

Мини-обзор генома бактерии *Methylococcus capsulatus*

Дрождев Алексей

Факультет биоинженерии и биоинформатики МГУ им. М.В. Ломоносова
Ленинские горы МГУ 1 стр. 73, г. Москва, 119234, Российская федерация
Dr.alexey7@kodomo.fbb.msu.ru

Резюме

Был проанализирован геном бактерии *Methylococcus capsulatus*, использующей метан в качестве источника углерода. При помощи программы Excel были проанализированы длины белков бактерии, их распределение. Также мы рассмотрели пересечения генов на цепях и обнаружили, что между прямой и обратной цепью имеются различия в минимальной длине пересечения. Было обнаружено, что большая часть пересечений имеет одинаковую небольшую длину. Были получены данные о квазиоперонах каждой из цепочек, существенных различий между ними обнаружено не было.

Ключевые слова:

Methylococcus capsulatus, геном, пересечение генов, квазиоперон, excel

Введение

Methylococcus capsulatus – вид грамотрицательных бактерий, которые используют метан в качестве источника углерода. Клетки инкапсулированы и часто располагаются по две. Помимо метана могут окислять метанол, формальдегид и муравьиную кислоту. Оптимальная температура роста бактерии составляет около 37°C, может жить в условиях, в которых мало молекулярного кислорода. Способны фиксировать атмосферный азот.^{[3][4]}

Метан является гораздо более сильным парниковым газом, чем диоксид углерода, поэтому данный вид бактерий имеет большое экологическое значение для сокращения количества этого газа, выделяющегося в атмосферу земли.^[1]

Также используется в сельском хозяйстве для производства кормов из природного газа.^[2]

Материалы и методы

Информация о геноме данной бактерии была получена с сайта NCBI:

<https://www.ncbi.nlm.nih.gov/genome/genomes/1005> (Штамм Bath, база данных GenBank)

Для анализа данных была использована программа MS Office Excel, с её помощью были созданы все приведенные таблицы и диаграммы.

Для анализа длин белков рассматривались строки со значением “CDS” в колонке “feature” и “with_protein” в колонке “class”.

При анализе генов рассматривались строки со значением “gene” в столбце “feature”.

Значения столбца “class” интерпретировались следующим образом:

“protein_coding” – белок

“pseudogene” – не экспрессируется

“tRNA”, “misc_RNA”, “rRNA” – РНК

Результаты и обсуждение

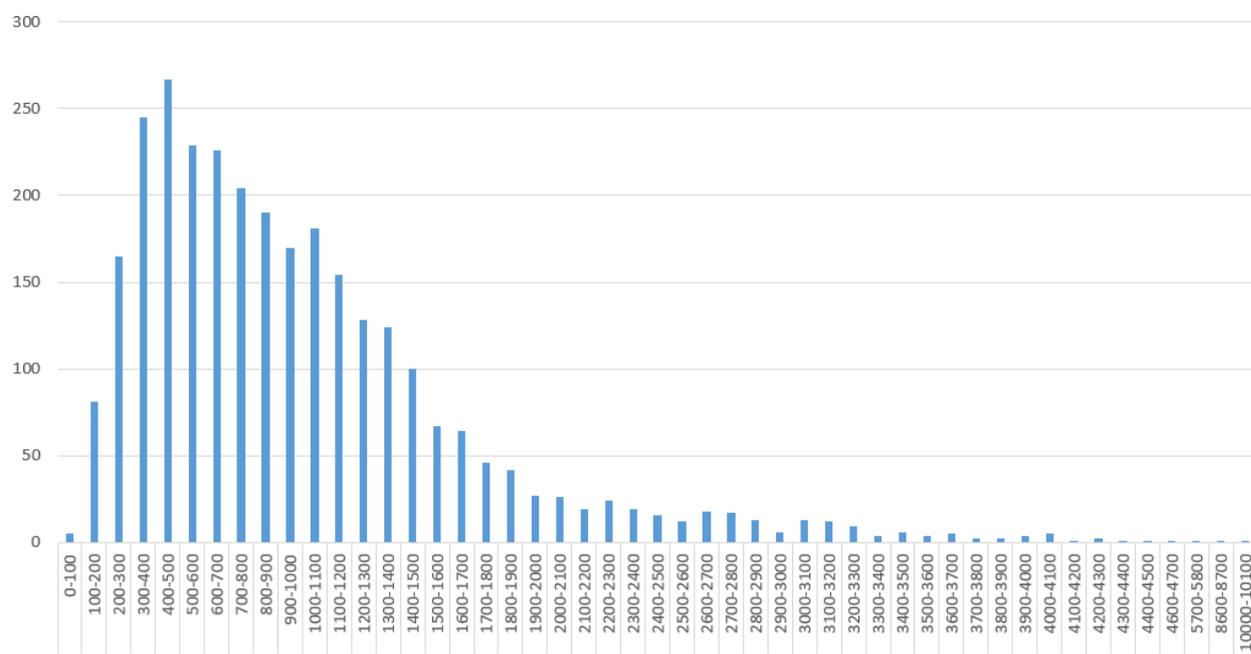
Длины белков

Геном *Methylococcus capsulatus* содержит 2960 белков. Их длина изменяется от 25 до 3350 аминокислотных остатков. Самый длинный белок в данном геноме – гемагглютинин-зависимый белок. Средняя длина белка составляет 333 а.к., в то время как медианная – 277 а.к., из чего можно сделать вывод, что распределение длин белков не будет нормальным. Реальное распределение можно увидеть на гистограмме 1.

Среднеквадратичное отклонение составляет 243 а.к., из чего можно сделать вывод, что белковый состав бактерии будет весьма разнообразным.

Минимальная длина	25
Максимальная длина	3350
Среднее	332,9873
Среднеквадратичное отклонение	243,4927
Медиана	277

Таблица 1. Статистические данные о длинах белков



Гистограмма 1. Распределение длин белков

Распределение

Было подсчитано количество генов, кодирующих рибосомальные, транспортные и гипотетические белки, а также тРНК и рРНК.

	Белки			
	Рибосомальные	Транспортные	Гипотетические	Остальные
количество	63	136	815	1946

Таблица 2. Распределение белков по категориям

	РНК		
	тРНК	рРНК	Остальные
количество	46	6	2

Таблица 3. Распределение РНК

Также было проанализировано распределение генов, кодирующих белки, РНК и псевдогены по цепочкам.

Цепь	Число генов белков	Число псевдогенов	Число генов РНК
Прямая	1503	30	33
Комплементарная	1457	13	21

Таблица 4. Распределение по цепочкам

На прямой цепи оказалось существенно больше псевдогенов и генов, кодирующих РНК.

Пересечения

Было подсчитано число пересечений генов для каждой цепи в отдельности. На прямой цепи было обнаружено 276 пересечений. Средняя длина пересечения составила 10,72 нуклеотида. Наименьшее перекрытие составило 1 нуклеотид. Самое большое пересечение составило 533 нуклеотида, оно возникло при перекрытии гена, кодирующего рРНК, и гена, кодирующего белок. Следующее за ним составило 169 нуклеотидов и возникло на стыке гена, кодирующего белок, и не экспрессирующегося гена. Отбросив эти два крайних значения, была построена диаграмма длин пересечений генов прямой цепи. По горизонтальной оси отложены координаты начала пересечения, по вертикальной – длина пересечения.



Диаграмма 1. Длины пересечений на прямой цепи

Как видно из диаграммы, большинство пересечений (157 из 276) имело длину 4.

Подобный анализ был также проведен для генов, расположенных на комплементарной цепи, на ней число пересечений составило 249, средняя длина которых составила 14,85 нуклеотидов, что существенно больше, чем на прямой цепи. Интересно, что самое

большое пересечение (1434 нуклеотида) также было образовано в результате перекрытия гена, кодирующего рРНК и гена, кодирующего белок, а следующее за ним (157 нуклеотидов) – также перекрытием не экспрессирующегося гена и гена, кодирующего белок. Отбросив эти крайние значения, была построена аналогичная первая диаграмма.

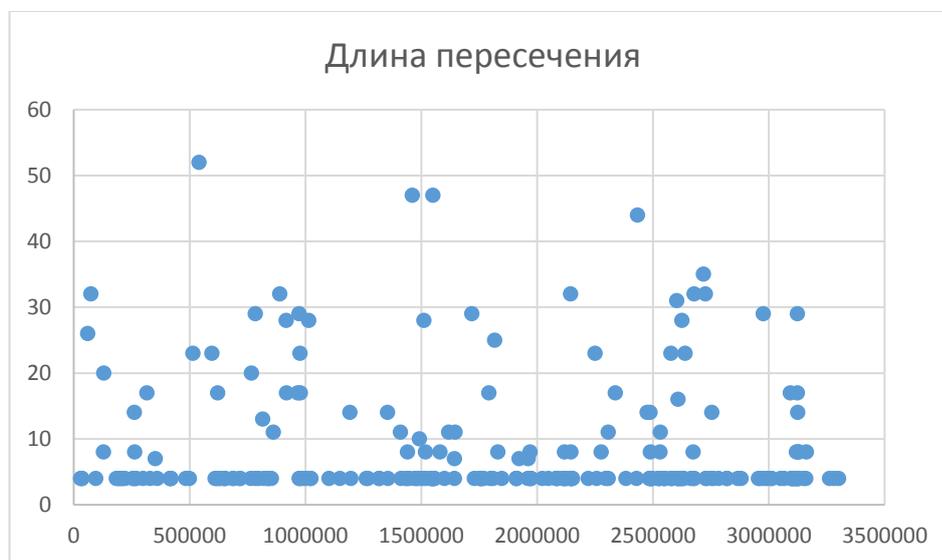


Диаграмма 2. Длины пересечений на комплементарной цепи

Интересно, что в отличие от основной, на комплементарной цепи не было пересечений длины 1.

Большинство пересечений (171 из 246) также имело длину 4.

Квазиопероны

На прямой цепи было выявлено 678 квазиоперонов, их средняя длина составила 2,31 гена, максимальная длина – 24 гена.

На комплементарной цепи было обнаружено 658 квазиоперонов, их средняя длина составила 2,27 гена, максимальная длина – 32 гена.

Таким образом, существенных различий по количеству и длине квазиоперонов между цепочками выявлено не было.

Заключение

Геном *Methylococcus capsulatus* может оказаться интересным объектом для дальнейшего изучения. Были обнаружены различия между прямой и обратной цепями днк, такие как отсутствие пересечений длины 1 на комплементарной цепи, различное количество генов, кодирующих РНК и псевдогенов. Также может представлять интерес тот факт, что большинство пересечений имело длину в 4 нуклеотида. Объяснение данных фактов может быть целью дальнейших исследований.

Сопроводительные материалы

Сопроводительные результаты можно найти на моем сайте:

<https://kodomofbb.msu.ru/~dr.alexey7/term1/block4.html>

Благодарности

Я хотел бы поблагодарить Марию Точилкину за помощь в подготовке базы данных для отчета, а также Ивана Черных за помощь в принятии решений и моральную поддержку.

Список литературы

1. <https://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.0020303>
2. <https://www.newscientist.com/article/2112298-food-made-from-natural-gas-will-soon-feed-farm-animals-and-us/>
3. https://ru.wikipedia.org/wiki/Methylococcus_capsulatus
4. https://en.wikipedia.org/wiki/Methylococcus_capsulatus