

Карагандинский государственный медицинский университет

УДК 579:577.2:616-036.22

На правах рукописи

БАБЕНКО ДМИТРИЙ БОРИСОВИЧ

**Исследование конкордантности и дискриминирующей способности
молекулярных методов, используемых в субвидовом типировании
микроорганизмов**

6D110100 – Медицина

Диссертация на соискание степени
доктора философии (PhD)

Научный руководитель:
д.м.н., профессор,
Азизов Илья Сулейманович

Научный консультант:
MD, доктор PhD
Antonella Chesca

Республика Казахстан
Караганда, 2016

СОДЕРЖАНИЕ

НОРМАТИВНЫЕ ССЫЛКИ.....	4
ОБОЗНАЧЕНИЯ И СОКРАЩЕНИЯ	5
1 СТАНОВЛЕНИЕ МОЛЕКУЛЯРНОЙ ЭПИДЕМИОЛОГИИ КАК НАУКИ.....	11
1.2 Приложения молекулярной эпидемиологии	15
1.3 Область применения методов типирования	17
1.4 Критерии оценки методов внутривидового типирования	19
1.5 Обзор методов типирования	21
1.5.1 Фенотипические методы типирования	21
1.5.2 Молекулярно-генетические методы типирования.....	23
2 МАТЕРИАЛЫ И МЕТОДЫ	30
2.1 Методы бактериологического исследования	30
2.2 Молекулярно-генетические методы типирования <i>S.aureus</i>	30
2.3 Spa, MLST, MVLST, MLVA, PFGE, cgMLST, wgMLST и panGenome типирование на основе полногеномных данных <i>S.aureus</i> : <i>in silico</i> ис- следование.....	33
2.4 Кластеризация результатов типирования.....	38
2.5 Оценка дискриминирующей мощности методов внутривидового типирования и их согласованность (конкордантность).....	39
3 РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ	40
3.1 Характеристика клинических изолятов <i>S.aureus</i> . Определение чувствительности к антибиотикам	40
3.2 Spa типирование клинических штаммов <i>S.aureus</i>	43
3.3 MVLVA типирование клинических штаммов <i>S.aureus</i>	47
3.4 Результаты сравнения данных Spa и MLVA типирований.....	49
3.5 Характеристика полногеномных данных <i>S.aureus</i> , включенных в <i>in silico</i> исследование.....	51
3.6 <i>In silico</i> анализ на определение PFGE, MLVA, MLST, MVLST, spa, cgMLST, wgMLST, PanGenome типов используя полногеномные данные <i>S.aureus</i> . Оценка дискриминирующей способности методов типирования и конкордантности между полученными данными.....	51
3.7 Дискриминирующая способность и конкордантность методов типирования	68
3.8 Параметры, влияющие на дискриминирующую способность метода типирования	70
3.9 Параметры, влияющие на показатель конкордантности между данными, полученными различными методами типирования.....	81
3.10 Оценка возможностей генотипических методов на клинических штаммах <i>S.aureus</i> с фенотипическими изменениями за 2 – 7-месячный период.....	87
ЗАКЛЮЧЕНИЕ	93
ПРАКТИЧЕСКИЕ РЕКОМЕНДАЦИИ.....	95

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ	96
ПРИЛОЖЕНИЕ А.....	114
ПРИЛОЖЕНИЕ Б.....	115

НОРМАТИВНЫЕ ССЫЛКИ

Руководство Института Клинических и Лабораторных Стандартов: «Clinical and Laboratory Standards Institute: Performance standards for antimicrobial susceptibility testing. Twentieth informational supplement document. 2010; 30 (1):108–14».

Заключение этической комиссии КГМУ № 13 от 11.11.2013 года;

ГОСТ 7.32-2001 (изменения от 2006 г.). Отчет о научно-исследовательской работе. Структура и правила оформления;

ГОСТ 7.1-2003. Библиографическая запись. Библиографическое описание. Общие требования и правила составления.

ОБОЗНАЧЕНИЯ И СОКРАЩЕНИЯ

CDS – Coding DNA Sequences
CLSI – Clinical & Laboratory Standards Institute
MALDI-TOF – Matrix Assisted Laser Desorption/Ionization Time of Flight
MIST – Microbial In Silico Typer (software)
MLEE – Multilocus enzyme electrophoresis
MLST – MultiLocus Sequence Typing
MLVA – Multiple-Locus Variable number tandem repeat Analysis
MRSA – Methicillin-resistant Staphylococcus aureus
MST – Minimum Spanning Tree
MVLST – Multi-Virulence-Locus Sequence Typing
PFGE – Pulsed Field Gel Electrophoresis
RFLP – Restriction Fragment Length Polymorphism
SNP – Single-Nucleotide Polymorphism
ST – Sequence Type
UPGMA – Unweighted Pair-Group Method Using Arithmetic Averages
VNTR – Variable Number Tandem Repeat
WGS – Whole Genome Sequencing
ДНК – Дезоксирибонуклеиновая кислота
МИК – Минимальная Ингибирующая Концентрация
МЛСТ – Мультилокусное Секвенирование Типирование
МЛЭФ – Мультилокусный Электрофорез Ферментов
п.о. – Пар Оснований (длина нуклеотидной последовательности)
ПГЭ – Пульс-Гель Электрофорез
ПДРФ – Полиморфизм Длин Рестрикционных Фрагментов
ПЦР – Полимеразная Цепная Реакция

ВВЕДЕНИЕ

Актуальность проблемы

Основным предназначением внутривидового типирования микроорганизмов является оценка взаимосвязи между изолятами микроорганизмов определенного вида. Основными задачами эпидемиологических исследований являются: установление источника инфекционных заболеваний, определение путей и механизмов их распространения, подтверждение или исключение эпидемиологического характера вспышки, а также оценка эффективности мер эпидемиологического контроля. Именно субвидовое типирование позволяет в большей степени разрешить многие вопросы, связанные с эпидемиологией инфекционных патогенов, благодаря возможности проведения различий эпидемиологически связанных изолятов от несвязанных с эпидемическим процессом, на основании установления клонального характера распространения патогенов [1, 2]. Благодаря лабораторным методам внутривидового типирования микроорганизмов появилась возможность количественной оценки эффективности систем наблюдения и контроля за инфекционными заболеваниями, обеспечивая проведение анализа в разработке стратегии управления в области общественного здравоохранения.

При всем разнообразии методов типирования патогенов, в основе которых лежат различные принципы, отличающиеся как по времени выполнения исследования и трудозатратности, надежности, воспроизводимости и цене, так и способностью к дифференцировке изолятов [3], все методы могут быть подразделены на фенотипические и генотипические. В связи с наличием достаточно большого количества недостатков в фенотипических методах исследования, таких как трудоемкость, время выполнения исследования, вариабельность результатов, относительно невысокая разрешающая способность [4, 5], генетические методы типирования в настоящее время заняли доминирующую позицию в решении различных задач микробиологии и эпидемиологии [6].

Выбор подходящего метода молекулярного типирования существенно зависит от решения проблемы и эпидемиологической ситуации, в которой выбранный метод будет использоваться, а также времени и географического масштаба его использования. Следует отметить, что представители определенного вида микроорганизмов, вызывающие заболевания у человека, представляют собой достаточно гетерогенное сообщество, поэтому важно, чтобы метод обладал способностью типирования всех микроорганизмов данного вида [7]. Одним из ключевых параметров при выборе метода типирования с целью изучения вспышек инфекций является дискриминирующая мощность, позволяющая проводить необходимое различие эпидемиологически связанных изолятов от штаммов, неассоциированных со вспышкой. В идеальном случае, метод субвидового типирования должен различать близкородственные штаммы, позволяя выявить

пути и механизмы передачи микроорганизмов в восприимчивой популяции, что имеет ключевое значение при разработке стратегии по предотвращению дальнейшего диссеминирования инфекции.

Таким образом, на фоне большого разнообразия молекулярно-генетических методов типирования микроорганизмов, а также накопленных данных, полученных в результате типирования, представляет определенный интерес провести сравнительный анализ современных методов внутривидового типирования.

Цель исследования

Сравнительная оценка конкордантности и дискриминирующей способности современных молекулярно-генетических методов субвидового типирования микроорганизмов.

Задачи исследования

1. Сформировать представительную коллекцию клинических штаммов *S.aureus*, аннотировать и определить чувствительности к антимикробным препаратам, а также получить полногеномные данные *S.aureus*, включая информацию о месте и времени изоляции, используя NCBI GenBank, NCBI PubMed, Pathosystems Resource Integration Center – PATRIC базы данных.

2. Провести внутривидовое типирование клинических штаммов *S.aureus* методами MLVA и Spa типирования, а также *in silico* типирование методами MLST, spa, MVLST, MLVA, PFGE, cgMLST, wgMLST и panGenome, используя полногеномные данные *S.aureus*.

3. Оценить дискриминирующую мощность методов субвидового типирования *S.aureus* методом Симпсона.

4. Оценить конкордантность (согласованность) между результатами, полученными различными методами типирования на уровне внутривидовых типов и близкородственных групп/комплексов/кластеров методом Ранда.

5. Изучить параметры, влияющие на показатели конкордантности и дискриминирующей способности методов внутривидового типирования микроорганизмов.

Научная новизна

- Впервые в Казахстане проведено внутривидовое типирование коллекции клинических штаммов *S.aureus* методами MLVA и Spa, с целью определения разрешающей мощности каждого метода и уровня согласованности полученных результатов.

- Впервые получена сравнительная оценка 7 современных молекулярно-генетических методов типирования, включая подходы на основе полногеномного секвенирования, таких как cgMLST, wgMLST и panGenome, на большой коллекции геномов (n = 4976), что позволило ранжировать методы внутривидового типирования микроорганизмов по дискриминирующей способности, а также оценить показатели согласованности результатов на уровне внутривидовых типов и близкородственных групп/комплексов/кластеров.

- Проведена комплексная оценка факторов, оказывающих влияние на дискриминирующую способность метода типирования и конкордантность результатов, полученных разными методами типирования.

Практическая значимость

Проведенные в диссертационной работе исследования расширяют представления о преимуществах и недостатках каждого метода внутривидового типирования микроорганизмов, позволяя провести адекватный выбор метода в зависимости от типа задач практического здравоохранения, включая расследование как локальных вспышек, так и глобальных эпидемий.

Внедрения результатов в практику

По результатам диссертационного исследования получено 1 свидетельство о государственной регистрации прав на объект авторского права (Приложение А). Методы расчета оценки конкордантности и дискриминирующей способности методов типирования микроорганизмов, используемые в диссертационной работе, внедрены в учебный процесс кафедры микробиологии и кафедры молекулярной биологии и медицинской генетики. Метод кластеризации Spa типов в близкородственные внутривидовые группы стафилококков внедрен в лаборатории коллективного пользования КГМУ (Приложение Б).

Основные положения диссертации, выносимые на защиту

1. Уровень резистентности *S.aureus*, выделенных в Центрально-Казахстанском регионе к макролидам (азитромицин) составил 25%. Уровень устойчивости к цефалоспорином (цефазолин, цефакситин) отмечался у 18% изолятов стафилококков. Резистентность к оксацилину наблюдалось на уровне 18%, обусловленная метициллин-устойчивыми штаммами стафилококка (MRSA), подтвержденная исследованием на наличие *mecA* гена.

2. Кластерный анализ результатов Spa и MLVA типирований изолятов *S.aureus* установил доминирование двух клональных групп/комплексов, структура которых указывает региональную циркуляцию представителей определенных внутривидовых групп. Обнаруженные t002, t030, t008 Spa типы указывают на принадлежность к клональным группам/комплексам *S.aureus* с глобальным эпидемическим распространением.

3. Дискриминирующая мощность MLVA и Spa методов, полученная при типировании коллекции клинических штаммов *S.aureus* (n = 98) была на уровне ~ 90-92%. Конкордантность результатов двух методов типирования составила 20,7% на уровне внутривидовых типов и 86,2% на уровне близкородственных групп/комплексов.

4. *In silico* анализ 4976 полногеномных данных *S.aureus* определил 100% дискриминирующую способность panGenome, wgMLST и cgMLST, 98% MLVA, 92% MVLST, 90% Spa и 85,5% MLST методов типирования. Согласованность результатов методов типирования не превысила 50% на уровне типов, за исключением cgMLST и wgMLST (85%). На уровне близкородственных групп/комплексов/кластеров конкордантность между данными типирования была в диапазоне от 66% до 95%. Исключение

составляет согласованность между MLVA типированием и другими методами субвидового типирования (39-61%).

5. Параметрами, влияющими на дискриминирующую способность схемы типирования, являются: количество локусов в схеме типирования, показатели межлокусной конкордантности и дискриминирующая мощность каждого локуса, входящего в схему типирования. Ведущая роль в определении разрешающей способности принадлежит показателю «количество маркеров» в схеме типирования. Коллекция микроорганизмов, которая подверглась внутривидовому типированию, также оказывает влияние на показатель Симпсона.

6. Все этапы кластерного анализа данных типирования микроорганизмов, включая выбор расчета матрицы генетических дистанций, алгоритм кластеризации с построением дендрограммы и правила формирования групп или значение cut-off влияют на группирование микроорганизмов и как следствие на показатель конкордантности между результатами внутривидового типирования микроорганизмов.

Связь диссертации с другими научно-исследовательскими работами

Диссертация выполнялась на базе лаборатории коллективного пользования НИЦ КГМУ в рамках научно-исследовательского проекта, финансируемого МОН РК «Многоцентровое исследование антибиотикорезистентности метициллин-устойчивых штаммов *Staphylococcus aureus* в Центральном Казахстане» (2012-2014) (№ госрегистрации 0112РК00812).

Апробация диссертации

Основные положения и результаты работы докладывались на Международной конференции молодых ученых «Мир науки и молодежь: достижения и перспективы», Караганда (Казахстан), 26 февраля 2015 года; 16 Международном конгрессе по инфекционным заболеваниям – ISID, Кейпт-Таун (ЮАР), 2-5 апреля, 2014 года; 25 Европейском конгрессе клинической микробиологии и инфекционных заболеваний – ECCMID, Копенгаген (Дания), 25-28 апреля 2015 года; 7 Международном конгрессе азиатско-тихоокеанского общества по инфекционному контролю – APSIC, Тайбэй (Тайвань), 26-29 марта 2015 года; XVII Международном конгрессе по антимикробной терапии – МАКМАХ (IACMAC), Москва (Россия), 20–22 мая 2015 г.; 17 Международном конгрессе по инфекционным заболеваниям – ISID, Хайдарабад (Индия), 2-5 марта 2016 года; Ближне-восточном и южно-азиатском конгрессе по эпигенетике и геномике инфекционных заболеваний – Амман (Иордания), 7-8 марта 2016 года; 26 Европейском конгрессе клинической микробиологии и инфекционных заболеваний – ECCMID, Амстердам (Нидерланды), 9-12 апреля 2016 года; Международной конференции молодых ученых «Мир науки и молодежь: новые пути развития», Караганда (Казахстан), 12 апреля 2016 года; XIX Международной медико-биологической конференции молодых исследователей «Фундаментальная наука и клиническая медицина – человек и его здоровье», Санкт-Петербург (Россия), 26 апреля, 2016; XVIII

Международном конгрессе по антимикробной терапии – МАКМАХ (IASMAC), Москва (Россия), 25-27 мая 2016 г.

Предзащита проходила на совместном заседании научно-экспертной комиссии «Медико-биологических наук», НИЦ КГМУ и кафедр смежных дисциплин КГМУ, протокол № 12 от 17 июня 2016 года.

Публикации

По материалам диссертации опубликовано 42 работ, в том числе главы в 2 сборниках по лабораторной диагностике, опубликованных издательским домом Lambert (2015, 2016 гг.). В изданиях, рекомендованных Комитетом по контролю в сфере образования и науки, опубликовано 5 работ. В изданиях, входящих в базу данных Scopus и/или Thomson Reuters опубликовано 13 публикаций, в том числе 6 статьи и 7 тезисов. Получено свидетельство об интеллектуальной собственности № 156 от 07.04.2016 года.

Структура и объем диссертации

Диссертация изложена на 117 страницах, состоит из введения, основной части, заключения и списка использованных источников, включающих 228 ссылок на работы. В работе содержится 36 таблиц и 48 рисунков и 2 приложения.

1 СТАНОВЛЕНИЕ МОЛЕКУЛЯРНОЙ ЭПИДЕМИОЛОГИИ КАК НАУКИ

Эпидемиология – это наука о распространении этиологически значимых агентов заболеваний в восприимчивой популяции [8, 9]. Уже на ранних этапах эпидемиологических исследований использовали, наряду с классическими эпидемиологическими подходами, молекулярные методы доступные на тот момент, как например, определение титра антител в сыворотке пациентов с лептоспирозом [10] или уровня липидов при сердечно-сосудистых заболеваниях [11]. Лишь в последние десятилетия молекулярная эпидемиология перешла в самостоятельную область исследования.

Впервые упоминания об использовании методов молекулярной эпидемиологии в литературе относятся к 1970-80 годам [12-14]. Это представляло собой комбинацию «сложных» и «продвинутых» лабораторных методов с традиционными эпидемиологическими методами исследования [12, с.480; 14, с.590]. Стало очевидным, что такой подход имеет ряд преимуществ в эпидемиологических исследованиях, что подтверждается все большим числом работ с использованием молекулярно-эпидемиологических подходов [13, 15, 16].

В отношении терминологии, до сих пор в литературе нет единого определения «молекулярная эпидемиология» [17]. Классическим определением считается вариант, предложенный Центром по Контролю за Заболеваниями (Center for Disease Control – CDC, Атланта, 1994), в котором «молекулярная эпидемиология» представляет собой «...использование различных биохимических и молекулярных методов в типировании и субтипировании патогенов», преследующая достаточно простую цель – видовую идентификацию с последующим внутривидовым типированием микроорганизма на основе методов, предложенных биохимией и молекулярной биологией. Эволюция понимания достаточно молодой, но активно развивающейся дисциплины «молекулярная эпидемиология» на различных этапах развития науки представлена в таблице 1.

В определение «молекулярная эпидемиология» включен термин «молекулярный», под которым подразумевается использование методов молекулярной биологии с целью детекции молекул нуклеиновых кислот и аминокислот и их оценки, и «эпидемиология», изучающая возникновение и распространение факторов, ассоциированных с заболеванием в популяции людей. Отличительной особенностью молекулярной эпидемиологии является интеграция показателей биомаркеров с методами эпидемиологического исследования для уточнения диагноза заболевания, установления факторов воздействия, детекция маркеров восприимчивости к заболеванию [27, 28].

Технологические прорывы в молекулярной биологии и генетике внесли значительный вклад в интеграцию лабораторных методов в современные молекулярно-эпидемиологические исследования.

Таблица 1 – Предложенные определения «молекулярная эпидемиология» в различные периоды становления как самостоятельной дисциплины

Определение	Год	Автор	Ссылка
«Применение сложных методов в эпидемиологическом исследовании биологического материала» (С.463)	1977	Higginson J.	[12]
«Молекулярная эпидемиология – наука, в которой используются биологические маркеры или проводят биологические измерения в эпидемиологических исследованиях» (С.13)	1993	Schulte P.A.	[18]
«Применение молекулярной биологии для изучения эпидемиологии инфекционных заболеваний». (С.65)	1994	Tompkins L.S.	[19]
«Использование молекулярных биомаркеров в эпидемиологии» (С.5)	1994	McMichael A.J.	[20]
«Молекулярно-эпидемиологическое исследование предполагает выявление воздействия предполагаемого возбудителя с последующим биологическим эффектом в группе особей, входящих в популяцию» (С. 763)	1995	Groopman J.D., Kensler T.W., Links J.M.	[21]
«Анализ нуклеиновых кислот и белков при изучении здоровья и заболеваний в человеческой популяции» (С.407)	1996	Hall A	[22]
«Молекулярная эпидемиология использует молекулярные подходы для определения заболевания и его доклинических состояний с целью количественной оценки воздействия и раннего биологического эффекта, а также наличие генов восприимчивости» (С.633)	1997	Shpilberg O., Dorman J.S., Ferrell R.E. и др.	[23]
«К практическим задачам молекулярной эпидемиологии относятся: выявление микропаразитов, ответственных за инфекционные заболевания, определение источников, биологические взаимоотношения и пути передачи, а также гены, ответственные за вирулентность, антигенность и резистентность к лекарственным препаратам» (С.806)	1999	Levin B.R., Lipsitch M., Bonhoeffer S.	[24]
«Молекулярная эпидемиология включает в себя эпидемиологические исследования, в которых факторы риска, исходы, влияющие факторы измеряются с помощью биомаркеров» (С.448)	2000	Boffetta P.	[25]
«Молекулярная эпидемиология является осью, которая объединяет знания молекулярного уровня и понимание болезни на популяционном уровне» (С.1)	2011	Schulte P. A., Rothman N., Hainaut P., Smith M. T., Boffetta P., Perera F. P.	[26]

Включение молекулярных инструментов в эпидемиологию инфекционных заболеваний привело к возможности более точно описывать генетическую изменчивость патогенов и оценить взаимосвязь этой изменчивости с распространением инфекции и развитием заболевания. Использование биомаркеров в эпидемиологических исследованиях добавляет уверенность, точность и специфичность в дизайне исследования путем более точного установления как заболевания, так и воздействующего фактора [28].

Несмотря на то, что на начальном этапе молекулярной эпидемиологии использовались иммунологические и биохимические методы с целью дифференцировки патогенных микроорганизмов, современные методы, основанные на достижениях в геномной технологии, позволили более точно и детально оценивать патогенов, а также анализировать пути их передачи в восприимчивой популяции [29]. Молекулярно-генетические подходы обеспечивают способ описания микроорганизмов на внутривидовом уровне путем определения уникальных структур в молекулах, таких как поверхностные белки, или выявления уникальных генетических изменений в ДНК и/или РНК, на основании которых производится установление взаимосвязи изолятов внутри вида [30]. Современные достижения в области генетического анализа позволяют проводить различие между организмами на уровне генов/геномов, идентифицировать измененную структуру или функцию как в патогене, так и в восприимчивом организме, что позволяет идентифицировать лица с наибольшей восприимчивостью к заболеваниям. Текущая практика хранения и обмена данными генетического анализа в виде последовательностей генов/геномов в доступных базах данных, таких как NCBI GenBank [31], European Nucleotide Archive (ENA) [32], DNA Data Bank of Japan (DDBJ) [33] повысило потенциал молекулярной эпидемиологии и предоставило уникальный подход в совместном использовании и описании глобальной эпидемиологии штаммов патогенных микроорганизмов.

1.1 Роль молекулярно-генетических методов в эпидемиологии

На текущий момент существует множество молекулярно-генетических подходов, которые отличаются по принципу метода и молекулярному объекту исследования (ДНК, РНК, белки или метаболиты). Тип информации, которую можно почерпнуть также варьируется. Одни тесты позволяют обнаружить текущее состояние (наличие/отсутствие) молекулярного маркера, другие методы оценивают вариацию либо в пространственной структуре молекулярной мишени, либо в последовательности генетического маркера. Методологические подходы в типировании также различаются по достоверности, надежности, стоимости и возможностью обеспечения высокой пропускной способности.

Принципиальной особенностью в определении видовой принадлежности патогенов на основе молекулярно-генетических методов является отсутствие зависимости от культуральных свойств на питательных средах. Достаточно наличие следовых концентраций нуклеиновых кислот в образце для

установления как родовой, так и видовой характеристики микроорганизма. Это позволяет описать и проанализировать не только наличие или отсутствие патогенов в среде, но и взаимоотношения с другими представителями микроорганизмов. Именно метагеномные исследования, целью которых является изучение всех геномов в конкретной среде или образце, изменили наше представление о возможном обитании микроорганизмов в том или ином биотопе, экотопе, а также их взаимодействия в ассоциации [34, 35].

Задачи эпидемиологических исследований, включая предположения и проверку гипотезы, а также описание и расследование эпидемических вспышек, дополняются и усиливаются молекулярно-генетическими методами исследования. Эпидемиологические исследования по изучению закономерности возникновения и распространений заболеваний в популяции людей с использованием современных лабораторных методов диагностики способны своевременно выявить возникающую угрозу для общественного здравоохранения, установить этиологического агента эпидемического процесса [36, 37], что в свою очередь, позволит разработать эффективные стратегические подходы в профилактике.

Современные молекулярные методы выполняют важную роль на всех этапах эпидемического исследования - будь то место и время возникновения эпидемического события, этиологический фактор или восприимчивое лицо (таблица 2).

Таблица 2 – Роль методов внутривидового типирования в эпидемиологических исследованиях

Место и время эпидемического процесса	<ul style="list-style-type: none"> • Возможность идентификации резервуара инфекции вследствие высокой чувствительности молекулярных методов. • Возможность уточнения локализации эпидемического процесса за счет улучшенной детекции инфицированных лиц, включая лица с субклиническим проявлением инфекции. • Более точное определение времени экспозиции воздействия этиологического агента. • Возможность тестирования образцов после длительного хранения, что позволит определить тенденцию инфекции во времени. • Установление первоисточника.
Этиологический агент	<ul style="list-style-type: none"> • Возможность проследить происхождение микроорганизмов. • Возможность установления происхождения генов, в особенности, ассоциированных с трансмиссией факторов резистентности и вирулентности. • Детекция механизмов передачи генетических факторов патогенности, вирулентности и резистентности.
Восприимчивое лицо	<ul style="list-style-type: none"> • Установление генетических маркеров, связанных с восприимчивостью к инфекции. • Проведение оценки ответной реакции хозяина на внедрение инфекционного агента, на основании определения уровня медиаторов ответной реакции.

1.2 Приложения молекулярной эпидемиологии

Контроль за пищевыми (токсико)инфекциями. Одним из приложений молекулярной эпидемиологии является создание сети лабораторий общественного здравоохранения (PulseNet), координируемой Центром по контролю и профилактике заболеваний, целью которой является контроль за пищевыми (токсико)инфекциями [38]. Основным методом субвидового типирования патогенов является пульс-гель электрофорезом (ПГЭ) на основе которого получают генетические фингерпринты (электрофоретические паттерны) болезнетворных бактерий, выделенных от людей с пищевой (токсико)инфекцией. Все результаты пульс-гель электрофоретического типирования сохраняются в базе данных, доступной для дальнейшего анализа, на основании которых можно группировать патогены в различные кластеры исходя из временных и географических данных, что в свою очередь позволит определить источник инфекции [38, с.22].

Географическое распространение инфекционных заболеваний. Основным вопросом, возникающим на ранних этапах исследования потенциальной вспышки, является наличие возможной эпидемической связи наблюдаемых заболеваний в данном регионе со случаями, возникшими в других географических районах. Определение степени географического распространения эпидемически-значимых клональных групп и степень их родства является важным аспектом молекулярной эпидемиологии. Кроме того, наблюдаемые случаи могут не относиться к эпидемиологическому процессу, являясь отдельными или "спорадическими" случаями. Данный аспект может быть также разрешен с помощью молекулярно-генетических подходов [39, 40].

Айкембаев и др. в своей работе [41] использовали MLVA и SNP типирование штаммов *B.anthraxis* с целью описания генетических профилей, а также выявление доминантных клональных групп и их географическое распространение на территории Казахстана. На основании данных генетического анализа, исследователям удалось установить генеалогическую связь казахстанских штаммов *B.anthraxis* с европейской генетической линией.

Распространение инфекционного процесса во времени. Данные о возникновении и распространении инфекционных заболеваний во времени являются ключевым знанием, на основе которого будет формироваться представление об инфекционных заболеваниях, вызванных тем или иным патогеном в человеческой популяции с точки зрения истории, а также эволюции инфекционных агентов и их изменений в ответ на селективное давление. Данная информация в совокупности с пониманием об уровне полиморфизма(ов) и скорости мутаций является ключевым моментом в понимании механизмов возникновения устойчивости микроорганизма, что является в свою очередь ценной информацией, касающейся разработки стратегии вакцинации и/или разработки новых антибактериальных препаратов.

Группа исследователей под руководством Feikin и др. [42] показала, что за период 1928-2008 гг. в США наблюдается эпидемиологическое распространение пневмококковой инфекции с общей тенденцией на увеличение

заболеваемости. В тоже время, авторами отмечается связь снижения уровня заболеваемости данной инфекцией с введением сульфаниламидных антибиотиков в 1936 году и последующего введения пенициллина [42, с.551].

Механизм передачи инфекций. Методы, используемые молекулярной эпидемиологией, позволяют установить механизм передачи инфекции, а также определить долевое соотношение инфекций, вызванных определенными типами и/или клональными комплексами микроорганизмов. Примером, того как на практике было проведено расследование механизма передачи инфекционного агента, является работа группы исследователей под руководством Холмса по изучению внутрибольничной вспышки, ассоциированной с *Burkholderia cerasia* [43]. Данный патоген известен как этиологический фактор кистозного фиброза инфекционной природы, имеющий, как правило, общего источника внутрибольничных вспышек. Была показана связь возникновения внутрибольничной вспышки среди пациентов без кистозного фиброза, связанная с общим источником. Именно, благодаря методам внутривидового типирования, таким как RFLP и PFGE Холмс и др. смогли определить тип *B. cerasia* и установить путь передачи данного агента между пациентами с кистозным фиброзом и без кистозного фиброза в условиях стационара [43, с.1200].

Нозокомиальные инфекции. Молекулярно-генетические инструменты молекулярной эпидемиологии позволяют проводить расследования внутрибольничных вспышек, определяя источник инфекционного процесса, что имеет ключевое значение при регистрации увеличение количества случаев инфекционных заболеваний. Информация об источнике инфекций позволит определить профилактические мероприятия для предотвращения возникновения дополнительных случаев инфекционных заболеваний. Примером того как была расследована внутрибольничная вспышка, вызванная *Enterobacter cloacae* в Кампинасе (Бразилия), является работа Tresoldi A.T. и др. [44], когда с помощью пульс-гель электрофореза исследователи позволили идентифицировать клональную структуру и установить источник инфекции. Генетический профиль изолятов, выделенных от новорожденных был идентичен с PFGE профилем штаммов микроорганизмов, выделенных из раствора для парентерального питания для новорожденных, определив его в качестве источника инфицирования.

Динамика инфекционного процесса. Молекулярная эпидемиология имеет решающее значение в понимании динамики распространения инфекционных болезней. Благодаря данной информации появляется возможность разработать эффективную стратегию по предупреждению распространения инфекционного процесса. Для осуществления эффективных мер профилактики, необходимо определить механизм передачи, патогенность/вирулентность инфекционного агента, потенциальные восприимчивые лица, резервуары возбудителя. Именно на данной информации будут выстраиваться этапы профилактического контроля инфекционного заболевания. Роль молекулярно-генетических технологий крайне важна в

определении динамики передачи тех или иных внутривидовых групп, которые позволяют более точно и тонко проводить дифференцировку клинических изолятов. Одним из первых и наиболее популярных методов типирования, применяемого в эпидемиологических исследованиях являлся ПДРФ на основе анализа IS6110 [45, 46].

Исследование, проведенное в Южной Корее, где оценивалась популяционная структура *M. tuberculosis* с множественной лекарственной устойчивостью (XDR TB) на основе ПДРФ и сполиготипирования/MIRU-VNTR типирования, указало на высокий уровень кластеризации изолятов [47], позволив предложить гипотезу о приобретении резистентности к лекарственным средствам посредством неадекватного лечения, а не путем распространения устойчивых штаммов, указывая на проблему общественного здравоохранения.

1.3 Область применения методов типирования

Выбор метода типирования зависит от двух основных параметров:

1. популяционная структура микроорганизма;
2. географический и временной масштаб исследуемой популяции микроорганизмов.

Популяционная генетика микроорганизмов. Скорость эволюционных изменений, внутривидовое разнообразие и генетическая структура популяции сильно различаются среди бактериальных патогенов человека [48]. В зависимости от частоты рекомбинации и/или точечных мутаций, популяционная структура микроорганизмов может быть представлена от клональной, в которой образовавшие, путем бинарного деления клетки идентичны исходным родительским клеткам и диверсификация микроорганизмов происходит исключительно путем мутационного генетического «дрифта», до панмиктической (от др.-греч. πᾶν - всё и μίξις - смешение), в которой дочерние клетки генетически отличаются от родительских вследствие высокой частоты случайного межгеномного обмена.

На сегодняшний день большинство бактериальных патогенов имеют преимущественно клональную структуру, характеризующуюся наличием доминирующих клональных комплексов, в которых наблюдается сильное "неравновесное сцепление" (linkage disequilibrium) – неслучайное распределение частот аллелей разных локусов, обусловленное не только тесным генетическим сцеплением генов, но и наличием адаптивного преимущества конкретной комбинации аллелей, частота которой в результате возрастает в сравнении с ожидаемой при случайном распределении [48, с.134 - 52]. Тем не менее, у некоторых видов микроорганизмов с панмиктической популяционной структурой, также может быть обнаружено переходное состояние клональности, когда происходит экологическое изолирование клональных комплексов или вследствие более быстрого распространения близкородственных штаммов, входящих в клональную структуру, во время эпидемического процесса в сравнении со скоростью рекомбинации в пределах

популяции. Например, в случаях с менингококком, который обычно имеет панмиктическую популяционную структуру, могут быть идентифицированы эпидемические клоны, распространяющиеся в географически ограниченных территориях [49]. В свою очередь, может возникнуть ситуация, когда независимый эволюционный процесс среди генетически разнородных изолятов может привести к приобретению общих фенотипических признаков под действием селективного давления. Сложившаяся конвергенция и оценка изолятов на основе типирования по данному признаку может привести к размыванию данных об эпидемическом процессе и к искаженному выводу о клональной структуре.

Понимание генетических основ популяционной структуры и молекулярных механизмов эволюционного процесса бактерий является необходимым условием для выбора соответствующего метода внутривидового типирования. Например, в одних случаях метод типирования может обладать достаточной разрешающей силой в отношении патогенов с панмиктической популяционной структурой, но не сможет провести различие среди изолятов с клональной популяционной структурой.

Масштаб эпидемиологических исследований. Диапазон применения методов субвидового типирования очень широк, начиная от анализа образцов, выделенных от одного источника (хозяина), до анализа коллекции собранной из различных регионов земного шара. В первом случае методы могут использоваться с целью дифференцировки рецидива инфекционного процесса от реинфекции, вызванной новым изолятом того же вида.

Молекулярно-генетические методы типирования при исследовании эпидемической вспышки способны определить:

1. Степень эпидемического распространения инфекционного агента в популяции;
2. Источник инфекции и пути распространения инфекционного агента;
3. Степень влияния профилактических мероприятий, направленных на сдерживание эпидемического процесса, а также скорость распространения инфекционного агента.

Методы субвидового типирования имеют различные уровни применения (рисунок 1). Важно понимать "эпидемиологическую шкалу", как ключевой параметр при выборе подходящего молекулярно-генетического метода.

Например, при исследованиях эволюционного развития микроорганизмов, как правило, используют методы типирования на основе локусов/генов с умеренно консервативной структурой (например, гены домашнего хозяйства, *housekeeping genes*). С одной стороны, такие методы типирования являются своего рода референс методом для других методов типирования, позволяя проводить калибрование других методов типирования по скорости эволюционных изменений. С другой стороны, такие методы, как правило, обладают невысокой разрешающей способностью в отношении коллекции изолятов, собранной за относительно короткий промежуток времени из локальных географических районов.

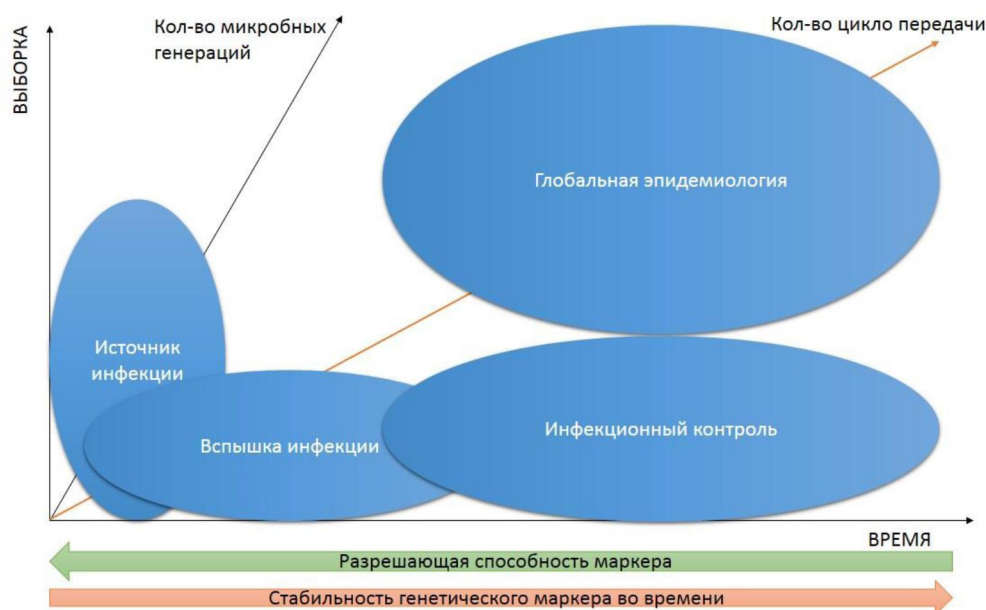


Рисунок 1 – Пространственно-временная шкала различных точек приложений при выборе подходящего метода типирования микроорганизмов

Для исследований локальных вспышек более предпочтительны методы с высоким разрешением. В таких методах используются в качестве мишени гены или участки генома, обладающие более высокой эволюционной скоростью – как правило, это поверхностные белки или участки генома с тандемными повторами. При локальном эпидемиологическом надзоре необходимо найти компромиссное решение между дискриминирующей мощностью и стабильностью генетического маркера, на основе которого будет базироваться метод субвидового типирования.

Эпидемиологические исследования с использованием молекулярно-генетических подходов, в конечном счете, должны обеспечить понимание и описание внутривидовых групп (типов, вариантов) с использованием критериев, на основе которых происходит типирование. Идентичность изолятов является мерой «равенства», определенное конкретным методом типирования. Но можно ли утверждать, что указанные внутривидовые типы на самом деле представляют собой реально идентичные или схожие и связанные штаммы микроорганизмов? В конечном счете, все сравнения тождественности штаммов имеют дело с вероятностью того, что молекулярный кластер/группа/подтип, определенный методом типирования, является лишь инструментом оценки истинной эпидемиологической связи или эволюционного развития, если речь идет о систематике или филогенетике.

1.4 Критерии оценки методов внутривидового типирования

Перед использованием молекулярных методов типирования в эпидемиологических исследованиях следует провести оценку и проверить на соответствие критериям, разработанным ESCMID исследовательской комиссией по генетическим маркерам (ESGEM) [53].

Типируемость – способность метода определить внутривидовой тип для протестированных изолятов, выраженная в процентном соотношении. Это количество детектируемых методом внутривидовых типов изолятов к общему количеству включенных в исследование штаммов микроорганизмов. Другими словами, способность метода определять внутривидовой тип для каждого изолята включенного в исследование.

Стабильность – критерий, относящийся к используемому маркеру и оценивается для каждого вида организмов. Как правило, этот параметр проверяется в лаборатории последовательно, проводя внутривидовое типирование одних и тех же изолятов. В идеале, стабильность следует оценивать и в естественных условиях. Так как скорость рекомбинации и/или мутации различаются в зависимости от вида микроорганизма. Другими словами, это способность распознавать клональное родство между изолятами, несмотря на накапливающиеся генетические различия.

Воспроизводимость – способность метода типирования присваивать один и тот же тип в отношении одного и того же изолята при повторном тестировании. Воспроизводимость может зависеть от многих параметров, включая интерпретацию результатов. Все эти потенциально влияющие параметры должны быть учтены во время оценки результатов. Кроме того, необходимо использовать референс метод для контроля качества каждого прогона тестирования. При тестировании большой коллекции микроорганизмов, межлабораторная воспроизводимость является ключевой. Другими словами, способность обеспечивать идентичными результатами при исследовании изолята в разных лабораториях.

Дискриминирующая мощность – способность метода назначать два разных типа двум эпидемиологически несвязанным штаммам в исследуемой популяции. Этот параметр может быть оценен с помощью индекса разнообразия Симпсона [54], который выражает вероятность назначения двух различных типов для двух штаммов, отобранных случайным образом из большой коллекции эпидемиологически неродственных изолятов [7, с.5]. Определение доверительного интервала позволяет заключить о статистической значимости различия дискриминирующей силы, рассчитанной на параметре Симпсона, для разных методов типирования. Другими словами, это способность обеспечивать различие штаммов, ассоциированных с эпидемиологическим процессом от эпидемиологически несвязанных изолятов на субвидовом уровне. Пригодность метода типирования в эпидемиологических исследованиях не обязательно должна быть связана с высокой дискриминирующей мощностью.

Эпидемиологическая согласованность (конкордантность) – способность метода типирования определять эпидемически связанные изоляты, выделенные при различных вспышках в единый клон или клональный комплекс (группу). Этот параметр должен быть проверен на нескольких хорошо стандартизованных методах. По сути, это способность различать эпидемиологически родственные штаммы, от тех, которые ими не являются, и

результаты молекулярного типирования сверяются и согласуются с данными эпидемиологического исследования. Например, при эпидемической вспышке выделенные штаммы могут использоваться в качестве эталона при сверке результатов типирования других микроорганизмов.

Простота интерпретации полученных данных – способность к однозначной, недвусмысленной интерпретации полученных данных.

Простота в исполнении – способность быть воспроизведенным в не специализированных лабораториях, не требующих сложной подготовки специалистов для проведения исследования.

Высокая пропускная способность – способность обрабатывать большое количество изолятов за определенный промежуток времени и зависит как от простоты процедуры, так и возможности автоматизации выполнения анализов.

Стоимость и доступность – зависит от стоимости и доступности реагентов и оборудования, а также от наличия необходимых навыков для выполнения.

1.5 Обзор методов типирования

1.5.1 Фенотипические методы типирования

Типирование на основе чувствительности к антибактериальным препаратам. Определение чувствительности к антимикробным препаратам осуществляется в клинических микробиологических лабораториях с целью назначения адекватной химиотерапии. Подход, основанный на антибиотикограмме, был также использован для типирования патогенов в качестве маркера клональности при условии достаточной дискриминирующей силы и стабильности во времени. В то же время, генетические детерминанты устойчивости могут располагаться на плазмидах и/или мобильных генетических элементах, что, в свою очередь, может привести к горизонтальному переносу генов резистентности и/или возможной их потере за достаточно короткий промежуток времени [55]. Более того, уровень экспрессии генов резистентности регулируется достаточно сложными внутриклеточными системами, зависящими от многих факторов [56, 57]. Тем не менее существуют работы по успешному применению спектра чувствительности к антибактериальным препаратам в качестве маркера в эпидемиологических исследованиях [58, 59]. Blanc D.S. с соавт. в своей работе показали применение данных чувствительности к антибиотикам, оцененные по зонам задержки роста, с целью типирования метициллин-устойчивых стафилококков, продемонстрировав хорошую корреляцию с данными методом риботипирования [60]. С оценки глобального распространения специфических мультирезистентных клонов были созданы международные базы данных, содержащие данные о чувствительности к антибиотикам, а также клиническое и географическое происхождение изолятов [61].

Серотипирование широко использовалось в качестве эталонного метода при типировании сальмонелл и кишечной палочки [62, 63]. Основными недостатками данного подхода являются вариабельность в показателях

типирования, в стабильности и дискриминирующей мощности вследствие возможной перекрестной реакции, что часто приводило к осложнению в интерпретации полученных результатов. Разрешающая способность серотипирования зависит от видовой принадлежности микроорганизма [64]. Рекомбинационные события приводят к изменению серотипа, что сказывается на воспроизводимости показателей типирования (например, трансформация *S. pneumoniae*, трансдукции *Salmonella*). В целом, конкордантность антигенных вариантов с генетическими маркерами достаточно слабая, даже у видов с клональной популяционной структурой, как например, *Escherichia coli* [65] или *Legionella pneumophila* [66]. Для некоторых схем серотипирования были обнаружены альтернативные генотипические методы, как например для *E. coli* [67] и *C. trachomatis* [68].

Метод **фаготипирования** основан на определении характера литического спектра бактериальных изолятов, подвергающихся определенному набору бактериофагов. Данный метод был разработан для небольшого количества патогенных видов микроорганизмов, являясь технически достаточно сложным и времязатратным. К тому же, данный метод продемонстрировал невысокую стабильность, типизируемость и воспроизводимость, проявляя вариабельность в дискриминирующей способности данного метода. Благодаря этим недостаткам, фаготипирование было заменено молекулярными подходами в субвидовом типировании. На текущем этапе роль бактериофагов рассматривают как трансмиссивный фактор вирулентности, который можно отследить с помощью анализа бактериальной ДНК [69].

Мультилокусный электрофорез ферментов (multilocus enzyme electrophoresis – MLEE) – фенотипический метод субвидового типирования, в основе которого лежит детекция изменений в наборе из 10-35 ферментов домашнего хозяйства, вариация которых детектируется с помощью геля электрофореза [70, 71]. Когда-то МЛЭФ рассматривался в качестве эталонного метода для определения филогенетической структуры бактериальных популяций [71, с.79]. На текущем этапе произошла замена МЛЭФ метода типирования генотипическим эквивалентом в виде мультилокусного секвенирования типирования.

Масс-спектрометрия первоначально использовалась с целью анализа состава сложных смесей химических или органических (макро) молекул. Было показано, что разработанный новый вариант масс-спектрометрии – Матрично-активированная лазерная десорбция/ионизация, (от англ. Matrix Assisted Laser Desorption/Ionization Time of Flight, MALDI-TOF) – может быть применена к целым клеткам, позволяя анализировать органические молекулы, входящие в состав клетки и определять их сложный спектр, которые впоследствии сравниваются с масс-спектрами базы данных. MALDI-TOF как метод быстрой и надежной видовой идентификации микроорганизмов был успешно подтвержден в ряде работ [72-74]. Одним из главных преимуществ метода MALDI-TOF масс-спектрометрии является отсутствие предварительной подготовки микроорганизма для анализа. Достаточно произвести нанесение

бактериальной колонии на пластину, смешав с соответствующей матрицей. Было показано, что условия культивирования на плотных питательных средах, включая состав среды, температурные условия и фаза роста и размножения микроорганизмов, оказывают незначительный эффект в формировании масс-спектров [75, 76]. Обладая достаточной разрешающей способностью, масс-спектрометрия может использоваться и для внутривидового типирования микроорганизмов [77], хотя в отношении некоторых видов микроорганизмов разрешающая способность MALDI-TOF была не высокой [78].

1.5.2 Молекулярно-генетические методы типирования

Полиморфизм длины рестрикционных фрагментов (ПДРФ) основан на фрагментировании хромосомной ДНК с использованием ферментов-рестриктаз в сочетании с гибридизацией с мечеными зондами и относится к первому поколению генотипирования микроорганизмов, используемый в 1980-х годах [79, 80]. Успешное использование этого подхода иллюстрируется на примере IS6110 типирования *M.tuberculosis*, стандартизованного более 15 лет назад и с успехом применяется и по сей день многими исследователями [81-83].

Пульс-гель электрофорез (ПГЭ), разработанный и используемый на протяжении последних двух десятилетий, является логическим продолжением метода, основанного на полиморфизме длин рестрикционных фрагментов. В основе ПГЭ лежит протокол с использованием макрорестрикции и последующим электрофоретическим разделением в пульсирующем электромагнитном поле. На современном этапе рассматривается в качестве «золотого стандарта» при эпидемиологических расследованиях вспышек пищевых и внутрибольничных инфекций бактериальной природы [84].

При использовании ПГЭ бактериальный геном анализируется без использования зондов. Бактериальная хромосома фрагментируется редкощепящими рестриктазами, которые распознают специфические последовательности ДНК длиной 5-8 оснований, формируя 10-30 фрагментов. Критически важным моментом при проведении данной методики является недопущение механического повреждения хромосомной ДНК. В связи с этим, лизис бактериальных клеток и последующая фрагментация геномной ДНК проходит в иммобилизованном состоянии в агарозном геле. Полученные фрагменты ДНК прогоняются в специальном агарозном геле, где происходит их миграция в соответствии с их молекулярными размерами. Такой подход позволяет разделить фрагменты ДНК до 800 кбайт. Точечные мутации, делеции, вставки или приобретение плазмид могут привести к незначительным различиям в электрофоретических профилях, позволяющие включать их в одну генетическую группу (кластер) эпидемиологически связанных штаммов. Эти генетические события, как правило, приводят к 2-3 различиям в паттернах бэндов. На основании этих данных было принято правило по интерпретации результатов типирования методом ПГЭ: разница в профилях для тесно связанных изолятов составляет 2-3 бэнда, для потенциально связанных штаммов – различие в 4-6 фрагментов. Различие в 7 и более фрагментах

свидетельствует о большой генетической гетерогенности и отсутствии эпидемиологической связи [85].

Сильными сторонами пульс-гель электрофореза являются универсальность и высокая разрешающая сила. Тем не менее, этой методике присущ ряд недостатков, таких как продолжительность выполнения метода типирования, порой занимающая 2-4 дня, трудоемкость, необходимое высокое качество и стандартизированное количество ДНК, требующее в целом стандартизацию протоколов для межлабораторной верификации результатов [86]. Несмотря на успешное применение стандартных подходов ПЭГ в типировании пищевых инфекций, процесс стандартизации испытывал трудности в отношении клинических штаммов, таких как метициллин-резистентный золотистый стафилококк (MRSA), связанные с плохой межлабораторной воспроизводимостью [87]. Несмотря на сложности, позже были разработаны, стандартизованы и распространены протоколы исследования MRSA методом PFGE в США и Европе. Была создана база данных и номенклатура по типированию метициллин-устойчивых штаммов методом пульс-гель электрофореза [88].

Полимеразная цепная реакция (ПЦР) – широко распространенный диагностический метод детекции, основанный на амплификации специфической последовательности нуклеиновых кислот. Способность данного метода выявлять специфические участки геномной ДНК патогенов, позволила стать популярным методом диагностики инфекционных заболеваний. Дальнейшее развитие ПЦР подхода и создание новых дизайнов праймеров и зондов позволили применять полимеразную цепную реакцию как метод типирования, в частности PCR-RFLP [89], PCR-ribotyping [90], randomly amplification of polymorphic DNA (RAPD) [91], repetitive extragenic palindromic PCR, (Rep-PCR) [92], enterobacterial repetitive intergenic consensus – PCR (ERIC-PCR) [93], PCR-sequencing [94] и другие в эпидемиологических исследованиях.

ПЦР-полиморфизм длины рестрикционных фрагментов (ПЦР-ПДРФ) является разновидностью метода типирования, основанным на полиморфизме длины рестрикционных фрагментов, в котором рестрикционный анализ проводят на продуктах ПЦР-амплификации. Полученный профиль бэндов после рестрикции и электрофореза используют для сравнения паттернов микроорганизмов данного вида. Важным моментом в данном методе является выбор мишени амплификации – ген должен иметь вариабельную последовательность с консервативными фланкирующими участками, на которых будут отжигаться праймеры. На текущий момент ПЦР-ПДРФ метод применяется в эпидемиологическом исследовании различных родов и видов микроорганизмов [95].

Случайная амплификация полиморфной ДНК (Random Amplification of Polymorphic DNA – RAPD-PCR) является методом типирования микроорганизмов с использованием коротких случайных праймеров, которые гибридизируются с достаточной аффинностью к геномной ДНК-последовательности при низкой температуре отжига. Существуют вариации

RAPD-PCR, такие как ПЦР с произвольными праймерами (arbitrarily primed PCR - AP-PCR), ПЦР-фингерпринт (DNA amplification fingerprinting, DAF), отличающиеся, главным образом, длиной праймеров и условиями термоциклирования. Продукты амплификации разделяются с помощью агарозного гель-электрофореза. Полученные электрофоретические паттерны используются при сравнении бактериальных штаммов. Метод достаточно прост в исполнении и не требует предварительной информации о ДНК последовательности [96]. Отсутствие строгой специфичности праймеров обеспечивает гибкостью данный метод, что позволяет применять его на достаточно широкой группе микроорганизмов [97, 98]. Дополнительными преимуществами RAPD-PCR являются сравнительно небольшое время необходимое на выполнение исследования, а также использование небольшого количества бактериальной ДНК. В то же время наблюдается зависимость дискриминирующей мощности метода от количества праймеров и их последовательности, а также условий амплификации. Также недостатком такого подхода в типировании микроорганизмов является низкая межлабораторная воспроизводимость [99, 100].

Стратегия **ПЦР повторяющихся межгенных консенсусных последовательностей Энтеробактерий (Enterobacterial Repetitive Inter-genic Consensus sequences - ERIC-PCR)** основана на использовании специфических геномных последовательностей с консервативными центральными инвертированными повторами (126 пар оснований), расположенными в межгенных районах бактериальных геномов. Положение этих элементов в геномах энтеробактерий варьируется в зависимости от видовой и внутривидовой принадлежности микроорганизма и успешно используется в качестве генетического маркера [93, с.2294; 101]. ERIC-PCR был особенно полезен в качестве метода типирования полирезистентных штаммов энтеробактерий и быстрой идентификации пандемического клона ST131 кишечной палочки [102].

Мультилокусный анализ tandemных повторов (multilocus variable number of tandem repeat analysis – MLVA). Было обнаружено, что в геномах прокариот широко встречаются генетические элементы с повторяющимися группами нуклеотидных блоков [87]. Вариабельность в повторных элементах образуется за счет ошибочной репликации - «проскальзывание вилок репликаций» (slipped-strand mispairing), что, в результате, приводит к увеличению или укорочению tandemного региона за счет вставок или выпадений повторов [103]. Tandемные повторы различаются как по локализации в геноме, так и по размеру и сложности. Мультилокусный анализ tandemных повторов с вариабельным числом звеньев (MLVA-анализ) основан на определении количества повторов с помощью мультиплексной ПЦР с фланкирующими праймерами с последующим определением размера продуктов ПЦР в агарозном геле или с помощью капиллярного электрофореза высокого разрешения. Для каждого локуса, в зависимости от количества tandemных повторов или молекулярной массы ПЦР продукта, присваивается

цифровой код. Метод типирования, основанный на тандемных повторах представляет особый интерес для генотипирования патогенов с высокой степенью однородности генома, таких как *B.anthraxis* [104] и *M.tuberculosis* [105, 106]. В целом, MLVA типирование продемонстрировало хорошую корреляцию с другими методами генотипирования, а ее техническая простота и быстрота делают данный подход привлекательным. Потенциальным недостатком MLVA типирования является слишком быстрая скорость эволюционных изменений, что может привести к нарушению эпидемиологической конкордантности.

Методы типирования, основанные на секвенировании. Изменчивость бактериальных геномов обеспечивается мутационными и/или рекомбинационными процессами. Измененная первичная структура ДНК может использоваться в методах молекулярного типирования. На сегодняшний день, ПЦР технология позволяет накопить большое количество ДНК матрицы путем амплификации, которая может быть отсекументирована, предварительно очистив ее от ПЦР компонентов.

Основным преимуществом метода типирования на основе первичной последовательности гена(ов) над методами, использующими электрофоретическое разделение, является получение однозначных, стандартизованных и портативных данных используя секвенирование геномных мишеней.

Однолокусное секвенирование типирование (ОЛСТ) – метод типирования, основанный на секвенировании одного гипервариабельного локуса и имеющий преимущество над мультилокусным секвенированием типированием по параметрам трудозатратности, требуемого времени на выполнение исследования и стоимости анализа одного образца. Данный подход используется для типирования *Staphylococcus aureus* на основе анализа гипервариабельной X области *Spa* гена, который содержит переменное число 24 п.н. тандемных повторов. Являясь разновидностью методов, основанных на секвенировании, данный методический подход совмещает в себе такие практические преимущества как быстрота, воспроизводимость, простота использования и портативность [107]. Вариабельность X области *Spa* гена имеет как микровариации, так и макровариации, которые в целом имеют хорошую согласованность с вариациями всего генома. Данное свойство позволяет использовать *Spa* типирование как в локальной, так и глобальной эпидемиологии [108]. Исследования, проведенные на большой коллекции, собранной за относительно большой период времени, показали, что *Spa* типирование обладает высокой разрешающей способностью и хорошей конкордантностью с MLST и PFGE [109, 110]. Сложным моментом является анализ полученных данных *Spa* типирования, и на текущий момент решается с помощью коммерческих программных продуктов, таких как Bionumerics и Ridom *Spa* typer, позволяя проводить анализ в полуавтоматическом режиме, используя данные, полученные с анализирующего прибора (секвенатора).

Мультилокусное секвенирование типирование (МЛСТ) – метод, основанный на детекции изменений в последовательностях фрагментов длиной 400-500 п.о. генов домашнего хозяйства (house-keeping genes). Первоначально данный метод был разработан для эпидемиологического анализа гипервирулентных штаммов менингококков [111]. Каждая последовательность локуса рассматривается в качестве аллели; совокупность аллелей, определенных для каждого гена, включенных в методику МЛСТ, формируют профиль, именуемый сиквенс тип (sequence type - ST). Одной из особенностей МЛСТ является то, что изменения, наблюдающиеся в аллелях, будь то однонуклеотидные замены или рекомбинационные явления, включая горизонтальный перенос генов (horizontal gene transfer - HGT), рассматриваются в качестве одиночных генетических событий.

Генетическое родство, сходство может быть выявлено с использованием различных методов кластеризации:

а) методом невзвешенного попарного среднего (Unweighted Pair-Group Method Using Arithmetic Averages - UPGMA) с использованием матриц расстояний, содержащих попарные различия аллельных профилей;

б) деревьев минимальных дистанций, в которых все элементы соединены таким образом, что суммарная дистанция всех звеньев дерева является минимальной;

в) с использованием eBURST, или более оптимизированной версии goeBURST алгоритма, который позволяет на основании близкородственных сиквенс типов вывести закономерность эволюционного происхождения изолятов с клональной популяционной структурой, позволяющей сформировать клональные комплексы[112].

На данный момент в базах МЛСТ типирования находятся данные для более чем 40 патогенных бактерий и грибов. Такие базы данных курируются специалистами, поддерживая высокое качество данных, формируя основу для стандартизированной номенклатуры МЛСТ-типирования. Некоторыми исследователями МЛСТ типирование рассматривается в качестве "золотого стандарта" для определения внутривидовой структуры популяции многих микроорганизмов [113, 114]. Данный метод нашел применение при проведении эпидемиологических исследований различных масштабов, а также при изучении популяционной структуры и эволюционного развития бактерий [115].

Недостатком мультилокусного секвенирования типирования является относительно невысокая разрешающая способность в сравнении с другими методами, как например PFGE, MLVA, что делает его непригодным в эпидемиологических исследованиях на коллекции бактериальных изолятов, собранных в течении короткого времени и на ограниченной территории, например, при расследовании внутрибольничных вспышек. Кроме того, в силу особенности технологии секвенирования, используемый в данной методике, МЛСТ метод является трудоемким и время затратным, что ограничивает его использование при эпидемиологическом исследовании, где требуется высокая пропускная способность.

Важным моментом при разработке МЛСТ схемы для последующего использования его в эпидемиологических исследованиях является выбор генов и их количество для достижения достаточной разрешающей способности, позволяющей различать штаммы с ранней дивергенцией. Для этой цели необходимо выбрать гены, находящиеся под более сильным селективным давлением, например, гены вирулентности. В данном случае метод типирования, основанный на генах вирулентности с более высокой вариабельностью в сравнении с генами домашнего хозяйства называется мультилокусное секвенирование типирование на основе генов вирулентности (Multi-virulence-locus sequence typing – MVLST) [116].

Однонуклеотидный полиморфизм (ОНП, англ. Single nucleotide polymorphism - SNP) – метод, основанный на детекции изменений нуклеотидов (А, Т, G или С) в последовательности ДНК при сравнении с диким типом, возникающим вследствие случайных мутаций, горизонтального переноса генов или рекомбинационных процессов. Методы генотипирования, основанные на однонуклеотидных полиморфизмах применяются для типирования изолятов, имеющих однородную, генетически консервативную структуру генома (например, *M.tuberculosis*, *E.coli* O157: H7 или *B.anthraxis*). Детекция наиболее часто встречающихся полиморфизмов, ассоциированных с МЛСТ типом, является более простой и экономически эффективной альтернативой к полноценной МЛСТ схеме типирования. Филогенетическая структура, оцененная на SNP, имеет схожую с МЛСТ филогенетикой, но с меньшей дискриминацией. Еще одним ограничением ОНП подхода является невозможность детектировать новые аллельные варианты. Тем не менее SNP генотипирование может применяться для скрининга эпидемического процесса или детекции клональных комплексов с глобальным распространением [117].

Микрочипы (Microarrays). С момента вступления микробиологии в постгеномную эру, типирование микроорганизмов в большей степени основывается на анализе полногеномных данных. На текущий момент секвенирование всего генома является дорогостоящим и, соответственно, ограниченным мероприятием для практических лабораторий. В такой ситуации, платформа анализа генома на чипах является хорошей альтернативой полногеномному исследованию, позволяющей обеспечивать в некоторых случаях исчерпывающей информацией о бактериальном геноме, включая генотипирование. Микрочипы представляют собой твердую подложку с нанесенными на ней ДНК зондами с флуоресцентной меткой, на которой происходит гибридизация мишеней исследуемого образца с получением количественных или качественных сигналов [118]. Технология на чипах может быть использована для широкого спектра применений, включая типирование микроорганизмов [119, 120]. Основным преимуществом данной технологии является способность получить описательную характеристику штамма, включая наличие генов вирулентности и/или резистентности, а также генотип изолятов [121]. Microarray технология успешно применялась в эпидемиологических исследованиях при типировании бактерий различных

видов, включая *S. enterica* [122], *S. aureus* [123], *E. coli* [124], *Campylobacter jejuni* [125] и других. Поскольку наблюдается тенденция к улучшению технологии микрочипов с параллельным снижением стоимости анализа, платформа на чипах становится распространенной не только в исследовательских лабораториях, но и клинических [126].

Полногеномное секвенирование (Whole Genome Sequencing - WGS). Успехи в разработке методов полногеномного секвенирования приводят к более широкому применению этой технологии за счет снижения экономических затрат и времени секвенирования генома. На текущий момент стоимость и время секвенирования типичного бактериального генома составляет порядка \$ 100-400 за 1-5 дней. Полногеномное секвенирование все больше становится конкурентоспособным по сравнению с традиционными диагностическими методами, включая метод культивирования бактерий. Данный метод позволяет определить MLST и SNP данные исходя из полногеномных данных и, по мнению некоторых исследователей, в течение нескольких лет может стать основным молекулярно-эпидемиологическим методом исследования [127]. Тем не менее, комплексный подход, сочетающий «традиционные» методы и полногеномное секвенирование является, на текущий момент, необходимым, особенно, в условиях достаточно высоких затрат на полногеномное секвенирование. При этом секвенирование генома является реальной альтернативой в случаях с некультивируемыми, опасными и медленно растущими бактериями [128-130].

Таким образом, методы внутривидового типирования микроорганизмов играют важную роль в биомедицинских исследованиях, позволяя решать вопросы, связанные с антибиотикорезистентностью, вирулентностью, популяционной структурой, филогенетикой и в расследовании эпидемиологических вспышек. На текущий момент предложено большое количество разнообразных методик внутривидового типирования микроорганизмов, в основе которых лежат различные принципы действия. Для каждого, имеющегося в арсенале молекулярной эпидемиологии, метода, в том числе вновь разработанной методики типирования, научное сообщество дает оценку методу по ряду параметров, с целью понимания насколько успешно и в каких случаях данная разработка может быть использована. Одним из главных параметров является дискриминирующая сила метода типирования, который оценивается с помощью показателя Симпсона.

Использование различных подходов в типировании микроорганизмов различными исследовательскими группами привело к тому, что на данный момент, в публикациях и базах данных накопилось и продолжает пополняться большое количество результатов типирования, что в свою очередь, приводит к необходимости проведения сравнительного анализа методов типирования по их разрешающей силе, а также уровня согласованности (конкордантности) результатов типирования.

2 МАТЕРИАЛЫ И МЕТОДЫ

2.1 Методы бактериологического исследования

За период 2014 г. была собрана коллекция микроорганизмов вида *S.aureus* (n=98) из различных отделений, включая педиатрическое, хирургическое, кожно-венерологическое, пульмонологическое и др. отделений стационаров городов: Караганда, Астана Жезказган и Семей. Материалами исследования являлись образцы крови, гнойные выделения ран, моча, мокрота, мазки из носа, зева, конъюнктивы, бронхоальвеолярный лаваж и другие образцы, ассоциированные с инфекционным процессом.

Микробиологические исследования проводились на базе микробиологической лаборатории НИЦ КГМУ. Выделение чистых культур из клинического материала проводили по общепринятой схеме [131-133]. Первичная идентификация вида микроорганизмов проводилась в соответствии с идентификационными ключами «Определителя бактерий Берджи» [134]. Подтверждение в видовой идентификации штаммов, а также установление вида в сомнительных случаях проводилось на основе время-пролетной масс-спектрометрии с использованием MALDI-TOF Microflex (Bruker, Daltoniks). Идентифицированная культура клинических штаммов *S.aureus* вносилась в базу данных WhoNet [135], а сами штаммы хранились в среде с добавлением глицерина при -80°C. Перед использованием штаммы восстанавливали на мясо-пептоном бульоне в течение 18-24 часов [136].

Анализ на чувствительность к антибиотикам проводился количественным методом – методом серийных разведений с использованием 96 луночных планшетов (двукратное разведение - от 512 мкг/мл до 0,1 мкг/мл). Процесс приготовления раствора антибиотика в бульоне и нанесение микробной взвеси *S.aureus*, предварительно стандартизовав по 0,5 МакФарланду [137] проводился с использованием роботизированной станции Tecan 100. Результаты чувствительности к антибиотикам учитывались спектрофотометрически, дополняя базу данных WhoNet со штаммами. Анализ антибиотикограммы проводился в соответствии с критериями CLSI и экспертными правилами.

2.2 Молекулярно-генетические методы типирования *S.aureus*

Культивирование микроорганизмов *S.aureus* осуществлялось на жидкой питательной среде в течение 18 часов в колбах (500мл.), содержащих 120 мл. питательной среды при 37°C используя термостатируемый шейкер (100 об/мин), доведя оптическую плотность культуры до OD440=0.5. Культура клеток микроорганизмов трижды отмывалась от питательной среды центрифугированием (при t = 4°C). Отмытый осадок бактериальных клеток ресуспендировали в физиологическом растворе до OD440 = 10.

Выделение бактериальной ДНК проводилось, используя коммерческий комплект реагентов для экстракции ДНК экспресс-методом «ЭДЕМ» в соответствии с инструкцией. Оценка качества выделенной ДНК проводилась на основе метода прямой спектрофотометрии [138] с использованием

спектрофотометра Nanophotometr P-class (P-330) (IMPLEN, GMBH, Munich, Germany).

Выделенная ДНК клинических штаммов *S.aureus* использовалась в качестве объекта молекулярно-генетического типирования методами Spa [139] и MLVA на основе Staphylococcal Interspersed Repeat Unit (SIRU) маркеров [140].

Spa типирование стафилококков проводилось в соответствии с методикой, предложенной Frenay H. с соавт. [141], опубликованной на сайте ridom.de. Амплификация гипервариабельного участка Spa гена проводилась с использованием праймеров - spa-1113f 5'-TAAAGACGATCCTTCGGTGAGC-3' и spa-1514r 5'-CAGCAGTAGTGCCGTTTGCTT-3', которые отжигались на фланкирующие участки фрагмента Spa гена. Анализ ПЦР-продуктов проводился с использованием 2% агарозного геля, в котором оценивалась приблизительная длина продуктов амплификации, результат которого оценивался на гель-документирующей системе Gel-Doc XR+ (Bio-Rad).

На следующем этапе осуществлялась процедура ферментативной обработки продуктов амплификации экзонуклеазой I и щелочной фосфатазой креветки (Fermentas, Литва). Секвенирование амплифицированного участка Spa гена проводили с использованием набора BigDye Terminator v3.1 Cycle Sequencing Kit на приборе ABI PRISM 3500 Genetic Analyzer (Applied Biosystems, США) (рисунок 2).

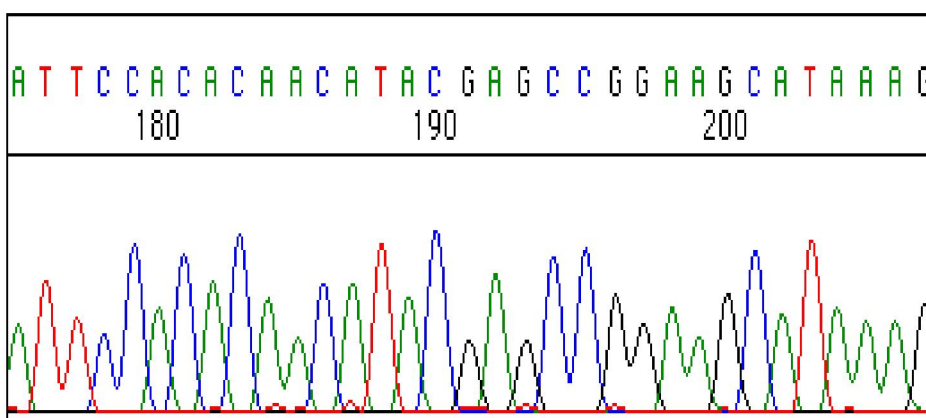


Рисунок 2 – Пример данных, получаемых при секвенировании по Сэнгеру, используя набор BigDye Terminator v3.1 (секвенатор Genetic Analyzer)

Полученная последовательность сохранялась в fasta формате и импортировалась в программу Geneious [142] для анализа полученной последовательности фрагмента Spa гена путем поиска фланкирующих специфических сигнатур и установления качественного и количественного состава повторов и их порядок в последовательности (рисунок 3).

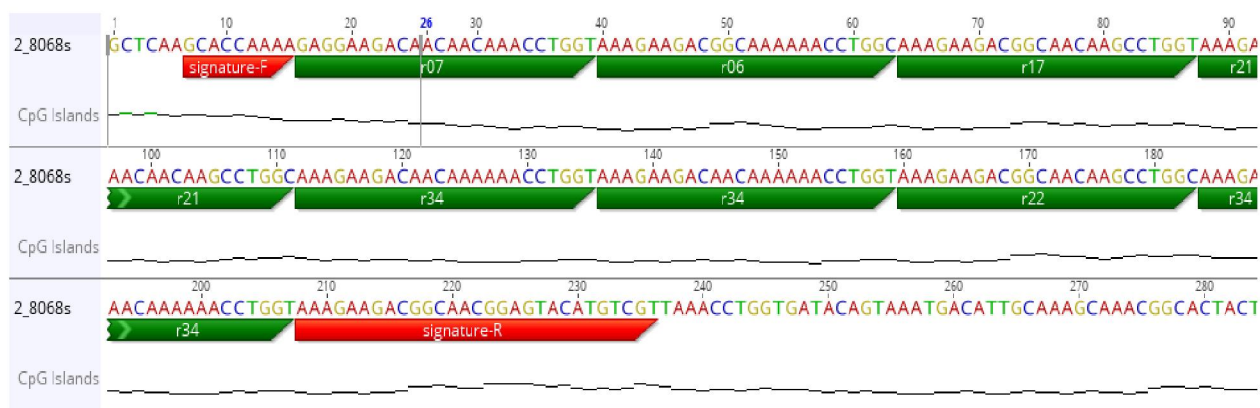


Рисунок 3 – Анализ полученной Spa последовательности штамма *S.aureus*, произведенный в программе Geneious с целью определения количества и положения Spa повторов

Полученный состав Spa повторов сопоставлялся с базой данных spa.ridom.de для определения Spa типа (рисунок 4).

spa.ridom.de/spatypes.shtml	
t161	00-00-01-01-13-17-31
t162	14-44-12-17-17-23-18-17
t163	04-20-17-20-17-45-16-34
t164	07-06-17-21-34-34-22-34
t165	07-23-12-34-46-12-47-12-23
t166	04-44-33-31-12-16-34-16-12-25-22-34

Рисунок 4 – Определение Spa типа исходя из повторов, полученных после секвенирования участка Spa гена

MLVA типирование клинических штаммов *S.aureus* проводилось по 6 локусам - SIRU 01, SIRU 05, SIRU 07, SIRU 13, SIRU 15, SIRU 21, согласно методике, предложенной Ikawaty R. и др. [143] с использованием двух мультиплексных ПЦР реакций, амплифицируя 3 локуса в каждой реакции. Размер ПЦР продуктов определялся с помощью капиллярного электрофореза высокого разрешения с автоматической флуоресцентной детекцией (рисунок 5). Полученные молекулярные массы ампликонов использовались для формирования MLVA профилей, состоящих из 6-значного цифрового кода, соответствующих количеству повторов в каждом локусе.

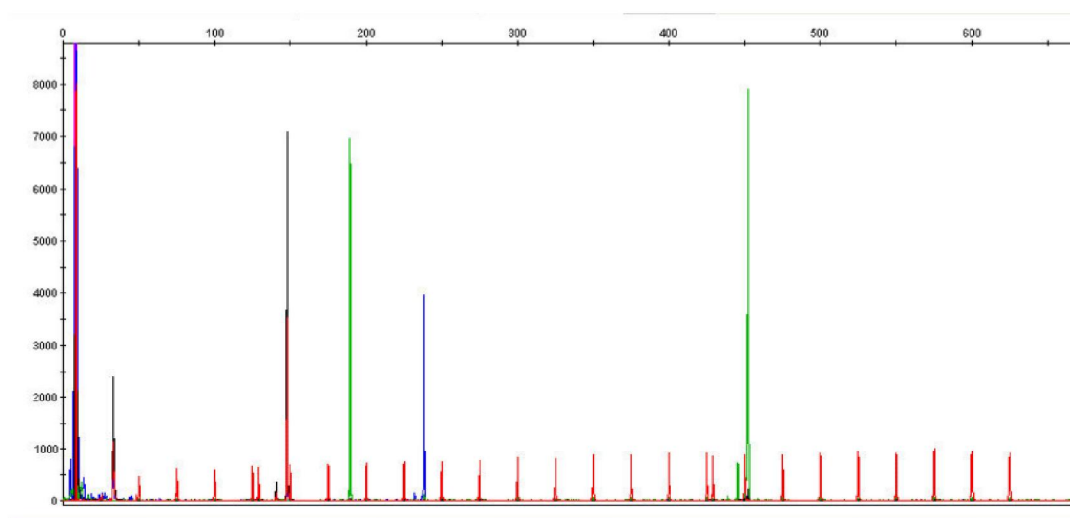


Рисунок 5 - Характер пиков продуктов амплификации с мечеными праймерами для установления молекулярной массы (капиллярный электрофорез высокого разрешения Genetic analyzer 3500)

2.3 Spa, MLST, MVLST, MLVA, PFGE, cgMLST, wgMLST и panGenome типирование на основе полногеномных данных *S.aureus*: *in silico* исследование

Spa типирование на основе полногеномных последовательностей *S.aureus* проводилось с использованием программы SeqSphere+ (Ridom, Germany). Принцип работы данной программы заключается в поиске в геномной последовательности фланкирующих сигнатур F-GCMCCAAAA и R-AAAGAAGAYGGYAACGGRGTAYATGTCGT, между которыми находятся повторы, определяющие Spa тип. Последовательность повторов идентифицируется на основе BLAST выравнивания [144]. На текущий момент существуют более 700 уникальных повторов, формирующих свыше 16000 Spa типов.

MLST типирование, целью которого являлось определение сиквенс типа (ST), проводилось в соответствии со схемой, которая была разработана Mark Enright [145], включающая в себя 7 локусов генов домашнего хозяйства - Carbamate kinase (arc), Shikimate dehydrogenase (aro), Glycerol kinase (glp), Guanylate kinase (gmK), Phosphate acetyltransferase (pta), Triosephosphate isomerase (tpi), Acetyl coenzyme A acetyltransferase (yqi). Поиск генов домашнего хозяйства и установление MLST типа при использовании геномных данных *S.aureus* проводились с использованием Microbial *in silico* typing (MIST) программы, разработанной Peter Kruczkiewicz и др. [146]. Принцип работы данной программы состоит в поиске последовательностей локусов сходных с базой данных аллелей 7 генов на основе BLAST алгоритма. В результате 100% совпадения происходит присвоение аллельного номера для каждого локуса, что, в конечном счете, формирует цифровой профиль MLST сиквенс типа из семи значений. Схема работы программы MIST представлена на рисунке 6.

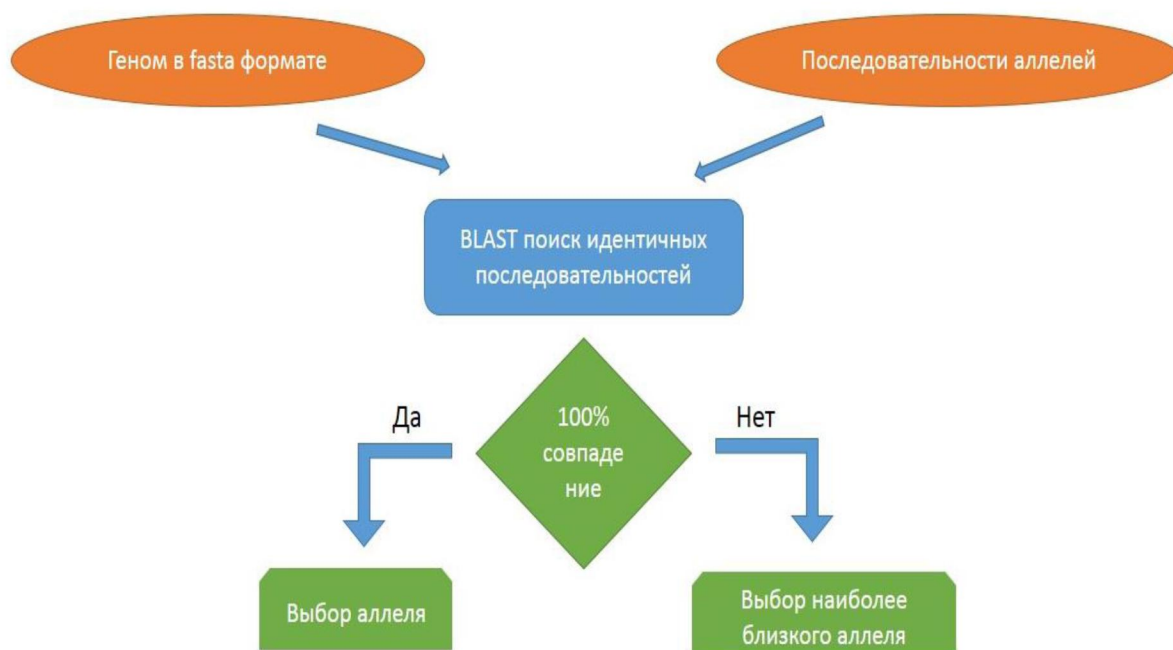


Рисунок 6 – Принцип работы MIST программы для определения MLST типа

Определение MVLST типа проводилось на основе методики, включающей вариабельные локусы 6 генов вирулентности: autolysin (alt), Type VII secretion system protein EssC (essC), lipase (geh), c-hemolysin (hlgA), serine protease (htrA) и Ser-Asp rich protein (sdrC). На основе праймеров, предложенных Bindhu Verghese и соавт. [147] была проведена симуляционная ПЦР с целью получения фрагментов амплификации. Полученные данные ПЦР амплификации использовались для создания BLAST базы данных, по которой, в дальнейшем, производился поиск идентичных или сходных последовательностей фрагментов генов вирулентности в геномах *S.aureus*. Для каждой новой последовательности локуса присваивался новый номер, формируя профиль MVLST типа.

MLVA типирование осуществлялось путем симуляции e-PCR с использованием праймеров, указанных в таблице 1, предложенных Schouls L. M. и соавт. [148], используя primer3 программу [149], включенная в программный пакет Genenious. Пример MLVA анализа, выполненного на геноме *Staphylococcus aureus* strain 2395 USA500 представлен на рисунке 7. Как видно из рисунка, в Geneious была создана база праймеров, которые использовались для симуляции полимеразной цепной реакции. В результате были получены продукты амплификации в размере от 129 до 479 пар оснований, размер которых может быть определен на практике, используя капиллярный электрофорез высокого разрешения (капиллярный секвенатор).

Таблица 3 - Последовательность праймеров, используемых для MLVA типирования *S.aureus*

Наименование праймера	Последовательность праймеров	Длина (п.о.)	% GC	T _m
VNTR09_01Ff	ATAAGCATTGAAACCATTTATGATG	24	29,2	54
VNTR09_01r	GCAACTTCTTAAAACAAAATATTG	24	25	52,2
VNTR21_01r	AGCAATGAATCAATAATTTTCA	22	22,7	50,7
VNTR21_01Vf	GTCGATAAAGCATAAAGCTTT	21	33,3	52,6
VNTR24_01Pf	CAGCAGTAGTGCCGTT	16	56,3	53,6
VNTR24_01r	GTAACGGCTTCATCCA	16	50	50,2
VNTR61_01Nf	AATGCACATGAAACACTAATT	21	28,6	52,3
VNTR61_01r	GGTCAAGAATATTTAAAATCAATT	24	20,8	50,3
VNTR61_02f	CTGTGAAGTTAGATAGATGAGTTT	24	33,3	54,2
VNTR61_02r	GCAATTAACGATTTCTTCAC	20	35	51,2
VNTR63_01Ff	TGAAGATGTAGTAGGAATGTTAGT	24	33,3	54,8
VNTR63_01r	AGAAAAAGCTAAAGAAGTTGAA	22	27,3	52,2
VNTR67_01Pf	CGTGAATCTCTTTTATAAGAGTGT	24	33,3	54,8
VNTR67_01r	CCCTCCTATTAATATATATACCGT	24	33,3	52,3
VNTR81_01Nf	TTTGGATATGAAGCGAGA	18	38,9	50,4
VNTR81_01r	CATATGTTCGCAGTACCATC	19	47,4	53,1

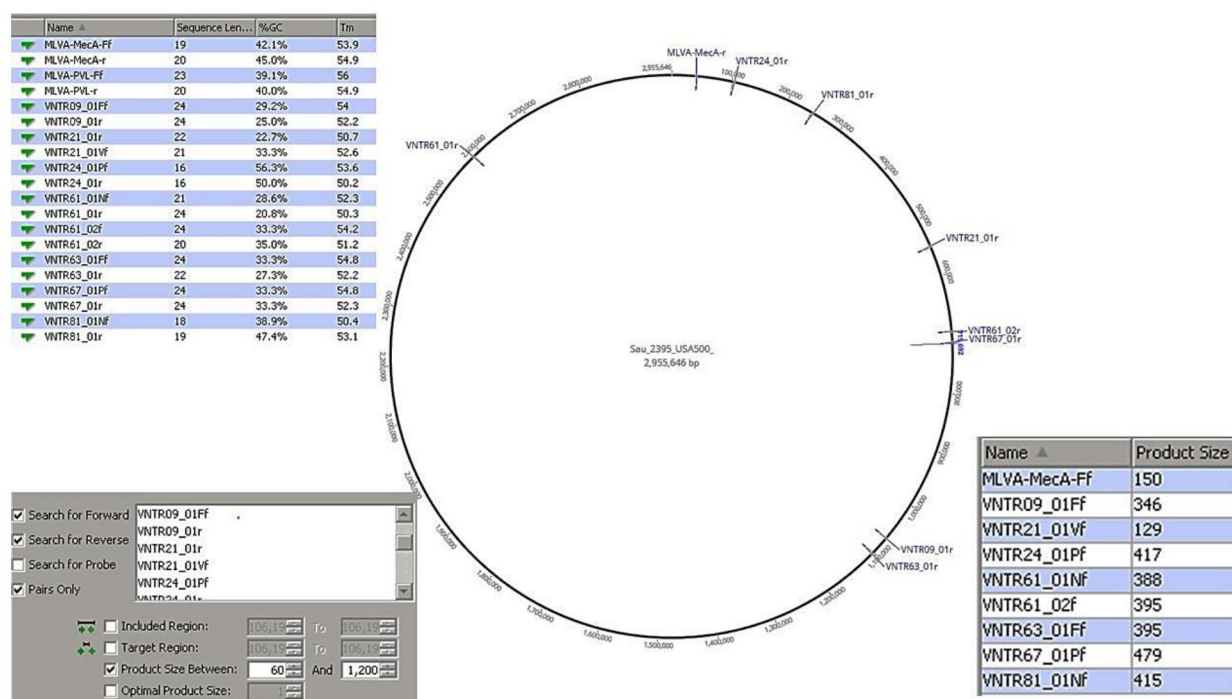


Рисунок 7 – Симуляция ПЦР амплификации с использованием праймеров для MLVA типирования *S.aureus*

На основании полученных размеров продуктов амплификации формировался MLVA профиль, отражающий количество повторов в каждом локусе, путем вычитания фланкирующих участков с последующим делением на длину повтора. Данные о размерах фланкирующих участков и длин повторов для каждого локуса представлены в таблице 4.

Таблица 4 – Показатели, необходимые для вычисления количества повторов в локусе при формировании MLVA профиля

Наименование маркера	Длина повтора	Длина фланкирующих участков
VNTR09	9 п.о.	238 п.о.
VNTR21	21 п.о.	108 п.о.
VNTR24	24 п.о.	153 п.о.
VNTR61_01	61 п.о.	241 п.о.
VNTR61_02	61 п.о.	207 п.о.
VNTR63	63 п.о.	269 п.о.
VNTR67	67 п.о.	145 п.о.
VNTR81	81 п.о.	91 п.о.

Например, после ПЦР амплификации (e-PCR) с использованием генома *S. aureus* 2395 USA500 получились продукты размерами: для VNTR09 – 346 п.о., для VNTR21 – 129 п.о., для VNTR24 – 417 п.о., для VNTR61_01 – 388 п.о., для VNTR61_02 – 395 п.о., для VNTR63 – 395 п.о., для VNTR67 – 479 п.о. и для VNTR81 – 496 п.о. После вычитания соответствующих длин фланкирующих участков, устанавливаем количество tandemных повторов в зависимости от кратности длины каждого повтора, получаем MLVA профиль 12-1-11-2-3-2-5-4.

Симуляция PFGE *in silico* осуществлялась в 2 этапа. На первом этапе геномы *S. aureus* подвергались рестрикционному расщеплению с помощью SmaI фермента с сайтом распознавания в 6 п.о. – CCC[^]GGG [150]. Полученные фрагменты подвергались электрофоретическому разделению с образованием электрофоретического паттерна (количество полос и их расположение на геле). Данная процедура выполнялась на базе программы Geneious. Полученное электрофоретическое изображение анализировалось в TotalLab (Nonlinear Dynamics, UK). Процесс PFGE анализа полногеномных данных *S. aureus in silico* представлен на рисунке 8.

Метод полногеномного мультилокусного секвенирования типирования (cgMLST, wgMLST) является стандартизированной процедурой типирования, разработанной Alexander Mellmann (Inst. Hygiene) и Dag Harmsen (University Munster), которая включает в себя 1861 основных генов (core genes), входящие в wgMLST схему и 706 добавочных генов (accessory genes). В результате типирования были использованы 2 схемы, включающие 2567 генов. Анализ полногеномного MLST типирования на основе геномных последовательностях *S. aureus* проводился с использованием SeqSphere+ программы. Принцип работы cg/wgMLST в программе SeqSphere+ представлен на рисунке 9.

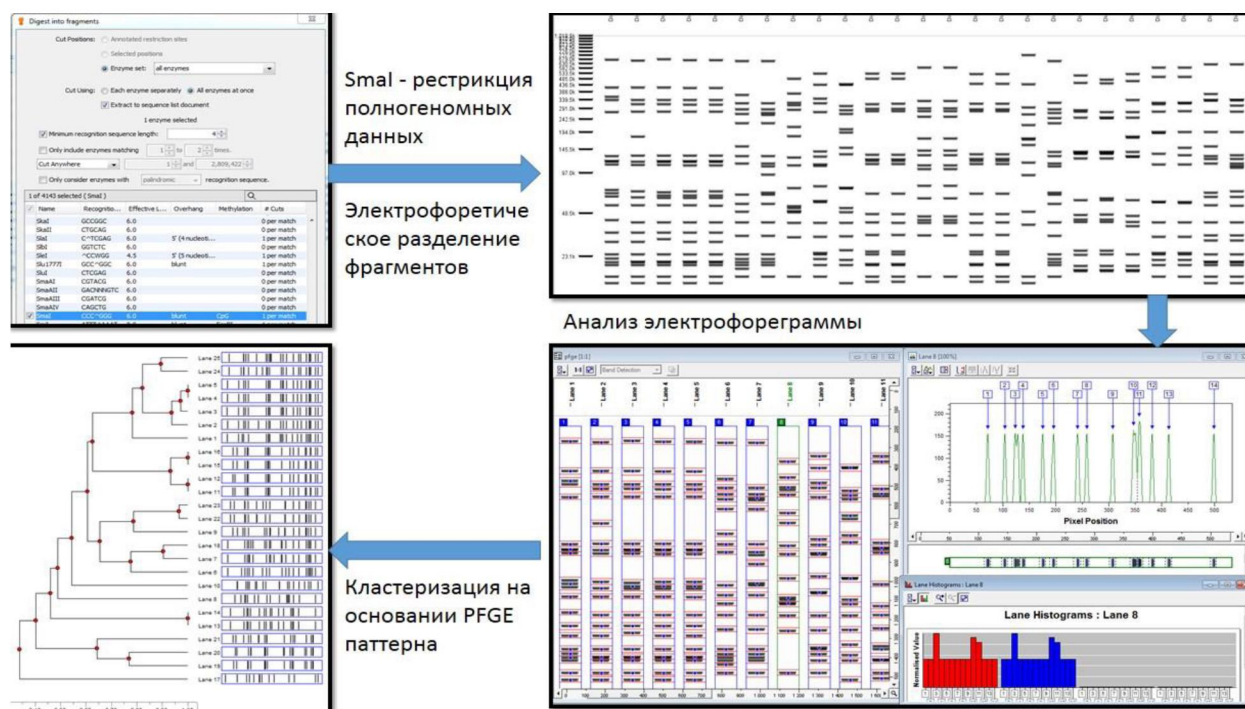


Рисунок 8 – Схема *in silico* PFGE типирования геномных данных *S.aureus* с использованием двух программ: вверху – Geneious, внизу TotalLab

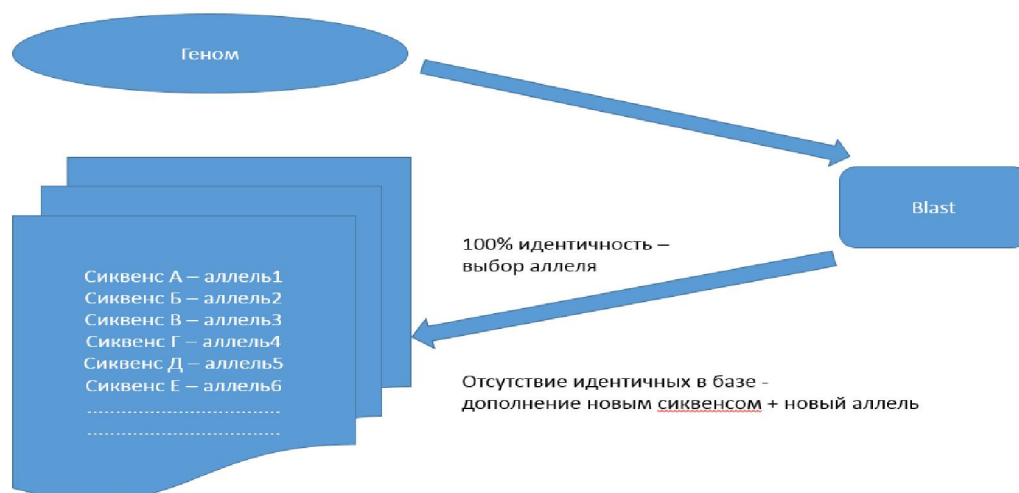


Рисунок 9 – Схематический принцип работы программы SeqSphere+ для полногеномного MLST анализа

Анализ данных полногеномного MLST типирования проводился с помощью *porppr* пакета R статистики [151]. Для выбора локусов, входящих в схемы cgMLST и wgMLST были выбраны 0% и 5% cut-off, соответственно, при использовании полногеномных данных *S.aureus*, которые являются на текущий момент завершенными (n = 116). Таким образом, cgMLST схема включала 1107 генов, а wgMLST - 2097 генов. При анализе стафилококковых геномов, состоящих из скаффолдов и/или контигов (незавершенные геномы) критерии

выбора локусов для cgMLST и wgMLST были 0,5% и 5% cut-off, что позволило сформировать схему из 1333 генов для cgMLST и 1715 генов для wgMLST. Критерий cut-off для cgMLST был изменен с 0% до 0,5% в случае с незавершенными геномами вследствие наличия разрывов (gaps) между контигами и/или скаффолдами что, в свою очередь, приводит к исключению ряда локусов при BLAST поиске методом cgMLST типирования.

Метод типирования на основе пангеномного сравнения геномов *S.aureus* (PanGenome typing) проводился с помощью R/Bioconductor package «FindMyFriends» [152]. Предварительно была произведена детекция транскрибируемых участков генома (coding DNA sequences – CDS), их размера и позиция в геноме при помощи программы Prodigal, переводя из нуклеотидной последовательности в аминокислотную [153]. Результаты анализа геномов сохранялись в виде fasta файлов, которые в дальнейшем использовались при пангеномном анализе. Из всей совокупности генов геномов формировались группы-ортологи с 85 % идентичностью на основе алгоритмов CD-HIT [154, 155] и kmer cosine similarity [152, с.2]. В результате была сформирована матрица бинарных значений (1 – наличие группы-ортолога в геноме, 0 – отсутствие), которая использовалась в дальнейшем при кластеризации изолятов *S.aureus*.

2.4 Кластеризация результатов типирования

Группирование изолятов *S.aureus* проводилось методом кластеризации, путем построения дендрограммы или деревьев минимальных дистанций (minimum spanning tree – MST) на основе данных, полученных различными методами типирования.

Оценка генетического различия между различными Spa типами проводилась методом попарного выравнивания Spa повторов с использованием алгоритма Needleman–Wunsch [156]. На основе полученного элаймента рассчитывалась матрица генетических дистанций, где за единицу генетического события принимались точечные мутации (mismatches) или вставка/выпадение (InDels). Кластеризация проводилась на основе алгоритма деревьев минимальных дистанций. В spa группу вошли spa-типы, отличавшиеся на 1 событие.

Результаты типирования методами MLST, MVLST, MLVA, cgMLST, wgMLST, представленные в виде аллельных профилей, анализировались с использованием формулы расчета генетического расстояния, предложенной Nei М. с соавт. [157, 158]:

$$1 - \frac{1}{r} \sum_j^r \sum_j^{m_j} \sqrt{x_{ij} y_{ij}}$$

Кластеризация и визуализация аллельных данных осуществлялась с использованием eBURST/MST алгоритма [159]. Формирование кластеров, комплексов осуществлялось при использовании двух правил:

1. Максимальная дистанция в группе/комплексе;
2. Минимальное количество изолятов в группе.

Для методов типирования, включающих ограниченное количество маркеров в схеме типирования максимальная дистанция для сформированной группы является 1, при условии не менее 2 изолятов в группе.

Формирование кластеров для схем с большим количеством локусов в схеме типирования, например, cgMLST и wgMLST проводилось на основании значения 10% cut-off.

Кластеризация результатов типирования методом пульс-гель электрофореза осуществлялась на сравнении паттернов бэндов, которые кодировались в виде бинарных данных, где 1 – наличие полосы, 0 – отсутствие полосы на геле. Оценка генетических дистанций бинарных данных проводилась методом, предложенным Jaccard P. (1912) [160, 161]. В результате, формирование кластеризации проводилось на основе алгоритма невзвешенного попарного среднего (Unweighted Pair-Group Method Using Arithmetic Averages – UPGMA) на основании полученной матрицы генетических дистанций [162, 163]. Формирование PFGE групп проводилось при 80% cut-off [164].

Данные PanGenome типирования представляют собой бинарную таблицу генетических групп-ортологов, где 1 – присутствие гена в ортолог-группе, 0 – отсутствие гена. Кластеризация и визуализация бинарных данных осуществлялась с использованием eBURST/MST алгоритма; группирование осуществлялось с использованием 1% cut-off.

2.5 Оценка дискриминирующей мощности методов внутривидового типирования и их согласованность (конкордантность)

Оценка дискриминирующей способности метода типирования проводилась с использованием индекса Симпсона по формуле, используя «*vegan*» r package [165]):

$$D = \sum_{i=1}^S \frac{n_i(n_i - 1)}{N(N - 1)}$$

Параметр конкордантности оценивался на основании скорректированного коэффициента Ранда [166] («*vegclust*» r package [167]).

$$\begin{aligned} ARI(X, Y) &= \frac{Index - ExpectedIndex}{MaxIndex - ExpectedIndex} \\ &= \frac{2(N_{00}N_{11} - N_{01}N_{10})}{(N_{00} + N_{01})(N_{01} + N_{11}) + (N_{00} + N_{10})(N_{10} + N_{11})} \end{aligned}$$

Расчет 95% доверительного интервала для показателей разрешающей мощности метода внутривидового типирования проводился путем resampling jackknife («*resample*» r package [168]).

3 РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

3.1 Характеристика клинических изолятов *S.aureus*. Определение чувствительности к антибиотикам

Метициллин-устойчивый золотистый стафилококк (MRSA) является этиологическим агентом множества тяжелых инфекционных заболеваний с высоким риском смертности [169]. Хотя, по мнению некоторых исследований, в ряде случаев наблюдается носительство данного агента, роль которого до сих пор не ясна [170, 171]. Тем не менее, в последние годы наблюдается тенденция к увеличению распространения устойчивых штаммов *S.aureus* к антимикробным препаратам [172]. Как было показано в ряде работ, изоляты *S.aureus* способны приобретать генетические элементы, в частности SCCmecA и *mecC* кассеты, несущие гены патогенности, вирулентности и устойчивости к антибактериальным препаратам, в результате горизонтального переноса [173, 174].

Результаты молекулярно-эпидемиологических исследований, изучающие распространенность *S.aureus*, в том числе MRSA, показали, что для стафилококка характерна клональная популяционная структура. При этом наблюдаются различия между разными клонами комплексами [175, 176]. Отмечается, что благодаря своим особенностям, некоторые клоны *S.aureus* обладают преимуществом и способны к более быстрому распространению в госпиталях, являясь тем самым более частыми этиологическими агентами внутрибольничных вспышек [177-179].

На начальном этапе нашего исследования было необходимо фенотипически оценить коллекцию клинических штаммов *S.aureus*. На момент исследования микробиологическая коллекция *S.aureus* была представлена 98 изолятами, собранными из различных стационаров Караганды, Астаны, Жезказгана и Семей. Долевое соотношение клинических штаммов золотистого стафилококка, полученных из различных городов Казахстана, представлено на рисунке 10.

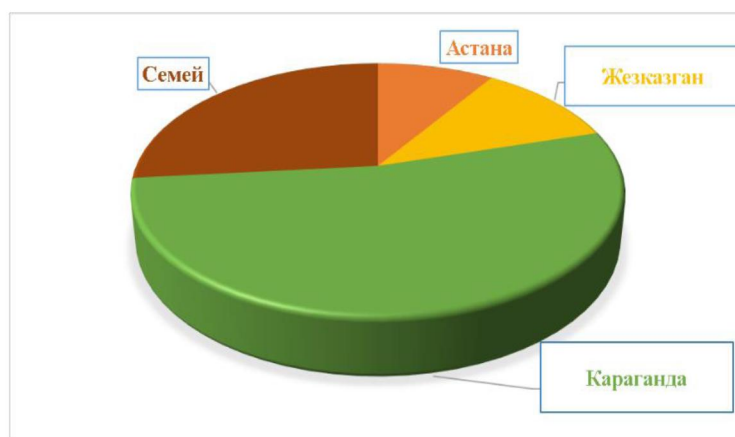


Рисунок 10 – Долевое соотношение изолятов *S.aureus*, выделенных из различных городов Республики Казахстан

Как представлено на рисунке 10, более половины (64,3%) штаммов были получены из стационаров Карагандинского региона, включая 53,1% штаммов из Караганды и 11,2% из Жезказгана. Долевые пропорции штаммов из Семей и Астаны составили 25,6% и 9,2%, соответственно.

Клинические образцы, из которых происходило выделение *S.aureus*, а также структура профилей стационаров представлена в таблице 5.

Таблица 5 – Профиль отделений и клинический материал, из которых были выделены клинические штаммы *S.aureus*

Клинический материал	Профиль отделений						
	Дермато- венеро- логия	Лор	Офталь- мология	Педи- атрия	Пульмо- нология	Терапия	Хирур- гия
Бронхо- альвеолярный лаваж	0	0	0	0	1	0	1
Гнойное отделяемое	0	1	0	0	0	3	31
Мазок из зева	0	15	0	16	0	1	4
Соскоб с кожи	1	0	0	0	0	1	0
Мазок с конъюнктивы	0	0	3	1	0	0	1
Кровь	0	0	0	0	0	1	0
Мокрота	0	1	0	4	2	1	1
Моча	0	0	0	0	0	2	0
Выделения из носа	0	1	0	0	0	0	0
Раневое отделяемое	0	0	0	0	0	0	3
Мазок из уха	0	2	0	0	0	0	0

Для оценки уровня устойчивости к антибактериальным препаратам было проведено фенотипическое исследование методом серийных разведений с целью получения количественных показателей резистентности клинических штаммов *S.aureus*. Оценка минимальной ингибирующей концентрации (МИК) антибиотика проводилась спектрофотометрически. Определение устойчивости к антибиотику устанавливалось в соответствии с руководством Клинической лаборатории Института стандартов (Clinical Laboratory Standards Institute - CLSI) [180], которые включены в программу WhoNet.

Результаты антибиотикограммы *S.aureus* представлены в таблице 6.

Таблица 6 - Чувствительность к антибиотикам штаммов *S.aureus*, определенная методом серийных разведений

Антибиотик	Количество штаммов	%R	%I	%S	%R 95%C.I.
Oxacillin	98	17,3	0	82,7	10.7-26.6
Cefazolin	98	17,3	0	82,7	10.7-26.6
Cefoxitin	88	18,2	0	81,8	11.1-28.2
Gentamicin	98	10,2	0	89,8	5.3-18.4
Streptomycin	92	3,3	1,1	95,7	0.9-10.0
Tobramycin	86	7	0	93	2.9-15.2
Rifampin	98	2	0	98	0.3-7.8
Ciprofloxacin	98	11,2	5,1	83,7	6.0-19.6
Lincomycin	98	15,3	0	84,7	9.1-24.3
Azithromycin	98	24,5	0	75,5	16.6-34.4
Linezolid	96	0	0	100	0.0-4.8
Vancomycin	98	2	0	98	0.3-7.8
Fusidic acid	97	0	0	100	0.0-4.7

Одним из методов, подтверждающих устойчивость к метициллину, оксациллину является определение генетического маркера *mecA* гена [181]. В результате ПЦР исследования 17 штаммов *S.aureus* обладали *mecA* геном, что позволило их отнести к группе MRSA. Электрофоретическая картина результатов ПЦР амплификации представлена на рисунке 11.

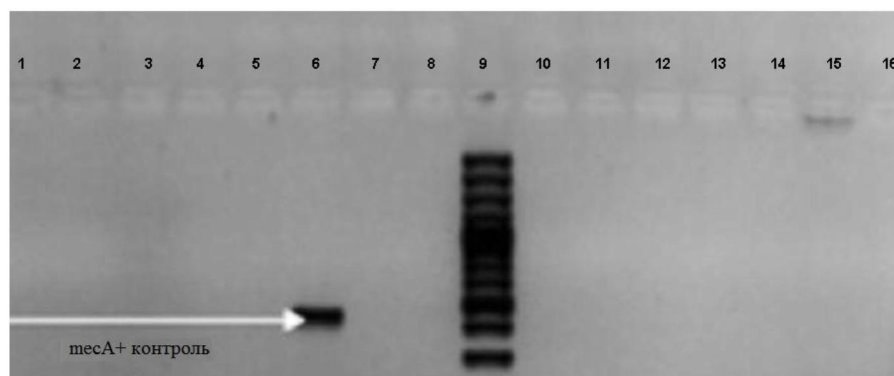


Рисунок 11 – Электрофоретическая детекция *mecA* гена. 9 линия – маркер молекулярных размеров, 6 – положительный контроль, 1-5, 7, 8, 10-16 линии клинических штаммов *S.aureus* (результат отрицательный)

По результатам лабораторного анализа по определению чувствительности к антибактериальным препаратам, выяснилось, что высокой антибактериальной активностью обладали ванкомицин, линезолид, рифампицин, фузидиевая кислота, препараты группы аминогликозидов (стрептомицин, тобрамицин). В то же время, резистентность к азитромицину, представителю макролидов,

наблюдалась в 24,5 % случаев (95% ДИ 16,6-34,4), являясь следствием индуцибельного характера возникновения резистентности к макролидам [182-184]. Также наблюдался достаточно высокий процент устойчивых штаммов к целофаспоринам - цефазолин 17,3% (95% ДИ 10.7-26.6), цефокситин 18,2% (95% ДИ, 11.1-28.2).

Существуют два основным механизма формирования устойчивости к макролида грам-положительных бактерий: 1. модификация мишени действия, в частности, 50S субъединица бактериальной рибосомы [185], 2. Активация системы эффлюкса.[186], Модификационный процесс мишени включает в метилировании аденина в 23S-рибосомальной РНК ферментом метилазой. На данный момент описано свыше 30 генов-метилаз (erm — erythromycin ribosome methylation) для различных видов микроорганизмов [187]. Эффлюкс[188] система обеспечивает выброс антибиотиков посредством либо белков семейства АТФ-связывающих кассет, либо посредством мембранных транспортных белков, обеспечивающих перенос веществ из клетки наружу, обеспечивая нормальную функциональную активность рибосом. Большинство макролид-резистентные гены ассоциированы с мобильными генетическими элементами, обеспечивая распространение устойчивости среди микроорганизмов [188, с.3885].

Таким образом, оценка чувствительности к антимикробным препаратам штаммов *S.aureus*, выделенных из различных стационаров городов: Караганда, Астана, Жезказган и Семей позволила выявить 17 штаммов MRSA. Устойчивость к макролидам (азитромицин) составила 25%, а к цефалоспоринам (цефазолин, цефакситин) порядка 18%.

3.2 Spa типирование клинических штаммов *S.aureus*

Типирование коллекции штаммов *S.aureus* (n=98) методом секвенирования фрагмента Spa-гена проходило в 2 стадии:

1. Амплификация гипервариабельного региона Spa-гена методом ПЦР. Контроль за эффективностью амплификации осуществлялся методом гель-электрофореза (рисунок 12).
2. Определение последовательности продуктов амплификации методом капиллярного секвенирования (по методу Sanger).

Как видно из электрофоретической картины амплификации Spa гена, в 31 случае из 36 было проведено успешное амплифицирование фрагмента Spa гена. Исходя из маркера молекулярных масс, размеры продуктов ПЦР амплификации варьировал от 300-420 нуклеотидов.

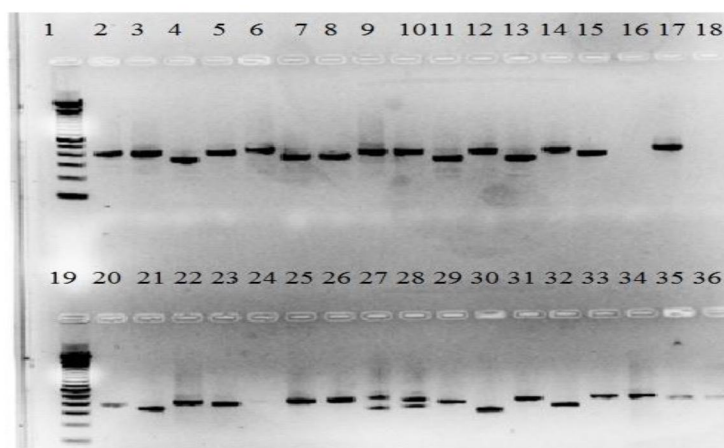


Рисунок 12 – Длины амплифицированных фрагментов переменного региона Spa исследованных штаммов

Примечание - 1, 19 – Маркеры молекулярных размеров, 2-15, 17, 20-36 – клинические образцы *S. aureus*. 18 – отрицательных контроль. 16, 24 – отсутствие амплификации Spa гена

Секвенирование продуктов ПЦР исследования позволило установить первичную последовательность ДНК гипервариабельного участка Spa гена. Биоинформатический подход позволил определить Spa тип штаммов *S. aureus* исходя из полученной последовательности X региона Spa гена.

Из 98 образцов 94 изолята были успешно проанализированы с помощью Spa типирования. На рисунке 13 представлена структура Spa типов *S. aureus*.

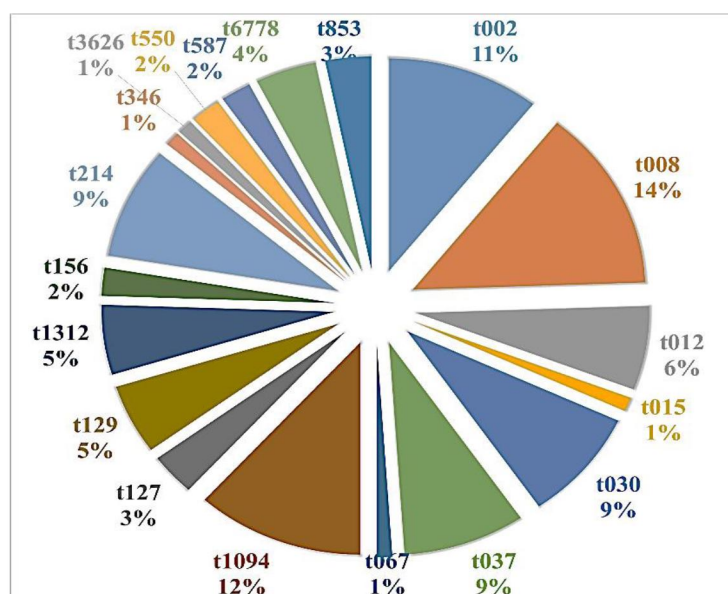


Рисунок 13 – Структура Spa типов, полученных на коллекции штаммов *S. aureus* (n = 94)

Как видно из таблицы 7, доминирующими Spa типами являлись t008 (13.8%), t1094 (11.7%), t002 (10.6%).

Таблица 7 – Результаты Sra типирования коллекции клинических изолятов *S.aureus*

Город	Sra тип	Sra повторы	Количество
Астана	t002	26-23-17-34-17-20-17-12-17-16	2
Жезказган	t002	26-23-17-34-17-20-17-12-17-16	2
Караганда	t002	26-23-17-34-17-20-17-12-17-16	1
Семей	t002	26-23-17-34-17-20-17-12-17-16	5
Жезказган	t008	11-19-12-21-17-34-24-34-22-25	2
Караганда	t008	11-19-12-21-17-34-24-34-22-25	5
Семей	t008	11-19-12-21-17-34-24-34-22-25	6
Караганда	t012	15-12-16-02-16-02-25-17-24-24	4
Семей	t012	15-12-16-02-16-02-25-17-24-24	2
Жезказган	t015	08-16-02-16-34-13-17-34-16-34	1
Астана	t030	15-12-16-02-24-24	1
Караганда	t030	15-12-16-02-24-24	6
Семей	t030	15-12-16-02-24-24	1
Астана	t037	15-12-16-02-25-17-24	1
Караганда	t037	15-12-16-02-25-17-24	4
Семей	t037	15-12-16-02-25-17-24	3
Семей	t067	26-23-17-34-17-20-17-12-17	1
Жезказган	t1094	23-17-34-17-20-17-12-17-16	3
Караганда	t1094	23-17-34-17-20-17-12-17-16	7
Семей	t1094	23-17-34-17-20-17-12-17-16	1
Караганда	t127	07-23-21-16-34-33-13	2
Семей	t127	07-23-21-16-34-33-13	1
Караганда	t129	15-12-24	4
Семей	t129	15-12-24	1
Жезказган	t1312	04-20-12-17-17	2
Караганда	t1312	04-20-12-17-17	2
Семей	t1312	04-20-12-17-17	1
Караганда	t156	07-23-12-33-22-17	2
Астана	t214	26-23-17-34-17-20-17-12-17-16-16	3
Жезказган	t214	26-23-17-34-17-20-17-12-17-16-16	1
Караганда	t214	26-23-17-34-17-20-17-12-17-16-16	2
Семей	t214	26-23-17-34-17-20-17-12-17-16-16	2
Караганда	t346	07-23-12-34-12-12-23-02-12-23	1
Караганда	t3626	07-23-12-21-17-34-34-34-33	1
Астана	t550	08-17-34-16-34	2
Караганда	t587	14-12-23-02-02-34-34	2
Караганда	t6778	26-23-16-34-17-20-17-12-17-16	3
Семей	t6778	26-23-16-34-17-20-17-12-17-16	1
Караганда	t853	07-23-34-34-12-12-23-02-12-23	3

В то же время, на практике принято проводить группирование близкородственных Sra типов в группы, для того чтобы определить насколько гетерогенная структура *S.aureus* циркулирует в том или ином регионе.

Было проведено сравнение профилей Sra типов, на основании которых формировались Sra группы, в которые вошли Sra типы, отличающиеся на одно генетическое событие (мутация или делеция/инсерция).

Таким образом, в результате кластеризации методом деревьев минимальных дистанций были сформированы 4 sraCC, включающие от 2 до 5 sra типов (рисунок 14).

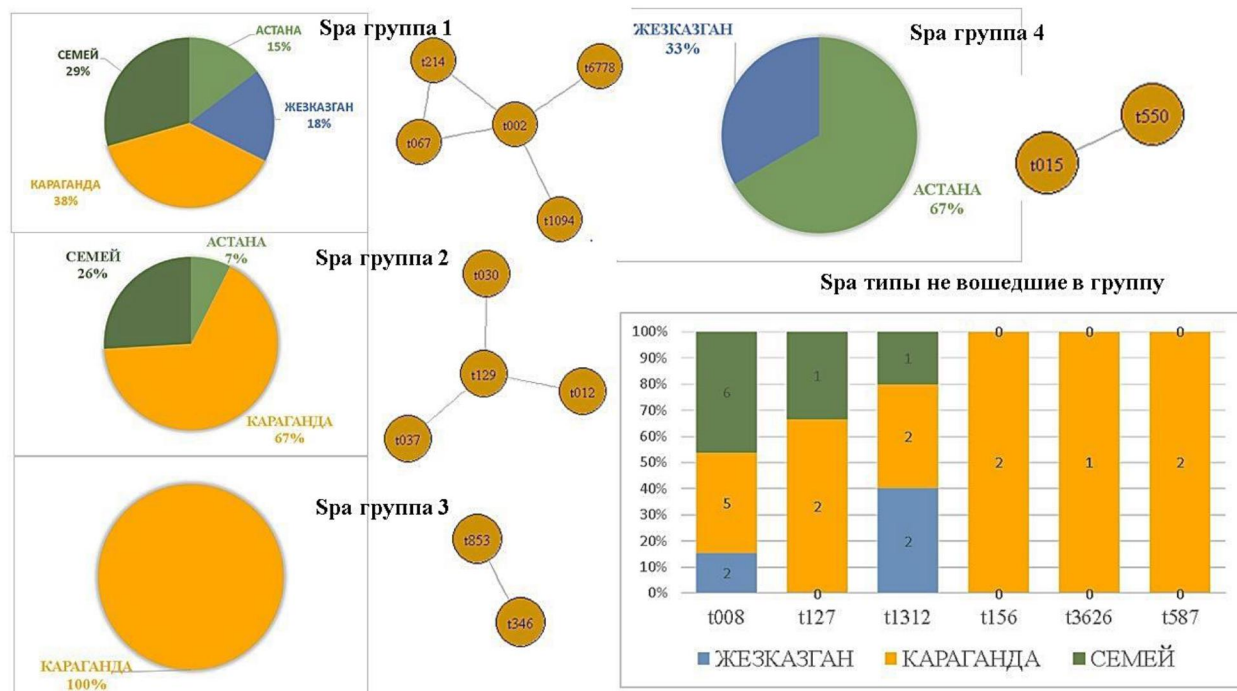


Рисунок 14 – Кластеризация Sra типов в группы методом деревьев минимальных дистанций

В результате Sra типирования 94 штаммов *S.aureus* 33 изолята (35%) принадлежали 1 sra группе (sraCC/t002), 27 изолятов *S.aureus* (28.7%) вошли во вторую sra группу (sraCC/t030). Sra тип t008 был определен в 13 изолятах *S.aureus* (13.8%).

Кластеризация Sra типов изолятов, выделенных из различных регионов Республики Казахстан (Астана, Караганда, Жезказган и Семей), в Sra группы указывает на доминирование двух Sra клональных групп и Sra типа t008, что позволяет сделать вывод циркуляции определенных клональных групп в Республике Казахстан. Полученные результаты согласуются с данными авторов соседних стран – России [189, 190] и Китая [191, 192]. Таким образом, циркуляция определенных Sra-типов, распространенных в Казахстане и в соседних странах, а также во многих других регионах земного шара [193-195] позволяет говорить о вовлеченности в глобальную эпидемиологию изолятов *S.aureus*, выделенных в Республике Казахстан.

Оценка разрешающей способности Spa типирования, как метода внутривидового типирования стафилококков, составила 0,929 (95% ДИ; 0,913-0,944), что указывает на достаточно высокую дискриминирующую силу данного метода типирования.

3.3 MLVA типирование клинических штаммов *S.aureus*

Внутривидовое типирование коллекции штаммов *S.aureus* методом анализа тандемных повторов (MLVA типирование) проводилось с использованием 6 SIRU маркеров - SIRU 01, SIRU 05, SIRU 07, SIRU 13, SIRU 15, SIRU 21. В результате данного исследования были выявлены 16 MLVA типов (MLVA профилей). Наиболее распространенными MLVA профилями были 295-566-305-305-368-342 (21,4%), 348-266-360-179-368-342 (18,4%) и 348-246-360-179-368-342 (12,2%). Распределение MLVA профилей *S.aureus* представлено на рисунке 15.

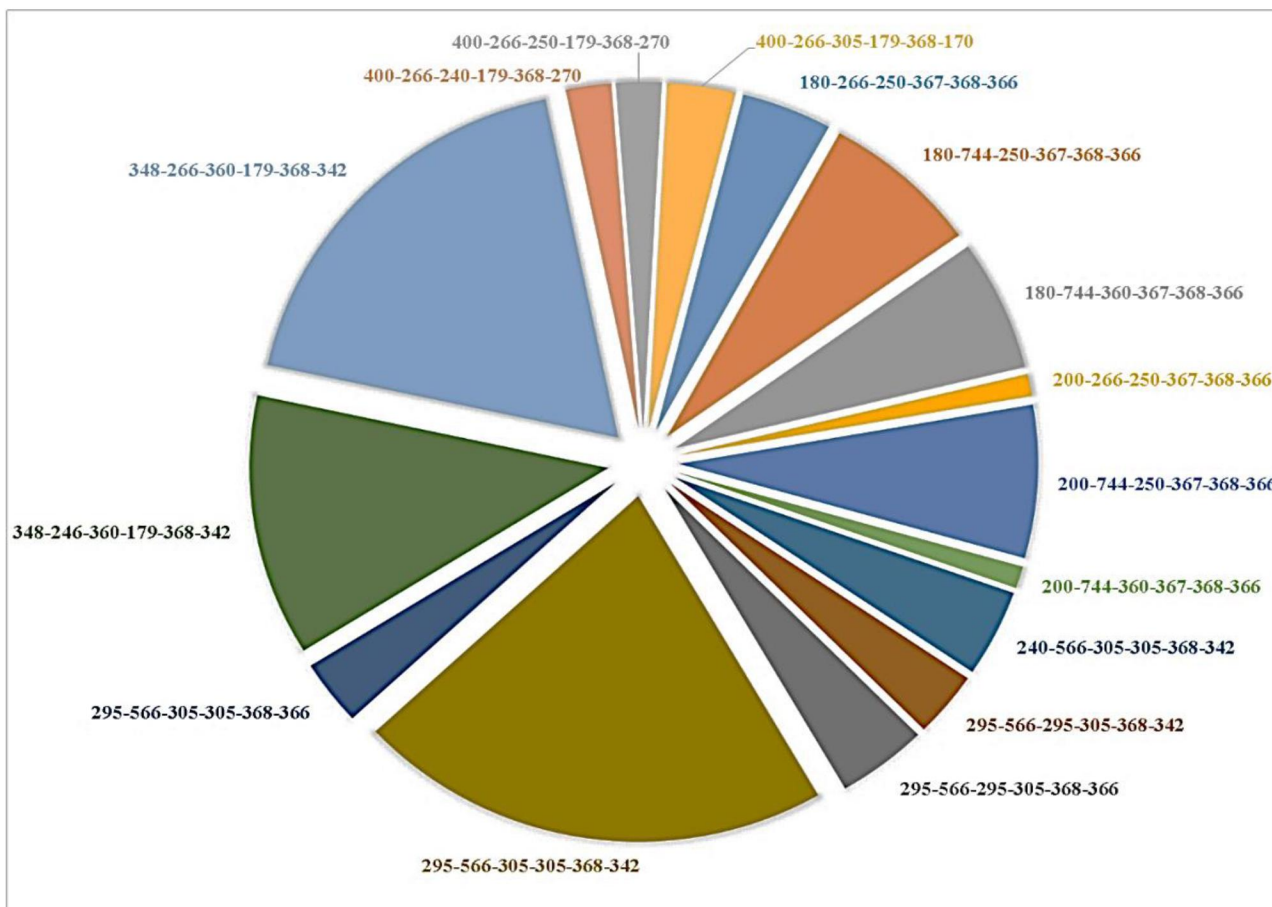


Рисунок 15 – Структура MLVA типов клинических изолятов *S.aureus* (n=98)

Данные о количестве штаммов *S.aureus* с определенным MLVA профилем, а также место их выделения представлены в таблице 8.

Таблица 8 - Количественная характеристика изолятов *S.aureus* с определенным MLVA типов и местом изоляции

Город	MLVA профиль	Количество
Караганда	180-266-250-367-368-366	1
Семей	180-266-250-367-368-366	3
Караганда	180-744-250-367-368-366	5
Семей	180-744-250-367-368-366	2
Караганда	180-744-360-367-368-366	6
Семей	200-266-250-367-368-366	1
Жезказган	200-744-250-367-368-366	3
Караганда	200-744-250-367-368-366	2
Семей	200-744-250-367-368-366	2
Жезказган	200-744-360-367-368-366	1
Караганда	240-566-305-305-368-342	4
Караганда	295-566-295-305-368-342	1
Семей	295-566-295-305-368-342	2
Семей	295-566-295-305-368-366	4
Астана	295-566-305-305-368-342	5
Жезказган	295-566-305-305-368-342	5
Караганда	295-566-305-305-368-342	6
Семей	295-566-305-305-368-342	5
Жезказган	295-566-305-305-368-366	1
Караганда	295-566-305-305-368-366	2
Астана	348-246-360-179-368-342	2
Караганда	348-246-360-179-368-342	7
Семей	348-246-360-179-368-342	3
Караганда	348-266-360-179-368-342	14
Семей	348-266-360-179-368-342	4
Караганда	400-266-240-179-368-270	2
Караганда	400-266-250-179-368-270	2
Астана	400-266-305-179-368-170	2
Жезказган	400-266-305-179-368-170	1

Кластеризация штаммов *S.aureus* на основе MLVA профилей осуществлялась методом построения деревьев минимальных дистанций, представленная на рисунке 16. Формирование MLVA комплексов проводилось на основании правила: максимальное отличие в комплексе =1, минимальное количество штаммов формирующие комплекс = 2. В результате были получены 4 MLVA комплекса.

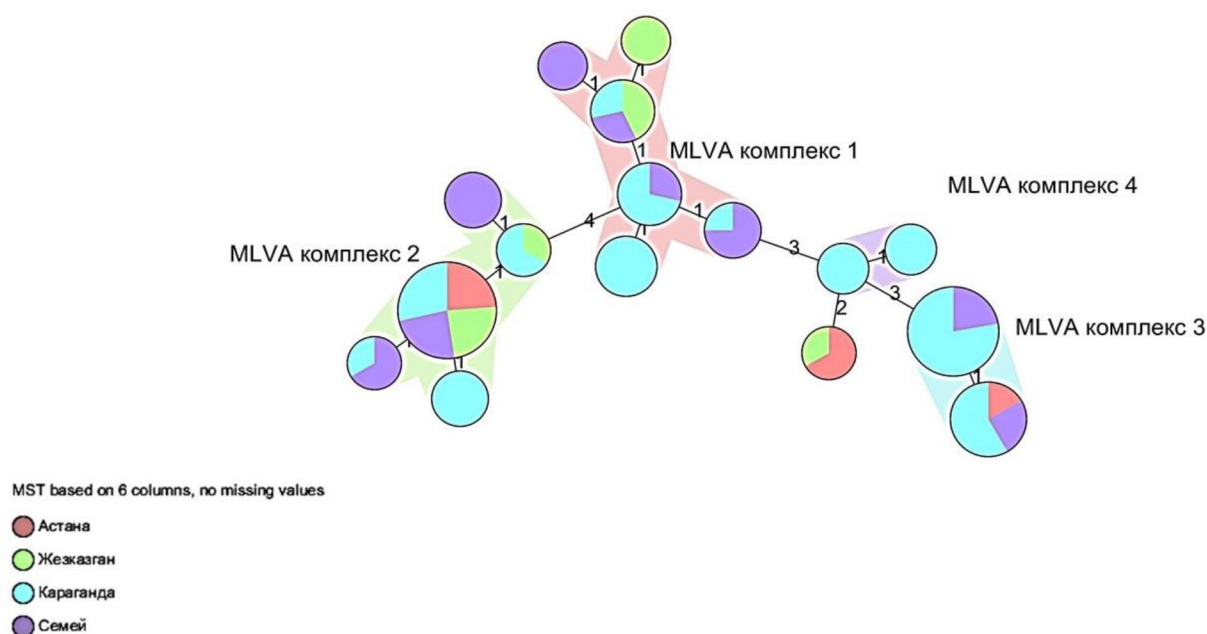


Рисунок 16 – Кластеризация клинических штаммов *S.aureus* методом MLVA типирования

Анализ данных MLVA типирования коллекции *S.aureus*, полученных из городов: Астана, Караганда, Жезказган и Семей позволяет сделать заключение о распространении и взаимной циркуляции штаммов *S.aureus* в данном регионе.

Оценка разрешающей способности метода MLVA составила 0,898 (95% ДИ, 0,869-0,926), что говорит о достаточно высокой разрешающей способности метода типирования, основанного на тандемных повторах.

3.4 Результаты сравнения данных Spa и MLVA типирований

Сравнительная оценка дискриминирующей силы двух методов типирования клинических изолятов золотистого стафилококка представлена в таблице 9. Несмотря на более высокий показатель разрешающей способности Spa типирования в сравнении с MLVA типированием, перекрывающийся 95% доверительные интервалы не позволяют заключить о значимом различии в разрешающей способности данных методов типирования.

Таблица 9 – Оценка дискриминирующей мощности методов типирования *S.aureus*

Метод типирования	Количество различных типов	Индекс Симпсона	95% ДИ
MLVA	16	0,898	(0,869-0,926)
Spa	17	0,923	(0,906-0,939)

Оценка согласованности результатов Spa и MLVA типирования проводилась на основе расчета скорректированного коэффициента Ранда (таблица 10). Данные Spa типирования и MLVA типирования имеют слабую согласованность 20,7% (95% ДИ; 14,2-27,3%) на уровне внутривидовых типов. В то же время, группирование близкородственных типов в группы/комплексы позволило добиться 86,2% (95% ДИ; 75,8-97%) конкордантности.

Таблица 10 – Согласованность результатов Spa и MLVA типирования

Внутривидовые типы и группы	Spa тип	Spa группа	MLVA тип
Spa группа	0,406 (0,305-0,502)	-	-
MLVA тип	0,207 (0,142-0,273)	0,479 (0,372-0,590)	-
MLVA комплекс	0,323 (0,263-0,384)	0,862 (0,758-0,970)	0,444 (0,347-0,544)

Таким образом, результаты исследования на 98 клинических изолятах *S.aureus*, собранных за 2014 год из различных стационаров Астаны, Караганды, Жезказгана и Семей указывают на межрегиональную циркуляцию определенных внутривидовых типов *S.aureus*, кластеризованных в 4 основные группы, полученных как Spa типированием, так и MLVA типированием. Основные Spa типы, такие как t002, t003, t008, которые выявлены в нашем исследовании, также распространены во многих регионах земного шара [196-198], указывая на вовлеченность в глобальную эпидемиологию клинических изолятов *S.aureus*, выделенных в Республике Казахстан.

Дискриминирующая способность двух методов субвидового типирования *S.aureus* (Spa и MLVA типирование) была схожей и составила порядка ~90-92%.

Анализ на конкордантность результатов различных методов типирования, оцененный с помощью скорректированного показателя Ранда, показал слабую согласованность между Spa и MLVA типами (20,7%), но достаточно высокую (86%) на уровне комплексов.

Слабая согласованность результатов типирования на уровне внутривидовых типов, но хорошая согласованность на уровне близкородственных групп (комплексов) объясняется причинами:

- Маркеры Spa и MLVA типирования обладают разными скоростями эволюционных изменений, что привело к невысокой согласованности результатов типирования, используя типы при сравнении.
- Коллекция клинических штаммов *S.aureus*, собранная за относительно небольшой промежуток времени в ограниченном географическом ареале (регион Центрального Казахстана и город Семей), а также клональная

популяционная структура золотистого стафилококка объясняет 86% конкордантность результатов типирования на уровне групп/комплексов.

3.5 Характеристика полногеномных данных *S.aureus*, включенных в *in silico* исследование

На момент исследования (август 2016 года) база данных NCBI GenBank включала 116 завершенных, аннотированных геномов *S.aureus*. Последовательности геномов использовались для определения PFGE, MLVA, MLST, MVLST, spa, cgMLST, wgMLST и panGenome типов с помощью биоинформатических методов.

Анализ литературных данных, а также информации, указанной GenBank при включении генома *S.aureus* в базу данных, о времени и месте изоляции штамма *S.aureus* указывает на разобщенность коллекции изолятов золотистого стафилококка ($n = 116$) в пространственно-временном измерении (рисунок 17).

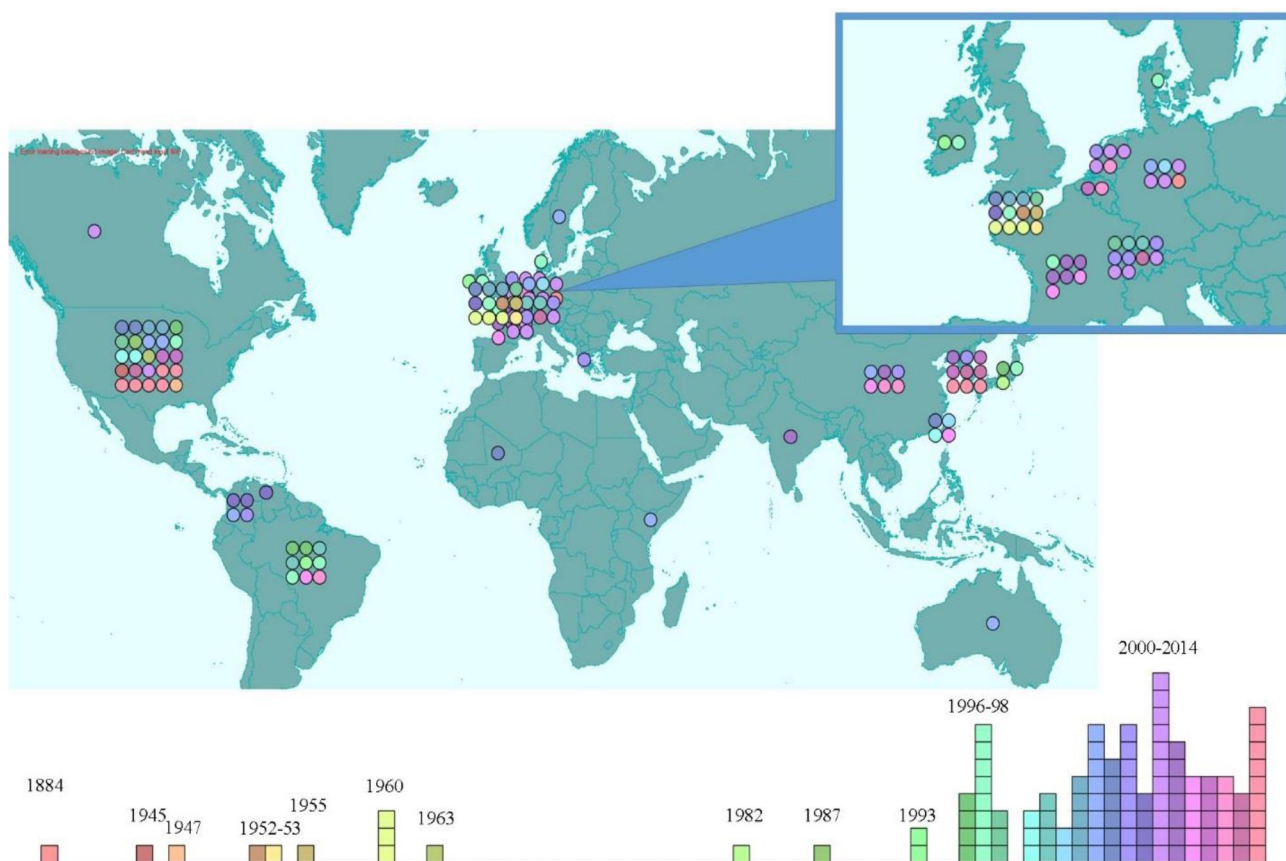


Рисунок 17 – Пространственно-временная характеристика коллекции изолятов *S.aureus*, для которых были определены полногеномные последовательности

3.6 *In silico* анализ на определение PFGE, MLVA, MLST, MVLST, spa, cgMLST, wgMLST, PanGenome типов используя полногеномные данные *S.aureus*. Оценка дискриминирующей способности методов типирования и конкордантности между полученными данными

Типирование *S.aureus* методом пульс-гель электрофореза (PFGE) - *in silico* исследование. Хромосомная ДНК 116 штаммов *S.aureus* подверглась

расщеплению *in silico* с помощью фермента SmaI, распознающего CCC[^]GGG последовательность. В результате рестрикции были получены фрагменты, которые разделялись в виртуальном электрофореze в диапазоне от 50 до 750 килобаз с формированием полос в геле от 6 до 12 (в среднем 9 бэндов) (рисунок 18).

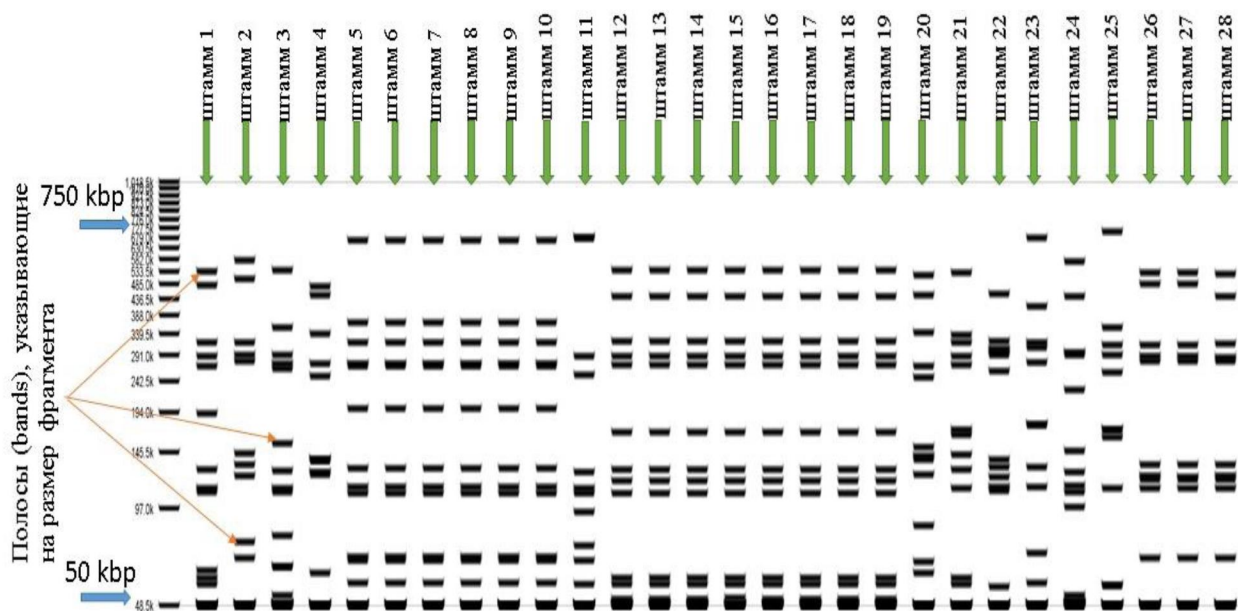


Рисунок 18 – Пример формирования электрофоретического профиля
пульс-гель электрофореза

Валидационная оценка пульс-гель электрофореграммы, полученной биоинформатическими методами, проводилась в сравнении с пульс-форезограммами, взятыми из публикации, где использовались штаммы, которые на текущий момент являются полностью отсековированы и аннотированы [199, 200]. В работе Hiramatsu K. и соавт., проводилось PFGE типирование в диапазоне от 50 до 400 килобаз, включающее 4 штамма, которые исследовались *in silico* в нашей работе. В исследовательской работе Sieradzki K. с соавт. проводился анализ штаммов *S.aureus* методов PFGE в диапазоне 48,5-630 кб., который включал 3 общих с нашим исследованием штаммов *S.aureus*. Как показывает сравнение электрофорезограмм, в обоих случаях наблюдается хорошее совпадение бэндов, полученных в лабораторных условиях и *in silico* (рисунки 15 и 16).

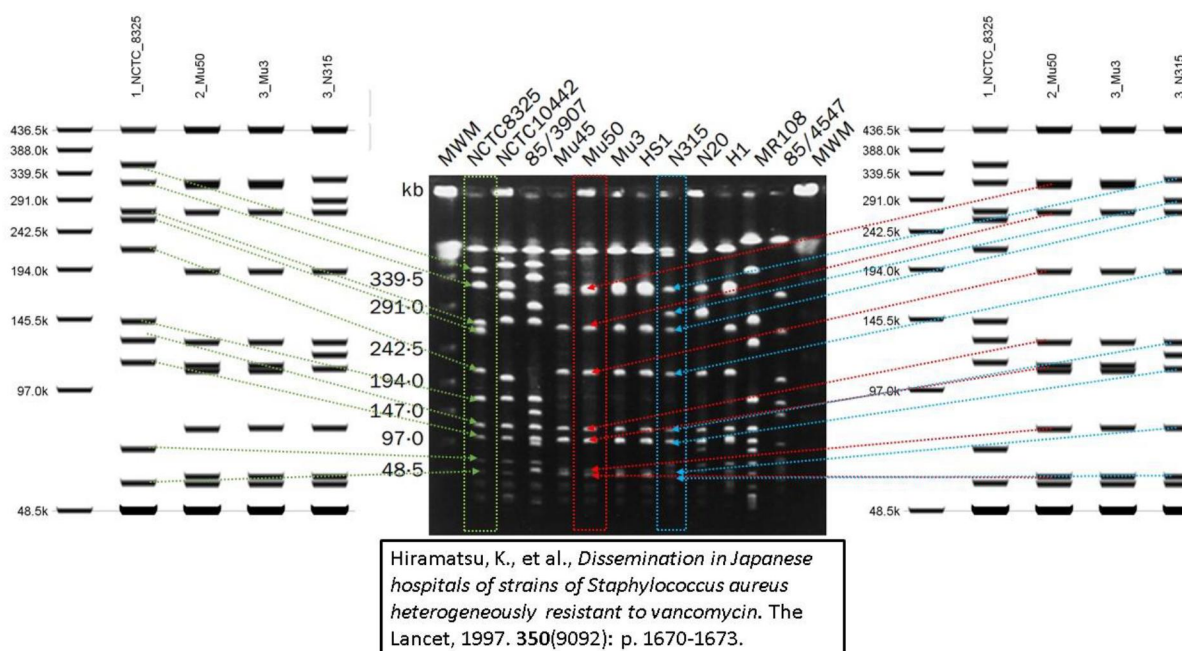


Рисунок 19 – Сравнение пульс-гель паттернов, полученных в лабораторных условиях и *in silico* 4 общих изолятов *S.aureus* (SmaI рестрикция)

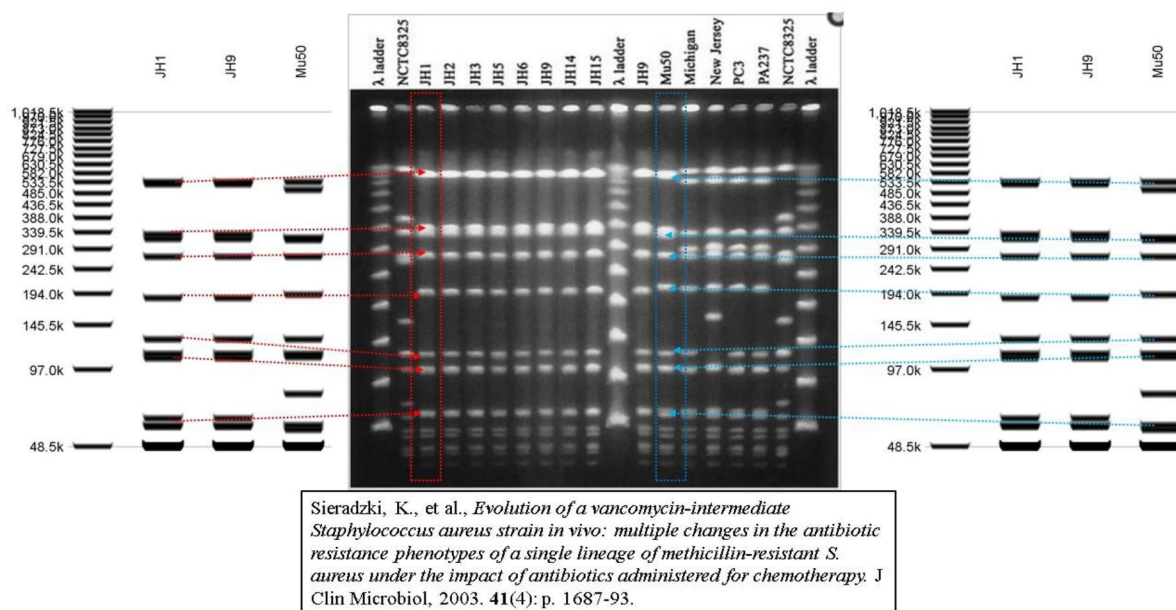


Рисунок 20 – Сравнение пульс-гель паттернов, полученных в лабораторных условиях и *in silico* 3 общих изолятов *S.aureus* (SmaI рестрикция)

Сформированные *in silico* PFGE профили 116 штаммов *S.aureus* были кластеризованы с целью получения пульс-групп, включающих близкородственные пульс-типы. Формирование групп проводилось на основе Jaccard /UPGMA алгоритма для построения дендрограммы, на которой 80% cut-off определял PFGE группы (рисунок 21).

Пульс-гель электрофорез позволил детектировать из 116 штаммов 93 уникальных пуль-типов. Оценка разрешающей способности данного метода типирования (индекс Симпсона) составила 0,991 (95% ДИ, 0,984-0,998).

Типирование *S.aureus* методом мультилокусного VNTR анализа (MLVA) – *in silico* исследование. Определение внутривидового типа *S.aureus* на основе Multiple Loci VNTR Analysis (MLVA) осуществлялось путем симуляции e-PCR с использованием 8 пар праймеров VNTR09, VNTR21, VNTR24, VNTR61-01, VNTR61-02, VNTR63, VNTR67 и VNTR81, предложенных Leo M. Schouls и др. [148, с.5087]. Полученные размеры продуктов амплификации по 8 локусам в последующем пересчитывались на количество повторов в каждом локусе для формирования MLVA профиля. С целью оценки уровня согласованности результатов MLVA типирования, полученные *in silico* и данные, представленные Leo M. Schouls в MLVA базе данных для стафилококков (количество изолятов *S.aureus*, которые на данный момент полностью отсекарованы и аннотированы, а также присутствуют в базе данных составляет 14 изолятов).

Результаты сравнения MLVA профилей, полученных в нашей работе биоинформатическим подходом и результатами, представленные Leo M. Schouls с соавт. указаны в таблице 11.

Таблица 11 – Сравнение MLVA профилей, полученных в лабораторных условиях и *in silico*

Штамм <i>S. aureus</i>	VNTR locus								MLVA профиль, полученный в данной работе
	V09 01	V21 01	V24 01	V61 01	V61 02	V63 01	V67 01	V81 01	
RF122	8	1	3	0	1	1	1	4	8-1-3-0-1-1-1-4
N315	14	2	11	2	1	6	3	5	14-2-11-2-1-6-3-5
NCTC8325	12	1	12	3	3	2	9	4	12-1-12-3-3-2-9-4
USA300	12	1	11	3	3	2	5	3	12-1-11-3-3-2-5-3
USA300_TCH1516	12	1	9	3	3	2	5	3	12-1-9-3-3-2-5-3
Str. Newman	12	1	11	3	3	1	5	4	12-1-11-3-3-1-5-4
COL	12	1	11	3	3	3	5	4	12-1-11-3-3-3-5-4
JH1	14	2	11	2	1	5	3	5	14-2-11-2-1-5-3-5
JH9	14	2	11	2	1	5	3	5	14-2-11-2-1-5-3-5
MRSA252	19	1	12	99*	2	2	2	3	19-1-12-отр-2-2-2-3
MSSA476	11	1	9	1	1	2	5	4	11-1-9-1-1-2-5-4
MW2	9	1	9	0	2	1	7	4	9-1-9-0-2-1-7-4
Mu3	14	4	11	2	1	6	3	5	14-4-11-2-1-6-3-5
Mu50	14	4	11	2	1	6	3	5	14-4-11-2-1-6-3-5
Примечание - 99* - отсутствие продуктов амплификации									

Как показывает сравнение, результаты, полученные в нашей работе, полностью совпадают с результатами, полученными вышеуказанными авторами, что также свидетельствует о высокой достоверности данных типирования методом MLVA, полученных *in silico*.

В то же время, метод, основанный на ПЦР амплификации участков с tandemными повторами, выявил случаи, когда не наблюдается накопление продуктов амплификации для некоторых локусов. В нашем *in silico* эксперименте в случаях с VNTR21, VNTR24, VNTR61-01, VNTR61-02, VNTR63, VNTR81 наблюдались отрицательные результаты ПЦР в 8, 1, 4, 2, 1 и 1 образцах, соответственно (таблица 12).

Таблица 12 – Количество негативных ПЦР реакций, а также индекс Симпсона, рассчитанный для каждого локуса MLVA схемы типирования *S.aureus*

MLVA локус	Количество образцов	Количество образцов с отрицательной ПЦР	Количество различных MLVA типов	Индекс Симпсона	95% ДИ
VNTR09	116	0	11	0.714	[0.65 - 0.778]
VNTR21	108	8	4	0.391	[0.299 - 0.484]
VNTR24	115	1	9	0.808	[0.77 - 0.846]
VNTR61_01	112	4	7	0.751	[0.71 - 0.791]
VNTR61_02	114	2	4	0.655	[0.614 - 0.697]
VNTR63	115	1	7	0.723	[0.67 - 0.777]
VNTR67	116	0	8	0.788	[0.759 - 0.817]
VNTR81	115	1	7	0.742	[0.703 - 0.781]

О возможном отрицательном результате ПЦР амплификации некоторых локусов MLV A схемы типирования *S.aureus* также указывается в ряде работ [201-203].

Полученные MLVA профили, включающие 8 маркеров, использовались с целью группирования MLVA комплексов, включающих MLVA типы, отличающихся на 1 маркер. Кластеризация данных MVLA типирования проводилась на основе eBURST/MST алгоритма, которая представлена на рисунке 22.

MLVA типы = 81
 MLVA комплексы = 12
 Количество образцов в
 MLVA комплексах = 90

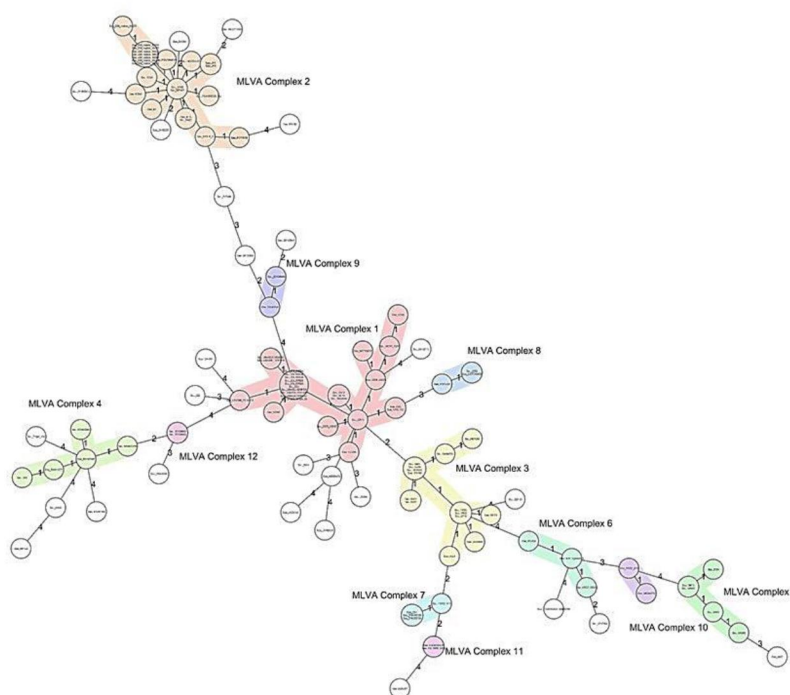


Рисунок 22 - Кластеризация 116 штаммов *S.aureus* на основании MLVA типирования методом деревьев минимальных дистанция (minimum spanning tree - MST)

Примечание - Цветом выделены MLVA комплексы

В результате кластеризации 90 изолятов *S.aureus* 116 были распределены в 12 MLVA комплексов. Остальные 26 штаммов *S.aureus* не были включены в тот или иной MLVA комплекс.

Суммарный дискриминирующий индекс MLVA типирования, полученный на 116 штаммах *S.aureus* методом Симпсона составил 0,986 (95%ДИ; 0,977 – 0,995).

Типирование *S.aureus* методом мультилокусного секвенирования типирования (MLST) – *in silico* исследование.

Поиск аллельных вариантов 7 локусов MLST схемы в полногеномной последовательности *S.aureus* проводился с помощью программы MIST. База аллельных вариантов локусов и таблица MLST типов, определенная на аллельных вариантах локусов была взята с официальной базы данных.

В 113 из 116 геномах *S.aureus* были найдены аллельные последовательности локусов и определены MLST типы. В двух случаях были обнаружены 2 новых аллельных варианта *argS* гена, что не позволило установить MLST тип и в одном случае были успешно определены все аллели локусов, но комбинация маркеров являлась новой в сравнении с базой MLST типов, что также не позволило установить MLST тип. Тем не менее, все MLST профили использовались как в кластеризации, так и для определения

разрешающей способности данного метода типирования. Топ 11 детектированных MLST типов представлены в таблице 13.

Таблица 13 – Топ 11 MLST типов, определенных на полногеномных данных 116 изолятов *S.aureus*

MLST тип	Количество изолятов	% от общего количества
ST8	22	18.97
ST 5	14	12.07
ST 239	13	11.21
ST 228	8	6.90
ST 398	5	4.31
ST 15	4	3.45
ST 59	4	3.45
ST 72	4	3.45
ST 105	3	2.59
ST 30	3	2.59
ST 80	3	2.59

Как представлено в таблице 13 наиболее популярными сиквенс типами *S.aureus* являются ST8, ST5 и ST239. Согласно литературным данным, данные MLST типы являются основными представителями внутрибольничных инфекций, вызванных стафилококками, которые обладают множественной устойчивостью к антимикробным препаратам и имеют глобальное распространение [204, 205].

Результаты кластеризации MLST типов в клональные комплексы (MLST CC) представлены на рисунке 23.

MLST типы = 39
MLST комплексы = 13
Количество образцов в MLST комплексах = 104

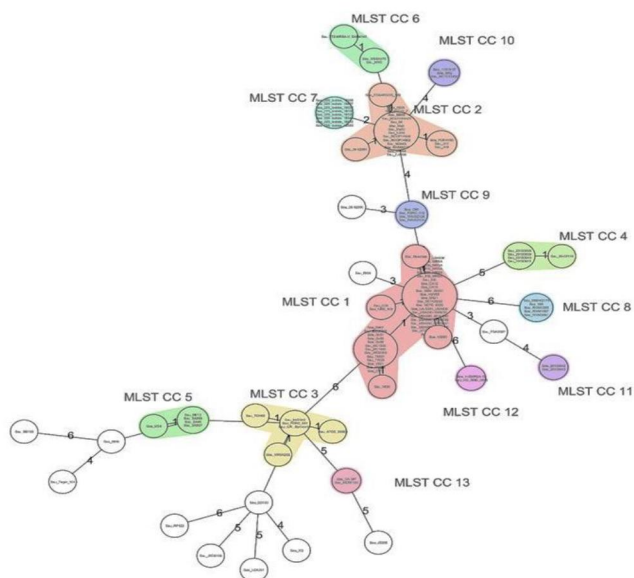


Рисунок 23 – eBURST/MLST кластеризация MLST профилей 116 штаммов *S.aureus*

В результате группирования 39 различных MLST типов *S.aureus*, полученных на данных геномной последовательности ДНК были получены 13 MLST клональных комплексов, включающих 104 штамма стафилококка.

Разрешающая способность каждого маркера, входящего в MLST схему типирования, а также количество аллельных вариантов, обнаруженных в результате MLST типирования 116 *S.aureus* представлены в таблице 14.

Таблица 14 – Дискриминирующая способность маркера, входящего в состав MLST схемы типирования *S.aureus*

MLST локус	Коли- чество образцов	Количество образцов необнаруженные в базе данных	Количество различных аллелей для локуса	Индекс Симпсона	95% ДИ
Arcc	114	2	14	0,789	(0,746 - 0,832)
Aroe	116	0	18	0,781	(0,726 - 0,835)
Glpf	116	0	14	0,434	(0,318 - 0,55)
Gmk	116	0	13	0,722	(0,665 - 0,779)
Pta	116	0	18	0,77	(0,713 - 0,828)
Tpi	116	0	22	0,813	(0,756 - 0,871)
Yqil	116	0	20	0,82	(0,765 - 0,875)

Следует отметить, что MLST схема типирования является стандартизированной методикой, которая курируется и постоянно пополняется новыми данными. Единая база данных MLST типов *S.aureus* определяет международную номенклатуру, по которой происходит определение MLST типа *S.aureus*. Несмотря на явное преимущество данного подхода, может возникнуть ситуация, когда на момент исследования может не оказаться аллельного варианта маркера MLST схемы или комбинации аллелей в базе данных. В таком случае, только после включения нового аллельного варианта или комбинации аллельных маркеров и присвоения нового сиквенс типа с последующим включением его в базу данных, можно будет определить новый вариант MLST типа.

Суммарный дискриминирующий индекс MLST схемы типирования оцененный на 116 штаммах *S.aureus* составил 0,929 (0,905 - 0,954).

Типирование *S.aureus* методом мультилокусного секвенирования типирования на основе генов вирулентности (Multi-virulence-locus sequence typing - MVLST) – *in silico* исследование.

В 2012 году Bindhu Verghese с соавт. предложили альтернативную MLST методику типирования *S.aureus* путем секвенирования 6 генов вирулентности (alt, essC, geh, hlgA, htrA и sdrC) с целью повышения дискриминирующей способности метода типирования [147, с.1819]. Предложенные праймеры для амплификации фрагментов генов были использованы для определения участков амплификации, последовательности которых сравнивались. Для каждого нового аллельного варианта локуса присваивался новый номер, а для каждой

комбинации аллелей маркеров присваивался новый MVLST тип. В результате были получены для всех 116 штаммов *S.aureus* профили MVLST типирования, на основании которых проводилась кластеризация MVLST типов в MVLST комплексы, представленные на рисунке 24.

MVLST типы = 52
MVLST комплексы = 15
Количество образцов в
MVLST комплексах = 97

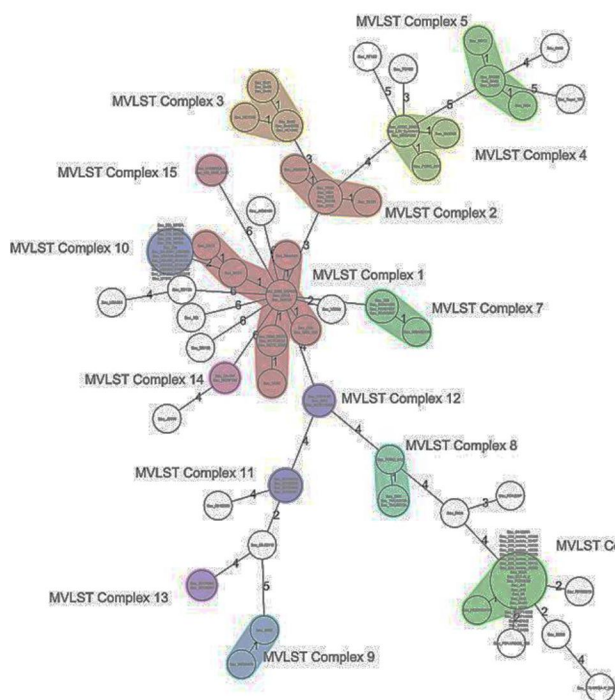


Рисунок 24 – eBURST/MST кластеризация MVLST профилей 116 штаммов *S.aureus*

В результате группирования 52 различных MVLST типов *S.aureus*, полученных в результате *in silico* анализа полногеномных данных были получены 15 MVLST комплексов, включающих 97 изолятов стафилококка.

Разрешающая способность каждого маркера, входящего в MVLST схему, а также количество аллельных вариантов, обнаруженных в результате MLST типирования 116 образцов *S.aureus* представлены в таблице 15.

Таблица 15 – Дискриминирующая способность маркера, входящего в состав MVLST схемы типирования *S.aureus*

MVLST локус	Количество образцов	Количество отрицательных образцов	Количество различных аллелей для локуса	Индекс Симпсона	95% ДИ
alt	109	7	27	0,912	(0,882 - 0,943)
EssC	116	0	26	0,88	(0,845 - 0,915)
geh	116	0	26	0,893	(0,864 - 0,922)
hlgA	116	0	18	0,802	(0,756 - 0,847)
htrA	115	1	24	0,856	(0,816 - 0,895)
sdrC	115	1	24	0,855	(0,814 - 0,897)

Как видно из таблицы 15, в некоторых случаях наблюдается отсутствие детекции маркера MVLST схемы типирования. Тем не менее, правило формирования позволяет объединять MVLST профили с различающимися на 1 аллель изолятов, входящих в MVLST комплекс. Следует отметить, что каждый маркер MVLST типирования обеспечивает высокую разрешающую способность свыше 80%. Суммарный индекс Симпсона для данного метода типирования составил 0,943 (95% ДИ; 0,916 – 0,97).

Типирование *S.aureus* методом Spa типирования – *in silico* исследование. Spa типирование основано на изучении вариации в полиморфной части протеина А (spa ген). В данной методике используют секвенирование одного локуса, что позволяет отнести его к однолокусному секвенированию типированию. Метод, в основе которого лежит секвенирование фрагмента гена обладает такими же преимуществами, как и метод MLST. Более того, в данном случае происходит секвенирование одного локуса, что обеспечивает преимущество как в стоимости, так и во времени выполнения одного анализа.

Определение Spa типа, используя данные последовательности геномов *S.aureus*, проводилось биоинформатическими методами. В результате анализа 116 стафилококковых геномов, в 3 случаях не удалось установить Spa тип, так как на момент исследования последовательности spa повторов не было в базе данных Spa типов. В результате исследования были получены 50 различных spa типов. Топ 10 spa типов представлены в таблице 16.

Таблица 16 – Часто встречающиеся spa типы, полученные на данных полногеномных последовательностях *S.aureus*

Spa тип	Spa повторы	Количество изолятов	% от общего количества
t008	11-19-12-21-17-34-24-34-22-25	16	14.16
t002	26-23-17-34-17-20-17-12-17-16	14	12.39
t037	15-12-16-02-25-17-24	11	9.73
t1003	26-17-20-17-34-17-20-17-12-17-16	8	7.08
t044	07-23-12-34-34-33-34	3	2.65
t084	07-23-12-34-34-12-12-23-02-12-23	3	2.65
t324	07-23-12-12-17-20-17-12-12-17	3	2.65
t437	04-20-17-20-17-25-34	3	2.65
t011	08-16-02-25-34-24-25	2	1.77
t021	15-12-16-02-16-02-25-17-24	2	1.77

С целью установления групп, была проведена кластеризация Spa типов. В Spa группу вошли Spa типы, отличающиеся на 1 генетическое событие. В результате было получено 10 spa групп, включающих 96 штаммов стафилококков (рисунок 25).

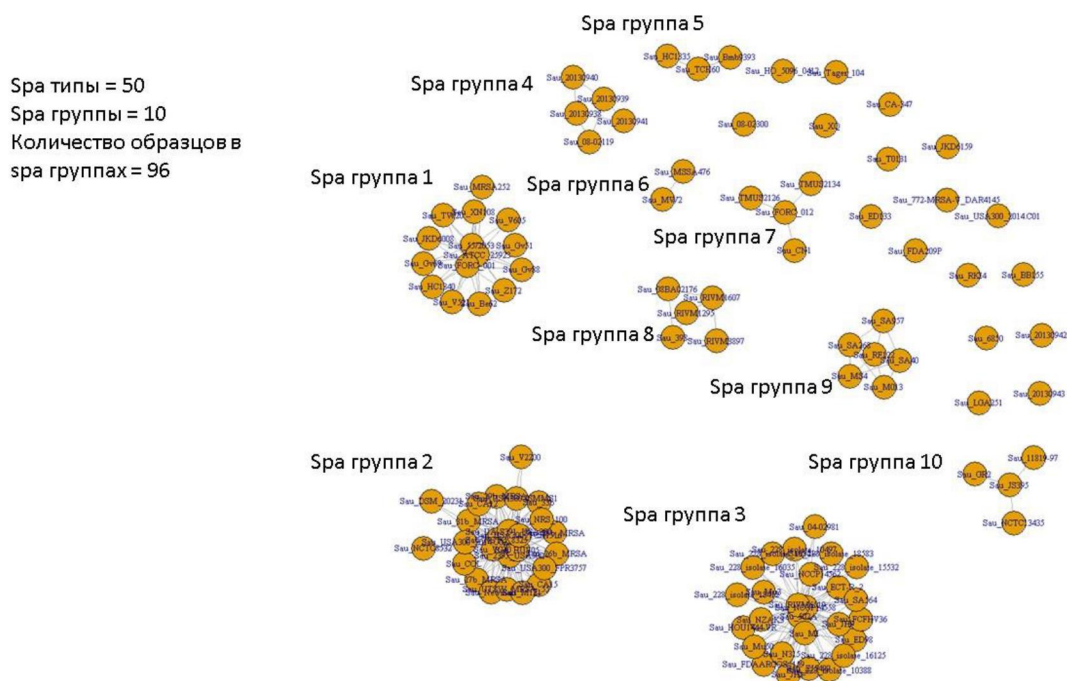


Рисунок 25 – Кластеризация spa типов в группы

Данный метод типирования стафилококков показал 95% (95%ДИ; 93-97%) дискриминирующую мощность.

Типирование *S.aureus* методами cgMLST и wgMLST – *in silico* исследование. Предложенный Shana R. Leopold с соавт. в 2014 году новый метод типирования cgMLST (MLST⁺ в программе SeqSphere) *S.aureus* представлял собой полногеномный вариант аллельного типирования [206]. В качестве референса был использован геном *S. aureus* изолят COL (GenBank accession no. NC_002951). Проведя cgMLST анализ с 39 данными полногеномными последовательностями *S.aureus* методом попарного сравнения генов в геноме при использовании параметров $\geq 90\%$ идентичности и 100% покрытия генов получили набор генов-маркеров (n=1861), вошедших в схему cgMLST типирования. Впоследствии с целью исключения путаницы в наименовании cgMLST был переименован в wgMLST. На данный момент wgMLST схема типирования *S.aureus* курируется Alexander Mellmann (Inst. Hygiene) и Dag Harmsen (University Munster). В то же время, исходя из определения core genome в cgMLST схему, следует включить гены, которые встречаются во всех геномах данного вида микроорганизмов. В результате проведенного нами анализа 116 геном *S.aureus* 1107 генов вошли в core геном. Для wgMLST схемы были выбраны гены, которые встречались в 95% и выше, что в результате дало возможность получить 2097 локусов.

Характеристика 10 маркеров с наибольшей разрешающей способностью, входящих в cgMLST и wgMLST схемы типирования представлены в таблицах 17 и 18, соответственно.

Таблица 17 – Топ 10 локусов cgMLST с наивысшей разрешающей способностью

Локус	Белок	Ген	Длина гена	N геномов	Различные типы	Индекс Симпсона (95% ДИ)
SACOL 0164	gramicidin S synthetase 2 related protein	-	7176	116	53	0,953 (0,933 - 0,972)
SACOL 1390	DNA topoisomerase IV subunit A	parC	2403	116	45	0,953 (0,938 - 0,968)
SACOL 0520	DNA polymerase III subunits gamma and tau	dnaX	1698	116	36	0,926 (0,903 - 0,949)
SACOL 1490	penicillin-binding protein 2	pbp2	2184	116	37	0,926 (0,898 - 0,955)
SACOL 0196	Gfo/Idh/MocA family oxidoreductase	-	1077	116	32	0,925 (0,898 - 0,951)
SACOL 0514	glutamate synthase	gltB	4500	116	53	0,923 (0,885 - 0,962)
SACOL 0261	drug transporter	-	1377	116	34	0,921 (0,897 - 0,946)
SACOL 2663	PTS system, fructose-specific IIABC components	-	1953	116	34	0,921 (0,895 - 0,947)
SACOL 2624	acetyl-CoA synthetase	-	1599	116	38	0,92 (0,892 - 0,948)
SACOL 0515	glutamate synthase	gltD	1464	116	34	0,919 (0,892 - 0,945)

Таблица 18 – Топ 10 локусов wgMLST с наивысшей разрешающей способностью

Локус	Белок	Ген	Длина гена	N геномов	Различные типы	Индекс Симпсона (95% ДИ)
SACOL0164	gramicidin S synthetase 2 related protein	-	7176	116	53	0,953 (0,933 - 0,972)
SACOL1390	DNA topoisomerase IV subunit A	parC	2403	116	45	0,953 0,938 - 0,968)
SACOL0006	DNA gyrase, A subunit	gyrA	2664	115	39	0,946 (0,929 - 0,962)
SACOL1644	competence protein ComEC/Rec2	-	2148	112	34	0,946 (0,93 - 0,962)
SACOL0547	transcription-repair coupling factor	mfd	3507	113	40	0,941 (0,916 - 0,965)
SACOL1791	FtsK/SpoIIIE family protein	-	3825	111	37	0,939 (0,919 - 0,959)
SACOL0233	sorbitol dehydrogenase	gutB	1056	115	35	0,938 (0,917 - 0,96)
SACOL0115	glycosyl transferase, group 1 family protein	-	1167	114	32	0,935 (0,917 - 0,953)
SACOL0020	sensory box histidine kinase YycG	yycG	1827	114	40	0,93 (0,904 - 0,956)
SACOL2725	ABC transporter permease	-	1881	115	37	0,928 (0,901 - 0,954)

Группирование 116 штаммов *S.aureus* на основании данных cgMLST и wgMLST осуществлялось методом кластеризации с использованием eBURST/MST алгоритма, на котором изоляты с 10% и менее различиями формировали кластеры. Графическое изображение дендрограммы дерева минимальных дистанций с выделенными кластерами *S.aureus* представлены на рисунках 26 и 27.

cgMLST типы = 105
 cgMLST кластеры = 14
 Количество образцов в
 cgMLST кластерах = 88

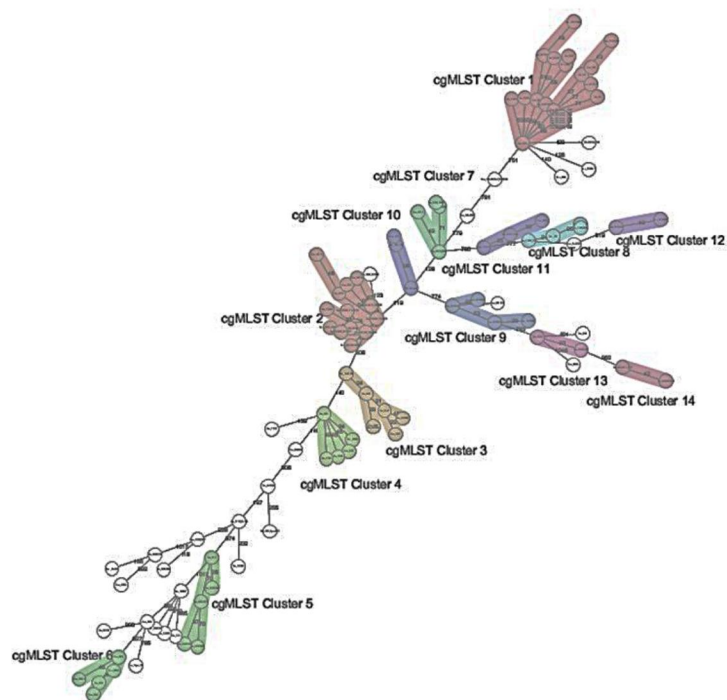
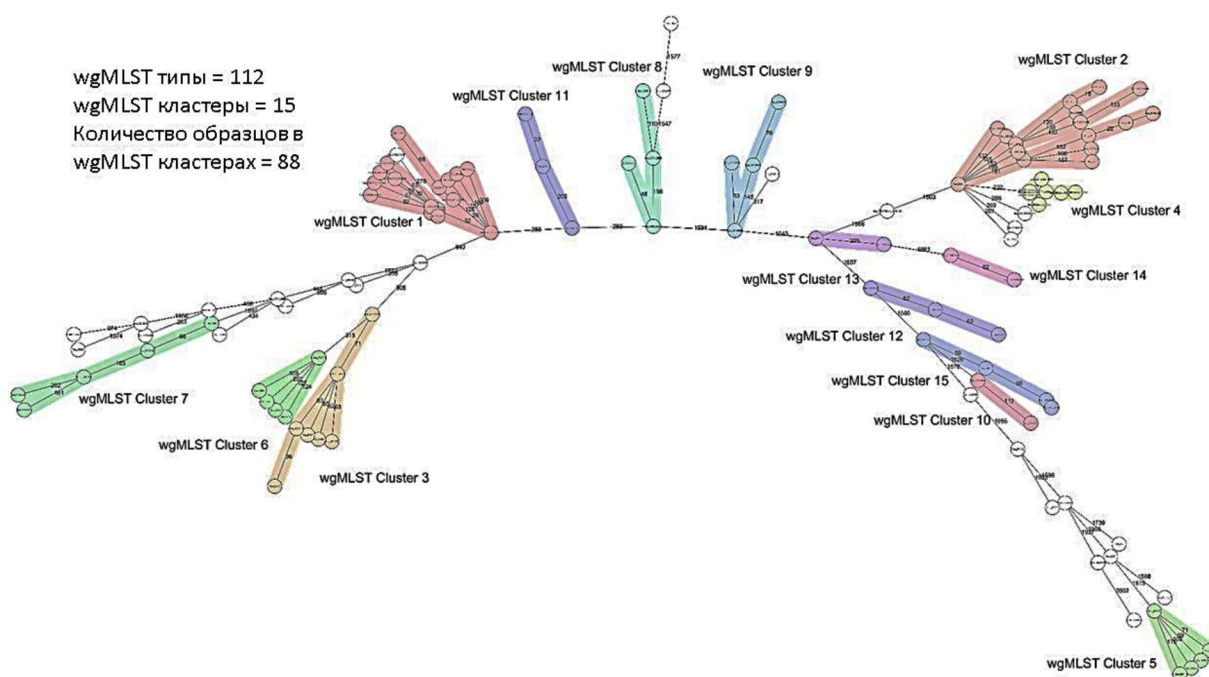


Рисунок 26 – Кластеризация 116 изолятов *S. aureus* на основании cgMLST типировании



wgMLST типы = 112
 wgMLST кластеры = 15
 Количество образцов в
 wgMLST кластерах = 88

Рисунок 27 – Кластеризация 116 изолятов *S. aureus* на основании wgMLST типировании

В результате типирования методами cgMLST и wgMLST были получены 105 и 112 различных внутривидовых типов, соответственно, что позволило

добиться разрешающей мощности 99,5% (95% ДИ; 98,9 – 100%) и 99,9% (95% ДИ; 99,8 – 100%).

Типирование *S.aureus* методом PanGenome – *in silico* исследование.

В сравнительной геномике микроорганизмов, panGenome определяется исходя из распределения генов по группам-ортологам на основе их сходства. Термин «panGenome» ввел Duccio Medini в 2004 году, а год спустя Hervé Tettelin опубликовал первую работу по исследованию panGenome *Streptococcus agalactiae* [207].

Как правило, при исследовании пангенома микроорганизмов, все гены распределены на 3 группы, в зависимости от их распространенности среди геномов. К основной группе (core genes) относятся гены, представленные во всех геномах, дополнительная или вспомогательная группа генов (accessories или dispensable genes), которые встречаются в некоторых, но не во всех геномах и уникальная группа (singletone или unique genes) – гены, встречающиеся только в одном геноме.

На текущий момент существуют несколько программ, имеющие различные принципы формирования групп-ортологов, необходимых для анализа пангенома [208-212].

В данном исследовании был использован «FindMyFriends» R/Bioconductor пакет при анализе 116 штаммов *S.aureus* используя параметры анализа по умолчанию, за исключением установления 85% cut-off для установления гомологии и определения ортолог-группы. В результате анализа 116 геномов *S.aureus* было определено 10339 групп-ортологов, на основании которых и проводился анализ дискриминирующей мощности, а также кластеризация и оценка конкордантности, полученных данных с результатами других методов типирования. Результаты распределения генов в пангеноме представлены на рисунке 28.

Примечательным является тот факт, что при использовании panGenome анализа 116 штаммов core genome был представлен 1115 генами, в то время как в cgMLST схеме, которые также должны быть представлены во всех 116 геномах количество генов было схожим и включало 1107. При этом следует отметить, что определение core genome происходило на основании использования принципиально различающихся алгоритмов.

EBURST/MST алгоритм кластеризации бинарных данных генетических групп-ортологов 116 штаммов *S.aureus* позволил определить кластеры на основании 1% cut-off. Графическое изображение дендрограммы дерева минимальных дистанций с выделенными кластерами *S.aureus* представлены на рисунке 29.

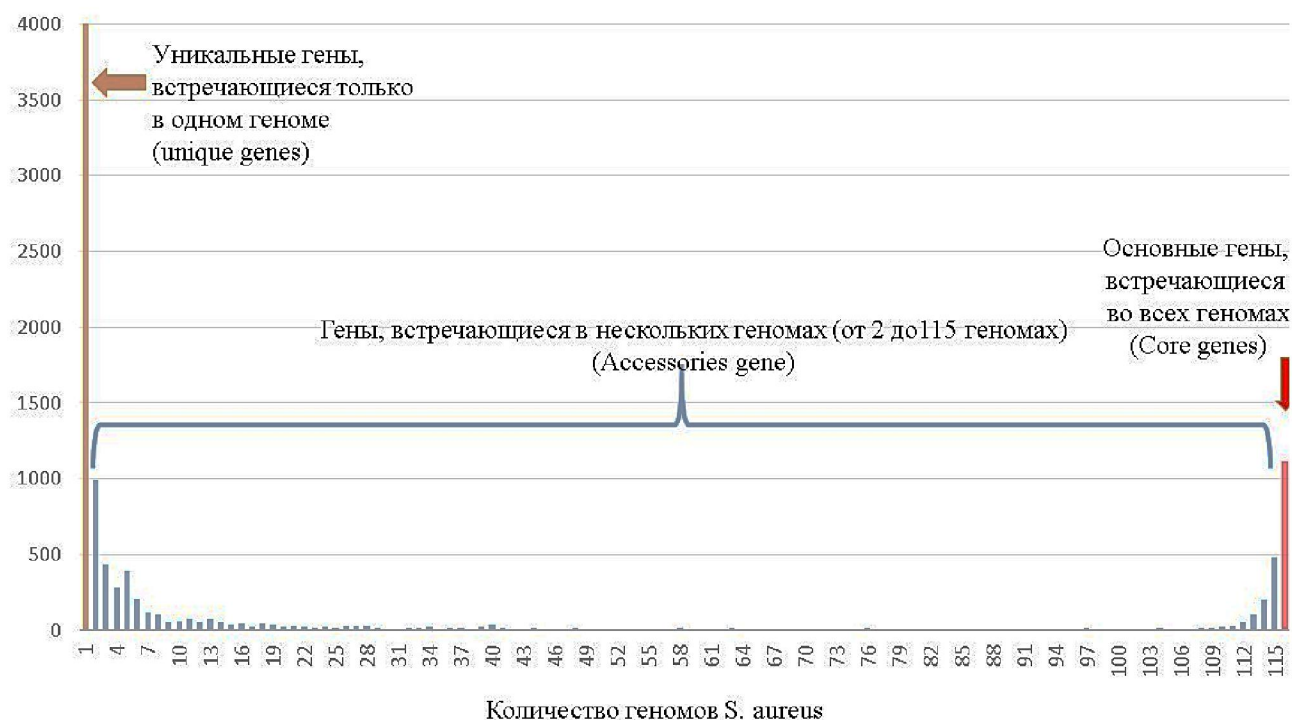


Рисунок 28 – Три группы генов, полученных в результате пангеномного анализа 116 *S. aureus*

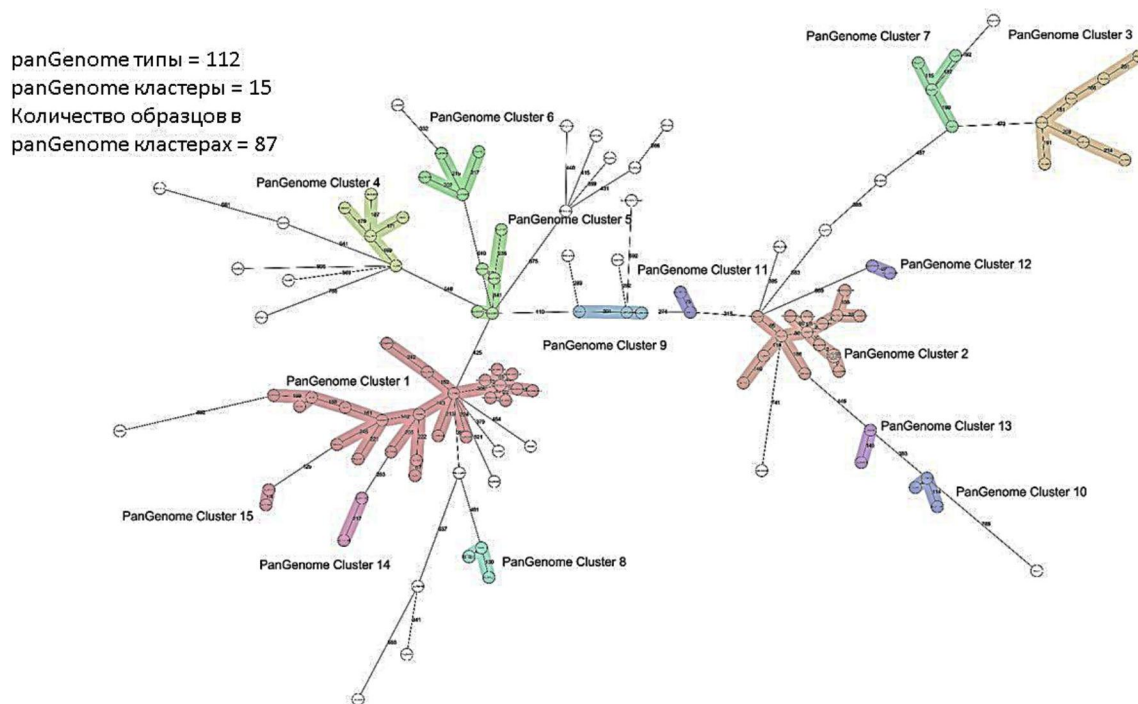


Рисунок 29 – Кластеризация штаммов *S. aureus* при использовании panGenome типирования

Пангеномный анализ 116 штаммов *S. aureus* позволил определить 112 различных внутривидовых типов, что определяет данный метод, как высокоразрешающий с 99,9% (95% ДИ; 99,7-100%) дискриминирующей мощностью.

3.7 Дискриминирующая способность и конкордантность методов типирования

Биоинформатическое исследование полногеномных данных 116 изолятов *S.aureus* с целью установления внутривидовых типов, применяя различные современные методы генотипирования, позволило определить один из самых важных параметров метода типирования – дискриминирующую мощность. Результаты *in silico* исследования по оценке разрешающей способности восьми современных молекулярно-генетических методов представлено в таблице 19.

Таблица 19 – Разрешающая мощность современных молекулярно-генетических методов типирования

Метод типирования	Количество различных типов	Индекс Симпсона	95% ДИ
panGenome	112	0,999	(0,997-1,000)
wgMLST	112	0,999	(0,998-1,000)
cgMLST	102	0,995	(0,989-1,000)
PFGE	90	0,991	(0,984-0,998)
MLVA	79	0,986	(0,977-0,995)
Spa	50	0,95	(0,93-0,97)
MVLST	52	0,943	(0,916-0,97)
MLST	38	0,929	(0,905 - 0,954)

Методы типирования *S.aureus*, основанные на использовании полногеномной последовательности, такие как panGenome, wgMLST и cgMLST, а также использование редкощелящих эндонуклеаз, узнающих специфические сайты рестрикции по всему геному (PFGE типирование), показали наибольшую разрешающую способность типирования в 99% и выше. Метод, в основе которого используются локусы с tandemными повторами, продемонстрировал чуть ниже дискриминирующую мощность (98,5%). Spa, MVLST и MLST как методы типирования, основанные на секвенировании небольшого числа локусов, обладают 95% и ниже разрешающей силой. В то же время следует отметить, что перекрывающиеся 95% доверительные интервалы не позволяют говорить о значимом различии в дискриминирующей мощности.

Оценка согласованности между данными на уровне внутривидовых типов и близкородственных внутривидовых групп/комплексов/кластеров, полученных различными методами, представлена в таблицах 20 и 21. В целом, согласованность результатов на уровне типов не превышает 50%, за исключением в паре PFGE-MLVA (64%) и cgMLST-PFGE (74%), что отражает низкую конкордантность результатов различных методов типирования, не позволяя предсказать с достаточно высокой вероятностью на основании данных, полученных одним методом типирования, данные для другого метода типирования. В то же время, следует отметить, что группирование близкородственных типов в группы/комплексы/кластеры позволило достичь более высоких показателей согласованности результатов, достигнув в некоторых случаях 80% и более (таблица 21).

Таблица 20 – Оценка конкордантности результатов, полученных различными методами типирования на уровне внутривидовых типов

Внутривидовые типы <i>S.aureus</i>	PFGE тип	MLVA тип	MLST тип	MVLST тип	Spa тип	cgMLST тип	wgMLST тип
PFGE тип	-	-	-	-	-	-	-
MLVA тип	0.641	-	-	-	-	-	-
MLST тип	0.207	0.298	-	-	-	-	-
MVLST тип	0.249	0.338	0.449	-	-	-	-
Spa тип	0.308	0.428	0.526	0.419	-	-	-
cgMLST тип	0.714	0.436	0.13	0.157	0.191	-	-
wgMLST тип	0.18	0.121	0.024	0.03	0.035	0.292	-
panGenome тип	0.207	0.14	0.027	0.034	0.041	0.046	0

Таблица 21 - Оценка конкордантности результатов, полученных различными методами типирования на уровне близкородственных групп/клональных комплексов/кластеров

Внутривидовые группы <i>S.aureus</i>	PFGE-группа	MLVA комплекс	MLST комплекс	MVLST комплекс	Spa группа	cgMLST кластер	wgMLST кластер
PFGE-группа	-	-	-	-	-	-	-
MLVA комплекс	0.297	-	-	-	-	-	-
MLST комплекс	0.198	0.613	-	-	-	-	-
MVLST комплекс	0.357	0.657	0.429	-	-	-	-
Spa группа	0.264	0.809	0.596	0.696	-	-	-
cgMLST кластер	0.411	0.668	0.43	0.853	0.698	-	-
wgMLST кластер	0.49	0.568	0.492	0.678	0.554	0.845	-
panGenome кластер	0.407	0.649	0.422	0.737	0.69	0.885	0.753

3.8 Параметры, влияющие на дискриминирующую способность метода типирования

Коллекция изолятов *S.aureus*, включенная в исследование. Одним из важных параметров, влияющих на показатели дискриминирующей мощности метода субвидового типирования микроорганизмов, является коллекция микроорганизмов, подвергшаяся типированию. Исследования, проведенные различными авторами на различных выборках изолятов, указывают на различающиеся параметры разрешающей способности методов типирования (таблица 22).

Таблица 22 – Оценка разрешающей способности методов типирования, использовавшихся в различных исследованиях

Описание выборки	n	PFGE	MLVA	SIRU	Spa	MLST
США 2008-2009 [213]	200	0,76 (0,7 - 0,82)	-	-	0,87 (0,83 - 0,92)	0,84 (0,79- 0,88)
Глобальная коллекция 1943-2006 [140]	104	0,998 (0,997- 1,0)	-	0,992 (0,987- 0,998)	0,976 (0,964 - 0,988)	0,917 (0,885 - 0,95)
Глобальная коллекция с 1957 года [214]	198	0,959 (0,949 - 0,969)	-	-	0,973 (0,963- 0,983)	0,948 (0,934 - 0,962)
Нидерланды 1988-2005 [203]	85	0,887 (0,847 - 0,926)	-	0,781 (0,71 - 0,852)	0,69 (0,627- 0,752)	0,831 (0,787- 0,875)
Нидерланды, Германия [215]	134	-	0,789 (0,735 - 0,844)	-	0,76 (0,701- 0,818)	-
Нидерланды 2005-2007 [148]	529	0,979 (0,975 - 0,983)	0,982 (0,979 - 0,985)	-	0,949 (0,94 - 0,959)	-
Казахстан 2014 [Наше исследование]	94	-	-	0,898 (0,869- 0,926)	0,923 (0,906 - 0,939)	-
Глобальная коллекция 1884-2014 [Наше <i>in silico</i> исследование]	116	0,991 (0,983- 0,998)	0,985 (0,976- ,995)	-	0,95 (0,930 - 0,970)	0,926 (0,9 - 0,952)

Как указано в таблице, значения дискриминирующей способности методов и их ранжирование по данному признаку меняется в зависимости от исследования. В то же время следует отметить, что общим недостатком является ограниченная коллекция изолятов микроорганизмов, которые подверглись внутривидовому типированию. Данное ограничение не позволяет провести значимое различие в дискриминирующей способности методов типирования и однозначно ранжировать по данному параметру.

С целью решения данной проблемы было решено провести *in silico* типирование, используя 6970 геномов *S.aureus*, доступных на момент исследования.

***In silico* типирование 6970 геномов *S.aureus*.** 6970 полногеномных последовательностей *S.aureus*, представленных в виде контигов, скаффолдов были получены с GenBank базы данных. В 5604 геномах *S.aureus* (80,4%) удалось успешно детектировать Spa типы. Анализ на определение MLST типов позволил определить сиквенс типы в 6912 геномах (99,2%). MVLST анализ полногеномных данных определил MVLST типы в 6882 геномах (98,7%). MLVA профили были успешно обнаружены в 6290 геномах (90,2%). cgMLST и wgMLST типы были определены для 5202 геномов *S.aureus* (74,5%), при условии включения локусов с количеством не более 0,5% отрицательных результатов (1332 локусов) в cgMLST и не более 5% отрицательных результатов для локуса (1714 локусов) в wgMLST схеме типирования. Таблица, для которой были успешно определены внутривидовые типы на основе Spa, MLST, MLVA, MVLST, cgMLST, wgMLST методов типирования включала 4976 геномов. После определения геномов, для которых были успешно определены основные внутривидовые типы, различными методами типирования, был проведен panGenome анализ. Решение о выполнении типирования на основе пангеномного анализа в последнюю очередь было принято исходя из ресурсоемкости и времени необходимого для выполнения данного исследования – пул, содержащий свыше 900 тысяч генов 4976 геномов, необходимо было определить по группам-ортологам. В результате, были сформированы 48578 групп-ортологов на основе 85 % cut-off схожести.

Анализ 4976 геномов *S.aureus* биоинформатическими методами позволил определить 251 MLST тип, 507 Spa типов, 924 MVLST типа, 1344 MLVA типа, 4057 cgMLST типов, 4221 wgMLST тип и 4970 panGenome типов. На основании полученных внутривидовых типов была рассчитана дискриминирующая мощность каждого метода типирования, которая представлена в таблице 23. Результаты согласованности между данными, полученные различными методами типирования, как на уровне внутривидовых типов, так и на уровне близкородственных групп/комплексов/кластеров показаны в таблице 24.

Таблица 23 – Дискриминирующая мощность методов субвидового типирования *S.aureus*

Метод типирования	Количество типов	Индекс Симпсона	95% ДИ
PanGenome	4970	1	(1,000-1,000)
wgMLST	4221	1	(1,000-1,000)
cgMLST	4057	1	(1,000-1,000)
MLVA	1344	0,979	(0,977-0,980)
MVSLT	924	0,919	(0,914-0,925)
Spa	507	0,907	(0,902-0,913)
MLST	251	0,854	(0,848-0,859)

Таблица 24 – Оценка согласованности результатов, полученных методами типирования *S.aureus*

Метод типирования	panGenome	wgMLST	cgMLST	MLST	MVLST	Spa*	MLVA
panGenome	0	0.88166480	0.88034520	0.66565030	0.78012950	0.65857880	0.38715910
wgMLST	0.00163121	0	0.95698500	0.73442500	0.85357340	0.72063950	0.42693130
cgMLST	0.00119544	0.84513270	0	0.77486160	0.85597420	0.76116170	0.45216510
MLST	0.00000847	0.00229015	0.00312795	0	0.82659780	0.90374400	0.61026850
MVLST	0.00001658	0.00395918	0.00546281	0.34239160	0	0.81855280	0.48064000
Spa	0.00001426	0.00340656	0.00466653	0.40535450	0.38802870	0	0.56137000
MLVA	0.00006620	0.00576403	0.00750489	0.17084670	0.16272570	0.31035170	0
Примечание - Левая нижняя часть таблицы указывает на согласованность между типами. Правая верхняя часть указывает на конкордантность между группами. * в случае со Spa типами, в формировании Spa групп участвовали Spa типы с более 3 повторами. В связи с исключением 42 геномов <i>S.aureus</i> с 3 и менее повторами в Spa типах из анализа, конкордантность между Spa группами с результатами других методов типирования оценивалась на 4934 геномах <i>S.aureus</i>							

Анализ, проведенной на большой выборке позволил получить более однозначную оценку разрешающей способности каждого метода типирования (на что указывает 95% доверительный интервал), что в свою очередь позволил провести ранжирование методов типирования исходя из дискриминирующей способности. Как и прежде, методы типирования, использующие полногеномные данные, такие как panGenome, wgMLST, cgMLST продемонстрировали наивысшую разрешающую способность. Типирование, основанное на выявлении различий в tandemных повторах (MLVA типирование) показало 98% дискриминирующую мощность. Разрешающая сила MVLSST и spa типирования была чуть выше 90% (92% и 90%, соответственно). Наименьшей дискриминирующей силой обладал метод MLST (85,5%).

Дискриминирующая способность маркера и их количество в схеме типирования. Ряд работ, включая наши исследования (см. таблицу 22) указывают, что наименьшей дискриминирующей способностью обладает MLST, несмотря на 7 маркеров, используемых в данном методе типирования. Анализ схемы MLST типирования показал, что причиной столь сравнительно невысокой разрешающей способности явилась высокая конкордантность между маркерами, в частности между *arcc*, *glpf*, *pta*, *yqio* локусами (таблица 25).

Таблица 25 – Межлокусная согласованность (конкордантность) в MLST схеме типирования *S.aureus*

Локусы в MLST схеме	<i>arcc</i>	<i>aroe</i>	<i>gmk</i>	<i>pta</i>	<i>tpi</i>
<i>Arcc</i>	-	-	-	-	-
<i>Aroe</i>	0.744	-	-	-	-
<i>Glpf</i>	0.135	0.26	-	-	-
<i>Gmk</i>	0.706	0.874	-	-	-
<i>Pta</i>	0.746	0.928	0.876	-	-
<i>Tpi</i>	0.75	0.898	0.847	0.927	-
<i>Yqio</i>	0.599	0.77	0.716	0.786	0.784

Высокая межлокусная конкордантность означает, что при добавлении каждого такого локуса в схему типирования не наблюдается значительное увеличение разрешающей способности метода в целом. Был проведен анализ упрощенной схемы MLST, включающей 4 локуса (*arcc*, *glpf*, *pta*, *yqio*) предварительно исключив локусы с высокой межлокусной конкордантностью. Индекс Симпсона данной MLST схемы достиг показателя 0.838 (95%ДИ; 0,832-0,843), что сопоставимо с классической схемой MLST типирования, включающей 7 маркеров – 0,854 (95%ДИ; 0,848-0,859). Более того, согласованность, оцененная на скорректированном показателе Ранда (adjusted Rand) между MLST схемами с 4 и 7 локусами была: 94% на уровне типов и 95% на уровне клональных комплексов.

Таким образом, высокая межлокусная согласованность в схеме типирования является одним из факторов, влияющим на разрешающую силу метода в целом.

Результаты межлокусной конкордантности MLVA и MVLST схем типирования представлены в таблицах 25 и 26.

Таблица 26 – Межлокусная согласованность MLVA схемы типирования *S.aureus*

Локусы в MLVA схеме	Локус 9	Локус 21	Локус 24	Локус 61_1	Локус 61_2	Локус 63	Локус 67
Локус 21	0,138	-	-	-	-	-	-
Локус 24	0,154	0,011	-	-	-	-	-
Локус 61_1	0,276	0,258	0,194	-	-	-	-
Локус 61_2	0,335	0,596	0,146	0,447	-	-	-
Локус 63	0,236	0,382	0,079	0,269	0,329	-	-
Локус 67	0,332	0,268	0,252	0,401	0,452	0,341	-
Локус 81	0,172	0,117	0,092	0,209	0,234	0,249	0,336

Таблица 27 – Межлокусная согласованность MVLST схемы типирования *S.aureus*

Локусы в MVLST схеме	alt	essC	geh	hlga	htra
Essc	0,522	-	-	-	-
Geh	0,725	0,64	-	-	-
Hlga	0,656	0,578	0,703	-	-
Htra	0,804	0,554	0,751	0,69	-
Sdrc	0,699	0,574	0,712	0,759	0,734

В схеме MLVA типирования наблюдается в среднем низкая межлокусная конкордантность 0.261 (95%ДИ; 0,211 – 0,311), в то время как усредненная межмаркерная согласованность в методе MVLST составила 0,673 (95%ДИ; 0,627 – 0,72). Несмотря на большее количество локусов в MLVA типировании (n =8) в сравнении с MVLST схемой (n=6), а также более низкой межлокусной согласованностью, дискриминирующая мощность MLVA обладает лишь на 6% выше дискриминирующей мощностью (97,9% для MLVA против 91,9% для MVLST). Таким образом, помимо межлокусной конкордантности, на дискриминирующую способность метода типирования в целом оказывает влияние разрешающая сила отдельно взятого маркера, входящего в схему типирования. Для MLVA схемы усредненный показатель разрешающей способности локусов составил 0,694 (95% ДИ; 0,609 – 0,774), в то время как для MVLST этот показатель был выше 0,827(95% ДИ; 0,788-0,867).

Схематическое изображение факторов, влияющих на дискриминирующую мощность метода, представлено на рисунке 30.

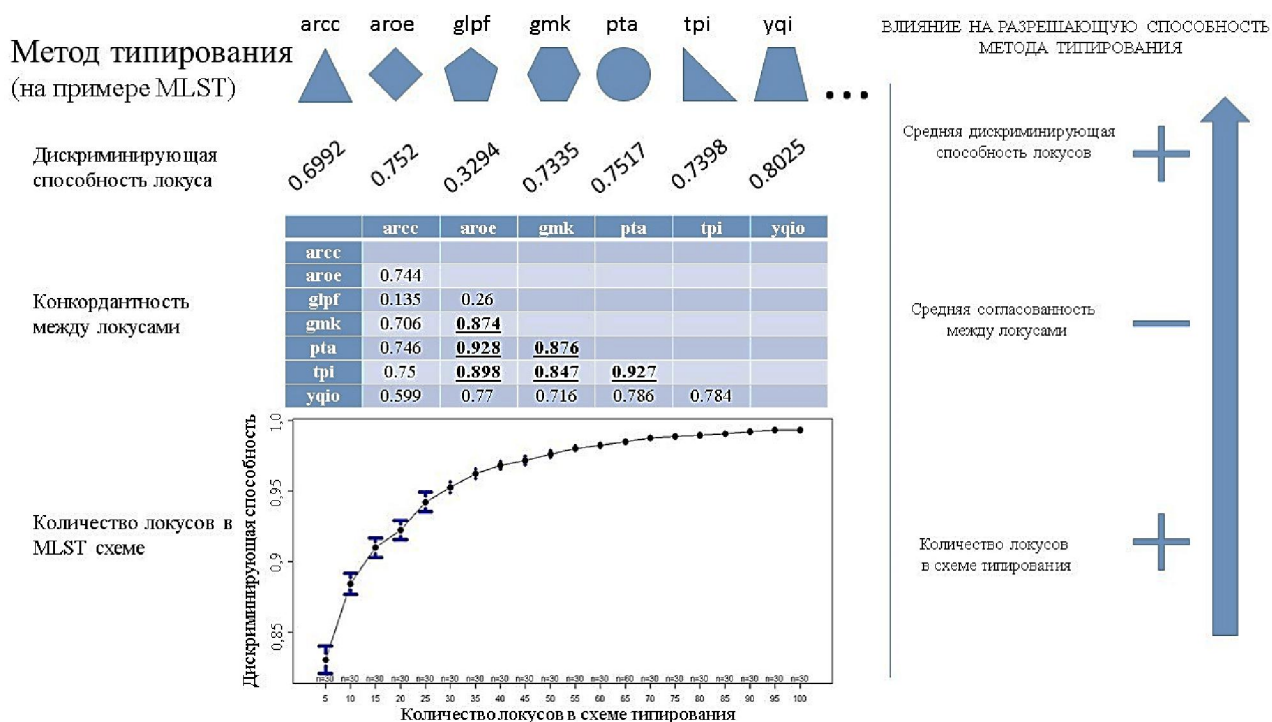
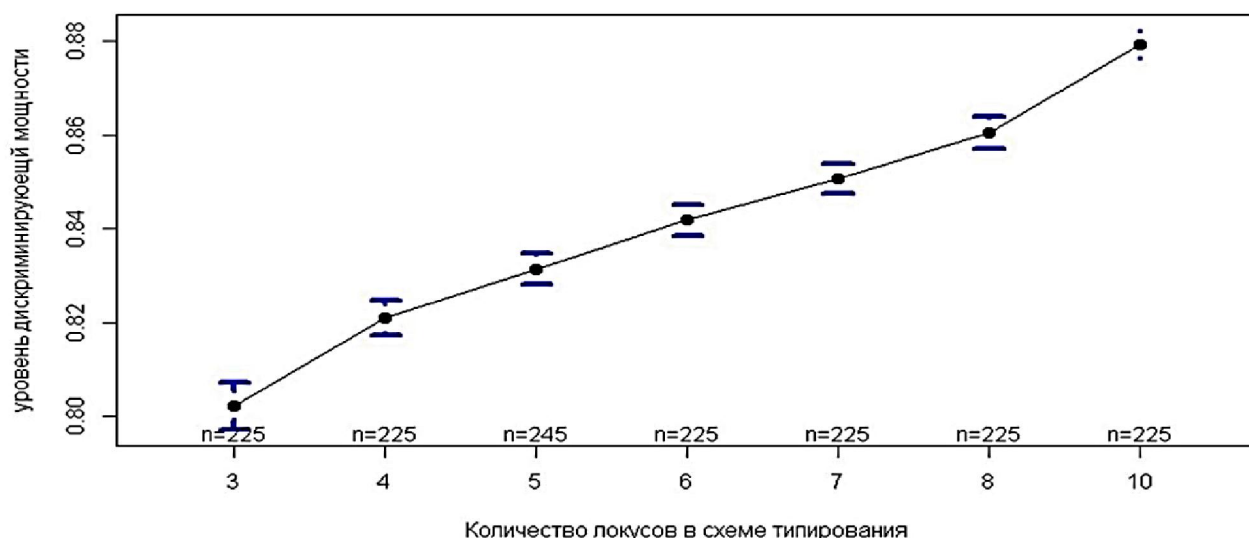


Рисунок 30 – Параметры, определяющие разрешающую мощность метода типирования

Популярность некоторых методов типирования *S.aureus*, таких как Spa, MLST, MVLST зависит от ряда параметров, в том числе благодаря хорошим показателям разрешающей способности при использовании ограниченного количества маркеров в схеме типирования [216-218]. Было проведено исследование на определение зависимости разрешающей способности с использованием ограниченного количества локусов (от 3 до 10 генов), взятых из схемы sgMLST с целью получения уровня дискриминирующей силы метода в целом от количества локусов, их разрешающей способности, а также уровня конкордантности между ними. В результате была получена линейная зависимость разрешающей силы метода типирования от всех трех параметров. Полученная модель позволяет объяснить такую зависимость на 76% (скорректированный коэффициент детерминации $R^2 = 0,762$), при этом все параметры, включенные в данную модель, имели р-значение ниже 0,001, что позволяет с большей уверенностью сделать заключение о влиянии предикторов (количество локусов, средняя межлокусная согласованность, средняя разрешающая способность локусов) на уровень разрешающей способности метода в целом (рисунок 31).

Оценка относительного вклада каждого из трех параметров в данную модель показана на рисунке 32. Наибольшим влиянием на дискриминирующую способность метода в целом обладает параметр «количество локусов» с

вкладом в 54,4%. Вторым по уровню относительного влияния на дискриминирующую мощность метода типирования являлся усредненный показатель разрешающей способности локусов, входящих в метод субвидового типирования (28,8%). На долю третьего параметра «межлокусная конкордантность» пришлось 16,8%.



$$SID = 0.482 + 0.01N(loci) - 0.286 * \frac{\sum_1^{(N(loci) * (N(loci) - 1)) / 2} adj.Rand}{(N(loci) * (N(loci) - 1)) / 2} + 0.696 * \frac{\sum_1^{N(loc)} SI}{N(loci)},$$

где

SID – *Simpson Index Diversity* (Дискриминирующая сила метода типирования;

N(loci) – количество локусов в схеме типирования;

$\frac{\sum_1^{(N(loci) * (N(loci) - 1)) / 2} adj.Rand}{(N(loci) * (N(loci) - 1)) / 2}$ – усредненный показатель конкордантности между локусами на основе скорректированного параметра Ранда (*adj.Rand*);

$\frac{\sum_1^{N(loc)} SI}{N(loci)}$ – усредненный показатель дискриминирующей силы локусов, входящие в схему типирования (*SI* – индекс Симпсона)

Рисунок 31 – Математическая модель линейной зависимости разрешающей мощности метода от трех параметров при использовании ограниченного количества локусов

Следует отметить, что принципиально важным моментом является тот факт, что данная модель обладает предсказательной способностью в случае с методом типирования *S.aureus*, который имеют в своей схеме небольшое количество локусов (маркеров), как правило, от 3 до 10. Использование данного диапазона количества маркеров в методе типирования позволяет достичь уровня разрешающей мощности метода от 80 до 88%. В то же время, в качестве исключения, была обнаружена комбинация схемы типирования из 3 локусов, позволившая достичь разрешающую способность в 90% при показателях

межлокусной конкордантности 40% и разрешающей способности локусов в среднем 73% (таблица 28).

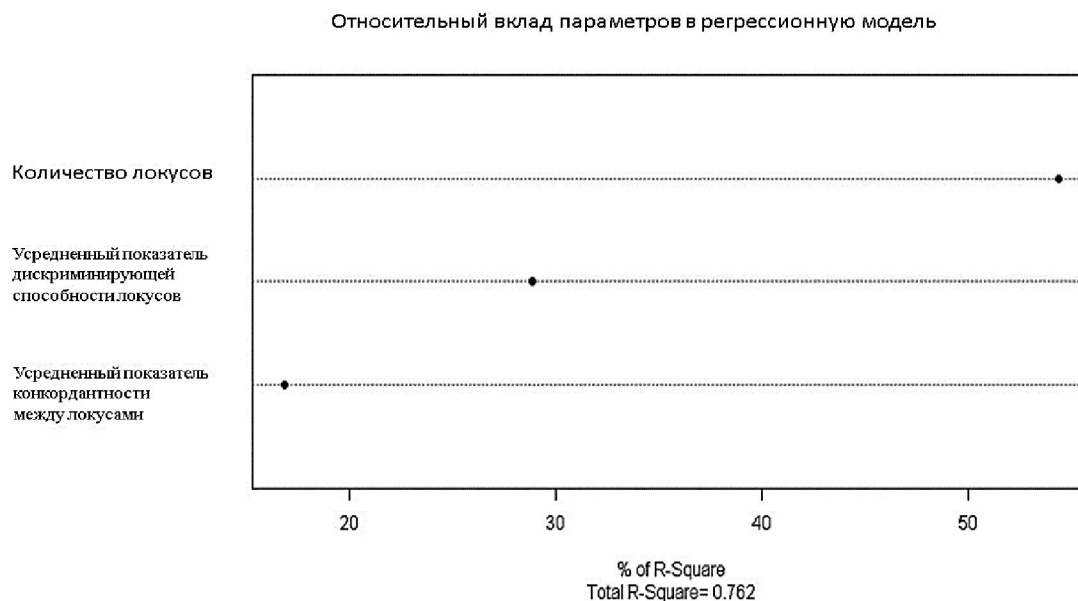


Рисунок 32 – Относительный вклад каждого параметра в регрессионную модель

Таблица 28 – Локусы cgMLST, позволившие достичь 90% уровня дискриминирующей способности метода

Наименование локуса	Ген	Длина гена (п.о.)	Наименование продукта гена	Количество аллелей в базе Ridom
SACOL1075	purC	705	phosphoribosylaminoimidazole-succinocarboxamide synthase	104
SACOL1390	parC	2403	DNA topoisomerase IV subunit A	350
SACOL2440	-	807	hypothetical protein	138

При включении в схему типирования большего количества (сотни, тысячи) локусов/маркеров/генов зависимость разрешающей способности метода в целом от количества маркеров, их дискриминирующей силы и конкордантности приобретает логарифмический характер (рисунок 33).

Как показывает кривая зависимости, 90% разрешающая способность метода достигается в среднем при включении 8 локусов в схему типирования, хотя существуют схемы из 3-7 локусов, позволяющие получать индекс Симпсона 0,9. Дискриминирующая способность в 95%, сопоставимая со Sra и MVLST типированием, достигается при использовании 25 локусов. Для достижения 99% дискриминирующей мощности метода, схожей с таковой при

использовании MLVA и PFGE методов типирования, потребуется включение 65-85 локусов в схему типирования. 99,9% разрешающая сила метода типирования возможна при использовании 200-250 локусов.

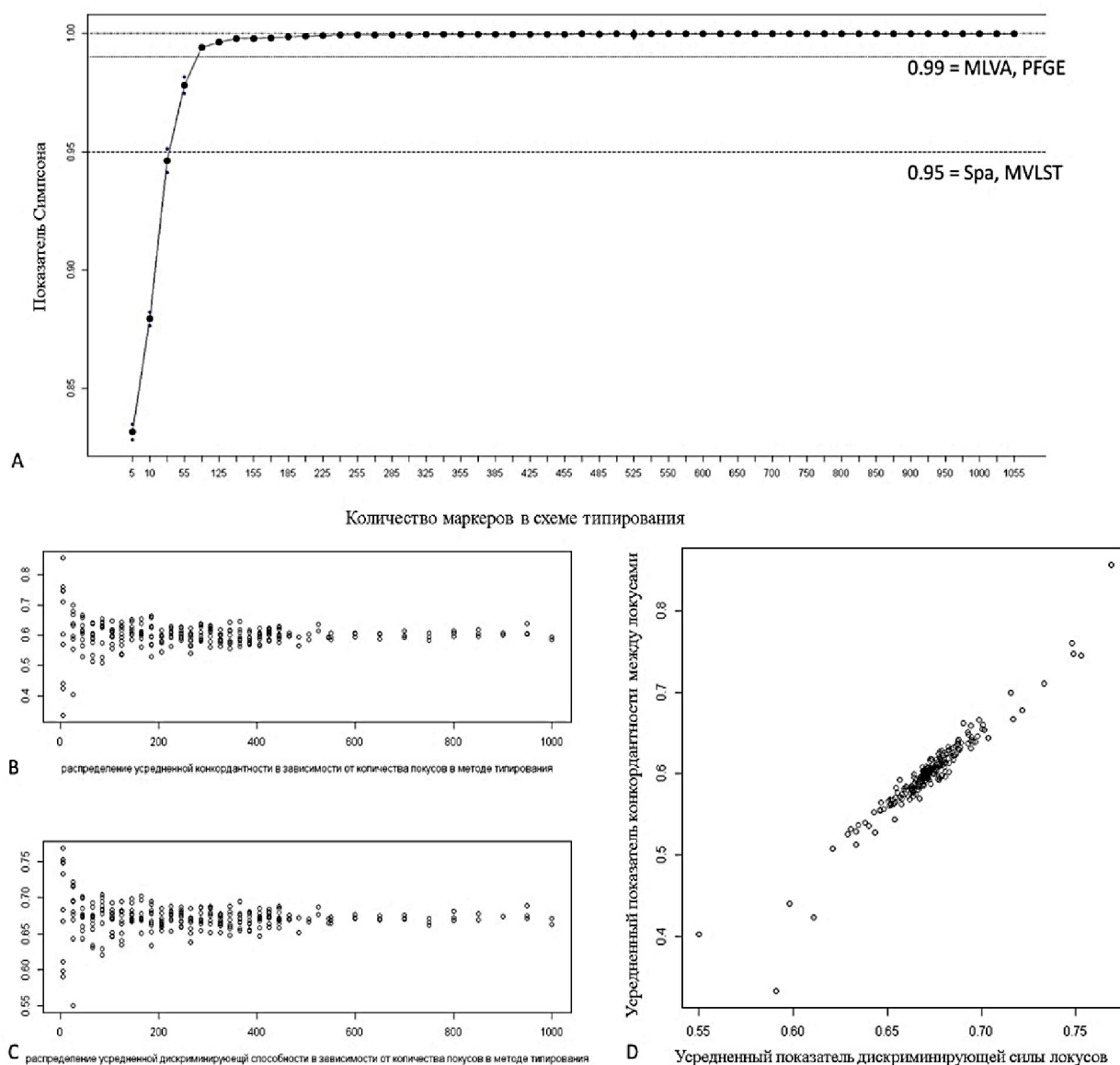


Рисунок 33 – Логарифмический характер зависимости дискриминирующей способности от 3 предикторов

Последующее увеличение количества локусов не позволяет увеличить разрешающую способность метода, даже включив 1000 маркеров. В нашем случае лишь sgMLST схема типирования, состоящая более чем из 1300 локусов, позволила достичь 100% разрешающую силу, оцененную индексом Симпсона. При этом следует отметить, что с увеличением количества локусов, достигая 1000 маркеров, средний уровень межлокусной конкордантности стремится к 0,627 (95% ДИ; 0,626-0,628) показателю скорректированного индекса Ранда, а усредненный показатель индекса Симпсона для локусов, входящих в схему типирования – 0,671 (95%ДИ; 0,661-0,681) (рисунок 33 В, С).

Более того, наблюдается прямая зависимость между показателями конкордантности и дискриминирующей способностью локусов в схеме типирования (рисунок 33 D). Данный факт означает, что между локусами с высокой разрешающей способностью наблюдается высокая согласованность.

С целью установления доли генов, обладающих высокой разрешающей способностью, а также уровня конкордантности между ними, был проведен анализ всех локусов, входящих в cgMLST схему типирования. Как показало исследование, в геноме *S.aureus* более половины генов (57,2%) являются гипервариабельными с 75% и выше индексом Симпсона. Дальнейший анализ показал, что между данными генами наблюдается высокий уровень конкордантности. В 75% генах с высоким показателем Симпсона, конкордантность достигла 80% и выше (рисунок 34).

Таким образом, модель зависимости дискриминирующей способности метода, включающего сотни маркеров, указывает на логарифмический характер, которая достигает 99,9% при использовании 200-250 локусов в схеме типирования.

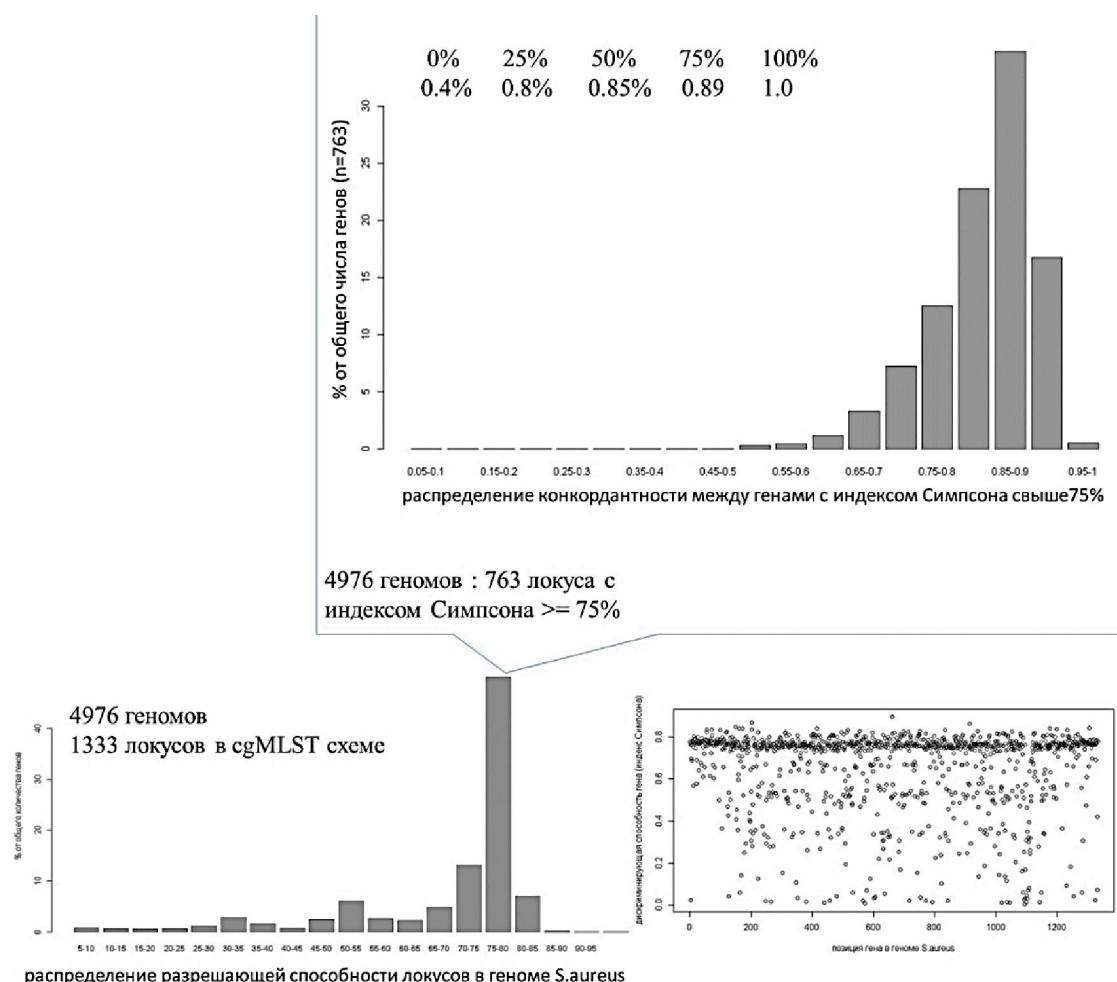


Рисунок 34 – Показатели дискриминирующей мощности генов, входящих в геном *S.aureus*, а также уровень конкордантности между ними

Математическая модель зависимости дискриминирующей способности метода, включающего сотни/тысячи маркеров, представлена на рисунке 35. Скорректированный коэффициент детерминации (adj.R^2) данной прогностической модели составил 95,4%, объясняя явную зависимость разрешающей силы метода от 3 параметров. Значения p для всех трех параметров составило менее 0,05. Исходя из данной модели, уровни относительного влияния предикторов изменились. Доля параметра «количество локусов» стала более определяющей (85,9%). На долю усредненного показателя межлокусовой конкордантности пришлось около (13,9%).

$$SID = 0.8822 + \lg(\lg(0.3201 * N(\text{loci}))) - 0.0001 * \frac{\sum_{i=1}^{N(\text{loci})} \frac{(N(\text{loci}) * (N(\text{loci}) - 1)) / 2}{adj.Rand} + 0.0239 * \frac{\sum_{i=1}^{N(\text{loci})} SI}{N(\text{loci})},$$

где

SID – Simpson Index Diversity (Дискриминирующая сила метода типирования);

$N(\text{loci})$ – количество локусов в схеме типирования;

$\frac{\sum_{i=1}^{N(\text{loci})} \frac{(N(\text{loci}) * (N(\text{loci}) - 1)) / 2}{adj.Rand}}{(N(\text{loci}) * (N(\text{loci}) - 1)) / 2}$ – усредненный показатель конкордантности между локусами на основе скорректированного параметра Панда (adj.Rand);

$\frac{\sum_{i=1}^{N(\text{loci})} SI}{N(\text{loci})}$ – усредненный показатель дискриминирующей силы локусов, входящие в схему типирования (SI – индекс Симпсона)

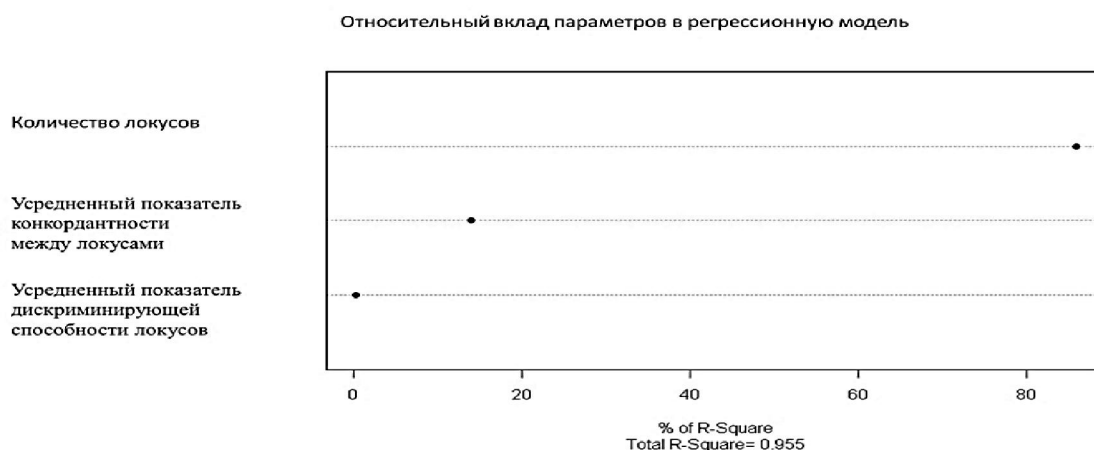


Рисунок 35 – Математическая модель логарифмической зависимости разрешающей мощности метода от трех параметров при использовании большого количества локусов, а также относительный вклад каждого из параметров

Доля усредненного показателя разрешающей способности локусов, входящих в предсказательную модель составила 0,2%. Такое изменение в относительной роли данных параметров в предсказании разрешающей мощности метода в целом отчасти объясняется тем фактом, что увеличение числа маркеров в схеме типирования приводит к снижению вариации, двух параметров, таких как средние показатели конкордантности и дискриминирующей способности локусов, входящих в метод типирования.

Таким образом, на величину дискриминирующей способности метода в целом, в первую очередь, влияет параметр «количество локусов». При использовании небольшого количества локусов наблюдается линейная

зависимость, в то время как, добавление большего количества локусов в схему типирования (десятки и сотни) приводит к снижению тенденции увеличения разрешающей мощности метода в целом, достигнув максимального значения индекса Симпсона 0,999 при использовании 200-250 локусов в схеме типирования. Дальнейшее увеличение маркеров не приводит к увеличению показателя Симпсона, что, вероятно, объясняется включением в схему типирования все большего количества локусов с высокой межлокусной конкордантностью, доля которых в геноме составляет 57,2%, что, в свою очередь, не позволяет значительно увеличить разрешающую способность метода в целом.

3.9 Параметры, влияющие на показатель конкордантности между данными, полученными различными методами типирования

Как было показано ранее в наших исследованиях, а также в работах некоторых авторов [109 с.91; 219, 220] группирование близкородственных типов *S.aureus* в комплексы/кластеры может привести к повышению согласованности результатов типирования, полученных разными методами. На практике процесс формирования групп осуществляется кластеризацией, которая в свою очередь состоит из нескольких этапов и зависит от некоторых параметров. Стадии формирования групп представлены на рисунке 36.

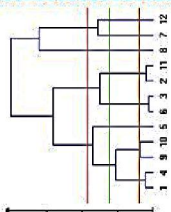
Тип данных	Бинарные (наличие/отсутствие маркера) (пример: PFGE, AFLP, RAPD)	Числовые (ко-доминирующие маркеры) (например, microstallites, allozymes, MLST)
Расчет матрицы генетических дистанций (коэффициенты сходства/различий)	Simple Matching $S_{SM} = \frac{a+d}{a+b+c+d}$ Jaccard (1901) $S_J = \frac{a}{a+b+c}$ Dice (1945), Sorensen (1948) $S_S = \frac{2a}{2a+b+c}$	Euclidean distance $D_{eu} = \sqrt{\sum (X_i - Y_i)^2}$ Rogers distance (1972) $D_R = \sqrt{\frac{\sum (X_i - Y_i)^2}{2}}$ Nei distance (1983) $D_N = 1 - \sum \sqrt{X_i Y_i}$
Методы кластеризации	Иерархическая кластеризация Complete-linkage $D(c_1, c_2) = \max_{x_1 \in c_1, x_2 \in c_2} D(x_1, x_2)$ Single-linkage $D(c_1, c_2) = \min_{x_1 \in c_1, x_2 \in c_2} D(x_1, x_2)$ Average linkage (UPGMA) $D(c_1, c_2) = \frac{1}{ c_1 + c_2 } \sum_{x_1 \in c_1, x_2 \in c_2} D(x_1, x_2)$ Ward $TD_{c_1 \cup c_2} = \sum_{x \in c_1 \cup c_2} D(x, \mu_{c_1 \cup c_2})^2$	Деревья минимальных дистанций / (go)leBurst алгоритм Метод k-средних $V = \sum_{i=1}^k \sum_{x_j \in S_i} (x_j - \mu_i)^2$ Метод четкой кластеризации С-средних $E = \sum \sum \mu_j^m \cdot \ x_j - c_j\ ^2$ Анализ главных компонент Дискриминантный анализ главных компонент
Определение групп	Выбор значение Cut-off на дендрограмме 	Правило формирования групп 1) Максимальная дистанция в группе 2) Минимальное количество (изолятов/генотипов в группе)

Рисунок 36 – Схема определения внутривидовых групп микроорганизмов

Данные, получаемые при использовании современных молекулярно-генетических методов типирования, в основе которых лежат различные

методические принципы, можно разделить на 2 группы – бинарные (PFGE, panGenome) и буквенно-числовые (SNP, MLST, MLVA). Исходя из типа данных происходит выбор метода расчета генетических дистанций, являющийся следующим этапом формирования кластеризации.

Выбор метода расчета матрицы генетических дистанций (индекс сходства/различия). На текущий момент существует множество предложенных коэффициентов для расчета генетических дистанций (матрица сходства/различий). David Mayer и Christian Buchta в своей работе указывают на более 20 индексов расчетов матрицы генетических дистанций для бинарных данных [221].

Для определения влияния использования различных методов расчет генетических дистанций, были использованы 6 различных алгоритмов, включая популярные Jaccard и Dice коэффициенты [160, с.5486; 222, 223] анализа бинарных данных, полученных в результате *in silico* PFGE типирования.

Корреляционный анализ рассчитанных матриц показал хорошую корреляцию между Gower и Simple matching ($R^2=1$), Cosine и Dice ($R^2=0,99$), Dice и Jaccard ($R^2=0,98$), Cosine и Jaccard ($R^2=0,97$). В то же время в некоторых случаях наблюдается невысокая корреляция между данными (рисунок 37).

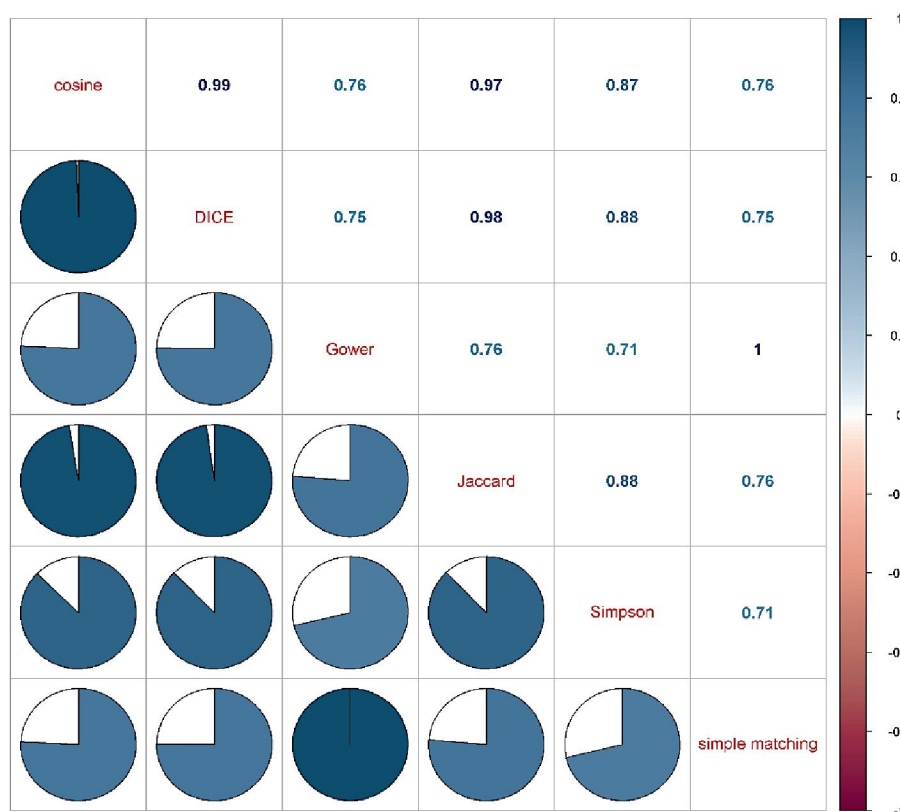


Рисунок 37 – Корреляционная таблица матриц генетических дистанций, рассчитанных различными коэффициентами

На основании полученных результатов, представленные в виде матрицы генетического сходства, рассчитанные различными алгоритмами, штаммы

S.aureus группировались на дендрограмме, используя complete-linkage кластеризацию. В качестве порогового значения включения в группу использовался показатель 20% cut-off различия. Результаты кластеризации представлены на рисунке 38.

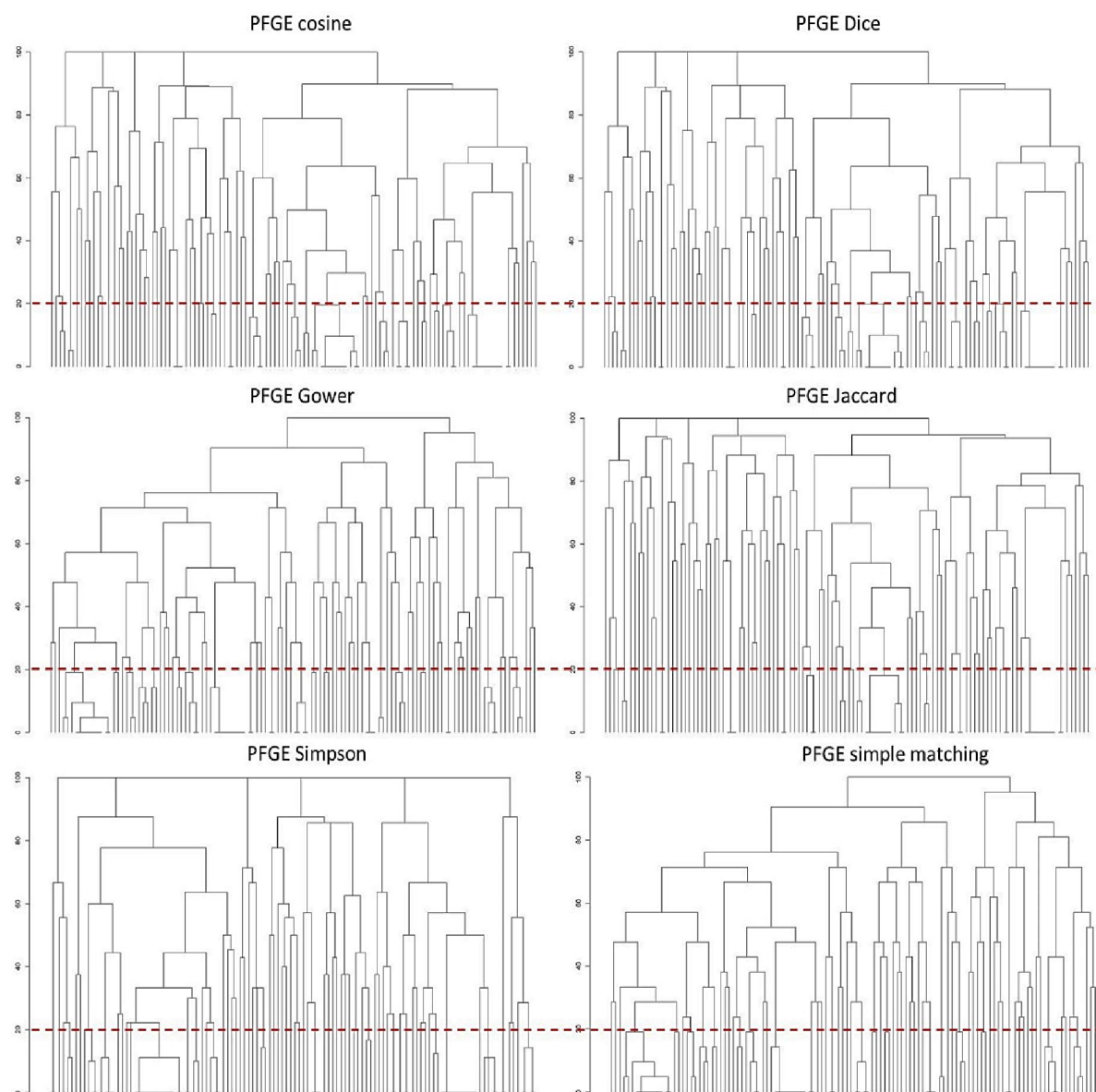


Рисунок 38 – Кластеризация и группирование данных PFGE типирования *S.aureus*, используя различные методы расчета генетических дистанций

Результаты конкордантности, представленные в таблице 29, показывают, что в некоторых случаях согласованность результатов PFGE типирования при использовании различных методов расчета матрицы генетических расстояний может снижаться до 78%, указывая на достаточно большое различие в формировании PFGE групп, используя различные алгоритмы расчетов. Верхне-правые значения обозначают среднее значение относительных различий в длинах ветвей дендрограмм; нижне-левые значения – совпадение/отличие в топологии дендрограмм.

Таблица 29 – Результаты согласованности данных PFGE типирования, используя различные методы расчета матрицы генетических расстояний

Алгоритмы генетических дистанцы	Cosine	DICE	Gower	Jaccard	Simpson
DICE	0.996	-	-	-	-
Gower	0.960	0.964	-	-	-
Jaccard	0.802	0.798	0.778	-	-
Simpson	0.982	0.986	0.950	0.784	-
Simple matching	0.960	0.964	1.000	0.778	0.950

Сравнительный анализ дендрограмм, полученных при использовании различных алгоритмов расчета матриц генетических дистанций показал различия как в топологии, так и в длине ветвей дендрограмм, что в свою очередь привело к вариации включения изолятов в группы (таблица 30).

Таблица 30 - Сравнительная оценка дендрограмм, полученных при использования различных методов расчета генетических дистанций

Алгоритмы генетических дистанцы	Cosine	Dice	Gower	Jaccard	Simpson	Simple matching
Cosine	-	0.008	0.115	0.319	0.113	0.1158
Dice	отл.	-	0.117	0.317	0.118	0.1174
Gower	отл.	отл.	-	0.462	0.18	совп.
Jaccard	отл.	совп.	отл.	-	0.322	0.316
Simpson	отл.	отл.	отл.	отл.	-	0.098
Simple matching	отл.	отл.	совп.	отл.	отл.	-

Примечание - Верхне-правые значения обозначают усредненные показатели различий в длинах ветвей дендрограмм; ниже-левые – совпадение/различие топологий дендрограмм

Таким образом, выбор метода расчета генетического расстояния между изолятами микроорганизмов на основании данных методов типирования, в конечном счете, может повлиять на конкордантность результатов типирования.

Выбор метода кластеризации. Кластерный анализ представляет собой пошаговый процесс последовательного объединения объектов (в данном случае изолятов микроорганизмов) на основе их сходства, определенного формулой расчета генетического сходства/различия, в группы/кластеры. Результат этого процесса, как правило, изображается в виде дендрограммы. На данный момент существуют различные алгоритмы кластеризации, включая такие как Ward, UPGMA и eBURST [159, с.1524; 224, с.240; 225, с.1424].

Влияние использования 6 различных методов расчета кластеризации, таких как полной связи (complete-linkage), Уорда (Ward), наименьшей связи (single-linkage), eBURST, и методом попарного невзвешенного среднего (unweighted pair group method with arithmetic mean, UPGMA) с последующим построением дендрограммы, было проанализировано с использованием данных MLVA типирования 116 штаммов *S.aureus*, полученных *in silico* (рисунок 39).

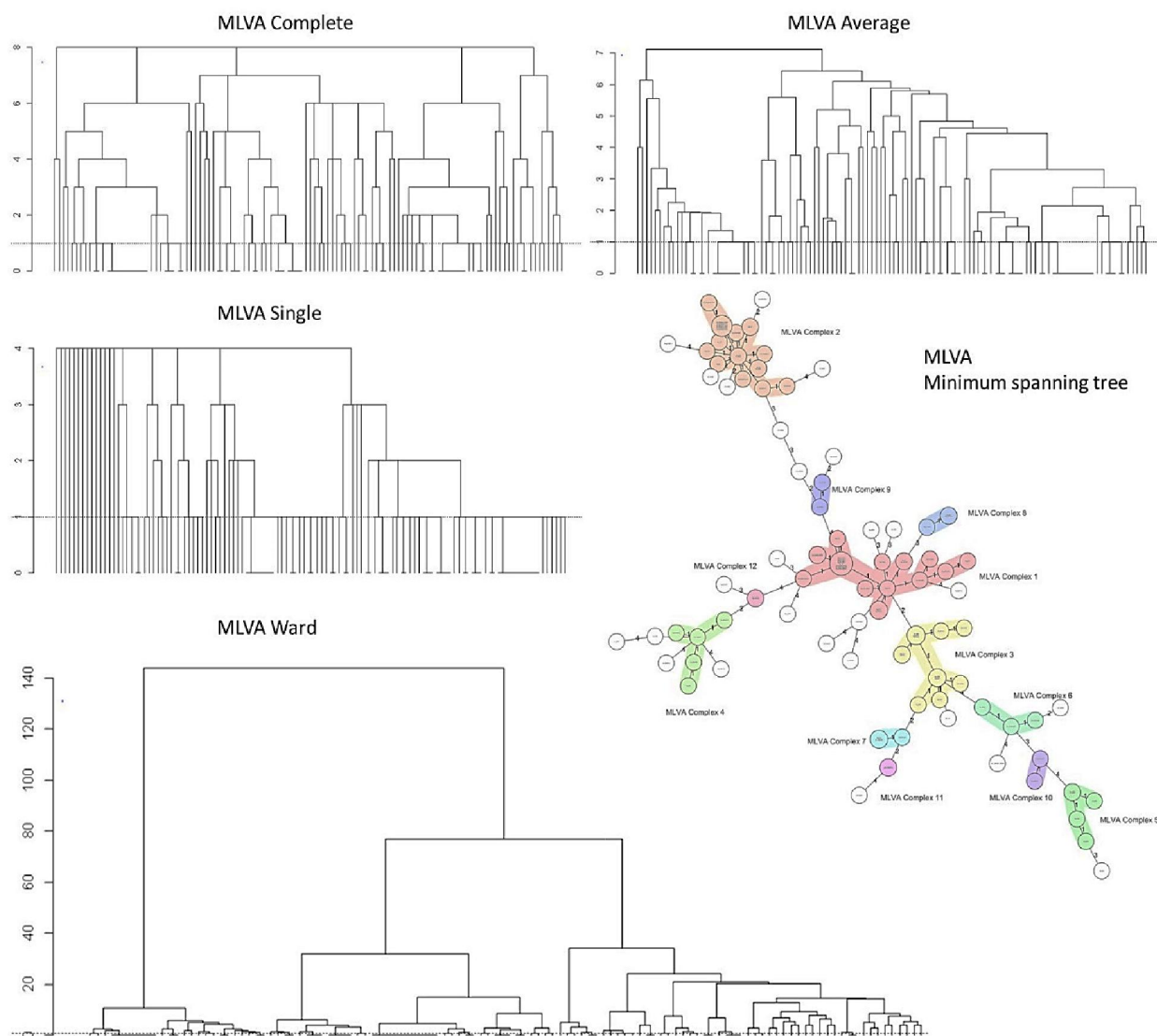


Рисунок 39 – Различные методы кластеризации штаммов *S.aureus* на основе данных MLVA типирования

Результаты конкордантности полученных данных представлены в таблице 31. Данные, представленные в таблице, указывают на вариацию конкордантности от 25% до 100%. Сравнительный анализ дендрограмм, полученных различными методами кластеризации данных MLVA типирования также показал различия как в топологии, так и длине ветвей (таблица 32).

Таблица 31 – Согласованность результатов MLVA типирования *S.aureus* при использовании различных алгоритмов кластеризации

Методы кластеризации	MST	Ward	Average	Single
Minimum spanning tree (eBurst)	-	-	-	-
Dendrogram(ward clustering)	0,246	-	-	-
Dendrogram(average clustering)	0,450	0,638	-	-
Dendrogram(single clustering)	1,000	0,246	0,450	-
Dendrogram(complete clustering)	0,450	0,638	1,000	0,450

Таблица 32 – Сравнительная оценка дендрограмм, полученных при использовании различных алгоритмов кластерного построения дендограм

Методы кластеризации	Ward	Average	Single	Complete
Dendrogram(ward clustering)		0,184	0,416	1,71
Dendrogram (average clustering)	отл.		0,287	2,11
Dendrogram(single clustering)	отл.	отл.		3,28
Dendrogram(complete clustering)	отл.	отл.	отл.	
Примечание - Верхне-правые значения обозначают усредненные показатели различий в длинах ветвей дендрограмм; ниже-левые – совпадение/различие топологий дендрограмм				

Таким образом, при использовании различных алгоритмов кластеризации данных методов типирования с последующей визуализацией на дендрограмме возникают изменения конкордантности между полученными данными вследствие различий в длинах ветвей дендрограммы, а также топологии, что в свою очередь при выбранном уровне cut-off приведет к формированию различных групп.

Cut-off значение при формировании близкородственных групп /комплексов/кластеров. Завершающим этапом определения внутривидовых близкородственных групп после выполнения кластеризации является определение правил или значения cut-off. На влияющий характер выбора показателя cut-off на конкордантность результатов указывается в работе [226].

Было проанализировано изменение параметра конкордантности между данными cgMLST и wgMLST типировании исходя из различных показателей cut-off (рисунок 40).

При установлении 10% cut-off для формирования кластеров согласованность результатов между cgMLST и wgMLST методами типирования составила 84%. В то же время, при изменении значений 10% cut-off для cgMLST и 9% cut-off для wgMLST конкордантность достигается 95%. В случае с 13% и 11% значениями cut-off для cgMLST и wgMLST наблюдается 100% согласованность результатов.

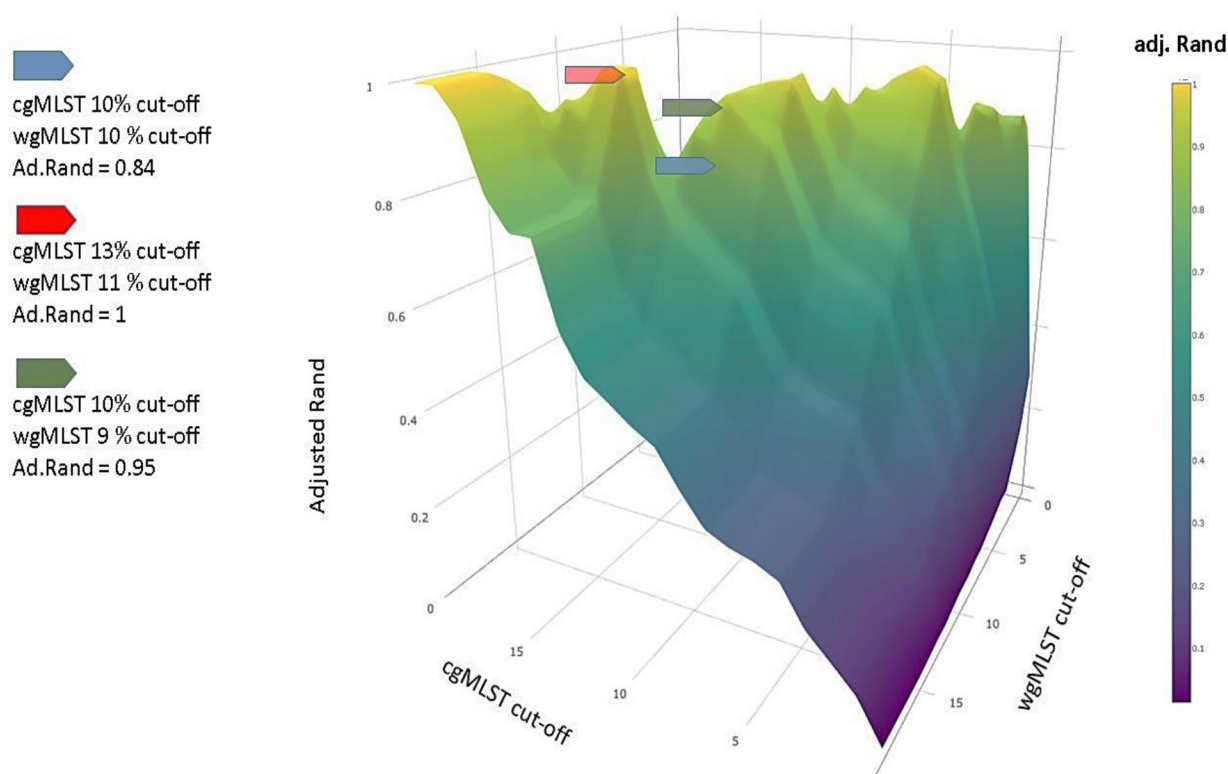


Рисунок 40 – 3D поверхность, отражающая изменение показателей конкордантности от изменения параметров cut-off

Таким образом, показатели cut-off, которые позволяют сформировать внутривидовые близкородственные группы на дендрограмме, являются одним из главных параметров, влияющих на уровень согласованности между данными, полученными различными методами типирования.

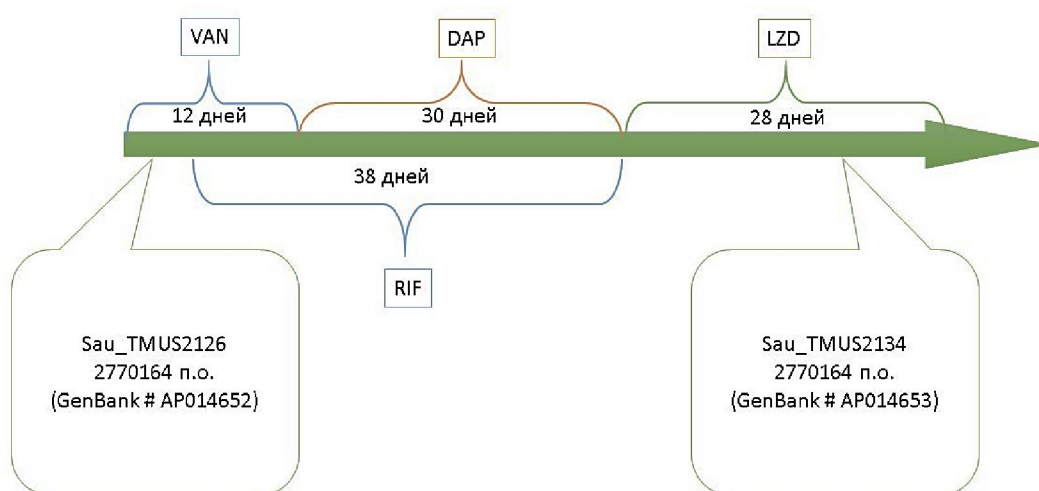
3.10 Оценка возможностей генотипических методов на клинических штаммах *S.aureus* с фенотипическими изменениями за 2 – 7-месячный период

В GenBank NCBI на момент исследования (август 2016 г.) находилось 116 полных, завершенных и аннотированных геномов *S.aureus*. Для 10 изолятов *S.aureus* описаны клинические случаи, в которых указывается изменение фенотипических свойств, таких как появление устойчивости на фоне лечения антимикробными препаратами или появление факторов вирулентности при переходе инфекционного процесса из острой стадии в хроническую. Отмечается, что данные свойства микроорганизмов появились спустя несколько месяцев. Авторы данных исследований решили провести полногеномное секвенирование с целью установления изменений в геноме, повлекших появление фенотипической изменчивости.

Геномные последовательности данных изолятов *S.aureus* были проанализированы *in silico* с целью установления внутривидовых типов, определенных различными методами типирования и определения

возможностей методов типирования в способности провести различие штаммов, с приобретенными фенотипическими особенностями за 2 – 7-месячный период.

Случай 1. Tetsuo Yamaguchi с соавт. провели полное секвенирование 2 штаммов, которые были выделены от пациента с операцией на аорте, после которой возникло воспаление в средостении [227]. Бактериологическое исследование определило, что этиологическим агентом является золотистый стафилококк, обладающий резистентностью к метициллину (MRSA). На этапе начала лечения был выделен штамм *S.aureus*, который обладал чувствительностью к даптомицину. Спустя 2 месяца был получен изолят *S.aureus*, который обладал устойчивостью к даптомицину. Схема лечения пациента в послеоперационный период представлена на рисунке 41.



VAN – ванкомицин, DAP – даптомицин, RIF – рифампицин, LZD – линезолид

Рисунок 41 – Курс антибактериальной терапии пациента, на фоне которой у *S.aureus* развилась резистентность

Авторы провели полногеномное секвенирование данных изолятов, с целью изучения изменения в геноме, ассоциированные с возникновением резистентности к даптомицину за столь небольшой промежуток времени. В результате сравнения двух геномов, анализ показал различие в 4 нуклеотида.

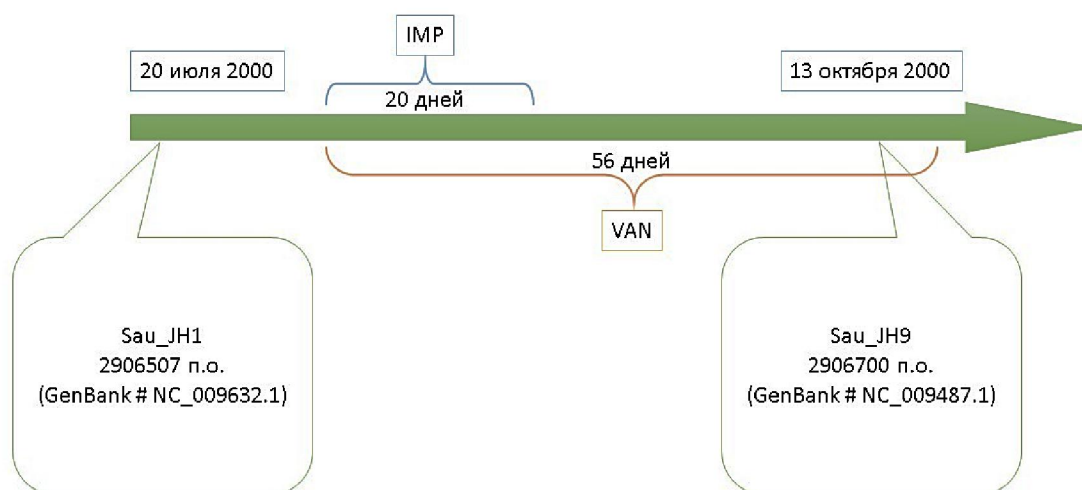
In silico исследование данных геномов *S.aureus* с целью определения достаточности разрешающей способности различных методов типирования для двух штаммов представлено в таблице 33.

В данном случае, штаммы золотистого стафилококка, с изменениями в 4 нуклеотида в геноме, накопившиеся за 2 месячный период смогли быть различены лишь с использованием методов cgMLST и wgMLST типирования. Остальные методы внутривидового типирования не смогли провести разделения данных изолятов *S.aureus*.

Таблица 33 – Сравнение результатов типирования двух изолятов *S.aureus*

Штамм	Страна выделения	Год	MLST тип	Sra тип	MVLST профиль	MLVA профиль	PFGE тип	Cg MLST тип	Wg MLST тип	Pan Genome тип
Sau_TMUS2126	Корея	2014	72	t324	4, 3, 1, 3, 4, 4	12-1-11-1-3-1-1-5	82	93	type 100	101
Sau_TMUS2134	Корея	2014	72	t324	4, 3, 1, 3, 4, 4	12-1-11-1-3-1-1-5	82	94	type 101	101

Случай 2. Клинические изоляты *S.aureus* были выделены от пациента с бактериемией [200, с.1688]. Бактериологическое исследование установило принадлежность данных штаммов к метициллин-резистентным стафилококкам. На фоне лечения имипенемом и ванкомицином у штамма *S.aureus* JH1 спустя 2,5 месяца развилась устойчивость к ванкомицину (*S.aureus* JH9) (рисунок 42).



IMP – имипенем, VAN – ванкомицин

Рисунок 42 – Курс антибактериальной терапии пациента, на фоне которой у *S.aureus* развилась резистентность

Полногеномное сравнение показало, что за 2,5 месячный период в геноме *S.aureus* JH9 произошли порядка 100 генетических событий (70 SNP, 6 делеций, 23 инсерций, 1 транслокация) в сравнении с геномом *S.aureus* JH1.

Биоинформатический анализ геномов *S.aureus*, на основе которого были получены внутривидовые типы *S.aureus* различными методами типирования для оценки достаточности разрешающей способности данных методов в случае с 2 клиническими штаммами *S.aureus* представлен в таблице 34.

Методы, основанные на таргетном секвенировании (MLST, MVLST, spa), а также на тандемных повторах (MLVA) и сайтах рестрикции (PFGE) не смогли установить накопленное различие в геноме порядка 0,001%. В данном случае,

методы, в основе которых используются полногеномные данные (cgMLST, wgMLST и panGenome) позволили провести различие между данными изолятами *S.aureus*.

Таблица 34 – Сравнение результатов типирования двух изолятов *S.aureus*

Штамм	Страна выделения	Год	MLST тип	Spa тип	MVLST профиль	MLVA профиль	PFGE тип	Cg MLST тип	Wg MLST тип	Pan Geno me тип
Sau_JH1	USA	2000	105	t002	7, 5, 5, 5, 5, 6	14-2-11-2-1-5-3-5	56	63	65	64
Sau_JH9	USA	2000	105	t002	7, 5, 5, 5, 5, 6	14-2-11-2-1-5-3-5	56	64	66	65

Случай 3. В своей работе французские исследователи изучали фенотипическую изменчивость клинических штаммов *S.aureus*, вызванную адаптационными механизмами с приобретением факторов вирулентности в период перехода от острого в хроническую фазу инфекционного процесса [228]. На данный момент, геномы клинических изолятов *S.aureus*, выделенные от пациентов в постоперационный период и спустя 3 месяца в 1 случае, 7 месяцев во втором случае и 5 месяцев в третьем случае, были секвенированы и размещены в NCIB GenBank (# CP012978, CP012979, CP012970, CP012972, CP012974, CP012976).

Анализ полногеномных последовательностей с целью определения количества однонуклеотидных замен, инсерций и делеций показал, что в геноме Sau_20130939 (в сравнении с Sau_20130938) за 2–3 -месячный период произошли порядка 114 генетических событий, включающих 34 однонуклеотидные замены, 36 делеций и 44 инсерций, что составило 0,05% различие в геномах. В случае с изолятами Sau_20130942 и Sau_20130943, в последнем за 6-7-месячный период произошли 102 генетических события (33 SNPs, 27 делеций и 42 инсерций), что составляет 0,007% различие в геномах. Для пары геномов Sau_20130940 и Sau_20130941 за 5 месячный период произошло накопление в 47 однонуклеотидные мутаций, 63 делеции и 27 инсерции, что в итоге, привело к 0,5% различию.

Возможности методов типирования провести различие данных штаммов *S.aureus*, в геномах которых произошли генетические изменения за период в несколько месяцев, представлены в таблице 35.

В данном случае методы полногеномного типирования (cgMLST, wgMLST и panGenome) смогли провести различие данных изолятов. Следует отметить, что в 1 из 3 случаев MLVA типирование позволило провести различие между штаммами *S.aureus* в случае с пациентом №1.

В то же время, если провести сравнение штаммов, выделенных от различных пациентов с различием во времени 7–8 месяцев, но в одном госпитале, например, штаммы *S.aureus* 20130938 и *S.aureus* 20130940, которые

с долей вероятности могут быть отнесены к представителям внутрибольничной инфекции (выделенные в одной больнице с общим MLST, MVLST профилем). Сравнение геномов данных штаммов определило 2,14% различие, включая 378 однонуклеотидных замен, 51 делецию и 378 инсерций. При таком различии в геномах *S.aureus* методы типирования, такие как Spa, MLVA, PFGE, а также методы, основанные на данных полногеномной последовательности (cgMLST, wgMLST и panGenome) смогли провести различие. Методы, использующие частичное секвенирование, такие как MLST и MVLST не показали достаточной разрешающей способности для проведения дифференцировки данных штаммов *S.aureus* (таблица 36).

Несмотря на то, что MVLST типирование показало большую разрешающую мощность в сравнении со Spa типированием на коллекции 4976 штаммов *S.aureus*, характер изменений в геноме и их количество не позволили определить накопленное различие, в то время как Spa исследование разделило их.

Таким образом, несмотря на такой важный параметр метода типирования как дискриминирующая мощность, в геномах микроорганизмов могут произойти изменения, которые могут быть детектированы методом типирования, обладающим меньшей разрешающей способностью.

Таблица 35 – Сравнение результатов типирования клинических изолятов *S.aureus*

Пациент №1 (мужчина 26 лет)										
Штамм	Страна выделения	Время выделения	ST	Spa тип	MVLST профиль	MLVA профиль	PFGE тип	cgMLST тип	wgMLST тип	PanGenome тип
Sau_20130938	Франция	2009-05	15	t385	25, 31, 23, 2, 1, 1	7-1-11-3-2-4-1-3	71	81	88	89
Sau_20130939	Франция	2009-08	15	t385	25, 31, 23, 2, 1, 1	7-1-11-1-2-5-1-3	71	82	89	90
Пациент №2 (мужчина 80 лет)										
Sau_20130942	Франция	2009-05	25	t078	11, 14, 24, 2, 20, 21	12-1-10-0-2-6-4-3	6	6	6	6
Sau_20130943	Франция	2009-12	25	t078	11, 14, 24, 2, 20, 21	12-1-10-0-2-6-4-3	6	7	7	7
Пациент №3 (женщина 82 лет)										
Sau_20130940	Франция	2010-01	15	t084	25, 31, 23, 2, 1, 1	7-1-12-3-2-4-1-3	70	79	86	87
Sau_20130941	Франция	2010-06	15	t084	25, 31, 23, 2, 1, 1	7-1-12-3-2-6-1-2	70	80	87	88

Таблица 36 – Сравнение результатов типирования клинических изолятов *S.aureus*

Штамм	Страна выделения	год	ST	Spa тип	MVLST профиль	MLVA профиль	PFGE тип	cgMLST тип	wgMLST тип	PanGenome тип
Sau_20130938	Франция	2009-05	15	t385	25, 31, 23, 2, 1, 1	7-1-11-3-2-4-1-3	71	81	88	89
Sau_20130940	Франция	2010-01	15	t084	25, 31, 23, 2, 1, 1	7-1-12-3-2-4-1-3	70	79	86	87

ЗАКЛЮЧЕНИЕ

Диссертационная работа выполнялась на базе Лаборатории Коллективного Пользования КГМУ (Караганда, Казахстан), большая часть работы по проведению внутривидового типирования *S.aureus* и их последующего анализа выполнялась самим автором.

Первичным этапом обозначенного направления исследования являлся сбор коллекции штаммов *S.aureus*, выделенных от пациентов различных стационаров городов: Астана, Караганда, Жезказган и Семей. Получение чистой культуры микроорганизмов и установление видовой принадлежности проводилось с использованием классических бактериологических методов лабораторного исследования. Подтверждение видовой идентификации проводилось на основе масс-спектрометрии MALDI-TOF технологии. Оценка чувствительности к антимикробным препаратам проводилась количественным методом (метод серийного разведения) с использованием роботизированной станции Тесап. Результаты чувствительности к антимикробным препаратам оценивались на основе руководства CLSI, которые были прописаны в программе WhoNet.

На следующем этапе коллекция клинических изолятов *S.aureus* подверглась внутривидовому типированию методами Spa (Shopsin B с соавт., 1999) и MLVA (Hardy, K. с соавт., 2004) с последующей кластеризацией для установления гетерогенности внутривидовых групп, а уровня клонального распространения *S.aureus* на территории Центрального Казахстана.

Оценка дискриминирующей способности методов Spa и MLVA типирования, а также уровня конкордантности результатов методов типирования на уровне внутривидовых типов и внутривидовых групп рассчитывалась на основе индекса Симпсона и скорректированного показателя Ранда. Сравнительная оценка современных молекулярно-генетических методов типирования, таких как MLST, Spa, MVLST, MLVA, cgMLST, wgMLST, panGenome, полученная на большой выборке глобальной коллекции изолятов *S.aureus*, была получена в *in silico* исследовании на основе полногеномных данных (n=6970) *S.aureus* с использованием различных биоинформатических подходов. На заключительном этапе исследования был проведен подробный анализ параметров, влияющих на разрешающую способность метода типирования, а также уровень согласованности результатов, полученных при использовании различных методов субвидового типирования микроорганизмов на примере *S.aureus*.

На основании результатов диссертационной работы можно сформулировать следующие выводы:

1. Уровень устойчивости клинических изолятов *S.aureus* (n=98) Центрально-Казахстанского региона к макролидам (азитромицин) составляет 25%, а к цефалоспорином (цефазолин, цефакситин) – порядка 18%. Резистентность к оксациллину наблюдается на уровне 18% штаммами MRSA, что подтверждается ПЦР исследованием на наличие *mecA* гена.

2. Кластеризация данных Spa типирования и MLVA типирования изолятов *S.aureus* позволило установить 4 внутривидовых комплекса/группы с доминированием двух клональных комплексов. Структура комплексов, включающая изоляты из различных городов Казахстана взаимную циркуляцию представителей определенных внутривидовых групп. В то же время, обнаружение Spa типов t002, t030, t008 позволяет говорить о вовлеченности в глобальную эпидемиологию изолятов *S.aureus*, выделенных в Республике Казахстан.

3. Дискриминирующая мощность MLVA и Spa методов, полученная при типировании коллекции клинических штаммов *S.aureus* (n = 98) была на уровне ~90-92%.

4. Конкордантность результатов двух методов типирования (Spa и MLVA) составила 20,7% на уровне внутривидовых типов и 86,2% на уровне близкородственных групп, указывая на хорошую согласованность для внутривидовых групп, определенных данными методами типирования.

5. *In silico* анализ 4976 полногеномных данных глобальной коллекции *S.aureus*, позволил ранжировать 8 современных молекулярно-генетических методов типирования *S.aureus* по их дискриминирующей способности. Методы, основанные на полногеномной последовательности микроорганизмов, такие как panGenome, wgMLST, cgMLST продемонстрировали наивысшую разрешающую способность (100%). Метод типирования на tandemных повторах показал 98% дискриминирующую мощность. Разрешающая сила MVLSST и Spa типирования была чуть выше 90% (92% и 90%, соответственно). Наименьшей дискриминирующей силой обладал MLST метод (85,5%).

6. Конкордантность данных методов типирования на уровне типов не превысила 50%, за исключение cgMLST и wgMLST (85%). Согласованность результатов, сгруппированных в близкородственные группы/комплексы/кластеры хорошую согласованность в диапазоне от 66% до 95%, за исключение согласованности между MLVA типированием и другими методами (39-61%).

7. Параметрами, влияющими на дискриминирующую способность схемы типирования, являются: количество локусов в схеме типирования, показатели межлокусной конкордантности и дискриминирующая мощность каждого локуса, входящие в схему типирования. Ведущая роль в определении разрешающей способности принадлежит показателю «количество маркеров» в схеме типирования. Кроме того, коллекция микроорганизмов, на которой выполнялось исследование, также влияет на показатель Симпсона, отражающий дискриминирующую мощность метода типирования.

8. Все этапы кластерного анализа данных, такие как выбор расчета матрицы генетических дистанций, алгоритмы кластерной сборки для построения дендрограммы, а также показатель cut-off или правила, на которых происходит формирование внутривидовых групп, оказывают влияние на показатель конкордантности между результатами внутривидового типирования микроорганизмов.

ПРАКТИЧЕСКИЕ РЕКОМЕНДАЦИИ

Большое количество разнообразных методов типирования микроорганизмов, а также отсутствие общепринятого референс метода, который может быть использован в качестве «золотого стандарта», оставляет за исследователем право выбора метода типирования. В то же время, выбор метода типирования должен исходить не только из характеристик метода, но также целей и задач исследовательской работы. Метод, основанный на секвенировании генов домашнего хозяйства является подходящим для изучения популяционной структуры микроорганизмов, в то время как эпидемиологическое расследование внутрибольничных вспышек потребует метод типирования с большей дискриминирующей способностью.

Методы субвидового типирования микроорганизмов, основанные на секвенировании, являются более предпочтительными в силу возможной стандартизации результатов. Создание международной базы данных результатов типирования позволит обеспечить международную стандартизованную номенклатуру внутривидовых типов микроорганизмов, в которой сопоставляя собственные результаты с данными других исследователей можно получить эпидемиологическую динамику тех или иных клонов в глобальном масштабе.

Однолокусное секвенирование типирование на основе маркера, совмещающее в себе вариации как в нуклеотидной последовательности, так в количестве tandemных повторов, как например Sra типирование, обеспечивает ряд преимуществ, таких как стандартизация полученных данных, портативность результатов, межлабораторная воспроизводимость, высокая дискриминирующая способность при уменьшенных затратах времени и стоимости на выполнение исследования в сравнении с мультилокусным секвенированием типированием.

Методы, основанные на анализе полногеномных данных, такие как cgMLST, wgMLST, panGenome типирование, помимо результатов генотипирования высокого разрешения, могут предоставить информацию о наличии генетических детерминант устойчивости, вирулентности и патогенности. В то же время, анализ полногеномных данных потребует от исследователя наличие биоинформатических навыков и инструментов для работы.

Активно развивающийся язык программирования для статистической обработки данных – R statistics, включающий пакеты по биоинформатическому анализу (R/Bioconductor packages), может являться хорошей альтернативой дорогим коммерческим программным продуктам, таких как CLC Genomic Workbench, Geneious, Lasergene DNASTar, Bionumerics и многим другим.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

- 1 Tenover F.C., Arbeit R.D., Goering R.V. How to select and interpret molecular strain typing methods for epidemiological studies of bacterial infections: a review for healthcare epidemiologists. Molecular Typing Working Group of the Society for Healthcare Epidemiology of America // Infect Control Hosp Epidemiol. - 1997. - Vol. 18, No 6. - P. 426-439.
- 2 MacCannell D. Bacterial strain typing // Clin Lab Med. - 2013. - Vol. 33, No 3. - P. 629-650.
- 3 Foxman B., Zhang L., Koopman J.S. et al. Choosing an appropriate bacterial typing technique for epidemiologic studies // Epidemiol Perspect Innov. - 2005. - Vol. 2. - P. 10-17.
- 4 Duarte A., Seliwiorstow T., Miller W.G. et al. Discriminative power of Campylobacter phenotypic and genotypic typing methods // J. Microbiol Methods. - 2016. - Vol. 125. - P. 33-39.
- 5 Keelara S., Scott H.M., Morrow W.M. et al. Comparative phenotypic and genotypic characterization of temporally related nontyphoidal Salmonella isolated from human clinical cases, pigs, and the environment in North Carolina // Foodborne Pathog Dis. - 2014. - Vol. 11, No 2. - P. 156-164.
- 6 Шагинян И.А. Роль и место молекулярно-генетических методов в эпидемиологическом анализе внутрибольничных инфекций // Клиническая микробиология и антимикробная химиотерапия. - 2000. - Т. 2, № 3. - С. 82-95.
- 7 Struelens M.J. Consensus guidelines for appropriate use and evaluation of microbial epidemiologic typing systems // Clin Microbiol Infect. - 1996. - Vol. 2, No 1. - P. 2-11.
- 8 Gordis L. Epidemiology. - 5th edn. – Philadelphia: PA, 2013. - 416 p.
- 9 Hennekens C.H., Buring J.E. Epidemiology in medicine. - 1st edn. – Philadelphia: PA, 1987. - 247 p.
- 10 Heath C.W., Alexander A.D. Galton M.M. Leptospirosis in the United States. Analysis of 483 cases in man, 1949, 1961 // N Engl J. Med. - 1965. - Vol. 273, No 17. - P. 915-922.
- 11 Truett J., Cornfield J., Kannel W. A multivariate analysis of the risk of coronary heart disease in Framingham // J. Chronic Dis. - 1967. - Vol. 20, No 7. - P. 511-524.
- 12 Higginson J. The role of the pathologist in environmental medicine and public health // Am J Pathol. - 1977. - Vol. 86, No 2. - P. 460-484.
- 13 Lower G.M., Nilsson T., Nelson C.E. et al. N-acetyltransferase phenotype and risk in urinary bladder cancer: approaches in molecular epidemiology. Preliminary results in Sweden and Denmark // Environ Health Perspect. - 1979. - Vol. 29, No - P. 71-79.

- 14 Perera F.P., Weinstein I.B. Molecular epidemiology and carcinogen-DNA adduct detection: new approaches to studies of human cancer causation // J. Chronic Dis. - 1982. - Vol. 35, No 7. - P. 581-600.
- 15 Summers W.C. Molecular epidemiology of DNA viruses: applications of restriction endonuclease cleavage site analysis // Yale J. Biol Med. - 1980. - Vol. 53, No 1. - P. 55-59.
- 16 Harris T.J., Underwood B.O., Knowles N.J. et al. Molecular approach to the epidemiology of swine vesicular disease: correlation of variation in the virus structural polypeptides with serological properties // Infect Immun. - 1979. - Vol. 24, No 3. - P. 593-599.
- 17 Foxman B., Riley L. Molecular epidemiology: focus on infection // Am J. Epidemiol. - 2001. - Vol. 153, No 12. - P. 1135-1141.
- 18 Schulte P.A., Perera F.P. Molecular epidemiology: principles and practices. - San Diego: Academic Press, 1993. – 588 p.
- 19 Tompkins L.S. Molecular epidemiology: development and application of molecular methods to solve infectious disease mysteries. - Washington: American Society for Microbiology, 1994. - P. 63–73.
- 20 McMichael A.J. Invited commentary-"molecular epidemiology": new pathway or new travelling companion? // Am J. Epidemiol. - 1994. - Vol. 140, No 1. - P. 1-11.
- 21 Groopman J.D., Kensler T.W., Links J.M. Molecular epidemiology and human risk monitoring // Toxicol Lett. - 1995. - Vol. 82-83. - P. 763-769.
- 22 Hall A. What is molecular epidemiology? // Trop Med Int Health. - 1996. - Vol. 1, No 4. - P. 407-408.
- 23 Shpilberg O., Dorman J.S., Ferrell R.E. et al. The next stage: molecular epidemiology // J. Clin. Epidemiol. - 1997. - Vol. 50, No 6. - P. 633-638.
- 24 Levin B.R., Lipsitch M., Bonhoeffer S. Population biology, evolution, and infectious disease: convergence and synthesis // Science. - 1999. - Vol. 283, No 5403. - P. 806-809.
- 25 Boffetta P. Molecular epidemiology // J. Intern. Med. - 2000. - Vol. 248, No 6. - P. 447-454.
- 26 Schulte P.A., Rothman N., Hainaut P. et al. Molecular epidemiology: linking molecular scale insights to population impacts // IARC Sci Publ. - 2011. - No 163. - P. 1-7.
- 27 Committee on Biological Markers of the National Research Council. Biologic markers in environmental health research // Environ Health Perspect. - 1987. - Vol. 74. - P. 3-9.
- 28 Schulte P.A. A conceptual and historical framework for molecular epidemiology. In: Schulte PA, Perrera FP Molecular epidemiology: principles and practices. - New York: Academic, 1993. – 44 p.

- 29 Lipsitch M. Evolution in health and disease // Trends Microbiol. - 1997. - Vol. 5, No 8. - P. 303-305.
- 30 Thompson R.C., Constantine C.C., Morgan U.M. Overview and significance of molecular methods: what role for molecular epidemiology? // Parasitology. - 1998. - Vol. 117. - P. S161-175.
- 31 Clark K., Karsch-Mizrachi I., Lipman D.J. et al. GenBank // Nucleic Acids Res. - 2016. - Vol. 44, No D1. - P. 67-72.
- 32 Silvester N., Alako B., Amid C. et al. Content discovery and retrieval services at the European Nucleotide Archive // Nucleic Acids Res. - 2015. - Vol. 43. - P. 23-29.
- 33 Mashima J., Kodama Y., Kosuge T. et al. DNA data bank of Japan (DDBJ) progress report // Nucleic Acids Res. - 2016. - Vol. 44, No D1. - P. 51-57.
- 34 Biswas S., McDonald M., Lundberg D.S. et al. Learning Microbial Interaction Networks from Metagenomic Count Data // J. Comput Biol. - 2016. - Vol. 23, No 6. - P. 526-535.
- 35 Levy M., Thaïs C.A., Elinav E. Metagenomic cross-talk: the regulatory interplay between immunogenomics and the microbiome // Genome Med. - 2015. - Vol. 7. - P. 120-129.
- 36 Hosseinkhani F., Jabalameli F., Nodeh Farahani N. et al. Variable number of tandem repeat profiles and antimicrobial resistance patterns of *Staphylococcus haemolyticus* strains isolated from blood cultures in children // Infect Genet Evol. - 2016. - Vol. 38. - P. 19-21.
- 37 Struelens M.J., Hawkey P.M., French G.L. et al. Laboratory tools and strategies for methicillin-resistant *Staphylococcus aureus* screening, surveillance and typing: state of the art and unmet needs // Clin Microbiol Infect. - 2009. - Vol. 15, No 2. - P. 112-119.
- 38 CDC. Centers for Disease Control and Prevention PulseNet. – 2012, August - 20. - 84 p. <http://www.cdc.gov/pulsenet>.
- 39 Xiao L., Kehoe P.G. Molecular epidemiology of infectious diseases - expanding horizons for IJMEG // Int. J. Mol Epidemiol Genet. - 2010. - Vol. 1, No 3. - P. 245-247.
- 40 Riley L.W. Molecular epidemiology of infectious diseases: principles and practices. - Washington: ASM, 2004. – 62 p.
- 41 Aikembayev A.M., Lukhnova L., Temiraliyeva G. et al. Historical distribution and molecular diversity of *Bacillus anthracis*, Kazakhstan // Emerg Infect Dis. - 2010. - Vol. 16, No 5. - P. 789-796.
- 42 Feikin D.R., Klugman K.P. Historical changes in pneumococcal serogroup distribution: implications for the era of pneumococcal conjugate vaccines // Clin. Infect Dis. - 2002. - Vol. 35, No 5. - P. 547-555.

- 43 Holmes A., Nolan R., Taylor R. et al. An epidemic of burkholderia cepacia transmitted between patients with and without cystic fibrosis // J. Infect Dis. - 1999. - Vol. 179, No 5. - P. 1197-1205.
- 44 Tresoldi A.T., Padoveze M.C., Trabasso P. et al. Enterobacter cloacae sepsis outbreak in a newborn unit caused by contaminated total parenteral nutrition solution // Am J. Infect Control. - 2000. - Vol. 28, No 3. - P. 258-261.
- 45 Said H.M., Krishnamani K., Omar S.V. et al. Evaluation of Semiautomated IS6110-Based Restriction Fragment Length Polymorphism Typing for Mycobacterium tuberculosis in a High-Burden Setting // J. Clin Microbiol. - 2016. - Vol. 54, No 10. - P. 2547-2552.
- 46 Otal I., Martin C., Vincent-Levy-Frebault V. et al. Restriction fragment length polymorphism analysis using IS6110 as an epidemiological marker in tuberculosis // J. Clin Microbiol. - 1991. - Vol. 29, No 6. - P. 1252-1254.
- 47 Shamputa I.C., Lee J., Allix-Beguec C. et al. Genetic diversity of Mycobacterium tuberculosis isolates from a tertiary care tuberculosis hospital in South Korea // J. Clin Microbiol. - 2010. - Vol. 48, No 2. - P. 387-394.
- 48 Deplano A., Struelens M. Marqueurs épidémiologiques. – Paris: ESKA, 2006. – 248 p.
- 49 Tibayrenc M., Ayala F.J. How clonal are Neisseria species? The epidemic clonality model revisited // Proc Natl Acad Sci U S A. - 2015. - Vol. 112, No 29. - P. 8909-8913.
- 50 Smith J.M., Smith N.H., O'Rourke M. et al. How clonal are bacteria? // Proc Natl Acad Sci U S A. - 1993. - Vol. 90, No 10. - P. 4384-4388.
- 51 Ilczyszyn W.M., Sabat A.J., Akkerboom V. et al. Clonal Structure and Characterization of Staphylococcus aureus Strains from Invasive Infections in Paediatric Patients from South Poland: Association between Age, spa Types, Clonal Complexes, and Genetic Markers // PLoS One. - 2016. - Vol. 11, No 3. - P. e0151937- e0151941.
- 52 Vaez H., Moghim S., Nasr Esfahani B. et al. Clonal Relatedness among Imipenem-Resistant Pseudomonas aeruginosa Isolated from ICU-Hospitalized Patients // Crit Care Res Pract. - 2015. - Vol. 2015. - P. 983207-983214.
- 53 Van Belkum A., Tassios P.T., Dijkshoorn L. et al. Guidelines for the validation and application of typing methods for use in bacterial epidemiology // Clin Microbiol Infect. - 2007. - Vol. 13. - P. 1-46.
- 54 Hunter P.R., Gaston M.A. Numerical index of the discriminatory ability of typing systems: an application of Simpson's index of diversity // J. Clin. Microbiol. - 1988. - Vol. 26, No 11. - P. 2465-2466.
- 55 Koonin E.V. Horizontal gene transfer: essentiality and evolvability in prokaryotes, and roles in evolutionary transitions // F1000Res. - 2016. - Vol. 5. – P. 114-115.

- 56 Sacha P., Wieczorek P., Ojdana D. et al. Expression of MexAB-OprM efflux pump system and susceptibility to antibiotics of different *Pseudomonas aeruginosa* clones isolated from patients hospitalized in two intensive care units at University Hospital in Bialystok (northeastern Poland) between January 2002 and December 2009 // *APMIS*. - 2014. - Vol. 122, No 10. - P. 931-940.
- 57 Dietrich L.E., Teal T.K., Price-Whelan A. et al. Redox-active antibiotics control gene expression and community behavior in divergent bacteria // *Science*. - 2008. - Vol. 321, No 5893. - P. 1203-1206.
- 58 Moore J.E., Goldsmith C.E. Phenotyping of *Campylobacter jejuni* and *Campylobacter coli* by a quantitative antibiogram [MIC] typing scheme using Euclidean distances [QATED] // *BMC Microbiol*. - 2001. - Vol. 1. - P. 13-18.
- 59 Rossney A.S., Coleman D.C., Keane C.T. Antibiogram-resistogram typing scheme for methicillin-resistant *Staphylococcus aureus* // *J. Med Microbiol*. - 1994. - Vol. 41, No 6. - P. 430-440.
- 60 Blanc D.S., Petignat C., Moreillon P. et al. Quantitative antibiogram as a typing method for the prospective epidemiological surveillance and control of MRSA: comparison with molecular typing // *Infect Control Hosp Epidemiol*. - 1996. - Vol. 17, No 10. - P. 654-659.
- 61 Fluit A.C., Bruggen J.T., Aarestrup F.M. et al. Priorities for antibiotic resistance surveillance in Europe // *Clin Microbiol Infect*. - 2006. - Vol. 12, No 5. - P. 410-417.
- 62 Herikstad H., Motarjemi Y., Tauxe R.V. Salmonella surveillance: a global survey of public health serotyping // *Epidemiol Infect*. - 2002. - Vol. 129, No 1. - P. 1-8.
- 63 Fratamico P.M., DebRoy C., Liu Y., et al. Advances in Molecular Serotyping and Subtyping of *Escherichia coli* // *Front Microbiol*. - 2016. - Vol. 7. - P. 644-651.
- 64 Camargo D.R., Pais F.S., Volpini A.C. et al. Revisiting molecular serotyping of *Streptococcus pneumoniae* // *BMC Genomics*. - 2015. - Vol. 16, Suppl 5. - P. 1-5.
- 65 Ochman H., Selander R.K. Evidence for clonal population structure in *Escherichia coli* // *Proc Natl Acad Sci U S A*. - 1984. - Vol. 81, No 1. - P. 198-201.
- 66 Cazalet C., Jarraud S., Ghavi-Helm Y. et al. Multigenome analysis identifies a worldwide distributed epidemic *Legionella pneumophila* clone that emerged within a highly diverse species // *Genome Res*. - 2008. - Vol. 18, No 3. - P. 431-441.
- 67 Coimbra R.S., Grimont F., Lenormand P. et al. Identification of *Escherichia coli* O-serogroups by restriction of the amplified O-antigen gene cluster (rfb-RFLP) // *Res Microbiol*. - 2000. - Vol. 151, No 8. - P. 639-654.
- 68 Stothard D.R., Jones R.B. Chlamydia trachomatis omp1 genotyping // *J. Infect. Dis*. - 2001. - Vol. 183, No 10. - P. 1542-1543.

- 69 Goerke C., Matias y Papenberg S., Dasbach S. et al. Increased frequency of genomic alterations in *Staphylococcus aureus* during chronic infection is in part due to phage mobilization // *J. Infect. Dis.* - 2004. - Vol. 189, No 4. - P. 724-734.
- 70 Stanley T.G. Wilson I. Multilocus enzyme electrophoresis: a practical guide // *Mol Biotechnol.* - 2003. - Vol. 24, No 2. - P. 203-220.
- 71 Musser J.M., Kroll J.S., Granoff D.M. et al. Global genetic structure and molecular epidemiology of encapsulated *Haemophilus influenzae* // *Rev Infect Dis.* - 1990. - Vol. 12, No 1. - P. 75-111.
- 72 Cherkaoui A., Hibbs J., Emonet S. et al. Comparison of two matrix-assisted laser desorption ionization-time of flight mass spectrometry methods with conventional phenotypic identification for routine identification of bacteria to the species level // *J. Clin. Microbiol.* - 2010. - Vol. 48, No 4. - P. 1169-1175.
- 73 Altun O., Botero-Kleiven S., Carlsson S. et al. Rapid identification of bacteria from positive blood culture bottles by MALDI-TOF MS following short-term incubation on solid media // *J. Med. Microbiol.* - 2015. - Vol. 64, No 11. - P. 1346-1352.
- 74 Karger A. Current developments to use linear MALDI-TOF spectra for the identification and typing of bacteria and the characterization of other cells/organisms related to infectious diseases // *Proteomics Clin Appl.* - 2016. - Vol. 10, No 9-10. - P. 982-993.
- 75 Balazova T., Makovcova J., Sedo O. et al. The influence of culture conditions on the identification of *Mycobacterium* species by MALDI-TOF MS profiling // *FEMS Microbiol Lett.* - 2014. - Vol. 353, No 1. - P. 77-84.
- 76 Goldstein J.E., Zhang L., Borrer C.M. et al. Culture conditions and sample preparation methods affect spectrum quality and reproducibility during profiling of *Staphylococcus aureus* with matrix-assisted laser desorption/ionization time-of-flight mass spectrometry // *Lett Appl Microbiol.* - 2013. - Vol. 57, No 2. - P. 144-150.
- 77 Christner M., Trusch M., Rohde H. et al. Rapid MALDI-TOF mass spectrometry strain typing during a large outbreak of Shiga-Toxigenic *Escherichia coli* // *PLoS One.* - 2014. - Vol. 9, No 7. - P. e101924- e101930.
- 78 Lasch P., Fleige C., Stammler M. et al. Insufficient discriminatory power of MALDI-TOF mass spectrometry for typing of *Enterococcus faecium* and *Staphylococcus aureus* isolates // *J. Microbiol Methods.* - 2014. - Vol. 100. - P. 58-69.
- 79 Jarcho J. Restriction fragment length polymorphism analysis // *Curr Protoc Hum Genet.* - 2001. – Vol. 2. - P. 27-34.
- 80 Enayat M.S. Restriction fragment length polymorphism // *Methods Mol Biol.* - 2002. - Vol. 187. - P. 29-35.
- 81 Sukkasem S., Yanai H., Mahasirimongkol S. et al. Drug resistance and IS6110-RFLP patterns of *Mycobacterium tuberculosis* in patients with recurrent

tuberculosis in northern Thailand // *Microbiol Immunol.* - 2013. - Vol. 57, No 1. - P. 21-29.

82 Chauhan D.S., Sharma V.D., Parashar D. et al. Molecular typing of *Mycobacterium tuberculosis* isolates from different parts of India based on IS6110 element polymorphism using RFLP analysis // *Indian J. Med Res.* - 2007. - Vol. 125, No 4. - P. 577-581.

83 Mendiola M.V., Martin C., Otal I. et al. Analysis of the regions responsible for IS6110 RFLP in a single *Mycobacterium tuberculosis* strain // *Res Microbiol.* - 1992. - Vol. 143, No 8. - P. 767-772.

84 Michelon D., Felix B., Vingadassalon N. et al. PFGE standard operating procedures for *Listeria monocytogenes*: harmonizing the typing of food and clinical strains in Europe // *Foodborne Pathog Dis.* - 2015. - Vol. 12, No 3. - P. 244-252.

85 Tenover F.C., Arbeit R.D., Goering R.V. et al. Interpreting chromosomal DNA restriction patterns produced by pulsed-field gel electrophoresis: criteria for bacterial strain typing // *J. Clin. Microbiol.* - 1995. - Vol. 33, No 9. - P. 2233-2239.

86 Murchan S., Kaufmann M.E., Deplano A. et al. Harmonization of pulsed-field gel electrophoresis protocols for epidemiological typing of strains of methicillin-resistant *Staphylococcus aureus*: a single approach developed by consensus in 10 European laboratories and its application for tracing the spread of related strains // *J. Clin. Microbiol.* - 2003. - Vol. 41, No 4. - P. 1574-1585.

87 Van Belkum A., Leeuwen W., Kaufmann M.E. et al. Assessment of resolution and intercenter reproducibility of results of genotyping *Staphylococcus aureus* by pulsed-field gel electrophoresis of *Sma*I macrorestriction fragments: a multicenter study // *J. Clin. Microbiol.* - 1998. - Vol. 36, No 6. - P. 1653-1659.

88 Ostojic M. Epidemiologic genotyping of methicillin-resistant *Staphylococcus aureus* (MRSA) by pulsed-field gel electrophoresis (PFGE) // *Bosn J. Basic Med Sci.* - 2008. - Vol. 8, No 3. - P. 259-265.

89 Hawash Y., Ghonaim M.M. Al-Shehri S.S. An Improved PCR-RFLP Assay for Detection and Genotyping of Asymptomatic *Giardia lamblia* Infection in a Resource-Poor Setting // *Korean J. Parasitol.* - 2016. - Vol. 54, No 1. - P. 1-8.

90 Fawley W.N., Knetsch C.W., MacCannell D.R. et al. Development and validation of an internationally-standardized, high-resolution capillary gel-based electrophoresis PCR-ribotyping protocol for *Clostridium difficile* // *PLoS One.* - 2015. - Vol. 10, No 2. - P. e0118150- e0118158.

91 Marialouis X.A., Santhanam A. Antibiotic Resistance, RAPD- PCR Typing of Multiple Drug Resistant Strains of *Escherichia Coli* From Urinary Tract Infection (UTI) // *J. Clin. Diagn Res.* - 2016. - Vol. 10, No 3. - P. 5-9.

92 Kim Y.J., Chon J.W., Lim J.S. et al. Traceback investigation for *Salmonella* contamination at egg processing plants in South Korea: prevalence, antibiotic resistance, and epidemiological tracing by rep-PCR fingerprinting // *J. Food Sci.* - 2015. - Vol. 80, No 4. - P. 759-764.

- 93 Ardakani M.A., Ranjbar R. Molecular typing of uropathogenic *E. coli* strains by the ERIC-PCR method // *Electron Physician*. - 2016. - Vol. 8, No 4. - P. 2291-2296.
- 94 Leung M.H., Bryson K., Freystatter K. et al. Sequotyping: serotyping *Streptococcus pneumoniae* by a single PCR sequencing strategy // *J. Clin. Microbiol.* - 2012. - Vol. 50, No 7. - P. 2419-2427.
- 95 Pereyre S., Touati A., Petitjean-Lecherbonnier J. et al. The increased incidence of *Mycoplasma pneumoniae* in France in 2011 was polyclonal, mainly involving *M. pneumoniae* type 1 strains // *Clin. Microbiol Infect.* - 2013. - Vol. 19, No 4. - P. 212217- 212222.
- 96 Roetzer A., Diel R., Kohl T.A. et al. Whole genome sequencing versus traditional genotyping for investigation of a *Mycobacterium tuberculosis* outbreak: a longitudinal molecular epidemiological study // *PLoS Med.* - 2013. - Vol. 10, No 2. - P. 1001387-1001393.
- 97 Diba K., Makhdoomi K., Mirhendi H. Molecular characterization of *Aspergillus* infections in an Iranian educational hospital using RAPD-PCR method // *Iran J. Basic Med Sci.* - 2014. - Vol. 17, No 9. - P. 646-650.
- 98 Lanzone V., Tofalo R., Fasoli G. et al. Food borne bacterial models for detection of benzo[a]pyrene-DNA adducts formation using RAPD-PCR // *Microb Biotechnol.* - 2016. - Vol. 9, No 3. - P. 400-407.
- 99 Madadgar O., Tadjbakhsh H., Salehi T.Z. et al. Evaluation of random amplified polymorphic DNA analysis and antibiotic susceptibility application in discrimination of *salmonella typhimurium* isolates in Iran // *New Microbiol.* - 2008. - Vol. 31, No 2. - P. 211-216.
- 100 Lin A.W., Usera M.A., Barrett T.J. et al. Application of random amplified polymorphic DNA analysis to differentiate strains of *Salmonella enteritidis* // *J. Clin. Microbiol.* - 1996. - Vol. 34, No 4. - P. 870-876.
- 101 Khosravi A.D., Hoveizavi H., Mohammadian A. et al. Genotyping of multidrug-resistant strains of *Pseudomonas aeruginosa* isolated from burn and wound infections by ERIC-PCR // *Acta Cir Bras.* - 2016. - Vol. 31, No 3. - P. 206-211.
- 102 Wei G., Pan L., Du H. et al. ERIC-PCR fingerprinting-based community DNA hybridization to pinpoint genome-specific fragments as molecular markers to identify and track populations common to healthy human guts // *J. Microbiol Methods.* - 2004. - Vol. 59, No 1. - P. 91-108.
- 103 Levinson G. Gutman G.A. Slipped-strand mispairing: a major mechanism for DNA sequence evolution // *Mol Biol Evol.* - 1987. - Vol. 4, No 3. - P. 203-221.
- 104 Okutani A., Tunglag H., Boldbaatar B. et al. Molecular epidemiological study of *Bacillus anthracis* isolated in Mongolia by multiple-locus variable-number tandem-repeat analysis for 8 loci (MLVA-8) // *Jpn. J. Infect Dis.* - 2011. - Vol. 64, No 4. - P. 345-348.

- 105 Nikolayevskyy V., Trovato A., Broda A. et al. MIRU-VNTR Genotyping of *Mycobacterium tuberculosis* Strains Using QIAxcel Technology: A Multicentre Evaluation Study // *PLoS One*. - 2016. - Vol. 11, No 3. - P. 149435- 149440.
- 106 Bolado-Martinez E., Candia-Plata Mdel C., Zenteno-Cuevas R. et al. Proposal of a five MIRU-VNTR panel to screen clinical isolates of *Mycobacterium tuberculosis* in Mexico // *Enferm Infecc Microbiol Clin*. - 2015. - Vol. 33, No 9. - P. 609-612.
- 107 Aires-de-Sousa M., Boye K., de Lencastre H. et al. High interlaboratory reproducibility of DNA sequence-based typing of bacteria in a multicenter study // *J. Clin. Microbiol*. - 2006. - Vol. 44, No 2. - P. 619-621.
- 108 Koreen L., Ramaswamy S.V., Graviss E.A. et al. Spa typing method for discriminating among *Staphylococcus aureus* isolates: implications for use of a single marker to detect genetic micro- and macrovariation // *J. Clin. Microbiol*. - 2004. - Vol. 42, No 2. - P. 792-799.
- 109 O'Hara F.P., Suaya J.A., Ray G.T. et al. Spa Typing and Multilocus Sequence Typing Show Comparable Performance in a Macroepidemiologic Study of *Staphylococcus aureus* in the United States // *Microb Drug Resist*. - 2016. - Vol. 22, No 1. - P. 88-96.
- 110 Hallin M., Deplano A., Denis O. et al. Validation of pulsed-field gel electrophoresis and spa typing for long-term, nationwide epidemiological surveillance studies of *Staphylococcus aureus* infections // *J. Clin Microbiol*. - 2007. - Vol. 45, No 1. - P. 127-133.
- 111 Maiden M.C., Bygraves J.A., Feil E. et al. Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms // *Proc Natl Acad Sci U S A*. - 1998. - Vol. 95, No 6. - P. 3140-3145.
- 112 Francisco A.P., Vaz C., Monteiro P.T. et al. PHYLOViZ: phylogenetic inference and data visualization for sequence based typing methods // *BMC Bioinformatics*. - 2012. - Vol. 13. - P. 87-92.
- 113 Larsen M.V., Cosentino S., Rasmussen S. et al. Multilocus sequence typing of total-genome-sequenced bacteria // *J. Clin. Microbiol*. - 2012. - Vol. 50, No 4. - P. 1355-1361.
- 114 Diosque P., Tomasini N., Lauthier J.J. et al. Optimized multilocus sequence typing (MLST) scheme for *Trypanosoma cruzi* // *PLoS Negl Trop Dis*. - 2014. - Vol. 8, No 8. - P. 3117- 3122.
- 115 Dan T., Liu W., Song Y. et al. The evolution and population structure of *Lactobacillus fermentum* from different naturally fermented products as determined by multilocus sequence typing (MLST) // *BMC Microbiol*. - 2015. - Vol. 15. - P. 107-112.
- 116 Chen Y., Zhang W., Knabel S.J. Multi-virulence-locus sequence typing identifies single nucleotide polymorphisms which differentiate epidemic clones and

outbreak strains of *Listeria monocytogenes* // J. Clin. Microbiol. - 2007. - Vol. 45, No 3. - P. 835-846.

117 Best E.L., Fox A.J., Frost J.A. et al. Identification of *Campylobacter jejuni* multilocus sequence type ST-21 clonal complex by single-nucleotide polymorphism analysis // J. Clin. Microbiol. - 2004. - Vol. 42, No 6. - P. 2836-2839.

118 Li P.C. Overview of Microarray Technology // Methods Mol Biol. - 2016. - Vol. 1368. - P. 3-4.

119 Li Y. Establishment and Application of a Visual DNA Microarray for the Detection of Food-borne Pathogens // Anal Sci. - 2016. - Vol. 32, No 2. - P. 215-218.

120 Garaizar J., Rementeria A., Porwollik S. DNA microarray technology: a new tool for the epidemiological typing of bacterial pathogens? // FEMS Immunol Med Microbiol. - 2006. - Vol. 47, No 2. - P. 178-189.

121 Zimenkov D.V., Kulagina E.V., Antonova O.V. et al. Simultaneous drug resistance detection and genotyping of *Mycobacterium tuberculosis* using a low-density hydrogel microarray // J. Antimicrob Chemother. - 2016. - Vol. 71, No 6. - P. 1520-1531.

122 Madaczak G., Szych J., Wasiak M. Results of *Salmonella enterica* subsp. *enterica* serotype identification by *Salmonella* Check&Trace microarray in international External Quality Assurance Systems // Med Dosw Mikrobiol. - 2015. - Vol. 67, No 1. - P. 15-22.

123 Rieg S., Jonas D., Kaasch A.J. et al. Microarray-based genotyping and clinical outcomes of *Staphylococcus aureus* bloodstream infection: an exploratory study // PLoS One. - 2013. - Vol. 8, No 8. - P. 71259- 71264.

124 Geue L., Monecke S., Engelmann I. et al. Rapid microarray-based DNA genoserotyping of *Escherichia coli* // Microbiol Immunol. - 2014. - Vol. 58, No 2. - P. 77-86.

125 El-Adawy H., Hotzel H., Tomaso H. et al. Detection of genetic diversity in *Campylobacter jejuni* isolated from a commercial turkey flock using *flaA* typing, MLST analysis and microarray assay // PLoS One. - 2013. - Vol. 8, No 2. - P. 51582-51592.

126 Patel A., Cheung S.W. Application of DNA Microarray to Clinical Diagnostics // Methods Mol Biol. - 2016. - Vol. 1368. - P. 111-132.

127 Schurch A.C., Soolingen D. DNA fingerprinting of *Mycobacterium tuberculosis*: from phage typing to whole-genome sequencing // Infect Genet Evol. - 2012. - Vol. 12, No 4. - P. 602-609.

128 Koser C.U., Ellington M.J., Cartwright E.J. et al. Routine use of microbial whole genome sequencing in diagnostic and public health microbiology // PLoS Pathog. - 2012. - Vol. 8, No 8. - P. 1002824- 1002827.

129 Aanensen D.M., Feil E.J., Holden M.T. et al. Whole-Genome Sequencing for Routine Pathogen Surveillance in Public Health: a Population Snapshot of Invasive *Staphylococcus aureus* in Europe // MBio. - 2016. - Vol. 7, No 3. – P. 54-58.

- 130 Williams O.M., Abeel T., Casali N. et al. Fatal nosocomial MDR TB identified through routine genetic analysis and whole-genome sequencing // *Emerg Infect Dis.* - 2015. - Vol. 21, No 6. - P. 1082-1084.
- 131 Скала Л.З. Практические аспекты современной клинической микробиологии. – М., 1997. – 312 с.
- 132 Reisner B.S. Specimen Processing // *Manual of Clinical Microbiology.* - 1999. - Vol. 7. - P. 64-104.
- 133 Клиническая лабораторная аналитика / под ред. В.В. Меньшикова. – М.: Агат-Мед, 2003. – 805 с.
- 134 William B.W. Bergey's Manual of Systematics of Archaea and Bacteria. Editor(s). - New Jersey, 2015. – 854 p.
- 135 Agarwal A., Kapila K., Kumar S. WHONET Software for the Surveillance of Antimicrobial Susceptibility // *Med J. Armed Forces India.* - 2009. - Vol. 65, No 3. - P. 264-266.
- 136 Perry S.F. Freeze-drying and cryopreservation of bacteria // *Mol Biotechnol.* - 1998. - Vol. 9, No 1. - P. 59-64.
- 137 Andrews J.M. Determination of minimum inhibitory concentrations // *J. Antimicrob Chemother.* - 2001. - Vol. 48, Suppl 1. - P. 5-16.
- 138 Shim S.M., Kim J.H., Jung S.E. et al. Multilaboratory assessment of variations in spectrophotometry-based DNA quantity and purity indexes // *Biopreserv Biobank.* - 2010. - Vol. 8, No 4. - P. 187-192.
- 139 Shopsin B., Gomez M., Montgomery S.O. et al. Evaluation of protein A gene polymorphic region DNA sequencing for typing of *Staphylococcus aureus* strains // *J. Clin. Microbiol.* - 1999. - Vol. 37, No 11. - P. 3556-3563.
- 140 Conceicao T., Aires de Sousa M., Lencastre H. Staphylococcal interspersed repeat unit typing of *Staphylococcus aureus*: evaluation of a new multilocus variable-number tandem-repeat analysis typing method // *J. Clin. Microbiol.* - 2009. - Vol. 47, No 5. - P. 1300-1308.
- 141 Frenay H.M., Bunschoten A.E., Schouls L.M., et al. Molecular typing of methicillin-resistant *Staphylococcus aureus* on the basis of protein A gene polymorphism // *Eur J. Clin. Microbiol Infect Dis.* - 1996. - Vol. 15, No 1. - P. 60-64.
- 142 Kearse M., Moir R., Wilson A. et al. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data // *Bioinformatics.* - 2012. - Vol. 28, No 12. - P. 1647-1649.
- 143 Ikawaty R., Willems R.J., Box A.T. et al. Novel multiple-locus variable-number tandem-repeat analysis method for rapid molecular typing of human *Staphylococcus aureus* // *J. Clin. Microbiol.* - 2008. - Vol. 46, No 9. - P. 3147-3151.
- 144 Altschul S.F., Gish W., Miller W. et al. Basic local alignment search tool // *J. Mol. Biol.* - 1990. - Vol. 215, No 3. - P. 403-410.

- 145 Enright M.C., Day N.P., Davies C.E. et al. Multilocus sequence typing for characterization of methicillin-resistant and methicillin-susceptible clones of *Staphylococcus aureus* // J. Clin. Microbiol. - 2000. - Vol. 38, No 3. - P. 1008-1015.
- 146 Kruczkiewicz P., Mutschall S., Barker D. et al. MIST: a tool for rapid in silico generation of molecular data from bacterial genome sequences. - 2013. – P. 1-8 <https://bitbucket.org/peterk87/microbialinsilicotyper/wiki/Home#markdown-header-paper>.
- 147 Verghese B., Schwalm N.D., Dudley E.G. et al. A combined multi-virulence-locus sequence typing and *Staphylococcal* Cassette Chromosome *mec* typing scheme possesses enhanced discriminatory power for genotyping MRSA // Infect Genet Evol. - 2012. - Vol. 12, No 8. - P. 1816-1821.
- 148 Schouls L.M., Spalburg E.C., Luit M. et al. Multiple-locus variable number tandem repeat analysis of *Staphylococcus aureus*: comparison with pulsed-field gel electrophoresis and spa typing // PLoS One. - 2009. - Vol. 4, No 4. - P. 5082- 5088.
- 149 Untergasser A., Cutcutache I., Koressaar T. et al. Primer3--new capabilities and interfaces // Nucleic Acids Res. - 2012. - Vol. 40, No 15. - P. 115-119.
- 150 McDougal L.K., Steward C.D., Killgore G.E. et al. Pulsed-field gel electrophoresis typing of oxacillin-resistant *Staphylococcus aureus* isolates from the United States: establishing a national database // J. Clin. Microbiol. - 2003. - Vol. 41, No 11. - P. 5113-5120.
- 151 Kamvar Z.N., Brooks J.C., Grunwald N.J. Novel R tools for analysis of genome-wide population genetic data with emphasis on clonality // Front Genet. - 2015. - Vol. 6. - P. 208-212.
- 152 Pedersen T.L. FindMyFriends: Microbial Comparative Genomics in R. // R package version 1.2.2. - 2016. - P.1-12. <https://github.com/thomasp85/FindMyFriends>.
- 153 Hyatt D., Chen G.L., Locascio P.F. et al. Prodigal: prokaryotic gene recognition and translation initiation site identification // BMC Bioinformatics. - 2010. - Vol. 11. - P. 119-123.
- 154 Fu L., Niu B., Zhu Z. et al. CD-HIT: accelerated for clustering the next-generation sequencing data // Bioinformatics. - 2012. - Vol. 28, No 23. - P. 3150-3152.
- 155 Li W., Godzik A. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences // Bioinformatics. - 2006. - Vol. 22, No 13. - P. 1658-1659.
- 156 Murata M. Three-way Needleman--Wunsch algorithm // Methods Enzymol. - 1990. - Vol. 183. - P. 365-375.
- 157 Nei M., Tajima F., Tateno Y. Accuracy of estimated phylogenetic trees from molecular data. II. Gene frequency data // J. Mol. Evol. - 1983. - Vol. 19, No 2. - P. 153-170.

- 158 Tateno Y., Nei M., Tajima F. Accuracy of estimated phylogenetic trees from molecular data. I. Distantly related species // *J. Mol. Evol.* - 1982. - Vol. 18, No 6. - P. 387-404.
- 159 Feil E.J., Li B.C., Aanensen D.M. et al. eBURST: inferring patterns of evolutionary descent among clusters of related bacterial genotypes from multilocus sequence typing data // *J. Bacteriol.* - 2004. - Vol. 186, No 5. - P. 1518-1530.
- 160 Carrico J.A., Pinto F.R., Simas C. et al. Assessment of band-based similarity coefficients for automatic type and subtype classification of microbial isolates analyzed by pulsed-field gel electrophoresis // *J. Clin. Microbiol.* - 2005. - Vol. 43, No 11. - P. 5483-5490.
- 161 Jaccard P. The distribution of the flora in the alpine zone // *New Phytologist.* - 1912. - Vol. 11, No 2. - P. 37-50.
- 162 Sokal R.R., Michener C.D. A statistical method for evaluating systematic relationships // *University of Kansas scientific bulletin.* - 1958. - Vol. 38, No 6. - P. 1409-1438.
- 163 Hennekinne J.A., Kerouanton A., Brisabois A. et al. Discrimination of *Staphylococcus aureus* biotypes by pulsed-field gel electrophoresis of DNA macro-restriction fragments // *J. Appl Microbiol.* - 2003. - Vol. 94, No 2. - P. 321-329.
- 164 Harastani H.H., Araj G.F., Tokajian S.T. Molecular characteristics of *Staphylococcus aureus* isolated from a major hospital in Lebanon // *Int J. Infect Dis.* - 2014. - Vol. 19. - P. 33-38.
- 165 Oksanen J. Community Ecology Package // R package version 2.3-5/r2997. - 2015. – P. 1-291. <https://R-Forge.R-project.org/projects/vegan>.
- 166 Hubert L., Arabie P. Comparing partitions // *Journal of Classification.* - 1985. - Vol. 2. - P. 193–218.
- 167 De Caceres M., Font X., Oliva F. The management of vegetation classifications with fuzzy clustering // *Journal of Vegetation Science.* - 2010. - Vol. 21, No 6. - P. 1138-1151.
- 168 Hesterberg T. Resample: Resampling Functions // R package version 0.4. - 2015. – P. 1-25. <https://CRAN.R-project.org/package=resample>.
- 169 Moreillon P. New and emerging treatment of *Staphylococcus aureus* infections in the hospital setting // *Clin. Microbiol Infect.* - 2008. - Vol. 14, Suppl 3. - P. 32-41.
- 170 Davis K.A., Stewart J.J., Crouch H.K. et al. Methicillin-resistant *Staphylococcus aureus* (MRSA) nares colonization at hospital admission and its effect on subsequent MRSA infection // *Clin. Infect Dis.* - 2004. - Vol. 39, No 6. - P. 776-782.
- 171 Lucet J.C., Paoletti X., Lolom I. et al. Successful long-term program for controlling methicillin-resistant *Staphylococcus aureus* in intensive care units // *Intensive Care Med.* - 2005. - Vol. 31, No 8. - P. 1051-1057.

172 SENTRY Antimicrobial Program Participants Group. Global aspects of antimicrobial resistance among key bacterial pathogens. Results from the 1997-1999 SENTRY antimicrobial program // Clin. Infect. Dis. - 2001. - Vol. 32, Suppl 2. - P. 81-155.

173 Garcia-Lozano T., Marti-Gonzalez R., Pascual-Pla F.J. et al. Clinical significance of molecular analysis of the chromosomal cassette SCCmecA in methicillin-resistant strains of Staphylococcus aureus // Med Clin (Barc). - 2015. - Vol. 145, No 2. - P. 91-92.

174 Laurent F., Chardon H., Haenni M. et al. MRSA harboring mecA variant gene mecC, France // Emerg Infect Dis. - 2012. - Vol. 18, No 9. - P. 1465-1467.

175 King J.M., Kulhankova K., Stach C.S. et al. Phenotypes and Virulence among Staphylococcus aureus USA100, USA200, USA300, USA400, and USA600 Clonal Lineages // mSphere. - 2016. - Vol. 1, No 3. - P. 124-129.

176 Rasmussen G., Monecke S., Ehricht R. et al. Prevalence of clonal complexes and virulence genes among commensal and invasive Staphylococcus aureus isolates in Sweden // PLoS One. - 2013. - Vol. 8, No 10. - P. 77477- 77482.

177 Ugolotti E., Bandettini R., Marchese A. et al. Molecular characterization of hospital-acquired methicillin-resistant Staphylococcus aureus strains in pediatric outbreaks using variable tandem repeat analysis with spa and ClfB typing // Diagn Microbiol Infect Dis. - 2011. - Vol. 69, No 2. - P. 213-217.

178 Blok H.E., Troelstra A., Kamp-Hopmans T.E. et al. Role of healthcare workers in outbreaks of methicillin-resistant Staphylococcus aureus: a 10-year evaluation from a Dutch university hospital // Infect Control Hosp Epidemiol. - 2003. - Vol. 24, No 9. - P. 679-685.

179 Sa-Leao R., Santos Sanches I., Dias D. et al. Detection of an archaic clone of Staphylococcus aureus with low-level resistance to methicillin in a pediatric hospital in Portugal and in international samples: relics of a formerly widely disseminated strain? // J. Clin. Microbiol. - 1999. - Vol. 37, No 6. - P. 1913-1920.

180 Clinical and Laboratory Standards Institute: Performance standards for antimicrobial susceptibility testing // Twentieth informational supplement document. - 2010. - Vol. 30, No 1. - P. 108-114.

181 Zahan N.A., Hossain M.A., Musa A.K. et al. PCR for mecA gene of methicillin resistant Staphylococcus aureus // Mymensingh Med J. - 2009. - Vol. 18, No 1. - P. 21-26.

182 Inouye Y., Morifuji K., Masumi N. et al. Susceptibility of methicillin-resistant Staphylococcus aureus clinical isolates to various antimicrobial agents. III. Novel, inducible resistance to macrolide-lincosamide-streptogramin B (MLS) antibiotics // Hiroshima J. Med Sci. - 1993. - Vol. 42, No 3. - P. 101-107.

183 Nakajima Y., Janosi L., Endou K. et al. Inducible resistance to a 16-membered macrolide, mycinamicin, in Staphylococcus aureus resistant to 14-

membered macrolides and streptogramin B antibiotics // J. Pharmacobiodyn. - 1992. - Vol. 15, No 6. - P. 319-324.

184 Hashimoto H., Oshima H., Mitsuhashi S. Drug resistance of staphylococci. IX. Inducible resistance to macrolide antibiotics in *Staphylococcus aureus* // Jpn. J. Microbiol. - 1968. - Vol. 12, No 3. - P. 321-327.

185 Saito T., Mitsuhashi S. Methylated adenine in the 23S ribosomal RNA of *Staphylococcus aureus* carrying temperature-inducible resistance to macrolide antibiotics // Microbiol Immunol. - 1979. - Vol. 23, No 11. - P. 1127-1129.

186 Roberts M.C. Resistance to macrolide, lincosamide, streptogramin, ketolide, and oxazolidinone antibiotics // Mol. Biotechnol. - 2004. - Vol. 28, No 1. - P. 47-62.

187 Schwendener S., Perreten V. New MLSB resistance gene *erm* (43) in *Staphylococcus lentus* // Antimicrob Agents Chemother. - 2012. - Vol. 56, No 9. - P. 4746-4752.

188 Roberts M.C. Update on macrolide-lincosamide-streptogramin, ketolide, and oxazolidinone resistance genes // FEMS Microbiol Lett. - 2008. - Vol. 282, No 2. - P. 147-159.

189 Farrell D.J., Flamm R.K., Sader H.S. et al. Ceftobiprole activity against over 60,000 clinical bacterial pathogens isolated in Europe, Turkey, and Israel from 2005 to 2010 // Antimicrob Agents Chemother. - 2014. - Vol. 58, No 7. - P. 3882-3888.

190 Khokhlova O.E., Hung W.C., Wan T.W. et al. Healthcare- and Community-Associated Methicillin-Resistant *Staphylococcus aureus* (MRSA) and Fatal Pneumonia with Pediatric Deaths in Krasnoyarsk, Siberian Russia: Unique MRSA's Multiple Virulence Factors, Genome, and Stepwise Evolution // PLoS One. - 2015. - Vol. 10, No 6. - P. 128017-128026.

191 Chen Y., Liu Z., Duo L. et al. Characterization of *Staphylococcus aureus* from distinct geographic locations in China: an increasing prevalence of *spa*-t030 and SCCmec type III // PLoS One. - 2014. - Vol. 9, No 4. - P. 96255- 96259.

192 Fenner L., Widmer A.F., Dangel M. et al. Distribution of *spa* types among methicillin-resistant *Staphylococcus aureus* isolates during a 6 year period at a low-prevalence University Hospital // J. Med. Microbiol. - 2008. - Vol. 57, No 5. - P. 612-616.

193 Soliman R.S., Phillips G., Whitty P. et al. Distribution of methicillin-resistant *Staphylococcus aureus spa* types isolated from health-care workers and patients in a Scottish university teaching hospital // J. Med. Microbiol. - 2009. - Vol. 58, No 9. - P. 1190-1195.

194 Stefani S., Chung D.R., Lindsay J.A. et al. Methicillin-resistant *Staphylococcus aureus* (MRSA): global epidemiology and harmonisation of typing methods // Int J. Antimicrob Agents. - 2012. - Vol. 39, No 4. - P. 273-282.

- 195 Adler A., Chmelnitsky I., Shitrit P. et al. Molecular epidemiology of methicillin-resistant *Staphylococcus aureus* in Israel: dissemination of global clones and unique features // *J. Clin. Microbiol.* - 2012. - Vol. 50, No 1. - P. 134-137.
- 196 Nulens E., Stobberingh E.E., Smeets E. et al. Genetic diversity of methicillin-resistant *Staphylococcus aureus* in a tertiary hospital in The Netherlands between 2002 and 2006 // *Eur J. Clin. Microbiol Infect Dis.* - 2009. - Vol. 28, No 6. - P. 631-639.
- 197 Miko B.A., Hafer C.A., Lee C.J. et al. Molecular characterization of methicillin-susceptible *Staphylococcus aureus* clinical isolates in the United States, 2004 to 2010 // *J. Clin. Microbiol.* - 2013. - Vol. 51, No 3. - P. 874-879.
- 198 Shittu A.O., Okon K., Adesida S. et al. Antibiotic resistance and molecular epidemiology of *Staphylococcus aureus* in Nigeria // *BMC Microbiol.* - 2011. - Vol. 11. - P. 92-96.
- 199 Hiramatsu K., Aritaka N., Hanaki H. et al. Dissemination in Japanese hospitals of strains of *Staphylococcus aureus* heterogeneously resistant to vancomycin // *Lancet.* - 1997. - Vol. 350, No 9092. - P. 1670-1673.
- 200 Sieradzki K., Leski T., Dick J. et al. Evolution of a vancomycin-intermediate *Staphylococcus aureus* strain in vivo: multiple changes in the antibiotic resistance phenotypes of a single lineage of methicillin-resistant *S. aureus* under the impact of antibiotics administered for chemotherapy // *J. Clin. Microbiol.* - 2003. - Vol. 41, No 4. - P. 1687-1693.
- 201 Tenover F.C., Vaughn R.R., McDougal L.K. et al. Multiple-locus variable-number tandem-repeat assay analysis of methicillin-resistant *Staphylococcus aureus* strains // *J. Clin. Microbiol.* - 2007. - Vol. 45, No 7. - P. 2215-2219.
- 202 Sobral D., Schwarz S., Bergonier D. et al. High throughput multiple locus variable number of tandem repeat analysis (MLVA) of *Staphylococcus aureus* from human, animal and food sources // *PLoS One.* - 2012. - Vol. 7, No 5. - P. 33967-33971.
- 203 Ikawaty R., Brouwer E.C., Jansen M.D. et al. Characterization of Dutch *Staphylococcus aureus* from bovine mastitis using a Multiple Locus Variable Number Tandem Repeat Analysis // *Vet Microbiol.* - 2009. - Vol. 136, No 3-4. - P. 277-284.
- 204 Feil E.J., Nickerson E.K., Chantratita N. et al. Rapid detection of the pandemic methicillin-resistant *Staphylococcus aureus* clone ST 239, a dominant strain in Asian hospitals // *J. Clin. Microbiol.* - 2008. - Vol. 46, No 4. - P. 1520-1522.
- 205 Enright M.C., Robinson D.A., Randle G. et al. The evolutionary history of methicillin-resistant *Staphylococcus aureus* (MRSA) // *Proc Natl Acad Sci U S A.* - 2002. - Vol. 99, No 11. - P. 7687-7692.
- 206 Leopold S.R., Goering R.V., Witten A. et al. Bacterial whole-genome sequencing revisited: portable, scalable, and standardized analysis for typing and detection of virulence and antibiotic resistance genes // *J. Clin. Microbiol.* - 2014. - Vol. 52, No 7. - P. 2365-2370.

- 207 Tettelin H., Masignani V., Cieslewicz M.J. et al. Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: implications for the microbial "pan-genome" // *Proc Natl Acad Sci U S A.* - 2005. - Vol. 102, No 39. - P. 13950-13955.
- 208 Chen F., Mackey A.J., Stoeckert C.J. et al. OrthoMCL-DB: querying a comprehensive multi-species collection of ortholog groups // *Nucleic Acids Res.* - 2006. - Vol. 34. - P. 363-368.
- 209 Fouts D.E., Brinkac L., Beck E. et al. PanOCT: automated clustering of orthologs using conserved gene neighborhood for pan-genomic analysis of bacterial strains and closely related species // *Nucleic Acids Res.* - 2012. - Vol. 40, No 22. - P. 172- 177.
- 210 Chaudhari N.M., Gupta V.K., Dutta C. BPGA- an ultra-fast pan-genome analysis pipeline // *Sci Rep.* - 2016. - Vol. 6. - P. 24373-24380.
- 211 Page A.J., Cummins C.A., Hunt M. et al. Roary: rapid large-scale prokaryote pan genome analysis // *Bioinformatics.* - 2015. - Vol. 31, No 22. - P. 3691-3693.
- 212 Snipen L., Liland K.H. Micropan: an R-package for microbial pan-genomics // *BMC Bioinformatics.* - 2015. - Vol. 16 - P. 79-82.
- 213 Rodriguez M., Hogan P.G., Satola S.W. et al. Discriminatory Indices of Typing Methods for Epidemiologic Analysis of Contemporary *Staphylococcus aureus* Strains // *Medicine (Baltimore).* - 2015. - Vol. 94, No 37. - P. 1534- 1541.
- 214 Faria N.A., Carrico J.A., Oliveira D.C. et al. Analysis of typing methods for epidemiological surveillance of both methicillin-resistant and methicillin-susceptible *Staphylococcus aureus* strains // *J. Clin. Microbiol.* - 2008. - Vol. 46, No 1. - P. 136-144.
- 215 Brandt K.M., Mellmann A., Ballhausen B. et al. Evaluation of multiple-locus variable number of tandem repeats analysis for typing livestock-associated methicillin-resistant *Staphylococcus aureus* // *PLoS One.* - 2013. - Vol. 8, No 1. - P. 54425- 54430.
- 216 Goudarzi M., Fazeli M., Goudarzi H. et al. Spa Typing of *Staphylococcus aureus* Strains Isolated from Clinical Specimens of Patients with Nosocomial Infections in Tehran, Iran // *Jundishapur J. Microbiol.* - 2016. - Vol. 9, No 7. - P. 35685- 35691.
- 217 Saunders N.A., Holmes A. Multilocus sequence typing (MLST) of *Staphylococcus aureus* // *Methods Mol. Biol.* - 2014. - Vol. 1085. - P. 113-130.
- 218 Didi J., Lemee L., Gibert L. et al. Multi-virulence-locus sequence typing of *Staphylococcus lugdunensis* generates results consistent with a clonal population structure and is reliable for epidemiological typing // *J. Clin. Microbiol.* - 2014. - Vol. 52, No 10. - P. 3624-3632.
- 219 Mellmann A., Weniger T., Berssenbrugge C. et al. Based Upon Repeat Pattern (BURP): an algorithm to characterize the long-term evolution of

Staphylococcus aureus populations based on spa polymorphisms // BMC Microbiol. - 2007. - Vol. 7. - P. 98-105.

220 Strommenger B., Kettlitz C., Weniger T. et al. Assignment of *Staphylococcus* isolates to groups by spa typing, SmaI macrorestriction analysis, and multilocus sequence typing // J. Clin. Microbiol. - 2006. - Vol. 44, No 7. - P. 2533-2540.

221 Meyer D., Buchta C. Distance and Similarity Measures // R package version 0.4-16. - 2016. – P.1-10. <https://CRAN.R-project.org/package=proxy>.

222 Gilpin B., Robson B., Lin S. et al. Pulsed-field gel electrophoresis analysis of more than one clinical isolate of *Campylobacter* spp. from each of 49 patients in New Zealand // J. Clin. Microbiol. - 2012. - Vol. 50, No 2. - P. 457-459.

223 Anderson K.M., Abbott J., Zhao S. et al. Molecular Subtyping of Shiga Toxin-Producing *Escherichia coli* Using a Commercial Repetitive Sequence-Based PCR Assay // J. Food Prot. - 2015. - Vol. 78, No 5. - P. 902-911.

224 Ward J.H. Hierarchical Grouping to Optimize an Objective Function // Journal of the American Statistical Association. - 1963. - Vol. 58. - P. 236–244.

225 Sokal R.R., Michener C.D. A statistical method for evaluating systematic relationships // University of Kansas Science Bulletin. - 1958. - Vol. 38. - P. 1409–1438.

226 Malachowa N., Sabat A., Gniadkowski M. et al. Comparison of multiple-locus variable-number tandem-repeat analysis with pulsed-field gel electrophoresis, spa typing, and multilocus sequence typing for clonal characterization of *Staphylococcus aureus* isolates // J. Clin. Microbiol. - 2005. - Vol. 43, No 7. - P. 3095-3100.

227 Yamaguchi T., Suzuki S., Okamura S. et al. Evolution and single-nucleotide polymorphisms in methicillin-resistant *Staphylococcus aureus* strains with reduced susceptibility to vancomycin and daptomycin, based on determination of the complete genome // Antimicrob Agents Chemother. - 2015. - Vol. 59, No 6. - P. 3585-3587.

228 Vogel V., Falquet L., Calderon-Copete S.P. et al. Short term evolution of a highly transmissible methicillin-resistant *Staphylococcus aureus* clone (ST228) in tertiary care hospital // PLoS One. - 2012. - Vol. 7, No 6. - P. 38969-38975.

ПРИЛОЖЕНИЕ А

Свидетельство о государственной регистрации прав на объект авторского права

Авторлық құқық объектісіне құқықтарды мемлекеттік тіркеу туралы	
ҚҰДЛҚ	
№ <u>156</u>	7 ақпан 2014 ж.
<p>Қазақстан Республикасы Әділет министрінің Зияткерлік меншік құқығы комитетінде авторлардың өтініші бойынша авторлары Илья Сулейманович Азизов, Елена Алексеевна Захарова, Аліёна Владимировна Лавриненко, Дмитрий Борисович Бабенко болып табылатын авторлық құқықпен қорғалатын объектіге айрықша мүлдіктік құқықтар «Программа для управления роботом Tesla Freedom EVO 150 «EVO-MIC» (ЭЕМ- та арналған бағдарлама) атауымен тіркелгені құпандырылды.</p> <p>Авторлардың өтініші бойынша авторлық құқықпен қорғалатын объектіге айрықша мүлдіктік құқықтар және 2014 жылғы 13 қаңтарда жасалған объекті И.С. Азизова, Е.А. Захарова, А.В. Лавриненко, Д.Б. Бабенко тиесілі және авторлар жоғарыда көрсетілген объектіні жасаған кезде басқа адамдардың зияткерлік меншік құқығы бұзылмағандығына кепілдік береді.</p> <p>Тізімде 2014 жылғы 7 ақпан жасалған № <u>156</u> жазба бар.</p> <p>Тораға А. Естаев</p>	
<p>СВИДЕТЕЛЬСТВО о государственной регистрации прав на объект авторского права</p> <p>№ <u>156</u> 7 ақпан 2014 г.</p> <p>Настоящим удостоверяется, что в Комитете по правам интеллектуальной собственности Министерства Юстиции Республики Казахстан зарегистрированы исключительные имущественные права на объект авторского права под названием «Программа для управления роботом Tesla Freedom EVO 150 «EVO-MIC» (программа для ЭВМ), авторами которого по заявлению авторов являются Азизов Илья Сулейманович, Захарова Елена Алексеевна, Лавриненко Аліёна Владимировна, Бабенко Дмитрий Борисович.</p> <p>По заявлению авторов исключительные имущественные права на объект авторского права, созданный 13 января 2014 года, принадлежат Азизову И.С., Захаровой Е.А., Лавриненко А.В., Бабенко Д.Б. и авторы гарантируют, что при создании вышеуказанного объекта не были нарушены права интеллектуальной собственности других лиц.</p> <p>Запись в реестре за № <u>156</u> от 7 ақпан 2014 года имеется.</p> <p>Председатель А. Естаев</p> <p>ИС 0011349</p>	

ПРИЛОЖЕНИЕ Б

Акты внедрения результатов научно-исследовательской работы

Ф КГМУ 7-06/02

ПП КГМУ 7/02

Утверждаю
Проректор по стратегическому
развитию, науке и
международному
сотрудничеству КГМУ

А.А. Турмухамбетова Турмухамбетова А.А.
«7» «11» 2016

Акт
внедрения результатов
научно-исследовательской работы № _____

Карагандинский государственный медицинский университет, кафедра молекулярной биологии и медицинской генетики

(наименование учреждения, где внедряется)

Наименование предложения: Оценка конкордантности результатов, полученные различными методами типирования микроорганизмов, рассчитанная на основе скорректированного индекса Ранда (1971г.)

Работа включена: расчет проводится на основе функции «concordance» программного пакета Miquel De Caceres «vegclust».

Внедрения в инициативном порядке, заимствование из методических рекомендаций

(Журнальных статей, статей, диссертаций.)

Форма внедрения: Внедрение метода

(Внедрение метода, аппарата, в лечебно-профилактическое учреждение лекции,

Семинары, подготовка на рабочем месте и прочее – указать),

Ответственный за внедрение и исполнитель: Бабенко Дмитрий Борисович

Эффективность внедрения: научно-исследовательская

(Лечебно-диагностическая, экономическая,

Семинары социальная - указать)

Предложения и замечания учреждения, осуществляющего внедрения

Сроки внедрения 2016

Председатель комиссии
Члены ответственные за внедрение

Б.Ж.Култанов
Д.Б.Бабенко



Ф КГМУ 7-06/02
ПП КГМУ 7/02

Утверждаю
Проректор по стратегическому
развитию, науке и
международному
сотрудничеству КГМУ
А.А. Турмухамбетов Турмухамбетов А.А.
« 8 » 11 2016

Акт
внедрения результатов
научно-исследовательской работы № _____

Карагандинский государственный медицинский университет, кафедра микробиологии
(наименование учреждения, где внедряется)

Наименование предложения: Оценка дискриминирующей способности метода
внутривидового типирования микроорганизмов на основе коэффициента разнообразия
Симпсона (1949 г.).

Работа включена: заимствована из функции «Simpson» программного пакета Robin K. S.
Hankin «unbt».

Внедрения в инициативном порядке, заимствование из методических рекомендаций
(Журнальных статей, статей, диссертаций.)

Форма внедрения: Внедрение метода
(Внедрение метода, аппарата, в лечебно-профилактическое учреждение, лекции,
Семинары, подготовка на рабочем месте и прочее – указать).

Ответственный за внедрение и исполнитель: Бабенко Дмитрий Борисович

Эффективность внедрения: научно-исследовательская
(Лечебно-диагностическая, экономическая,
Семинары социальная- указать)

Предложения и замечания учреждения, осуществляющего внедрения

Сроки внедрения 2016

Председатель комиссии
Члены ответственные за внедрение

Г.А.Абдуллина
Д.Б.Бабенко



Ф КГМУ 7-06/02
ПП КГМУ 7/02

Утверждаю
Проректор по стратегическому
развитию, науке и
международному
сотрудничеству КГМУ
Душесов Турмухамбетова А.А.
«5» 11 2016

Акт
внедрения результатов
научно-исследовательской работы № _____

Карагандинский государственный медицинский университет, Лаборатория
коллективного пользования

(наименование учреждения, где внедряется)

Наименование предложения: Анализ Spa типов на основе алгоритма Needleman-Wunsch,
позволяющего проводить глобальное выравнивание (global alignment) повторов Spa типов.
Работа включена: заимствована из программного кода для R statistics Alanocallaghan
пользователя www.github.com ресурса с модификацией кода для анализа Spa повторов в
Spa time.

Внедрения в инициативном порядке, заимствование из методических рекомендаций

Журнальных статей, статей, диссертаций.)

Форма внедрения: Внедрение метода

(Внедрение метода, аппарата, в лечебно-профилактическое учреждение лекции,

Семинары, подготовка на рабочем месте и прочее – указать).

Ответственный за внедрение и исполнитель: Бабенко Дмитрий Борисович

Эффективность внедрения: научно-исследовательская

Лечебно-диагностическая, экономическая,

Семинары социальная- указать)

Предложения и замечания учреждения, осуществляющего внедрения

Сроки внедрения 2016

Председатель комиссии
Члены ответственные за внедрение

Л.Л. Ахмалтдинова
Д.Б.Бабенко

