

# Обзор протеома бактерии *Gemmatimonas phototrophica* AP64

Кондратенко Наталья Дмитриевна<sup>1</sup>

<sup>1</sup>Факультет биоинженерии и биоинформатики МГУ им. М. В. Ломоносова

## РЕЗЮМЕ:

В данной работе был исследован протеом бактерии *Gemmatimonas phototrophica* AP64. Была изучена особенность распределения белков по длине, особенности распределения кодирующих последовательностей по продукту (РНК, белок или псевдоген) и положению на цепи.

## КЛЮЧЕВЫЕ СЛОВА:

Протеом, белок, *Gemmatimonas phototrophica* AP64, геном

## 1 ВВЕДЕНИЕ

### Систематическое положение *Gemmatimonas phototrophica*

Домен	<i>Prokaryota</i>
Царство	<i>Bacteria</i>
Тип	<i>Gemmatimonadetes</i>
Класс	<i>Gemmatimonadetes</i>
Порядок	<i>Gemmatimonadales</i>
Семейство	<i>Gemmatimonadaceae</i>
Род	<i>Gemmatimonas</i>
Вид	<i>G. phototrophica</i>

Геном бактерии *Gemmatimonas phototrophica*, штамм AP64, был секвенирован в 2015 году. Она была обнаружена в изоляции в пресноводном Лебедином озере (в оригинале Swan Lake) в западной пустыне Гоби. Основываясь на 96.1% идентичности гена 16S рРНК к бактерии *Gemmatimonas aurantiaca* T-27T данный вид был отнесен к типу *Gemmatimonadetes*. *Gemmatimonas phototrophica* имеет функциональные реакционные центры фотосинтеза II типа, но не ассимилирует органический углерод, из чего можно сделать вывод, что бактерия ведет фотогетеротрофный образ жизни [1]. Открытие данного краснопигментированного штамма, продуцирующего бактериохлорофилл а имело большое значение, так как ранее было зафиксировано только 6 из 30 описанных типов бактерий, продуцирующих бактериохлорофилл: *Cyanobacteria*, *Proteobacteria*, *Chlorobi*, *Chloroflexi*, *Firmicutes* и *Acidobacteria*. Секвенирование генома дало основания для того, чтобы считать *Gemmatimonadetes* седьмым бактериохлорофиллоносным типом бактерий и

может дать новое представление об эволюции бактериального фотосинтеза [2]. В данной работе проведен анализ распределения кодирующих последовательностей данной бактерии, что может быть полезно для дальнейших исследований генома данной бактерии.

## 2 МЕТОДЫ

Были получены данные о геноме данной бактерии на сайте NCBI: *Gemmatimonas phototrophica* AP64, complete genome, GCA\_000695095.2\_ASM69509v2\_feature\_table.txt.gz.

На основе результатов таблицы была построена плоская таблица.

Для каждого сочетания в колонках «#features» и «class» было посчитано число строк с такими значениями. Далее была построена гистограмма длин белков, с помощью встроенных функций была подсчитана и описана статистика: максимальная и минимальная длины белка, средняя длина, среднестатистическое отклонение от среднего значения. С помощью СЧЁТЕСЛИМН было посчитали число генов белков, псевдогенов и генов РНК, на прямой и комплементарной цепи. Все упомянутые выше действия были произведены в программе Microsoft Office Excel.

## 3 РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

### 3.1 Белки

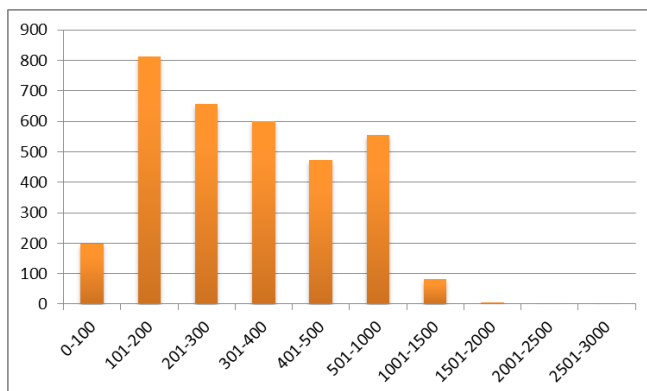
Геном бактерии *Gemmatimonas phototrophica* содержит 3388 последовательностей, кодирующих белок. Длина белков варьировалась от самой маленькой в 43 аминокислоты до самой большой в 2787 аминокислот. Средняя длина была около 353 аминокислот. Статистические данные приведены в Таблице 1.

**Таблица 1.** Статистические данные о длине белков бактерии *G. phototrophica*.

Максимальная длина	43
Минимальная длина	2787
Среднее значение	352.7576741
Среднеквадратичное отклонение от среднего	236.1325385
Медиана	303

Чаще всего в протеоме данной бактерии можно найти белки длины от 100 до 200 аминокислот, что видно на приведенной ниже Диаграмме 1.

Диаграмма 1. Распределение белков по длине.



[https://kodomo.fbb.msu.ru/~natalia\\_kondratenko/term1/pr13.xlsx](https://kodomo.fbb.msu.ru/~natalia_kondratenko/term1/pr13.xlsx)

## 6 СПИСОК ЛИТЕРАТУРЫ

- [1] Zeng, Y., Selyanin, V., Lukes, M., Dean, J. (2015) *International journal of systematic and evolutionary microbiology* : Characterization of the microaerophilic, bacteriochlorophyll a-containing bacterium *Gemmatimonas phototrophica* sp. nov., and emended descriptions of the genus *Gemmatimonas* and *Gemmatimonas aurantiaca*.
- [2] Zeng, Y., (2014) *Proceedings of the National Academy of Science of the United States of America*: Functional type 2 photosynthetic reaction centers found in the rare bacterial phylum Gemmatimonadetes.

### 3.2 Распределение последовательностей, кодирующих РНК, белки и псевдогены по цепям ДНК

Всего в геноме бактерии *G. phototrophica* AP64 3965 кодирующих последовательностей, при этом белки кодируются 3388 последовательностями, а РНК – 54 последовательностями, также в геноме насчитывается 523 псевдогена. В Таблице 2 приведены данные о распределении кодирующих последовательностей между прямой и обратной цепями ДНК. Как можно видеть из данной таблицы, на прямой и обратной цепях ДНК находится примерно одинаковое количество кодирующих последовательностей.

Таблица 2. Распределение кодирующих последовательностей.

Цепь	Белки	Псевдогены	Гены РНК
Прямая	1626	263	26
Обратная	1762	260	28

## 4 ЗАКЛЮЧЕНИЕ

- (1) В геноме бактерии *Gemmatimonas phototrophica* AP64 наиболее распространены белки в диапазоне от 100 до 200 аминокислот. В целом, чаще всего встречаются белки с длиной от 100 до 400 аминокислот, но есть и маленькие (43 аминокислоты), и большие (2787 аминокислот).
- (2) Кодирующие последовательности встречаются на прямо и обратной цепях ДНК с приблизительно равной частотой

## 5 СОПРОВОДИТЕЛЬНЫЕ МАТЕРИАЛЫ

Файл со всеми таблицами, на основе которых была написана эта статья, доступен по ссылке: