

## ОБЗОР КОДИРУЮЩИХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ ГЕНОМА БАКТЕРИИ *ENTEROCOCCUS HIRAE* ATCC 9790

АННА ПАГО

Московский Государственный Университет им. М. В. Ломоносова, Факультет биоинженерии и  
биоинформатики ГСП-1, Ленинские горы, д. 1, стр. 73, Москва, Россия, 119991  
annapago@fbb.msu.ru

В данной работе исследовались некоторые характеристики кодирующих последовательностей бактерии *Enterococcus hirae* ATCC 9790, с целью получения статистических данных о длинах и функциях белков, а также распределения генов по комплиментарным цепям.

*Ключевые слова:* геном, *Enterococcus hirae*.

### 1. Введение

*Enterococcus hirae* – это ГРАМ-позитивная каталазотрицательная молочно-кислая бактерия<sup>1</sup>, которая является факультативным анаэробом. Данную бактерию используют для производства кисло-молочных продуктов. Также она уже более 40 лет она является объектом проведения различных исследований. За это время она меняла своё название 2 раза, с *Streptococcus faecalis* на *Streptococcus faecium*, и позже на *Enterococcus hirae*.

Исследование 2012 года показало, что в геноме этой бактерии присутствует эндемическая плазида, состоящая из 28 699 п.о., которую исследователи назвали pTG9790. Она кодирует 33 гипотетических белка, но только четверо похожи на белки с уже известными функциями. Один из них — это белок, отвечающий за инициацию процесса образования спор, что весьма интересно, учитывая то, что *Enterococcus hirae* спор не образует<sup>4</sup>. С этим видом связано возникновение задержек в развитии молодых куриц, а также, возможно, эндокардита<sup>5</sup> и септицемии у людей<sup>6</sup>.

### 2. Методы

Данные, анализируемые в этой работе, были взяты с сайта NCBI<sup>7</sup>. С помощью программы LibreOffice Calc была построена гистограмма, отображающая распределение длин белков, получены максимальное, минимальное и среднее значения, стандартное отклонение и медиана, построена таблица распределения генов белков, РНК и псевдогенов по двум цепям ДНК. Также было посчитано количество белков и РНК с различными функциями.

### 3. Результаты и обсуждение

#### 3.1 Анализ длин белков

В процессе исследования генома *Entrococcus hirae* ATCC 9790 была построена гистограмма распределения длин белков (рис. 1). Минимальное, среднее, максимальное значения, стандартное отклонение и медиана представлены в таблице 1. Хотя длины белков варьируются от 30 до 2161 а.о., из гистограммы видно, что длина большинства белков не превышает 500 а.о.



Рис.1. Распределение длин белков

Максимальное значение	Минимальное значение	Стандартное отклонение	Среднее значение	Медиана
2161	30	196,58	274,5	229

Таблица 1. Статистические данные

#### 3.2 Распределение по комплиментарным цепям ДНК

Результаты исследования распределения генов, кодирующих белки, РНК и псевдогены по цепям ДНК представлено в таблице 2. Видно, что геном содержит всего один псевдоген.

Цепь	Белки	РНК	Псевдогены
+	1405	60	0
-	1350	29	1
<b>Сумма</b>	<b>2755</b>	<b>89</b>	<b>1</b>

Таблица 2. Распределение по цепям ДНК

### 3.3 Распределение кодирующих последовательностей по функциям

Исследовалось количество гипотетических белков, белков-транспортёров и гипотетических белков, а также количество псевдогенов, рибосомальных и транспортных ДНК. Как можно увидеть из таблицы 3, большинство кодирующих последовательностей — это белки, причём больше трети из них гипотетические, а транспортных РНК почти в четыре раза больше рибосомальных.

<b>Белки</b>	2755	%
Транспортёры	152	5,34%
Рибосомальные	64	2,25%
Гипотетические	1047	36,8%
Остальные	1492	52,44%
<b>Псевдогены</b>	1	0,035%
<b>РНК</b>	89	3,13%
Транспортные	71	2,5%
Рибосомальные	18	0,63%
<b>Всего</b>	<b>2845</b>	<b>100%</b>

Таблица 3. Распределение по функциям

## 4. Сопроводительные материалы

Файл в формате Excel

### **Благодарности**

Выражаю благодарность преподавателям информатики факультета Биоинженерии и Биоинформатики за предложение провести эту работу и всю помощь, предоставленную в выполнении этого задания.

### **Источники**

---

<sup>1</sup> <https://jb.asm.org/content/194/18/5126>

<sup>2</sup> <http://mic.microbiologyresearch.org/content/journal/micro/10.1099/mic.0.026385-0>

<sup>3</sup> <http://www.findpatent.ru/patent/247/2476592.html>

<sup>4</sup> <https://jb.asm.org/content/194/18/5126>

<sup>5</sup> <https://jcm.asm.org/content/40/7/2689>

<sup>6</sup> <https://web.archive.org/web/19991128143611/http://link.springer.de/link/service/journals/10096/bibs/8017008/80170576.htm>

<sup>7</sup> <https://www.ncbi.nlm.nih.gov/>