
Обзор генома и протеома бактерии *Paenibacillus odorifer*

Михайлова Екатерина

Факультет биоинженерии и биоинформатики, МГУ

РЕЗЮМЕ

В данной работе содержится анализ генома и протеома бактерии *Paenibacillus odorifer*. Посчитаны GC-skew и частоты стоп-кодона, построена гистограмма длин белков, показано распределение генов по цепям ДНК.

Ключевые слова: геном, протеом, бактерия, *Paenibacillus odorifer*

1 ВВЕДЕНИЕ

Paenibacillus odorifer - грамположительная, спорообразующая бактерия, была выделена из лука-порея и других продуктов питания [1]. Показано, что данная бактерия способствует порче молока [2,3], ее споры стабильны к термической обработке, а адаптации, связанные с холодовым шоком [6], позволяют расти при температурах охлаждения на протяжении всего хранения молока, вызывая протеолиз и липолиз [2-5]. Также *P. odorifer* способна к нитрат и нитрит редукции [7]. Целью работы является изучение генома и протеома данной бактерии.

2 МЕТОДЫ

Информация о геноме и протеоме получена из базы данных NCBI [8].

Для построения графика 1 была использована программа GenSkew[9]. Данная программа строит график интегрального значения (cumulative GC-skew) в определенном диапазоне(window), по позициям идут с некоторым шагом(step). Также на основе графика выводятся значения максимума и минимума. Window и step программа выбирает автоматически, на основе предложенного генома.

Код, с помощью которого получены данные из таблицы 1, был написан на языке Python, вместе с пояснениями его можно найти по ссылке в разделе “Сопроводительные материалы”. Для составления рисунка 1 была использована программа Google sheets,

Геном *Paenibacillus odorifer* состоит из одной кольцевой хромосомы и имеет длину 6,8 Mb, что является средним значением для рода *Paenibacillus* [10]. GC - состав данной бактерии равен 44,1 % [11], это немного меньше GC-состава бактерии *Paenibacillus sp. JDR-2* (50,3 %) из того же рода [12]. Не очень высокий GC-состав согласуется с тем, что бактерия не является термофильной.

4.2 GC-skew

GC-skew представляет собой соотношение между нуклеотидами G (гуанин) и C (цитозин) и вычисляется по формуле:

$$GC\ skew = (n(G) + n(C)) / (n(G) - n(C))$$

где $n(G)$, $n(C)$ - число нуклеотидов G и C.

С его помощью можно найти точку начала и терминции репликации, для этого вычисляется cumulative GC-skew, который представляет собой сумму всех GC-skew посчитанных ранее. Известно, что минимуму GC-skew соответствует точка начала репликации (oriC), а максимуму - точка терминции репликации (ter).

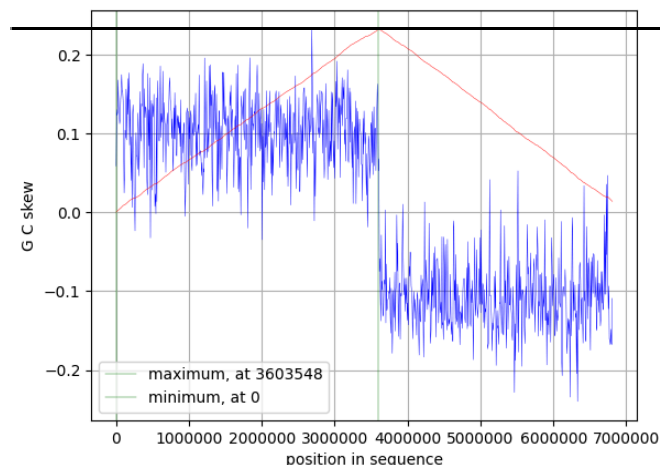
С помощью программы[9] построен график cumulative GC-skew вдоль генома *Paenibacillus odorifer*, window и step равны 6812 п.н.

Минимум cumulative GC-skew равен 0, а максимум – 3603548. Значит предположительные oriC и ter соответственно равны 0 и 3603548.

3 РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

4 СОПРОВОДИТЕЛЬНЫЕ МАТЕРИАЛЫ

4.1 Стандартные данные о геноме



4.3 Частоты стоп-кодонов в кодирующих последовательностях

Частоты стоп-кодонов приведены в таблице 1.

Данные подтверждают гипотезы, приведенные в статье[13]: высокая частота кодона ТАА может быть связана тем, что ТАА может узнаваться и фактором терминации RF1 и фактором терминации RF2, в то время как TGA и TAG – только одним.

Также бактерия предпочитительнее использует ТАА при низком GC-составе, что как раз характерно для *Paenibacillus odorifer*.

Таблица 1

стоп-кодон	TAG	TAA	TGA
количество	1292	3355	1260

4.4 Гистограмма длин белков

Больше всего белков длиной 280-300 аминокислотных остатков

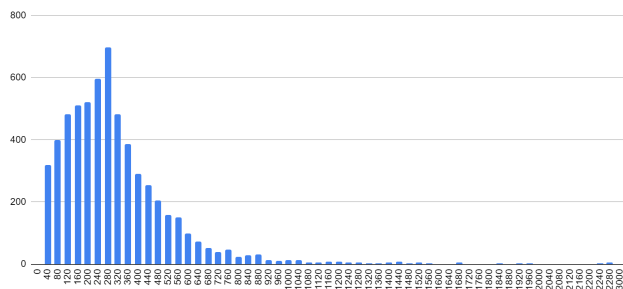


Рисунок 2. Гистограмма длин белков

Таблица 1

стоп-кодон	TAG	TAA	TGA
------------	-----	-----	-----

количество 1292 3355 1260

4.5 Распределение генов по цепям

В таблице 2

Таблица 2

название	на “+” цепи	на “-” цепи	степень случайности
gene protein coding	2841	3018	0,022
pseudogene	37	22	0,067
rRNA	18	12	0,362
tRNA	40	47	0,520

ACKNOWLEDGEMENTS

REFERENCES

1. Berge, O., Guinebretiere, M.H., Achouak, W., Normand, P. and Heulin, T. (2002) *Paenibacillus graminis* sp. nov. and *Paenibacillus odorifer* sp. nov., isolated from plant roots, soil and food. *Int J Syst Evol Microbiol* 52, 607–616.
2. Beno SM, Cheng RA, Orsi RH, Duncan DR, Guo X, Kovac J, Carroll LM, Martin NH, Wiedmann M. 2020. *Paenibacillus odorifer*, the predominant *Paenibacillus* species isolated from milk in the United States, demonstrates genetic and phenotypic conservation of psychrotolerance but clade-associated differences in nitrogen metabolic pathways. *mSphere* 5:e00739-19
3. Ivy RA, Ranieri ML, Martin NH, Bakker HC, Xavier BM, Wiedmann M, Boor KJ. 2012. Identification and characterization of psychrotolerant sporeformers associated with fluid milk production and processing. *Appl Environ Microbiol* 78:1853–1864.
4. Doll EV, Scherer S, Wenning M. 2017. Spoilage of microfiltered and pasteurized extended shelf life milk is mainly induced by psychrotolerant spore-forming bacteria that often originate from recontamination. *Front Microbiol* 8:135.
5. Ranieri M, Huck J, Sonnen M, Barbano D, Boor K. 2009. High temperature, short time pasteurization temperatures inversely affect bacterial numbers during refrigerated storage of pasteurized fluid milk. *J Dairy Sci* 92:4823–4832.
6. Phadtare S. 2004. Recent developments in bacterial cold-shock response. *Curr Issues Mol Biol* 6:125–136.
7. Buxton R. 2011. *Nitrate and nitrite reduction test protocols*. American Society for Microbiology, Washington, DC. Accessed 1 October 2019.
8. https://ftp.ncbi.nlm.nih.gov/genomes/all/GCF/000/758/725/GCF_000758725.1_ASM75872v1
9. <https://genskew.csb.univie.ac.at/webskew>
10. <https://www.ncbi.nlm.nih.gov/genome/browse#!/overview/Paenibacillus>

11. <https://www.ncbi.nlm.nih.gov/genome/?term=Paenibacillus+odorifer>

12. Lightfield J, Fram NR, Ely B (2011) Across Bacterial Phyla, Distantly-Related Genomes with Similar Genomic GC Content Have Similar Patterns of Amino Acid Usage. PLoS ONE 6(3): e17677. doi:10.1371/journal.pone.0017677
13. Alexander T Ho, Laurence D Hurst, Variation in Release Factor Abundance Is Not Needed to Explain Trends in Bacterial Stop Codon Usage, *Molecular Biology and Evolution*, 2021;,msab326,<https://doi.org/10.1093/molbev/msab326>