

ПОЛНОГЕНОМНОЕ СЕКВЕНИРОВАНИЕ ГЕНОМА ШТАММА *LACTOBACILLUS FERMENTUM* 630

Алексеева А.Е.¹, Точилина А.Г.¹, Бруснигина Н.Ф.¹, Солнцев Л.А.¹, Белова И.В.¹,
Ефимов Е.И.¹, Соловьева И.В.¹

¹ФБУН «Нижегородский научно-исследовательский институт эпидемиологии и микробиологии имени академика И.Н. Блохиной Роспотребнадзора Российской Федерации», Нижний Новгород, Россия (603950, Нижний Новгород, ул. Малая-Ямская, 71), e-mail: mazepavn@mail.ru

Проведено полногеномное секвенирование штамма *Lactobacillus fermentum* 630, выделенного из кишечника здорового человека, на высокопроизводительном секвенаторе второго поколения MiSeq. В результате выравнивания и объединения коротких чтений было получено 390 контигов и 250 скаффолдов, общая длина нуклеотидной последовательности составила 1,830,322 нуклеотидов. В результате биоинформационного анализа было определено значительное количество однонуклеотидных полиморфизмов, инсерций и делеций относительно референсного штамма *Lactobacillus fermentum* F6. Филогенетический анализ установил, что наиболее близкородственным является штамм *Lactobacillus fermentum* 39. При аннотировании с помощью сервера RAST было выявлено наличие 1668 белок кодирующих последовательностей, 58 транспортных РНК и 19 рибосомальных РНК. Также были обнаружены последовательности мобильных элементов, фаговых интеграз, генов, ответственных за антибиотикорезистентность.

Ключевые слова: *Lactobacillus fermentum*, полногеномное секвенирование, филогенетический анализ, аннотация генома

WHOLE GENOME SEQUENCING OF *LACTOBACILLUS FERMENTUM* STRAIN 630

Alekseeva A.E.¹, Tochilina A.G.¹, Brusnigina N.F.¹, Solntsev L.A.¹, Belova I.V.¹,
Efimov E.I.¹, Soloveva I.V.¹

I.N. Blokhina Research Institute of Epidemiology and Microbiology, Nizhny Novgorod, Russia (603950, Nizhny Novgorod, Malaya-Yamskaya street, 71), e-mail: mazepavn@mail.ru

A whole genome sequencing of *Lactobacillus fermentum* strain 630 isolated from the gut of a healthy person was conducted on the next-generation sequencer MiSeq. As a result, of alignment and assembly of short reads 250 scaffolds were obtained, total length of the nucleotide sequence was 1,830,322 nucleotides. Bioinformatic analysis revealed a significant number of single nucleotide polymorphisms, insertions and deletions relative to the reference strain *L. fermentum* F6. Phylogenetic analysis has established that the strain *L. fermentum* 39 is the most closely related one. Using RAST annotation server the presence of 1668 protein-coding sequences, 58 tRNAs and 19 rRNA were revealed. Also a sequence of transposable elements, prophages and the genes responsible for antibiotic resistance were detected.

Keywords: *Lactobacillus fermentum*, whole genome sequencing, phylogenetic analysis, genome annotation

Микроорганизмы рода *Lactobacillus* широко распространены в природе, а некоторые виды являются важнейшими представителями микробиома человека [3]. Положительная роль бактерий данного рода в сохранении гомеостаза организма достоверно доказана [2, 3, 4], что обуславливает их использование при производстве кисломолочных продуктов, обогащении различных продуктов питания, создании пробиотиков в форме лекарственных препаратов и БАД к пище.

В настоящее время актуальным является изучение не только фенотипических, но и генетических свойств промышленно-перспективных штаммов лактобацилл с

использованием различных молекулярно-генетических методов, в том числе основанных на секвенировании нуклеиновых кислот.

Современные технологии массивного параллельного секвенирования (next-generation sequencing – NGS) позволяют получать информацию о полногеномных последовательностях широкого спектра бактерий и вирусов, и проводить молекулярно-генетические исследования на качественно новом уровне [1].

Цель исследования: изучение структуры генома штамма *Lactobacillus fermentum 630* с использованием метода полногеномного секвенирования (whole-genome shotgun sequencing).

Материалы и методы.

Объектом исследования являлся штамм *L. fermentum 630* из Государственной коллекции лактобацилл Нижегородского НИИЭМ им. академика И.Н. Блохиной, выделенный из кишечника здорового человека.

Видовая принадлежность исследуемого штамма подтверждена путем изучения его биохимических свойств с использованием тест-систем API 50 CHL (Biomerieux, Франция) и масс-профилей рибосомальных белков, полученных с помощью времяпролетного MALDI масс-спектрометра Autoflex (Bruker Daltonics, Германия). Клетки исследуемого микроорганизма имеют вид неподвижных полиморфных грамположительных палочек, не образующих спор и капсул.

Выделение и очистку ДНК исследуемого штамма проводили с использованием набора «Ампли-Сенс ДНКсорб В» (ЦНИИЭ, Москва). Концентрацию ДНК в образцах определяли с помощью флуориметра Qubit (Invitrogen, Австрия). Подготовку библиотеки ДНК для секвенирования осуществляли с использованием набора Nextera XT (Illumina, США) на 500 циклов согласно инструкции производителя. Секвенирование генома проводили на платформе NGS MiSeq (Illumina, США) в режиме ресеквенирования. В качестве референса служила полногеномная последовательность штамма *L. fermentum F6* (номер GenBank NC_021235.1). Для выравнивания нуклеотидных последовательностей использовали программу Burrows-Wheeler Aligner (BWA), для поиска однонуклеотидных полиморфизмов, инсерций и делеций – программное обеспечение Genome Analysis Toolkit (GATK). Визуализацию и анализ полученных данных проводили с помощью программного обеспечения UGENE [6]. Аннотацию генома проводили с использованием сервера RAST (Rapid Annotation using Subsystem Technology) [7]. Филогенетическое дерево строили с помощью программы Mega 6.0 [8] методом ближайших соседей (neighbor-joining), предварительно проведя множественное выравнивание нуклеотидных последовательностей 19 геномов *L. fermentum*, задепонированных в базе данных NCBI (National Center for

Biotechnology Information) и *L. fermentum* 630 с помощью сервера RALPHY 1.10 (Reference sequence Alignment based Phylogeny builder) [5].

Результаты и обсуждение

В результате секвенирования было получено 2 237 480 коротких чтений длиной до 250 н.о. Вторичный анализ позволил получить 390 контигов, которые были объединены в 250 скаффолдов. Общая длина последовательности составила 1,830,322 пар нуклеотидов, средний уровень покрытия - 54,6. С помощью программного обеспечения GATK было обнаружено 15076 однонуклеотидных полиморфизмов, 213 инсерций и 253 делеции относительно референсного штамма *L. fermentum* F6.

При аннотировании с помощью сервера RAST в геноме было выявлено 1668 белок кодирующих последовательностей, 58 последовательностей тРНК и 19 рРНК.

В базе данных NCBI представлены нуклеотидные последовательности геномов 15 штаммов *L. fermentum* в виде контигов и скаффолдов (на 12 ноября 2015 года). В таблице дана сравнительная характеристика структур генома *L. fermentum* 630 и неполных геномов штаммов *L. fermentum*, задепонированных в базе данных Whole Genome Shotgun contigs (WGS) NCBI.

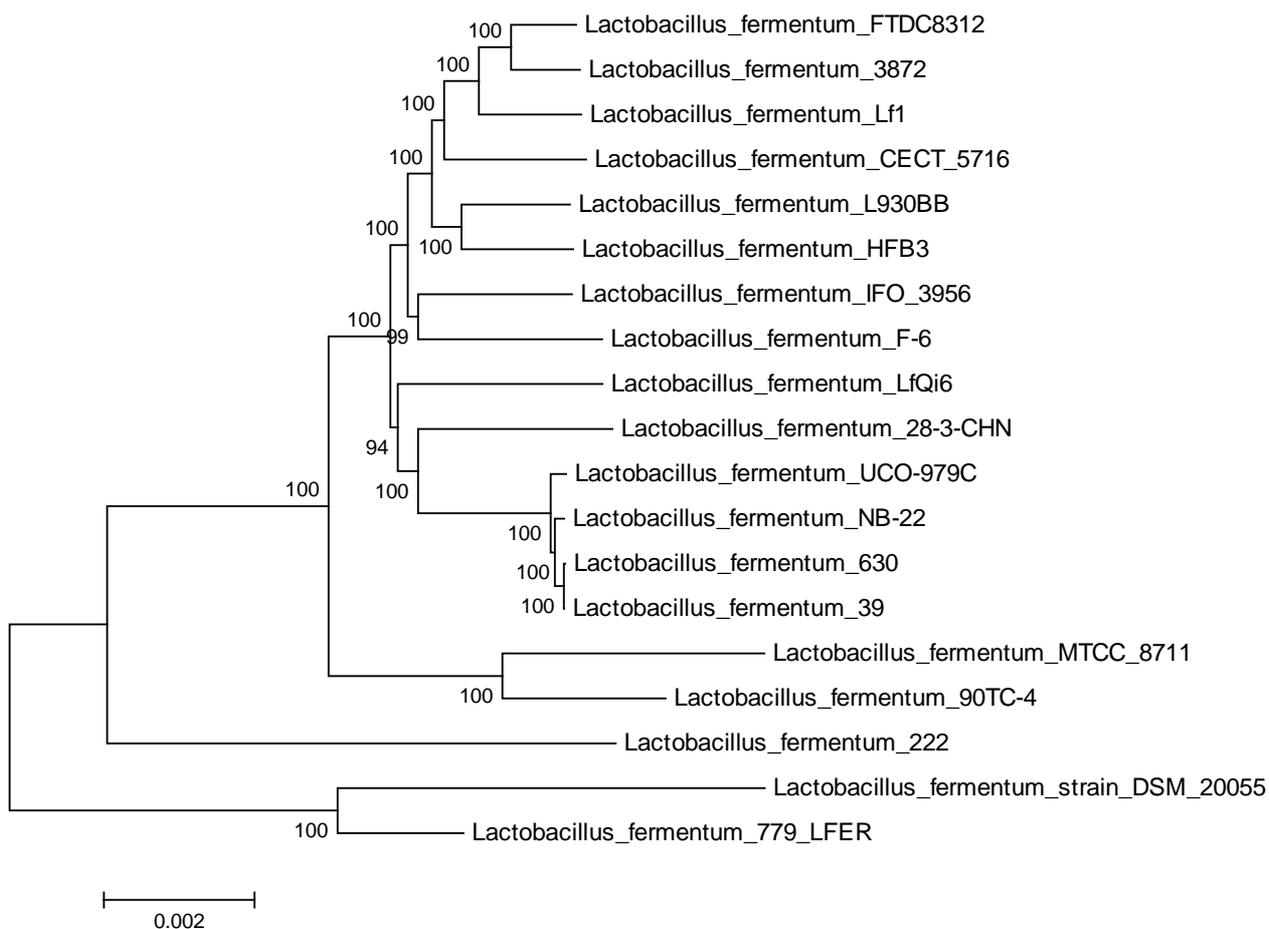
Сравнительная характеристика структур геномов *L. fermentum* 630 и задепонированных в базе данных WGS NCBI

Штамм <i>L. fermentum</i>	Размер, Mb	% ГЦ	Число белок кодирующих последовательностей	Кол-во тРНК	Кол-во рРНК
630	1,83032	52,45	1668	58	19
ATCC 14931	1,86701	52,60	1656	56	4
28-3-CHN	2,02652	52,00	1806	53	4
FTDC8312	1,96655	51,80	1718	16	2
NB-22	2,01131	51,80	1826	31	4
LfQi6	2,1998	51,50	1847	56	8
DSM 20055	1,90005	52,4	1713	47	3
Lf1	1,81565	52,50	1633	57	15
MTCC8711	2,5663	46,65	1911	57	15
39	1,82966	51,60	1658	51	13
90 TC-4	1,82248	51,90	1571	54	11
L930BB	1,97184	52,10	1767	56	7
779_LFER	1,93581	52,10	1729	49	15
HFB3	2,04336	51,80	1382	58	20
UCO-979C	2,01183	51,90	1492	57	13
222	1,95041	52,10	1864	54	4

Следует отметить, что данные, полученные в нашем исследовании, совпадают с результатами аналогичных проектов, представленных в таблице. Высокопроизводительное

секвенирование генома штамма *L. fermentum* 630 позволило определить 80-90% нуклеотидной последовательности.

При проведении филогенетического анализа исследуемого штамма с последовательностями геномов *L. fermentum*, представленных в базе данных NCBI, было установлено что штамм *L. fermentum* 630 находится на одной ветке со штаммом *L. fermentum* 39. Филогенетически близкородственными являются также штаммы *L. fermentum* NB-22 и *L. fermentum* UCO-979C (рисунок).



Филогенетическое дерево штаммов L. fermentum, построенное на основании анализа нуклеотидных последовательностей геномов методом ближайших соседей. Цифрами указана статистическая значимость порядка ветвления (в %), определенная с помощью бутстрап-анализа 1000 альтернативных деревьев

В структуре генома с помощью сервера RAST обнаружены нуклеотидные последовательности 21 мобильного элемента, 2 генов, кодирующих фактор элонгации G, обеспечивающего устойчивость к тетрациклину, и 2 генов, кодирующих белки семейства бета-лактамаз, гена хлорамфеникол ацетилтрансферазы, детерминирующего устойчивость к хлорамфениколу, а также гены белков семейства фаговых интеграз.

Выводы:

1. С использованием высокопроизводительного секвенирования на секвенаторе MiSeq получена молекулярно-генетическая характеристика штамма *L. fermentum 630*.
2. В геноме *L. fermentum 630* выявлено значительное количество однонуклеотидных полиморфизмов, а также инсерций и делеций относительно референсного штамма *L. fermentum F6*
3. Наиболее близкородственным, по результатам филогенетического анализа является штамм *L. fermentum 39*.
4. Обнаруженные последовательности мобильных элементов, фаговых интеграз, генетических детерминант антибиотикорезистентности послужат основой для создания генетической карты генома и обнаружения генетических детерминант, обуславливающих уникальные фенотипические свойства.

Список литературы

1. Алексеева А.Е. Бруснигина Н.Ф. Полногеномное секвенирование в системе эпидемиологического надзора за инфекционными заболеваниями// ВИНТИ РАН. – 2015. – № 130-B2015. – 14с.
2. Белова И.В. Авторский синбиотик с антимикотической активностью: основные биологические свойства, эффективность применения// Медицинский альманах. – 2011. – № 4. – С. 84-88.
3. Бондаренко В.М., Грачева Н.М., Мацулевич Т.В. Дисбактериозы кишечника у взрослых. – М.: КМК Scientific Press, 2003. – 206 с.
4. Соловьева И.В. Профилактика формирования устойчивых дисбиотических состояний микрофлоры желудочно-кишечного тракта студентов с помощью нового симбиотика// Медицинский альманах. – 2011. – № 2. – С. 112-114.
5. Bertels F. Automated reconstruction of whole genome phylogenies from short sequence reads// Mol. Biol. Evol. – 2014. – vol. 31, № 5. – P. 1077-1088.
6. Okonechnikov K. Unipro UGENE: a unified bioinformatics toolkit//Bioinformatics. – 2012. – №28. – P. 1166-1167.
7. Overbeek R. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST)// Nucleic Acids Res. – 2014. – №42. (Database issue). – D: 206-214. [Электронный ресурс]. URL: <http://doi.org/10.1093/nar/gkt1226>. (дата обращения: 01.11.2015)
8. Tamura K. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0// Mol. Biol. Evol. – 2013. – vol .30, № 12. – P. 2725-2729.

Рецензенты:

Смирнов В.Ф., д.б.н., профессор кафедры физиологии и биохимии человека и животных Федерального государственного автономного образовательного учреждения высшего образования «Национальный исследовательский государственный Нижегородский университет им. Н.И. Лобачевского» Министерства образования России, г.Нижний Новгород;

Заславская М.И., д.б.н., профессор кафедры микробиологии и иммунологии Государственного бюджетного образовательного учреждения высшего профессионального образования «Нижегородская государственная медицинская академия» Министерства образования и социального развития России, г.Нижний Новгород.