trp	C	G	2881455	-	Rv2561	Tyr	Cys	A	G
444351	-	Rv0366c	Asn	Lys	G	T	2893238	-	Rv2568c	Arg	Arg	C	A
450640	-	Rv0373c	Ala	Ala	C	T	2945004	межген.				C	G
472705	межген.				T	C	2945518	-	Rv2616	Arg	Arg	G	T
479632	fadE7	Rv0400c	Thr	Asn	G	T	2960231	PE_PGRS46	Rv2634c	Gly	Gly	G	T
488037	pks6	Rv0405	Glu	Asp	G	T	2968913	межген.				T	C
501535	thiO	Rv0415	Val	Phe	G	T	2979141	-	Rv2657c	Arg	Gly	G	C
534691	-	Rv0446c	Trp	Ter	C	T	3004091	-	Rv2687c	Arg	Trp	G	A
535695	ufaA1	Rv0447c	Gly	Gly	C	T	3029610	-	Rv2716	Ala	Thr	G	A
558501	icl1	Rv0467	Asp	Asp	C	T	3030594	nrdR	Rv2718c	Ala	Val	G	A
563420	-	Rv0472c	Glu	Gly	T	C	3061615	ftsK	Rv2748c	Met	Val	T	C
584017	-	Rv0493c	Ile	Thr	A	G	3073868	thyA	Rv2764c	Thr	Ala	T	C
584438	-	Rv0493c	Pro	Ser	G	A	3077039	PPE43	Rv2768c	Gly	Val	C	A
589536	-	Rv0499	Leu	Leu	G	A	3101119	-	Rv2791c	Arg	Ser	G	T
600116	-	Rv0508	Glu	Lys	G	A	3102800	truB	Rv2793c	Glu	Ala	T	G
637922	-	Rv0546c	Gln	Gln	C	T	3108055	-	Rv2799	Ala	Ala	C	T
657895	-	Rv0566c	Val	Ile	C	T	3108674	-	Rv2800	Arg	Arg	A	C
664249	-	Rv0571c	Gln	Gln	C	T	3165074	-	Rv2854	Val	Ala	T	C
675673	PE_PGRS7	Rv0578c	Ala	Thr	C	T	3175702	-	Rv2864c	Ile	Val	T	C
709226	-	Rv0613c	Thr	Ile	G	A	3178445	межген.				C	G
726703	recC	Rv0631c	Arg	Met	C	A	3191027	cdsA	Rv2881c	Leu	Leu	G	A
736710	mmaA4	Rv0642c	Asn	Ser	T	C	3218343	dacB2	Rv2911	Arg	Gln	G	A
752343	mkl	Rv0655	Glu	Gly	A	G	3236230	smc	Rv2922c	Arg	Leu	C	A
752802	межген.				A	C	3243630	межген.				G	A
764995	rpoC	Rv0668	Ala	Ala	C	G	3260301	ppsC	Rv2933	Gly	Gly	A	G
771273	lpqP	Rv0671	Asp	Gly	A	G	3261380	ppsC	Rv2933	Asn	Ser	A	G
773809	echA4	Rv0673	Glu	Glu	G	A	3314629	-	Rv2962c	Trp	Ter	C	T
796509	-	Rv0696	Gly	Cys	G	T	3320128	pca	Rv2967c	His	His	G	A
812808	rpsN1	Rv0717	Trp	Ter	G	A	3331535	-	Rv2975c	Leu	Leu	G	A
862987	-	Rv0769	Arg	Arg	C	T	3351926	-	Rv2994	Ser	Pro	T	C
863805	-	Rv0770	Phe	Leu	T	C	3354896	serA1	Rv2996c	Glu	Glu	C	T
868174	-	Rv0774c	Asp	Asn	C	T	3400627	-	Rv3040c	Thr	Thr	C	G
885542	-	Rv0791c	Ser	Cys	G	C	3418328	-	Rv3057c	Asp	Ala	T	G

919393	-	Rv0825c	Thr	Arg	G	C	3418330	-	Rv3057c	His	His	G	A
942479	межген.				T	C	3423184	-	Rv3060c	Thr	Thr	C	T
991515	-	Rv0890c	Gly	Asp	C	T	3426795	ligB	Rv3062	Ser	Ser	C	G
1040050	pstS2	Rv0932c	Ala	Ala	C	T	3429202	cstA	Rv3063	Tyr	Asp	T	G
1057788	-	Rv0948c	Lys	Thr	T	G	3443144	pknK	Rv3080c	Ile	Val	T	C
1073033	-	Rv0959	Gly	Gly	T	G	3454986	tgS4	Rv3088	Ala	Glu	C	A
1087279	accD2	Rv0974c	Lys	Glu	T	C	3460419	-	Rv3091	Tyr	Cys	A	G
1093322	межген.				C	T	3460986	-	Rv3092c	Pro	Leu	G	A
1099691	repD	Rv0983	Ser	Thr	G	C	3542262	-	Rv3174	Leu	Arg	T	G
1101205	mscL	Rv0985c	Leu	Leu	C	T	3548641	-	Rv3179	Tyr	His	T	C
1103249	-	Rv0987	Ala	Ala	C	T	3564400	-	Rv3195	Gly	Ser	G	A
1107434	межген.				A	T	3572800	-	Rv3200c	Asp	Asp	A	G
1107917	-	Rv0990c	His	Gln	G	T	3594124	-	Rv3217c	Ala	Thr	C	T
1142266	mfD	Rv1020	Leu	Leu	A	C	3601494	-	Rv3225c	Pro	Leu	G	A
1144664	eno	Rv1023	Arg	Gln	G	A	3610441	tgS3	Rv3234c	Arg	His	C	T
1159119	-	Rv1034c	Gln	Gln	C	T	3621770	-	Rv3242c	Arg	Trp	G	A
1159162	-	Rv1034c	Gly	Ala	C	G	3628148	межген.				G	C
1161721	PPE15	Rv1039c	Val	Asp	A	T	3643392	-	Rv3263	Gly	Gly	C	T
1162274	PPE15	Rv1039c	Ala	Thr	C	T	3643630	-	Rv3263	Gly	Ser	G	A
1177994	-	Rv1056	Tyr	His	T	C	3682023	nei	Rv3297	Thr	Met	C	T
1190154	PE_PGRS19	Rv1067c	Glu	Lys	C	T	3730327	PPE54	Rv3343c	Phe	Phe	G	A
1199547	-	Rv1075c	Pro	Leu	G	A	3732656	PPE54	Rv3343c	Ala	Gly	G	C
1248059	ephC	Rv1124	Phe	Phe	C	T	3732658	PPE54	Rv3343c	Leu	Phe	C	G
1251033	ppdK	Rv1127c	Gly	Gly	G	A	3732660	PPE54	Rv3343c	Leu	Leu	A	G
1272010	межген.				C	T	3732664	PPE54	Rv3343c	Val	Val	G	A
1278583	межген.				A	C	3734518	PPE54	Rv3343c	Phe	Leu	G	T
1303095	fbiC	Rv1173	Leu	Leu	G	A	3776706	-	Rv3365c	Ala	Thr	C	T
1324724	-	Rv1184c	Met	Ile	C	A	3794884	межген.				G	A
1354437	-	Rv1211	Ala	Ala	A	C	3825560	-	Rv3406	Leu	Leu	C	G
1389738	-	Rv1248c	Asp	Asp	G	A	3847684	PPE59	Rv3429	Gly	Arg	G	A
1404169	cyp130	Rv1256c	Lys	Asn	T	G	3847684	-	Rv3430c	Pro	Pro	G	A
1428506	-	Rv1278	Ala	Ser	G	T	3850197	межген.				G	A
1442733	межген.				C	T	3859376	-	Rv3439c	Gly	Glu	C	T
1445474	межген.				G	A	3873392	cut3	Rv3451	Leu	Arg	T	G
1446923	argS	Rv1292	Ile	Ser	T	G	3875511	-	Rv3454	Ala	Ala	G	A
1452071	thrB	Rv1296	Gly	Gly	C	A	3889569	-	Rv3472	Arg	Cys	C	T
1480024	-	Rv1318c	Phe	Leu	G	T	3894032	межген.				T	A
1481185	-	Rv1319c	Asp	Glu	A	C	3895103	PPE60	Rv3478	Gly	Gly	G	C

1512031	mbtK	Rv1347c	Asp	Ala	T	G	3895111	PPE60	Rv3478	Thr	Lys	C	A
1519847	-	Rv1353c	Gly	Arg	C	G	3895113	PPE60	Rv3478	Ala	Thr	G	A
1546530	-	Rv1373	Glu	Glu	G	A	3895129	PPE60	Rv3478	Leu	Arg	T	G
1575783	nlhH	Rv1399c	Pro	Pro	A	C	3895137	PPE60	Rv3478	Leu	Ile	C	A
1586249	-	Rv1410c	Gln	Gln	C	T	3895157	PPE60	Rv3478	Ile	Met	A	G
1613960	gap	Rv1436	Ala	Ala	G	T	3895257	PPE60	Rv3478	Glu	Gln	G	C
1651306	-	Rv1463	Ala	Ala	C	T	3895691	межген.				C	G
1688024	lipL	Rv1497	Ala	Ala	G	C	3903530	cpsA	Rv3484	Ala	Ala	G	C
1692795	межген.				G	C	3913455	lprN	Rv3495c	Ile	Ile	A	G
1718761	-	Rv1524	Gly	Gly	C	T	3926151	fadD17	Rv3506	Pro	Gln	C	A
1734994	-	Rv1534	Ala	Ala	C	T	3936761	межген.				A	G
1752992	PPE21	Rv1548c	Arg	Arg	T	C	3938168	ilvX	Rv3509c	Ala	Val	G	A
1755519	plsB1	Rv1551	Met	Ile	G	T	3973954	hsaG	Rv3535c	Gly	Arg	C	T
1762615	mmpL6	Rv1557	Ala	Val	C	T	3983271	fadE28	Rv3544c	Ile	Leu	T	G
1770906	treX	Rv1564c	Leu	Leu	C	A	4022129	-	Rv3579c	Leu	Leu	G	A
1775312	межген.				A	C	4026593	radA	Rv3585	Val	Val	G	A
1777213	bioF1	Rv1569	Ala	Gly	C	G	4028931	-	Rv3586	Leu	Leu	G	A
1779243	-	Rv1572c	His	Pro	T	G	4038287	clpC1	Rv3596c	Asn	Asn	G	A
1793769	межген.				C	T	4038968	clpC1	Rv3596c	Gln	Gln	C	T
1797027	nadB	Rv1595	Thr	Ile	C	T	4057146	межген.				C	T
1805948	hisI	Rv1606	Thr	Ile	C	T	4060334	esxW	Rv3620c	Glu	Glu	T	C
1815604	-	Rv1615	Ala	Thr	G	A	4075957	-	Rv3636	Ala	Asp	C	A
1829645	-	Rv1627c	Gly	Gly	C	G	4088346	cspA	Rv3648c	Thr	Thr	G	A
1831340	polA	Rv1629	Arg	Trp	C	T	4096985	-	Rv3659c	Asp	Ala	T	G
1834836	rpsA	Rv1630	Met	Thr	T	C	4135112	-	Rv3693	Met	Ile	G	A
1894878	-	Rv1668c	Ile	Ile	G	A	4136581	-	Rv3694c	Val	Val	C	T
1895174	-	Rv1668c	Asp	Asn	C	T	4156467	dnaQ	Rv3711c	Val	Ala	A	G
1899351	-	Rv1674c	Glu	Gly	T	C	4163944	-	Rv3720	His	Arg	A	G
1943592	-	Rv1716	Pro	Leu	C	T	4174564	-	Rv3727	Ile	Thr	T	C
1947903	-	Rv1722	Val	Leu	G	T	4209511	lpqH	Rv3763	Glu	Asp	G	C
1951800	межген.				T	G	4220174	lipE	Rv3775	Asp	Asn	G	A
1981056	межген.				C	T	4233299	-	Rv3786c	Thr	Ile	G	A
2007545	-	Rv1773c	Ala	Ala	T	G	4264218	-	Rv3802c	Val	Phe	C	A
2026008	PPE25	Rv1787	Phe	Leu	C	G	4293072	papA1	Rv3824c	Leu	Phe	G	A
2026025	PPE25	Rv1787	Gln	Arg	A	G	4299480	pks2	Rv3825c	Ser	Ser	G	A
2037716	eccA5	Rv1798	Arg	Arg	T	C	4315384	glpQ1	Rv3842c	Asp	Gly	T	C
2049645	межген.				C	G	4316046	-	Rv3843c	Val	Ala	A	G
2053411	межген.				G	C	4320050	межген.				A	G

2053682	mgtC	Rv1811	Ile	Ile	C	T	4325029	межген.				T	C
2071576	-	Rv1825	Pro	Ser	C	T	4340330	espE	Rv3864	Leu	Val	T	G
2077253	gcvB	Rv1832	Thr	Thr	G	A	4363069	eccE1	Rv3882c	Ile	Val	T	C
2097990	ureC	Rv1850	Ala	Ala	A	C	4371779	-	Rv3888c	Val	Leu	C	A
2108838	межген.				C	T	4373496	espG2	Rv3889c	Val	Val	C	G
2111513	-	Rv1864c	Glu	Glu	C	T	4376098	межген.				G	A
2123181	-	Rv1873	Ser	Leu	C	T	4385530	-	Rv3900c	Pro	Arg	G	C
2128908	glnA3	Rv1878	Ser	Phe	C	T	4388606	-	Rv3903c	Thr	Thr	G	A
2144617	-	Rv1897c	Pro	Leu	G	A	4390753	esxF	Rv3905c	Ser	Ser	C	G
2156196	furA	Rv1909c	Leu	Leu	G	A	4391553	-	Rv3906c	Pro	Pro	C	T
2163444	PPE34	Rv1917c	Asn	Asp	T	C	4394450	-	Rv3909	Val	Leu	G	C
2207525	межген.				C	T	4397368	-	Rv3910	Met	Leu	A	C
2216248	mce3F	Rv1971	Pro	Arg	C	G	4408156	gid	Rv3919c	Leu	Arg	A	C
2231100	-	Rv1987	Ala	Ala	C	T	4376098	межген.				G	A

Приложение 5. Специфические полиморфизмы для представителей генетического семейства Ural. Серым цветом отмечены полиморфизмы, исключенные из анализа в ходе процесса валидации.

Позиция	Ген	№ локуса	Ам-та реф	Ам-та мут	Нукл. реф	Нукл. мут.	Позиция	Ген	№ локуса	Ам-та реф	Ам-та мут	Нукл. реф	Нукл. мут.
9841	межген.				C	T	2147906	cinA	Rv1901	Thr	Ile	C	T
15517	trpG	Rv0013	Arg	Gly	C	G	2169879	PPE35	Rv1918c	Phe	Cys	A	C
43722	leuS	Rv0041	Pro	Leu	C	T	2173033	lppF	Rv1921c	Trp	Arg	A	G
101092	ctpA	Rv0092	Val	Val	C	A	2181026	-	Rv1928c	Pro	Pro	G	C
202675	mce1D	Rv0172	Ile	Thr	T	C	2185342	fadE17	Rv1934c	Asp	Asn	C	T
217306	bglS	Rv0186	Val	Val	C	G	2186785	echA13	Rv1935c	Gln	Gln	C	T
295644	межген.				C	T	2211600	mce3B	Rv1967	Gln	Glu	C	G
306639	cobQ1	Rv0255c	Glu	Gly	T	C	2216035	mce3F	Rv1971	Glu	Gly	A	G
314338	aac	Rv0262c	Cys	Arg	A	G	2225237	межген.				T	C
317540	oplA	Rv0266c	Ala	Ala	G	C	2225238	межген.				A	C
328464	межген.				C	T	2237059	ctpG	Rv1992c	Thr	Asn	G	T
355803	mycP3	Rv0291	Ala	Ser	G	T	2251179	-	Rv2005c	Pro	Pro	C	T
367824	PPE5	Rv0304c	Ser	Ser	A	G	2280734	-	Rv2033c	His	Tyr	G	A
368079	PPE5	Rv0304c	Thr	Thr	A	G	2300664	pks12	Rv2048c	Pro	Leu	G	A
378407	-	Rv0309	Gln	His	G	T	2316510	-	Rv2060	Val	Ile	G	A
407877	-	Rv0339c	Glu	Gly	T	C	2316510	-	Rv2059	Gly	Asp	G	A
422879	dnaJ1	Rv0352	Arg	His	G	A	2361492	helZ	Rv2101	Leu	Arg	T	G
423722	hspR	Rv0353	Thr	Thr	C	T	2396770	-	Rv2136c	Ser	Ser	G	C
433515	PPE8	Rv0355c	Leu	Leu	G	A	2411730	murC	Rv2152c	Ser	Ser	G	C
447411	-	Rv0370c	Arg	Arg	G	T	2441970	-	Rv2179c	Phe	Phe	G	A
553613	lpdC	Rv0462	Ala	Ser	G	T	2480676	ephD	Rv2214c	Leu	Leu	G	A
585891	-	Rv0495c	Glu	Gln	C	G	2540441	cyp124	Rv2266	Leu	Pro	T	C
591505	-	Rv0500B	Arg	Ser	C	A	2559151	-	Rv2286c	Gly	Asp	C	T
609254	-	Rv0517	Pro	Leu	C	T	2625052	-	Rv2345	Arg	Gln	G	A
625536	межген.				A	G	2631226	plcA	Rv2351c	Asp	Tyr	C	A
738902	mmaA2	Rv0644c	Asp	Asn	C	T	2748087	valS	Rv2448c	Ser	Ser	G	A
758322	atsD	Rv0663	Asp	Gly	A	G	2793706	lipQ	Rv2485c	Leu	Phe	G	A
783601	fusA1	Rv0684	Arg	Arg	A	C	2851736	межген.				C	G
800204	межген.				T	C	2930057	межген.				G	C
849486	PPE12	Rv0755c	Asn	Lys	G	C	2954571	-	Rv2627c	Leu	Val	A	C
864468	-	Rv0771	Glu	Gly	A	G	2988815	-	Rv2672	Asn	Lys	C	G
893895	pepC	Rv0800	Ala	Val	C	T	3021962	-	Rv2709	Glu	Lys	G	A

954906	dapC	Rv0858c	Arg	Arg	C	T	3056663	-	Rv2743c	Cys	Cys	G	A
1017836	-	Rv0913c	Phe	Cys	A	C	3086742	-	Rv2779c	Phe	Val	A	C
1057990	межген.				G	A	3103461	pptT	Rv2794c	Tyr	Tyr	G	A
1105587	-	Rv0988	Trp	Arg	T	C	3104336	-	Rv2795c	Glu	Glu	T	C
1164571	межген.				A	G	3124014	-	Rv2817c	Ser	Ala	A	C
1237536	-	Rv1111c	Pro	Pro	C	G	3186664	dipZ	Rv2874	Ala	Ala	A	C
1375446	-	Rv1232c	Leu	Leu	C	T	3198496	tsf	Rv2889c	Ile	Ile	G	A
1409766	-	Rv1261c	Thr	Thr	C	T	3218122	-	Rv2910c	Ile	Ser	A	C
1441794	-	Rv1288	Val	Val	A	G	3233940	ftsY	Rv2921c	Ala	Gly	G	C
1442141	-	Rv1288	Thr	Arg	C	G	3257233	ppsC	Rv2933	Asp	Asn	G	A
1466779	atpD	Rv1310	Ile	Ile	C	T	3257531	ppsC	Rv2933	Arg	His	G	A
1568018	metK	Rv1392	Asp	Asp	C	T	3290870	lppX	Rv2945c	Ser	Ser	G	C
1605170	-	Rv1429	Ala	Val	C	T	3309275	-	Rv2956	Pro	Arg	C	G
1622092	межген.				C	T	3391074	-	Rv3031	Val	Leu	G	C
1626857	zwf2	Rv1447c	Val	Ile	C	T	3406045	fecB	Rv3044	Ala	Thr	G	A
1634806	межген.				C	T	3422687	-	Rv3060c	Thr	Ile	G	A
1645173	-	Rv1459c	Trp	Cys	C	A	3442970	pknK	Rv3080c	Ala	Thr	C	T
1647444	-	Rv1461	His	His	C	T	3447800	virS	Rv3082c	Asp	Glu	G	T
1670814	-	Rv1480	Gly	Gly	C	T	3469694	smpB	Rv3100c	Ala	Ala	G	T
1710354	-	Rv1518	Thr	Thr	C	G	3471151	ftsE	Rv3102c	Arg	Arg	G	A
1719823	-	Rv1524	Lys	Lys	A	G	3541237	-	Rv3172c	Lys	Ile	T	A
1728022	pks5	Rv1527c	Gly	Arg	C	G	3574504	-	Rv3201c	Arg	Trp	G	A
1732204	adh	Rv1530	Thr	Ala	A	G	3585577	-	Rv3208	Cys	Arg	T	C
1747327	-	Rv1546	Leu	Leu	C	T	3589537	-	Rv3212	Pro	Pro	C	T
1793790	межген.				A	C	3646964	rmlD	Rv3266c	Leu	Leu	C	G
1808795	trpE	Rv1609	Asp	Ala	A	C	3666905	accA3	Rv3285	Thr	Thr	C	T
1849381	lysX	Rv1640c	Ile	Ile	G	A	3867404	eccC4	Rv3447c	Glu	Ter	C	A
1860528	pheT	Rv1650	Cys	Cys	T	C	3868379	eccD4	Rv3448	Arg	Trp	C	T
1862521	PE_PGRS30	Rv1651c	Thr	Thr	C	T	3875230	-	Rv3454	Thr	Ala	A	G
1867614	argJ	Rv1653	Ala	Ala	C	T	3917305	mce4B	Rv3498c	Val	Phe	C	A
1867838	argJ	Rv1653	Ser	Asn	G	A	3999707	-	Rv3559c	Ala	Ala	G	A
1872211	argG	Rv1658	Ala	Ala	G	A	4095320	межген.				G	A
2008217	-	Rv1774	Val	Gly	T	G	4129917	-	Rv3688c	Ser	Ser	C	A
2053454	mgtC	Rv1811	Leu	Leu	G	A	4168771	-	Rv3723	Ser	Leu	C	T
2057591	-	Rv1815	Pro	Ser	C	T	4280072	-	Rv3815c	Ala	Ala	C	A
2059831	-	Rv1817	Gly	Gly	T	G	4309617	-	Rv3835	Ala	Ser	G	T
2063685	bacA	Rv1819c	Leu	Leu	C	T	4310449	-	Rv3836	Ala	Thr	G	A
2063911	bacA	Rv1819c	Ile	Thr	A	G	4355141	eccD1	Rv3877	Glu	Asp	A	C

2066470	secA2	Rv1821	Gly	Asp	G	A	4364826	межген.				G	C
2100906	ureG	Rv1852	Asp	Asp	C	T	4376445	eccC2	Rv3894c	Pro	Pro	G	A
2123182	-	Rv1873	Ser	Ser	A	C	4390391	-	Rv3903c	Ala	Ala	C	G

Приложение 6. Праймеры, используемые для определения полногеномной последовательностей штамма CTRI-2 и SP28

название праймера	5'-3' последовательность	название праймера	5'-3' последовательность
TB2-f1	GCGAATCACTTTGACGTGC	TB52-f1	GCAGACCTTATCGAGCCCG
TB2-r1	CGGGAGGAGCTTGTCTGAGG	TB52-f2	TGGGCGCGGAGGTCATCG
TB2-f2	ACGATATTGAGCGGCTTTGC	TB52-r	TGTTTGGGCACGTTGTCTGG
TB2-r2	GCGCTCTTTGTCTGAGTCCC	TB53-f1	CTTCCAAGAACACGACCCCG
TB2-r1n	ACCTTCCCCTATCAACTCG	TB53-f2	TAGGTCGCGCCCATCACC
TB2-r2n	TGTTAGTTCGGCCATAACG	TB53-r1	GTTGCTGGTGTTTTTTCGGGC
TB4-f1	CTCACTCAAGGTCCATTCGG	TB53-r2	CCTGGGAGATGTCTGTTTCTG
TB4-r1	GCCAGCACCACCACCAACG	TB54-f1	GGAACAGCTTATGGGCAACG
TB4-f2	CCGTGCTCCTACACCCCG	TB54-f2	GAGCTGACACCGCTGGACCC
TB4-r2	GTGAAGTAGGCGACCACCG	TB54-R2	GTGTTGGTGCCTTGCAGCG
TB5-f1	GCAGCACGGCGTTGGTCTG	TB54-R1	ATGCACACTCGAATCGGC
TB5-r1	GCGGCGGTGTTCTCTACG	TB55-f1	GAAAAGCCCGAGGATAGCC
TB5-f1	GTCGGACATCACATGGCGG	TB55-f2	GGAATTGAAGAAGCCCGACG
TB5-r2	TGCTGCTGGGCGGATGCG	TB55-r2	AACGTGATGGCTTCACCC
TB8-f1	TTCCAGACCCTCGTCGAGC	TB55-r1	GTTCAATTCTTCTGCCCGCC
TB8-f2	CGCCTTGCTGGATGCTGG	TB56-f2	CGGAATCTGCAATGGCGGC
TB8-f3	CAATCCCACCGCTACCGC	TB56-f1	CAGGTAGACTAAAGCCACCG
TB8-f4	CGACGACTCCGCCGAAGC	TB56-r2	AATGGGCTGTTCAACACGGG
TB8-f5	ACCAGGTGTGTCTTGAAGG	TB56-r1	GTCGTCCAACACCGGCACG
TB8-f6	CGAAGAGCGAGCCGCCAC	TB57-f1	GTGTTCCGCCGCGCTACG
TB8-f7	ACCGTTACCGCCATTGCCG	TB57-f2	ATTGCCGCCGTTGCCGCC
TB8-f8	GCCGACGTTGATTATGCTCG	TB57-r1	GCTGTTACCACCGGCGG
TB8-r8	AACGGATTTAGCCAGTCTCG	TB57-r2	GGACGGGCGGATTCGGGG
TB8-r7	GTGGAGGGACGAGCATAATC	TB57-F3	AGTGGTAACGCTGCGGTG
TB8-r6	TCAACGCGATCAACGAGCC	TB57-r3	AGGCGGGATCCTGTTTGGC
TB8-r5	CCAGCGGAATCCTGTCAGG	TB58-f1	GGCGGGATTCTGTACGGC
TB8-r3	ATTCACGAACGGCAGTGCAC	TB58-f2	AATGGCGGGAATGCTGGGC
TB8-r2	CAAAATGTGCTCAATGTGGTG	TB58-f3	TACTGGCGCGACGGTGGG
TB8-r1	GGAACGGCGGACTGTTCGG	TB58-r1	ATGTGCAGATTGCCCATAGG
TB11-f1	GTCTTGCGCTCCAGGAGTC	TB58-r2	TCCCATTTTGGCCGAATAGC
TB11-f2	GCCAGCATCAGATCGGCG	TB58-r3	ACCCATGCCACCCGTGCC
TB11-f3	GCCTCCACGAGTGGGGGC	TB59-f1	CGATTTCCACAAGCAGCCG
TB11-r1	CACGGTATCTACAGACTCGG	TB59-f2	TATGCCGGATTCACTCTGGG
TB11-r2	AAAATGGCCCCACCGTTGC	TB59-f3	ACCGTCACCGCCAACACC
TB12-f1	ACCAAACCTCGGCTTACCTC	TB59-r1	AGGGTGTGACGAGGTGCC
TB12-r1	TGACGGTGATCTCCCTGAG	TB59-r3	AGCCGCGCTGCGACTGTG
TB12-f2	GTGCGACTGGTTGGCGAGG	TB59-f4	CCGTTCCCGACCAATCCG
TB12-r2	TTGCCAGACGAATCCGACC	TB59-r2	TAGAGTGTGTGCTCGACGTG
TB12-r2n	TGCCAGACGAATCCGACC	TB60-f1	CGCCGACTCCAGCAGACG
TB12-f2n	TGCGACTGGTTGGCGAGG	TB60-f2	TCGGACACTGTCTGCGGGT
TB13-f1	ACGACAAGACCACTCGGGG	TB60-f3	GGTTTGCACGCTGTTGCCG

TB13-f2	CGTCTGCTGCTGGTGGTG	TB60-r1	TGAGGAGCGCGTCATGTGC
TB13-r1	TGTGCTGAATGCCCAACCC	TB60-r2	GTTTCACGAGCAGTTTGTGC
TB13-f3	CGTGGAGTAGCAGTTGCTG	TB60-r3	ACCCTGCAAGGTTTCAGCG
TB13-r2	ACCAGTGGCAACCTTGATTG	TB60-f4	CGTTGCCATCGAAGAAGCTG
TB13-f4	GCGACGGCAGCAGCACCC	TB61-f1	CGATCATCGACACATTGTCC
TB13-r3	GAATGCCGAAGGGTTGAGC	TB61-f2	CCCTTGTTGTGTTGGACCGB
TB13-r4	CGCGATGGCCGAATTACTG	TB61-r1	ATCGGTATTCATGCGGGCG
TB14-f1	CGAATCATTGGCACGTCTAC	TB61-r2	CCGAGCGTGGTGAGCAGG
TB14-f2	GGTTCGCCGTGTTGTCCG	TB62-f1	AAATGGCGACGTAGGGGTG
TB14-r1	GTTTCATGATCGGTTCTGTC	TB62-f2	ACCGAGATGTTTCGACGACG
TB14-r2	TCAACTGTGAACAGAGTTTGC	TB62-r	CATCGACTTGAAGACCCG
TB14-r3	CCACGGCACTCTCTTCAGC	TB63-f1	TGAGCCTCAATCGGATCCC
TB14-r4	GACCGGCTCAAGCAGCAGC	TB63-f2	ACGACCCATCCTGATTAACC
TB14-f3	CTGTCTCGCCAGGGTGG	TB63-r1	GTTAGGCGAGGTTAGTAATGG
TB14-f4	CGCTGAGGGCATGGTCGC	TB63-r2	ACGCCAGGGACGTGATGTC
TB14-r5	GGTGGTGATGGCGGTTTGG	TB64_f1	CGATGGATGCGGTGATCGC
TB14-r6	GCTACGGCGGTACAGTGG	TB64_f2	CATGCGTCGGTTCACCTCA
TB14-f5	TCGTGCTGGAACCGTTGAC	TB64_r1	CTGGGGTTGAAGCATTGGC
TB14-f6	GAGCGGACGCATTGAAGCC	TB64_r2	TGCTGGACGAGGTGGGAG
TB14-r7	GCCGGTGTACGGGTACG	TB65_f1	GGTTGCTCAGGGCTTTCAG
TB15-f1	ATACAGCCAGTCCACCCGC	TB65_r1	GCCCCGCTGCATGACAAG
TB15-r1	ACCGGCTCCGACTCGTTGG	TB66-f1	CCAGGACACCACACGGCG
TB15-f2	GACAGTGCCTGCTGGGATG	TB66-f2	GACCGTGAAATGTGCAGGTG
TB15-r2	CCTCACGCACGAAGACGAC	TB66-r1	CGCCGACGGAGCCAGTGG
TB16-f1	ATCGGCAGCAACGCTCGGG	TB66-r2	TGTGGGAGACCTCGGTGC
TB16-r1	CGCAGTTCCCTCGAAAAGCC	TB66-f3	CCGACGCCACCGCTGCTG
TB16-f2	GACGCCAGCAGCTCACCC	TB66-r3	GCGTCTGTTCCAGGTAGG
TB16-r2	CGTCACCGACGAGCAGGC	TB67-f1	GCCGAAGATCACGAGCTGG
TB17-f1	CCACTCGCTGCCAGGCTC	TB67-r1	CGCAGGAATACCAACAGATC
TB17-r1	TGGACGGTGGGCGAGCGG	TB67-r2	CAGTGCCTTTGGGGGTGG
TB17-f2	CGTCTCGGCGTGCCTGGGEXT	TB69_f1	TAGGTGGTGGTGGTCTTGC
TB17-r2	GCTGGACTACACGCTGCC	TB69_r1	GTAGGGTGTTCATCGCTGGC
TB21-f1	GCAGATCAAGATCGCCGTC	TB69_r2	CGACACAGGTGGAGGCTG
TB21-r1	GGCGAAAGTCGATCATCCG	TB71_f1	GGTATTTCGACATCCGAGGC
TB21-f2	GGAGAAATCCTATGGGTTGC	TB71_f2	CGGCAATCATCCCCTTCCC
TB21-r2	CGACCTTGGCGTTGACACC	TB71_r1	TGTGGGTGCGCAGACCTAC
TB22-f1	CAAGGCGATCCAGTCGGCG	TB71_r2	CCATGCGGGTGTGGTGG
TB22-r1	TGCGATGTTTAACCTTACGG	TB72_f1	AGTAGGCGATGTGCGACTC
TB22-f2	AGCGCAATTAGTGGCCAG	TB72_f2	CCGAGGCCACCCGCATC
TB22-r2	GCGCCGATTCGATCTGCCG	TB72_r1	TCTGGCTCAGCTGGGGTG
TB23-f1	TGCGGTGTACTTGTCATCAG	TB72_r2	GTCGGAGAATCCAGACACC
TB23-f2	TGACTCGGGTGGGAGGTGG	TB73_f1	CAGTCATCCAATCGTGCTG
TB23-f3	GTGATCAGTGGGCTGCTGC	TB73_f2	TCATCGAGTCCAGGGTGTT
TB23-f4	TTGGGAGGCATGCGATGTC	TB73_r1	TAAATAGCTGAGCGTAGACC

TB23-f5	AAACGGAGGGCAGGGAACG	TB73_r2	CTGATCGGCTGCCTGCTC
TB23-f6	CGTGCTGCTGCTCGGTGAG	TB74_f1	CGACAACAGAGATCCTGGC
TB23-r1	AGCACTTCGGACAACGCTC	TB74_f2	CTGCTCAACGCCATCAATGC
TB23-r2	CCCTTGCCGCCGATGTGG	TB74_r1	GTCTACGTGCGTGCCACC
TB23-r3	CGTCCACGACGCTTCTAC	TB74_r2	CCGCCGGAGATGACAATCT
TB23-r4	CCCGTTTTGGCCAGCAGC	TB75_f1	AGCGCGTGTACCAGCAGCA
TB23-r5	CCCAACAACCAGCCACCTG	TB75_f2	CGGTGGACATCGACTTGCG
TB24-f1	AACGGAGGCAACGGTGGC	TB75_r1	GTCGTATCAGGCACTCAAG
TB24-f2	GGCAACGACGGCAGTGGC	TB75_r2	CACCCAGGAACAGATTGCC
TB24-f3	CGGTGGTGGAGCCTTGCTG	TB76_f1	TGCACGCAGTTCGACCACA
TB24-f4	GCGTCGTGCGTAGCGTTGC	TB76_r1	GACGGCCTCTACGACTGC
TB24-f5	GTTGCCTCCGTTCCCACCC	TB77-f1	GCGGTGCCGTGGAAGTCG
TB24-f7	GCCGTTGGAGCCGTTAACC	TB77-r1	GTGGTCGGGGAGGGTGGC
TB24-r3	CCACTGTCGCCAGCAACGC	TB77-f2	TTCGCCGTAACCCCAACC
TB24-r4	GCGTGTGGACTTGTGACGG	TB77-f0	AATTGTGAGTCCCCAGTGC
TB24-r5	GCTGGTGGCGACGCTTCC	TB77-f3	CGAGCCACTTGTGGAACCG
TB24-r2	GGTCTTGGTGGGGCTGCG	TB77-f4	GTCGTCAGGCTCTCCGTC
TB24-r1	TTTTCGCACTCGACTTGAGC	TB77-r3	TTGACGGCTTTGACATTAACG
TB26-f1	CCAGACTCCGGTGCCCCC	TB77-r4	CGGTGACGACGGCAAGGG
TB26-r1	TCACGCTACTCGTACCTGG	TB77-r5	GGAGGCATTGGTATCACCG
TB26-f2	ACCGTCTGGATCAAGGGCG	TB78_f1	CGCACGTCAACGCACCCA
TB26-r2	GGCGGCGGTCACAACGGC	TB78_r1	TGGGTGGCGGCGCTGAC
TB27-f1	CGATCAAGGTTGCCGAGGG	TB78-f2	TCCCTGGCGTGGAAGTGG
TB27-r1	ACCCCGGAGCGCTGATCG	TB78-r2	GGTTGGCCTTCAGCTGGTC
TB27-f2	GAGCACCCCTACACGCCG	TB80-f1	GTTCAGTGATTACGCCACGC
TB27-r2	GCCGGACTCGAATTCCAGC	TB80-f2	CATTTTCATCACAGCTTCGTGC
TB28-f1	AGAACGGTGGGGCGGGCG	TB80-f3	GCCGTTTCGCTGCTTTTGG
TB28-f2	AGAACGGTGGGGCGGGCG	TB80-r1	GGGTTGTTCGAGTAGTGCG
TB28-r1	GCCGATGTTGCCGTTGTGG	TB80-r2	CGACGCCGATACTAAACAGG
TB28-r2	GTTAATTCCGCCGAAGCCG	TB80-r3	CGCCCAGGCTGAACAATCG
TB28_f3	GGTGGTTCGCGGGGTGGG	TB82-f1	ATGACTGTGAGGTAGCTGTG
TB28_r3	AGCCGATGTCGCCCTGCC	TB82-f2	CAGTTCGGGTTCCGGCAG
TB29-f1	CGACGGTGATGGGGGAGC	TB82-r2	GGCACAACACGGGTGTAGG
TB29-f2	CCGATGGTGATGGGGGAGC	TB84_f1	CAACGCCAGCATCGAGAAG
TB29-r1	CGACCAGCACTATCAAGCG	TB84_f2	CCCAACAGCACGAACACCA
TB29-r2	CAGAACGTGCTCGATGCG	TB84_r1	CGTGCCGAGCGCTACCTC
TB29-f1n	GGAAGCTCATACCAAAGAGG	TB84_r2	GAACGCTATGAACGCCTGA
TB29-r1n	CGCCCGCATCTCAACATCG	TB85-f1	TTCAGCAACGAGTCGATGTG
TB31-f1	GTACATGAGGAAGTGGTGC	TB85-f2	TACGGCAGCAGCGGTAACG
TB31-f2	AAGTGGAGGAGTTCGTGCG	TB85-r2	CTCAGCTAGGCAGCAATCC
TB31-r1	CCGGCCTACCCAACATTCC	TB85-r1	TTGGACCATGTGCTGACCG
TB31-r2	GCGTACCCGACCCAATCAC	TB86-f1	CTTCGGTCAGTGACAGTACG
TB32-f1	TGGACTTGGGGTTGCACCG	TB86-f2	ATCAAGTGGGACCAGGATGC
TB32-f2	GTAGTTTGGCTCGCCGCTG	TB86-r1	TGACGGCTGGTGACCAAGG

TB32-r1	ACCAATGTCGAGCAGCAGG	TB86-r2	GCCGCGACTATCGTTGAGG
TB32-r2	CGGTGGGTTGTTGTACGGC	TB87_f1	CGGCACCGACCAGCGACT
TB33-f1	GCTTCGATGATCAATTTGGG	TB87_f2	CGCAGCCGAGACATAACAA
TB33-f2	GGACATAAGCCACTCGTCG	TB87_r1	TGGCGGGTCTGGCTATCG
TB33-r1	GGCAACGGCGGCAACGGG	TB87_r2	GTTGGTTCGTCGTCGTGGTG
TB33-r2	GGCAGCGTTTCATGACCAG	TB88_f1	GACGACCTGAACGTTGTAG
TB35-f2	GGTCAACGTCTCGGCGGC	TB88_f2	GACCGACTGGGTGGCCGT
TB35-f1	GCGTATCAGGCGGTGAGCG	TB88_r1	CGTCGGCTGCGTGAATCGT
TB35-r1	TTCAGGTTTTTCGACATGGC	TB91_f1	CGGGCAGCCTTGACATCC
TB35-r2	CGCAAGCTGTAGTAGACGC	TB91_f2	CGGGCTTGCTCGAACTCAG
TB35-f3	CGTCGGCGGTGTCCGGC	TB91_r1	TGCAGCAGCGTACGGAACA
TB35-r3	CATTGCCGAACAACCACCC	TB91_r2	CACACAGACTCTCGACGAC
TB35-r4	AAGTTGGTCGGCCAGGGC	TB92-f1	GAGCACCGTGTTCTGTCTAC
TB36-f1	TTGGCACGCACGGGCAGG	TB92-f2	GGTGTGTTGTCGAGCAAGCC
TB35-f2	AACGTCTCGGCGGCAG	TB92-r1	CCGATATTTGGTGGTGATCC
TB36-r1	CAAGGACGCCGATCATTCC	TB92-r2	TGGTCCTTGTTGAGGTTGG
TB36-r2	CGAAAACCGTGGTGTACCC	TB93_f1	AGATCACCGAACCGCAAAGA
TB36-f1n	GCAGGCGTTGACTGGCGG	TB93_r1	CACCGCCGACGTAGTTGG
TB36-r3	CGCCGCCATTGGTGCTGC	TB93_r2	GAGGACGCTGGTTACGGTC
TB36-f3	CAGCGGTGGCAGCGGTGG	TB94_f1	GCAACGCCATATCGGTGTC
TB36-r2n	GGATATTGCCATCCCTCGA	TB94_f2	CTAAGATCGGCACTCTACC
TB36-f4	TGGAGGAGGAACCGGTGGG	TB94_r1	CCGTCCGCCACTAGATCC
TB36-r4	CGACACCGAACAACAGCCC	TB94_r2	GCTTCGGGTGTAGATCCTG
TB37-f1	ACTGCGCGATCGTTGAGGG	TB95_f1	GACCGCGAACGTTGCATCC
TB37-f2	CAAGAACGTCGTCGTCGGG	TB95_f2	CCAGCGTCTGCGAGCACT
TB37-r1	TGGCATCGGCGGCAACGG	TB95_r1	GAGCGGCCTTGCGTAACC
TB37-r2	GCGACGGCGTTTCATCAGC	TB95_r2	CGGAAGTCGATGTCGTTGG
TB38-f1	GTTGTTCGGGTTCCCTACGG	TB95_f1	GCAGGTGCTGCTGGACCG
TB38-f2	GCGTTGGGCCCTGCTCGAC	TB95_f2	CGATTTGGTCTCCAGCCCC
TB38-r1	GGCGGCATTCATCAGCAG	TB95_r1	CCCACCAGCAGCACCACC
TB38-r2	CAAGCGGTGCTGGATGTGG	TB95_r2	GTCGATGTCGTTGGTGCCC
TB40-f1	CGCACCCGGCTGATCGAG	TB96_f1	GTTTCGCGACGCTCTACAAG
TB40-f2	GCGATGGGCGGCTACGAC	TB96_f2	TAGCTACCTCGACGGTGTC
TB40-r1	CGACGACGGAACCTTCTGGC	TB96_r1	CTGCGGGTCCCAATCGTC
TB40-r2	TGCCCAGTTCGTTCAAACC	TB96_r2	GCCTGTTCTTGCTGGTTGG
TB41-f1	TGGGGCTACCATGTCCGGC	TB97-f1	ATGTGCTGAGGTCCAAGGG
TB41-f2	GGGTGTTTCGTGCTGGCGG	TB97-f2	AATGATGGTGTGTTGTAGG
TB41-r1	ACGTCACGACACAGACTCC	TB97-r1	TGGAGCAGTAAGTGTGGG
TB41-r2	ACCGGACACTTAGCGCCAC	TB97-r2	ATCAACGCGCCCACACAGC
TB42-f1	AACTGTTGTTACCAATCCG	TB97-f3	ACTGCTGACCGACACTACC
TB42-f2	TTGCCGACGCACATAGCCC	TB97-r3	TCGACAGCTTCACTGTGCC
TB42-r1	CCAGCAAATCCGATAGGTCC	TB97-f4	GCCGATGGGGGTGTTGAAG
TB42-r2	GTCGAAACCCATGTCCATCG	TB97-r4	GTTGAGCAACCTCGGTGGC
TB43-f1	TGGACTIONGCCGAGGAAGG	TB97-f5	TAGCGCCGAAGCTCCCATC

TB43-f2	TGCTACGCAACGTCGGGC	TB97-f6	GTGGTCGTGCTTTTGCTGC
TB43-r1	CTCGGCGATCAGCACGGC	TB97-r5	GGCTCAGGCGTCGTCAACG
TB43-r2	GCGGCGTTTCACCAGCAG	TB97-r6	CGTCGTCTCCGAGACTGGG
TB43-r3	GCAATGGCGGAATTGGTGG	TB97-r7	CGGCACTGACGGCATCACC
TB43-r4	CTGGTGGGCAACTCATTGG	TB97-f7	GACCAACCCGCCATGACCG
TB43-f3	TCGTACTIONGGTACGGTCCC	TB98-f1	CGTTACACGATCAGTTTGTCC
TB43_f4	CCGAACACGCCGAAGCCG	TB98-f2	GGTCAAGATGGCGGTCTGC
TB43_r5	GTTGGAGGCAAAGTCGGGG	TB98-r2	GCCGTTTGTGCTGTCAGCG
TB44-f1	GGCTGGTGTCCATCATCGG	TB98-r1	CTCGGCATTCCCTACGACG
TB44-f2	GCCGATTACGACTTGGTGC	TB98-f3	CAAACCAGCGTCCTGATCG
TB44-r1	GCGCTCAGTGCCCAAATGG	TB98-r3	GGTTTGC GGCTCGATATTGC
TB44-r2	TCCCTTGCAGGCCCTCG	TB99-f1	AAGGGAGAACAGCCATGTCCG
TB45-f1	GATAATCGGTTGGTTCGGCG	TB99-f2	GCCGAGAGCCACACCGTC
TB45-f2	GCGGGATGTTGGTGGGTAC	TB99-f3	GCAACGACGGCAACAATACC
TB45-r1	TTTGACAGCCATGCTCGGG	TB99-r1	ACGGAGCGGTGGAGTACG
TB45-r2	TGGCGGCGTTTCATGACCG	TB99-r2	CCGACGTTGCCGATGAGC
TB45-r3	GCAACGGTGGCAACGGCG	TB99-r3	CCACGACGTTGCCACGC
TB45-f3	TCTGCCCGTCACCGCC	TB100-f1	CGAGGTGATCCATGAGCAG
TB45-f0	TGCTGCTGGGGCTACGTCG	TB100-f2	CGACTAGCTGACCTGACGG
TB45-f4	TGGAACCGCCAAGTCCTGC	TB100-r1	CGAATCGTAGTTCCTGCGTC
TB45-f5	CCTGGGTTACCGTTTGGCC	TB100-r2	CCCTTAGACCAGATTGCCAG
TB45-f6	CGGACCAGCCATTACCAAC	TB101_f1	GTATCTGCGGCCACTACAG
TB45-f7	CAGACACCGCCGTTAGGG	TB101_f2	AGGCTCCGCTGAGTTGGTC
TB45-f8	CGTGCTGGGACAAGTCCTC	TB101_r1	AAGACCGCGCAGAACGACC
TB45-r4	CACCAACGAGGGCGATGGC	TB101_r2	GATCCGCGTGGCTATGTGC
TB45-r5	GGTGGCAACTCCCAAGTGG	TB101_f1	GTATCTGCGGCCACTACAG
TB45-r6	GATAACGGAGGCGATGGAG	TB101_f2	GCTGAGTTGGTCATAGATGG
TB45-r7	TCTCGTGCCTTGATTGTCGC	TB101_r1	CGGTACGCGAGATTCCCC
TB45-r8	TATCTGGGATTGATCAGTGTG	TB101_r2	ACGGTGCAGTTATGTCGGG
TB45-r9	GATGGTAGACCGGGTAACTG	TB104_f1	CATGAGCGGTTTCGTGCAGG
TB46-f1	TCTGGCGGTGGTGTCTGC	TB104_f2	GCTCGATTTGCTGAATGCG
TB46-f2	ACGCTGCACTGATTGGGGG	TB104_r1	TATCTAGGGTCGGCGGTGG
TB46-r1	AATGGATGAGATGCGCCCG	TB104_r2	AGGTGGCTGTCGCGTTCG
TB46-r2	GCGATGGATGGGTTGGCGG	TB105_f1	CCGTCACACGCTTTCGAGG
TB46-F3	CGTTGCCGCCGCTATTGCC	TB105_f2	TCCCAGGCACCGAAGACC
TB46-R3	AGCAGTCGGTGGCAAAGGC	TB105_r1	CGCTGGACTACGACATCGC
TB46-R4	GGTCGGCTCGTATGTCAGC	TB105_r2	AGTCATCCTCGACGAGACC
TB47-f	ATTGTGCGCTTGCTGCTGG	TB103-f1	TGACGTGGCGTGGCTGACC
TB47-r	CGACATGGGATACCGGGC	TB103-f2	GTGGACACCTGGCTGGCTG
TB48-f1	GATGGAGTGGGTGTTCCGGG	TB103-f3	CTCCGTAGGCCAGCCAG
TB48-f2	GTGTCCCAATGCGGTTCCC	TB103-r3	CCACCGAACCAGACCAGCC
TB48-r1	AGGTGCCAGATGTCGTTCCG	TB103-r1	GAAAGGGGTGTCAGATGTCCG
TB48-r2	GCGTTTCACAGCCAGTTCG	TB103-f4	GGACCCGTAGATCAAGCCG
TB49-f1	GAGTTGCTTGAGGTCTGCC	TB103-r2	CATGACGAGGCTCCACAGG

TB49-r1	GACCGTCTTGCCGACCCC
TB51-f1	CCTGGTGGCATCTGTCAGC
TB51-r1	AGTGGCGGCAACGGTATCG
TB51-f2	GTTGGCGAGTATGTTGATAGC
TB51-r2	GCGGGTTGTTCGGCAATGG