

Мини-обзор генома бактерии *Streptomyces globisporus* C-1027

Юдина А.С.^{1*}

РЕЗЮМЕ

В своей работе я анализировала протеом бактерии *Streptomyces globisporus* C-1027. Этот микроорганизм представляет интерес для науки, т.к. способен синтезировать противоопухолевый агент C-1027.

С развитием компьютерных технологий описывать геномы и протеомы различных организмов стало проще, используя такие программы как Excel. Для анализа и наглядного представления данных из генома бактерии *Streptomyces globisporus* C-1027 в представленной работе использована данная программа.

Результаты моей работы отражены в таблицах и гистограмме.

1 ВВЕДЕНИЕ

Streptomyces globisporus C-1027 - грам-положительная бактерия, относится к порядку Actinomycetales к семейству Streptomycetaceae. Обитает в почве. Работы по изучению данного организма ведутся в Фармакологическом центре исследования микроорганизмов в Китае (Center for Culture Collection of Pharmaceutical Microorganisms in China). Интерес для науки бактерия представляет из-за продукции антибиотика C-1027, который широко используется в качестве противоопухолевого агента. (материал взят из Wikipedia.org)

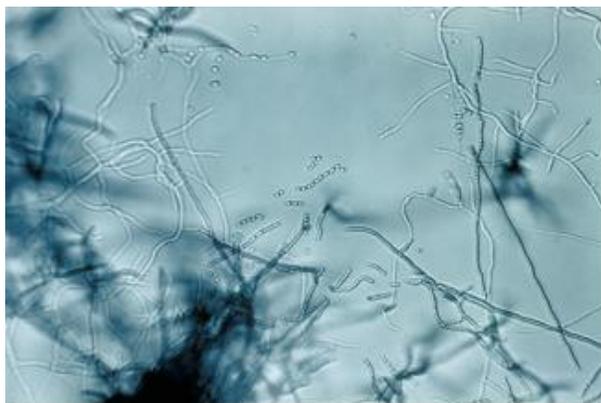


Рис. 1. Внешний вид бактерий рода *Streptomyces* (взято из Wikipedia.org). Детальное описание морфологии рассматриваемого штамма приведено в статье «A new macromolecular antitumor antibiotic, C-1027»

При исследовании данного микроорганизма в Институте медицинской биотехнологии в Китае было выявлено следующее:

- Ферменты, выделяемые бактерией, подавляют жизнедеятельность большинства Грам-положительных бактерий, но никак не воздействует на представите-

лей рода *Mycobacterium* и Грам-отрицательных бактерий.

- Ферменты способны подавлять деление опухолевых клеток. Это было доказано экспериментально на культуре клеток карциномы, затем на лабораторных мышках. В некоторых случаях C-1027 по своей активности превосходил известные противоопухолевые агенты. (результаты исследований Института медицинской биотехнологии)

В данной работе я сделала акцент на анализе протеома бактерии, т. к. детальное изучение продуктов генной экспрессии, закономерности их распределения в соответствии с различными параметрами могут подсказать более эффективный метод получения C-1027 агента.

2 РЕЗУЛЬТАТЫ

В своей работе я рассматриваю геном микроорганизма по двум критериям:

- 1) Распределение генов в геноме, т.е. как много генов транскрибируется с прямой и обратной цепей.
- 2) Наиболее часто встречаемые длины белков.

2.1 Распределение генов по цепям

2.1.1 Геном бактерии *Streptomyces globisporus* C-1027 представлен кольцевой хромосомой и двумя плазмидами (SGLP1 и rSGL1). Гены закодированы на прямой и обратной цепях.

Длина генома примерно 7 608 110 п.н., количество генов - 6879. То есть около 905 генов приходится на 1 миллион п.н.. Можно сказать, что гены на цепи расположены достаточно плотно.

Количество генов белков и генов РНК по категориям сведено в таблицу 1.

ncRNA	tRNA	tmRNA	rRNA	CDS
2	67	1	18	6791

Таблица 1. Число генов *Streptomyces globisporus* C-1027 по категориям.

Из таблицы видно, что большая часть генов кодирует белки, почти в сто раз меньше генов, кодирующих транспортные РНК, также имеются рибосомальные, не кодирующие и «транспортно-матричная» РНК. «Транспортно-матричная»

*To whom correspondence should be addressed.

РНК предназначена для завершения трансляции пептида с транспортной РНК, у которой по каким-то причинам не функционирует стоп-кодон. Такой способ завершения трансляции является у бактерий достаточно редким, в большинстве случаев для этих целей используется АТФ-зависимые протеазы.

(взято из статьи из журнала *Genes Dev.*)

Эти данные я получила из базы данных NCBI, упорядочив их в плоскую таблицу.

2.1.2 Как уже упоминалось выше, гены закодированы на прямой и обратной цепи ДНК. Я составила таблицу 2, иллюстрирующую их распределение.

Strand	RNA	CDS
+		47 3380
-		41 3411

Таблица 2. Распределение генов *Streptomyces globisporus* C-1027 по цепям ДНК.

Гены по комплементарным цепям распределены очень равномерно с незначительным превышением генов CDS на обратной цепи и генов RNA на прямой цепи. Такое распределение является интересной особенностью данного микроорганизма.

2.2 Распределение длин белков

Следующий критерий, на который я обратила внимание, это длины белков, производимых микроорганизмом.

Белковых продуктов получается меньше, чем генов в геноме, так как гены РНК не кодируют белки и, как видно из таблицы, взятой из базы NCBI, у некоторых последовательностей CDS нет конечного белкового продукта. В большинстве случаев, это псевдогены, такие гены, в которых произошла поломка; у *Streptomyces globisporus* C-1027 псевдогенов 149 (вычла из общего числа генов количество генов РНК и конечных белковых продуктов).



Рис.2. Распределение длин белков *Streptomyces globisporus*

Делать какие-либо выводы о микроорганизме, основываясь на длинах его белков, удобно, располагая гистограммой длин белков (Рис.2). Для ее построения я использовала программу Excel.

На гистограмме виден один пик, это означает, что больше всего белков находится в диапазоне длин от 60 до 390 аминокислотных остатков. Самое большое количество в интервале 210-240.

Число белков резко возрастает, а затем очень медленно снижается. Разброс наблюдается от 29 до 13403 аминокислотных остатков, причем, начиная с длины в 1350 остатков, белков становится очень мало, и для наглядности последние столбцы пришлось объединить в один.

3 ОБСУЖДЕНИЕ И ЗАКЛЮЧЕНИЕ

Сведем все замеченные особенности генома *Streptomyces globisporus* C-1027 в один список.

- 905 генов на 1 млн. пар нуклеотидов. Гены достаточно плотно расположены на цепи, если сравнить с расположением генов в эукариотических клетках. Я предполагаю, что это нормально для микроорганизмов, т.к. в клетке без ядра важно, чтобы генетический материал был доступно организован и не занимал много места.
- В распределении генов по категориям встречается единственная «транспортно-матричная» РНК, судя по материалам других статей, такой тип РНК встречается довольно редко. Как уже было сказано выше, эта РНК служит для завершения трансляции пептидов с транспортной РНК. Эту же функцию в других организмах выполняют ферменты пептидазы. Возможно, в дальнейших исследованиях стоит обратить внимание на то, почему данный штамм отказался от использования стандартных методов прерывания трансляции.
- Интересным, на мой взгляд, фактом является равномерное распределение генов на прямой и обратной цепях ДНК. Предполагаю, что так происходит для дополнительного уменьшения размера генома микроорганизма.
- Наличие последовательностей CDS, которые не производят белковые продукты, тоже заслуживают внимания. Я предполагаю, что это псевдогены, с которых из-за поломки (вставки в середину стоп-кодона, например) больше не возможно синтезировать белок. Хотелось бы попытаться изучить свойства этих псевдогенов. Возможно, они кодируют ферменты, увеличивающие активность антибиотика, производимого данным штаммом. Или же, если удастся «включить» эти гены, можно будет точно установить родство этих микроорганизмов с другими группами. В статье из журнала «The journal of antibiotics» опи-

сывается каким образом открытый штамм был отнесен к виду *globisporus*, но с исправленными псевдогенами этому нашлось бы больше подтверждений.

- Распределение белков по длинам, я считаю, достаточно обычным, т.к. у многих бактерий из других родов наблюдается схожая картина гистограммы. Заслуживает внимания, я считаю, наличие такой максимальной длины белка - 13403 аминокислотных остатков. Было бы интересно выделить его и исследовать его свойства. Скорее всего, он выполняет строительную функцию, так как для каталитической он слишком большой, по сравнению с другими белками бактерии.

Анализ генома с использованием компьютерных программ полезен тем, что позволяет охарактеризовать свойства микроорганизма с помощью статистических методов. Данные, полученные на их основе, могут стать основанием для новых гипотез, доказать или опровергнуть высказанные ранее предположения.

4 СОПРОВОДИТЕЛЬНЫЕ МАТЕРИАЛЫ

Для наглядности в статье приведены только конечные результаты работы в программе Excel. Работы с расчетными формулами и промежуточными результатами доступны по ссылкам:

1. Плоская таблица с описанием генома - http://kodomofbb.msu.ru/~stacyud/term1/Yudina_pr12_table.xlsx
2. Файл с материалами для гистограммы и таблиц - [http://kodomofbb.msu.ru/~stacyud/term1/Yudina_pr13\(2\).xlsx](http://kodomofbb.msu.ru/~stacyud/term1/Yudina_pr13(2).xlsx)

5 МАТЕРИАЛЫ И МЕТОДЫ

1. Геном скачан из базы данных Genome на сайте NCBI: NCBI Genome => Browse by organism => Search by organism.
2. Для построения таблиц, гистограммы и анализа данных мной использовалась программа Excel из пакета MsOffice.
3. При работе в программе я использовала функции: МАКС, МИН, СЧЁТЕСЛИМН для построения гистограммы; СЧЁТЕСЛИМН, СЧЁТЕСЛИ для построения таблиц.
4. Плоская таблица с основными характеристиками генома построена на основе таблицы из базы Genome. Из оригинала я удалила неинформативные строки *gene* с помощью фильтра, удалила пустые столбцы, отсортировала строки по координатам старт-кодона.

БЛАГОДАРНОСТИ

Работа состоялась благодаря преподавателям биоинформатики на Факультете биоинженерии и биоинформатики МГУ им М. В. Ломоносова.

Отдельные благодарности Алексеевскому Андрею Владимировичу, поставившему данную задачу.

СПИСОК ЛИТЕРАТУРЫ

- Gottesman S. (1998) *Genes Dev* 1338-47. The ClpXP and ClpAP proteases degrade proteins with carboxy-terminal peptide tails added by the SsrA-tagging system.
- Hu Jilan (1988) *The journal of antibiotics* VOL. XLI NO. II: A new macromolecular antitumor antibiotic, C-1027. Institute of Medicinal Biotechnology, Chinese Academy of Medical Sciences, Beijing, People's Republic of China.
- Wikipedia.org.(2016) *Streptomyces globisporus*