

Обзор генома и протеома бактерии *Kribbella flavida* DSM 17836

Вольхин Илья

Факультет Биоинженерии и биоинформатики, Московский
Государственный Университет им. Ломоносова, Ленинские горы 1-73,
Москва, Россия

Ключевые слова: *Kribbella flavida*, длина белков.

Резюме

Отчёт посвящён протеому и геному бактерии *Kribbella flavida*, закономерностям, связанным с размером белков и расположением генов.

Введение

Kribbella flavida – вид бактерий из семейства *Nocardioideae*. Это грамположительный свободноживущий почвенный аэробный организм, относящийся к актиномицетам. Была открыта в 1979 году, но

только в 1999 вынесена в отдельный род [1]. В 2010 году был описан весь геном этой бактерии [2].

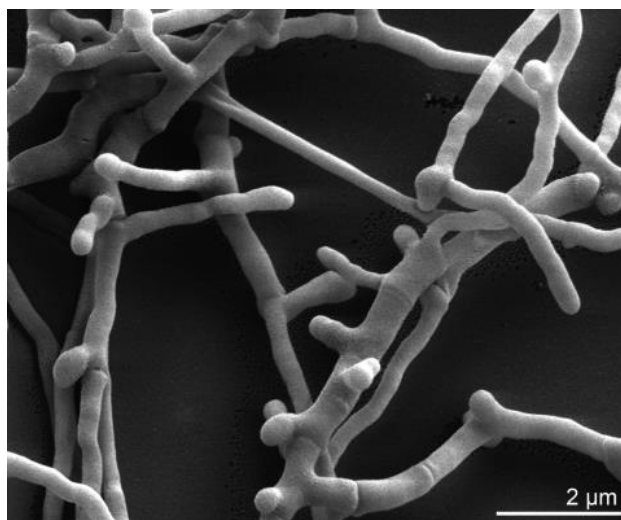


Рис1: *Kribbella flavida* [2]

Материалы и методы

Информация о геноме была получена из базы данных NCBI и банка данных GenBank[3]. Анализ генома был проведён с помощью Microsoft Excel 2007. Для получения информации о количестве генов разных типов были использованы сводные таблицы. Статистические данные получены с использованием функций СРОТКЛ, СРЗНАЧ, МЕДИАНА, МИН, МАКС, СЧЁТЕСЛИМН и возможностей гистограмм.

Результаты

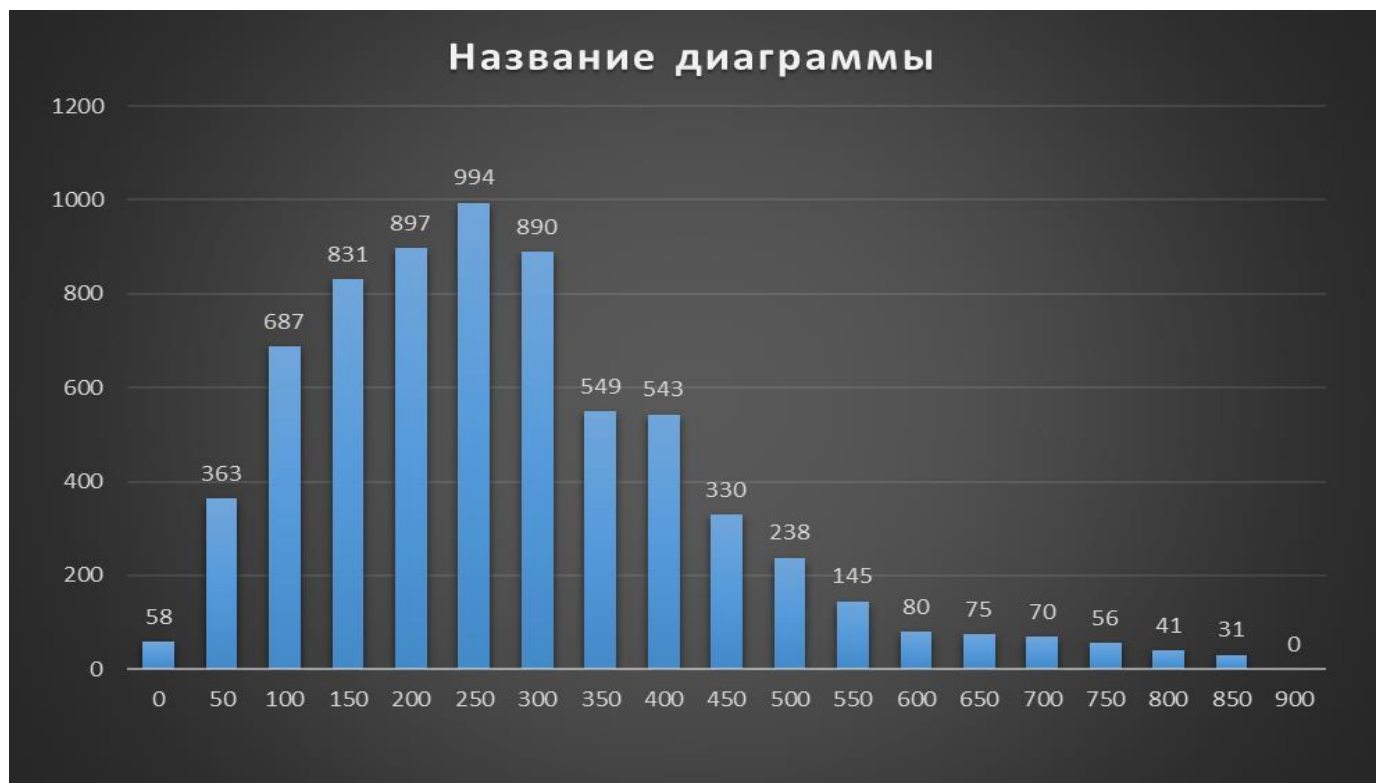
Статистика длин белков

Был проведён анализ длин белков, на основе которого была построена

гистограмма (рис.2).

Как видно из рис.2, количество белков, имеющих длину в данном интервале, сначала резко возрастает, достигая максимума в районе 200-300 аминокислотных остатков, а затем убывает.

Рис.2: гистограмма распределения белков по длине.



Распределение последовательностей

Был проведён анализ распределения по геному различных кодирующих последовательностей, а именно: кодирующих белки, кодирующих РНК и псевдогенов. Результаты суммированы в табл.1 (представлены для двух цепей ДНК: прямой и обратной).

Столбец1	Прямая	Обратная	Всего
Белки	3479	3464	6943
Псевдогены	65	78	143
РНК	26	39	65

Табл.1

Обсуждение

Распределение белков по длинам вполне предсказуемо: за резким подъёмом графика следует менее резкий спад (как в распределении Максвелла), однако данная бактерия демонстрирует интересную особенность: белков длиной 300-350 аминокислотных остатков в полтора раза больше, чем белков длиной 350-400 а.о. Количество белков на интервалах 350-400 а.о. и 400-450 а.о. практически совпадают.

В геноме бактерии резко преобладают гены, кодирующие белки. Меньшее число содержится псевдогенов и совсем небольшое число участков, кодирующих РНК.

Сопроводительные материалы

Файл Excel2.xlsx, в котором получены рис.2 и табл.1, находится по ссылке):
<http://kodomo.fbb.msu.ru/~volkhin2/block4.ht>

Список литературы

1. <https://www.ncbi.nlm.nih.gov/pubmed/10319>
2. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3035274/>
3. <https://www.ncbi.nlm.nih.gov/nuccore/283807292>