

МАЛЕНЬКАЯ, ДА УДАЛЕНЬКАЯ *PANTOEA ANANATIS* AJ13355

Нестеренко Екатерина
Факультет биоинженерии и биоинформатики
Московский Государственный Университет имени М.В. Ломоносова
Москва, Россия
catherine.nesterenko@kodofo.fbb.msu.ru

*Обзор протеома бактерии **Pantoea Ananatis** AJ13355. Эта бактерия – выгодный объект для проведения различных опытов и исследований в области биотехнологии и биоинженерии. В данной работе проведено аналитическое изучение генома и белков этого организма.*

*Ключевые слова: бактерия, биотехнологии, биоинженерия, *Pantoea Ananatis*, Excel.*

1. Введение.

Pantoea Ananatis AJ13355 – это грамотрицательная бактерия из семейства Enterobacteriaceae с многообещающими характеристиками, новый объект биотехнологии. Бактерии этого семейства имеют желтые пигменты, подвижны и могут образовывать слизистые колонии. Они широко распространены во многих средах. Их можно обнаружить в воде, в почве, это также эндо- и экзо- паразиты животных и растений.

Данный вид выживает и размножается в кислой среде, в присутствии насыщающих концентраций L-Глу и устойчив к высоким концентрациям ряда других органических кислот. В настоящее время выведен штамм на основе AJ13355, который продуцирует L-Глу.

В данной работе были исследованы длины и другие характеристики белков этого организма.

2. Материалы и методы.

Информация о белках данного организма была получена следующим путем.

- На сайте NCBI была найдена база данных Genome
- В ней находился файл с геномом изучаемой бактерии

Данные из файла были преобразованы в таблицу в программе Excel, с помощью которой и далее проводился анализ.

2.1 Работа с белками в таблице

Гены различных белков и РНК были найдены различными путями.

2.1.1. Гены белков в целом

В колонке #features написано CDS и в колонке class пусто, или в колонке #features написано gene, а в колонке class - написано protein_coding. Две стоящие рядом строки в данном файле кодируют один и тот же продукт. Убедиться в этом можно из колонки product_accession. Удалять дубликаты не было особой причины, а от этого могла нарушиться изначальная целостность таблицы.

2.1.2. Гены транспортных белков

Их число – 348. Отбирались они в колонке name с помощью фильтра по слову «transport».

2.1.3. Гены рибосомальных белков

Количество – 77. Схема отбора та же, по слову «ribosomal». Главное следить, чтобы в отбор не попали ribosomal RNA.

2.1.4. Гены гипотетических белков

Их количество – 594. Критерий отбора слово «hypotetical»

2.2. Гены РНК

В геноме встречаются гены двух видов РНК: транспортных (78 штук) и рибосомальных (23 штуки). Других видов нет.

2.3. Таблицы и рисунки.

Таблицы и рисунки оформлялись в программе Excel.

3. Результаты и обсуждения.

3.1. *Различные белки и РНК*

В таблице 1 данные о количестве белков и РНК приведены наглядно. Большое количество гипотетических белков, возможно, говорит о немалом проценте мутаций.

РНК	Количество
Транспортные (tRNA)	78
Матричные (rRNA)	23
Другие	0
Белки	Количество
Рибосомальные	77
Транспортные	348
Гипотетические	594
Другие	3148

Таблица 1. Количество РНК и белков

3.2. *Сортировка белков по длине*

На рис.1 представлена гистограмма длинны белков. Как можно заметить, наибольшее количество белков имеют длину от 171 нуклеотидов до 340. Не самые длинны белки. Такие белки удобны для быстрой сборки, скорее всего они просто устроены и необходимы бактерии в большом количестве. Если посмотреть, что это за белки в прикрепленной таблице из базы данных, то можно заметить, что в большинстве своем это белки «домашнего хозяйства», связанные с работой ДНК, АТФ, РНК и работой метаболизма.

Вторая группы по длине – это группа от 341 до 510. В основном в ней содержатся специфические белки. Много транспортных, а так же ответственных за работу разнообразных ферментов.

Белки длиннее 1530 встречаются редко.

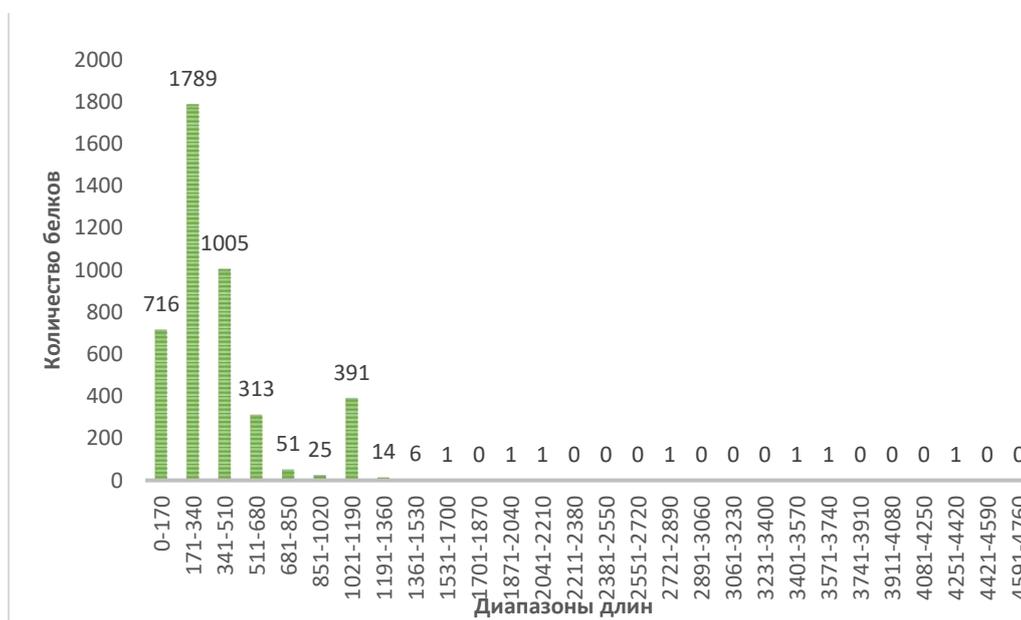


Рис.1 Гистограмма длин белков

3.3. Данные о длинах

Средняя длина	Максимальная	Минимальная	Медиана	Стандартное отклонение
335,94	4385	36	296	222,64

Таблица 2. Данные о длинах

Самый длинный белок – 4385 нуклеотидов. Он называется «Большой повторяющийся белок YeeJ» (англ. «large repetitive protein YeeJ»). Его функция – транспорт адгезивов. Адгезив — вещество, способное соединять материалы путём поверхностного сцепления.

Самые маленькие белки имеют длину 36 нуклеотидов. Это гипотетический белок YhcO (англ. «hypothetical protein YhcO») и 50S рибосомальный белок L34 RpmH (англ. «50S ribosomal protein L34 RpmH»).

3.4. Данные генах

#feature	class	Общее количество	На прямой цепи	На обратной
gene	protein_coding	4067	2053	2014
CDS		4067	2053	2014
gene	rRNA	22	9	13
rRNA		22	9	13
gene	tRNA	78	44	34
tRNA		78	44	34

Таблица 3. Данные о генах.

В таблице 3 приведены данные о генах, псевдогенах и генах, кодирующих РНК и их количество на прямой и комплементарной цепях. Также в нее включены CDS и сами РНК.

4. Заключение

Итак, бактерия *Pantoea Ananatis* AJ13355 – интереснейший объект для изучения, и ей определенно стоит уделить больше внимания. Огромное количество белков и свойства этого организма уже принесли пользу человеку, и принесут еще.

Благодарности

Благодарю преподавателей за ценные и своевременные указания к работе.

Сопроводительные материалы

http://kodomofbb.msu.ru/~catherine.nesterenko/term1/pr13_Nesterenko.xlsx

Список литературы

1. <https://www.ncbi.nlm.nih.gov/pubmed/22159605>
2. <http://www.dissercat.com/content/izuchenie-faktorov-vliyayushchikh-na-biosintez-l-asparagिनovoi-kisloty-v-pantoea-ananatis-i->
3. https://en.wikipedia.org/wiki/Pantoea#cite_note-0-1

4. <https://bmcgenomics.biomedcentral.com/articles/10.1186/1471-2164-15-404>
5. <https://www.ncbi.nlm.nih.gov/protein/BAK11977.1>
6. <https://ru.wikipedia.org/wiki/%D0%90%D0%B4%D0%B3%D0%B5%D0%B7%D0%B8%D0%B2%D1%8B>