

УДК 577.21:579.873.21:579.258

# Транскриптом *Mycobacterium tuberculosis* при заражении мышей с разной генетической чувствительностью к инфекции

Т. А. Скворцов<sup>1</sup>, Д. В. Игнатов<sup>1</sup>, К. Б. Майоров<sup>2</sup>, А. С. Апт<sup>2</sup>, Т. Л. Ажикина<sup>1</sup><sup>1</sup>Институт биоорганической химии им. акад. М.М. Шенякина и Ю.А. Овчинникова РАН, 117997, Москва, ул. Миклухо-Маклая, 16/10<sup>2</sup>Центральный научно-исследовательский институт туберкулеза РАМН, 107564, Москва, Яузская аллея, 2

\*E-mail: timofey@ibch.ru

Поступила в редакцию 18.12.2012

**РЕФЕРАТ** Изучение экспрессии генов с помощью массированного секвенирования в настоящее время находит широкое применение в различных областях биологии, включая микробиологию. В данной работе приведены результаты анализа полногеномной экспрессии *Mycobacterium tuberculosis* при персистенции в тканях легкого инфицированного хозяина. Мышей двух линий с генетически различной чувствительностью к туберкулезу инфицировали патогенными бактериями *M. tuberculosis*. Библиотеки кДНК микобактерий, полученные из инфицированных тканей, были подвергнуты массированному секвенированию и сравнительному анализу. Установлено, что структура транскриптома *M. tuberculosis* при развитии инфекции указывает на активизацию липидного метаболизма, метаболизма аминокислот, переход к анаэробному дыханию, повышение экспрессии факторов модуляции иммунного ответа. Важный результат нашей работы – выделение 209 генов, уровень экспрессии которых повышался в ходе развития инфекционного процесса в организмах хозяев обеих линий – Commonly Upregulated Genes (CUG). Наибольший интерес представляют гены, относящиеся к функциональным категориям липидного метаболизма и клеточной стенки и клеточных процессов. Мы полагаем, что продукты этих генов необходимы *M. tuberculosis* для преодоления иммунного ответа организма хозяина и тем самым представляют потенциальные мишени для разработки лекарственных средств.

**КЛЮЧЕВЫЕ СЛОВА** *Mycobacterium tuberculosis*, транскриптом *in vivo*, Commonly Upregulated Genes, RNA-seq, туберкулез.

**СПИСОК СОКРАЩЕНИЙ** CC – клонирование идентичных последовательностей (Coincidence Cloning); CUG – Commonly Upregulated Genes; PDIM – димикоцерозат фтиоцерола (phthiocerol dimycocerosate).

## ВВЕДЕНИЕ

Эксперименты, в ходе которых экспрессия организма изучается на полногеномном уровне, генерируют гигантские объемы информации и имеют практически безграничные возможности применения. К примеру, данные, полученные в подобных экспериментах, могут использоваться для оценки эффективности антибиотиков или изучения изменений бактериального метаболизма в ходе инфекционного процесса. Среди инфекций, вызываемых бактериальными патогенами, туберкулез занимает первое место по уровню смертности, унося каждый год около 1.5 млн жизней. Неудивительно, что первое исследование транскриптома его возбудителя, *Mycobacterium tuberculosis*, было проведено в течение года с момента опубликования полной нуклеотидной последовательности

генома этой бактерии [1, 2]. Уже через 5 лет появилось значительное число публикаций, описывающих результаты использования микроэрреев для транскриптомного анализа микобактерий в различных условиях [3, 4]. Тем не менее, хотя технологии микроэрреев и массированного секвенирования стали быстро применяться в микробиологических исследованиях, будучи адаптированы в том числе и к изучению полногеномной экспрессии микобактерий, большинство подобных работ выполнены *in vitro*. В то же время анализ экспрессии генов микобактерий в ходе развития инфекционного процесса *in vivo*, представляющий собой наибольший научный интерес, достаточно сложен [5, 6]. В данной работе мы описываем результаты изучения экспрессии генов *M. tuberculosis* в ходе модельного инфекционного процесса в мы-

шах двух линий, различающихся по чувствительности к туберкулезу. Так как различия в генотипе животных непосредственно влияют на экспрессию генов патогена, увеличение экспрессии ряда бактериальных генов в более неблагоприятных условиях (в организме хозяина, устойчивого к инфекции) указывает на важность этих генов для адаптации *M. tuberculosis* к защитным механизмам хозяина.

## ЭКСПЕРИМЕНТАЛЬНАЯ ЧАСТЬ

### Модельная инфекция и выделение РНК

В качестве модельных организмов использовали мышей инбредных линий I/StSnEgYCit (I/St) и C57BL/6YCit (B6), содержащихся в стандартных условиях в виварии ЦНИИТ РАМН в соответствии с положениями приказа Минздрава СССР № 755 от 12 августа 1977 и NIH Office of Laboratory Animal Welfare Assurance № A5502-11. Корм и вода животным были доступны *ad libitum*. Все экспериментальные процедуры одобрены комитетом по биоэтике ЦНИИТ.

Самки мышей обеих линий в возрасте 2,5–3,0 мес. инфицировали вирулентным штаммом *M. tuberculosis* H37Rv с помощью системы аэрозольного заражения (Glas-Col, Terre Haute, США) из расчета 100–200 КОЕ на мыш. Инфицированные животные были умерщвлены через 4 и 6 нед. с момента заражения, после чего ткани легких немедленно использовали для выделения РНК. Суммарную РНК выделяли при помощи набора реактивов SV Total RNA Isolation System (Promega, США), после чего образцы РНК обрабатывали ДНКазой I (MBI Fermentas, Литва) для удаления следовых количеств ДНК.

### Синтез кДНК

кДНК синтезировали с использованием эффекта template-switch (Clontech, США) по методике [7]. кДНК синтезировали с помощью обратной транскриптазы PowerScript II (Clontech, США) в условиях, рекомендованных фирмой-производителем. Олигонуклеотидные праймеры BR (5'-AAGCAGTGGTATC AACGCAGAGTAC(N)9) и SMART (5'-AAGCAGTGGTATCAACGCAGAGTACGCrGrGrG) добавляли к 2 мкг суммарной РНК каждого из образцов в буферном растворе объемом 11 мкл. Полученную смесь инкубировали в течение 2 мин при 70°C, а затем помещали в лед на 10 мин. Во время нахождения на ледяной бане в смесь вносили 11 мкл раствора, содержащего 4 мкл 5×буфера для обратной транскриптазы, 2 мкл раствора 100 мМ DTT, 2 мкл раствора 10 мМ каждого dNTP, 1 мкл (200 ед. акт.) обратной транскриптазы PowerScript (Clontech, США). Параллельно с реакцией обрат-

ной транскрипции (RT+) проводили контрольную реакцию (RT-) без добавления обратной транскриптазы. Реакционные смеси RT+ и RT- инкубировали при 37°C в течение 10 мин, затем 40 мин при 42°C. Реакционные смеси инкубировали при 37°C (10 мин), затем при 42°C (120 мин). Параллельно с реакцией обратной транскрипции (RT+) проводили контрольную реакцию (RT-) без добавления обратной транскриптазы. Для препаративного синтеза кДНК проводили 30 циклов ПЦР (95°C, 20 с; 64°C, 20 с; 72°C, 2 мин) с использованием праймера 5S (5'-GTGGTATCAACGCAGAGT). Амплифицированную кДНК очищали набором реактивов QIAquick PCR Purification kit (Qiagen, США).

### Клонирование идентичных последовательностей

Идентичные последовательности (Coincidence Cloning) клонировали согласно [8]. Геномную ДНК *M. tuberculosis* H37Rv и образцы суммарных (т.е. синтезированных на матрицах суммарной РНК) кДНК фрагментировали эндонуклеазами рестрикции RsaI и AluI (MBI Fermentas, Литва). Полученные фрагменты геномной ДНК и кДНК лигировали с соответствующими супрессионными адаптерами (структуры адаптеров приведены в [7]). Смесь, содержащую 100 нг образца геномной ДНК с адаптерами и 100 нг соответствующего образца кДНК с адаптерами в 2 мкл гибридационного буфера НВ (50 мМ HEPES, pH 8.3; 0.5 М NaCl; 0.02 мМ EDTA, pH 8.0), инкубировали при 99°C в течение 5 мин (денатурация), а затем при 68°C в течение 18 ч (ренатурация). Затем в реакционную смесь добавляли 100 мкл предварительно нагретого до 68°C гибридационного буфера НВ. Полученный раствор (1 мкл) использовали для проведения двустадийной ПЦР-амплификации.

Первую стадию ПЦР осуществляли в реакционном объеме 25 мкл, содержащем 10 пмоль внешнего праймера Т7. После преинкубации в течение 5 мин при 72°C проводили 20 циклов амплификации (94°C, 30 с; 66°C, 30 с; 72°C, 90 с). Вторую стадию амплификации проводили с использованием внутренних праймеров (94°C, 30 с; 68°C, 30 с; 72°C, 90 с; 25 циклов) [7]. В качестве матрицы для второй стадии амплификации использовали амплификат первой стадии, разведенный в 10 раз. Амплификат второй стадии очищали с помощью набора реактивов QIAquick PCR Purification kit (Qiagen, США) и использовали для массированного секвенирования.

### Массированное секвенирование

Перед секвенированием одинаковые количества амплификатов различных продуктов Coincidence Cloning (по 500 нг каждого амплификата) объединяли

в один образец. Нуклеотидные последовательности определяли параллельным пиросеквенированием с помощью генетического анализатора GS FLX (454 Roche, Германия). Полученные последовательности (прочтения) проверяли на наличие в них последовательностей адаптеров. Прочтения с укороченными, неправильными или отсутствующими адаптерными последовательностями исключали из дальнейшего анализа. Оставшиеся последовательности (190031 прочтение) были разделены на три группы (библиотеки) в зависимости от нуклеотидной последовательности их адаптеров: CC4(RES) и CC6(RES), представляющие собой библиотеки кДНК из легких инфицированных мышей линии В6 (устойчивой к туберкулезу) на 4-й и 6-й нед. с момента заражения соответственно; CC6(SUS) – библиотеки кДНК из легких инфицированных мышей линии I/St (чувствительной к туберкулезу) на 6-й нед. с момента заражения. Файл в формате FASTA, содержащий нуклеотидные последовательности из всех трех групп, использовали для дальнейшей работы.

Файлы в формате FASTA могут быть предоставлены авторами по запросу.

### Картирование последовательностей и статистический анализ

Полученные последовательности картировали на последовательность генома *M. tuberculosis* H37Rv (сборка GenBank AL123456.2) с помощью программы blastn, входящей в пакет программ NCBI BLAST+, со следующими параметрами: -perc identity 95 и -evalue 0.01. В результате были отобраны последовательности, имеющие не менее 95% идентичности всей длины с участками генома *M. tuberculosis* H37Rv. Все последовательности длиной менее 40 нуклеотидов были исключены из обработки. Из дальнейшего анализа исключили также последовательности, картируемые в геноме *M. tuberculosis* не уникально (в два места и более). Прочтения, картируемые на межгенные последовательности, из дальнейшего анализа не исключали, так как это изменило бы размеры библиотек и привело к ошибкам при последующем статистическом анализе. Затем для каждого гена и межгенного участка определяли количество прочтений в каждой библиотеке и проводили сравнительный анализ представленности фрагментов кДНК, соответствующих бактериальным генам и межгенным участкам при помощи алгоритма Audic-Claverie [9]. Различия в экспрессии генов (межгенных участков) признавали значительными, если количество прочтений, картируемых на ген (межгенный участок), по крайней мере в одной из двух сравниваемых библиотек, было не менее 20 и значение  $p < 0.05$ .

### РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Для выявления особенностей профиля экспрессии *M. tuberculosis*, коррелирующих с прогрессией инфекции, на разных этапах инфекционного процесса мы провели сравнительный количественный и качественный анализ последовательностей, транскрибирующихся при заражении мышей, генетически чувствительных (неэффективный иммунный ответ) и резистентных (эффективный ответ) к этим бактериям.

Мы сравнили транскриптомы *M. tuberculosis* H37Rv при инфекции мышей двух линий, I/StSnEgYCit (I/St) и C57BL/6YCit (B6). Эти линии мышей были подробно описаны ранее [10]. У мышей линии B6 устойчивость к *M. tuberculosis* выше, чем у мышей линии I/St, что выражается в менее агрессивном течении инфекционного процесса у мышей линии B6 и более длительной выживаемости инфицированных животных.

Самок мышей обеих линий спустя 4 и 6 нед с момента аэрогенного заражения бактериями *M. tuberculosis* умерщвляли и выделяли из их легких суммарную РНК. Образцы суммарной РНК из тканей легкого мышей I/St и B6 использовали для синтеза кДНК, обогащенной впоследствии фрагментами бактериальной кДНК методом клонирования идентичных последовательностей [8]. Всего получили три библиотеки последовательностей, характеризующие транскриптомы *M. tuberculosis* из тканей мышей линии I/St на 6-й нед. с момента инфекции (CC6(SUS)) и из тканей мышей линии B6 на 4-й и 6-й нед. с момента инфекции (CC4(RES) и CC6(RES) соответственно). Нуклеотидные последовательности фрагментов кДНК из данных библиотек определяли с помощью массированного 454 пиросеквенирования. Общая схема эксперимента приведена на *рисунке*, общая характеристика проанализированных библиотек приведена в *табл. 1*. Всего определили нуклеотидные последовательности 190031 фрагмента кДНК. Из них в CC4(RES) определили 73410 последовательностей, 75655 – в образце CC6(SUS), 40966 – в CC6(RES). Полученные последовательности картировали на последовательность генома *M. tuberculosis* H37Rv (сборка GenBank AL123456.2) с помощью программы blastn из программного пакета NCBI BLAST+.

По результатам картирования было установлено, что в образце CC4(RES) 14990 (20.42%) последовательностей идентичны последовательностями *M. tuberculosis*, в образце CC6(SUS) – 43618 (57.65%) и в образце CC6(RES) – 34234 (83.57%). Приведенные результаты свидетельствуют о том, что удалось достигнуть значительного обогащения образцов кДНК бактериальными последовательностями.

Из 4012 генов и семи псевдогенов *M. tuberculosis* в образце CC4(RES) экспрессировались 1012 (25.2%)

Таблица 1. Результаты секвенирования и картирования библиотек CC4(RES), CC6(RES) и CC6(SUS)

Библиотека	CC4(RES)	CC6(SUS)	CC6(RES)
Прочтений, всего	73410	75655	40966
Mtb-специфических прочтений, уникальных	14990	43618	34234
Mtb-специфических прочтений (уникальных), % от всех	20.4	57.7	83.6
Генов экспрессируется (прочтений > 0)	1012	1353	1940
Генов экспрессируется, % от общего количества генов	25.2	33.7	48.3
IGR экспрессируется (прочтений > 0)	164	221	376
IGR экспрессируется, от общего количества IGR, %	5.3	7.2	12.3

от общего количества генов), в образце CC6(SUS) – 1353 (33.7%), в образце CC6(RES) – 1940 генов (48.3%). Ни в одном из образцов не экспрессировались 1428 генов (35.5%), в то время как 469 (11.7%) генов экспрессировались в каждом из образцов.

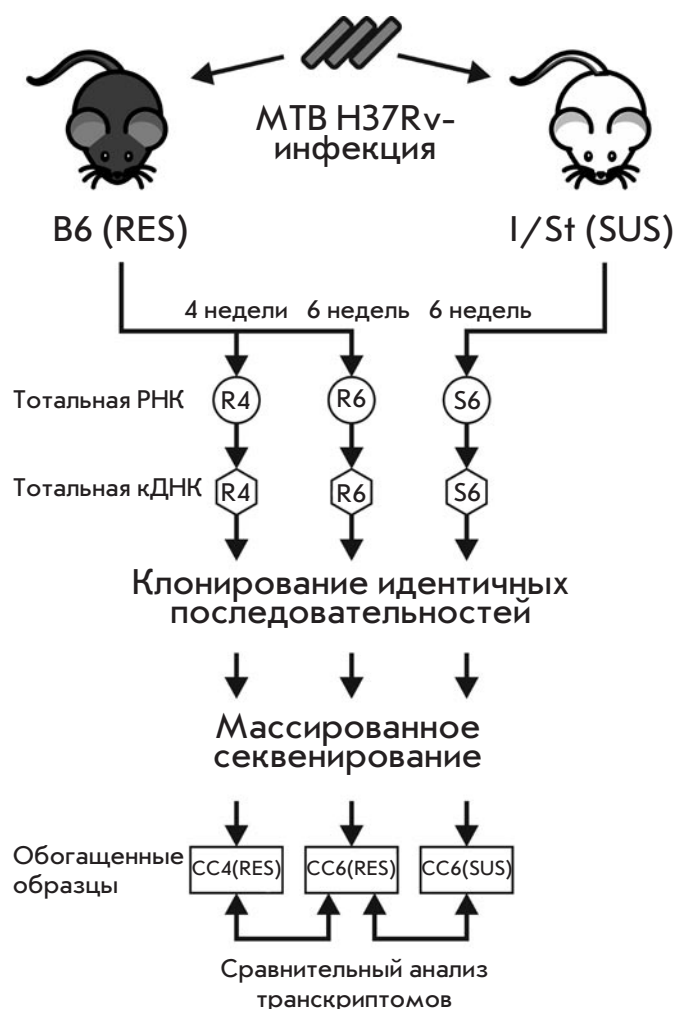
### Некодирующие РНК

Новейшие исследования с применением технологий массивного секвенирования показали, что бактериальный транскриптом значительно более сложен, чем считалось ранее. Показано, что некодирующие РНК представляют значительную и важную часть транскриптома и регулируют различные клеточные функции, среди которых можно выделить репликацию, энергетический метаболизм и регуляцию экспрессии факторов вирулентности у ряда бактериальных патогенов [11].

Мы провели поиск транскриптов из локусов, расположенных в межгенных областях, так как подобная локализация свидетельствует о потенциальной принадлежности этих транскриптов к группе некодирующих РНК. На основании данных о структурной организации генома мы выделили 3069 межгенных областей. После анализа данных пиросеквенирования установили, что в образцах CC4(RES), CC6(SUS) и CC6(RES) присутствовали 164 (5.3%), 221 (7.2%) и 376 (12.3%) транскриптов межгенных областей соответственно. При этом транскрипты из 27 (0.9%) локусов были представлены в каждом из трех образцов, и ни разу не встретились транскрипты 2490 (81.1%) локусов.

Данные об экспрессии межгенных областей, полученные в ходе эксперимента, сопоставили с опубликованной и представленной в базах данных информацией о локализации генов малых РНК.

Мы выявили значительный уровень экспрессии последовательностей из ряда межгенных участ-



Общая схема эксперимента по сравнению транскриптомов *M. tuberculosis*. RES – генетически устойчивая линия мышей; SUS – генетически чувствительная линия мышей; CC – библиотека, обогащенная бактериальной кДНК

ков, в частности из IGR3987 и IGR0629, в образце CC6(SUS) и из IGR1186 в образце CC6(RES). Результаты нашего эксперимента показали также экспрессию из межгенных участков IGR3987, IGR0629 и IG1136, что коррелирует с данными, полученными Arnvig *et al.* [12] (согласно обозначениям, используемым в [12], – межгенные участки IGR2975, IGR0479 и IGR0858 соответственно). Мы полагаем, что обнаруженная нами экспрессия межгенных последовательностей может свидетельствовать о потенциальной локализации в данных локусах генов малых РНК и, принимая во внимание их дифференциальную экспрессию, может быть отражением компенсаторной реакции патогена на воздействия внешней среды.

### Гены, экспрессия которых повышается при развитии инфекции

Мы сравнили транскриптом при развитии инфекции в генетически устойчивой линии мышей (CC6(RES) и CC4(RES)) и в одной временной точке в генетически различных линиях мышей (CC6(RES) и CC6(SUS)). Целью этого сравнения был поиск генов, экспрессия которых повышается при развитии инфекции, т.е. в мышцах линии В6 на 6-й нед. после заражения по сравнению с другими точками. Сравнение CC6(RES) и CC4(RES) позволило нам выявить 226 генов, экспрессия которых повышена в ходе развития инфекции в тканях мышей В6. В результате сравнения CC6(RES) и CC6(SUS) обнаружили 253 гена, экспрессия которых повышена в образце CC6(RES).

Сравнение CC6(RES) и CC4(RES) позволило выявить только 17 генов, экспрессия которых в CC6(RES) выше, чем в CC4(RES), тогда как экспрессия 44 генов в CC6(RES) была выше, чем в CC6(SUS). Вероятно, эти результаты отражают тот факт, что первое сравнение характеризует динамику изменения экспрессии генов патогена во времени в пределах одного микроокружения. Во втором случае выявляются различия между двумя разными микроокружениями, что отражается в большем количестве генов, экспрессия которых повышена в CC6(RES).

Гены, экспрессия которых в образце CC6(RES) повышена только по сравнению с CC4(RES), относятся преимущественно к категориям cell wall and cell processes, intermediary metabolism and respiration и lipid metabolism. Белковые продукты 12 из 17 генов обнаружены во фракции клеточной мембраны и/или клеточной стенки, где они выполняют преимущественно транспортные и защитные функции. Так, ген *embA* кодирует индолилацетилинотол-арабинозилтрансферазу EmbA, участвующую в синтезе арабиана. Мутации в этом гене приводят к устойчивости к этамбутолу. Ген *Rv3273* кодирует карбонатдегидратазу, функции которой связа-

ны с транспортом сульфатов (TubercuList). Анализ с использованием баз данных KEGG Pathways (<http://www.genome.jp/kegg/pathway.html>) и TBCYC (<http://tbcyc.tdb.org/>) не выявил метаболических путей, которые активируются на более поздних стадиях развития инфекции. Этот факт может быть либо следствием случайных флуктуаций в экспрессии генов патогена, его реакцией на случайные изменения свойств микроокружения, либо отражать небольшие, но существенные различия функциональной активности *M. tuberculosis* в различных временных точках.

Сравнение образцов CC6(RES) и CC6(SUS) обнаружило в CC6(RES) больше генов, экспрессия которых в этом образце была выше, чем в CC6(SUS). Усиление энергетического обмена выражалось в повышенной экспрессии генов трех субъединиц NADH-дегидрогеназы (*nuoH*, *nuoI*, *nuoL*); большей активности цикла трикарбоновых кислот (*acn*), а также в повышенной экспрессии гена *Rv1916*. Ген *Rv1916* является второй частью гена *aceA* (*icl2*), разделенного у *M. tuberculosis* H37Rv на два модуля, *Rv1915* и *Rv1916* (*aceAa* и *aceAb*), которые экспрессируются по отдельности. Среди других важных отличий можно отметить также повышенную экспрессию генов, продукты которых отвечают за метаболизм и катаболизм липидов и аминокислот (*lipV*, *lipF*, *Rv2531c*), а также ферментов, участвующих в репарации ДНК (*recO*, *recB*). Подобная картина достаточно предсказуема, так как микроокружение резистентного хозяина представляет собой враждебную среду обитания, что объясняет необходимость большей активности систем репарации. Повышенная экспрессия липолитических ферментов (*lipF*, *lipV*, *plcA*), ферментов цикла трикарбоновых кислот и *aceAb* могут указывать на большую степень использования липидов в качестве источника энергии и углерода.

Мы сосредоточились на поиске генов *M. tuberculosis*, повышенная экспрессия которых не зависит от генетических особенностей организма хозяина. Эти гены образуют некий базовый набор, отражающий универсальную компенсаторную реакцию патогена на неблагоприятные условия окружающей среды. Далее эти гены обозначены как CUG – Commonly Upregulated Genes. Определено 209 генов, экспрессия которых увеличена в обоих сравнениях (табл. 2). Согласно результатам транспозонного мутагенеза 44 гена *M. tuberculosis* H37Rv относятся к незаменимым (essential, [13]); необходимость еще трех генов (*Rv3569c*, *Rv3537*, *Rv3563*) для выживания в макрофагах мыши показана ранее (TubercuList, <http://tuberculist.epfl.ch>).

Мы сформировали группы генов CUG по их принадлежности к функциональным категориям (Tu-

Таблица 2. Гены группы CUG

Ген	Функциональная категория (согласно TubercuList)
<i>Rv0028, Rv0074, Rv0269c, Rv0274, Rv0281, Rv0421c, Rv0428c, Rv0433, Rv0448c, Rv0455c, Rv0492A, Rv0525, Rv0597c, Rv0695, Rv1179c, Rv1186c, Rv1203c, Rv1232c, Rv1419, Rv1428c, Rv1828, Rv1835c, Rv1868, Rv1998c, yfiH, Rv2974c, Rv3030, Rv3031, Rv3205c, Rv3272, Rv3519, Rv3627c, Rv3651, Rv3662c, Rv3703c, Rv3753c, Rv0026, Rv0061, Rv0140, Rv0141c, Rv0145, Rv0332, Rv0712, Rv0785, Rv0998, Rv1514c, Rv1515c, wbbL2, Rv1760, Rv2077A, Rv2135c, Rv2466c, Rv2699c, Rv2751, Rv2823c, Rv3067, Rv3090, Rv3094c, Rv3510c</i>	CH – conserved hypotheticals
<i>Rv0051, Rv0309, lprL, Rv0621, Rv0876c, lytB2, irtA, Rv1687c, secA2, Rv2209, Rv2265, mmpL7, Rv3194c, Rv3658c, embC, espE, ponA1, Rv0072, narK3, iniA, cpsY, lpqR, pstS1, Rv0996, kdpC, Rv1097c, sugB, Rv1431, Rv1667c, Rv2136c, Rv2203, efpA, rip, Rv2963, lpqF</i>	CWaCP – cell wall and cell processes
<i>Rv0161, ndhA, Rv0526, menH, Rv0805, lipU, glyA1, dapE, atpF, atpH, Rv1432, frdB, cmk, plcD, lipJ, cobK, cobS, cysK1, cysE, gdh, gabT, miaA, ilvC, guaB2, cyp142, hsaD, Rv0089, Rv0331, aspC, hemA, Rv0567, atsA, gltA2, Rv0943c, Rv1096, Rv1106c, narH, thrB, hisB, ilvG, rocD1, plcB, phoH1, ggtB, lepA, Rv2499c, dapF, purU, kstD, folP1</i>	IMaR – intermediary metabolism and respiration
<i>end, fusA1, polA, lysX, helZ, spoU, ppiB, thrS, Rv3201c</i>	IP – information pathway
<i>Rv0095c, Rv0920c, Rv2791c</i>	ISaP – insertion sequences and phages
<i>fadD10, nrp, fadD7, fadE4, fadD2, fadD12, pks17, pks12, mbtF, mbtE, mbtC, TB7.3, accA3, fadE27, fadD17, accD4, mmaA3, mmaA1, fadE19, Rv2613c, fadD26, ppsC, ppsD, fadD19, fadE31, fadE32, pks13</i>	LM – lipid metabolism
<i>PPE8, PE_PGRS19, PE16, PPE34, PPE50, PE2, PPE64</i>	PE/PPE – PE/PPE protein families
<i>pknA, senX3, trcR, Rv1359, fhaA, Rv0465c, Rv3066, Rv3736</i>	RP – regulatory proteins
<i>Rv2645, Rv2818c</i>	U – unknown
<i>treS, mce2C, Rv1026, ephB, vapB16, Rv2581c, vapC3, cinA, virS</i>	VDA – virulence, detoxification, adaptation

bercuList) и сравнили их распределение с распределением всех генов *M. tuberculosis*. В целом эти распределения оказались сходными, за исключением генов категории липидного метаболизма, что может свидетельствовать в пользу их важности в процессах бактериальной адаптации.

Чуть меньше трети генов приходилось на две категории – conserved hypotheticals (59 генов) и unknown (два гена). Несмотря на отсутствие известных функций, гены данной категории представляют потенциальные терапевтические мишени, так как невысокая степень их гомологии с генами других микроорганизмов означает, что они характерны именно для микобактерий или конкретно для *M. tuberculosis*, определяя, возможно, вирулентные свойства.

В пользу того, что микобактерии оказываются в бедных питательными веществами условиях, указывает высокий уровень экспрессии генов различных систем захвата и накопления питательных веществ, например, фосфата (*pstS1*), железа (*irtA, mbtC, mbtE, mbtF*). На нехватку фосфата указывает и повышенная экспрессия гена *senX3*, сенсорного компонента двухкомпонентной регуляторной системы *senX3/regX3*, активирующей так называемый «строгий ответ» (stringent response) в условиях дефици-

та фосфата. На переход к использованию в качестве основного источника энергии и углерода указывает экспрессия генов метаболизма липидов (*fadD, fadE, lipU, lipJ*). Еще одна особенность генов группы CUG – повышенная экспрессия генов, продукты которых тем или иным образом связаны с метаболизмом аминокислот (*aspC, hisB, thrB, thrS*). Причина подобного явления не вполне ясна, так как стимуляция экспрессии ферментов может быть вызвана как отсутствием необходимых аминокислот (и соответственно необходимостью их синтеза), так и их наличием (и возможностью для бактерии использовать их).

На переход к характерному для латентной инфекции анаэробному нитратному дыханию указывает повышенная экспрессия генов *narH* и *narK3* [14]. К генам группы CUG мы отнесли также гены *atpF* и *atpH*, хотя обычно сообщают о снижении их экспрессии при развитии инфекции, так как энергетические потребности патогена снижаются по мере того, как он входит в состояние латентной инфекции [15, 16].

Функция белков семейства PE/PPE не вполне понятна. Считается, что они необходимы для создания антигенной вариабельности у микобактерий [17]. Тем не менее гены *Rv0152c* и *Rv0355c* экспрессируются на высоком уровне в образце CC6(RES), и их экспрес-

сия обнаружена в образцах CC4(RES) и CC6(SUS), а ген *Rv3135* относится к числу незаменимых у *M. tuberculosis* H37Rv, что может указывать на какие-либо дополнительные функции, помимо обеспечения антигенной вариабельности.

Наконец, стоит отметить ген *secA2*. Этот ген кодирует транслоказу SecA2 – компонент вспомогательной транспортной системы Sec *M. tuberculosis*, обеспечивающей, помимо прочего, секрецию супероксиддисмутазы SodA и каталазы KatG. Живая вакцина, созданная на основе мутанта *M. tuberculosis* по гену *secA2*, показала высокую эффективность и безопасность при испытаниях на животных [18].

Недавно методом микроэрреев провели сравнительный анализ профилей экспрессии 17 представителей комплекса *M. tuberculosis* в активированных и неактивированных макрофагах мыши [19]. В результате удалось выделить 280 генов (168 с универсально повышенной и 112 с универсально пониженной экспрессией), изменение экспрессии которых не зависело от штамма и статуса активации макрофага. Мы сравнили гены группы CUG со 168 генами с универсально повышенной экспрессией, взятыми из работы [19], и выделили восемь (*Rv0140*, *Rv0145*, *atsA*, *Rv2466c*, *fadD26*, *ilvC*, *Rv3067* и *kstD*), представленных в обоих списках. Столь небольшое совпадение можно объяснить тем, что а) инфекция макрофагов в культуре представляет собой достаточно упрощенную модель по сравнению со сложными взаимоотношениями между клетками-хозяина и патогена в ходе инфекционного процесса в организме-хозяина и б) результаты анализа экспрессии с помощью микроэрреев могут достаточно сильно отличаться от результатов, полученных с помощью массивированного секвенирования. К примеру, Ward *et al.* [20] пишут о расхождении результатов, полученных этими двумя методами.

Тем не менее расхождения в результатах, полученных нами и Nomolka *et al.* [19], не являются кардинальными. Функциональный анализ генов с повышенной экспрессией показал, что они ассоциированы с такими внутриклеточными стрессовыми воздействиями, как гипоксия, действие активных форм кислорода и азота, ремоделирование клеточной стенки и метаболизм жирных кислот. В качестве примера можно указать гены, ассоциированные с нехваткой железа (гены кластера *mtbA-F*), а также гены, вовлеченные в биосинтез валина и изолейцина (*ilvB-ilvN-ilvC*) и димикоцерозатов фтиоцерола (PDIM) клеточной стенки (*ppsA-D*).

### Продукты генов группы CUG как потенциальные терапевтические мишени

После анализа доступной научной литературы и баз данных было выделено шесть генов, представляющих

собой потенциальные или уже предложенные лекарственные мишени. Продукты этих генов (*hisB*, *aspC*, *PPE50*, *Rv1026*, *ilvC* и *Rv1186c*) упомянуты как точки, привлекательные для терапевтического воздействия, так как нарушения их функциональной активности оказывали максимальный дестабилизирующий эффект на метаболизм *M. tuberculosis*. К примеру, аспартат-аминотрансфераза AspC, кодируемая геном *aspC* (*Rv0337c*), в метаболической сети микобактериальной клетки идентифицирована как фермент, инактивация которого влияла на большое число других белков *M. tuberculosis* и тем самым эффективно разрушала большое число биохимических циклов [21]. Белковые продукты двух других генов, *Rv1186c* и *PPE50* (*Rv3135*), вошли в список наиболее привлекательных потенциальных терапевтических мишеней, составленный на основе многочисленных данных об экспрессии, вовлеченности в различные метаболические пути и структурной гомологии с белками других бактерий и человека [22]. Гены *Rv1026*, *hisB* (*Rv1601*) и *Rv3001c*, как полагают [23], кодируют продукты, пригодные для создания специфических ингибиторов. Белок, кодируемый геном *Rv1601* (*hisB*), независимо рассматривается в качестве потенциальной терапевтической мишени [24]. Стоит отметить, что экспрессия гена *Rv1026*, кодирующего пирофосфатазу, повышена в макрофагах и легких инфицированных мышей [25, 26], а также в условиях ингибирования трансляции у микобактерий [27]. Недавно показали, что недостаток полифосфатов, обусловленный гидролитической активностью *Rv1026*, может приводить к изменению содержания жирных кислот в клеточной стенке *M. smegmatis*, влияя тем самым на способность к перемещению (sliding motility) и образованию биопленок [28]. Таким образом, можно ожидать, что и продукты остальных генов группы CUG могут быть использованы для создания лекарственных средств или диагностики туберкулеза.

### ЗАКЛЮЧЕНИЕ

Инфекционные заболевания, вызываемые внутриклеточными патогенными бактериями, представляют собой серьезную медицинскую проблему. Ход инфекционного процесса зависит не только от защитных механизмов хозяина, но и от специфической экспрессии бактериальных генов. Изменения экспрессии генов в ответ на различные реакции иммунной системы хозяина необходимы для выживания и размножения патогенных бактерий. Изучение изменений профиля транскрипции *M. tuberculosis* под воздействием различных стимулов и условий внешней среды позволяет описать адаптационные механизмы, необходимые бактерии для эффективного заражения организма-хозяина.

Изучение профилей транскрипции *M. tuberculosis* в различных условиях позволило выявить базовый набор генов (CUG), экспрессия которых повышается в ходе развития инфекции и не зависит от генетических особенностей организма-хозяина. Экспрессия генов из этого ряда может рассматриваться как универсальная реакция микобактерий на различные стрессовые факторы внешней среды. Накопление и анализ данных об экспрессии генов *M. tuberculosis* позволит значи-

тельно упростить разработку эффективных подходов к диагностике и лечению туберкулеза. ●

*Работа поддержана Министерством образования и науки РФ (соглашение № 8308); РФФИ (грант № 11-04-01325); Программой поддержки ведущих научных школ России (проект НШ-1674.2012.4); Программой Президиума РАН «Молекулярная и клеточная биология».*

#### СПИСОК ЛИТЕРАТУРЫ

- Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D., Gordon S.V., Eiglmeier K., Gas S., Barry C.E., 3<sup>rd</sup>, et al. // *Nature*. 1998. V. 393. № 6685. P. 537–544.
- Wilson M., DeRisi J., Kristensen H.H., Imboden P., Rane S., Brown P.O., Schoolnik G.K. // *Proc. Natl. Acad. Sci. USA*. 1999. V. 96. № 22. P. 12833–12838.
- Butcher P.D. // *Tuberculosis (Edinb)*. 2004. V. 84. № 3–4. P. 131–137.
- Kendall S.L., Rison S.C., Movahedzadeh F., Frita R., Stoker N.G. // *Trends Microbiol*. 2004. V. 12. № 12. P. 537–544.
- Скворцов Т.А., Ажикина Т.Л. // *Биоорган. химия*. 2010. Т. 36. № 5. С. 596–606.
- Скворцов Т.А., Ажикина Т.Л. // *Биоорган. химия*. 2012. Т. 38. № 4. С. 391–405.
- Игнатов Д.В., Скворцов Т.А., Майоров К.Б., Апт А.С., Ажикина Т.Л. // *Acta Naturae*. 2010. Т. 2. № 3 (6). С. 57–62.
- Azhikina T.L., Skvortsov T.A., Radaeva T.V., Mardanov A.V., Ravin N.V., Apt A.S., Sverdlov E.D. // *Biotechniques*. 2010. V. 48. № 2. P. 139–144.
- Audic S., Claverie J.M. // *Genome Res*. 1997. V. 7. № 10. P. 986–995.
- Kondratieva E., Logunova N., Majorov K., Averbakh M., Apt A. // *PLoS One*. 2010. V. 5. № 5. P.e10515.
- Arnvig K., Young D. // *RNA Biology*. 2012. V. 9. № 4. P. 427–436.
- Arnvig K.B., Comas I., Thomson N.R., Houghton J., Boshoff H.I., Croucher N.J., Rose G., Perkins T.T., Parkhill J., Dougan G., et al. // *PLoS Pathog*. 2011. V. 7. № 11. P.e1002342.
- Sassetti C.M., Boyd D.H., Rubin E.J. // *Mol. Microbiol*. 2003. V. 48. № 1. P. 77–84.
- Shi L., Sohaskey C.D., Kana B.D., Dawes S., North R.J., Mizrahi V., Gennaro M.L. // *Proc. Natl. Acad. Sci. USA*. 2005. V. 102. № 43. P. 15629–15634.
- Stokes R.W., Waddell S.J. // *Future Microbiol*. 2009. V. 4. № 10. P. 1317–1335.
- Waddell S.J. // *Drug Discov. Today: Dis. Mech*. 2010. V. 7. № 1. P. e67–e73.
- Karboul A., Mazza A., Gey van Pittius N.C., Ho J.L., Brousseau R., Mardassi H. // *J. Bacteriol*. 2008. V. 190. № 23. P. 7838–7846.
- Hinchey J., Jeon B.Y., Alley H., Chen B., Goldberg M., Derrick S., Morris S., Jacobs W.R., Jr., Porcelli S.A., Lee S. // *PLoS One*. 2011. V. 6. № 1. P. e15857.
- Homolka S., Niemann S., Russell D.G., Rohde K.H. // *PLoS Pathog*. 2010. V. 6. № 7. P.e1000988.
- Ward S.K., Abomoelak B., Marcus S., Talaat A.M. // *Front. Microbiol*. 2010. V. 2. P.
- Raman K., Vashisht R., Chandra N. // *Mol. BioSystems*. 2009. V. 5. № 12. P. 1740–1751.
- Raman K., Yeturu K., Chandra N. // *BMC Systems Biol*. 2008. V. 2. № 1. P. 109.
- Kalapanulak S. High quality genome-scale metabolic network reconstruction of mycobacterium tuberculosis and comparison with human metabolic network: application for drug targets identification. Edinburgh: Univ. of Edinburgh, 2009.
- Nisa S.Y. ParA: a novel target for anti-tubercular drug discovery. Wellington: Victoria Univ. of Wellington, 2010.
- Srivastava V., Jain A., Srivastava B.S., Srivastava R. // *Tuberculosis*. 2008. V. 88. № 3. P. 171–177.
- Srivastava V., Rouanet C., Srivastava R., Ramalingam B., Loch C., Srivastava B.S. // *Microbiology*. 2007. V. 153. № 3. P. 659–666.
- Boshoff H.I.M., Myers T.G., Copp B.R., McNeil M.R., Wilson M.A., Barry C.E. // *J. Biol. Chem*. 2004. V. 279. № 38. P. 40174–40184.
- Shi T., Fu T., Xie J. // *Curr. Microbiology*. 2011. V. 63. № 5. P. 470–476.