

Обзор протеома бактерии *Leptospirillum Ferriphilum* ML-04

Ocheredko Elena

School of bioengineering and bioinformatics, MSU

ABSTRACT

В обзоре анализируется протеом бактерии *Leptospirillum ferriphilum*, штамм ML-04, выполняемый в рамках курса биоинформатики; анализ проводился с помощью программы Microsoft office excel 2007. В результате проведённой работы выяснилось следующее:

- В геноме исследуемой бактерии преобладают белки, имеющие длину около 300 аминокислот;
- Количество генов, кодирующих белки, во много раз превышает количество генов, кодирующих РНК;
- Вероятнее всего, распределение генов по цепочкам ДНК не случайно.

1 INTRODUCTION

Любой человек, занимающийся научной деятельностью, должен уметь писать статьи. И студенты факультета биоинженерии и биоинформатики МГУ им. Ломоносова – не исключение. Данный обзор представляет собой одно из домашних заданий курса биоинформатики. Большой научной ценностью он не обладает, но содержит все элементы, которые присутствуют в подобных работах.

2 METHODS

- Исходные данные были получены на сайте <ftp://ftp.ncbi.nlm.nih.gov/genomes/Bacteria/>;
- Microsoft office excel 2007 – программа, с помощью которой была сделана данная работа. Были использованы функции «счѣтесли» (с её помощью было посчитано количество генов на каждой цепочке ДНК) и «биномраспр» (с помощью этой функции была посчитана вероятность случайного распределения генов по цепочкам ДНК), а также команда построения диаграммы.

3 RESULTS

Результаты проделанной работы представлены на Рисунке 1 и в Таблице 1. Расчѣты приведены в файле, на листе «расчѣты». Исходя из полученных данных, можно сделать следующие выводы:

- В геноме бактерии *Leptospirillum ferriphilum* большинство белков (1033 белка, или 41,8% от общего числа белков) имеют длину около 300 аминокислот (Рис.1);
- Количество белков длиной более 1300 аминокислот очень мало (12 белков, или 0,48% от общего числа белков) (Рис.1);
- В геноме бактерии *Leptospirillum ferriphilum* количество генов, кодирующих белки, примерно в 46 раз больше числа генов, кодирующих РНК (Таблица 1);
- Вероятность случайного распределения генов белков по цепочкам составляет 91%; большая часть генов (51%) находится на прямой цепи. Такая вероятность достаточно велика, чтобы предположить, что данное распределение случайно (Файл, лист «расчѣты»).
- Вероятность случайного распределения генов РНК по цепочкам составляет 34%. Большая часть генов (54%) находится на обратной цепи (Файл, лист «расчѣты»).

4 DISCUSSION

- Примерно 0.5% белков не являются полипептидами в строгом смысле этого слова, т.к. их длина менее 30 аминокислотных остатков;
- Большинство белков исследуемой бактерии имеют длину, равную средней длине белка у бактерий (300 – 400 аминокислотных остатков)(<http://geektimes.ru/post/181850/>);
- Полученные величины вероятности распределения генов РНК по цепочкам ДНК позволяют предположить, что такое распределение не случайно (С другой стороны, такая низкая вероятность

случайного распределения может быть обусловлена тем, что выборка имела очень маленький размер).

SUPPLEMENTARY MATERIALS



Рисунок 1. Распределение длин белков бактерии *Leptospirillum ferriphilum*

	прямая цепь	обратная цепь
гены белков(количество)	1269	1202
гены РНК (количество)	25	29

Таблица 1. Распределение генов по цепям ДНК у бактерии *Leptospirillum ferriphilum*

Эти данные, а также исходные данные и расчёты, доступны на <http://kodom0.fbb.msu.ru/~elena0ch/term1/Excel.html> в разделе «Практикум 15».

ACKNOWLEDGEMENTS

Выражаю благодарность Андрею Алексеевскому и Ивану Русинову, которые ведут данный блок практикума курса биоинформатики.

REFERENCES

<http://www.ncbi.nlm.nih.gov/>

<http://geektimes.ru/post/181850/>