

# Мини-обзор протеома бактерии *Pseudomonas putida* KT2440

Потапенко Елена<sup>1</sup>

<sup>1</sup>Факультет биоинженерии и биоинформатики МГУ, 119991, Москва, Ленинские горы 1, стр. 73

## АННОТАЦИЯ

В данном обзоре исследованы геном и протеом бактерии *P. putida* KT2440. Изучены распределение длин белков, положение кодирующих последовательностей на прямой и обратной цепях; число РНК- и белок-кодирующих генов (по группам).

**Ключевые слова:** *Pseudomonas putida*, протеом, геном, Excel.

## 1 INTRODUCTION

*P. putida* – непатогенная Грам-отрицательная сапротрофная почвенная бактерия. Может расщеплять стирол до биоразлагаемых пластиков, может быть использована в переработке полистирольных смол, которые долго считались биологически неразлагаемыми. Легко поддается биоинженерным манипуляциям, поэтому используется в фармацевтике и сельском хозяйстве для синтеза различных органических веществ [3]. Полное таксономическое положение штамма представлено в таблице 1.

**Таблица 1.** Таксономическое положение *P. putida*.

Царство	Bacteria
Тип	Proteobacteria
Класс	Gammaproteobacteria
Порядок	Pseudomonadales
Семейство	Pseudomonadaceae
Род	<i>Pseudomonas</i>
Вид	<i>P. putida</i>

## 2 МАТЕРИАЛЫ И МЕТОДЫ

Хромосомная таблица получена с сайта NCBI [1]. Обработка данных производилась с помощью программы Excel 2016. Использовались математические, логические и статистические функции, фильтр из меню “data”, а также построение гистограммы.

Соответствующими генам белков считались строчки, в колонках “class” которых значится “CDS”.

Гены, лежащие на прямой и обратной цепях – имеющие знак “+” и “-” в колонке “strand” соответственно.

Гены рибосомальных белков – содержащие в столбце «name» подстроку «ribosomal» и подстроку «protein»; транспортные белки – содержащие в том же столбце подстроку «transport»; гипотетические – содержащие там же подстроку «hypothetical».

Также использовались скрипты на Python 2, в основном чтобы найти среди названий белков конкретные функции. Так, для подсчета аминокислот-тРНК-синтетаз из списка названий белков выбирались названия, содержащие подстроку «tRNA».

Возможные «лишние» попавшиеся названия убирались вручную.

## 3 РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Геном состоит из одной кольцевой хромосомы длиной 6.0072 миллионов пар оснований, содержащей 5786 генов, из которых 5564 – белок-кодирующие [2]. Таким образом, на тысячу пар оснований приходится в среднем один кодирующий участок.

### 3.1 Распределение длин белков.

Как видно из гистограммы на рисунке 1, большая часть белков бактерии имеет длину от 70 до 350 аминокислотных остатков, такие числа типичны для большинства бактерий.



**Рисунок 1.** Распределение длин белков в протеоме *P. putida*

Статистические данные для длин белков *P. putida* представлены в таблице 2. Стоит заметить, что самый длинный белок длиной 2478 а.о. относится к hypothetical protein, также как и несколько других «портящих статистику» белков, в то время как все белки с известными функциями имеют длину в пределах 1000 а.о.

**Таблица 2.** Статистические данные для длин белков *P. putida*.

Параметр	Значение
Минимальная длина	29
Максимальная длина	2478
Средняя длина	301
Стандартное отклонение	212

Функциональное распределение белков показано в таблице 3. Среди 58 рибосомальных белков 22 принадлежат субъединице 30S и 34 белка к субъединице 50S. Кроме того, в состав субъединиц рибосом входят 15 рибосомальных РНК. Остальные РНК бактерии – одна некодирующая и 51 транспортная. Транспортных тРНК 51 по количеству триплетов, используемых бактерией для кодирования аминокислот. Кодированных аминокислот 20, но некоторые кодируются не одним триплетом. Также в протеоме есть аминоацил-тРНК-синтетазы, например, tyrosyl-tRNA synthase и methionyl-tRNA synthetase. Их количество почти вдвое больше числа

**Таблица 3.** Разные типы белков.

Тип белка	Число белков такого типа
рибосомальные	58
транспортные	156
гипотетические	341
остальные	1154

**Таблица 4.** Разные типы РНК.

Тип РНК	Число РНК такого типа
рибосомальные	58
транспортные	156
гипотетические	341
остальные	1154

Много в протеоме и белков-транспортёров. На их примере хорошо видно, что субъединицы одного фермента (или отдельные ферменты, выполняющие последовательные функции) лежат в геноме на одной цепи на небольшом расстоянии друг от друга. Например, на участке длиной около 10000 нуклеотидов лежит несколько транспортёров разных сахаров (маннозы, лактозы, аскорбиновой кислоты и еще нескольких). Можно предположить, что бактерии удобнее экспрессировать подряд гены со сходными функциями, или что эволюционно сходные транспортёры были получены из одного транспортёра дубликацией с последующими изменениями (первый вариант кажется мне более вероятным, но не исключающим второй).

По белкам-транспортёрам можно предположить, к каким условиям бактерия может приспосабливаться: например, у рассматриваемой *P. putida* есть транспортёр кобальта, а также транспортёры из семейства NRAMP, выводящие ионы тяжелых металлов. Такие белки обеспечивают устойчивость к ионам, которые для других организмов являются ядовитыми и позволяют освоить новые среды обитания.

### 3.2 Расположение разных типов генов на прямой и обратной цепях ДНК

Число генов каждого типа, расположенных на прямой и комплементарной цепях хромосомы, приведены в таблице 3.

**Таблица 3.** Статистика генов на прямой и обратной цепях.

Тип гена	Прямая цепь	Обратная цепь
Protein_coding	801	908
Pseudogene	23	18
All types of RNA	37	31

Стоит заметить, что у всех трех типов генов распределение по цепям можно считать случайным. Соответствующие расчеты можно найти в сопроводительных материалах.

### СОПРОВОДИТЕЛЬНЫЕ МАТЕРИАЛЫ

Все вычисления собраны в один файл Excel, который Вы можете скачать по ссылке: [http://kodomu.fbb.msu.ru/~potapenko/term1/pr14\\_potapenko.xlsx](http://kodomu.fbb.msu.ru/~potapenko/term1/pr14_potapenko.xlsx) Пример скрипта на Python, использовавшегося для сортировки функций белков: <http://kodomu.fbb.msu.ru/~potapenko/term1/script.py>

### ССЫЛКИ

- [1] [ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA/000/007/565/GCA\\_000007565.2\\_AS\\_M756v2](ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA/000/007/565/GCA_000007565.2_AS_M756v2)
- [2] <https://www.ncbi.nlm.nih.gov/genome/?term=kt2440>
- [3] [https://en.wikipedia.org/wiki/Pseudomonas\\_putida](https://en.wikipedia.org/wiki/Pseudomonas_putida)