

Мини-обзор протеома бактерии *Streptomyces avermitilis* MA-4680

Бормотов Игорь^{1*}

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

РЕЗЮМЕ

Данная работа представляет собой анализ протеома бактерии *Streptomyces avermitilis* MA-4680. В процессе выполнения работы мною было установлено число генов, кодирующих белки, транспортные, рибосомальные и другие типы РНК, их распределение по прямой и комплементарной цепям, а также распределение длин белков в геноме бактерии.

1 ВВЕДЕНИЕ

Широко распространенные преимущественно почвенные бактерии. В природе встречается 364 разновидности рода Стрептомицеты, многие из которых являются промышленными производителями антибиотиков и вторичных метаболитов, таких как противогрибковых, противовирусных и средств от опухолей, и т.д.

Еще одной важной особенностью Стрептомицетов является длина их цепи ДНК (~8 млн. пар нуклеотидов, что вдвое больше чем у большинства бактерий), представленной линейной хромосомой с отличными теломерами



Рисунок 1. *Streptomyces avermitilis*

Таблица 1. Систематическое положение *Streptomyces avermitilis*

Домен:	<i>Bacteria</i>
Тип:	<i>Actinobacteria</i>
Порядок:	<i>Actinomycetales</i>
Род:	<i>Streptomycetaceae</i>
Семейство:	<i>Streptomyces</i>
Вид:	<i>S. avermitilis</i>

2 МАТЕРИАЛЫ И МЕТОДЫ

При выполнении работы я воспользовался файлом с сайта NCBI [feature_table.txt](#), содержащим необходимую информацию о геноме *Streptomyces avermitilis* MA-4680. Также для обработки данных мною была использована программа Microsoft Office Excel 2007. При работе с программой использовались статистические и логические функции, такие как фильтр, СУММ, СЧЁТЕСЛИ и другие.

3 РЕЗУЛЬТАТЫ

Протеом *Streptomyces avermitilis* MA-4680 представлен кольцевой молекулой ДНК (9, 025, 608 п.н.).

3.1 Количественное и процентное содержание генов, кодирующих белки и различные типы РНК

При анализе генома *S. avermitilis* я выяснил, что рассматриваемая бактериальная хромосома содержит 7,680 генов, кодирующих белки, что составляет около 98,8% всех генов, содержащихся в протеоме, тРНК кодируются примерно 0,9% генов, а рибосомальные и некоторые другие молекулы РНК кодируются 0,2 и 0,1% генов соответственно.

Таблица 2. Содержание генов в протеоме *S. avermitilis* MA-4680

Типы молекул	Число генов	Соотношение (%)
Белки	7680	98,8
тРНК	72	0,9
рРНК	18	0,2
другие РНК	3	>0,1

Учитывая, что на 9, 025, 608 п.н. приходится 7680 генов, примерная оценка числа генов на 1 млн п.н. составляет 851.

3.2 Распределение генов на прямой и комплементарной цепях

При подсчете генов, находящихся на противоположных цепях ДНК, я обнаружил, что прямая цепь содержит меньше генов, чем цепь, ей комплементарная.

Из таблицы 3 видно, что отклонение распределения генов по цепям от идеального (1:1) составляет около 65 случаев из 7680. Рассмотрев упрощенную модель случайного распределения с вероятностью происхождения каждого из двух событий 0,5, можно заметить, что из 148 выборок в среднем лишь 6 из них имеют то же или большее значение отклонения (каждая выборка содержит 3863 события). Такая низкая вероятность подобного отклонения (0.04) убеждает меня в том, что гены распределены по цепям неслучайно.

Таблица 3. Распределение генов по прямой и комплементарной цепям

Число генов	Прямая цепь	Комплементарная цепь
Белков	3775	3905
РНК	47	46

3.3 Распределение длин белков

В процессе исследования генома *S. avermitilis* MA-4680 я выяснил, что большинство закодированных в нем белков имеют длину от 200 до 300 аминокислотных остатков (таких белков в геноме 1748, что составляет около 23% всех закодированных белков), 21% составляют белки длиной от 100 до 200 аминокислотных остатков, белки, имеющие меньшую длину, составляют около 8%, а белки, состоящие более чем из 300 аминокислотных остатков – 48%. Более полные данные по распределению длин белков представлены на рисунке 2:

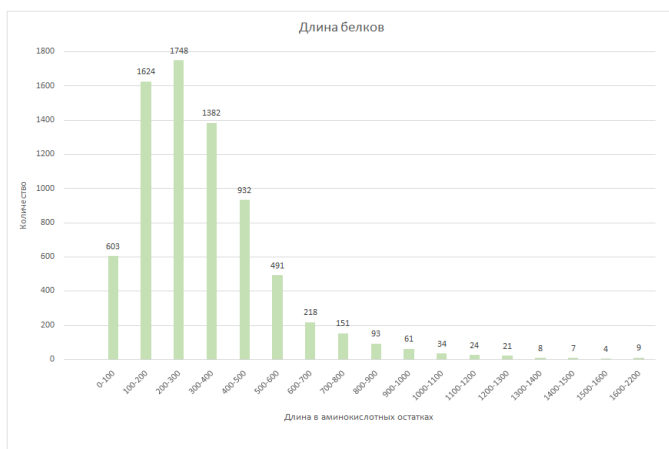


Рисунок 2 Распределение длин белков *S. avermitilis* MA-4680 (по горизонтальной оси отмечены длины белков в аминокислотных остатках, по вертикальной оси – количество белков, закодированных в геноме)

4 ОБСУЖДЕНИЕ

Как уже было отмечено в пункте 3.2, распределение генов по прямой и комплементарной цепям не совпадает со случайным распределением и скорее всего имеет какую-либо причину (к примеру, определенные гены могут формировать опероны или же могли произойти мутации).

Также неочевидно, почему в геноме закодировано именно 72 тРНК. Всего существует 61 вариант триплетов, кодирующих аминокислоты. Видимо, увеличенное количество тРНК связано с наличием неоднозначных пар оснований (образованием связей типа G-U и I-U / I-A / I-C).

5 ЗАКЛЮЧЕНИЕ

Полученное распределение длин белков можно считать нормой для бактериального организма. Более 57% входят в диапазон от 100 до 400 а.о. Чуть более 0,004 процента имеют длину более 1500 а.о. Максимальная длина белка 7747 а.о. Соотношение генов РНК в обоих цепях 1:1.

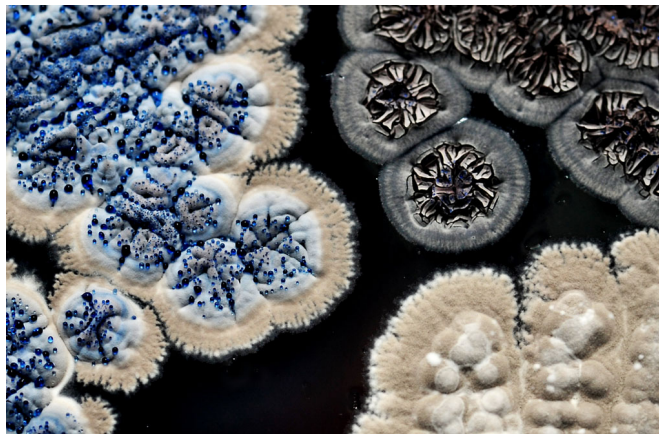


Рисунок 2. *Streptomyces avermitilis*

6 СОПРОВОДИТЕЛЬНЫЕ МАТЕРИАЛЫ

Ссылка на xlsx файл с расчетами и результатами:
http://kodomofbb.msu.ru/~bormotov/term1/Bormotov_pr13.xlsx

7 СПИСОК ЛИТЕРАТУРЫ

- https://ru.wikipedia.org/wiki/Streptomyces_avermitilis страница Википедии с информацией, касающейся *Streptomyces avermitilis* MA-4680
- <https://www.ncbi.nlm.nih.gov/genome/?term=Streptomyces%20avermitilis%20MA-4680%20=%20NBRC%2014893> - страница обзора организма на NCBI
- <https://www.ncbi.nlm.nih.gov/nuccore/BA000030.4> страница сайта NCBI с информацией, касающейся про-теома *Streptomyces avermitilis* MA-4680