



УДК 579.25

ГЕНОМНЫЙ АНАЛИЗ КЛИНИЧЕСКИХ ШТАММОВ *Mycobacterium tuberculosis* СЕМЕЙСТВА LAM

*Тарлыков П.В., Атавлиева С.Ш.

Национальный центр биотехнологии

Кургальжинское шоссе, 13/5, Нур-Султан, 010000, Казахстан

*tarlykov@biocenter.kz

АБСТРАКТ

Лекарственно-устойчивый туберкулез является глобальной проблемой человечества, в особенности, в странах Юго-Восточной Азии, Восточной Европы и СНГ. В Казахстане показатель лекарственно-устойчивого туберкулеза среди впервые выявленных пациентов составляет более 20%. Более того, согласно данным ВОЗ, Казахстан занимает второе место среди тридцати стран с высоким бременем заболевания. В связи с этим становится очевидной необходимость исследования природы резистентности локально циркулирующих штаммов микобактерий. Применение методов молекулярно-генетического типирования позволит отслеживать распространения возбудителя туберкулеза любого семейства, что, в свою очередь, позволит обеспечить персонализированный подход к лечению. В данном исследовании были изучены геномы трех лекарственно-устойчивых клинических изолятов Латинско-Средиземноморского семейства (LAM), циркулирующих в Казахстане. Мы использовали полногеномное секвенирование для изучения распределения и лекарственной устойчивости изолятов. Филогенетический анализ сгруппировал геномы, описанные в этом исследовании, с последовательностями из России, Узбекистана и Казахстана, также принадлежащими к семейству LAM. У одного изолята была идентифицирована широкая лекарственная устойчивость к семи противотуберкулезным препаратам. Наши результаты показали, что в Казахстане циркулируют как минимум два лекарственно-устойчивых генотипа семейства LAM. Таким образом, применение молекулярно-генетических методов становится все более востребованным и эффективным в борьбе с распространением туберкулезной инфекции.

Ключевые слова: *Mycobacterium tuberculosis*, геном, туберкулез, филогения, ДНК, мутация, антибиотикорезистентность

ВВЕДЕНИЕ

В Республике Казахстан туберкулёзу уделяется особое внимание. Туберкулёз является первым в списке социально-значимых заболеваний, утвержденных Постановлением Правительства Республики Казахстан от 4 декабря 2009 года № 2018 «Об утверждении перечня социально-значимых заболеваний». Это вызвано не только инфекционной природой самого заболевания, но и его распространением среди социально-незащищенных и неблагополучных слоев населения. Кроме того, несмотря на общее снижение заболеваемости туберкулезом, в последние годы в Казахстане наблюдается опасная тенденция активного распространения его антибиотикорезистентных форм. Согласно данным ВОЗ, Казахстан входит в двадцатку стран, лидирующих по количеству случаев туберкулеза с множественной лекарственной устойчивостью. В годовом отчете за 2016 и последующие годы ВОЗ отметил, что количество случаев туберкулеза с множественной лекарственной устойчивостью среди больных в Казахстане уменьшается медленнее, чем общее количество новых зарегистрированных случаев туберкулеза. Это основные причины, почему Казахстан вошел в список стран, где требуются активные действия ВОЗ в 2016-2021 годах для достижения прогресса в борьбе с лекарственной устойчивостью. Именно распространение антибиотикорезистентных форм туберкулеза является основным обоснованием для проведения геномных исследований штаммов *M. tuberculosis*, циркулирующих на территории Республики Казахстан.

Возбудитель туберкулеза *M. tuberculosis* – человеческий патоген, имеющий несколько генетических семейств (линий), которые изначально появились в разных географических регионах [1]. Линии, присутствующие в комплексе *M. tuberculosis* (МТВС), делятся на древние и современные. К древним филогенетическим линиям относятся: линия 1 (Индо-Океаническая), линия 5 (Западная Африка 1), линия 6 (Западная Африка 2), а недавно обнаруженная линия 7 (Эфиопия), по-видимому, занимает промежуточное положение между древними и современными [2]. Современные филогенетические линии включают линию 2 (Восточно-Азиатскую), линию 3 (Восточно-Афро-Индийскую) и линию 4 (Евро-Американскую), последняя включает в себя Латиноамериканско-Средиземноморское (LAM) генетическое семейство. Семейство LAM было впервые обнаружено на основе коллекции штаммов, происходящих из Латинской Америки и Средиземноморья [3, 4]. Миграции людей привели к недавнему распространению семейства LAM по всему миру. В 2014 г. семейство LAM уже наблюдалось в 47 странах с разной степенью распространенности, включая несколько стран Центральной Азии (Казахстан, Узбекистан и Туркменистан) и Россию [5]. На сегодняшний день LAM является наиболее распространенным семейством *M. tuberculosis* в Казахстане после семейства Beijing [6]. Локусы региона различия (Region of Difference -RD) разделили генетическое семейство LAM на несколько подсемейств, а именно RD-Rio, RD174 и RD115. Сублиния RD-Rio определяется делецией в 26 тыс. п.н. и обычно сопровождается большой делецией RD174 [7]. Другая сублиния характеризуется делецией RD115, которая включает ветвь LAM-RUS со специфической вставкой IS6110 в ген *plcA* [8].

В настоящем исследовании предоставлены данные полногеномного секвенирования (WGS) трех изолятов LAM из коллекции лекарственно-устойчивых изолятов *M. tuberculosis*, собранных в городе Нур-Султан, Казахстан.

В этом исследовании мы провели полногеномное секвенирование с анализом *in silico* эпидемиологических характеристик местных изолятов LAM.

Казахстан входит в число 30 стран с самым высоким бременем туберкулеза с множественной лекарственной устойчивостью (МЛУ) в мире [9]. Недавнее распространение лекарственно-устойчивых изолятов в стране связано с «успешным» распространением Central Asian/Russian сублинии *M. tuberculosis* Beijing [10]. В то же время, в ряде исследований сообщается о быстром приобретении лекарственной устойчивости в семействе LAM, включая изоляты KZN из Южной Африки и сублинии LAM-RUS, широко распространенной в Казахстане и соседних странах [8, 11]. В предыдущих исследованиях генетическое семейство LAM было обнаружено в пропорции 11% в выборке из 470 изолятов *M. tuberculosis* из 12 областей Казахстана ($51/470 = 10,85\%$) [6, 12]. В другом исследовании описываются местные изоляты LAM, представленные в основном сублинией RD115 LAM-RUS ($29/30 = 96,67\%$) [13]. Кроме того, недавнее исследование выявило первый изолят с делецией RD-Rio (LAM RD-Rio), не эндемичный для Центральной Азии [14]. Тем не менее, геномные данные по изолятам LAM, циркулирующим в Казахстане, ограничены. Небольшое количество местных коллекций изолятов МТВС были протестированы на локусы RD или инсерции IS6110 [13]. Кроме того, на сегодняшний день опубликованы только три полногеномные последовательности (WGS) изолятов LAM из Казахстана [5]. Дополнительные геномы предоставят больше данных о генетических вариациях, встречающихся в изолятах семейства лекарственно-устойчивых LAM, циркулирующих в Казахстане и соседних странах.

Полногеномное секвенирование ДНК, полученной из выращенной культуры МБТ, используется для получения информации по клинически значимой лекарственной устойчивости, исходя из мутаций в генах, ответственных за сопротивление. Выявление известных мутаций резистентности в течение недели с момента получения образца открывает перспективы для персональной терапии туберкулеза с различными формами лекарственной устойчивости. Кроме того, именно данные полногеномного секвенирования штаммов *M. tuberculosis*, полученных от больных, участвующих в цепи трансмиссии ТБ, позволяют выявить генетически различающиеся источники инфекции с высокой точностью.

Эпидемиология *M. tuberculosis*, включая линию 4, обычно изучается путем анализа однонуклеотидных полиморфизмов (ОНП), региона различия (RD), полиморфизма длин рестрикционных фрагментов IS6110 (RFLP), сполиготипирования и/или типирования микобактериальных диспергированных единиц переменного числа копий (MIRU- Mycobacterial Interspersed Repetitive Units) [15]. Некоторые из этих традиционных методов молекулярного генотипирования трудоемки и имеют различные ограничения [16].

В области изучения *M. tuberculosis* большинство исследований были сфокусированы на генетической составляющей этого патогена. Полная геномная последовательность лабораторного штамма *Mycobacterium tuberculosis* H37Rv была расшифрована в 1998 году, при этом было выявлено около 4000 открытых рамок считывания, из которых 16% были специфичны для микобактерий. Большая часть из этих предполагаемых белков была отнесена к факторам вирулентности и патогенности [18]. Предполагаемые факторы вирулентности, присущие *M. tuberculosis*, были изучены с использованием мутагенеза, а также методов сравнительной геномики для идентификации микобактериальных генов [19], необходимых для выживания *in vivo* [20]. Знания, полученные в результате

генетического анализа микобактерий, улучшили понимание молекулярно-генетических механизмов *M. tuberculosis*, однако статичность геномных данных накладывает некоторые ограничения, не позволяющие изучить особенности механизмов выживания и распространения этой бактерии, а также динамическую регуляцию биологических процессов, происходящих в клетке в разных условиях окружающей среды.

Материалы и методы

Образцы *M. tuberculosis* были выделены из мокроты пациентов г. Нур-Султана с клиническим подозрением на туберкулез. Была собрана коллекция из 28 штаммов микобактерий с лекарственной устойчивостью. Чувствительность к лекарствам тестировали с использованием системы культивирования Vactec MGIT 960 (Becton, Dickinson) в соответствии с протоколом производителя. ДНК экстрагировали с использованием протокола цетилтриметиламмония бромида (СТАВ) [17]. Качество ДНК проверяли с помощью набора для анализа двухцепочечной ДНК высокой чувствительности (dsDNA HS, Thermo) и флуориметра Qubit 2.0 (Thermo). Анализ однонуклеотидного полиморфизма (ОНП) проводили с использованием ПЦР в реальном времени на приборе CFX96 Touch (Bio-Rad) для обнаружения замены G на A в кодоне 103 *fbpC* (Rv0129c). Три изолята LAM с подтвержденным ОНП были секвенированы (№ 3538, 4142 и 4330).

Изоляты № 3538 и 4142 секвенировали с использованием платформы MiSeq (Illumina), а другой изолят LAM № 4330 с использованием платформы Ion Torrent (Thermo). Для секвенирования на MiSeq библиотеки со средним размером фрагментов 600 п.н. были приготовлены с использованием набора Nextera DNA Flex Library Prep (Illumina) в соответствии с инструкциями производителя. Библиотеку с баркодами для изолята № 4330 получили с использованием набора библиотеки фрагментов Ion Xpress Plus и набора адаптеров баркода 1-16 Ion Xpress (Thermo). Средний размер библиотеки 480 п.н. для библиотеки с лигированным адаптером с 400 основаниями для чтения был выбран по размеру с помощью агарозного геля E-Gel SizeSelect II (Invitrogen). Секвенирование проводилось на платформе Ion Torrent PGM с использованием набора для секвенирования Hi-Q (Thermo) и чипа 318 (Thermo), как было описано ранее [18]. Качество исходных данных секвенирования проверяли с помощью FastQC v.0.11.9 [19]. Необработанные прочтения последовательности, отфильтрованные с помощью Trimmomatic v.0.38 (оценка Phred>20), были использованы для дальнейшего анализа нуклеотидных вариаций и сборки *de novo* с помощью SPAdes v.3.14.1 [20]. Онлайн-инструмент PhyResSe использовался для проверки гетерорезистентности полученных данных WGS. Аннотации геномов выполнялись NCBI Prokaryotic Genome Annotation Pipeline для парных прочтений или Prokka annotation pipeline v.1.14.5 для односторонних прочтений. Для всего программного обеспечения использовались параметры по умолчанию.

Сполиготимирование *in silico* было выполнено с использованием программы SpoTyping 2.1, а определение областей делеции было выполнено с помощью TB-Profiler. Прочтения были сопоставлены с эталонным геномом H37Rv (номер доступа в Genbank NC_000962.3), для подтверждения IS6110 вставок и

сполигопрофилей с помощью набора праймеров Kamerbeek, J. и др. (1997), был использован Geneious Prime v.2019.2.1.

Матрица ОНП была получена путем сравнения полиморфизмов, обнаруженных между ранее изученными геномами, и последовательностями 81 изолята МТВС. Отфильтрованные последовательности прочтений были картированы с помощью программы BWA-MEM с референсной последовательностью генома *M. tuberculosis* H37Rv (NC_000962.3). ОНП были получены с использованием алгоритма UnifiedGenotyper (GATK v.3.8.1.0). До компиляции объединенных последовательностей, ОНП повторяющихся последовательностей и генов PE/PPE были отфильтрованы по геному (NC_000962.3) с помощью TB Variant Filter v.0.1.3. Вызовы вариантов с покрытием на каждое основание с глубиной менее 10x или значения Phred меньше 20 были удалены. Полученные ОНП с высокой степенью достоверности впоследствии были записаны для выравнивания в мульти-FASTA файл. Филогенетическое дерево было построено по методу максимального правдоподобия на основе объединенных ОНП с использованием программы RAxML v8.2.11 со 100 итерациями начальной загрузки. Мы использовали модель нуклеотидной замены General Time Reversible (GTR), реализованную в RAxML. Филогения была укоренена с использованием *Mycobacterium bovis* в качестве внешней группы. Филогенетическое дерево визуализировали с помощью программы FigTree v.1.4.4.

РЕЗУЛЬТАТЫ

Характеристики необработанных полногеномных последовательностей изолятов ЛАМ перечислены в таблице 1. По данным полногеномного секвенирования в среднем покрытие геномов составило 110. Прибор MiSeq генерирует считывание последовательности с обоих концов фрагмента (считывание парных концов); в то время как Ion Torrent производит одностороннее чтение. Секвенирование с обоих концов фрагмента дает прочтения, способные точно определять геномные перестройки и повторяющиеся элементы последовательности. Процедура подготовки библиотеки заняла два дня для обеих платформ, в то время как секвенирование MiSeq заняло больше времени, чем односторонняя технология Ion Torrent (38 часов против 8 часов). Наименьшее покрытие (~ 84x) было получено для ДНК, секвенированной в одностороннем направлении из изолята № 4330, как показано в таблице 1.

Таблица 1. Характеристики трех собранных геномов *M. tuberculosis*

№	Платформа	Инвентарный номер SRA	Размер генома (п.н.)	Содержание GC (%)	Покрываемость (x)	Количество контигов	Общее количество кодирующих последовательностей
3538	MiSeq	SRR11241401	4,427,028	65.24	130	404	4,392
4142	MiSeq	SRR11241400	4,351,634	65.58	116	243	4,237



4330	Ion Torrent	SRR112414 02	4,344,962	65.00	84	925	4,495
------	----------------	-----------------	-----------	-------	----	-----	-------

Все три образца содержали высоко достоверные мутации в различных генах, связанных с лекарственной устойчивостью, по сравнению с геномом эталонного штамма H37Rv (инвентарный номер NC_000962.3) (Таблица 2). Как следует из таблицы 2, изученные изоляты не имели смеси аллелей дикого типа и мутантных аллелей, также известной как гетерорезистентность. Также не было смешанных вызовов в аллелях, связанных с лекарственной устойчивостью, с меньшими аллелями, составляющими более 5% глубины чтения. Фенотипическая чувствительность изолятов была проверена на изониазид (INH), рифампин (RIF), стрептомицин (SM), этамбутол (EMB), амикацин (AMI), канамицин (KAN) и офлоксацин (OFX). Полногеномное секвенирование подтвердило результаты фенотипирования изолятов с множественной лекарственной устойчивостью (МЛУ) № 3538 и 4330 и изолят с широкой лекарственной устойчивостью (ШЛУ) № 4142. Было обнаружено, что генотипическое предсказание чувствительности *M. tuberculosis* к противотуберкулезным агентам коррелирует с фенотипической восприимчивостью. Это подтверждает прогнозируемое определение лекарственной устойчивости на основе полногеномного секвенирования, как ценный инструмент для клинического использования.

Таблица 2. Мутации, наблюдаемые в локусах, связанных с лекарственной устойчивостью изолятов *M. tuberculosis*

Изолят	Устойчивость к лекарствам							ETH
	AMI	EMB	INH	OFX	PZA	RIF	SM	
3538	-	-		-	-		-	-
Ген			<i>katG</i>			<i>rpoB</i>		
Замена			Ser315Thr G/C 944			Ser450Leu C/T 1349		
4142					-			
Ген	<i>rrs</i>	<i>embB</i>	<i>katG</i>	<i>gyrA</i>		<i>rpoB</i>	<i>rrs</i>	<i>fabG</i>
Замена	- A/G140 1	Met306Ile G/C 918	Ser315Thr G/C 944	Asp94Tyr G/T 280		His445Leu A/T 1334	- A/C 514	- C/T -15
4330	-			-	-			
Ген		<i>embB</i>	<i>katG</i>			<i>rpoB</i>	<i>rrs</i>	<i>fabG</i>
Замена		Met306Ile G/C 918	Ser315Thr G/C 944			His445Leu A/T 1334	- A/C 514	- C/T-15

AMI - амикацин; EMB - этамбутол; INH - изониазид; OFX - офлоксацин; PZA - пипразинамид; RIF - рифампицин; SM - стрептомицин; ETH - этионамид.

Уровень распространенности изолятов LAM в данном исследовании соответствует ранее опубликованным данным ($3/28 = 10,71\%$). Делеция большого участка генома RD115 наблюдалась у трех изученных изолятов (№ 3538, 4142 и 4330). Два из трех образцов были отнесены к ветви LAM-RUS на основании вставки IS6110 в гене *plcA* (№ 4142 и 4330). Преобладание изолятов LAM-RUS в местных образцах вызывает интерес, поскольку линия 4 является наиболее гетерогенной линией *M. tuberculosis*, состоящей из 10 различных сублиний, определяемых отсутствием конкретных локусов RD, называемых RD115, RD122, RD174, RD182, RD183, RD193, RD219, RD724, RD726 и RD761. Все три генома локальных изолятов LAM, секвенированные Stucki и др. (2016), также были определены большой делецией RD115 (Таблица 3). Два из них были изолятами LAM-RUS со специфической вставкой IS6110 в гене *plcA* (G04493 и G04546) [8]. Примечательно, что RD115 присутствовал во всех изолятах LAM-RUS (Таблица 3), в то время как два изолята с RD115 не принадлежат к семейству LAM-RUS. Географическое картирование семейства LAM-RUS показывает его распространенность в Северной Евразии, включая Россию и Казахстан. Также сообщалось о других странах, таких как Бразилия, Венесуэла, Эфиопия и Сьерра-Леоне.

Таблица 3. Генотип LAM и SIT в изолятах *M. tuberculosis* из Казахстана по данным полногеномного секвенирования

Изолят	Год	Сублиния*	RD	SIT	IS6110 (ген <i>plcA</i>) LAM-RUS	Устойчивость	Источник
3538	2014	L4.3.3	RD115	42	-	МЛУ	-
4142	2014	L4.3.3	RD115	42	+	ШЛУ	-
4330	2014	L4.3.3	RD115	42	+	МЛУ	-
G04485	2002	L4.3.3	RD115	42	-	МЛУ	[5]
G04546	2007	L4.3.3	RD115	254	+	Ч	[5]
G04493	2011	L4.3.3	RD115	444	+	Ч	[5]

* Классификация сублиний основана на Coll et al. [2]
МЛУ - изолят с множественной лекарственной устойчивостью; ШЛУ - изолят с широкой лекарственной устойчивостью; Ч - чувствительный изолят.

В текущем исследовании сполиготипирование *in silico* классифицировало все три изолята как международный тип сполиготипа (SIT- spoligotype international type) 42. SIT42 является основным предковым сполиготипом семейства LAM, также известным как прототип сполигопротеина LAM. Эти недавние результаты Mokrousov и др. указывают на то, что в коллекции изолятов *M. tuberculosis* из России, Беларуси и Казахстана в значительной степени доминировал LAM-RUS в SIT42. Кроме того, мы провели сполиготипирование *in silico* для трех локальных изолятов LAM, секвенированных Stucki и др. Наш анализ *in silico* предполагает, что, по крайней мере, четыре популяции сублиний семейства LAM, а именно SIT42, SIT42/LAM-RUS, SIT254/LAM-RUS и

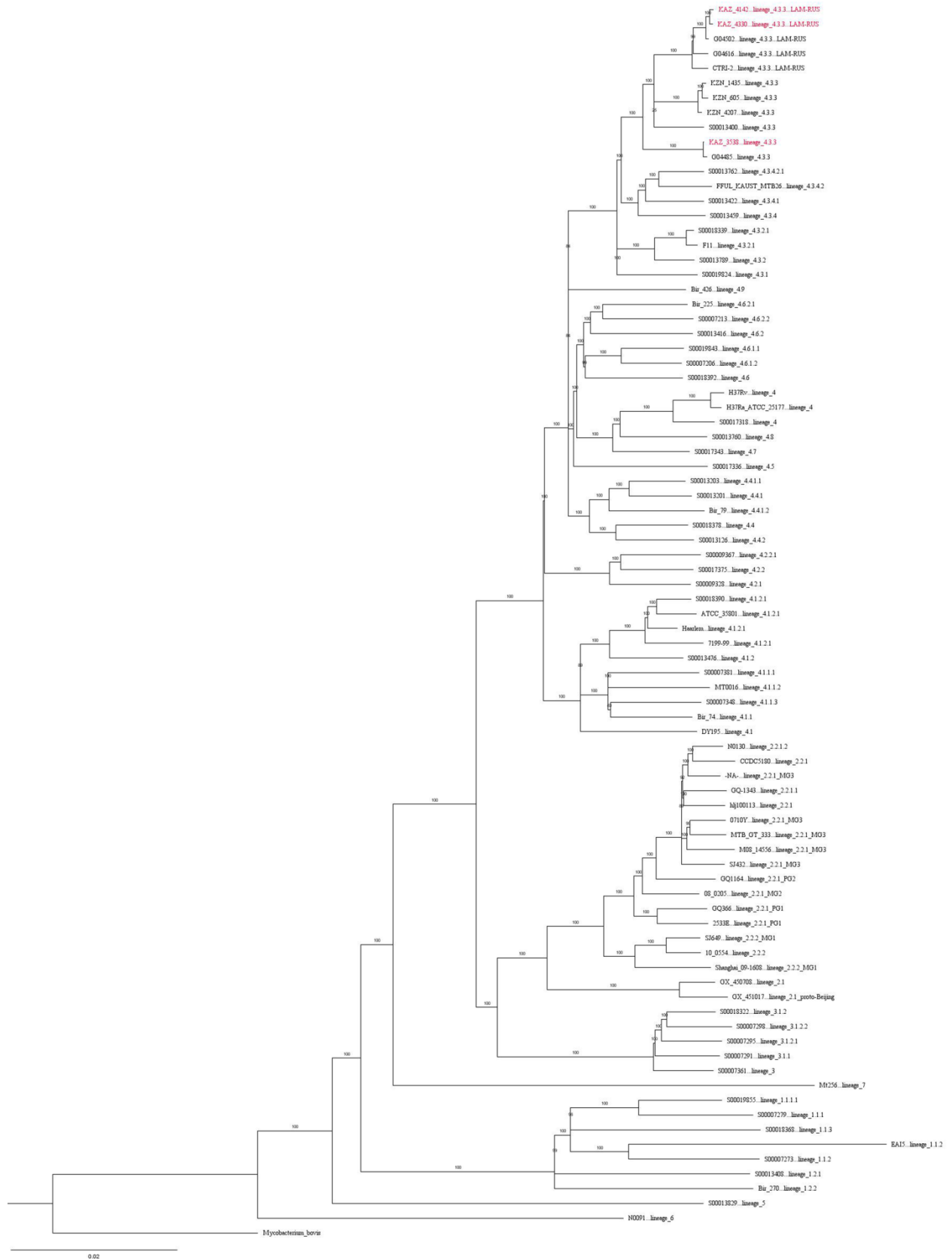


SIT444/LAM-RUS, циркулируют в Казахстане на основе данных полногеномного секвенирования. Из них два генотипа (SIT42 и SIT42/LAM-RUS) связаны с фенотипом МЛУ/ШЛУ.

Помимо сполиготипирования, наличие специфических ОНП является еще одной характеристикой изолятов LAM. Три изученных изолята несли одинаковый набор , специфичных для семейства LAM, включая замену G/A в *fbpC* 103 кодоне (Rv0129c), замену G/A в положении 8,040 (Rv0006), замену C/T в положении 403,364 (Rv0338c), замену G/A в положении 2,518,919 (Rv2245) и замену C/G в положении 3,426,795 (Rv3062). В результате, три изученных изолята были определены к генетической линии и сублинии 4.3.3 на основании результатов сполиготипирования *in silico* вместе с филогенетическим анализом ОНП.

ОБСУЖДЕНИЕ

В заключении, мы провели филогенетический анализ максимального правдоподобия трех отобранных штаммов. Филогенетическое дерево было построено на основе всех ОНП, извлеченных из 84 последовательностей геномной ДНК МТВС, включая данные полногеномного секвенирования из этого исследования (Рисунок 1, Таблица 4).



Сублинии указаны в соответствии с классификацией Coll et al.

Длины ответвлений пропорциональны заменам нуклеотидов, а топология основана на *Mycobacterium bovis*.

Красным выделены три изолята из этого исследования.

Рис. 1. Филогения с максимальной вероятностью трех изолятов *M. tuberculosis* из этого исследования и

81 репрезентативного генома штаммов МТВС.

Таблица 4. Список из 81 штамма МТВС, использованного в этом исследовании, с информацией о генотипах и их сублиниях

Изолят	Инвентарный номер	Генотип и сублиния*
S00007279	ERR117694	1.1.1
S00019855	ERR387008	1.1.1.1
EAI5	NC_021740	1.1.2
S00007273	ERR114427	1.1.2
S00018368	ERR351915+ERR386847	1.1.3
S00013408	ERR278529+ERR386888	1.2.1
Bir_270	ERR046881	1.2.2
GX_450708	SRR1710066	2.1
GX_451017	SRR1710070	2.1 proto-Beijing
hlj100113	SRS475367	2.2.1
CCDC5180	NC_017522	2.2.1
08_0205	SRR1710073	2.2.1
SJ432	SRR1710110	2.2.1
0710Y	ERR117454	2.2.1
M08_14556	ERR015616	2.2.1
MTB_GT_333	ERR234208	2.2.1
ERR019574	ERR019574	2.2.1
2533E	ERR234658	2.2.1
GQ366	ERR234133	2.2.1
GQ1164	ERR234116	2.2.1
GQ-1343	ERR234121	2.2.1.1
N0130	ERR234263	2.2.1.2
10_0554	SRR1710083	2.2.2
Shanghai_09-1608	SRS790114	2.2.2
SJ649	SRR1710111	2.2.2



S00007361	ERR114513	3
S00007291	ERR114453	3.1.1
S00018322	ERR351869+ERR386801	3.1.2
S00007295	ERR117696	3.1.2.1
S00007298	ERR114442	3.1.2.2
H37Rv	NC_000962	4
S00017318	ERR330651	4
H37Ra_ATCC_25177	NC_009525	4
DY195	ERR234201	4.1
Bir_74	ERR038277	4.1.1
S00007381	ERR114477	4.1.1.1
MT0016	SRR058116	4.1.1.2
S00007348	ERR117732	4.1.1.3
S00013476	ERR278597+ERR386956	4.1.2
7199-99	NC_020089	4.1.2.1
Haarlem	NC_022350	4.1.2.1
ATCC_35801	NC_020559	4.1.2.1
S00018390	ERR351937+ERR386869	4.1.2.1
S00009328	ERR228183	4.2.1
S00017375	ERR330708	4.2.2
S00009367	ERR228222	4.2.2.1
S00019824	ERR386977	4.3.1
S00013789	ERR294209	4.3.2
S00018339	ERR351886+ERR386818	4.3.2.1
F11	NC_009565	4.3.2.1
S00013400	ERR278521+ERR386880	L4.3.3
G04485	ERR1193792	L4.3.3
G04502	ERR1193802	L4.3.3 LAM-RUS
G04616	ERR1193863	L4.3.3 LAM-RUS
CTRI-2	NC_017524	L4.3.3 LAM-RUS
KZN_4207	NC_016768	4.3.3
KZN_605	NC_018078	4.3.3



KZN_1435	NC_012943	4.3.3
S00013459	ERR278580+ERR386939	4.3.4
S00013422	ERR278543+ERR386902	4.3.4.1
FFUL_KAUST_MTB26	ERR275206	4.3.4.2
S00013762	ERR294182	4.3.4.2.1
S00018378	ERR351925+ERR386857	4.4
S00013201	ERR270699	4.4.1
S00013203	ERR270701	4.4.1.1
Bir_79	ERR038282	4.4.1.2
S00013126	ERR270624	4.4.2
S00017336	ERR330669	4.5
S00018392	ERR351939+ERR386871	4.6
S00019843	ERR386996	4.6.1.1
S00007206	ERR123926	4.6.1.2
S00013416	ERR278537+ERR386896	4.6.2
Bir_225	ERR046839	4.6.2.1
S00007213	ERR123915	4.6.2.2
S00017343	ERR330676	4.7
S00013760	ERR294180	4.8
Bir_426	ERR072039	4.9
S00013829	ERR294249	5
N0091	ERR234254	6
Mt256	ERR181435	7
Canettii_6	SRR10177258	<i>M. canettii</i>
* Классификация по генетическим линиям и сублиниям основана на Coll et al.		

Используемый набор данных включал пятьдесят опубликованных геномных последовательностей, представляющих филогению *M. tuberculosis* линии 4 согласно классификации Coll и др. В результате сублиния LAM 4.3.3 была разделена на несколько ветвей. Последовательности №4142 и 4330 образуют отдельную ветвь 4.3.3 LAM-RUS вместе с опубликованными геномными последовательностями изолятов G04616 (Узбекистан), G04502 и CTRI-2 (Россия), а №3538 сгруппированы с изолятом G04485 (Казахстан). Филогения местных изолятов LAM-RUS соответствует более ранней гипотезе об основании бактериальной субпопуляции в Северной Евразии, которая получила широкое распространение в результате масштабных миграций людей во времена



Советского Союза. Данные полногеномного секвенирования сохранены в базе данных DDBJ/ENA/GenBank под инвентарным номером JABACH000000000, JABACG000000000 и JABACF000000000. Необработанные данные полногеномного секвенирования были представлены в NCBI SRA (инвентарные номера перечислены в таблице 1).

ВЫВОДЫ

Полученные нами данные отличаются от более ранних исследований тем, что основаны на полногеномном секвенировании. Кроме того, в рамках этой работы было проведено генотипическое прогнозирование лекарственной устойчивости в сочетании с эпидемиологическим генотипированием, которое охватило *in silico* сполиготипирование, присвоение локусов RD, а также вставки IS6110. Полученные геномы предоставили новые данные о генетических вариациях, встречающихся в МЛУ/ШЛУ изолятах семейства LAM, циркулирующих в Центральной Азии. Мы обнаружили два МЛУ/ШЛУ-ассоциированных генотипа семейства LAM, циркулирующих в Казахстане (SIT42 и SIT42/LAM-RUS). Появление приобретенной широкой и множественной лекарственной устойчивостью у местных изолятов семейства LAM вызывает особую озабоченность. Дальнейшее распространение этих устойчивых к лекарствам штаммов представляет угрозу региональной борьбе с туберкулезом и эффективности стандартизированной стратегии лечения. Одним из ограничений текущего исследования был сбор геномных последовательностей, который не является исчерпывающим для подтверждения происхождения и путей миграции изучаемых изолятов LAM. Дальнейшее сравнение полученных данных полногеномного секвенирования с другими циркулирующими изолятами сублинии LAM 4.3.3 необходимо для выявления генетических различий, ответственных за патогенность и эпидемиологию (инфицированность) изолятов LAM.

Финансирование

Данная работа выполнена в рамках финансирования грантового проекта Комитета науки Министерства образования и науки Республики Казахстан - ИРН AP09058045

ЛИТЕРАТУРА

1. de Almeida I.N., Vasconcellos S.E., de Assis Figueredo L.J., Dantas N.G., Augusto C.J., Hadaad J.P. Frequency of the *Mycobacterium tuberculosis* RD(Rio) genotype and its association with multidrug-resistant tuberculosis // BMC Infect Dis. – 2019. – Vol. 19, № 1. - P. 556. <https://doi.org/10.1186/s12879-019-4152-7>

2. Coll F., McNerney R., Guerra-Assuncao J.A., Glynn J.R., Perdigao J., Viveiros M. et al. A robust OHII barcode for typing *Mycobacterium tuberculosis* complex strains // Nat. Commun. – 2014. – Vol. 5. –P. 4812. <https://doi.org/10.1038/ncomms5812>
3. Sola C., Filliol I., Legrand E., Mokrousov I., Rastogi N. *Mycobacterium tuberculosis* Phylogeny Reconstruction Based on Combined Numerical Analysis with IS1081, IS6110, VNTR, and DR-Based Spoligotyping Suggests the Existence of Two New Phylogeographical Clades // J Mol Evol. - 2001. - Vol. 53, № 6. – P.680-689. <https://doi.org/10.1007/s002390010255>
4. van Soolingen D., Qian L., de Haas P.E., Douglas J.T., Traore H., Portaels F. et al. Predominance of a single genotype of *Mycobacterium tuberculosis* in countries of east Asia // J Clin Microbiol. – 1995. - Vol.33, № 12. P. 3234-3288. <https://doi.org/10.1128/JCM.33.12.3234-3238.1995>
5. Stucki D., Brites D., Jeljeli L., Coscolla M., Liu Q., Trauner A. et al. *Mycobacterium tuberculosis* lineage 4 comprises globally distributed and geographically restricted sublineages // Nat. Genet. – 2016. – Vol. 48, № 12. –P. 1535-1543. <https://doi.org/10.1038/ng.3704>
6. Klotoe B.J., Kacimi S., Costa-Conceicao E., Gomes H.M., Barcellos R.B., Panaiotov S. et al. Genomic characterization of MJTY/IIIJTY-TB in Kazakhstan by a combination of high-throughput methods predominantly shows the ongoing transmission of L2/Beijing 94-32 central Asian/Russian clusters // BMC Infect Dis. – 2019. – Vol. 9, № 1. - P. 553. <https://doi.org/10.1186/s12879-019-4201-2>
7. Lazzarini L.C., Huard R.C., Boechat N.L., Gomes H.M., Oelemann M.C., Kurepina N et al. Discovery of a novel *Mycobacterium tuberculosis* lineage that is a major cause of tuberculosis in Rio de Janeiro, Brazil // J Clin Microbiol. – 2007. – Vol. 45. № 12. – P. 3891-902. <https://doi.org/10.1128/JCM.01394-07>
8. Dubiley S., Kirillov E., Ignatova A., Stepanshina V., Shemyakin I. Molecular characteristics of the *Mycobacterium tuberculosis* LAM-RUS family prevalent in Central Russia // J Clin Microbiol. – 2007. – Vol. 45, № 12. – P. 4036-4038. <https://doi.org/10.1128/JCM.01217-07>
9. World Health Organisation. Global tuberculosis report 2019. 2019. URL: https://www.who.int/tb/publications/global_report/en/
10. Merker M., Barbier M., Cox H., Rasigade J.P., Feuerriegel S., Kohl T.A. et al. Compensatory evolution drives multidrug-resistant tuberculosis in Central Asia // eLife. – 2018. –Vol.7. 10.7554/eLife.38200 <https://doi.org/10.7554/eLife.38200>
11. Pillay M., Sturm A.W. Evolution of the Extensively Drug-Resistant F15/LAM4/KZN Strain of *Mycobacterium tuberculosis* in KwaZulu-Natal, South Africa // Clin Infect Dis. – 2007. – Vol.45, № 11. – P. 1409-1414. <https://doi.org/10.1086/522987>
12. Skiba Y., Mokrousov I., Ismagulova G., Maltseva E., Yurkevich N., Bismilda V. et al. Molecular snapshot of *Mycobacterium tuberculosis* population in Kazakhstan: a country-wide study // Tuberculosis (Edinb). - 2015. - Vol. 95, № 5. - P. 538-46. <https://doi.org/10.1016/j.tube.2015.04.012>
13. Mokrousov I., Vyazovaya A., Narvskaya O. *Mycobacterium tuberculosis* Latin American-Mediterranean family and its sublineages in the light of robust



evolutionary markers // J Bacteriol. – 2014. Vol. 196, № 10. – P. 1833-41.

<https://doi.org/10.1128/JB.01485-13>

14. Skiba Y., Mokrousov I., Nabirova D., Vyazovaya A., Maltseva E., Malakhova N. et al. *Mycobacterium tuberculosis* RD-Rio Strain in Kazakhstan // Emerg Infect Dis. – 2019.- Vol. 25, №3. – P. 604-606. <https://doi.org/10.3201/eid2503.181179>

15. Ei P.W., Aung W.W., Lee J.S., Choi G.E., Chang C.L. Molecular Strain Typing of *Mycobacterium tuberculosis*: a Review of Frequently Used Methods // J Korean Med Sci. – 2016. Vol. 31, № 11. –P. 1673-1683.

<https://doi.org/10.3346/jkms.2016.31.11.1673>

16. Mokrousov I. On sunspots, click science and molecular iconography // Tuberculosis (Edinb). – 2018. – Vol.110. –P.91-95.

<https://doi.org/10.1016/j.tube.2018.04.004>

17. van Soolingen D., Hermans P.W., de Haas P.E., Soll D.R., van Embden J.D. Occurrence and stability of insertion sequences in *Mycobacterium tuberculosis* complex strains: evaluation of an insertion sequence-dependent DNA polymorphism as a tool in the epidemiology of tuberculosis // J Clin Microbiol. – 1991. – Vol. 29, № 11. - P. 2578-2586. <https://doi.org/10.1128/JCM.29.11.2578-2586.1991>

18. Shevtsov A., Tarlykov P., Zholdybayeva E., Shevtsova E., Momykulov D., Sytnik I. et al. Draft Genome Sequence of the Live Vaccine Strain *Brucella abortus* // Genome Announc. – 2013. –Vol. 1, № 6. <https://doi.org/10.1128/genomeA.01101-13>

19. Andrews S. FastQC: a quality control tool for high throughput sequence data. URL: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>. 2019.

20. Nurk S., Bankevich A., Antipov D., Gurevich A.A., Korobeynikov A., Lapidus A. et al. Assembling single-cell genomes and mini-metagenomes from chimeric MDA products // Journal of computational biology: a journal of computational molecular cell biology. – 2013. –Vol.20, №10. - P. 714-737.

<https://doi.org/10.1089/cmb.2013.0084>

REFERENCES

1. de Almeida I.N., Vasconcellos S.E., de Assis Figueredo L.J., Dantas N.G., Augusto C.J., Hadaad J.P. Frequency of the *Mycobacterium tuberculosis* RD(Rio) genotype and its association with multidrug-resistant tuberculosis. *BMC Infect Dis.*, 2019, vol. 19, no. 1, pp. 556. <https://doi.org/10.1186/s12879-019-4152-7>

2. Coll F., McNerney R., Guerra-Assuncao J.A., Glynn J.R., Perdigo J., Viveiros M. et al. A robust SNP barcode for typing *Mycobacterium tuberculosis* complex strains. *Nat. Commun.*, 2014, vol. 5, pp. 4812. <https://doi.org/10.1038/ncomms5812>

3. Sola C., Filliol I., Legrand E., Mokrousov I., Rastogi N. *Mycobacterium tuberculosis* Phylogeny Reconstruction Based on Combined Numerical Analysis with IS1081, IS6110, VNTR, and DR-Based Spoligotyping Suggests the Existence of Two New Phylogeographical Clades. *J Mol Evol.*, 2001, vol. 53, no. 6, pp.680-689.

<https://doi.org/10.1007/s002390010255>



4. van Soolingen D., Qian L., de Haas P.E., Douglas J.T., Traore H., Portaels F. et al. Predominance of a single genotype of *Mycobacterium tuberculosis* in countries of East Asia. *J Clin Microbiol.*, 1995, vol.33, no. 12, pp. 3234-3288. <https://doi.org/10.1128/JCM.33.12.3234-3238.1995>
5. Stucki D., Brites D., Jeljeli L., Coscolla M., Liu Q., Trauner A. et al. *Mycobacterium tuberculosis* lineage 4 comprises globally distributed and geographically restricted sublineages. *Nat. Genet.*, 2016, vol. 48, no. 12, pp. 1535-1543. <https://doi.org/10.1038/ng.3704>
6. Klotoe B.J., Kacimi S., Costa-Conceicao E., Gomes H.M., Barcellos R.B., Panaiotov S. et al. Genomic characterization of MJTY/IIIJTY-TB in Kazakhstan by a combination of high-throughput methods predominantly shows the ongoing transmission of L2/Beijing 94-32 central Asian/Russian clusters. *BMC Infect Dis.*, 2019, vol. 9, no. 1, pp. 553. <https://doi.org/10.1186/s12879-019-4201-2>
7. Lazzarini L.C., Huard R.C., Boechat N.L., Gomes H.M., Oelemann M.C., Kurepina N et al. Discovery of a novel *Mycobacterium tuberculosis* lineage that is a major cause of tuberculosis in Rio de Janeiro, Brazil. *J Clin Microbiol.*, 2007, vol. 45, no. 12, pp. 3891-902. <https://doi.org/10.1128/JCM.01394-07>
8. Dubiley S., Kirillov E., Ignatova A., Stepanshina V., Shemyakin I. Molecular characteristics of the *Mycobacterium tuberculosis* LAM-RUS family prevalent in Central Russia. *J Clin Microbiol.*, 2007, vol. 45, no. 12, pp. 4036-4038. <https://doi.org/10.1128/JCM.01217-07>
9. World Health Organisation. Global tuberculosis report 2019. 2019. URL: https://www.who.int/tb/publications/global_report/en/
10. Merker M., Barbier M., Cox H., Rasigade J.P., Feuerriegel S., Kohl T.A. et al. Compensatory evolution drives multidrug-resistant tuberculosis in Central Asia. *eLife*, 2018, vol.7. 10.7554/eLife.38200 <https://doi.org/10.7554/eLife.38200>
11. Pillay M., Sturm A.W. Evolution of the Extensively Drug-Resistant F15/LAM4/KZN Strain of *Mycobacterium tuberculosis* in KwaZulu-Natal, South Africa. *Clin Infect Dis.*, 2007, vol.45, no. 11, pp. 1409-1414. <https://doi.org/10.1086/522987>
12. Skiba Y., Mokrousov I., Ismagulova G., Maltseva E., Yurkevich N., Bismilda V. et al. Molecular snapshot of *Mycobacterium tuberculosis* population in Kazakhstan: a country-wide study. *Tuberculosis (Edinb)*, 2015, vol. 95, no.5, pp. 538-546. <https://doi.org/10.1016/j.tube.2015.04.012>
13. Mokrousov I., Vyazovaya A., Narvskaya O. *Mycobacterium tuberculosis* Latin American-Mediterranean family and its sublineages in the light of robust evolutionary markers. *J Bacteriol.*, 2014, vol. 196, no. 10, pp. 1833-41. <https://doi.org/10.1128/JB.01485-13>
14. Skiba Y., Mokrousov I., Nabirova D., Vyazovaya A., Maltseva E., Malakhova N. et al. *Mycobacterium tuberculosis* RD-Rio Strain in Kazakhstan. *Emerg Infect Dis.*, 2019, vol. 25, no.3, pp. 604-606. <https://doi.org/10.3201/eid2503.181179>
15. Ei P.W., Aung W.W., Lee J.S., Choi G.E., Chang C.L. Molecular Strain Typing of *Mycobacterium tuberculosis*: a Review of Frequently Used Methods. *J*



Korean Med Sci., 2016, vol. 31, no. 11, pp. 1673-1683.

<https://doi.org/10.3346/jkms.2016.31.11.1673>

16. Mokrousov I. On sunspots, click science and molecular iconography. *Tuberculosis (Edinb.)*, 2018, vol.110, pp.91-95.

<https://doi.org/10.1016/j.tube.2018.04.004>

17. van Soolingen D., Hermans P.W., de Haas P.E., Soll D.R., van Embden J.D. Occurrence and stability of insertion sequences in *Mycobacterium tuberculosis* complex strains: evaluation of an insertion sequence-dependent DNA polymorphism as a tool in the epidemiology of tuberculosis. *J Clin Microbiol.*, 1991, vol. 29, no. 11, pp. 2578-2586. <https://doi.org/10.1128/JCM.29.11.2578-2586.1991>

18. Shevtsov A., Tarlykov P., Zholdybayeva E., Shevtsova E., Momynkulov D., Sytnik I. et al. Draft Genome Sequence of the Live Vaccine Strain *Brucella abortus*. *Genome Announc.*, 2013, vol. 1, no.6. <https://doi.org/10.1128/genomeA.01101-13>

19. Andrews S. FastQC: a quality control tool for high throughput sequence data. URL: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>. 2019.

20. Nurk S., Bankevich A., Antipov D., Gurevich A.A., Korobeynikov A., Lapidus A. et al. Assembling single-cell genomes and mini-metagenomes from chimeric MDA products. *Journal of computational biology: a journal of computational molecular cell biology*, 2013, vol.20, no.10, pp. 714-737.

<https://doi.org/10.1089/cmb.2013.0084>

LAM ТЕҢІЗІ *MYCOBACTERIUM TUBERCULOSIS* ТЕКТЕС КЛИНИКАЛЫҚ ШТАММДАРЫНЫҢ ГЕНОМДЫҚ ТАЛДАУЫ

*Тарлыков П.В., Атавлиева С.Ш.

Ұлттық биотехнология орталығы,

Корғалжын тас жолы, 13/5, Нұр-Сұлтан, 010000, Қазақстан

* pavel.tarlykov@gmail.com

ТҮЙІН

Дәріге төзімді туберкулез адамзаттың жаһандық проблемасы болып табылады, әсіресе, оңтүстік-шығыс Азия, Шығыс Еуропа және ТМД елдерінде. Қазақстанда алғаш рет анықталған пациенттер арасында дәріге төзімді туберкулез көрсеткіші 20% астам. Сонымен қатар, Дүниежүзілік денсаулық сақтау ұйымының деректеріне сәйкес Қазақстан Республикасы аурудың ауыртпалығы көрсеткіші бойынша жоғары отыз елдің ішінде екінші орында тұр. Мутациялардың әртүрлі жиынтықтары бар төзімді микобактериялық штаммдардың әртүрлі салаларда таралуы диагностикалық жүйелер мен препараттарды қолдану тиімділігіне айтарлықтай әсер етуі мүмкін. Осыған байланысты микобактериялардың жергілікті айналымдағы штаммдарының резистенттілігінің



табиғатын зерттеу қажеттілігі айқын болып отыр. Молекулалық-генетикалық әдістерін қолдану кез келген туберкулез топтары қоздырғышының таралуын қадағалауға мүмкіндік береді, бұл өз кезегінде емдеуге дербестендірілген тәсілді қамтамасыз етуге мүмкіндік береді. Адамдарға бейімделген туберкулез микобактериялары кешенінің штамдарына бастапқыда олардың географиялық таралуына байланысты жеті филогенетикалық тұқым кіреді. Осы зерттеуде Қазақстанда жиналған Латын Америкасы-Жерорта теңізі тұқымдастарының (LAM) үш дәріге төзімді клиникалық изоляттарының геномдары зерттелді. Біз изоляттардың таралуы мен дәріге төзімділігін зерттеу үшін толық геномдық реттілікті қолдандық. Филогенетикалық талдау осы зерттеуде сипатталған геномдарды (LAM) тұқымдасына жататын Ресей, Өзбекстан және Қазақстанның тізбектерімен топтастырды. Бір изолят жеті туберкулезге қарсы препаратқа кең дәрілік төзімділікті көрсетті. Біздің нәтижелеріміз Қазақстанда LAM тобының кемінде екі МЛУ/ШЛУ-қауымдастырылған генотипі таралғанын болжайды. Осылайша, молекулалық-генетикалық әдістерді қолдану туберкулез инфекциясының таралуына қарсы күресте сұранысқа ие және тиімді болып келеді.

Негізгі сөздер: *Mycobacterium tuberculosis*, геном, туберкулез, филогения, ДНҚ, мутация, антибиотикке төзімділік



GENOMIC ANALYSIS OF LAM FAMILY OF *MYCOBACTERIUM TUBERCULOSIS* CLINICAL STRAINS

*Tarlykov P., Atavliyeva S.

National Center for Biotechnology,

13/5, Korgalzhyn road, Nur-Sultan, Kazakhstan

*pavel.tarlykov@gmail.com

ABSTRACT

Drug-resistant tuberculosis is a global problem for humanity, especially in the countries of Southeast Asia, Eastern Europe and the CIS. In Kazakhstan, the rate of drug-resistant tuberculosis among newly diagnosed patients is more than 20%. Moreover, the Republic of Kazakhstan ranks second among the thirty countries with a high disease burden according to the World Health Organization. In this regard, it becomes obvious that it is necessary to study the nature of resistance of locally circulating strains of mycobacteria. The use of genetic typing methods will allow tracking the spread of tuberculosis pathogens of any family, which, in turn, will provide a personalized approach to treatment. The strains of the *Mycobacterium tuberculosis* complex capable of infecting humans include seven phylogenetic lineages, initially associated with their geographical distribution. In this study, the genomes of three drug-resistant clinical isolates of the Latin American Mediterranean family (LAM) circulating in Kazakhstan were studied. We used whole genome sequencing to study the distribution and drug resistance of isolates. Phylogenetic analysis grouped the genomes described in this study with sequences from Russia, Uzbekistan and Kazakhstan, also belonging to the LAM family. One isolate was identified as having extensive drug resistance to seven anti-TB drugs. Our results suggest that at least two drug-resistant genotypes of the LAM family are circulating in Kazakhstan. Thus, the use of genetic methods is becoming more and more popular and effective in combating the spread of tuberculosis infection.

Key words: *Mycobacterium tuberculosis*, genome, tuberculosis, phylogeny, DNA, mutation, antibiotic resistance