

ОБЗОР ГЕНОМА И ПРОТЕОМА БАКТЕРИИ RHODOSPIRILLUM RUBRUM ATCC 11170

Птицына Е.В.¹

¹Факультет биоинженерии и биоинформатики, Московский государственный университет имени М.В. Ломоносова; Россия, 119234, г. Москва, Ленинские горы, д. 1, стр. 73;

*e-mail: elena-pt@yandex.ru

Резюме

В данной статье представлен краткий обзор анализа генома и протеома бактерии *Rhodospirillum rubrum* ATCC 11170, выполненного с помощью программы Microsoft Office Excel. Изучено распределение числа генов белков и генов РНК по категориям и функциям, длин белков, оценена средняя плотность генов (888,9 gen/Mbh). Получено, что гены с вероятностью 64,4% распределены по прямой и обратным цепям хромосомы не случайно.

Ключевые слова: *Rhodospirillum rubrum*, геном, протеом, Excel.

Введение

Rhodospirillum rubrum – фототрофная пурпурная альфа-протеобактерия, обитающая в озерах, ручьях и стоячих водах. Она хорошо известна своими уникальными и активно изучающимися системами фиксации азота и окисления оксида углерода [1]. На *Rhodospirillum rubrum* была открыта пост-трансляционная модификация нитрогеназы [2].

В ближайшем будущем найдет *Rhodospirillum rubrum* применение и в промышленности: показана возможность использования этой бактерии в биокаталитических топливных элементах [3].

Цель нашей работы – провести краткий обзор генома и протеома штамма ATCC 11170 бактерии *Rhodospirillum rubrum*.

Материалы и методы

Полный геном бактерии был скачан с сайта <https://www.ncbi.nlm.nih.gov>, адрес файла:

ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA/000/013/085/GCA_000013085.1_ASM1308v1/GCA_000013085.1_ASM1308v1_feature_table.txt.gz [4].

Далее файл был импортирован в программу Microsoft Office Excel. В ходе выполнения анализа нами были использованы функции СУММ(), СЧЁТЕСЛИ(), СРЗНАЧ(), ДСТАНДОТКЛП(), МЕДИАНА(), СЧЁТЕСЛИМН(), БИНОМРАСП().

Результаты

1. Общая информация.

Геном бактерии представлен большой двухцепочечной молекулой ДНК (хромосомой) и плазмидой, содержит 4406557 пар нуклеотидов, 3917 генов. Оценка числа генов на 1 млн пар нуклеотидов показывает, что средняя плотность генов составляет 888.9 gen/Mbp.

2. Число генов белков, псевдогенов и генов РНК на прямой и комплементарной цепочках ДНК.

В таблице 1 представлено количество генов белков, псевдогенов и генов РНК в хромосоме и плазмиде одновременно и для хромосомы и для плазмиды по отдельности на прямой (+) и комплементарной (-) цепях. Расчёт показал, что на гены с вероятностью 64,4% распределены по прямой и обратным цепям хромосомы не случайно. Для плазмиды этот показатель значительно ниже: гены с вероятностью 23,9% распределены по прямой и обратным цепям плазмиды не случайно.

Таблица 1. Число генов белков, псевдогенов и генов РНК.

В хромосоме и плазмиде	Гены белков	Псевдогены	Гены РНК	Сумма
"+" цепь	1907	5	37	1949,00
"-" цепь	1934	4	30	1968
Сумма	3841	9	67	
Вероятность такого же или большего отклонения	38,06%			
Вероятность не случайного распределения генов	61,94%			
Хромосома	Гены белков	Псевдогены	Гены РНК	Сумма
"+" цепь	1880	5	37	1922,00
"-" цепь	1911	4	30	1945
Сумма	3791	9	67	
Вероятность получения такого же или большего отклонения	35,56%			
Вероятность не случайного распределения генов	64,44%			
Плазида	Гены белков	Псевдогены	Гены РНК	Сумма
"+" цепь	27	0	0	27,00
"-" цепь	23	0	0	23
Сумма	50	0	0	
Вероятность получения такого же или большего отклонения	76,01%			
Вероятность не случайного распределения генов	23,99%			

3. Число генов белков и генов РНК по категориям.

Белки и РНК нашей бактерии, как и следовало ожидать, очень разнообразны (Таблица 2).

Таблица 2. Гены белков и гены РНК по категориям, а также псевдогены.

Категории	Number	%
Гены белков	3841	98,06%
Рибосомальных	57	1,46%
Транспортных	241	6,15%
Трансфераз	277	7,07%
Гираз	2	0,05%
Киназ	115	2,94%
Синтаз	117	2,99%
Редуктаз	115	2,94%
Дегидрогеназ	110	2,81%
Гипотетических белков	777	19,84%
Других	2030	51,83%
Псевдогены	9	0,23%
Гены РНК	6	7
	5	1,40%
tРНК	5	1,40%
	1	
pРНК	2	0,31%
Других	0	0,00%
Генов	3917,00	

4. Длины белков.

Некоторые рассчитанные параметры для длин белков *Rhodospirillum rubrum* представлены в Таблице 3:

Таблица 3. Некоторые параметры длин белков.

Среднее значение	Стандартное отклонение	Медиана
336,8794585	218,1611854	295

Распределение длин белков бактерии имеет следующий вид (Рисунок 1):

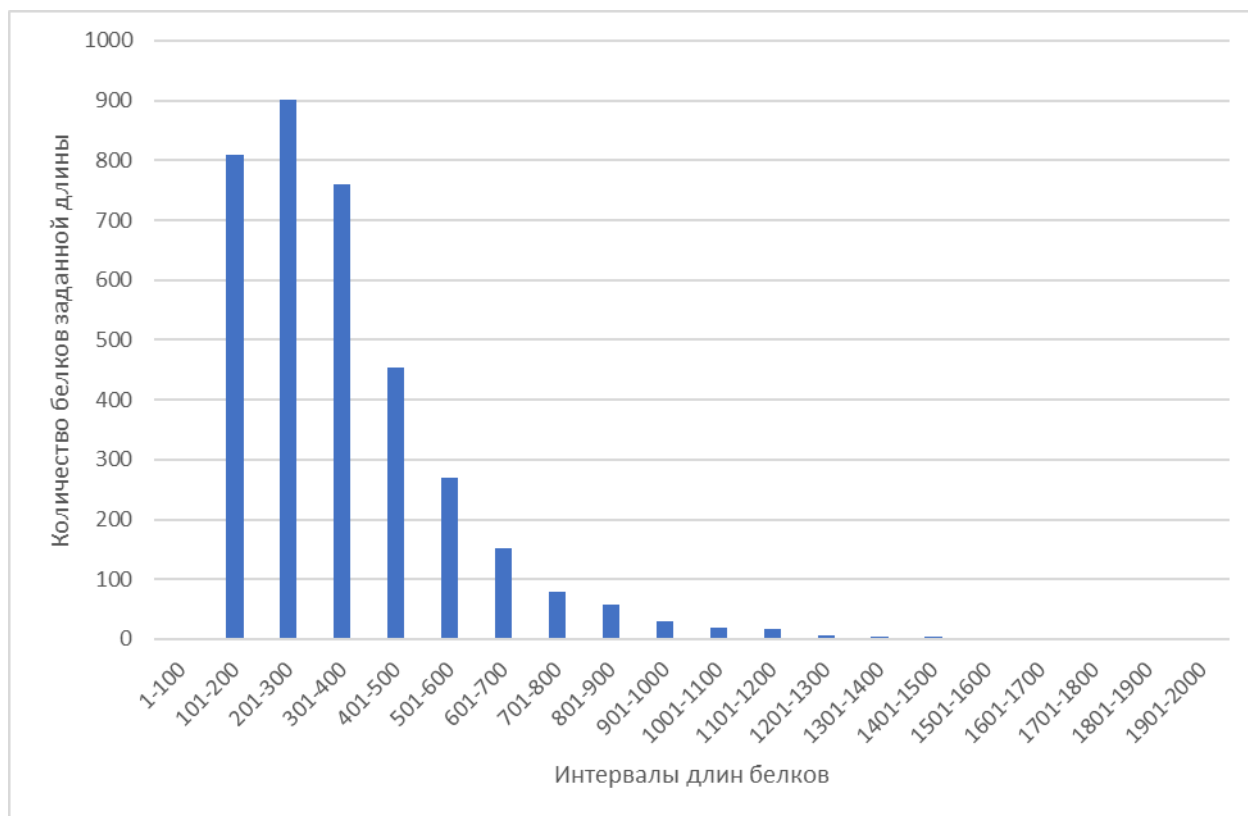


Рис. 1. Распределение длин белков *Rhodospirillum rubrum* ATCC 11170

Результаты и обсуждение

Распределение генов по прямой и обратным цепям в хромосоме, в отличие от плазмиды, не случайно. Количество генов белков намного превышает количество псевдогенов и генов рНК. Количество тРНК чуть больше количества рРНК. Больше всего в геноме достаточно длинных белков, чаще всего встречаются белки с длинами в интервале от 201 до 300 а.к.

Заключение

Rhodospirillum rubrum – важнейший объект для изучения, однако геном этой бактерии изучен недостаточно. В дальнейшем мы планируем продолжить исследования протеома *Rhodospirillum rubrum* и уменьшить долю гипотетических белков.

Сопроводительные материалы

Файл с расчетами Вы можете скачать по адресу: https://kodomo.fbb.msu.ru/~elena-pt/files/pr13_ptitsyna.xlsx

Благодарности

Мы благодарны всем сотрудникам класса биоинформатики факультета биоинженерии и биоинформатики МГУ имени М.В. Ломоносова.

СПИСОК ЛИТЕРАТУРЫ

- 1) Reslewic S. et al. Whole-genome shotgun optical mapping of *Rhodospirillum rubrum* //Applied and environmental microbiology. – 2005. – Т. 71. – №. 9. – С. 5511-5522.
- 2) Pope M. R., Murrell S. A., Ludden P. W. Covalent modification of the iron protein of nitrogenase from *Rhodospirillum rubrum* by adenosine diphosphoribosylation of a specific arginine residue //Proceedings of the National Academy of Sciences. – 1985. – Т. 82. – №. 10. – С. 3173-3177.
- 3) Sasaki S., Karube I. The development of microfabricated biocatalytic fuel cells //Trends in Biotechnology. – 1999. – Т. 17. – №. 2. – С. 50-52.
- 4) Геномы разных штаммов *Rhodospirillum rubrum*, взят штамм ATCC 11170 [Электронный ресурс] URL: <https://www.ncbi.nlm.nih.gov/genome/genomes/1016>, отсюда

ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA/000/013/085/GCA_000013085.1_ASM1308v1/GCA_000013085.1_ASM1308v1_feature_table.txt.gz (дата обращения: 20.12.2018)

Поступила в редакцию
21.12.2018

Сведения об авторах:

Птицына Елена Владимировна – студентка 1 курса Факультета биоинженерии и биоинформатики Московского государственного университета имени М.В. Ломоносова.
e-mail: elena-pt@yandex.ru