

Мини-обзор генома бактерии *Arthrobacter sp. A3*

Шпудейко П. С.¹

¹МГУ им. М. В. Ломоносова, факультет биоинженерии и биоинформатики, первый курс

РЕЗЮМЕ

В этой работе анализируется геном бактерии *Arthrobacter sp. A3*. Были определены размер генома, число генов, кодирующих белки и различные типы РНК, и их расположение на комплементарных цепях молекулы ДНК. Результаты анализа были представлены в виде таблиц и гистограммы, выполненных с помощью программы Excel из пакета MsOffice. Что помогло сделать выводы о некоторых закономерностях в строении генома данной бактерии.

1 ВВЕДЕНИЕ

Arthrobacter sp. A3 относится к роду *Arthrobacter*, семейству *Micrococcales*, порядку *Actinomycetales*, классу *Actinobacteria*, типу *Actinobacteria*, царству *Bacteria*.

Штаммы рода *Arthrobacter* широко распространены в окружающей среде (в частности в почве) и выделяются из клинического материала людей при бактериемии, различных инфекционных процессах.^[1]

Название рода *Arthrobacter* с греческого переводится, как "соединенная маленькая палочка", оно обусловлено характерными особенностями способа деления бактерий, т. н. "щелкающее" деление (snapping division). Это такой тип деления, когда вновь образуемые дочерние клетки как бы отскакивают друг от друга, приводя тем самым к угловому или V-образному расположению клеток.^[2]

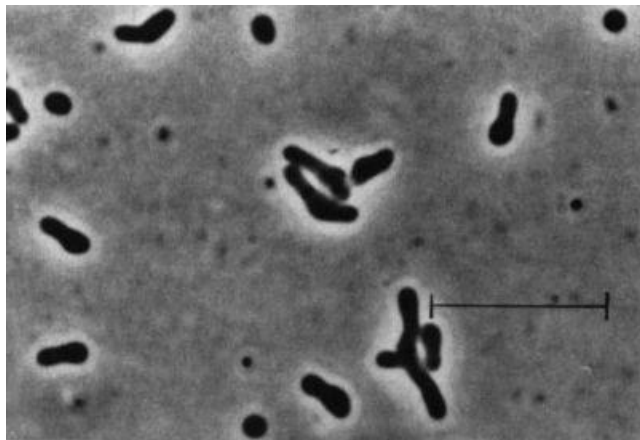


Рис. 1. Различные жизненные формы рода *Arthrobacter* (взято из microbewiki.kenyon.edu). В частности, представлено характерное V-образное расположение клеток.

Представители класса *Actinobacteria* являются дифтериеподобными бактериями. Они характеризуются плеоморфизмом (переменная форма) и изменчивостью при окраске по Граму. Генетически эти бактерии относятся к грамположительным облигат-

ным аэробам, обладающим сложным жизненным циклом, характеризующимся сменой формы клеток организмов. Так, при рассмотрении молодых культур под микроскопом, можно наблюдать клетки палочковидной формы, которые окрашиваются по Граму как грамотрицательные. А примерно через 30 часов клетки становятся очень короткими грамположительными кокками.^[2]

Представителей рода *Arthrobacter* используют для производства L-глутаминовой кислоты. В промышленности их выращивают на недорогих свекольных или тростниковых субстратах.^[3]

2 МАТЕРИАЛЫ И МЕТОДЫ

1. Описание генома бактерии было скачано с сайта NCBI: NCBI Genome => Browse by organism => Search by organism.

2. Для обработки и интерпретации данных использовалось программное обеспечение MsOffice Excel. При работе с данной программой в рамках данной работы применялись следующие встроенные функции: СЧЁТЕСЛИ, СУММ, ОКРУГЛ – для создания сводной таблицы генома; СЧЁТЕСЛИМН – для построения гистограммы; СЛУЧМЕЖДУ, СЧЁТЕСЛИ – для проверки гипотезы о случайном и равновероятном распределении генов.

3. Для получения данных, на основании которых были выполнены вышеуказанные действия, файл, скачанный с сайта NCBI, был преобразован в формат Excel, из него были удалены с помощью фильтра неинформативные строки "gene" и отсортированы строки по возрастанию порядкового номера стартового кодона.

3 РЕЗУЛЬТАТЫ

Геном бактерии *Arthrobacter sp. A3* представлен одной кольцевой хромосомой, состоящей из 4264061 пары нуклеотидов, и одной плазмидой из 194226 пар нуклеотидов. Он был впервые секвенирован 17 сентября 2014 года в Китайской Академии наук.^[4]

3.1 Число генов белков и РНК

В результате анализа было установлено, что кольцевая хромосома и плазида бактерии *Arthrobacter sp. A3* суммарно содержат 3983 гена и кодируют 3911 белков и 72 РНК (только на кольцевой хромосоме) различного типа: тРНК, рРНК, некодирующие РНК. Из 3911 белков кольцевая хромосома кодирует 3735 белков, а плазида – 176.

¹ [Описание генома](#)

Результат детального анализа соотношений между числом генов белков и РНК на кольцевой хромосоме сведён в Таблицу 1.

Таблица 1. Число генов бактерии *Arthrobacter sp.* АЗ по различным видам кодируемых ими продуктов.

	CDS	rRNA	tRNA	ncRNA
Amount	3911	19	52	1
%	98,19	0,48	1,31	0,03

Из таблицы видно, что большая часть генов бактерии *Arthrobacter sp.* АЗ кодирует синтез белков. Довольно большое (по отношению к другим типам РНК) количество генов приходится на транспортные РНК. Существует одна не кодирующая РНК, которая, как показывают исследования, предназначена для регуляции модификации хроматина, метилирования ДНК и транскрипционного сайленсинга. [5]

Также было посчитано, что на 1 млн пар нуклеотидов приходится примерно 934 гена.

3.2 Распределение генов по цепям ДНК

Так как молекула ДНК двухцепочечная, то выделяют прямую цепь ДНК (+) и обратную цепь ДНК (-). Распределение генов по цепям ДНК представлено в Таблице 2.

Таблица 2. Распределение генов по цепям ДНК.

Chain	CDS	RNA
+	1888	35
-	2023	37

3.3 Распределение длин белков

Данные по распределению длин белков представлены на гистограмме (Рис. 2).



Рис. 2. Гистограмма длин белков из протеома бактерии *Arthrobacter sp.* АЗ

Было выявлено, что довольно большой процент от общего числа белков обладает размерами от 60 до 510 аминокислотных

остатков (а именно 86,5%), 12,9% белков – от 510 до 3500 остатков, и размерами меньше 60 остатков – 0,6%.

По гистограмме можно определить, что наибольшее число белков (305 от общего числа) находится в диапазоне длин от 300 до 330 аминокислотных остатков. Заметно, что сначала происходит резко увеличивается число белков при увеличении числа аминокислотных остатков, а затем, начиная со значения 330, неравномерно убывает вплоть до 3500 (белка длиннее этого значения нет).

3.4 Гипотеза о распределении генов по цепям ДНК

Было предположено, что гены распределены по цепям ДНК случайно с вероятностью 0,5. Подсчитано, что реальное отклонение от ожидаемого распределения генов по цепям составляет 68 п. н. (общее число генов – 3983 поделено на 2). Сопоставив каждому гену либо 1, либо 0 100 раз, было установлено, что примерно в 0-5 случаях из 100 возможное случайное распределение генов превосходит либо является равным реальному отклонению.

Результат приведён в файле Excel на Листе «3» (см. «Сопроводительные материалы»).

4 ОБСУЖДЕНИЕ И ЗАКЛЮЧЕНИЕ

На основании приведённых результатов можно сделать следующие выводы:

- Гены по комплементарным цепям распределены почти равномерно, однако на обратной цепи число генов белков и РНК превосходит число генов на прямой цепи, в частности на 135 генов белков и на 2 гена РНК. Предположительно такое распределение имеет место, чтобы уменьшить размер генома микроорганизма.
- В результате анализа протеома микроорганизма было обнаружено, что действительное количество белковых продуктов отличается от числа генов в геноме. Это обусловлено тем, что гены РНК не кодируют белки и что существует 83 псевдогена (данные взяты из описания генома, скачанного с сайта NCBI), которые не дают конечного белкового продукта, так как они являются нефункциональными аналогами структурных генов и не кодируют белки. [6] Вполне возможно, что ранее эти гены «работали» и обеспечивали синтез белков, которые для этого вида в современном мире потеряли свою необходимость и были «выключены».
- На основании того, что лишь в 0-5% случаев возможно реальное распределение генов, можно сделать вывод, что гены распределены по цепям неслучайно и не равномерно у бактерии *Arthrobacter sp.* 3. Однако такую возможность нельзя полностью исключать, так как полностью возможность случайного распределения не исключается.

Невозможно переоценить достоинства использования программы Excel для анализа геномов микроорганизмов. С помощью этой программы можно выявить закономерности в распределении генов, выдвинуть или опровергнуть какие-либо гипотезы, сравнить данные о новом виде с уже известными.

СОПРОВОДИТЕЛЬНЫЕ МАТЕРИАЛЫ

По этой ссылке можно найти файл с расчетами и формулами, которые не были приведены в статье:

<http://kodomofbb.msu.ru/~polina.shpudeiko/term1/Arthrobacter.xlsx>

По этой ссылке находится файл с описанием генома, взятый с сайта NCBI:

http://kodomofbb.msu.ru/~polina.shpudeiko/term1/GCF_000755585.2_ASM75558v2_feature_table.xlsx

БЛАГОДАРНОСТИ

Хотелось бы поблагодарить преподавателей курса практической информатики факультета биоинженерии и биоинформатики МГУ им. М. В. Ломоносова, которые поставили данную задачу.

ССЫЛКИ НА ИСТОЧНИКИ ИНФОРМАЦИИ

1. [MP 4.2.0020-11. 4.2. Методы контроля. Биологические и микробиологические факторы. Фенотипическая идентификация бактерий рода *Corynebacterium*. Методические рекомендации", М., Федеральный центр гигиены и эпидемиологии Роспотребнадзора, 2011.](#)
2. microbewiki.kenyon.edu описание рода *Arthrobacter*
3. wikipedia.org описание рода *Arthrobacter*
4. ncbi.nlm.nih.gov описание генома бактерии
5. humbio.ru статья о некодирующих РНК
6. wikipedia.org статья о псевдогенах