

Обзор генома бактерии *Gemmatimonas phototrophica*

Шкарина Анастасия

Факультет биоинженерии и биоинформатики МГУ им. Ломоносова

РЕЗЮМЕ

В этой работе я исследовала геном бактерии *Gemmatimonas phototrophica*. Я проанализировала число генов на разных цепях ДНК, распределение генов по категориям и по длинам синтезируемых ими белков, проверила гипотезу о том, что белки распределены по цепочкам ДНК с вероятностью 0.5. Также я исследовала «квазиопероны» и пересекающиеся гены.

1 ВВЕДЕНИЕ

Gemmatimonas phototrophica один из двух видов, относящихся к роду *Gemmatimonas*, единственному в типе *Gemmatimonadetes* [1]. Это вид краснопигментированных полуаэробных бактерий, которые были найдены в пресноводном озере в пустыне Гоби. Бактерия *Gemmatimonas phototrophica* имеет полностью функционирующий фотосинтезирующий реакционный центр 2 типа (феофитин-хиноновый), но при этом она не усваивает неорганический углерод, что означает, что эта бактерия фотогетеротроф. В её геноме есть кластер фотосинтезирующих генов (PGC), из-за устройства и филогении которых сделали вывод, что в древности произошёл горизонтальный перенос этого кластера от пурпурных фототрофных бактерий. [2].

2 МЕТОДЫ

Для работы мне потребовались данные из базы NCBI. Обработка данных, построение гистограмм и таблиц была произведена с помощью Microsoft Office Excel 2013. Использовались функции «СЧЕТЕСЛИ», «СЛУЧМЕЖДУ», «ЕСЛИ», «МАКС» и другие.

3 РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЯ

3.1. Общие сведения о геноме бактерии.

Геном бактерии *Gemmatimonas phototrophica* состоит из двухцепочечной ДНК длиной 4716552 нуклеотидных пар. В нём 3442 гена. Следовательно, на 1 млн. н.п. приходится примерно 730 генов. В таблице 1 показано распределение генов по категориям.

3.2 Распределение генов по цепям.

Распределение генов по цепям показано в таблице 2. На прямой цепи 1652 гена, а на обратной – 1790 генов. Я проверила случайно ли данное распределение или нет. Разница между

количеством генов на прямой и обратных цепочках ДНК большая или равная реальной получается примерно в 20 случаях из 1000. Можно предположить, что данное распределение неслучайно.

Таблица 1. Таблица Распределения генов по категориям.

Категория	Количество
Гены белков	3388
Гены tRNA	47
Гены rRNA	6
Гены ncRNA	1

Таблица 2. Таблица числа генов белков и генов РНК на прямой и комплементарной цепочках ДНК.

	Число генов белков	Число генов РНК	Всего
Прямая цепь	1626	26	1652
Обратная цепь	1762	28	1790

3.3. Распределение белков по длинам.

На рисунке 1 представлена гистограмма распределения белков по длинам. Из него видно, что больше всего белков имеет длину в 100 – 200 аминокислот. Максимальная длина белка – 2787 аминокислот. Медиана равна 303 аминокислоты.

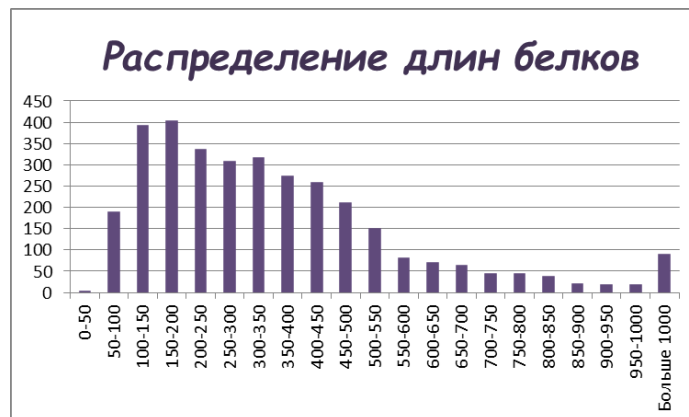


Рис. 1. Распределение белков по длинам. (По вертикали количество белков в данном диапазоне длин, по горизонтали диапазоны длин белков.)

3.4. «Квазиопероны» в геноме бактерии.

«Квазиопероном» – это максимальная последовательность генов, закодированных на одной цепочке с промежутками между генами не более порога 100 п.н. В геноме бактерии *Gemmatimonas phototrophica* 1935 квазиоперонов. Если увели-

читать порог между генами, то количество квазиоперонов увеличится. Наибольшее количество генов в квазиопероне 12. На рис. 2 представлена гистограмма количества генов в квазиоперонах.

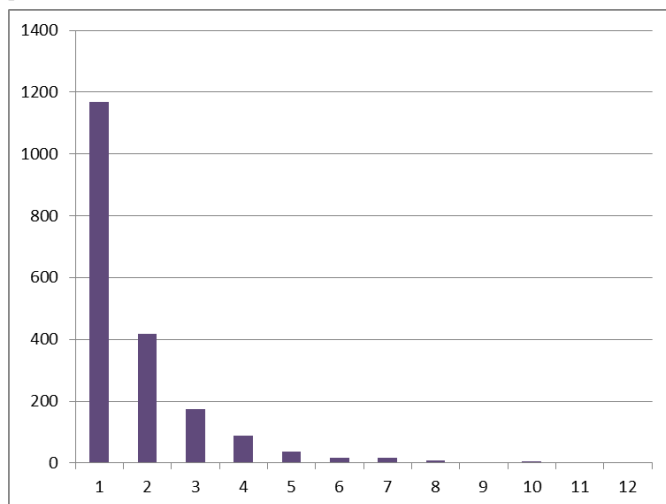


Рис. 2. Гистограмма распределения квазиоперонов по числу генов.

3.5. Данные о пересечениях генов.

В геноме бактерии 394 гена пересекающихся гена. В таблице 3 представлены данные об этих генах. Большинство(87,8%) пересекающихся генов лежат на одной цепи. У 73,6% генов нет сдвига рамки считывания, у 23,3% генов происходит сдвиг рамки считывания на 1 нуклеотид и у 3% рамка сдвигается на 2 нуклеотида.

Таблица 3. Данные о пересекающихся генов

Характеристика	Число генов
Всего пересекающихся генов	394
Пересекающиеся гены, лежащие на одной цепи	346
Пересекающиеся гены, лежащие на разных цепях	48
Гены со сдвигом рамки считывания на 1 нуклеотид	92
Гены со сдвигом рамки считывания на 2 нуклеотида	12
Гены без сдвига рамки считывания	290

Сопроводительные материалы

Ссылка на Excel файл с результатами:

http://kodomo.fbb.msu.ru/~anshkarina/files/Shkarina_pr14.xlsx

Ссылки:

[1]Wikipedia Gemmatimonas.

[2]Functional type 2 photosynthetic reaction centers found in the rare bacterial phylum Gemmatimonadetes. Zeng Y1, Feng F2, Medova H1, Dean J1, Koblizek M3.