

Мини-обзор генома бактерии *Acetobacter pasteurianus* NBRC 101655

З.Агаева

Факультет биоинженерии и биоинформатики, Московский государственный университет, Москва, РФ

РЕЗЮМЕ

Целью данной работы является анализ протеома бактерии *A. pasteurianus* NBRC 101655, вызывающей порчу вина. В ходе работы было получены следующие данные: количественная оценка генов белков и РНК в зависимости их функциональных категорий, зависимость длин белков от их доли в протеоме бактерии и количественное распределение генов по цепям ДНК.

Ключевые слова: *геном, штамм, Acetobacter pasteurianus, протеом*

1 ВВЕДЕНИЕ

Штамм *A. pasteurianus* NBRC 101655 был выделен 28 января 2016 года исследователями из университета Ямагучи, Япония. Данная грамотрицательная аэробная бактерия овоидной формы и с содержанием G+C равным 52.7589 %^[1] относится к семейству Acetobacteraceae класса альфа-протобактерий.

Домен	<i>Eubacteria</i>
Тип	<i>Proteobacteria</i>
Класс	<i>Alphaproteobacteria</i>
Порядок	<i>Rhodospirillales</i>
Семейство	<i>Acetobacteraceae</i>
Род	<i>Acetobacter</i>
Вид	<i>Acetobacter pasteurianus</i>

Таблица 1. Систематическое положение бактерии *A. pasteurianus*

Бактерия *A. Pasteurianus* относится к уксуснокислым бактериям, так как в процессе дыхания окисляет этанол до уксусной кислоты, а лактат и ацетат – до диоксида углерода и воды. В природе ее можно найти в богатых сахаром субстратах, таких как фрукты, цветы и овощи; также встречается в портящихся фруктах. Замечено, что в богатой кислородом среде ее рост и метаболические процессы ускоряются. Источникам углерода преимущественно являются этанол, маннит и глюкоза. Оптимальная температура обитания колеблется от 25 до 30 С, оптимальный диапазон pH - от 5.5 до 6.3. Установлено, что данная бактерия может послужить причиной порчи вина.^[2]

2 МАТЕРИАЛЫ И МЕТОДЫ

В данной работе были использованы следующие ресурсы и инструменты обработки данных:

- сайт NCBI^[1] для получения описания генома *A. Pasteurianus* NBRC 101655 (ссылка: GCA_000241585.3_ASM24158v3_feature_table.txt.gz);
- Office 2016 для обработки и анализа полученных данных;
- основной объем информации использовался из следующего файла: [13pr.xlsx](#)

Следует так же отметить, что определение категории гена (см. лист main_table) происходило следующим образом:

- Гены белков: в колонке feature стоит CDS, в колонке class - пустое значение
 1. Гены белков-транспортёров: гены белков такие, что в колонке name встречается слово transporter или transport protein
 2. Гены рибосомальных белков: гены белков такие, что в колонке name встречается слово ribosomal
 3. Гены гипотетических белков: гены белков такие, что в колонке name встречается слово hypothetical
 4. Гены, кодирующие другие белки – это все оставшиеся гены белков, количе-

- ство которых равно разности общего числа генов и числа генов белков вышеперечисленных категорий
- Гены РНК: в колонке feature стоит gene, колонке class – название РНК
 1. Гены тРНК такие, что в колонке class стоит tRNA
 2. Гены рРНК такие, что в колонке class стоит rRNA
 3. Гены тмРНК такие, что в колонке class стоит tmRNA

3 РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

1. Данные о количестве генов белков и РНК

В ходе работы были получены следующие данные о геноме *A. pasteurianus* NBRC 101655: общее количество генов РНК равно 71, белков – 2875. Всего генов – 2947. Размер генома данной бактерии равен 3.10099 млн пар нуклеотидов^[3], на 1 млн п.н. приходится приблизительно 950 генов. Более подробная информация приведена в Таблице 1 и Таблице 2 ниже.

категория	кол-во белков
рибосомальные белки	67
гипотетические белки	873
белки-транспортёры	165
остальные белки	1770

Таблица 1. Количество белков различных категорий

РНК	КОЛ-ВО
тРНК	55
рРНК	15
тмРНК	1

Таблица 2. Количество различных РНК

2. Соотношение количества белков разных длин



Как видно из гистограммы, большую часть протеома бактерии *A. pasteurianus* NBRC 101655 составляют белки длиной от 60 до 540 аминокислот, диапазон длины наиболее распространенных белков – от 120 до 180 аминокислот, это составляет примерно 15 % всего протеома. Небольшие белки (длиной меньше 60) и относительно крупные (длиной более 540 аминокислот) встречаются гораздо реже.

Средняя длина белков равна 322 аминокислотам, среднеквадратичное отклонение от среднего – 226, а медиана – 277. Гликозилтрансфераза – фермент, транспортирующий остатки моносахаридов от углевододонора на молекулу-акцептор, чаще всего спирт^[4] – имеет максимальную длину (2852 аминокислот); возможно это связано с тем, что ферменты имеют третичную либо четвертичную структуру. Самый маленький белок – пирролохинолинхинон, играющий роль окислительно-восстановительного агента в клетках, и длиной 28 аминокислот^[5].

3. Таблица числа генов по цепочкам ДНК

Судя по данным из Таблицы 3, обратная цепь ДНК содержит больше генов как РНК, так и белков по сравнению с прямой, однако логичное объяснение этому факту автор дать затрудняется. Также *A. pasteurianus* NBRC 101655 не имеет псевдогенов, то есть, по всей видимости, геном данной бактерии не претерпевал каких-либо мутаций, что кажется не очень правдоподобным.

цепь ДНК	гены белков	псевдогены	гены РНК
прямая цепь	1387	0	18
обратная цепь	1488	0	53

Таблица 3. Распределение количества ген белков и РНК по цепочкам ДНК

4 СОПРОВОДИТЕЛЬНЫЕ МАТЕРИАЛЫ

Ссылка на файл excel, в котором представлен опорный материал для создания данного мини-обзора: [13pr.xlsx](#)

5 СПИСОК ЛИТЕРАТУРЫ

- [1] https://www.ncbi.nlm.nih.gov/genome/904?genome_assembly_id=266072
- [2] http://wineserver.ucdavis.edu/industry/enology/winemicro/winebacteria/acetobacter_pasteurianus.html
- [3] <https://www.ncbi.nlm.nih.gov/genome/genomes/904>
- [4] <https://ru.wikipedia.org/wiki/%D0%93%D0%BB%D0%B8%D0%BA%D0%BE%D0%B7%D0%B8%D0%BB%D1%82%D1%80%D0%B0%D0%BD%D1%81%D1%84%D0%B5%D1%80%D0%B0%D0%B7%D1%8B>
- [5] https://ru.wikipedia.org/wiki/%D0%92%D0%B8%D1%82%D0%B0%D0%BC%D0%B8%D0%BD_B14