

Обзор протеома бактерии *Lactobacillus plantarum* ZJ316

Преображенская Ю.А.

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

РЕЗЮМЕ

Данная работа посвящена анализу протеома бактерии *Lactobacillus plantarum* ZJ316, направленному на установление распределения белков по длине, расположение их на + и – цепи ДНК, объединению в квазиопероны, установление степени перекрываемости генов; также осуществлялась проверка гипотезы случайности расположения генов, кодирующих РНК, на прямой и обратной цепях.

1 ВВЕДЕНИЕ

Lactobacillus plantarum ZJ316 - грамположительная факультативно анаэробная или микроаэрофильная бактерия. Подавляет развитие различных патогенных микроорганизмов в пробирке. Может метаболизировать широкий спектр растительных углеводов; устойчива к воздействию солей желчных кислот и низкому рН, антагонист в отношении патогенных микроорганизмов кишечника (1). В данном роде представлена 231 плаزمид. Геном *L.plantarum* штамма ZJ316 (BGI-Shenzhen, Shenzhen, China) секвенирован по методу дробовика. Штаммы молочнокислых бактерий, включая данный, используют в производстве медицинских препаратов — пробиотиков, предназначенных для восстановления нормальной микрофлоры кишечника и (после инфекционных заболеваний, антибиотикотерапии). Способна жить в очень кислой среде (рН около 0) и температуре выше 65 С, что делает ее одной из самых термоацидофильных организмов. Денитрифицирующая, метаногенная, нитрифицирующая, имеет восстановительный пентозофосфатный путь, метаболизирующая никотинат и аскорбат, сульфатредуцирующая бактерия (2).

Анализ протеома данной бактерии проводился с целью установить некоторые характеристики протеома с использованием программы Excel.

2 МАТЕРИАЛЫ И МЕТОДЫ

Исходная таблица с протеомом была получена со страницы ftp://ftp.ncbi.nlm.nih.gov/genomes/Bacteria/Lactobacillus_plantarum_ZJ316_uid18868 (3), использовались файлы NC_020229.ptt, NC_021904.ptt, NC_021912.ptt, NC_021903.ptt; NC_020229.rnt, NC_021904.rnt, NC_021912.rnt, NC_021903.rnt. Анализ данных и построение диаграммы были выполнены с помощью программы Microsoft Office Excel 2007. Были использованы функции «СЧЕТЕСЛИМН» для анализа распределения белков по длинам, а также логические функции для подсчета количества квазиоперонов и пересекающихся генов, «МЕДИАНА» для анализа средней длины протеинов, «БИНОМРАСПР» для проверки гипотезы случайной встречаемости генов RNA и CDS в зависимости от направления цепи (обратная (-) или прямая(+)).

3 РЕЗУЛЬТАТЫ

Далее можно ознакомиться с результатами проведенных исследований.

3.1 Плазмиды, представленные в данном роде

Анализируя данные о плаزمиде бактерий данного рода, подведем небольшие итоги:

- (1) В роде *Lactobacillus* представлена 231 плазмид.
- (2) В штамме *Lactobacillus plantarum* ZJ316 представлено 3 плазмиды, самая длинная из которых - 41.508Kb, самая короткая - 15.167Kb.
- (3) Средняя длина составила 31.930Kb.

3.2 Распределение генов по цепям

Результаты анализа распределения генов по цепям ДНК представлены в таблице 1.

Таблица 1. Распределение генов белков и РНК по цепям ДНК в протеоме бактерии *L.plantarum* ZJ316

Тип гена	Полный геном	Число генов		
		pLP-ZJ101 (плазмид)	pLP-ZJ102 (плазмид)	pLP-ZJ103 (плазмид)
CDS				
(+)	1575	5	37	12
CDS (-)	1584	12	16	35
RNA(+)	40			
RNA(-)	36			
Всего генов	3235	17	53	47

На основании данных мы можем составить промежуточные выводы:

- (1) Гены, кодирующие белки, преобладают на (-)-цепи.
- (2) Меньше всего генов расположено в плазмиде pLP-ZJ101.
- (3) Суммарно, количество генов белков больше количества генов служебных РНК.
- (4) В двух плаزمиде (pLP-ZJ101 и pLP-ZJ103) заметно, обратной (-) цепи для локализации генов белков.

- (5) Ни в одной плазмиде из указанных нет генов, кодирующих служебные РНК.
- (6) На прямой цепи встречается на 36 генов, кодирующих белки каналов (названия заканчиваются на –аза, –портер, –транспорт) больше, чем на обратной (данные в сопроводительном файле)

3.3 Распределение длин белков

Результаты анализа распределения белков по длине представлены на рисунке 1 и в сопроводительном файле. В протеоме *L.plantarum* ZJ316 всего 3260 белковых продуктов, наибольшее количество белков имеют длину от 0 до 100 аминокислот. Подавляющая часть белков (более 92%) входит в класс длин от 50 до 700 а-т. Белков, имеющих длину > 900 а-т, всего 1,33% (22 шт.), < 50 а-т – менее 1% (85 шт.).

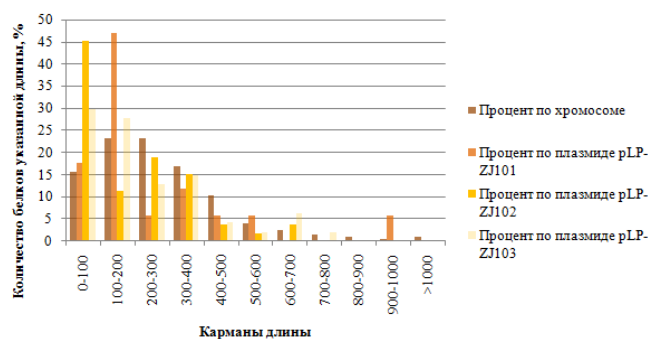


Рис.1. Распределение длин белков протеома *L.plantarum* ZJ316.

3.4 Количество квазиоперонов (возможных оперонов)

С помощью средств программы Excel было подсчитано, что количество квазиоперонов в хромосоме составляет 1058, в плазмиде pLP-ZJ101 – 6, pLP-ZJ102 – 16, pLP-ZJ103 – 11 (если брать за максимальное допустимое расстояние между генами одного оперона 100 п.н.). В данном анализе мы предполагали, что один ген также может быть опероном. При этом максимально возможное теоретическое количество квазиоперонов, равное количеству генов, составляет 3159. Если порог максимального расстояния между генами в квазиопероне изменить на 50 п.н., квазиоперонов оказывается 727 (т.е. генов, объединенных в квазиопероны с какими-то еще генами, становится меньше). Если пороговое расстояние = 200 п.н., квазиоперонов 1439. Результаты анализа для плазмид представлены в таблице 2 и таблице 3.

Таблица 2. Количество квазиоперонов при пороговом расстоянии, равном 50 п.н.

Цепь	Квазиопероны			
	Хромосома	pLP-ZJ101	pLP-ZJ102	pLP-ZJ103
(+)	364	1	9	3
(-)	363	2	2	6
Всего	727	3	11	9

Таблица 3. Количество квазиоперонов при пороговом расстоянии, равном 200 п.н.

Цепь	Квазиопероны			
	Хромосома	pLP-ZJ101	pLP-ZJ102	pLP-ZJ103
(+)	701	2	20	4
(-)	738	8	7	15
Всего	1439	10	27	19

3.5 Проверка гипотезы случайной встречаемости генов РНК в одной из цепей ДНК

В результате анализа числа генов РНК на прямой цепи, нами было выяснено, что вероятности встреч генов CDS и RNA на прямой и обратной цепях приближаются к статистической случайности (точное значение можно увидеть в сопроводительном файле), что было подсчитано с помощью вышеуказанных функций программы Excel. Далее на рисунке 2 составлен график, иллюстрирующий данное распределение.

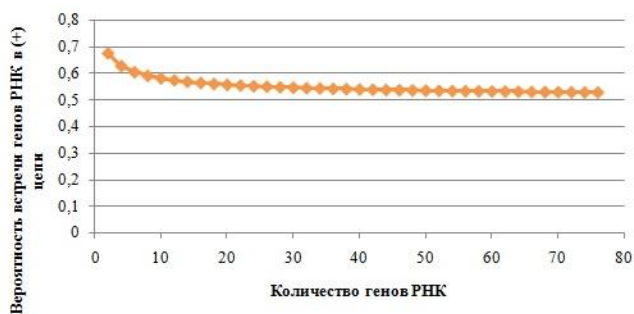


Рис.2. Проверка гипотезы случайной встречаемости генов РНК в одной из цепей ДНК *L.plantarum* ZJ316.

3.6 Пересечение генов

Анализу были подвергнуты хромосома, плазмиды pLP-ZJ101, pLP-ZJ102, pLP-ZJ103, в процессе чего выяснилось, что перекрывается 9,75% генов хромосомы (рассчитывалась разность между конечной и начальной координацией гена, затем, если значение было меньше нуля, то это свидетельствует о наличии перекрывания). На основе этих данных был составлен статистический обзор, представленный диаграммой ниже.

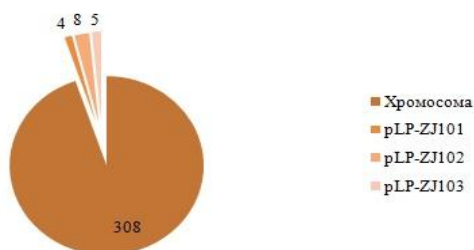


Рис.3. Статистический обзор по перекрытиям генов в хромосоме и плазидах *L. plantarum* ZJ316 (количество перекрывающихся генов)

ОБСУЖДЕНИЕ

В протеоме *Lactobacillus plantarum* ZJ316 представлено 3260 белков, 3 рибосомальные и 50 транспортных РНК. Количество генов, кодирующих белок или служебную РНК, на + и – цепи (соответственно 40 и 36) приблизительно одинаково и их распределение носит случайный характер. Средняя длина белка, кодируемого в бактериальной хромосоме – около 285 аминокислот, – близка к таковой у других бактерий и намного меньше таковой у дрожжей (средняя длина дрожжевого белка составляет около 460 а-т); в плазмиде pLP-ZJ101- 250 аминокислот, pLP-ZJ102-203 аминокислот, pLP-ZJ103-228 аминокислот. В целом длина белков описывается сдвинутым влево одномодальным распределением. Результаты по пересечению генов – около 9,74% всех генов пересекаются – согласуются с тем, что геном бактерий обычно довольно сильно компактизирован. Данные о количестве квазиоперонов - 1128 при общем количестве генов 3235 (на хромосоме) – свидетельствуют о том, что у данной бактерии небольшое количество генов может быть объединено в опероны, что не является типичным для прокариот; кроме того данные согласуются с результатами анализа пересечения генов. Довольно большое количество плазмид, представленное в данном роде, возможно, свидетельствует о высокой приспособленности к широкому спектру условий среды (например, отсутствие чувствительности к высоким температурам, ультрафиолету и т.д.).

ЗАКЛЮЧЕНИЕ

Стоит отметить, что в процессе проведенного исследования была обнаружена неочевидная закономерность – предпочтительное размещение генов канальных белков и белков-транспортёров на прямой цепи в хромосоме (установление эволюционной целесообразности данного признака не входило в рамки исследования).

Феномен генов, чья длина не делится на 3 объясняется тем, что данные гены кодируют РНК-последовательности, а для РНК триплетность необязательна.

СОПРОВОДИТЕЛЬНЫЕ МАТЕРИАЛЫ

Ссылка на xls- файл с таблицами и расчетами

<http://kodomo.fbb.msu.ru/~july.preobrazhencki/term1/excelblock4.html>

СПИСОК ЛИТЕРАТУРЫ

1. Данные о метаболизме *L. plantarum* ZJ316
<http://bmcvetres.biomedcentral.com/articles/10.1186/1746-6148-8-89>
2. Характеристика метаболизма *L. plantarum* ZJ316
<http://www.kegg.jp/pathway/lpt01120>
3. Страница сайта NCBI, на которой размещен геном *L. plantarum* ZJ316
ftp://ftp.ncbi.nlm.nih.gov/genomes/Bacteria/Lactobacillus_plantarum_ZJ316_uid18868