

УДК 616-092

Применение полногеномного анализа для определения молекулярных маркеров значимых генетических кластеров *Mycobacterium tuberculosis* в России

Мокроусов И.В.¹, Черняева Е.Н.¹, Вязовая А.А.¹, Журавлев В.Ю.²

¹ Федеральное бюджетное учреждение науки «Санкт-Петербургский научно-исследовательский институт эпидемиологии и микробиологии имени Пастера» Федеральной службы по надзору в сфере защиты прав потребителей и благополучия человека.

197101, Санкт-Петербург, ул. Мира, д. 14

² Федеральное государственное бюджетное научное учреждение «Санкт-Петербургский Научно-исследовательский институт фтизиопульмонологии» Министерства Здравоохранения Российской Федерации.

191036, Санкт-Петербург, Лиговский проспект, д. 2-4

Актуальность. Туберкулёз относят к числу «вновь вернувшихся» заболеваний, получивших эпидемическое распространение во многих регионах мира. Фундаментальный интерес представляет понимание генетических основ адаптации возбудителя заболевания *Mycobacterium tuberculosis* к организму хозяина.

Целью данного исследования было выявление молекулярных маркеров для значимых геновариантов *M. tuberculosis* в России на основе проведения полногеномного и биоинформационного анализа.

Материалы и методы. Полногеномное секвенирование и филогенетический анализ были проведены для репрезентативной выборки штаммов *M. tuberculosis* Восточно-Азиатской линии (древние и современные сублинии генотипа Beijing) и Евро-Американской линии (штаммы генотипов LAM, T и S).

Результаты. В результате филогенетического анализа геномных данных (4500 полиморфных нуклеотидных позиций), были выделены кластеры штаммов и определены их специфические однонуклеотидные полиморфизмы (ОНП, SNP). В частности, внутри сублинии LAM-RUS был выявлен эволюционно «молодой» кластер, включающий штаммы сполізотипов SIT252 и SIT266. Внутри штаммов семейства Beijing были выделены ветви современной (субтипы B0 и 94-32) и древней сублиний. Среди выявленных кластер-специфических однонуклеотидных полиморфизмов для групп внутри семейств Beijing и LAM наиболее широко были представлены категории генов «процессы, связанные с клеткой и клеточной стенкой» и «промежуточный метаболизм и дыхание». Вариации в генах категории «вирулентность» встречались только в штаммах Beijing B0/W148, древних Beijing и LAM-RUS (SIT252/266).

Выводы. Для эмерджентных, актуально или потенциально эпидемических вариантов *M. tuberculosis* генотипов LAM и Beijing, выявлены как филогенетически нейтральные, диагностические полиморфизмы, так и мутации в ряде генов вирулентности, адаптации, биосинтеза клеточной стенки, дыхания и липидного обмена.

Ключевые слова: *Mycobacterium tuberculosis*; молекулярная эпидемиология; полногеномное секвенирование; генотип Beijing.

Для цитирования: Мокроусов И.В., Черняева Е.Н., Вязовая А.А., Журавлев В.Ю. Применение полногеномного анализа для определения молекулярных маркеров значимых генетических кластеров *Mycobacterium tuberculosis* в России. Патогенез. 2019; 17(4): 43-49

DOI: 10.25557/2310-0435.2019.04.43-49

Для корреспонденции: Мокроусов Игорь Владиславович, e-mail: imokrousov@mail.ru

Финансирование. Исследование получило частичную поддержку Российского фонда фундаментальных исследований (грант 19-04-00263).

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Поступила: 10.08.2019

The use of whole-genome analysis for identifying molecular markers of significant genetic clusters of *Mycobacterium tuberculosis* in Russia

Mokrousov I.V.¹, Chernyaeva E.N.¹, Vyazovaya A.A.¹, Zhuravlev V.Y.²

¹ St. Petersburg Pasteur Institute,

Mira Str. 14, St. Petersburg 197101, Russian Federation

² St. Petersburg State Research Institute of Phthisiopulmonology,

Ligovskiy Prospekt 2-4, St. Petersburg 191036, Russian Federation

*Tuberculosis is one of the re-emerging diseases that have epidemically spread in many regions worldwide. Understanding of the genetic basis for *Mycobacterium tuberculosis* adaptation to the human host is of a fundamental interest.*

The aim of this study was to identify molecular markers for significant genetic clusters of *M. tuberculosis* in Russia based on a whole genome and bioinformatics analysis. Whole genome sequencing and phylogenetic analysis were performed for a representative sample of *M. tuberculosis* strains of the East Asian lineage (ancient and modern sublineages of the Beijing genotype) and the Euro-American lineage (strains of LAM, T and S genotypes). The phylogenetic analysis of genomic data (4500 polymorphic nucleotide positions) allowed to identify strain clusters and to determine their specific single nucleotide polymorphisms (SNPs). Specifically, an evolutionarily “young” cluster was identified within the LAM-RUS branch, including

strains of the SIT252 and SIT266 spoligotypes. Modern (subtypes B0 and 94-32) and ancient sublineages were identified within the Beijing genotype. Beijing and LAM cluster-specific SNPs were mainly found in gene categories "cell wall and cell processes" and "intermediary metabolism and respiration". Variations in the genes of the "virulence" category were found only in strains of Beijing B0/W148, ancient Beijing and LAM-RUS (SIT252/266) groups. To conclude, phylogenetically neutral polymorphisms as well as mutations in virulence, adaptation, cell wall biosynthesis, respiration, and lipid metabolism genes were identified for emerging, potentially epidemic variants of the LAM and Beijing genotypes of *M. tuberculosis*.

Key words: *Mycobacterium tuberculosis*, molecular epidemiology, whole genome sequencing, Beijing genotype.

For citation: Mokrousov I.V., Chernyaeva E.N., Vyazovaya A.A., Zhuravlev V.Y. [The use of whole-genome analysis for identifying molecular markers of significant genetic clusters of *Mycobacterium tuberculosis* in Russia]. *Patogenez* [Pathogenesis]. 2019; 17(4): 43-49 (in Russian)

DOI: 10.25557/2310-0435.2019.04.43-49

For correspondence: Mokrousov Igor Vladislavovich, e-mail: imokrousov@mail.ru

Funding. This study received partial support from the Russian Foundation for Basic Research (grant 19-04-00263).

Conflict of interest. The authors declare no conflict of interest.

Received: 10.08.2019

Введение

Туберкулёз (ТБ) относится к числу так называемых «вновь вернувшихся» заболеваний, получивших эпидемическое распространение во многих регионах мира. По оценке ВОЗ, треть человечества инфицирована *Mycobacterium tuberculosis*. Возросшая мобильность населения приводит к глобальной диссеминации наиболее опасных мультирезистентных и гипервирулентных штаммов *M. tuberculosis*. Современные молекулярные методы позволили понять функционирование генома возбудителя туберкулёза *M. tuberculosis* и оценить эволюцию этого микроорганизма [1]. Для вида *M. tuberculosis* характерно отсутствие горизонтального генного переноса и, как следствие этого – клональная и иерархическая структура популяции, состоящей, на глобальном уровне, из 8 филогенетических линий, менее крупных генетических семейств и компактных клональных кластеров, которые отличаются по своей клинической и/или эпидемиологической значимости. Некоторые генетические семейства, например, генотип *Beijing*, широко распространены в различных регионах мира, вероятно, по причине их генетического преимущества [2, 3]. Генотип разделяют на древние (предковые) и современные сублинии, при этом в России преобладают штаммы современной сублинии, прежде всего, следует отметить эпидемический кластер *Beijing* B0 [4, 5]. При этом, для штаммов древней сублинии в Японии и Корее ранее были описаны различия в особенностях развития лекарственной устойчивости и степени кластеризации как прокси-маркера трансмиссивности [6-8] и недавно был выявлен резистентный кластер древней сублинии генотипа *Beijing* в Западной Сибири [9].

M. tuberculosis может адаптироваться к воздействию факторов иммунного ответа организма хозяина и к селективному давлению антибиотиков путем точечных мутаций. В последние годы, благодаря относительной доступности полногеномного секвенирования, стало возможным проведение анализа большого количества геномов клинических изолятов. При этом следует отметить, что зачастую качество анализа и оригиналь-

ность выводов существенно отстают от впечатляющего массива обработанных геномных данных. Немалая часть публикуемых статей повторяет старые известные теории, выдвинутые в первое десятилетие 21 века, в частности в том, что касается генетических основ лекарственной устойчивости, роли компенсаторных мутаций, роли генов селекции второго порядка, генов вирулентности.

Фундаментальный интерес представляет понимание того, какие гены и каким образом меняются в процессе ко-адаптации микроорганизма и хозяина, какие факторы эволюции (они же факторы коадаптации с хозяином) определяют текущую и историческую филогеографию возбудителя туберкулёза. Накопление изменений в определенных генах в процессе эволюции отражает его взаимную адаптацию с популяцией человека и происходит по-разному в различных генетических семействах (линиях) *M. tuberculosis*. Также следует учитывать разные стартовые позиции штаммов на конкретной территории, то есть, является ли штамм (геновариант) историческим или новозанесенным. Возможно, что более древние варианты несут в своих геномах следы более долгой эволюции и коадаптации и характеризуются меньшей вирулентностью и трансмиссивностью. В то же время, у относительно недавно возникших современных геновариантов были различные условия происхождения (в разных человеческих популяциях при разных условиях окружающей среды), и, как следствие этого – разные парадигмы эволюции в сторону большей или пониженной вирулентности и других патогенетических особенностей.

Целью данного исследования было выявление молекулярных маркёров для значимых геновариантов возбудителя туберкулёза в России на основе проведения полногеномного и биоинформационного анализа.

Материалы и методы исследования

Изоляты *M. tuberculosis* были получены от больных туберкулёзом лёгких, проживающих в Санкт-Петербурге. Биологический материал собран с письменного

согласия пациентов Санкт-Петербургского НИИ фтизиопульмонологии, без раскрытия персональных данных (процедура регламентирована Этическим комитетом НИИ фтизиопульмонологии). Идентификация изолятов на уровне вида была проведена с использованием стандартных микробиологических и биохимических тестов.

Экстракция и очистка ДНК бактерий были проведены стандартным методом: обработка клеточного лизата лизоцимом, протеиназой К, цетилтриметиламмоний бромидом и додецилсульфатом натрия для разрушения клеточной стенки и вторичных метаболитов, протеинов и полисахаридов, с последующей экстракцией ДНК с использованием фенола и хлороформа, и осаждением в изопропанол [10].

Сполиготипирование проводили согласно стандартному протоколу, с использованием мембраны с иммобилизованными олигонуклеотидами на 43 спейсера локуса CRISPR *M. tuberculosis*, изготовленной в лаборатории молекулярной микробиологии НИИЭМ им. Пастера [11, 12]. Профили сполитипирования сравнивали с международной базой данных SITVIT2 [13].

Для дальнейшего исследования была сформирована репрезентативная выборка штаммов для проведения полногеномного секвенирования (ПГС) эпидемиологически и клинически значимых вариантов и генотипов возбудителя туберкулёза, циркулирующих в России: генотип *Beijing* (древние и современные штаммы, варианты B0 и A0; все штаммы имели сполитип SIT1) и генотип *LAM* (сублинии *RD-Rio*, *LAM-RUS*, сполитипы 252, 254, 266). Также были включены штаммы других генотипов в качестве внешней группы (сполитипы SIT40 и SIT34, относящиеся к Евро-Американской линии).

Полногеномное секвенирование проводили с использованием платформы MiSeq (Illumina), используя реагенты для получения парных прочтений длиной по 300 нуклеотидов. Для подготовки ДНК-библиотек для секвенирования использовали ультразвуковую фрагментацию ДНК и применяли набор NEBNext Ultra DNA Library Prep Kit for Illumina (New England Biolabs). Первичную обработку коротких нуклеотидных прочтений (fastq файлов) проводили с использованием программы Trimmomatic (<http://www.usadellab.org/cms/index.php?page=trimmomatic>) для удаления адаптеров и нуклеотидных прочтений низкого качества. Полученные файлы использовали для выравнивания на референсный геном *M. tuberculosis* H37Rv (NC_000962.3) для дальнейшего поиска геномных вариантов с использованием программы bowtie2, после чего для идентификации и аннотации нуклеотидных полиморфизмов использовали утилиты SAMtools (<http://samtools.sourceforge.net>). Онлайн-ресурсы Phyresse (<https://bioinf.fz-borstel.de/mchips/phyresse/>), TGS-TB (<https://gph.niid.go.jp/tgs-tb/>) и PhyTB (<http://pathogenseq.lshtm.ac.uk/phytblive/index.php>) использовали для дополнительной классификации секвенированных геномов.

Результаты исследования и обсуждение

Филогенетический анализ данных, полученных в результате полногеномного секвенирования, был проведен для коллекции штаммов, которая включала 19 штаммов *Beijing* (древние и современные сублинии), 11 штаммов *LAM* (сублинии *RD-Rio*, *LAM-RUS*), 2 штамма семейства *T* (сполитип SIT40) и 1 штамм семейства *S*. В результате филогенетического анализа геномных данных (4500 полиморфных нуклеотидных

Таблица 1

Распределение однонуклеотидных полиморфизмов, специфичных для подгрупп внутри генотипов *Beijing* и *LAM*

Функциональная категория (по <i>Tuberculist</i> EPFL)	<i>Beijing</i> , современная сублиния, B0-кластер	<i>Beijing</i> , современная сублиния, A0-кластер	<i>Beijing</i> , ранняя древняя сублиния	<i>LAM-RUS</i> , SIT252/266-кластер	<i>LAM-RUS</i> , SIT252	<i>LAM-RUS</i> , SIT266
Процессы, связанные с клеткой и клеточной стенкой	16	5	15	5	2	1
Консервативные гипотетические белки	18	2	2	5	1	1
Промежуточный метаболизм и дыхание	21	4	31	9	1	3
Липидный метаболизм	2	4	12	2	2	1
Информационные пути	2	3	5	4		2
Регуляторные белки	4	1		1		
Стабильные РНК				1		
Вирулентность, детоксификация, адаптация	3		10	1		
Всего снипов	66	19	78	28	6	8

Характеристика полиморфизмов, специфичных для кластера штаммов ранней древней сублинии генотипа *Beijing*

Ген	Позиция в геноме и мутация	Продукт гена (согласно https://mycobrowser.epfl.ch/)	Функциональная категория (согласно https://mycobrowser.epfl.ch/)
<i>dnaN</i>	3170A>G	ДНК полимеразы III	Информационные пути
<i>acpA</i>	36813A>G	Ацилпереносящий белок AcpA	Липидный метаболизм
<i>nrp</i>	114980C>A	Пептидсинтетаза Nrp	Липидный метаболизм
<i>gmhB</i>	138029C>A	D-глицеро-D-манно-гептоз 7-фосфаткиназа	Процессы, связанные с клеткой и клеточной стенкой
<i>oxcA</i>	143030C>A	Вероятная оксалил-СоА декарбоксилаза	Промежуточный метаболизм и дыхание
<i>yrbE1B</i>	198212C>A	Консервативный мембранный белок YrbE1B	Вирулентность, детоксификация, адаптация
<i>lprK</i>	204840C>A	Липопротеин Mse1E семейства Mse	Процессы, связанные с клеткой и клеточной стенкой
<i>mmpL11</i>	238686C>A	Вероятный консервативный трансмембранный транспортный белок MmpL11	Процессы, связанные с клеткой и клеточной стенкой
<i>varB25</i>	333363C>A	Возможный антитоксин VarB25	Вирулентность, детоксификация, адаптация
<i>rmlA</i>	398918C>A	dTDP-глюкозпирофосфорилаза	Промежуточный метаболизм и дыхание
<i>dnaJ1</i>	423459C>A	Вероятный белок-шаперон DnaJ1	Вирулентность, детоксификация, адаптация
<i>fadD30</i>	484395C>A	АМФ синтетаза жирных кислот	Липидный метаболизм
<i>thiG</i>	502406C>A	Белок биосинтеза тиазола	Промежуточный метаболизм и дыхание
<i>ctpH</i>	513669C>A	Возможная катион-транспортирующая АТФ-аза CtpH	Процессы, связанные с клеткой и клеточной стенкой
<i>mmpL2</i>	599670C>A	Вероятный консервативный трансмембранный транспортный белок MmpL2	Процессы, связанные с клеткой и клеточной стенкой
<i>menE</i>	634983C>A	O-сукцинил бензоат-СоА синтаза	Промежуточный метаболизм и дыхание
<i>cyp135B1</i>	660338C>A	Возможный цитохром P450 135B1	Промежуточный метаболизм и дыхание
<i>phoP</i>	851844C>A	транскрипционный регулятор PhoP	Регуляторные протеины
<i>adhB</i>	855199C>A	Возможная цинксодержащая алкогольдегидрогеназа НАД-зависимая AdhB	Промежуточный метаболизм и дыхание
<i>lpqU</i>	1143997C>A	вероятный консервативный липопротеин LpqU	Процессы, связанные с клеткой и клеточной стенкой
<i>coaA</i>	1219871C>A	Киназа пантотеновой кислоты	Промежуточный метаболизм и дыхание
<i>glpX</i>	1228027C>G	Фруктоз-1,6-бисфосфатаза GlpX	Промежуточный метаболизм и дыхание
<i>metE</i>	1259241C>G	Метионин синтаза	Промежуточный метаболизм и дыхание
<i>mmpL13b</i>	1273756C>G	Вероятный консервативный трансмембранный транспортный белок MmpL13b	Процессы, связанные с клеткой и клеточной стенкой
<i>fadD6</i>	1349375C>G	СоА синтетаза жирных кислот	Липидный метаболизм
<i>atpH</i>	1462302C>G	Дельта цепь АТФ-синтазы AtpH	Промежуточный метаболизм и дыхание
<i>rpoZ</i>	1565158C>G	ДНК-зависимая РНК-полимераза (цепь омега)	Информационные пути
<i>priA</i>	1578856C>G	Фактор репликации Y	Информационные пути
<i>wbbL2</i>	1720081C>G	Рамнозилтрансфераза WbbL2	Консервативные гипотетические белки
<i>cysA</i>	1930895C>T	D-серин/аланин/глицин транспортерный белок CysA	Процессы, связанные с клеткой и клеточной стенкой
<i>pknE</i>	1970117C>T	Трансмембранная серин-треонин киназа E	Регуляторные протеины
<i>tpx</i>	2183527C>T	тиолпероксидаза Tpx	Вирулентность, детоксификация, адаптация
<i>yrbE3B</i>	2209121C>T	Консервативный гипотетический интегральный мембранный белок YrbE3B	Вирулентность, детоксификация, адаптация
<i>ctpG</i>	2235800C>T	АТФ-аза G метал-катионного транспортера Р-типа	Процессы, связанные с клеткой и клеточной стенкой
<i>hspX</i>	2278544C>T	Белок теплового шока HspX	Вирулентность, детоксификация, адаптация
<i>acg</i>	2279152C>T	Консервативный белок Acg	Консервативные гипотетические белки

Продолжение табл. см. на стр. 47

Ген	Позиция в геноме и мутация	Продукт гена (согласно https://mycobrowser.epfl.ch/)	Функциональная категория (согласно https://mycobrowser.epfl.ch/)
<i>lipT</i>	2290732C>T	карбоксилэстераза LipT	Промежуточный метаболизм и дыхание
<i>mpa</i>	2376100C>T	микобактериальная протеасомная АТФ-аза	Процессы, связанные с клеткой и клеточной стенкой
<i>uspA</i>	2589661C>T	Сахар-транспортирующий мембранный белок-ABC-транспортер UspA	Процессы, связанные с клеткой и клеточной стенкой
<i>cysK1</i>	2609013C>T	О-ацетилсерин сульфгидролаза A	Промежуточный метаболизм и дыхание
<i>moeW</i>	2613117C>T	Белок биосинтеза молибдоптерина MoeW	Промежуточный метаболизм и дыхание
<i>amiA2</i>	2644643C>T	амидаза AmiA2	Промежуточный метаболизм и дыхание
<i>mbtJ</i>	2678523C>T	Возможная ацетилгидролаза MbtJ	Липидный метаболизм
<i>varC19</i>	2869016C>T	токсин VarC19	Вирулентность, детоксификация, адаптация
<i>relA</i>	2908252C>T	ГТФ-пирофосфокиназа RelA	Промежуточный метаболизм и дыхание
<i>speE</i>	2928883C>T	спермидинсинтаза SpeE	Промежуточный метаболизм и дыхание
<i>sthA</i>	3026472C>T	Растворимая пиридин трансгидрогеназа SthA	Промежуточный метаболизм и дыхание
<i>mgo</i>	3161688C>T	Малат-хинолон оксидоредуктаза Mgo	Промежуточный метаболизм и дыхание
<i>relF</i>	3177660C>T	атитоксин RelF	Вирулентность, детоксификация, адаптация
<i>viuB</i>	3204720C>T	Белок утилизации микобактерина ViuB	Промежуточный метаболизм и дыхание
<i>glnD</i>	3228926G>T	РП уридил-трансфераза	Промежуточный метаболизм и дыхание
<i>ppsA</i>	3245756G>T	поликетидсинтаза PpsA	Липидный метаболизм
<i>ppsD</i>	3266769G>T	поликетидсинтаза ppsD	Липидный метаболизм
<i>mas</i>	3281726G>T	Мембран-ассоциированная многофункциональная синтаза микосерозовой кислоты	Липидный метаболизм
<i>fadD22</i>	3299852G>T	Р-гидробензоил-АМФ-лигаза	Липидный метаболизм
<i>moaC1</i>	3479299G>T	протеин С биосинтеза молибденового кофактора	Промежуточный метаболизм и дыхание
<i>fbiA</i>	3640879T>C	Белок биосинтеза F420	Промежуточный метаболизм и дыхание
<i>pcd</i>	3674012T>C	дегидрогеназа пиперидин-6-карбоксилевой кислоты	Промежуточный метаболизм и дыхание
<i>lhr</i>	3679378T>C	АТФ-зависимая хеликаза	Информационные пути
<i>atsB</i>	3683721T>C	арилсульфатаза	Промежуточный метаболизм и дыхание
<i>moaX</i>	3709528T>C	белок слияния MoaD-MoaE	Промежуточный метаболизм и дыхание
<i>otsB2</i>	3787027T>C	Трегалоз-6-фосфат фосфатаза OtsB2	Вирулентность, детоксификация, адаптация
<i>amiD</i>	3789601T>C	ацилаза	Промежуточный метаболизм и дыхание
<i>eccD4</i>	3868372T>C	консервативный компонент EccD4 белка ESX	Процессы, связанные с клеткой и клеточной стенкой
<i>eccD4</i>	3869537T>C	консервативный компонент EccD4 белка ESX	Процессы, связанные с клеткой и клеточной стенкой
<i>truA</i>	3876244T>C	урацигидролаза	Информационные пути
<i>fadE31</i>	4003296T>C	ацил-СоА дегидролаза	Липидный метаболизм
<i>ask</i>	4153340T>C	аспартаткиназа	Промежуточный метаболизм и дыхание
<i>pks13</i>	4260414T>C	поликетидсинтаза	Липидный метаболизм
<i>fadD23</i>	4301135T>C	АМФ-синтаза жирных кислот	Липидный метаболизм
<i>glpQ1</i>	4315532T>G	глицеринфосфодиэфирфосфодиэстераза	Промежуточный метаболизм и дыхание
<i>gltD</i>	4330177T>G	L-глутамат синтаза	Промежуточный метаболизм и дыхание
<i>whiB6</i>	4338326T>G	Регуляторный белок WhiB	Регуляторные протеины
<i>espK</i>	4358125T>G	секреторный белок ESX-1	Процессы, связанные с клеткой и клеточной стенкой
<i>espB</i>	4360997T>G	Секретируемый субстрат протеина ESX-1	Процессы, связанные с клеткой и клеточной стенкой
<i>eccC2</i>	4376444T>G	Белок ESX-2 VII типа	Процессы, связанные с клеткой и клеточной стенкой
<i>trxB2</i>	4401802T>G	Тиоредоксин редуктаза	Промежуточный метаболизм и дыхание

позиций), были выделены кластеры штаммов и определены их специфические однонуклеотидные полиморфизмы (ОНП, SNP). В частности, внутри сублинии *LAM-RUS* был выявлен эволюционно «молодой» кластер, включающий штаммы сполиготипов SIT252 и SIT266. Внутри штаммов семейства *Beijing* были выделены ветви современной (субтипы B0 и 94-32) и древней сублиний.

Картирование коротких геномных прочтений на референсный геном штамма H37Rv позволило определить наличие крупной делеции *RD181* в геномах штаммов древних *Beijing*, т.е., они относились к несколько более поздней подгруппе ранних древних *Beijing*, согласно модифицированной эволюционной схеме для генотипа *Beijing* [14]. Анализ других штаммов выявил крупные геномные делеции для штаммов *LAM*, разделяющие их на две основные сублинии *RD174/RD-Rio* и *RD115*.

Соотнесение всего массива SNP, выявленных в изученных штаммах с известными генами, играющими роль в патогенезе и адаптации возбудителя туберкулёза, выявило кластер-специфические однонуклеотидные полиморфизмы для групп внутри семейств *Beijing* и *LAM* (табл. 1).

Наибольшее количество вариаций выявлено в штаммах эволюционно «молодых» генетического кластера *Beijing* B0/W148 и кластера *LAM-RUS* (сполиготипы SIT252, SIT266), но также и у штаммов древней сублинии *Beijing*. Больше количество ОНП (78) штаммов древней сублинии *Beijing* (табл. 2) может быть связано с их более длительной эволюцией. Также это может коррелировать с недавним выявлением высококорезистентного кластера ранней древней сублинии *Beijing* в Омской области Сибири [9] и описанием таких штаммов, выделенных от больных туберкулёзным спондилитом на Северо-Западе России [15].

Среди генов, содержащих кластер-специфические полиморфизмы, наиболее широко были представлены категории генов «процессы, связанные с клеткой и клеточной стенкой» и «промежуточный метаболизм и дыхание». Вариации в генах категории «вирулентность» встречались гораздо реже и только в штаммах *Beijing* B0/W148, древней сублинии *Beijing* и *LAM-RUS* (SIT252/266). Накопление определенного набора мутаций в патогенетически значимых локусах (и протеинах) может иметь отношение к особенностям «успешности» таких штаммов, представляющих эпидемические клоны возбудителя туберкулёза.

Анализ других типов мутаций (помимо однонуклеотидных замен) выявил потенциальные мутации сдвига рамки, как результат инсерций/делеций (инделов). Полиморфизмы, связанные с короткими инделами, чаще встречались в группах генов PE-PGRS, PPE (мультигенные полиморфные семейства с точно не определенными функциями) [16] и/или в межгенных участках. Анализ полиморфизма, ведущего к возможному сдвигу рамки считывания в результате коротких

(1-5 нуклеотидов) делеций или инсерций показал относительное преобладание таких мутаций у штаммов геноварианта *Beijing* B0/W148 (гены Rv1995, Rv0888, Rv1028c/kdpD, Rv3434c) и только двух мутаций у геноварианта *Beijing* 94-32 (Rv0888, Rv1971/mce3F). Особое внимание обращают на себя гены kdpD и mce3F. Ген kdpD относится к категории регуляторных белков; согласно ресурсу Tuberculist, он является членом двухкомпонентной регуляторной системы KDPD/KDPE и принадлежит к группе белков универсального стресса, ранее была сформулирована гипотеза о его роли в эволюции генотипа W148 [17]. Ген mce3F относится к функциональной категории «вирулентность, детоксификация, адаптация», также полагают, что белки Mce связаны с проникновением в клетку хозяина [18].

Заключение

Таким образом, для эмерджентных, актуально или потенциально эпидемических вариантов *M. tuberculosis* генотипов *LAM* и *Beijing*, выявлены как филогенетически нейтральные, диагностические полиморфизмы, так и мутации в ряде генов вирулентности, адаптации, биосинтеза клеточной стенки, дыхания и липидного обмена.

Ввиду значительного количества полиморфизмов у кластера российских штаммов древней сублинии генотипа *Beijing*, в том числе, в генах вирулентности, не исключено их эпистатическое взаимодействие, ведущее в повышенной трансмиссивности и вирулентности, но для подтверждения этой гипотезы необходимы дальнейшие исследования с использованием экспериментальных моделей.

Список литературы / References

1. Brosch R., Gordon SV., Marmiesse M., Brodin P., Buchrieser C., Eiglmeier K., Garnier T., Gutierrez C., Hewinson G., Kremer K., Parsons LM., Pym AS., Samper S., van Soolingen D., Cole S.T. A new evolutionary scenario for the Mycobacterium tuberculosis complex. *Proc. Natl. Acad. Sci. USA*. 2002; 99(6): 3684-3689. DOI: 10.1073/pnas.052548299
2. Hanekom M., Gey van Pittius N.C., McEvoy C., Victor T.C., Van Helden P.D., Warren R.M. Mycobacterium tuberculosis Beijing genotype: a template for success. *Tuberculosis (Edinb)*. 2011; 91(6): 510-523. DOI: 10.1016/j.tube.2011.07.005
3. Parwati I., van Crevel R., van Soolingen D. Possible underlying mechanisms for successful emergence of the Mycobacterium tuberculosis Beijing genotype strains. *Lancet Infect Dis*. 2010; 10(2): 103-111. DOI: 10.1016/S1473-3099(09)70330-5
4. Mokrousov I. Insights into the origin, emergence, and current spread of a successful Russian clone of Mycobacterium tuberculosis. *Clin. Microbiol. Rev.* 2013; 26(2): 342-360. DOI: 10.1128/CMR.00087-12
5. Mokrousov I. Mycobacterium tuberculosis phylogeography in the context of human migration and pathogen's pathobiology: Insights from Beijing and Ural families. *Tuberculosis (Edinb)*. 2015; 95 Suppl 1: 167-176. DOI: 10.1016/j.tube.2015.02.031
6. Iwamoto T., Yoshida S., Suzuki K., Wada T. Population structure analysis of the Mycobacterium tuberculosis Beijing family indicates an association between certain sublineages and multidrug resistance. *Antimicrob. Agents Chemother.* 2008; 52: 3805-3809. DOI: 10.1128/AAC.00579-08
7. Maeda S., Hang N.T., Lien L.T., Thuong P.H., Hung N.V., Hoang N.P., Cuong V.C., Hijikata M., Sakurada S., Keicho N. Mycobac-

- terium tuberculosis strains spreading in Hanoi, Vietnam: Beijing sublineages, genotypes, drug susceptibility patterns, and host factors. *Tuberculosis (Edinb)*. 2014; 94: 649-956. DOI: 10.1016/j.tube.2014.09.005
8. Shamputa I.C., Lee J., Allix-Béguet C., Cho E.J., Lee J.I., Rajan V., Lee E.G., Min J.H., Carroll M.W., Goldfeder L.C., Kim J.H., Kang H.S., Hwang S., Eum S.Y., Park S.K., Lee H., Supply P., Cho S.N., Via L.E., Barry C.E. 3rd. Genetic diversity of Mycobacterium tuberculosis isolates from a tertiary care tuberculosis hospital in South Korea. *J. Clin. Microbiol.* 2010; 48: 387-394. DOI: 10.1128/JCM.02167-09
 9. Mokrousov I., Vyazovaya A., Pasechnik O., Gerasimova A., Dymova M., Chernyaeva E., Tatarintseva M., Stasenko V. Early ancient sublineages of Mycobacterium tuberculosis Beijing genotype: unexpected clues from phylogenomics of the pathogen and human history. *Clin. Microbiol. Infect.* 2019; 25(8): 1039.e1-1039.e6. DOI: 10.1016/j.cmi.2018.11.024
 10. van Embden J., Cave M., Crawford J., Dale J.W., Eisenach K.D., Gicquel B., Small P. Strain identification on Mycobacterium tuberculosis by DNA fingerprinting: recommendations for a standardized methodology. *J. Clin. Microbiol.* 1993; 31: 406-409.
 11. Kamerbeek J., Schouls L., Kolk A., van Agterveld M., van Soolingen D., Kuijper S., Bunschoten A., Molhuizen H., Shaw R., Goyal M., van Embden J. Simultaneous detection and strain differentiation of Mycobacterium tuberculosis for diagnosis and epidemiology. *J. Clin. Microbiol.* 1997; 35(4): 907-14.
 12. Mokrousov I., Rastogi N. Spacer-Based Macroarrays for CRISPR Genotyping. *Methods Mol Biol.* 2015; 1311: 111-131. DOI: 10.1007/978-1-4939-2687-9_7
 13. Couvin D., David A., Zozio T., Rastogi N. Macro-geographical specificities of the prevailing tuberculosis epidemic as seen through SITVIT2., an updated version of the Mycobacterium tuberculosis genotyping database. *Infect. Genet. Evol.* 2019; 72: 31-43. DOI: 10.1016/j.meegid.2018.12.030
 14. Yin Q.Q., Liu H.C., Jiao W.W., Li Q.J., Han R., Tian J.L., Liu Z.G., Zhao X.Q., Li Y.J., Wan K.L., Shen A.D., Mokrousov I. Evolutionary History and Ongoing Transmission of Phylogenetic Sublineages of Mycobacterium tuberculosis Beijing Genotype in China. *Sci. Rep.* 2016; 6: 34353. DOI: 10.1038/srep34353
 15. Vyazovaya A., Mokrousov I., Solovieva N., Mushkin A., Manicheva O., Vishnevsky B., Zhuravlev V., Narvskaya O. Tuberculous spondylitis in Russia and prominent role of multidrug-resistant clone Mycobacterium tuberculosis Beijing B0/W148. *Antimicrob. Agents Chemother.* 2015; 59(4): 2349-2357. DOI: 10.1128/AAC.04221-14
 16. Brennan M.J. The Enigmatic PE/PPE Multigene Family of Mycobacteria and Tuberculosis Vaccination. *Infect. Immun.* 2017; 85(6): e00969-16. DOI: 10.1128/IAI.00969-16
 17. Merker M., Blin C., Mona S., Duforet-Freboung N., Lecher S., Willery E., Blum M.G., Rüsche-Gerdes S., Mokrousov I., Aleksic E., Allix-Béguet C., Antierens A., Augustynowicz-Kopeć E., Ballif M., Barletta F., Beck H.P., Barry C.E. 3rd., Bonnet M., Borroni E., Campos-Herrero I., Cirillo D., Cox H., Crowe S., Crudu V., Diel R., Drobniewski F., Fauville-Dufaux M., Gagneux S., Ghebremichael S., Hanekom M., Hoffner S., Jiao W.W., Kalon S., Kohl T.A., Kontsevaya I., Lillebæk T., Maeda S., Nikolayevskyy V., Rasmussen M., Rastogi N., Samper S., Sanchez-Padilla E., Savic B., Shamputa I.C., Shen A., Sng L.H., Stakenas P., Toit K., Varaine F., Vukovic D., Wahl C., Warren R., Supply P., Niemann S., Wirth T. Evolutionary history and global spread of the Mycobacterium tuberculosis Beijing lineage. *Nat. Genet.* 2015; 47(3): 242-249. DOI: 10.1038/ng.3195
 18. Zhang F., Xie J.P. Mammalian cell entry gene family of Mycobacterium tuberculosis. *Mol Cell Biochem.* 2011; 352(1-2): 1-10. DOI: 10.1007/s11010-011-0733-5

Сведения об авторах:

Мокроусов Игорь Владиславович — доктор биологических наук, заведующий лабораторией молекулярной эпидемиологии и эволюционной генетики Федерального бюджетного учреждения науки «Санкт-Петербургский научно-исследовательский институт эпидемиологии и микробиологии имени Пастера» Федеральной службы по надзору в сфере защиты прав потребителей и благополучия человека; <https://orcid.org/0000-0001-5924-0576>

Черняева Екатерина Николаевна — кандидат биологических наук, старший научный сотрудник лаборатории молекулярной эпидемиологии и эволюционной генетики Федерального бюджетного учреждения науки «Санкт-Петербургский научно-исследовательский институт эпидемиологии и микробиологии имени Пастера» Федеральной службы по надзору в сфере защиты прав потребителей и благополучия человека; <https://orcid.org/0000-0002-8306-3466>

Вязовая Анна Александровна — кандидат биологических наук, старший научный сотрудник лаборатории молекулярной эпидемиологии и эволюционной генетики Федерального бюджетного учреждения науки «Санкт-Петербургский научно-исследовательский институт эпидемиологии и микробиологии имени Пастера» Федеральной службы по надзору в сфере защиты прав потребителей и благополучия человека; <https://orcid.org/0000-0001-9140-8957>

Журавлев Вячеслав Юрьевич — кандидат биологических наук, руководитель направления «Лабораторная диагностика» Федерального государственного бюджетного научного учреждения «Санкт-Петербургский Научно-исследовательский институт фтизиопульмонологии» Министерства Здравоохранения Российской Федерации; <https://orcid.org/0000-0001-6906-6225>