

Обзор протеома возбудителя туляремии

Касьянова Мария

РЕЗЮМЕ:

В работе изучены различные свойства белков и генов бактерии вида *Francisella tularensis*, а также сделаны выводы о закономерности распределения генов в цепях ДНК.

1 ВВЕДЕНИЕ

Francisella tularensis – вид грамотрицательных бактерий, возбудящих туляремию (при этом заболевании могут появляться язвы, лихорадка, опухшие железы, пневмония и другие симптомы, зависящие от способа заражения) (1, 4). Является облигатным аэробом, культивируется, растет только на питательных средах, содержащих витамины, яичный желток, экстракты органов и тканей животных (2). Геном представлен кольцевой молекулой ДНК, содержащей 1892616 пар нуклеотидов (2). Носителями палочки туляремии являются зайцы, кролики, водяные крысы, полевки (1). Человек может заразиться при непосредственном контакте с этими животными, через пищу и воду, при укусах насекомых, аспирационным путем (1, 3). *Francisella tularensis* разрабатывалась и использовалась в качестве биологического оружия во время второй мировой войны несколькими странами (в том числе США) (3, 4). Сейчас изучается при разработке лекарств, а потому представляет интерес для рассмотрения (3).

2 МАТЕРИАЛЫ И МЕТОДЫ

Данные о геноме бактерии были взяты из [базы данных Genom на сайте NCBI](#) и обработаны при помощи программы Excel. Применялись импорт данных, форматирование, специальная вставка, различные формулы (ВПР, БИНОМРАСП, СЧЁТЕСЛИМН, СЧЁТ, ЕСЛИОШИБКА, ЕСЛИ, НАЙТИ и другие) и адресация. Для всех генов тРНК и рРНК были искусственно введены идентификаторы, позволившие осуществить связь таблиц и получить удобную плоскую таблицу. А в аналогичную таблицу генов класса CDS были добавлены столбцы с пометкой о недостаточной изученности функций соответствующих белков (столбец "hypothetical?") и их участия в создании рибосом (столбец "ribosomal?") (информация взята из названия кодируемых белков). Также несколько раз подсчитывалась вероятность случайного возникновения существующего распределения генов по 2 цепям ДНК при помощи функции БИНОМРАСП. По аргументам $k, n, 0.5$, ИСТИНА эта функция выдает с какой вероятностью хотя бы k испытаний из n будут успешными, если вероятность успеха каждого испытания независима и равна 50% (5). Будем считать что "испытание" № i успешно, если ген № i попал на первую цепь ДНК (при случайном распределении вероятность такого события равна 0.5), тогда значение БИНОМРАСП($k; n; 0.5; ИСТИНА$) показывает с какой вероятностью не менее k генов случайно оказались на первой цепи ДНК и не более $n-k$ на второй, однако, 2 цепи ДНК одинаковы, а значит результат необходимо умножить на 2.

3 РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЯ

3.1 Белки

В протеоме *Francisella tularensis* наименьшая длина белка составляет 32 аминокислоты, а наибольшая - 1422. На рисунке 1 представлено распределение белков по длинам и из него видно, что больше всего белков имеют длину в промежутке от 238 до 290 аминокислот.

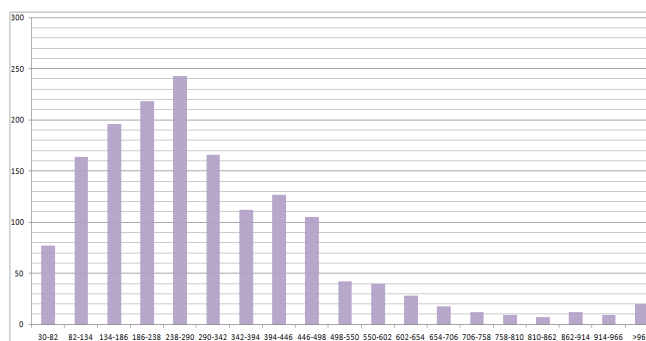


Рис. 1. Гистограмма длин белков.

Также можно заметить, что существование не всех белков выяснено достоверно, для некоторых существуют лишь гипотезы. Такие белки составляют 30,77% от общего количества (лист "гипотетические белки" сопроводительных материалов).

3.2 Гены

В геноме палочки туляремии существуют различные группы генов, более подробная информация об их количестве представлена в таблице 1. При заданном количестве генов рРНК расположение только на одной цепи ДНК наименее вероятно и, возможно, связано с какими-либо специфичными особенностями бактерии, но так как общее количество этих генов слишком мало, любое распределение с достаточно большой вероятностью могло возникнуть в ходе эволюции случайно. Внимание также привлекают кодирующие белок гены, большая часть которых лежит на прямой цепи ДНК. Можно предположить, что такое расположение по каким-то причинам выгодно для *Francisella tularensis*.

	гены, кодирующие белок	псевдогены	гены тРНК	гены рРНК
Количество на прямой цепи	883	103	20	0
количество на обратной цепи	721	97	18	10
вероятность такого распределения	0,00006	0,72377	0,87141	0,00195

Таблица 1. Группы генов.

Еще больший интерес вызывает распределение генов, кодирующих рибосомальные белки (таблица 2). Вероятность случайного возникновения такого расположения достаточно мала, чтобы продолжить поиски дополнительных факторов, влияющих на данный вид генов.

общее количество генов	количество генов на прямой цепи	количество генов на обратной цепи	вероятность такого распределения
62	52	10	0,000000057

Таблица 2. Гены рибосомальных белков.

4 ЗАКЛЮЧЕНИЕ

Изучив приведенные выше данные о генах бактерии *Francisella tularensis*, можно предположить, что их распределение по цепям ДНК не случайно. Определение факторов, влияющих на этот признак, может быть важно для понимания механизмов, лежащих в основе жизнедеятельности этого организма, а также для разработки лекарств против туляремии или исследований в других областях применения этой бактерии. Помимо этого в работе приведены некоторые характеристики белков.

СОПРОВОДИТЕЛЬНЫЕ МАТЕРИАЛЫ

[Таблицы с расчетами.](#)

СПИСОК ЛИТЕРАТУРЫ

1. Википедия, русская: [Туляремия](#). (14.12.2019 21:30)
2. Википедия, русская: [Francisella tularensis](#). (14.12.2019 21:30)
3. [Francisella tularensis: Stopping a biological weapon](#). (14.12.2019 21:30)
4. Microbewiki: [Francisella tularensis](#). (14.12.2019 21:30)
5. Служба поддержки office: [БИНОМРАСП](#). (14.12.2019 21:30)