

Мини-обзор протеома и генома бактерии *Streptomyces avermitilis* MA-4680=NBRC 14893

Потемкина Елена

Факультет биоинженерии и биоинформатики МГУ им.М.В. Ломоносова

РЕЗЮМЕ

Данная работа посвящена исследованию и анализу протеома и генома бактерии *Streptomyces avermitilis* MA-4680=NBRC 14893. Было определено распределение генов на прямой и обратной цепях, число генов белков и РНК по категориям. Так же был осуществлен анализ распределения белков бактерии по их длинам.

1 ВВЕДЕНИЕ

Бактерии рода *Streptomyces* широко распространены в природе, особенно в почве. Известны как важнейшие промышленные продуценты многих антибиотиков и других вторичных метаболитов антибактериальной, антигрибковой, антивирусной и противоопухолевой природы.



Рисунок 1. *Streptomyces avermectinius (avermitilis)* : sole source of the avermectins (credit: Kitasato Institute)

Streptomyces играют важную роль в почвенной среде, потому что разнообразные метаболические пути позволяют им разлагать плохо разлагающиеся остатки других организмов, включая хитин и лигноцеллюлозу. Особенностью представителей этого рода является то, что они содержат очень большие (примерно 8 миллионов пар нуклеотидов, что в два раза больше чем большинство геномов бактерий) линейные хромосомы, в отличие от многих других бактерий, которые имеют кольцевые хромосомы.

Вид *Streptomyces avermitilis* (рис. 1) является широко известным продуцентом авермектина, используемое как основное вещество лекарственных препаратов против инвазий, вызываемые нематодами и членистоногими.

2 МЕТОДЫ

Данные для исследования были взяты с сервера NCBI. Для обработки информации применялась программа Microsoft Office Excel 2007.

Для подсчета числа генов белков (CDS) и генов РНК на прямой и комплементарной цепи использовалась функция "СЧЁТЕСЛИМН".

Для определения числа генов белков и РНК по категориям использовалась функция "СЧЁТЕСЛИ".

Для распределения белков бактерии по длинам были выбраны карманы длин на основе данных о максимальной длине белка, минимальной длине белка, среднем значении длины и медиане значений. Для получения этой информации использовались функции "МИН", "МАКС", "СРЗНАЧ", "МЕДИАНА". С учётом полученных данных была построена гистограмма длин белков *Streptomyces avermitilis* MA-4680=NBRC 14893.

3 РЕЗУЛЬТАТЫ

3.1 Анализ генома

Геном *Streptomyces avermitilis* MA-4680=NBRC 14893 представлен линейной двуцепочечной молекулой ДНК и плазмидой SAP1.

Геном имеет размер примерно 9124287 пар нуклеотидов (далее п.н.) и содержит 7770 генов. Значит, на 1 млн. п.н. приходится приблизительно 852 гена. Их распределение по прямой и обратной цепям представлено в таблице 1. На (+) цепи - 3821 генов, на (-) - 3949.

Таблица 1. Распределение генов по прямой и обратной цепям в геноме *Streptomyces avermitilis* MA-4680=NBRC 14893

Цепь	CDS	RNA
(+)	3774	47
(-)	3903	46

Количества генов белков (CDS) и генов РНК (RNA) по категориям продемонстрировано в таблице 2.

Таблица 2. Количество генов белков (CDS) и генов РНК (RNA) *Streptomyces avermitilis* MA-4680=NBRC 14893 по категориям

Ген	Количество
CDS	7677
rRNA	18
tmRNA	1
ncRNA	2
tRNA	72

3.2 Анализ протеома

В геноме *Streptomyces avermitilis* MA-4680=NBRC 14893 7431 генов, кодирующих белки. Также в нем присутствуют последовательности CDS, называемые "pseudogene" (псевдогены), утратившие способность к экспрессии. Их количество считается как разность общего количества генов и суммы числа генов, кодирующих РНК, и числа конечных белковых продуктов и равно 246. Данные о длинах белков,

полученные и использованные в ходе обзора протеома приведены в таблице 3.

Таблица 3. Данные о длинах белков *Streptomyces avermitilis* MA-4680=NBRC 14893

Характеристика	Значение
Минимальная длина	29
Максимальная длина	7746
Среднее значение длины	336
Медиана значений длины	285

На основе полученных статистических данных были выбраны карманы длин шагом в 50 аминокислотных остатков (далее а.к.). Распределение белков *Streptomyces avermitilis* MA-4680=NBRC 14893 по длине представлено на гистограмме (рис. 2). Для того, чтобы гистограмма получилась наглядной, диапазон длин от 1300 а.к. до максимальной 7746 а.к. были объединены в общих столбец ">1300".

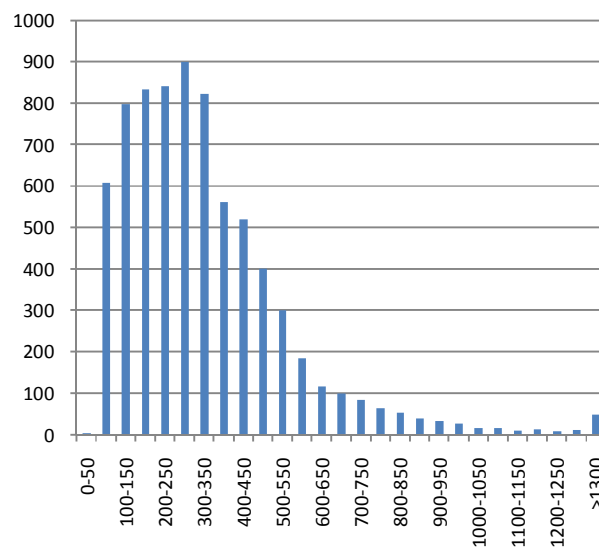


Рисунок 2. Распределение белков *Streptomyces avermitilis* MA-4680=NBRC 14893 по длине (по вертикали - количество белков, по горизонтали - диапазоны длин)

4 ОБСУЖДЕНИЕ

Высокая плотность расположения генов в ДНК у *Streptomyces avermitilis* MA-4680=NBRC 14893 типична для прокариотных организмов, которые, в отличие от эукариотных, не имеют интронов и почти вся их ДНК является кодирующей.

Распределение генов между прямой и обратной цепями практически одинаковое. Таким образом, можно предположить, что это распределение независимо, но нужно провести еще ряд расчетов для подтверждения гипотезы.

Кроме генов, кодирующих белки (CDS), рибосомальные РНК (rRNA), транспортные РНК (tRNA), в геноме присутствуют гены, кодирующие ранее неизвестные для меня "транспортно-матричные" РНК (tmRNA) и некодирующие РНК (ncRNA). tmRNA сочетает функции транспортной и матричной РНК и участвует в терминации трансляции по механизму, называемому транс-трансляцией. ncRNA играют роль в эпигенетическом сигнале.

Из-за присутствия в геноме псевдогенов число представленных CDS последовательностей не соответствует числу конечных белковых продуктов.

Исследование протеома показало, что большинство белков короткие. Это наглядно можно наблюдать по смещенной гистограмме распределения длин белков. Большая их часть имеет длину 250-300 аминокислотных остатков. Такое распределение вполне типично для многих бактерий.

Таким образом, хотя геном исследуемой бактерии представлен линейной двуцепочечной молекулой ДНК, что в не типично для прокариот, *Streptomyces avermitilis* MA-4680=NBRC 14893 проявляет признаки, характерные для прокариотных организмов.

5 СОПРОВОДИТЕЛЬНЫЕ МАТЕРИАЛЫ

Таблицы и расчеты по ссылке:

- http://kodomo.fbb.msu.ru/~potemkina_helen/Potemkina_pr13.xlsx
- http://kodomo.fbb.msu.ru/~potemkina_helen/pr12_2.xlsx

6 СПИСОК ЛИТЕРАТУРЫ

- <https://www.ncbi.nlm.nih.gov/genome/?term=Streptomyces+avermitilis>
- <https://ru.wikipedia.org/wiki/ТмРНК>
- [http://www.cell.com/trends/parasitology/fulltext/S1471-4922\(14\)00111-1](http://www.cell.com/trends/parasitology/fulltext/S1471-4922(14)00111-1)
- <http://medbiol.ru/medbiol/epigenetica/000913dc.htm>