

Метод высокопроизводительного секвенирования нового поколения для изучения полного генома *Staphylococcus* spp

Слизень В. В., Горецкий Д., Гудкова Е. И.

Учреждение образования «Белорусский государственный медицинский университет»,
г. Минск, Республика Беларусь

Реферат. В работе представлены результаты первого полногеномного секвенирования *Staphylococcus aureus* в Республике Беларусь. Исследован штамм *Staphylococcus aureus* BLR-DV, выделенный из мочи 73-летнего пациента с урологическим сепсисом. Высокопроизводительное секвенирование генома выполняли с использованием технологии MiSeq, Illumina и MiniOn, Nanopore. Сборка осуществлена 25 июня 2020 г. программами SPAdes v. 3.14; A5-miseq v. 20160825; Flye v. 2.7b; Canu v. 1.8. Покрытие генома составляло 400.0x. Секвенированный геном изолята *Staphylococcus aureus* BLR-DV размером 2,969,706 пар нуклеотидов загружен в генетический банк Gen Bank; код доступа CP058312.1. Данный изолят также содержал две плазмиды — pSTA-BLR-MRSA-DV-1 (код доступа Gen Bank — CP058313.1) и pSTA-BLR-MRSA-DV-2 ((код доступа Gen Bank CP058314.1) размером 2979 п. о. и 2348 п. о. Содержание G+C составляло 32,9 %. В функциональной организации генома установлено наличие 278 субсистем, которые кодируют 2839 белков, 77 молекул РНК.

Полученные экспериментальные данные по структуре полного генома *Staphylococcus aureus* sBLR-DV позволяют оценивать резистентность изолята к противомикробным лекарственным средствам, а также выявлять генетические особенности, важные для эпидемиологического типирования.

Ключевые слова: полный геном, *Staphylococcus aureus*, анализ полного генома.

Введение. Геном *Staphylococcus aureus* сложен и пластичен, что позволяет стафилококкам быстро приспосабливаться к внешним условиям, включая больничные, формировать устойчивость к противомикробным лекарственным средствам за счет мутаций и горизонтального переноса генов в составе

плазмид и стафилококковых хромосомных генетических кассет, которые могут содержать ген *mecA* или *mecC* и другие гены устойчивости [1].

Множественно лекарственно устойчивые стафилококки представляют угрозу для здоровья пациентов особенно в отделениях реанимации и интенсивной терапии. Определенные генетические клоны стафилококков могут получать эпидемическое распространение, что диктует необходимость углубленной генетической характеристики стафилококков и сравнения изолятов, выделяемых из разных источников и в разных географических регионах. Из-за ограничений типизирующей способности большинства генетических методов они часто не позволяют различить близкородственные штаммы и выявить существующее различие/сходство, в то время как полногеномное секвенирование обеспечивает получение полной генетической характеристики микроорганизма и выявление различий даже между близкородственными штаммами [2].

Высокопроизводительное секвенирование нового поколения активно используется в клинической практике в США, Нидерландах, Германии, Великобритании и Дании. Оно имеет важное значение в эпидемиологических исследованиях — выявлении источников инфекции, путей распространения возбудителей [3]. Высокопроизводительное секвенирование нового поколения позволяет выявлять у стафилококков новые генетические островки, несущие гены вирулентности — продукции токсинов и резистентности, более полно характеризовать патогеном [4].

В Российской Федерации впервые геном энтеротоксигенного *S. aureus* был секвенирован в 2014 г. [5]. Этот штамм вызвал массовую вспышку пищевой токсикоинфекции в 2013 г., в дальнейшем его использовали в качестве типового штамма при анализе эпидемических вспышек на территории России.

В Великобритании высокопроизводительное секвенирование нового поколения [6] *Staphylococcus spp.* используют для обнаружения путей заражения стафилококком пациентов в реанимации. Во Франции [7] с помощью полногеномного секвенирования были определены четыре линии *S. aureus*, причастные к вспышкам и спорадическим инфекциям в больницах Франции, и обнаружены их эволюционное происхождение.

Таким образом, существует необходимость в оптимизации технологии полногеномного секвенирования *Staphylococcus spp.* и внедрении ее в клиническую практику.

Цель работы — проведение первого полногеномного секвенирования *Staphylococcus aureus* в Республике Беларусь.

Материалы и методы. Был исследован штамм *Staphylococcus aureus BLR-DV*, выделенный из мочи 73-летнего пациента с урологическим сепсисом. Изолят был устойчив к бета-лактамам антибиотикам (амоксциллину, ампициллину, ампициллину-клавулановой кислоте, цефепиму, цефиксиму, цефотаксиму, цефокситину, цефтазидиму, эртапенему, имипенему, меропенему, пиперациллину), аминогликозидам (амикацину, гентамицину, тобрамицину), фторхинолонам, эритромицину, триметоприм-сульфаметоксазолу. Определение устойчивости проводили диско-диффузионным методом согласно стандарту Европейского комитета по определению чувствительности к антимикробным препаратам (European Committee on Antimicrobial Susceptibility Testing, EUCAST) [8].

С использованием описанного ранее метода мульти-праймерной ПЦР [9] установлено наличие *mecA* гена у данного изолята, что позволяет его отнести к варианту, резистентному к метициллину.

Выделение ДНК. Культуру *S. aureus BLR-DV*, выросшую на желточно-солевом агаре, суспендировали в 1xTAE буфере или суспендирующем буфере. Выделение ДНК проводили набором QIAGEN Blood&Cell Culture DNA Maxi Kit (CatNo./ID: 13362), который содержал кремниевые колонки, лизирующий раствор, протеиназу K, суспендирующие и отмывающие буферы. Оценку качества выделенной ДНК проводили спектрометрией и флуориметрией. В дальнейшие исследования включали ДНК с соотношением адсорбции на длинах волн 260 и 280 и 260 и 230 (A_{260}/A_{280} , A_{260}/A_{230}), в пределах 1,8–2,0 и 2,0–2,2. Определяли концентрацию ДНК и разводили исследуемую ДНК до концентрации — 1 нг с помощью 10мМ Tris-HCl (pH 7,5–7,8), либо воды.

Высокопроизводительное секвенирование нового поколения. Высокопроизводительное секвенирование генома *S. aureus BLR-DV* выполняли с использованием технологии MiSeq, Illumina и MiniOn, Nanopore. Все этапы подготовки проводили в соответствии с инструкцией фирмы-производителя с использованием наборов MiSeq, Illumina и MiniOn, Nanopore. Сборка осуществлена 25 июня 2020 г. с использованием программ SPAdes v. 3.14; A5-miseq v. 20160825; Flyev. 2.7b; Canu v. 1.8. Покрытие генома — 400.0x. Использовали приборы для полногеномного секвенирования: Illumina MiSeq; Oxford Nanopore MiniION. Для аннотирования использован подход Best-placed reference protein, программа GeneMarkS-2+, версия 4.11. и SeedViewer, версия 2.2.4. В ходе исследования использовали программу анализа геномов, фрагментов UniproUGEN v. 34

Результаты и их обсуждение. Изолят *S. aureus BLR-DV* с помощью ПЦР был идентифицирован как относящийся к MRSA в связи с присутствием гена *mecA*. Согласно результатам анализа полной нуклеотидной последовательности ДНК *S. aureus BLR-DV* было установлено, что геном штамма является уникальным и представлен одной кольцевой хромосомой размером 2,969,706 пар нуклеотидов (рисунок 1), а также данный изолят содержал две плазмиды — pSTA-BLR-MRSA-DV-1 и pSTA-BLR-MRSA-DV-2 размером 2979 п. о. и 2348 п. о. соответственно.

Нуклеотидные последовательности секвенированного генома и плазмид загружены в генетический банк GenBank, код доступа CP049108.1 [10], CP058313.1 [11], CP058314.1[12] соответственно. Содержание G+C составляет 32,9 %. В функциональной организации генома установлено наличие 278 субсистем, которые кодируют 2839 белков, 77 молекул РНК. На рисунке приведена кольцевая молекула ДНК секвенированного *S. aureus BLR-DV*.



Рисунок — Полный геном секвенированного *S. aureus BLR-DV*

Геном *S. aureus* имеет невысокое содержание G+C, содержание аденина и тимина в два раза больше, чем гуанина и цитозина, при этом в геноме часто встречаются динуклеотиды AA, TT, TA, AT — 12,36, 12,21, 9,53, 11,21 % соответственно, в то время как удельный вес других динуклеотидов варьировал в пределах — 2,56 — 6,48 % (таблица 1).

Таблица 1 — Количество динуклеотидов, встречающихся в геноме *S. aureus BLR-DV*

№	Нуклеотиды	Повторы	Процент	№	Нуклеотиды	Повторы	Процент
1	A	1 000 016	33,7	1	CG	76 016	2,56
2	C	489 320	16,5	2	CT	144 058	4,85
3	G	487 059	16,4	3	GA	157 468	5,3
4	T	993 311	33,4	4	GC	99 998	3,37
5	AA	366 955	12,36	5	GG	75 821	2,55
6	AC	156 356	5,27	6	GT	153 772	5,18
7	AG	143 794	4,84	7	TA	283 074	9,53
8	AT	332 911	11,21	8	TC	156 238	5,26
9	CA	192 518	6,48	9	TG	191 428	6,45
10	CC	76 726	2,58	10	TT	362 570	12,21

В геноме секвенированного *S. aureus* BLR-DV найдены детерминанты устойчивости к противомикробным лекарственным средствам. В положении 46 909–47 817 выявлена кассета — ant(6)-Ia, обеспечивающая устойчивость к стрептомицину, в положении 48 437–49 231 обнаружена кассета arh(3')-III_1, обеспечивающая устойчивость к амикацину. В положении 82 423–83 802 определено присутствие гена tet(K), детерминирующего устойчивость к доксициклину и тетрациклину, а также в положении 502 261–504 180 присутствует ген tet(M) размером 1920 нуклеотидов, который также контролирует устойчивость к тетрациклину. В положении 90 130–92 136 у секвенированного изолята присутствует ген mecA, обеспечивающий устойчивость ко всем бета-лактамам антибиотикам (исключая цефалоспорины 5 поколения). Также в геноме стафилококка присутствовали гены aac(6')-arh(2''), обеспечивающие устойчивость к амикацину, гентамицину, тобрамицину, а также ген blaZ (размером 888 нуклеотидов), связанный с устойчивостью к амоксициллину, ампициллину, пенициллину, пиперациллину (таблица 2). У изолята найдены детерминанты устойчивости к фторхинолонам, эритромицину, триметоприму-сульфаметоксазолу.

Функциональные группы генов, которые присутствуют в геноме *S. aureus*, и их количество приведены в таблице 2.

Таблица 2 — Функциональные группы генов в геноме *S. aureus*

Группа	Количество
Кофакторы, витамины, протезы, пигменты	185
Клеточная стена и капсула	105
Вирулентность, болезни и защита	92
Метаболизм калия — без подкатегории	9
Фаги, профаги, переносимые элементы, плазмиды	38
Мембранный транспорт	73
Приобретение железа и обмен веществ	41
Метаболизм рнк	118
Нуклозиды и нуклеотиды	101
Метаболизм белков	207
Деление клетки и клеточный цикл	40
Подвижность и хемотакси	1
Регуляция и клеточная сигнализация	48
Вторичный метаболизм	6
Метаболизм ДНК	92
Жирные кислоты, липиды, изопреноиды	100
Метаболизм азота	26
Покоящиеся формы	11
Дыхание	34
Ответ на стресс	73
Метаболизм ароматических соединений	5
Аминокислоты и производные	322
Метаболизм серы	16
Метаболизм фосфора	23
Углеводы	251
Всего	2017

Заключение. В ходе исследований отработана технология проведения полногеномного секвенирования *S. aureus* BLR-DV и алгоритм сборки и анализа генома, проведен подбор биоинформационных ресурсов, которые позволяют получать клинически значимую информацию (выявление маркеров устойчивости к противомикробным лекарственным средствам). Впервые секвенированный в стране геном штамма *S. aureus* BLR-DV имел размер 2,969,706 пар нуклеотидов, данный изолят содержал две плазмиды — pSTA-BLR-MRSA-DV-1 и pSTA-BLR-MRSA-DV-2 размером 2979 п. о. и 2348 п. о. Нуклеотидные последовательности секвенированного генома и плазмид загружены в генетический банк GenBank, код доступа CP049108.1 [10], CP058313.1 [11], CP058314.1 [12] соответственно.

Литература

1. Whole genome sequencing of methicillin-resistant *Staphylococcus aureus* / M. Kuroda [et al.] // *The Lancet*. — 2001. — Т. 357, № 9264. — С. 1225–1240.
2. Whole-genome sequencing of bacterial pathogens: the future of nosocomial outbreak analysis / S. Quainoo [et al.] // *Clinical microbiology reviews*. — 2017. — Т. 30, № 4. — С. 1015–1063.
3. Whole-genome sequencing for high resolution investigation of methicillin-resistant *Staphylococcus aureus* epidemiology and genome plasticity / D. J. SenGupta [et al.] // *J. Of clinical Microbiology*. — 2014. — Т. 52, № 8. — С. 2787–2796.
4. Whole genome sequencing of methicillin resistant *Staphylococcus aureus* / M. Kuroda [et al.] // *The Lancet*. — 2001. — Т. 357, № 9264. — С. 1225–1240.
5. Молекулярно-генетическая идентификация штамма *Staphylococcus aureus* — возбудителя пищевой токсикоинфекции при вспышке в Санкт-Петербурге в 2013 г. / Г. Г. Онищенко [и др.] // *Вестник РАМН*. — 2014. — № 9–10.
6. Whole-genome sequencing shows that patient-to-patient transmission rarely accounts for acquisition of *Staphylococcus aureus* in an intensive care unit / J. R. Price [et al.] // *Clinical infectious diseases*. — 2014. — Т. 58, № 5. — С. 609–618.
7. Routine whole-genome sequencing for outbreak investigations of *Staphylococcus aureus* in a national reference center / G. Durand [et al.] // *Frontiers in microbiology*. — 2018. — Т. 9. — С. 511.
8. The European Committee on Antimicrobial Susceptibility Testing. Breakpoint tables for interpretation of MICs and zone diameters. Version 8.1, valid from 2018-05-15. — Mode of access: http://www.eucast.org/fileadmin/src/media/PDFs/EUCAST_files/Breakpoint_tables/v_8.1_Breakpoint_Tables.pdf. — Date of access: 06.08.2018.
9. Генетическая идентификация устойчивых к метициллину *Staphylococcus spp.* / В. В. СлиZENь [и др.] // *Здравоохранение Беларуси*. — 2009. — № 10. — С. 46–48.
10. Accession to genome CP058312.1. — Mode of access: <https://www.ncbi.nlm.nih.gov/nucleotide/CP058312.1>. — Date of access: 09.11.2020.
11. Accession to *Staphylococcus aureus* strain BLR-DV plasmid pSTA-BLR-MRSA-DV-1, complete sequence CP058313.1. — Mode of access: <https://www.ncbi.nlm.nih.gov/nucleotide/CP058313.1>. — Date of access: 09.11.2020.
12. Accession to *Staphylococcus aureus* strain BLR-DV plasmid pSTA-BLR-MRSA-DV-1, complete sequence CP058314.1. — Mode of access: <https://www.ncbi.nlm.nih.gov/nucleotide/CP058314.1>. — Date of access: 09.11.2020.

New generation high performance sequencing method for study of *staphylococcus spp.* Full genome

Slizen V.V., Goretsky D., Gudkova E. I.

Belarusian State Medical University, Minsk, Republic of Belarus

The article represents the results of first whole genome sequencing of *Staphylococcus aureus* in Belarus. The *Staphylococcus aureus* BLR-DV strain isolated from the urine of a 73 year old patient with urological sepsis was studied. High performance genome sequencing was performed using MiSeq, Illumina and MiniOn, Nanopore technology. The assembly was carried out on June 25, 2020 by softwares SPAdes v. 3.14; A5-miseq v. 20160825; Flye v. 2.7b; Canu v. 1.8. Genome coverage 400.0x. The sequenced genome of the *Staphylococcus aureus* BLR-DV contained 2,969,706 base pairs and was uploaded to the GenBank, NCBI (accession code CP058312.1). This isolate also contained two plasmids — pSTA-BLR-MRSA-DV-1 (GenBank accession code CP058313.1) and pSTA-BLR-MRSA-DV-2 (GenBank access code CP058314.1) with the sizes 2979 bp и 2348 bp correspondingly. The G + C content was 32,9 %. The functional organization of the genome was represented by 278 gene cluster subsystems, which encoded 2839 proteins and 77 RNA molecules. The obtained experimental data on the structure of the *Staphylococcus aureus* BLR-DV whole genomemay have fundamental and clinical application — resistance assessment to antimicrobial drugs, as well as epidemiological surveillance of MRSA outbreaks.

Keywords: complete genome, *Staphylococcus aureus*, complete genome analysis.

Поступила 16.11.2020