

8. Asghar H., Diop O.M., Weldegebriel G. et al. Environmental surveillance for polioviruses in the Global Polio Eradication Initiative. *J. Infect. Dis.* 2014, 210 (1): 294-303.
9. Aylward R.B., Sutter R.W., Cochi S.T. et al. Risk Management in a Polio-Free World. *Risk Analysis.* 2006, 26 (6): 1441-1448.
10. Burns C.C., Dior O.M., Sutter R.W. et al. Vaccine-Derived Polioviruses. *J. Infect. Dis.* 2014, 210: S283-293.
11. Cochi S.L., Freeman A., Guirguis S. et al. Global polio eradication initiative: Lessons learned and legacy. *J. Infect. Dis.* 2014, 210 (1): S540-546.
12. Shrivastava S.R., Shrivastava P.S., Ramasamy J. Recommended Strategies to Move Closer Toward the Global Eradication of Polio: International Health Regulations. *Int. J. Preventive Medicine* 2016, 7-17.

Поступила 04.12.17

Контактная информация: Фельдблум Ирина Викторовна, д.м.н., проф.,
614990, Пермь, ул. Петропавловская, 26, р.т. (342)218-16-68

© КОЛЛЕКТИВ АВТОРОВ, 2018

*И.В.Мокроусов¹, О.А.Пасечник², А.А.Вязовая¹,
А.И.Блох², Е.Н.Черняева³, В.Л.Стасенко²*

О ВАЖНОСТИ ПРИМЕНЕНИЯ ЭВОЛЮЦИОННО НАДЕЖНЫХ МАРКЕРОВ ДЛЯ ДЕТЕКЦИИ ШТАММОВ MYCOBACTERIUM TUBERCULOSIS ГЕНЕТИ- ЧЕСКОГО СЕМЕЙСТВА LAM

¹НИИ эпидемиологии и микробиологии им. Пастера, Санкт-Петербург; ²Омский государственный медицинский университет; ³Санкт-Петербургский государственный университет

Цель. Клиническая и эпидемиологическая значимость генетического семейства LAM (Latin American Mediterranean) *Mycobacterium tuberculosis* определяет важность корректной детекции штаммов LAM. В настоящем исследовании комплекс молекулярных методов был использован для анализа штаммов LAM в популяции *M. tuberculosis* в Омской области Западной Сибири, для которой характерен высокий уровень заболеваемости резистентным туберкулезом. *Материалы и методы.* Изученная коллекция включала 207 штаммов *M. tuberculosis*, выделенных в Омской области в 2015 — 2016 гг. Методы исследования включали сполиготипирование, анализ полиморфизма Rv0129c 309G>A, специфического для семейства LAM, полногеномное секвенирование с последующим биоинформационным анализом. *Результаты.* В результате сравнения полученных профилей CRISPR-сполиготипирования с международной базой данных SITVIT_WEB 11 штаммов (5,3%) было отнесено к генотипу LAM. В то же время, в результате анализа филогенетического SNP в гене Rv0129c к генотипу LAM было отнесено 30 изолятов (14,5%). Для четырех изолятов, представляющих разные типы сполигопрофилей, было проведено полногеномное секвенирование. *Заключение.* Полученные результаты показывают ограниченность правил принятия решения, имплементированных в SITVIT_WEB для определения семейства LAM для штаммов с протяженными блоками делетированных спейсеров в локусе CRISPR или усеченными профилями сполготипирования. Для детекции штаммов LAM может быть рекомендован подход, включающий (1) первичное сполиготипирование, сравнение с SITVIT_WEB и обязательное уточнение интерпретации профилей в свете экспертного знания; (2) детекцию LAM-специфического SNP (например, с помощью PCR-RFLP).

Журн. микробиол., 2018, № 3, С. 60—66

Ключевые слова: *Mycobacterium tuberculosis*, эволюция, полногеномное секвенирование, сполиготипирование, генотип LAM (Latin-American Mediterranean)

I.V.Mokrousov¹, O.A.Pasechnik², A.A.Vyazovaya¹,
A.I.Blokh², E.N.Chernyaeva³, V.L.Stasenko²

ON IMPORTANCE OF USING EVOLUTIONARILY ROBUST MARKERS FOR DETECTION OF *MYCOBACTERIUM TUBERCULOSIS* STRAINS OF LAM GENETIC FAMILY

¹Pasteur St. Petersburg State Institute of Epidemiology and Microbiology; ²Omsk State Medical University; ³St. Petersburg State University, Russia

Aim. The clinical and epidemiological significance of the Latin American Mediterranean (LAM) genetic family of *Mycobacterium tuberculosis* determines the importance of the correct detection of LAM strains. In this study, a complex of molecular methods was used to analyze LAM strains in the population of *M. tuberculosis* in the Omsk region of Western Siberia, which is characterized by a high incidence of drug-resistant tuberculosis. *Materials and methods.* The collection included 207 strains of *M. tuberculosis*, isolated in the Omsk region in 2015 — 2016. The strains were subjected to spoligotyping, analysis of LAM-specific SNP Rv0129c 309G>A, and whole genome sequencing followed by bioinformatics analysis. *Results.* A comparison of the obtained CRISPR-spoligotyping profiles with the international SITVIT_WEB database, assigned 11 strains (5.3%) to the LAM genotype. At the same time, based on analysis of phylogenetic SNP in the gene Rv0129c, 30 isolates (14.5%) were assigned to LAM. Whole genome sequencing was performed for 4 isolates with different spoligotyping profiles. *Conclusion.* The results of this study show the limited utility of the decision rules implemented in SITVIT_WEB to define LAM family for isolates with long deleted blocks of spacers or abridged spoligoprofiles. The following approach can be recommended for detection of LAM isolates (1) primary spoligotyping, comparison with SITVIT_WEB, and mandatory interpretation in the light of expert knowledge; (2) detection of LAM-specific SNP (e.g., using PCR-RFLP).

Zh. Mikrobiol. (Moscow), 2018, No. 3, P. 60—66

Key words: *Mycobacterium tuberculosis*, evolution, whole genome sequencing, spoligotyping, LAM (Latin-American Mediterranean) genotype

ВВЕДЕНИЕ

Характерной особенностью вида *Mycobacterium tuberculosis* является строго клональная структура популяции как следствие отсутствия горизонтального генетического переноса. Помимо общебиологического смысла такого хода эволюции (обсуждение которого выходит за рамки данного сообщения), этот факт имеет и прикладное, клиническое значение применительно к реализации региональных программ борьбы с туберкулезом. В настоящее время общепризнано, что отдельные генотипы (генетические семейства, сублинии, клональные кластеры) *M. tuberculosis* отмечены повышенной вирулентностью, трансмиссивностью, или способностью быстрого развития устойчивости к противотуберкулезным препаратам. Структуру популяции возбудителя туберкулеза в России и других странах постсоветского пространства часто воспринимают через призму доминирования резистентных штаммов генетического семейства Beijing. В целом обоснованный, такой подход упрощает реальность как в количественном, так и в качественном аспектах. В том, что касается собственно структуры популяции, штаммы генетического семейства LAM (Latin American Mediterranean) также являются значимым компонентом популяции *M. tuberculosis*, особенно в Европейской части бывшего СССР, и составляют до 30 — 40% локальных популяций, например, в центральной России и Белоруссии [7, 15]. Более того, рассматривая патобиологические

особенности штаммов, нужно учитывать, что эмерджентные, потенциально или актуально эпидемические геноварианты отмечены не только среди представителей семейства Beijing, но также и LAM и еще одного (недостаточно оцененного) генотипа Ural.

В свою очередь, это подчеркивает необходимость применения четких, эволюционно надежных молекулярных маркеров для корректного определения таких генетических групп, имеющих клиническую и/или эпидемиологическую значимость. Эволюционно значимые геногруппы могут быть выявлены: (1) на основе детекции валидированных специфических SNP (single nucleotide polymorphism) или геномных делеций, представляющих уникальные и однонаправленные эволюционные события; (2) посредством филогенетического анализа (кластеризации) на основе множества независимых и нейтральных эволюционных маркеров, к которым могут быть отнесены полногеномные SNP или локусы VNTR (variable number of tandem repeats).

Генетическое семейство LAM *M. tuberculosis* впервые было описано в 2001 году в результате филогенетического анализа относительно глобальной коллекции сполиготипов [14]. Прототипным сполготипом LAM является SIT42, в профиле которого отсутствуют сигналы 21-24 и 33-36. Последующее применение эволюционно надежных маркеров подтвердило реальность LAM и позволило более адекватно определить его филогенетические границы. Также можно отметить, что LAM входит в крупную Евро-Американскую линию, хотя и определенную на основании SNP-анализа, но также имеющую характерный маркер на основании сполиготипирования, а именно отсутствие сигналов 33-36.

В наших предыдущих работах на коллекциях штаммов из различных регионов бывшего СССР (Северо-Запад РФ, Белоруссия, Казахстан) и Китая были валидированы ранее описанные специфические SNP в генах Rv0129c и Rv3062 для детекции штаммов LAM [8 — 11]. В настоящем исследовании мы применили комплекс молекулярных маркеров для анализа штаммов LAM в популяции *M. tuberculosis* в Омской области Западной Сибири, для которой характерен высокий уровень заболеваемости резистентным туберкулезом [1, 2], и расширили их верификацию с использованием полногеномного секвенирования следующего поколения: WGS (whole genome sequencing)/NGS (next generation sequencing).

МАТЕРИАЛЫ И МЕТОДЫ

Штаммы *M. tuberculosis* были выделены от больных туберкулезом легких, постоянно проживающих в Омской области. ДНК выделяли стандартным методом с использованием SDS, протеиназы K и цетилтриметиламмонийбромида для клеточного лизиса.

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. Профили сполиготипирования сравнивали с международной базой данных SITVIT_WEB (http://www.pasteur-guadeloupe.fr:8081/SITVIT_ONLINE/).

Анализ однонуклеотидного полиморфизма в гене Rv0129c 309G>A, специфического для семейства LAM, проводили методом ПЦР-ПДРФ [11].

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. Нуклеотидные прочтения выравнивали на референсный геном с использованием инструмента 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.

Дополнительно штаммы LAM были типированы на наличие специфической инсерции IS6110 в определенной позиции в гене *plcA* (позиция 2630571 в геноме штамма H37Rv, номер доступа 00962.3 в GenBank), которая является маркером подгруппы внутри LAM, ранее названной LAM-RUS [6].

РЕЗУЛЬТАТЫ

Изученная коллекция включала 207 штаммов *M. tuberculosis*, выделенных в Омской области в 2015-2016 гг. В результате сравнения полученных профилей сполиготипирования с базой данных SITVIT_WEB 11 штаммов (5,3%) было отнесено к генотипу LAM. В то же время, в результате анализа филогенетического SNP в гене *Rv0129c* к генотипу LAM было отнесено 30 изолятов (14,5%) (табл.). Дополнительно все 30 изолятов LAM были определены как LAM-RUS.

Для четырех изолятов, представляющих разные типы профилей внутри изученной выборки LAM, проведено полногеномное секвенирование с использованием технологии следующего поколения. Таким образом были изучены: штамм SIT42 (классический прототипный профиль для семейства LAM), штамм SIT254 с протяженным блоком делетированных спейсеров, штаммы SIT1451 и SIT_new с усеченными профилями гибридизации при сполиготипировании (табл.). Сравнение с базой сполиготипов SITVIT_WEB корректно отнесло первый из штаммов (SIT42) к LAM, в то время как штамм SIT254 был отнесен к семейству T5-RUS1, а для двух последних штаммов их филогенетический статус в SITVIT_WEB определен не был.

Для указанных 4 штаммов проведено полногеномное секвенирование и полученные файлы fastq были загружены в нуклеотидный архив NCBI, проект PRJNA401339. В результате выравнивания на референсный геном *M. tuberculosis* H37Rv для каждого штамма получен перечень геномных вариантов, включающий однонуклеотидные полиморфизмы и короткие инсерции и делеции (vcf файлы). Файлы fastq и vcf были проанализированы с использованием онлайн-ресурсов Phyresse, PhyTB и TGS-TB, с помощью которых была подтверждена принадлежность всех 4 штаммов к LAM. Кроме того, следует отметить в изученной омской выборке штамм с профилем SIT264 с протяженным блоком из 12 делетированных сигналов, отнесенный к семейству T1 в базе SITVIT_WEB. Однако ранее нами была показана принадлежность штам-

M. tuberculosis не является исключением. Подход на основе WGS/NGS позволяет выявить незначительные различия близкородственных геновариантов и успешно применяется как для филогеномных, так и для молекулярно-эпидемиологических исследований. Противоположным с точки зрения дискриминирующей способности является классический метод сполиготипирования. Практически единственным достоинством метода является его техническая простота и, следовательно, возможность быстрого анализа больших коллекций, а также наличие больших баз данных, собранных в результате 20 лет исследований. Прежде всего следует отметить глобальную базу данных SITVIT_WEB, созданную в Институте Пастера Гваделупы [5]. Последняя (все еще неопубликованная) версия этой базы (SITVIT2) содержит данные по более чем 110 000 изолятам, согласно диссертации D. Couvin [4]. Известными и серьезными ограничениями метода сполиготипирования являются как универсальные для всех бактерий (единый локус CRISPR, следовательно, не независимая эволюция спейсеров и т.о. неадекватность филогенетических построений на основе сполиготипов), так и специфические для *M. tuberculosis* (возможная гомоплазия как результат конвергентной эволюции, неопределенность интерпретации протяженных блоков делетированных сигналов). Глобальной проблемой анализа сполигопрофилей является и догматическое, некритическое восприятие крупных электронных ресурсов и баз данных.

Определение семейства LAM *M. tuberculosis* на основании сполиготипирования (в особенности профилей с протяженными блоками делетированных сигналов, например, SIT254, SIT264, SIT1451) имеет два существенных недостатка, как показано в настоящем и ряде предыдущих исследований в странах бывшего СССР. Во-первых, наличие усеченных профилей сполиготипирования (SIT1451) или профилей с длинными блоками делетированных сигналов (SIT264, SIT254) критически уменьшает количество признаков, пригодных для анализа, и не позволяет делать вывод о генотипе (семействе) штамма даже если алгоритм, применяемый в SITVIT_WEB, и позволяет (хотя и ошибочно) отнести такой штамм, например SIT254 или SIT264, к какой-либо подгруппе внутри семейства T (прежде всего к т.н. T5-RUS1, куда отнесены эти и близкие им производные сполиготипы в SITVIT_WEB). Во-вторых, сполиготип SIT803 представляет пример гомоплазии сполиготипов как результат конвергентной эволюции локуса CRISPR *M. tuberculosis*. В частности, как показано в данном исследовании (на основании полногеномного анализа) и в опубликованных работах по Северо-Западу России, Казахстану и Грузии (на основании анализа LAM-специфических SNP и мультилокусного VNTR-анализа) [11 — 13], штаммы SIT803 определенно относятся к семейству LAM. В то же время, анализ китайских штаммов с таким сполиготипом не выявил у них LAM-специфического SNP [9]. На построенной нами дендрограмме профилей 24-MIRU-VNTR 186 референс-штаммов *M. tuberculosis* разных генотипов (MIRU-VNTRplus.org) и 259 типов глобальной коллекции LAM [10] китайские штаммы SIT803 занимали положение далеко за пределами ветви LAM и располагались на периферии ветви семейства S (древо не показано). Все это подчеркивает необходимость осторожности в интерпретации сполигопрофилей с небольшим количеством сигналов гибридизации.

В то же время, даже в настоящее время все более широкого применения полногеномного анализа полезно сопоставлять новые данные со все еще применяемыми старыми маркерами и схемами. Не все лаборатории, особенно в странах с ограниченными ресурсами, могут использовать NGS в качестве

рутинного подхода. Конкретно для детекции штаммов LAM может быть рекомендован подход, включающий (1) первичное сполиготипирование, сравнение с SITVIT_WEB и обязательное уточнение интерпретации профилей в свете экспертного знания; (2) детекцию LAM-специфического SNP (например, с помощью PCR-RFLP).

Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 17-04-00367 «Популяция Mycobacterium tuberculosis в регионе Западной Сибири: актуальная молекулярная эпидемиология в контексте макроэволюционной реконструкции».

ЛИТЕРАТУРА

1. Пасечник О.А., Дымова М.А., Стасенко В.Л., Татаринцева М.П., Колесникова Л.П., Ляпина Е.С. Генетическое разнообразие лекарственно-устойчивых штаммов Mycobacterium tuberculosis в Омской области. Туберкулез и болезни легких. 2017, 7: 33-39.
2. Пасечник О.А., Руднева С.Н., Татаринцева М.П. Динамика эпидемиологических показателей по туберкулезу в Омской области. Туберкулез и болезни легких. 2015, 5: 139-140.
3. Barbosa C.B., Lazzarini L.C., Elias A.R. et al. Tuberculosis caused by RDRio Mycobacterium tuberculosis is not associated with differential clinical features. Int. J. Tuberc. Lung Dis. 2012, 16: 1377-1382.
4. Couvin D. Development and management of a global database of circulating genotypes of tubercle bacilli: Molecular methods and web-based tools for mapping, understanding and controlling the epidemic. PhD thesis, Université des Antilles et de la Guyane. Pointe-a-Pitre, Guadeloupe, 2014.
5. Demay C., Liens B., Burguière T. et al. SITVITWEB — a publicly available international multimer database for studying Mycobacterium tuberculosis genetic diversity and molecular epidemiology. Infect. Genet. Evol. 2012, 12 (4): 755-766.
6. Dubiley S., Kirillov E., Ignatova A. et al. Molecular characteristics of the Mycobacterium tuberculosis LAM-RUS family prevalent in Central Russia. J. Clin. Microbiol. 2007, 45 (12): 4036-4038.
7. Ignatova A., Dubiley S., Stepanshina V. et al. 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, 55 (10): 1413-1418.
8. Mokrousov I., Chernyaeva E., Vyazovaya A. et al. Next generation sequencing of Mycobacterium tuberculosis. Emerg. Infect. Dis. 2016, 22 (6): 1127-1129.
9. Mokrousov I., Jiao W.W., Wan K. et al. Stranger in a strange land: Ibero-American strain of Mycobacterium tuberculosis in Tibet, China. Infect. Genet. Evol. 2014, 26: 323-326.
10. Mokrousov I., Vyazovaya A., Iwamoto T. et al. Latin-American-Mediterranean lineage of Mycobacterium tuberculosis: Human traces across pathogen's phylogeography. Mol. Phylogenet. Evol. 2016, 99: 133-143.
11. Mokrousov I., Vyazovaya A., Narvskaya O. Mycobacterium tuberculosis Latin-American Mediterranean family and its sublineages: in the light of evolutionary robust markers. J. Bacteriol. 2014, 196 (10): 1833-1841.
12. Niemann S., Diel R., Khechinashvili G. et al. Mycobacterium tuberculosis Beijing lineage favors the spread of multidrug-resistant tuberculosis in the Republic of Georgia. J. Clin. Microbiol. 2010, 48 (10): 3544-3550.
13. Skiba Y., Mokrousov I., Ismagulova G. et al. Molecular snapshot of Mycobacterium tuberculosis population in Kazakhstan: a country-wide study. Tuberculosis. 2015, 95 (5): 538-546.
14. Sola C., Filliol I., Legrand E. et al. Mycobacterium tuberculosis phylogeny reconstruction based on combined numerical analysis with IS1081, IS6110, VNTR, and DR-based spoligo-typing suggests the existence of two new phylogeographical clades. J. Mol. Evol. 2001, 53 (6): 680-689.
15. Zalutskaya A., Wijkander M., Jureen P. et al. Multidrug-resistant Mycobacterium tuberculosis caused by the Beijing genotype and a specific T1 genotype clone (SIT No. 266) is widely transmitted in Minsk. Int. J. Mycobacteriol. 2013, 2: 194-198.

Поступила 25.10.17

Контактная информация: Мокроусов Игорь Владиславович, д.б.н.,
197101, Санкт-Петербург, ул. Мира, 14, р.т. (812)233-21-49