

Обзор протеома бактерии *Xylella fastidiosa* 9a5c

Щигал Ольга Евгеньевна

Московский государственный университет им. М. В. Ломоносова,
факультет Биоинженерии и биоинформатики, 2018 г.

<https://kodomo.fbb.msu.ru/~olga.shigal/index.html>

РЕЗЮМЕ

Данная работа была проведена с целью изучения протеома патогенной бактерии *Xylella fastidiosa* 9a5c. Были рассмотрены основные особенности генома бактерии: установлено распределение различных генов по прямой и комплементарной цепочкам, число белок- и РНК- кодирующих генов и псевдогенов, распределение длины белков.

Ключевые слова: протеом, бактерия, гены

ВВЕДЕНИЕ

Таблица 1. Систематическое положение бактерии

| | |
|--------|----------------------|
| Domain | Bacteria |
| Phylum | Proteobacteria |
| Class | Gamma proteobacteria |
| Order | Xanthomonadales |
| Family | Xanthomonadaceae |
| Genus | Xylella |

| | |
|---------|-----------------------------------|
| Species | <i>Xylella fastidiosa</i> |
| Stain | <i>Xylella fastidiosa</i> 9a5c |

Xylella fastidiosa – аэробная грамотрицательная бактерия монофилетического рода *Xylella*. Она является возбудителем цитрусового пестрого хлороза и передается исключительно насекомыми, питающимися ксилемной жидкостью. Этот организм был впервые идентифицирован в 1993 году как возбудитель болезни, поражающей разновидность сладких апельсинов, впервые замеченной в Бразилии в 1987 году.

Систематическое положение *Xylella fastidiosa* 9a5c приведено в таблице 1.

Цитрусовый пестрый хлороз наносит огромный ущерб коммерческим садам, что приводит к снижению урожая во

всем мире. В связи с этим *Xylella fastidiosa* 9a5c представляет большой интерес для исследования.

Секвенирование генома *X. fastidiosa* было осуществлено объединением более 30 исследовательских лабораторий в штате Сан-Паулу, Бразилия, и финансировалось Научным фондом этого штата (FAPESP).

Длина генома составляет 2,732 038 пар нуклеотидов.

В данной работе будут рассмотрены некоторые общие черты протеома *Xylella fastidiosa* 9a5c.

МАТЕРИАЛЫ И МЕТОДЫ

Основная часть данной работы велась с использованием программы Microsoft Excel. Для расчета использовались такие функции, как COUNTIF, VLOOKUP, СЧЕТЕСЛИМН, МАКС, МИН, ОКРУГЛ, СТАНДОТКЛ, СРЗНАЧ, а также функции построения диаграмм и составления сводных таблиц. Исходный файл был взят из базы данных NCBI:

ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCF/000/006/725/GCF_000006725.1_ASM672v1/GCF_000006725.1_ASM672v1_feature_table.txt.gz

ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCF/000/006/725/GCF_000006725.1_ASM672v1/GCF_000006725.1_ASM672v1_feature_table.txt.gz

РЕЗУЛЬТАТЫ

Таблица 2. Распределение генов по цепям

| | Гены белков | Псевдогены | Гены РНК |
|----------|-------------|------------|----------|
| Прямая | 1264 | 120 | 35 |
| Обратная | 1115 | 100 | 25 |

Таблица 3. Статистические данные белков

| | |
|--------------------|--------|
| Max | 3455 |
| Min | 38 |
| Average | 299,23 |
| Standart deviation | 249,7 |
| Median | 270 |

В таблице 2 показано распределение генов на прямой и обратной (комплементарной) цепи ДНК бактерии.

В таблице 3 приведены основные статистические данные о длинах белков *X. fastidiosa* 9a5c: максимальная и минимальная длины, среднее значение, стандартное отклонение и медиана.

Рис. 1. Гистограмма длин белков



На гистограмме длин белков из протеома *X. fastidiosa* 9a5c (рис.1) наглядно показано распределение белков разной длины.

Таблица 4. Распределение генов РНК по цепочкам ДНК

| | Прямая | Обратная |
|-------------|--------|----------|
| ncRNA | 1 | 0 |
| RNase_P_RNA | 2 | 0 |
| rRNA | 6 | 0 |
| SRP_RNA | 2 | 0 |
| tmRNA | 1 | 0 |
| tRNA | 24 | 25 |

В таблице 4 отображено распределение генов, кодирующих различные РНК, по прямой и обратной (комплементарной) цепям ДНК протеома бактерии.

ОБСУЖДЕНИЕ И ЗАКЛЮЧЕНИЕ

На основе данных из таблицы 2 можно сделать предположение о равномерном распределении генов на прямой и обратной цепи. Несмотря на то, что на

прямой цепи расположено больше белок-кодирующих, РНК-кодирующих генов и псевдогенов, это различие статистически невелико.

Опираясь на данные гистограммы (рис.1), можно говорить о том, что наибольшее число белков находится в диапазоне длин от 100 до 299, вдвое меньше – в диапазоне 300-499, а на остальные диапазоны приходится гораздо меньшее количество белков.

Минимальный показатель у диапазона 2000-3000 (всего 2 белка). Этот вывод можно подтвердить значением средней длины из таблицы 3 (Average = 299,23).

Согласно данным, представленным в таблице 4, в геноме *Xylella fastidiosa* 9a5c самым распространенным типом РНК является транспортная РНК (tRNA). Интересно, что остальных типов РНК в геноме более чем в 4 раза меньше, и кодирующие их гены расположены только на прямой цепи. Однако такое распределение РНК-генов на цепочках ДНК можно считать случайным совпадением.

СОПРОВОДИТЕЛЬНЫЕ МАТЕРИАЛЫ

<https://kodomo.fbb.msu.ru/~olga.shigal/pr13.xlsx>

БЛАГОДАРНОСТИ

Выражаю свою благодарность в подготовке данного отчета Салимгарееву Руслану и Боровиковой Ирине за консультации, ценные рекомендации и терпение, а также моей кошке за поддержание эмоционального спокойствия в процессе выполнения данной работы.

СПИСОК ЛИТЕРАТУРЫ

[1] <https://www.ncbi.nlm.nih.gov> – исходные таблицы и описание бактерии *X. fastidiosa* 9a5c

[2] <https://www.cabi.org> – систематика бактерии

[3] *Redak, Richard A. ; Перселл, Александр Х.; Лопес, Жоао РС; Вlua, Matthew J. ; Мизелл III, Рассел Ф. ; Андерсен, Питер С. (2003-12-03). «Биология ксилемных питательных насекомых-переносчиков *xylella fastidiosa* и их связь с эпидемиологией заболеваний». Ежегодный обзор энтомологии*

[4] Симпсон, АЖГ; Рейнах, ФК; Арруда, П; Абреу, ФА; и другие. (Июль 2000 г.) «Последовательность генома растительного возбудителя *Xylella fastidiosa* ». Nature.