

Биоинформатика

Мини-обзор генома бактерии *Proteus mirabilis* BB2000Литвин Анна^{1*}¹МГУ им М.В. Ломоносова, факультет биоинженерии и биоинформатики, первый курс

РЕЗЮМЕ

Целью данной работы было составить общее представление о геноме и протеоме бактерии *Proteus mirabilis* BB2000. Были проанализированы данные о бактериальных белках, организации генетического материала.

1 ВВЕДЕНИЕ

Proteus mirabilis BB2000 — штамм грамотрицательной, факультативно-анаэробной палочковидной бактерии. Систематическое положение организма представлено в Таблице 1.

Таблица 1. Систематическое положение *Proteus mirabilis*

Домен:	Bacteria
Тип:	Proteobacteria
Класс:	Gammaproteobacteria
Порядок:	Enterobacteriales
Семейство:	Enterobacteriaceae
Род:	<i>Proteus</i>
Вид:	<i>Proteus mirabilis</i>

Бактерия *Proteus mirabilis* BB2000 является частью нормальной кишечной флоры здорового человека, но также факультативным патогеном. Жгутиковое движение позволяет этому организму путешествовать через мочевыводящие пути в мочевой пузырь и почки. Пили позволяют прикрепляться к слизистой и эпителиальной оболочкам. В почках бактерия может привести к инфекции или к образованию камней у людей с постоянными катетерами или со сложной системой мочевыводящих путей (например, из-за травмы спинного мозга или анатомической патологии). Известно, что возбудитель использует уреазы, чтобы катализировать образование камней [2]. Также *P. mirabilis* чувствителен к большинству антибиотиков, кроме тетрациклина и нитрофурантоина [3].

Геном данного организма был впервые секвенирован в 2013 году [1].

В данной работе изучено расположение генов внутри генома (распределение по цепям ДНК, пересечение генов, число квазиоперонов), распределение длин белков.

2 МАТЕРИАЛЫ И МЕТОДЫ

Описание генома и протеома бактерии *Proteus mirabilis* BB2000 было скачано с сайта NCBI [1].

Файл формата gbk содержит общую информацию о геноме. Файлы формата ptt и mt содержат информацию о белках и РНК, такую как: длина белков и РНК, координаты генов, кодирующих данные

последовательности в хромосоме, расположение генов на прямой или комплементарной цепи ДНК.

Данные обрабатывались с помощью программного обеспечения Microsoft® Office Excel® 2007. Были использованы формулы для статистической характеристики выборки, математические, логические, текстовые формулы.

Некоторые данные были получены с помощью свободного программного обеспечения Python v2.7.10 ©.

3 РЕЗУЛЬТАТЫ

Геном *Proteus mirabilis* BB2000 состоит из одной кольцевой хромосомы, содержащей 3,846,754 пары оснований. В виде *Proteus mirabilis* обнаружено 9 плазмид, в роде *Proteus* обнаружено 15 плазмид [4].

Оказалось, что хромосома кодирует 3455 белков и 103 РНК (81 транспортную и 22 рибосомальных).

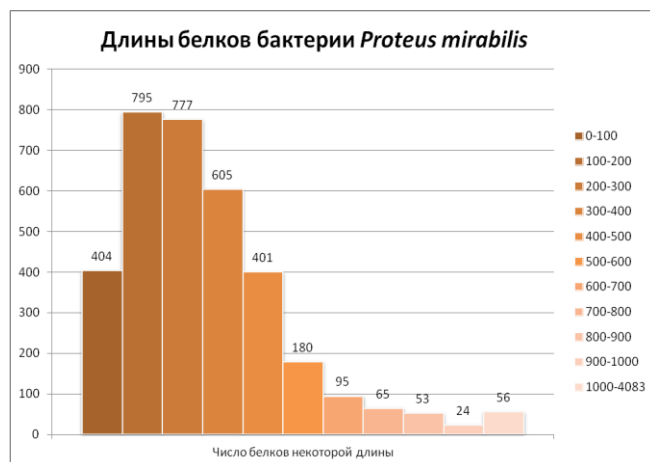
3.1 Длины белков

Таблица 2 отражает основные статистические характеристики диапазона длин белков. На рисунке 1 показана статистика распределения длин белков из протеома бактерии.

Таблица 2. Характеристики длин белков в протеоме бактерии

Величина	Значение в аминокислотах
Максимальная длина	4083
Минимальная длина	30
Средняя длина	313
Медиана длин	265

Рис. 1. Гистограмма длин белков.



Можно сказать, что наиболее распространены белки длиной 100 – 300 аминокислот, что соответствует представлениям о типичном бактериальном протеоме.

3.2 Анализ распределения генов внутри ДНК

В таблице 3 приведено распределение генов белков и РНК на прямой и комплементарной цепях ДНК.

Таблица 3. Распределение генов на прямой и комплементарной цепи ДНК

Гены	Прямая цепь		Комплементарная цепь	
РНК	60	58,25%	43	41,75%
Белки	1672	48,39%	1783	51,61%
Всего	1732	48,68%	1826	51,32%

3.3 Особенности расположения генов

Квазиоперон – группа генов, находящихся на одной цепи, расстояние между которыми не превышает определенного значения – интервала квазиоперона. Квазиоперон может состоять из одного гена.

Было рассмотрено взаиморасположение генов на цепи ДНК, и подсчитано количество квазиоперонов, которые можно на ней выделить.

На диаграмме (рисунок 2) представлена зависимость числа квазиоперонов от задаваемого интервала.

Число квазиоперонов закономерно уменьшается при увеличении интервала, так как увеличивается число генов, которые могут входить в один оперон.

Рис 2. Зависимость количества квазиоперонов от устанавливаемого значения интервала квазиоперона.



В результате анализа расположения генов было обнаружено 260 случаев их пересечения. Также оказалось, что длина гена всегда делится на 3, что соответствует принципу трансляции (три нуклеотида кодируют одну аминокислоту).

4 ОБСУЖДЕНИЕ

Характеристики белков являются стандартными для бактерий. Исходя из результатов пункта 3.2 мы можем сделать предположение о том, что гены распределены между цепями случайно с вероятностью 0,5, так как отклонение в распределении в случае нашего организма составляет 2,64%. Случаи пересечения генов пока не имеют объяснения, так как требуется провести анализ протеома *in vivo*.

5 СОПРОВОДИТЕЛЬНЫЕ МАТЕРИАЛЫ

http://kodomofbb.msu.ru/~litvinanna/term1/litvin_proteusproteins.xlsx - дополнительные материалы к обзору [таблица excel] с расчетами.

6 СПИСОК ЛИТЕРАТУРЫ

- [1] - ftp://ftp.ncbi.nlm.nih.gov/genomes/archive/old_refseq/Bacteria/ геном бактерии
- [2] - https://en.wikipedia.org/wiki/Proteus_mirabilis описание бактерии на Википедии
- [3] - <http://www.ncbi.nlm.nih.gov/genome/?term=Proteus+mirabilis+BB2000> описание бактерии на сайте NCBI
- [4] - http://kodomofbb.msu.ru/~litvinanna/term1/pr13_plasmids.php страница с работой по плазмидам