

ОБЗОР ГЕНОМА И ПРОТЕОМА БАКТЕРИИ *HALOARCUA MARISMORTUI* ATCC 43049

Жукова Надежда

Факультет биоинженерии и биоинформатики МГУ им. Ломоносова

20 декабря 2018 г.

РЕЗЮМЕ

Haloarcua marismortui - красный галофильный археон, обнаруженный в Мертвом море и процветающий в этой экстремальной среде благодаря нескольким адаптациям в структуре белка, метаболических стратегиях и физиологических реакциях, для изучения которых необходимо секвенирование его генома. [2][3] Лучшее понимание сетей регуляции генов, которые влияют на белок-белковые и белково-ДНК-взаимодействия, может также обеспечить основу для биотехнологических применений в будущем. В данной работе был проанализирован геном и протеом: получены количественные оценки для генов и белков с точки зрения их классификации по различным категориям, построена гистограмма распределения длин белков, рассмотрен вклад прямой и комплементарной цепей ДНК в кодирующие участки генома.

Ключевые слова: *Haloarcua marismortui*, обзор, Excel, геном

ВВЕДЕНИЕ

Domain:	Archaea
Phylum:	Euryarchaeota
Class:	Halobacteria
Order:	Halobacteriales
Family:	Halobacteriaceae
Genus:	Haloarcua
Species:	<i>Haloarcua marismortui</i>

Таблица 1: Систематическое положение *Haloarcua marismortui*

Этот организм обнаружен в Мертвом море, с высоким содержанием физиологического раствора, низкой растворимостью в кислороде и высокой интенсивностью света. Другие галофильные архейные организмы живут в солнечных соленых водах, таких как Большое Соленое Озеро, которые являются экстремальными средами с молярностью около 4,5 моль соли. Не проведено обширных исследований взаимодействия этого организма с другими организмами, которые живут в его окружающей среде. Не проводится обширных исследований его вклада в непосредственную окружающую среду или его влияния на окружающую среду. Однако известно, что окружающая среда воздействует на *H. marismortui*, увеличивая концентрацию соли в его клетках. Фактически, именно среда с высоким содержанием соли делает этот организм уникальным. [1]

Нет никаких известных болезней, которые вызваны этим организмом. Он не был признан патогенным.

H. marismortui имеет геном размером 4275 КБ, состоящий из девяти репликонов и 4242 кодирующих белок генов. Большая хромосома I представляет собой репликон размером 3132 т.п.н. Остальные восемь репликонов имеют меньший размер - от 33 до 410 кб. Эта организация, состоящая из двух частей генома, обычно встречается у всех членов этой группы организмов. Значение этого типа организации неизвестно.

МАТЕРИАЛЫ И МЕТОДЫ

Использованы данные с сайта NCBI:

Haloarcua marismortui ACTT 43049, complete genome, GCA_000011085.1_ASM1108v1_feature_table [4]. Данные были импортированы и форматированы в плоскую таблицу в Microsoft Office Excel, где также была выполнена их обработка: построение диаграммы, таблиц, статистический анализ.

Для подсчета числа генов белков по всем категориям использовались фильтры. По полученным данным было найдено количество различных генов. Для подсчета количества длин белков на заданных промежутках (с шагом в 100 а.о.) сначала была составлена таблица всех длин, а после была проанализирована частота встречаемости длин белков на данных промежутках. В результате была составлена гистограмма длин белков (см. Рисунок 1). Для подсчета числа генов белков РНК на прямой и обратной цепях ДНК были использованы фильтры «+», если ген белка находился на прямой цепи, и «-», если на комплементарной цепи. Результат можно увидеть в Таблице 3.

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Штамм *Haloarcua marismortui* ATCC 43049 содержит 4240 генов белка и 61 ген РНК. Количество генов белков, генов РНК, распределенное по цепям ДНК (прямой и комплементарной) представлено в Таблице 3.

	Прямая цепь	Обратная цепь	Всего
Белки	2103	2137	4240
РНК	32	29	61

Таблица 3: Распределение генов по прямой и обратной цепям ДНК

На Рисунке 1 представлена гистограмма распределения длин белков по различным диапазонам. Средняя длина белка синтезируемого исследуемой бактерией составляет около 284 аминокислоты. Результаты статистического анализа данных приведены в Таблице 2.

Минимальная длина гена:	93
Максимальная длина гена:	6921
Средняя длина гена:	854
Стандартное отклонение:	445,7608
Медиана:	727,5

Таблица 2: Статистические данные по длинам белков

Результаты распределения можно увидеть на Рисунке 1. Как видно из гистограммы, чаще всего у *Haloarcua marismortui* встречаются белки длиной 100-200 аминокислот. Дальше 500 аминокислот следует резкий спад количества белков. Можно обобщить, что в протеоме *Haloarcua marismortui*

превалируют белки со сравнительно низким числом аминокислот. Это объясняется примитивностью бактерий и отсутствием необходимости в сложных длинных белках. Характерное и интересное исключение из закономерности - длинный белок, состоящий из 6921 аминокислотных остатков.



Рисунок 1: Гистограмма длин белков

СПИСОК ЛИТЕРАТУРЫ

[1] NCBI: *Haloarcula marismortui*, Accessed June 4, 2007

http://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?mode=Info&id=2238&lvl=3&p=mapview&p=has_linkout&p=blast_url&p=genome_blast&lin=f&keep=1&srch-mode=1&unlock

[2] Baliga N, Bonneau R, Facciotti M, Pan M, Glusman G, Deutsch E, Shannon P, Chiu Y, Weng R, Gan R, Hung P, Date S, Marcotte E, Hood L, and Ng W. "Genome sequence of *Haloarcula marismortui*: A halophilic archaeon from the Dead Sea," *Genome Res.*, Nov 2004; 14: 2221 - 2234.

[3] Kennedy SP, Ng WV, Salzberg SL, Hood L, and DasSarma S. "Understanding the adaptation of *Halobacterium* species NRC-1 to its extreme environment through computational analysis of its genome sequence," *Genome Res.* 2001; 11: 1641-1650.

[4] Feature_table. NCBI. *Haloarcula marismortui* ATTC 43049 - ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA/000/011/085/GCA_000011085.1_ASM1108v1
GCA_000011085.1_ASM1108v1_feature_table.txt.gz

[5] <http://www.kodomo.fbb.msu.ru/~zhukonad99/term1/pr12-2.xlsx>

[6] <https://www.ncbi.nlm.nih.gov/pubmed/10459393?dopt=Abstract>

СОПРОВОДИТЕЛЬНЫЕ МАТЕРИАЛЫ

Материалы Excel, использованные в исследовании. [5]

БЛАГОДАРНОСТИ

Выражаю благодарность моим преподавателям Алексеевскому А.В., Спирину С.А. и Дмитрию Пензару за возможность овладеть большим количеством новых знаний и умений. Также хотелось бы поблагодарить моих однокурсниц Чумакову Милену и Шиляеву Марию и соседку по комнате Катину Екатерину за поддержку в трудное время. Отдельная благодарность моему психотерапевту Мухитовой Ю. и психиатру Задорожной М.С. за серьезную помощь в поддержании жизнедеятельности. Разумеется, нельзя не поблагодарить представителей рода *Blattella* (*Blattella germanica*) за несоизмеримый вклад в поддержание работоспособности автора в ночь с 20.12.2018 на 21.12.2018 путем обеспечения своим внезапным появлением в комнате №326 Филиала Дома Студентов выброса достаточного количества ад-рениалина в кровь автора.

КОНФЛИКТ ИНТЕРЕСОВ

Наиболее остро конфликтовали между собой интересы деда по дисциплине «Практическая биоинформатика», коллоквиума по дисциплине «Общая и неорганическая химия» и зачета по предмету «Безопасность жизнедеятельности», которые внезапно пришлось на один и тот же день – 21.12.2018. Однако конфликт был успешно разрешен в пользу депривации сна, которая по некоторым данным, к сожалению, еще не подкреплена моим личным опытом, имеет полезные свойства, в частности помогает в борьбе с депрессией. [6] И все же в ходе эксперимента установлено, что отсутствие сна прекрасно помогает справиться с проблемой выбора между горящими дедлайнами.