

Обзор протеома бактерии *Cyanothese sp. PCC 8801*

Князева Анастасия

МГУ имени М.В.Ломоносова, факультет биоинженерии и биоинформатики

РЕЗЮМЕ

Данная работа посвящена обзору протеома бактерии *Cyanothese sp. PCC 8801*. В ходе работы анализируется распределение длин белков из протеома бактерии и число генов белков и генов РНК на прямой и комплементарной цепочках ДНК. Работа выполнена с помощью программы Microsoft Excel.

1 ВВЕДЕНИЕ

Геном бактерии *Cyanothese sp. PCC 8801* состоит из трех плазмид (pP880101, pP880102, pP880103) и хромосомы. В данной работе изучалось распределение длин белков и распределение генов по прямой и обратной цепям. Была выдвинута гипотеза, о том, что гены распределены по цепочкам случайно с вероятностью 0.5.

2 МАТЕРИАЛЫ И МЕТОДЫ

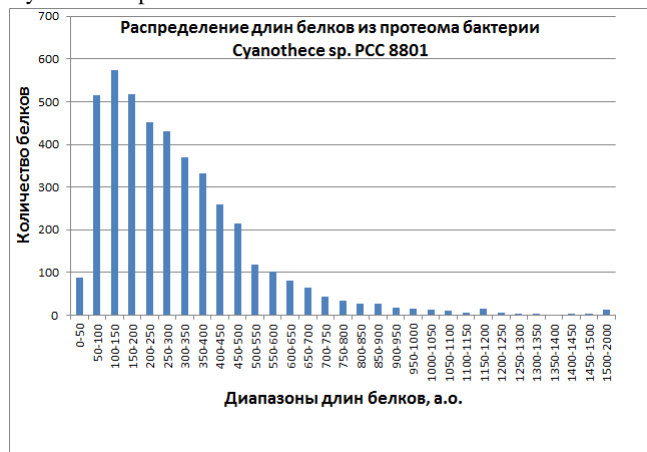
Исходные данные секвенирования полного генома бактерии *Cyanothese sp. PCC 8801* получены из базы данных National Centre of Biotechnology Information.

Для обработки данных использовалась программа Microsoft Excel 2013 (версия: 15.0.4454.1510). При построении гистограммы использовалась команда построения диаграмм и функция СЧЁТЕСЛИМН, которая была нужна для подсчета количества белков определенной длины. Для определения количества генов белков и РНК на прямой и обратной цепях использовались функции СЧЁТЕСЛИМН и СУММ. Для проверки гипотезы о распределении генов на прямой и обратной цепях использовалась функция БИНОМ.РАСП.

3 РЕЗУЛЬТАТЫ

3.1 Распределение длин белков

Рисунок 1. Распределение длин белков из протеома бактерии *Cyanothese sp. PCC 8801*.



Гистограмма длин белков из протеома бактерии *Cyanothese sp. PCC 8801* представлена на Рисунке 1.

Из гистограммы видно, что большинство белков имеет длину от 50 до 500 а.о. Наибольшее количество белков в диапазоне длины от 100 до 150 а.о.

3.2 Распределение генов по цепям

Число генов белков и генов РНК на прямой и комплементарной цепочках ДНК можно увидеть в Таблице 1.

Таблица 1. Распределение генов по цепям.

Цепь	Белки	РНК
Прямая цепь	2224	25
Обратная цепь	2143	28
Всего	4367	53

Из таблицы видно, что гены белков и РНК на прямой и обратной цепях распределены практически равномерно.

4 ОБСУЖДЕНИЕ

- 84% белков имеет длину от 50 до 500 а.о., что является средним значением длины белка прокариотического организма.
- 13% белков имеет длину от 100 до 150 а.о., количество белков большей длины постепенно уменьшается.
- Гены распределены на прямой и обратной цепях практически равномерно. Гипотеза о том, что гены распределены по цепочкам случайно с вероятностью 0.5, подтвердилась как для белков, так и для РНК (по расчетам вероятность для белков составила 0.89, для РНК – 0.39).

СОПРОВОДИТЕЛЬНЫЕ МАТЕРИАЛЫ

Knjasewa_calculation.xlsx – файл с гистограммой длин белков и с таблицей распределения генов по прямой и обратной цепям.

БЛАГОДАРНОСТИ

Выражаю благодарность А.В.Алексеевскому и И.Русинову, которые учат нас работать в программе Excel.

СПИСОК ЛИТЕРАТУРЫ

<ftp://ftp.ncbi.nlm.nih.gov/genomes/Bacteria/>