

Обзор протеома бактерии *Oenococcus oeni* PSU-1

Ногина Д.С.¹

¹Факультет биоинженерии и биоинформатики МГУ им. М.В.Ломоносова

Составлено 8 декабря, 2018

В данной работе представлен статистический анализ генома и протеома бактерии *Oenococcus oeni* PSU-1. Было установлено, что гены не случайно распределены по прямой и обратной цепи, и длина и количество оперонов на разных цепях различны. Были рассмотрены трансмембранные белки данного вида и выявлены общие закономерности, характерные для *O. oeni*.

трансмембранный белок | оперон | ацидофильные бактерии

Бактерия *Oenococcus oeni* - это ацидофил, в процессе метаболизма которого образуется лактат и этанол. Биотехнологическое значение данной бактерии состоит в том, что она осуществляет малолактозную ферментацию вина (1). Малолактозное брожение делает вино более кислым вследствие превращения малакта в лактат.

В 1961 Garvie (2) впервые предложил выделить вид *Leuconostoc oenos* из-за его устойчивости к низкому значению pH и интенсивного роста на среде, содержащей томатный сок. Позднее Dicks et al. (3) порекомендовали реклассифицировать *L. oeni* в новый род *Oenococcus*.

На момент написания известно, что геном *O. oeni* PSU-1 состоит из одной кольцевой хромосомы длиной 1,780,517 нуклеотидов. Содержание GC-пар составляет порядка 38%.

Известно, что вино обладает низким pH (от 2,5 до 4,0) и большой концентрацией спирта в результате первого этапа ферментации *Saccharomyces*. Также в вине относительно мало питательных веществ, причём в нём присутствуют полифенолы, ингибирующие рост бактерий (1). Эти обстоятельства делают *O. oeni* интересным объектом для изучения.

В данной работе были изучены закономерности расположения генов на прямой и обратной цепях хромосомы, распределение белков по длине и функциям для штамма PSU-1. Также были рассмотрены пересечения генов и предсказаны потенциальные опероны. Особое внимание уделено белковым кофакторам и трансмембранным белкам. Выяснено, что основные закономерности расположения трансмембранных доменов *O. oeni* отличаются от таковых *E. coli*.

Методы

Общие данные о генах бактерии были взяты из базы данных Genome NCBI (ABJ56155.1). Данные о трансмембранных участках, кофакторах и ЕС-коды белковых продуктов, а также о том, подтверждено ли состояние гена, были получены из базы Uniprot. Для обработки полученных данных использовалась программа Microsoft Excel 2016. С её помощью были построены все гистограммы и круговая диаграмма. В работе использовались функции МИН, МАКС, МЕДИАНА, СТАНДОТКЛОН.В, СРЗНАЧ, ЧАСТОТА, ОКРУГЛЕНИЕ, ДЛСТР, СЧЁТЕСЛИ и прочие. Для исследования взаимодействий белков квазиоперонов

использовалась база данных String.

Результаты и обсуждения

1. Общая информация о распределении генов

Данный штамм содержит 1691 белок-кодирующий ген, 122 псевдогена и 51 ген РНК, из которых 43 тРНК, 7 рРНК и 1 тмРНК. В дальнейшем под термином "гены" понимаются все гены, за исключением псевдогенов. Интересно наличие транспортно-матричной РНК в геноме *O. oeni*. Данная молекула участвует в транс-трансляции, связываясь с А-сайтом рибосомы в случае разрыва мРНК и навешивая на синтезируемый белок метку для протеолиза. Наличие подобного механизма может свидетельствовать о том, что у бактерии существует повышенный риск сбоя в процессе трансляции.

2. Распределение длин белков

На рис. 1 (Fig. 1) представлена гистограмма распределения длин белков. Медиана составляет 262 аминокислотных остатка, стандартное отклонение - около 186 аминокислотных остатков. Большая часть белков находится в диапазоне 50-350 остатков. Самым крупным белком (1475 аминокислот) является гликозил-гидролаза, ассоциированная с мембраной. Функция самого небольшого (22 аминокислоты) белка не определена.

3. Распределение генов на прямой и обратной цепи

Распределение генов на прямой и обратной цепи представлено в таблице 1 (Table 1). Была проверена гипотеза о случайности распределения генов по цепям ДНК. Для этого была использована встроенная функция Excel БИНОМРАСП. Было выяснено, что вероятность того, что гены распределены по цепям не случайно равна около 99,5%.

Significance Statement

Бактерия *Oenococcus oeni* - важный биотехнологический объект. Она участвует в подкислении вина. *O. oeni* - ацидофил, что делает её изучение интересным не только с практической, но и с теоретической точки зрения. Опираясь на данные для генома бактерии, мы обнаружили особенности строения её трансмембранных белков и расположения генов на хромосоме. Дальнейшие исследования помогут понять, как устроен метаболизм ацидофильных бактерий.

E-mail: nogina_daria@mail.ru

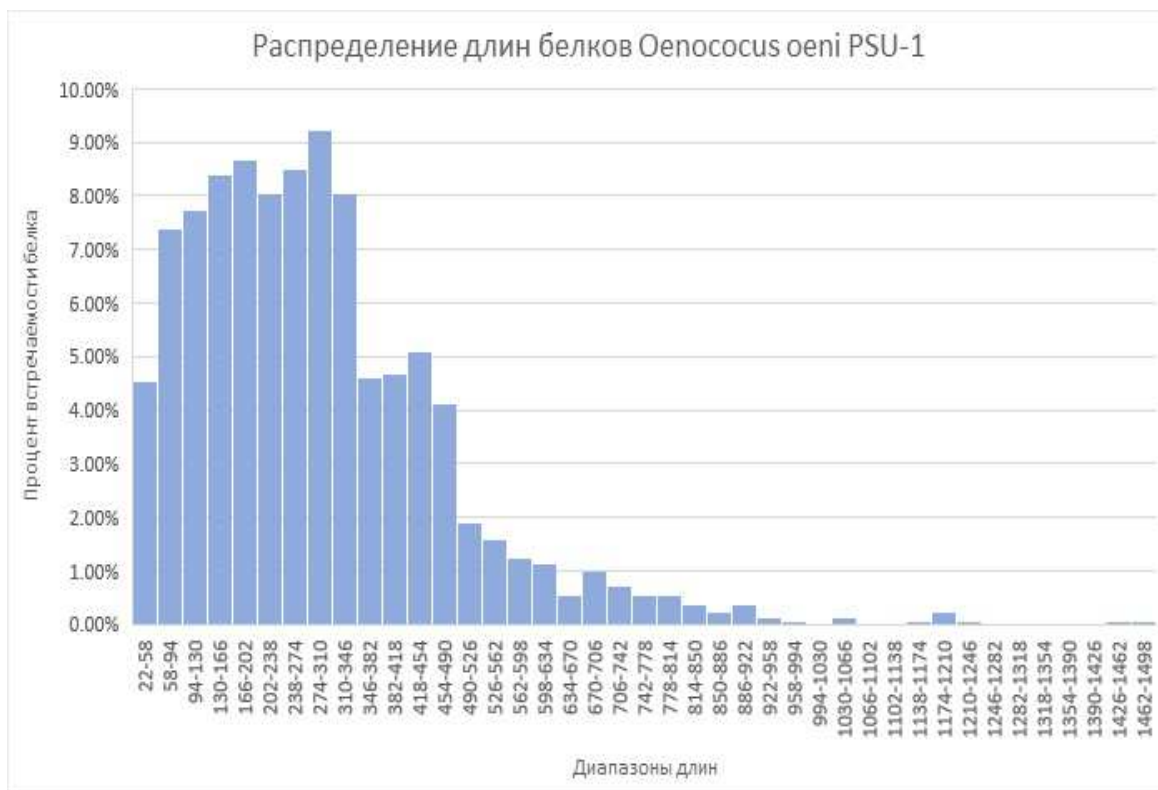


Fig. 1. Гистограмма распределения длин белков *O. oeni*.

Table 1. Число белок-кодирующих генов, псевдогенов и РНК-генов по прямой и комплементарной цепям ДНК.

Category	"+" chain	"-" chain	sum
1. protein-coding	882	809	1691
2. pseudogene	1862	1744	3606
3. RNA-gene	980	935	1915

4. Число и распределение квазиоперонов

Квазиоперон - это гипотетический оперон. Квазиопероны были найдены на основе межгенного расстояния. Максимально допустимым межгенным расстоянием для оперона в работе считалось расстояние в 50 нуклеотидов. Данное значение было выбрано на основе исследования (4). Было выяснено, что в сумме на прямой цепи 494 оперона, что в 1.3 раза меньше, чем на обратной (656 оперона). При этом максимальная длина оперона на прямой цепи составляет 18 генов, что больше, чем на обратной (8 генов). Объяснение данного явления может служить темой дальнейших исследований.

Большая часть квазиоперонов включает в себя один ген. На гистограммах (Fig. 2, Fig. 3) представлены распределения длин оперонов, состоящих из более чем одного гена, на прямой и обратной цепи.

Самые длинные опероны прямой и обратной цепи были протестированы на достоверность с помощью базы данных String. Белки, кодируемые опероном длины 18 прямой цепи, и их взаимодействия представлены на Fig. 4, опероном длины 8 обратной цепи соответственно на Fig. 5.

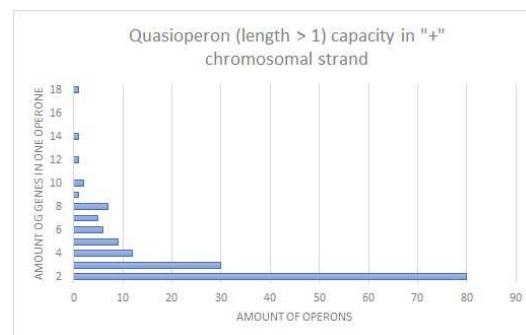


Fig. 2. Распределение длин квазиоперонов прямой цепи.

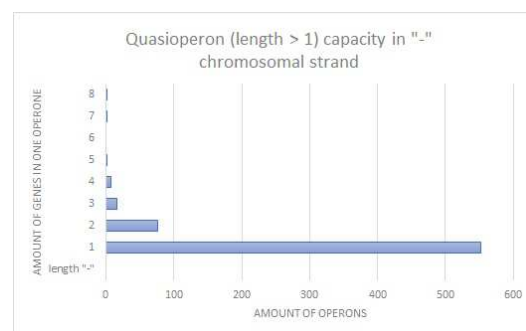


Fig. 3. Распределение длин квазиоперонов обратной цепи.

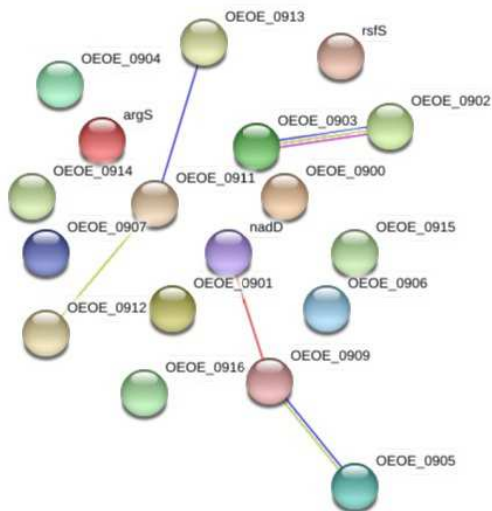


Fig. 4. Взаимодействие белков, кодируемых самым длинным опероном прямой цепи.

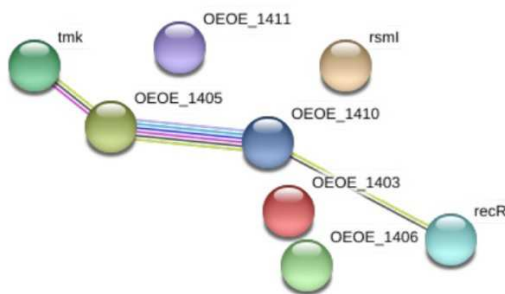


Fig. 5. Взаимодействие белков, кодируемых самым длинным опероном обратной цепи.

PROTEIN EXCISTENCE

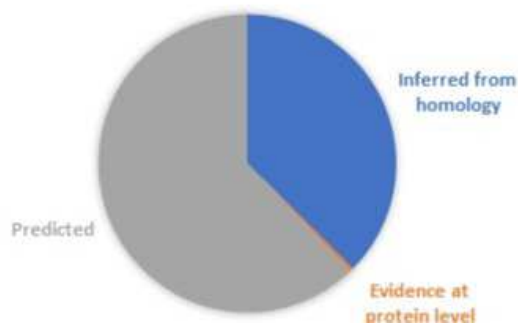


Fig. 6. Распределение белков *O. oeni* по категориям достоверности их существования согласно базе Uniprot.

Взаимодействие между белками квазиоперона длиной 18 генов небольшое. Это можно объяснить с одной стороны тем, что найденный квазиоперон слишком длинный для настоящего оперона. С другой стороны, нам доступно сравнительно мало информации об *O. oeni* (см. Fig. 6). Белки из "длинных квазиоперонов" обеих цепочек главным образом ответственны за репарацию ДНК.

5. Перекрытия генов

На графике (Fig. 7) отображена статистика пересечения генов на разноимённых цепях хромосомы. Данное явление наблюдается из-за тенденции бактерий к компактизации генома. На гистограмме видно, что большая часть перекрытий происходит без сдвига рамки считывания. Также количество перекрытий на одноимённых цепях больше, чем на разноимённых.

6. Общй анализ протеома

На основе данных, полученных из базы Uniprot, был проведён общй анализ протеома *O. oeni*. В результате анализа ЕС-номеров белков были получены тривиальные результаты (Fig. 8).

Распределение количества белков *O. oeni*, связывающихся с определёнными кофакторами, по сравнению с аналогичным распределением для *E. coli*, показывает в том числе то, что процент белков с марганцевыми кофакторами для *O. oeni* в 3.3 раза меньше, чем для *E. coli* (Fig. 9).

Самыми высокопредставленными семействами белков являются транспортные пермеазы и АВС-транспортёры. Данные результаты не являются уникальными для *O. oeni*. В результате подобного анализа *E. coli* были получено, что для этой бактерии данные семейства также являются самыми распространёнными.

7. Трансмембранные белки

Как было сказано выше, *O. oeni* - ацидофил. Для того, чтобы синтезировать АТР, используя внутренний ток протонов через F_0F_1 АТР-азу, бактерия должна поддерживать постоянный градиент рН на мембране. Вследствие этого трансмембранные белки играют ключевую роль в метаболизме ацидофилов, и их

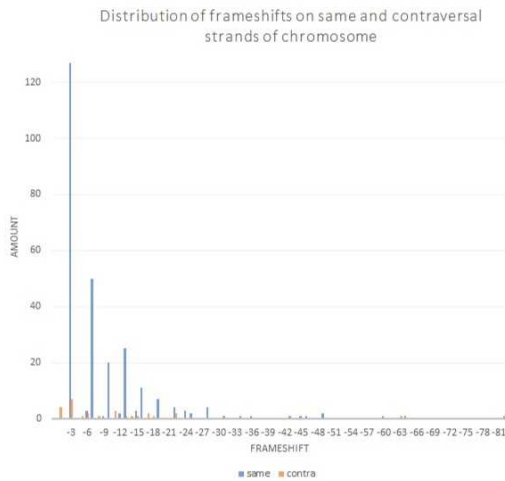


Fig. 7. Распределение пересечений генов на цепях хромосомной ДНК.

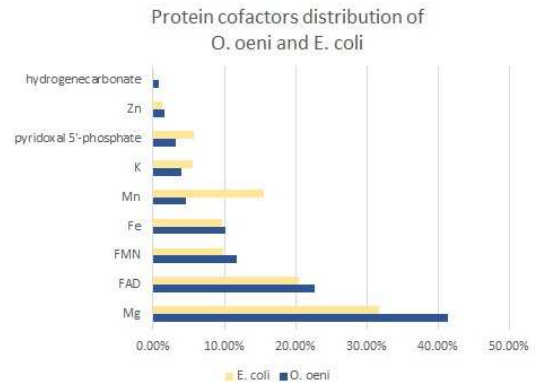


Fig. 9. Распределение числа белков, связывающихся с определёнными кофакторами.

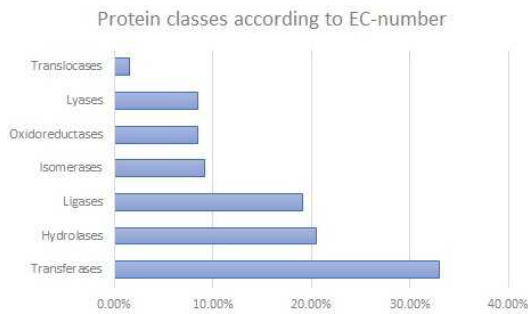


Fig. 8. Основные классы белков *O. oeni*.

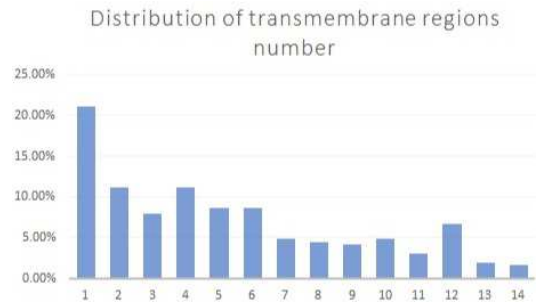


Fig. 10. Распределение белков, имеющих различное количество трансмембранных участков.

исследование может пролить свет на малоизученный метаболизм данных экстремофилов.

В работе были рассмотрены трансмембранные участки мембранных белков, в качестве источника информации служила база Uniprot. В результате анализа распределения белков по количеству трансмембранных доменов была получена гистограмма (Fig. 10). Большая часть белков имеет один трансмембранный домен. Было выдвинуто предположение о том, что в Uniprot не учтены посттрансляционные модификации, и подавляющая часть белков, учтённая как белки с одним трансмембранным доменом, - это белки с лидерным пептидом. Данное предположение основывалось на том, что строение лидерного пептида близко к строению трансмембранных участков.

Для проверки данной гипотезы был произведён анализ белков с одним "трансмембранным" доменом на предмет того, в какой части белка располагается интересующий нас участок. Сравнение полученных результатов с аналогичными результатами для *E. coli* представлены на Fig. 11. Заметно явное отличие между *E. coli* и *O. oeni* на участках, соответствующих около 70-80% от общей длины белка. Данное явление довольно любопытно и требует дальнейшего объяснения.

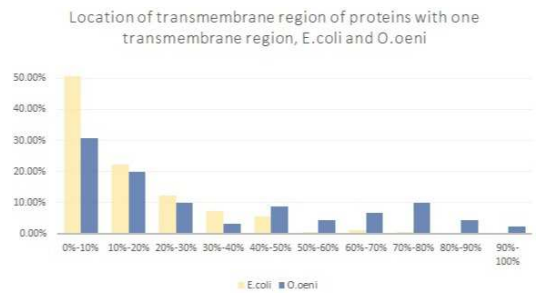


Fig. 11. Распределение белков по расположению единственного трансмембранного участка.

Заключение

Открытые базы данных открывают множество перспектив для "омикс"-исследований. В ходе анализа генома бактерии *Oenococcus oeni* PSU-I были выявлены её особенности и, в частности, отличия от *E. coli*. Автор полагает, что основное направление будущих исследований - это попытка объяснить необычное распределение белков по расположению трансмембранного участка.

Сопроводительные материалы

[ссылка на скачивание файла](#)

БЛАГОДАРНОСТИ

Автор глубоко признателен Сергею Александровичу Спирину за терпение и оптимизм.

1. David A. Mills, Helen Rawsthorne KM (2005) Genomic analysis of *oenococcus oeni* psu-1 and its relevance to winemaking. *FEMS Microbiology Reviews* 29(3):465–475.
2. Garvie E (1967) *Leuconostoc oenos* sp.nov. *J. Gen. Microbiol.* 48:431–438.
3. Dicks, L.M.T. DF, Collins M (1995) Proposal to reclassify *leuconostoc oenos* as *oenococcus oeni* (corrig) gen nov, comb nov. *Int. J. Syst. Bacteriol.* 45:395–397.
4. MJL de Hoon, S Imoto KK (2004) Predicting the operon structure of *bacillus subtilis* using operon length, intergene distance, and gene expression information. *Biocomputing* pp. 276–287.