

Обзор протеома бактерии *Ruminiclostridium thermocellum* DSM 1313.

Автор: Кузнецова Ксения Сергеевна

МГУ имени М. В. Ломоносова Факультет Биоинженерии и Биоинформатики.

Резюме

В данном обзоре приводятся результаты исследования протеома бактерии *Ruminiclostridium thermocellum* DSM 1313. Основной целью работы является анализ протеома этой бактерии. Все данные были получены с помощью Microsoft Excel 2013.

Введение

Ruminiclostridium thermocellum DSM 1313 также называют *Clostridium thermocellum* DSM 1313.

Классификация:

Надцарство: Procaryota

Царство: Bacteria

Тип: Firmicutes

Класс: Clostridia

Отряд: Clostridiales

Семейство: Ruminococcaceae

Род: *Ruminiclostridium* (*Clostridium*)

Вид: *Ruminiclostridium thermocellum* DSM 1313

Род *Clostridium* насчитывает около 150 различных видов анаэробов, которые широко распространены практически во всех местах обитания, в которых бескислородные органические соединения присутствуют, в том числе в почве, в водных отложениях и в желудочно-кишечных трактах животных и человека. Характерной чертой *clostridium* является форма их клеток, которая напоминает голень или стержень ("*clostridium*" по-гречески означает "веретено"). Эта форма объясняется наличием эндоспор, которые развиваются в условиях, неблагоприятных для вегетативного роста и неизлечимо надувают одиночные клетки. Многие виды эндоспор крайне крепки и могут выжить и в кипящей воде, и под воздействием воздуха. Споры прорастают в условиях, благоприятных для вегетативного роста, таких как анаэробные при наличии органических субстратов. С точки зрения эволюции, *clostridium* считаются наиболее древними бактериями.

Размер генома: 3561619

Количество генов в белке: 2917

Количество РНК генов: 72

Дата секвенирования генома: 16:01:2006

Материалы и методы

Для анализа и обработки данных использовалась программа Microsoft Excel 2013. Для получения данных использовались различные встроенные функции Excel (такие как гистограмма, сводная таблица и др.), а также формулы: логические, текстовые, статистические, ссылки и массивы и др. Исходные данные были взяты с сервера NCBI (ftp://ftp.ncbi.nlm.nih.gov/genomes/archive/old_refseq/Bacteria/).

Результаты

Гистограмма длин белков из протеома бактерии *Ruminiclostridium thermocellum* DSM 1313.

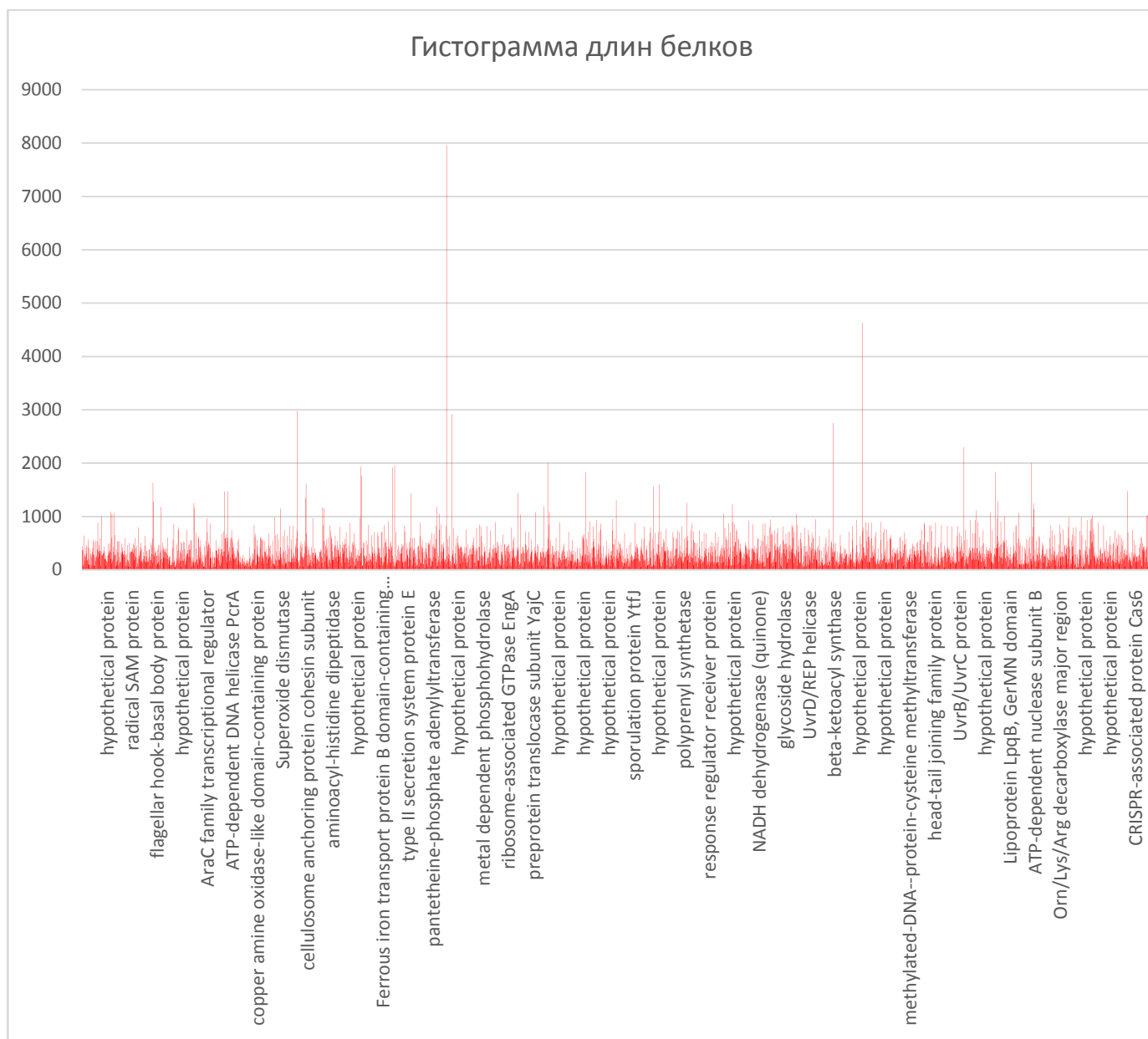


Рисунок 1. Гистограмма длин белков из протеома бактерии *Ruminiclostridium thermocellum* DSM 1313.

В ходе работы была составлена гистограмма длин белков бактерии (Рис. 1). Из гистограммы видно, что наибольшая длина белка 7955, а в основном длина белков не превышает 1000.

Гистограмма распределения количества белков от длины белков

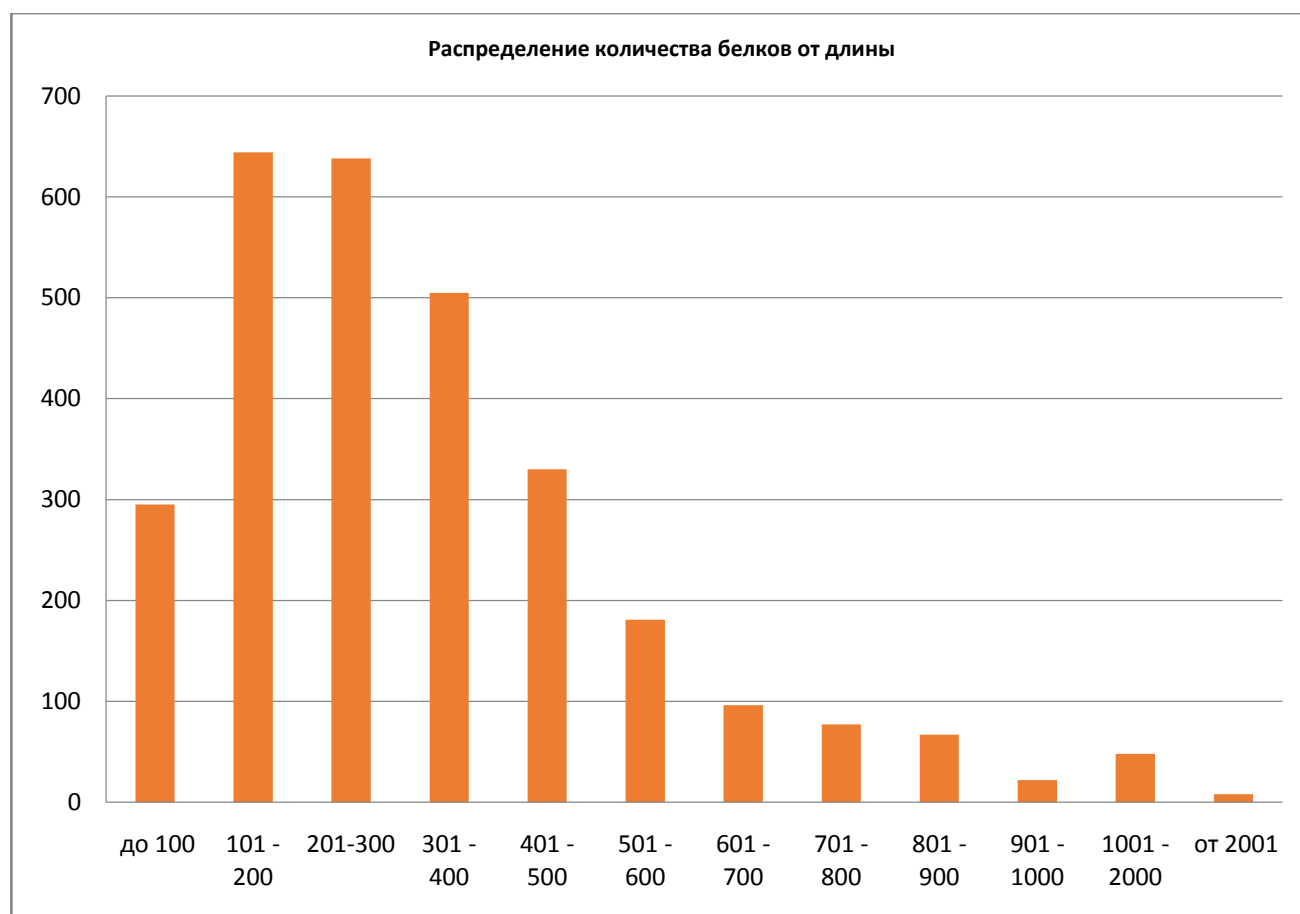


Рисунок 2. Гистограмма распределения количества белков от длины бактерии *Ruminiclostridium thermocellum* DSM 1313.

Из гистограммы (Рис. 2) видно, что чаще всего встречаются белки длиной в диапазоне от 100 до 300. Белки длиной до 100 встречаются реже, но тем не менее составляют более 10% от общего количества. В дальнейшем при увеличении длины белка происходит уменьшение количества белков.

Таблица распределения генов по цепям

	Число генов на прямой цепи	Число генов на комплементарной цепи
ДНК	1435	1476
РНК	39	33

Рисунок 3. Таблица распределения генов по цепям бактерии *Ruminiclostridium thermocellum* DSM 1313.

Проанализировав таблицу (Рис. 3), можем сделать вывод о том, что число генов на прямой цепи и на комплементарной цепи почти равно, и для ДНК, и для РНК. Это подтверждает гипотезу о том, что гены распределены между двумя цепями случайно.

Данные о геномах плазмид рода *Ruminiclostridium* и вида *Thermocellum*

Название плазмиды	Идентификатор записи с последовательностью в БД Refseq	Идентификатор записи с последовательностью в БД INSDC	Род организма-хозяина	Название вида	Название организма	Размер (Kb)	Белок	rRNA	tRNA	Other RNA	Gene	Pseudogene
pEBM130	NC_021665.1	KC788287	Ruminiclostridium	thermocellum	Ruminiclostridium thermocellum BL21	9.957	4	-	-	-	4	-
pEBM113	NC_021653.1	KC920664	Ruminiclostridium	thermocellum	Ruminiclostridium thermocellum BL21	9.892	5	-	-	-	5	-
pEBM107	NC_021652.1	KC920663	Ruminiclostridium	thermocellum	Ruminiclostridium thermocellum BL21	9.896	5	-	-	-	5	-

Рисунок 4. Геномы плазмид рода *Ruminiclostridium* и вида *Thermocellum*.

Из таблицы (Рисунок 4) видно, что в роде *Ruminiclostridium* и виде *Thermocellum* всего 3 плазмиды. Причем плазмиды pEBM113 и pEBM107 различаются только размером на 4 Kb.

Обсуждение

В ходе работы был проанализирован протеом бактерии *Ruminiclostridium thermocellum* DSM 1313, зависимость длин белков от их количества, распределение генов по цепям для ДНК и РНК этой бактерии. А также было выяснено, что данная бактерия не имеет плазмид или хромосом. Но в роде *Ruminiclostridium* есть три плазмиды.

Сопроводительные материалы

Ссылка на файл с расчетами <https://cloud.mail.ru/public/HpDH/UQx8q6rMN>

Благодарности

Выражаю благодарность преподавателям биоинформатики, которые помогли мне освоить Excel, позволивший выполнить данную работу.

Список литературы

1. Информация о бактерии с сайта NCBI [http://www.ncbi.nlm.nih.gov/genome/?term=txid1515\[orgn\]](http://www.ncbi.nlm.nih.gov/genome/?term=txid1515[orgn])
2. Информация о плазмидах с сайта anonymous@ftp.ncbi.nlm.nih.gov:/genomes/GENOME_REPORTS/plasmids.txt
3. Информация о данных с моего сайта http://kodomo.fbb.msu.ru/~ksenia1697/term1/pr13/EXCEL_2.html