

Мини-обзор

# Анализ бактерии Streptomyces clavuligerus

Фарниева В. А.

Факультет биоинженерии и биоинформатики, Московский государственный университет имени М.В. Ломоносова, Ленинские горы д.1 стр.73, 119234, Москва, Российская Федерация.

#### 1 Введение

Streptomyces clavuligerus - вид грамположительных бактерий, известных способностью продуцировать цефалоспорины и пенициллин [1].

S. clavuligerus ATCC 27064 (NRRL 3585, DSM 738) был впервые описан Хиггинсом и Костнером (1971), которые выделили его из образца почвы Южной Америки [2]. Его название относится к форме его спороносных гифальных ветвей: от латинского clavicula "маленькая дубинка" и igerus "несущий".

**Научная классификация**: *Streptomyces* относится к домену Бактерии, типу Actinomycetes, классу Actinomycetes, отряду Streptomycetes, семейству Streptomycetes, роду *Streptomyces*, виду *S. clavuligerus*.

#### 2 Материалы и Методы

В процессе исследования длины белков и GC-состава закодированных в геноме была использована таблица CDS from genome of Streptomyces clavuligerus. Предсказание функций гипотетических белков осуществлялось с помощью найденных идентификаторов отдельно взятых белков, для поиска последовательностей аминокислот ID белка вбивался в поисковую строку сайта NCBI Protein [3]. А для поиска схожих последовательностей использовался сайт NCBI BLAST, по типу поиска Вlastр вводилась найденная последовательность, после чего выводилась таблица находок. Для извлечения данных по ключевым словам и их структурирования был использован язык программирования Python [4].

### 3 Результаты

3.1 Длины белков закодированных в геноме Streptomyces clavuligerus

Наибольшее количество белков находится в пределах от 0 до 150 аминокислот. Это означает что белки с такой маленькой длинной могут выполнять такие простые функции, как связь с другими молекулами, а также участвуют в простых метаболических процессах и могут функционировать как гормоны или как другие сигнальные молекулы[5].

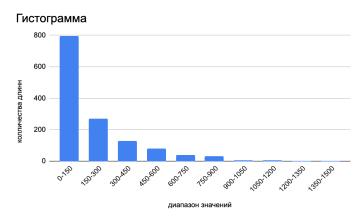
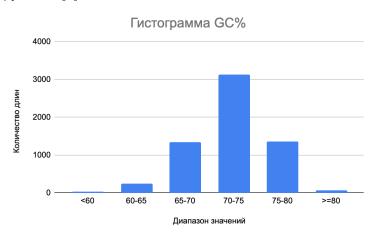


Рисунок 1. Гистограмма длин белков бактерии Streptomyces clavuligerus

#### 3.2 GC-состав (GC%) по CDS белка

GC-состав генома бактерии — это соотношение гуанина и цитозина (G и C) к общему количеству нуклеотидов в ДНК бактерии Streptomyces clavuligerus. Он выражается в процентах и играет важную роль в характеристике генома. У Streptomyces clavuligerus большинство последовательностей имеют содержание GC в пределах 70-75%, что может указывать на особенности генетической структуры. Такие уровни GC характерны для генетического материала некоторых бактерий, где высокая стабильность и структура ДНК играют ключевую роль в поддержании клеточных функций [6].



**Рисунок 2.** Гистограмма GC-состава (GC%) по CDS белка бактерии Streptomyces clavuligerus

#### 3.3 Гипотетические белки в геноме Streptomyces clavuligerus

Гипотетические белки — это белки, предсказанные на основе анализа геномных последовательностей, но их структура, функции и существование не подтверждены. Гипотетические белки *Streptomyces clavuligerus* представляют собой неизученный пласт информации, который может содержать важные функциональные особенности организма. Основываясь на результатах выполнения BLAST а именно таблицы находок содержащую в себе Total score (балл который учитывает количество совпадений и их расположение) E-value (показывает вероятность того, что совпадение между последовательностями произошло случайно) Identity (процент идентичности аминокислот которые совпали между собой) Length (длина последовательности): WP\_157849174.1 имеет сходство с глутаматным рецептором, можно сделать вывод, что гипотетический белок, вероятно, участвует в клеточной организации, возможно, через глутаматную систему, и его функция может быть связана с передачей сигналов или обработкой внешних клеточных стимулов [7].

WP\_230864229.1 белок, схожий с токсином из системы RelE/ParE в Streptomyces niveus, скорее всего, дает клетке возможность адаптироваться к стрессовым условиям и обеспечивает ее выживание в неблагоприятных ситуациях[8].

WP\_003956287.1 также схож с WP\_230864229.1 но с меньшим совпадением WP\_003953814.1 гипотетический белок, вероятно, является фосфатной гидролазой из семейства HAD, которая участвует в метаболизме фосфатов и сигнальных путях. Он может выполнять функции, связанные с разрушением фосфатных связей и регуляцией клеточных процессов [9].

## Благодарности.

Выражаю глубокую благодарность преподавательскому составу за качественное и понятное изложение материала, развитие необходимых профессиональных навыков, терпение и неизменную поддержку в решении сложных вопросов. Ваш вклад в подготовку высококвалифицированных специалистов невозможно переоценить!

#### Сопроводительные материалы.

- 1. Таблица особенностей CDS from genome of Streptomyces clavuligerus
- 2. Гистограмма GC-состава (GC%) по CDS белка ☐ GC%
- 3. Скрипт <a href="https://colab.research.google.com/drive/1x-pR8nx9occlHMvqyVaKCax2gcwhvU0V?usp=sharing">https://colab.research.google.com/drive/1x-pR8nx9occlHMvqyVaKCax2gcwhvU0V?usp=sharing</a>
- 4. Выдача Blast (таблица находок) 🖬 Blast

#### Список литературы.

- Feeney et al (2022). ActionBase: tools and protocols for researchers working on Streptomyces and other filamentous actinobacteria https://actinobase.org/index.php/Streptomyces\_clavuligerus?utm\_source
- Author(s). (Year). Title of the article. International Journal of Systematic and Evolutionary Microbiology, 21(4), 326. <a href="https://www.microbiologyresearch.org/content/journal/ijsem/10.1099/00207713-21-4-326">https://www.microbiologyresearch.org/content/journal/ijsem/10.1099/00207713-21-4-326</a>

- The National Center for Biotechnology Information advances science and health by providing access to biomedical and genomic information. <a href="https://www.ncbi.nlm.nih.gov/protein">https://www.ncbi.nlm.nih.gov/protein</a>
- 4. Скрипт <a href="https://colab.research.google.com/drive/1x-pR8nx9occlHMvqyVaKCax2gcwhvU0V?usp=sharing">https://colab.research.google.com/drive/1x-pR8nx9occlHMvqyVaKCax2gcwhvU0V?usp=sharing</a>
- Wang Z, Zuo Z, Li L, et al. Nrf2 in adipocytes. Arch Pharm Res. National Center for Biotechnology Information. (2024). BLAST: Basic Local Alignment Search Tool. National Institutes of Health. <a href="https://pubmed.ncbi.nlm.nih.gov/32189204/">https://pubmed.ncbi.nlm.nih.gov/32189204/</a>
- 6. Wikipedia contributors. (2024). *GC content*. Wikipedia, The Free Encyclopedia. <a href="https://en.wikipedia.org/wiki/GC-content">https://en.wikipedia.org/wiki/GC-content</a>
- The National Center for Biotechnology Information advances science and health by providing access to biomedical and genomic information https://www.ncbi.nlm.nih.gov/protein/1046926207
- 8. The National Center for Biotechnology Information advances science and health by providing access to biomedical and genomic information <a href="https://www.ncbi.nlm.nih.gov/protein/WP\_392683685.1?report=genbanklog\$=prottop&blast\_rank=90&RID=PVWGY62N013">https://www.ncbi.nlm.nih.gov/protein/WP\_392683685.1?report=genbanklog\$=prottop&blast\_rank=90&RID=PVWGY62N013</a>.
- 9. Кунин Е.В., Татусов Р.Л. (ноябрь 1994 г.). <a href="https://zenodo.org/records/1229874">https://zenodo.org/records/1229874</a>