

ИДЕНТИФИКАЦИЯ МИКРООРГАНИЗМОВ С ПОМОЩЬЮ ИНФРАКРАСНЫХ ФУРЬЕ-СПЕКТРОВ

А. Ю. Сунцова¹, Р. Р. Гулиев¹, Д. А. Попов², Т. Ю. Вострикова², Д. В. Дубоделов³, А. Н. Щеголихин¹, Б. К. Лайпанов⁵, Т. В. Припутневич³, А. Б. Шевелев^{1,4}✉, И. Н. Курочкин¹

¹ Институт биохимической физики имени Н. М. Эмануэля, Москва

² Национальный медицинский исследовательский центр сердечно-сосудистой хирургии имени А. Н. Бакулева, Москва

³ Национальный медицинский исследовательский центр акушерства, гинекологии и перинатологии имени В. И. Кулакова, Москва

⁴ Институт общей генетики имени Н. И. Вавилова, Москва

⁵ Московская государственная академия ветеринарной медицины и биотехнологии имени К. И. Скрябина, Москва

Актуальность развития быстрых методов идентификации патогенных биологических объектов растет в связи с массовым распространением в лечебных учреждениях микроорганизмов с лекарственной устойчивостью и в связи с разворачиванием центров мониторинга биологических угроз. Инфракрасная Фурье-спектроскопия является эффективной альтернативой масс-спектрометрии с точки зрения стоимости и портативности оборудования, экспрессности анализа. Целью работы было описать алгоритм идентификации микроорганизмов в чистых культурах, основанный на анализе колебательных инфракрасных Фурье-спектров культур (Fourier transform infrared, FTIR). Алгоритм основан на автоматизированном анализе спектров методом главных компонент. В отличие от известных в литературе, он позволяет идентифицировать бактерии вне зависимости от стадии роста культуры и состава среды. Обучающая база данных включала наиболее распространенные возбудители гнойно-септических инфекций человека: *Staphylococcus aureus* ($n = 67$), *Enterococcus faecalis* ($n = 10$), *Enterococcus faecium* ($n = 10$), *Klebsiella pneumoniae* ($n = 10$), *Escherichia coli* ($n = 10$), *Serratia marcescens* ($n = 10$), *Enterobacter cloacae* ($n = 10$), *Acinetobacter baumannii* ($n = 10$), *Pseudomonas aeruginosa* ($n = 10$) и *Candida albicans* ($n = 10$). Построенная модель успешно апробирована на серии клинических изолятов *Staphylococcus aureus*: в слепых испытаниях участвовало 10 штаммов с фенотипом лекарственной устойчивости MRSA (methicillin-resistant *Staphylococcus aureus*) и 10 чувствительных штаммов, а также смесь культур этих штаммов с клиническими изолятами *Pseudomonas aeruginosa*, *Escherichia coli* и *Klebsiella pneumoniae*.

Ключевые слова: спектроскопия, Фурье, ИК-спектр, метод главных компонент, идентификация, бактерии, дрожжи

✉ **Для корреспонденции:** Шевелев Алексей Борисович
ул. Косыгина, д. 4, г. Москва, 119334; shevel_a@hotmail.com

Статья получена: 01.08.2018 **Статья принята к печати:** 25.08.2018

DOI: 10.24075/vrgmu.2018.046

IDENTIFICATION OF MICROORGANISMS BY FOURIER-TRANSFORM INFRARED SPECTROSCOPY

Suntsova AY¹, Guliev RR¹, Popov DA², Vostrikova TY², Dubodelov DV³, Shchegolikhin AN¹, Laypanov BK⁵, Priputnevich TV³, Shevelev AB^{1,4}✉, Kurochkin IN¹

¹ Emanuel Institute of Biochemical Physics, Moscow

² Bakulev National Medical Research Center of Cardiovascular Surgery, Moscow

³ Kulakov National Medical Research Center for Obstetrics, Gynecology and Perinatology, Moscow

⁴ Vavilov Institute of General Genetics, Moscow

⁵ Skryabin Moscow State Academy of Veterinary Medicine and Biotechnology, Moscow

The need for novel techniques of rapid identification of pathogenic microorganisms arises from the massive spread of drug-resistant nosocomial strains and the emergence of centers for biohazard control. Fourier-transform infrared spectroscopy is a promising alternative to mass spectrometry as it is cost-effective, fast and suitable for field use. The aim of this work was to propose an algorithm for the identification of microorganisms in pure cultures based on the analysis of their Fourier transform infrared spectra. The algorithm is based on the automated principal component analysis of infrared spectra. Unlike its analogues described in the literature, the algorithm is capable of identifying bacteria regardless of the culture medium or growth phase. The training sample included the most prevalent causative agents of infections and sepsis in humans: *Staphylococcus aureus* ($n = 67$), *Enterococcus faecalis* ($n = 10$), *Enterococcus faecium* ($n = 10$), *Klebsiella pneumoniae* ($n = 10$), *Escherichia coli* ($n = 10$), *Serratia marcescens* ($n = 10$), *Enterobacter cloacae* ($n = 10$), *Acinetobacter baumannii* ($n = 10$), *Pseudomonas aeruginosa* ($n = 10$), and *Candida albicans* ($n = 10$). The model we built successfully passed a series of blind tests involving clinical isolates of 10 methicillin-resistant (MRSA) and 10 methicillin-sensitive (MSSA) *Staphylococcus aureus* strains as well as pair mixes of these cultures with clinical isolates of *Pseudomonas aeruginosa*, *Escherichia coli*, and *Klebsiella pneumoniae*.

Keywords: spectroscopy, Fourier, IR, principal component analysis, identification, bacteria, yeasts

✉ **Correspondence should be addressed:** Alexei B. Shevelev
Kosygina 4, Moscow, 119334; shevel_a@hotmail.com

Received: 01.08.2018 **Accepted:** 25.08.2018

DOI: 10.24075/brsmu.2018.046

Идентификация видовой принадлежности микроорганизмов важна в медицинской микробиологии. В частности, экспресс-идентификация патогена необходима для выбора схемы антибиотикотерапии при лечении гнойных инфекций и сепсиса. Эти заболевания характеризуются стремительным течением и служат одной из частых причин летальных исходов при проведении операций на сердечно-сосудистой системе [1], а также у родильниц и новорожденных детей [2].

В настоящее время во всем мире для окончательной идентификации возбудителей гнойно-воспалительных заболеваний, помимо традиционных микробиологических, широко используются новые методы, в частности, метод матричной лазерной десорбционной ионизационной времяпролетной масс-спектрометрии (MALDI-ToF MS). Из наиболее распространенных приборов можно назвать MALDI-ToF масс-спектрометры марок MALDI BioType (Bruker; Германия) и Vitek MS (Biomerieux; Франция). Они дают быстрые и достоверные результаты, однако имеют высокую стоимость и пока недоступны для повсеместного использования в лечебных учреждениях. Кроме того, приборы достаточно громоздки и неудобны для работы в полевых условиях, что препятствует их использованию структурами, отвечающими за биологическую безопасность населения.

В качестве возможной альтернативы MALDI ToF масс-спектрометрии для идентификации микроорганизмов может быть рассмотрен метод инфракрасной колебательной Фурье-спектроскопии (Fourier transformation Infrared, FTIR). Он позволяет анализировать химический состав культуры по содержанию любых типов макромолекул и низкомолекулярных соединений. Подготовка образцов культур к анализу тоже проста, как и в случае MALDI ToF масс-спектрометрии: достаточно поместить культуру на подложку из материала, прозрачного в ИК-области спектра и высушить. Спектроскопия FTIR может рассматриваться в качестве метода экспресс-диагностики. Однако ее использование в практической микробиологии ограничено изменчивостью FTIR-спектров культуры в зависимости от состава среды культивирования и фазы ее роста. Целью работы была разработка алгоритма идентификации микроорганизмов в чистых культурах, основанного на анализе FTIR-спектров культур, позволяющем надежно идентифицировать культуру в любой фазе роста на среде любого состава.

МАТЕРИАЛЫ И МЕТОДЫ

Штаммы патогенов человека

В работе использовали штаммы наиболее распространенных возбудителей гнойно-септических инфекций человека: *S. aureus* (20 изолятов с фенотипом MRSA и 47 изолятов без лекарственной устойчивости), *E. faecalis* ($n = 10$), *E. faecium* ($n = 10$), *K. pneumoniae* ($n = 10$), *E. coli* ($n = 10$), *S. marcescens* ($n = 10$), *E. cloacae* ($n = 10$), *A. baumannii* ($n = 10$), *P. aeruginosa* ($n = 10$), *S. epidermidis* ($n = 10$) и *C. albicans* ($n = 10$).

Возбудители гнойно-септических инфекций человека были выделены от больных, проходивших лечение в стационарах ННПЦССХ им. А. Н. Бакулева и НМИЦ АГП им. В. И. Кулакова. Выделение и типирование штаммов осуществляли по стандартной методике [3] с использованием дифференциально-диагностических сред: элективного солевого агара для стафилококков, агара для выделения энтерококков, агара Эндо-ГПМ для выделения

грамотрицательных палочек, включая энтеробактерии и *P. aeruginosa*, среды Сабуро и мясо-пептонной среды с добавлением 1% глюкозы для *C. albicans*. Подтверждение принадлежности изолятов к указанным видам патогенов проводили с помощью масс-спектрометра MALDI BioType (Bruker; Германия).

Изоляты, заранее депонированные при $-70\text{ }^{\circ}\text{C}$ с использованием системы длительной криоконсервации «Криобанк», первоначально высевали на кровяной агар в аэробных условиях при $37\text{ }^{\circ}\text{C}$ в течение суток.

Каждый изолят высевали на агаризованную и жидкую среду двух типов: с содержанием крови и без нее — всего четыре типа сред для каждого изолята. В качестве сред без содержания крови использовали желточно-солевую и мясо-пептонный бульон, бульон Эндо и среду Сабуро. Материал для анализа методом спектроскопии FTIR собирали через 12, 24 и 48 ч.

С целью исключения опасности инфицирования оператора FTIR-спектрометра, а также с целью консервации культур патогенов при закладке на хранение, до снятия спектров их инактивировали добавлением 70% спирта. Для этого из свежих жидких культур отбирали по 30 мкл, помещали в полипропиленовую микропробирку и добавляли 70 мкл 96% этанола, перемешивая пипетированием. Материал с поверхности агаризованных сред отбирали микробиологической петлей и суспендировали в 70% водном этаноле из расчета 30 мг биомассы на 100 мкл спиртового раствора.

Регистрация ИК-спектров культур изолятов

Для регистрации ИК-спектров изолятов использовали фиксированные в 70% водном этаноле суспензии культур патогенов. Процедура подготовки проб индивидуальных патогенов к измерению методом ИК-спектроскопии «на пропускание» состояла в том, что на поверхность стандартных (диаметр 12,5 мм, толщина 2 мм) ZnSe-подложек (Электростекло; Москва) микропипеткой наносили 5–10 мкл конкретной суспензии, которую затем высушивали на воздухе до полного испарения этанола (5–15 мин). Все спектры регистрировали с помощью FTIR-спектрометра Spectrum Two (Perkin-Elmer; США) в интервале волновых чисел $4000\text{--}600\text{ см}^{-1}$ при оптическом разрешении 4 см^{-1} (цифровое разрешение 1 см^{-1}). Для записи спектров культур изолятов ZnSe-подложку с нанесенной пробой патогена устанавливали вертикально в специальный держатель образцов «Microfocus» (Perkin-Elmer; США) на пути горизонтального зондирующего ИК-луча прибора и аккумулировали с усреднением 16 индивидуальных сканов (около 2 мин). Для записи фонового спектра прибора в идентичных условиях использовали чистую стандартную ZnSe-подложку, а сам фоновый спектр регулярно обновляли незадолго до регистрации ИК-спектра каждой новой пробы патогена.

Анализ данных

После исключения аномальных спектров полученные исходные данные состояли из 347 спектров чистых культур: 188 спектров *S. aureus* (39 — с MRSA-фенотипом, 48 — с MSSA-фенотипом (methicillin-sensitive *S. aureus*), 101 — не охарактеризованные по наличию фенотипа устойчивости к антибиотикам) и 169 спектров других патогенов (14 — *A. baumannii*, 32 — *C. albicans*, 8 — *E. cloacae*, 21 — *E. faecalis*, 20 — *E. faecium*, 11 — *E. coli*, 17 — *K. pneumoniae*,

18 — *P. aeruginosa*, 8 — *S. marcescens*, 188 — *S. aureus*, 10 — *S. epidermidis*); 10 спектров смешанных культур: 2 — *S. aureus* (MRSA) + *E. coli*, 2 — *S. aureus* (MSSA) + *E. coli*, 2 — *S. aureus* (MRSA) + *K. pneumoniae*, 2 — *S. aureus* (MSSA) + *K. pneumoniae*, 2 — *S. aureus* (MSSA) + *P. aeruginosa*. На рис. 1 представлены исходные спектры, использованные при анализе.

С помощью широко известных в спектроскопии рутинных алгоритмов все исходные спектры подвергали предварительной обработке для приведения всего набора к единому формату и устранения артефактов, вызванных, например, искривлением базовой линии спектров или флуктуациями в атмосфере содержания влаги и углекислого газа. Для этого исходные спектры нормировали на среднее пропускание, после чего рассчитывали первую производную от огибающей спектра, а принимаемые в расчет диапазоны спектральных данных ограничивали интервалами волновых чисел 600–1800 см⁻¹ и 2800–3000 см⁻¹. Поскольку исходные спектры изначально были хорошего качества и не требовали предварительного сглаживания, расчет производной проводили с помощью формулы симметричной разности по двум точкам для численного

дифференцирования. На рис. 2 представлены спектры после предобработки.

На основе предобработанных спектров строили модель, позволяющую идентифицировать присутствие в образце *S. aureus*. Кроме того, на основе спектров *S. aureus* по наличию MRSA/MSSA-фенотипа была построена модель, дискриминирующая MRSA- и MSSA-штаммы. Для построения обеих моделей использовали спектры только чистых культур. Спектры смешанных культур использовали только для валидации полученных моделей.

Обе модели строили с помощью пошагового применения метода главных компонент (principal component analysis, PCA) и линейного дискриминантного анализа (linear discriminant analysis, LDA) [4]. Линейный дискриминантный анализ [5] позволяет вычислить линию (или гиперплоскость), в проекции на которую две (или более) группы разделяются наиболее эффективно. Эффективность разделения групп определяется отношением межгрупповой дисперсии к внутригрупповой. Но из-за особенностей матричных вычислений, принятых в методе, при анализе спектральных данных он не используется напрямую, так как имеется большое количество коррелированных признаков и

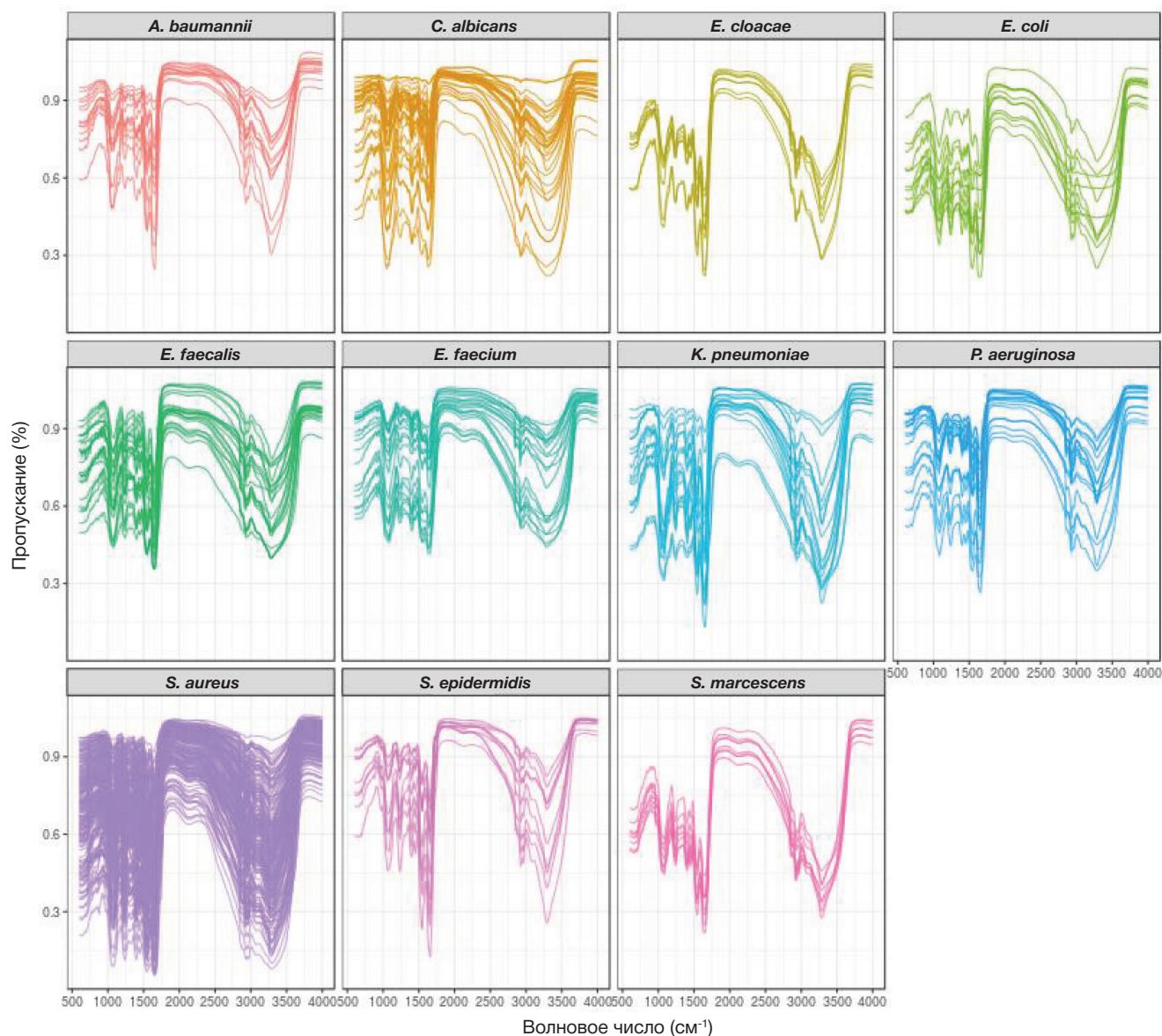


Рис. 1. Исходные инфракрасные Фурье-спектры бактериальных патогенов. Слева направо и сверху вниз в единицах пропускания представлены спектры *A. baumannii*, *C. albicans*, *E. cloacae*, *E. coli*, *E. faecalis*, *E. faecium*, *K. pneumoniae*, *P. aeruginosa*, *S. aureus*, *S. epidermidis*, *S. marcescens*

неинформативных участков. Поэтому перед применением LDA, как правило, используют метод главных компонент, в котором выделяют наиболее информативную и некоррелированную часть информации, имеющуюся в данных. В качестве оценки информативности в методе используется дисперсия: если на каком-то волновом числе дисперсия небольшая, то почти все спектры ведут себя в этой области одинаково и такой участок не несет полезной информации. Выделение информативной и некоррелированной информации в PCA выполняется также путем проецирования данных на соответствующие направления — главные компоненты. Фактически, модель, получаемая с помощью PCA-LDA, является проецированием спектральных данных на новое направление, что на практике означает расчет линейной комбинации с определенными коэффициентами.

Построенные модели валидировали методом скользящего контроля [6]. Суть его заключается в проведении серии слепых контролей: исходные спектры разбивают случайным образом на k частей, одна из частей (тестовая) отбрасывается, а к оставшимся $k - 1$ частям (обучающим) применяется PCA-LDA, и полученная модель

предсказывает значения для тестовой части так, как будто она не была известна на момент построения модели; данная процедура повторяется для каждой из k частей. Усреднив результаты на тестовых частях, можно получить оценку предсказательной модели на новых неизвестных спектрах. Однако при разбиении на k частей важно контролировать, чтобы спектры, относящиеся к одному и тому же изоляту, попадали только в одну часть. В противном случае оценка искусственно завышается. При построении модели, идентифицирующей *S. aureus*, использовали метод скользящего контроля с $k = 20$, т. е. из общей выборки спектров выделяли 20 частей таким образом, что спектры, относящиеся к одному изоляту, всегда попадали в одну и ту же часть. При построении модели, дискриминирующей MSSA- и MRSA-штаммы, использовали равное количество изолятов каждого типа ($k = 40$). Таким образом, в данном случае каждая часть соответствовала одному изоляту.

Все процедуры предобработки и анализа спектральных данных проводили в программной среде R [7] и интегрированной среде разработки RStudio. Для работы со спектральными данными использовали библиотеку hyperSpec [8], для построения и валидации классифицирующих

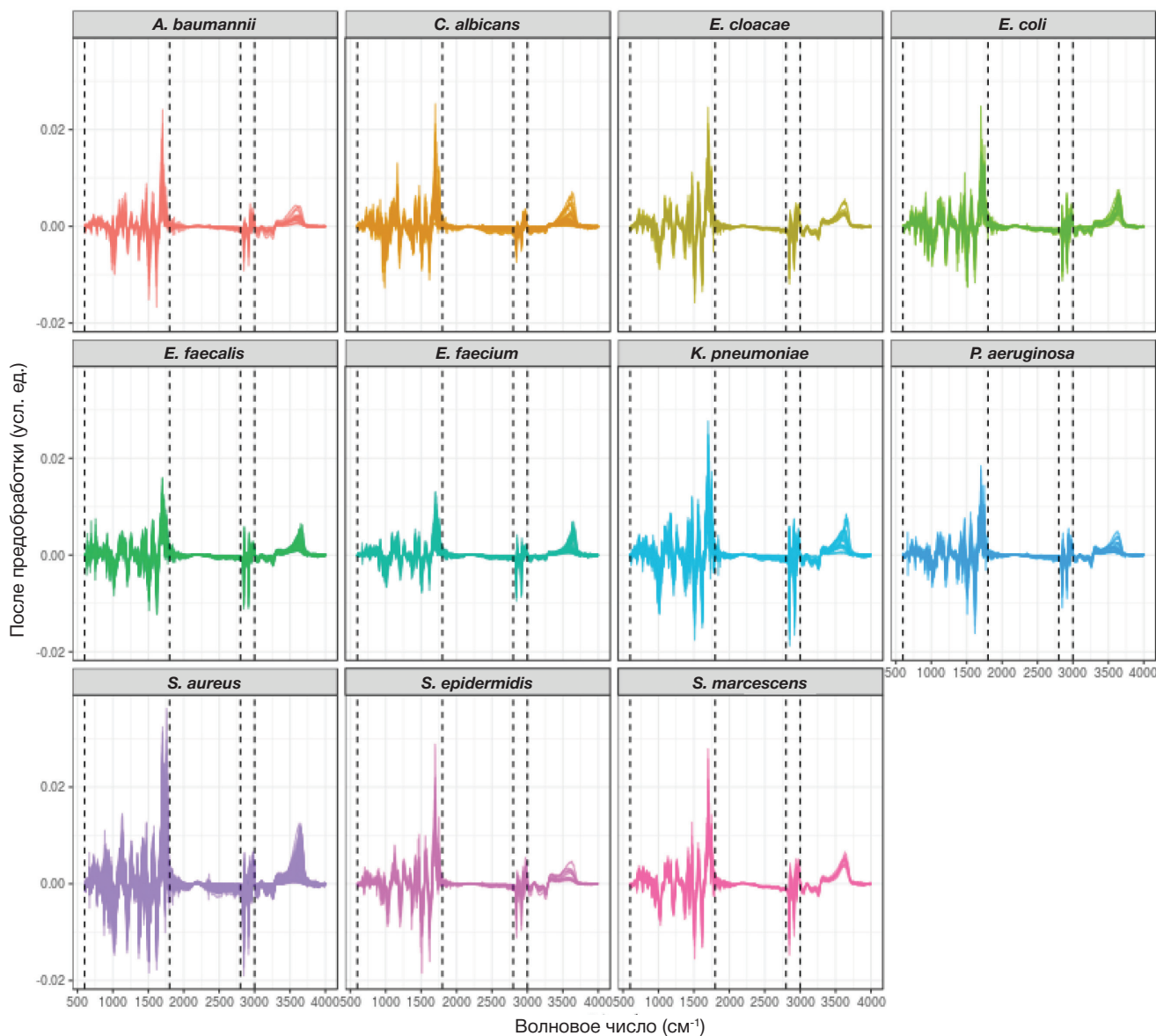


Рис. 2. Исходные Фурье-спектры после предобработки: нормализации на среднее и вычисления первой производной. Пунктирными линиями отмечены границы диапазонов, использовавшихся в анализе: 600, 1800, 2800, 3000 cm^{-1}

моделей — caret [9] и MASS [10]. Иллюстрации были получены с помощью библиотеки ggplot2 [11].

РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ

Идентификация золотистого стафилакокка

На основе полученных спектров 11 типов наиболее часто встречающихся возбудителей гнойно-септических инфекций человека: *S. aureus*, *S. epidermidis*, *E. faecalis*, *E. faecium*, *K. pneumoniae*, *E. coli*, *S. marcescens*, *E. cloacae*, *A. baumannii*, *P. aeruginosa* и *C. albicans* методом PCA-LDA была построена математическая модель, позволяющая идентифицировать *S. aureus*. Точность модели была оценена методом скользящего контроля по 20 блокам (спектры одного изолята попадали в один блок) и составила в среднем 98,4% ($\pm 4\%$). Однако более детальный анализ результатов применения метода скользящего контроля показал, что почти все ошибки были получены тогда, когда в тестовую часть попадал изолят *S. epidermidis*. Это говорит о том, что на уровне рода модель в существующем виде (с учетом размера и качественного состава штаммов в «обучающей» базе данных) работает хорошо. Для достоверной идентификации штаммов на уровне вида в обучающей выборке оказались недостаточно представлены спектры стафилококков других видов: всего в исходных данных, использованных для анализа, присутствовали спектры только двух изолятов *S. epidermidis*. Наихудшая точность,

полученная в методе скользящего контроля (81%), как раз приходится на случай, когда в тестовую часть попадают спектры изолята *S. epidermidis*, так как в обучающей части недостаточно данных для более достоверного различения *S. epidermidis* и *S. aureus*.

Можно предположить, что эти два родственника вида в принципе не различимы с помощью такого подхода. Но финальная модель, построенная по всем данным (без скользящего контроля, когда все спектры использовались при построении модели), демонстрирует способность различать эти два вида со 100% точностью. Эту способность также наглядно демонстрирует проекция спектральных данных на линейный дискриминант (разделяющую линию в методе PCA-LDA) (рис. 3).

Приведенная проекция на линейный дискриминант по сути результат произведения каждого спектра на вектор коэффициентов: если представить (предобработанный) спектр в виде вектора (набора значений), то линейный дискриминант является результатом линейной комбинации. Настройка модели заключается в подборе оптимальных коэффициентов. Значения данных коэффициентов для полученной модели представлены в виде графика на рис. 4.

Отображение коэффициентов позволяет получить грубую интерпретацию полученной модели: чем больше (по абсолютной величине) значение коэффициента, тем более значим соответствующий диапазон; более высокие (по абсолютной величине) значения (с учетом предобработки

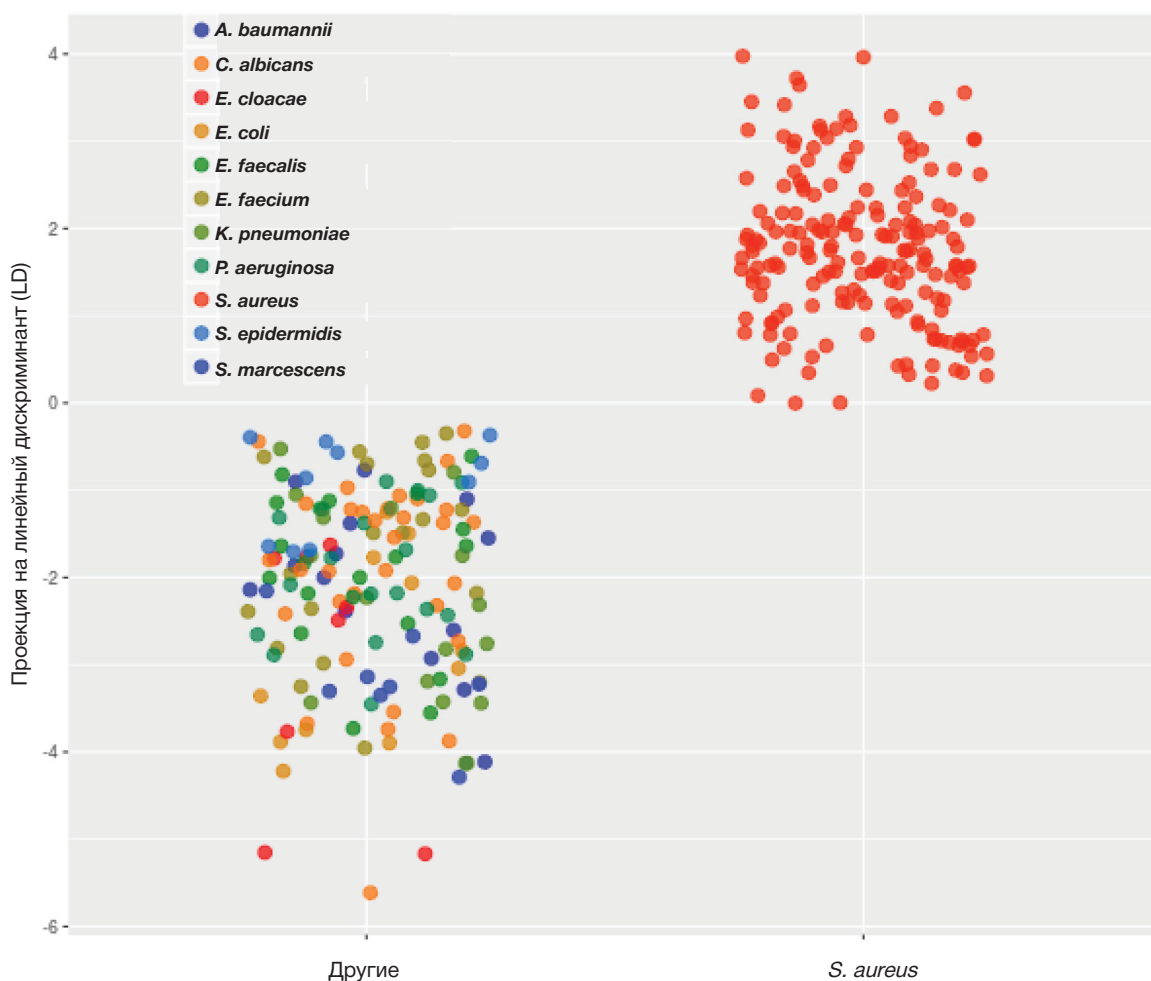


Рис. 3. Проекция исходных данных на линейный дискриминант (LD), полученный с помощью метода PCA-LDA. Цветами отмечены разные патогены. Для удобства восприятия спектры целевого патогена *S. aureus* и других разделены по оси абсцисс. Можно заметить, что группы хорошо разделяются по условию $LD > 0$

и вычисления производной) в областях, где коэффициенты отрицательны, говорят о том, что представленный спектр не относится к *S. aureus*, и наоборот.

Для более развернутой оценки возможности применения подхода в практике медицинской микробиологии была исследована способность полученной модели идентифицировать целевой патоген в смеси с другими бактериями. Для анализа использовали 2 штамма *S. aureus* с фенотипом MRSA и 3 штамма без лекарственной устойчивости. Микроорганизмы выращивали на кровяном бульоне в течение 24 или 48 ч. По окончании культивирования образцы готовили путем смешивания равных объемов культур *S. aureus* с культурами грамотрицательных палочек *E. coli*, *K. pneumoniae* и *P. aeruginosa*, выращенных на бульоне Эндо. Содержание бактериальных тел в единице объема среды не исследовали. Все образцы-смеси, фиксированные спиртом, были представлены на идентификацию финальной модели, упомянутой выше (табл.).

Во всех спектрах, кроме двух, относящихся к одному образцу, с высокой вероятностью был обнаружен целевой патоген, что говорит о надежности модели и перспективности использования подобного подхода на практике.

Предсказание наличия у изолятов *S. aureus* фенотипа лекарственной устойчивости

Важной задачей медицинской микробиологии является возможность определения антибиотикорезистентности патогена. В данной работе была предпринята попытка выявления фенотипа MRSA/MSSA у изолятов *S. aureus* по спектру FTIR. Построение классифицирующей модели происходило аналогичным образом, т. е. методом PCA-LDA с последующей валидацией методом скользящего контроля.

Такого же хорошего качества модели, как в случае идентификации *S. aureus*, получить не удалось. Точность модели, оцененная методом скользящего контроля, составила 73%. Проекция на линейный дискриминант представлена на рис. 5. Точки разделяются существенно хуже, чем в случае дискриминации *S. aureus* от других патогенов. Однако 80% спектров были идентифицированы точно. Это наблюдение позволяет предполагать, что за счет увеличения объема «обучающей» базы данных удастся повысить достоверность идентификации до практически приемлемого уровня.

ОБСУЖДЕНИЕ

Впервые сообщение об использовании FTIR для идентификации микроорганизмов было опубликовано в 1991 г. [12]. Затем вышли работы, посвященные идентификации бактерий во внешней среде, в частности, лактоацилл и возбудителей пищевых токсикоинфекций [13, 14]. В ряде исследований была впервые показана возможность использования FTIR для идентификации таких патогенов, как микобактерии и листерии [15–17]. В 2011 г в связи с появлением коммерчески доступных спектрометров производства Bruker (Германия) и Perkin-Elmer (США), позволяющих достоверно определять колебательные спектры нативных культур микроорганизмов, возросло число публикаций, посвященных использованию FTIR в микробиологии [18–20]. К этому времени относится появление групп, работающих в этой области за пределами Германии: в Польше [21], Великобритании [22, 23] и Нидерландах [24]. Последняя работа интересна тем, что в ней впервые проведена идентификация возбудителей сепсиса у человека, а также приводится сопоставление

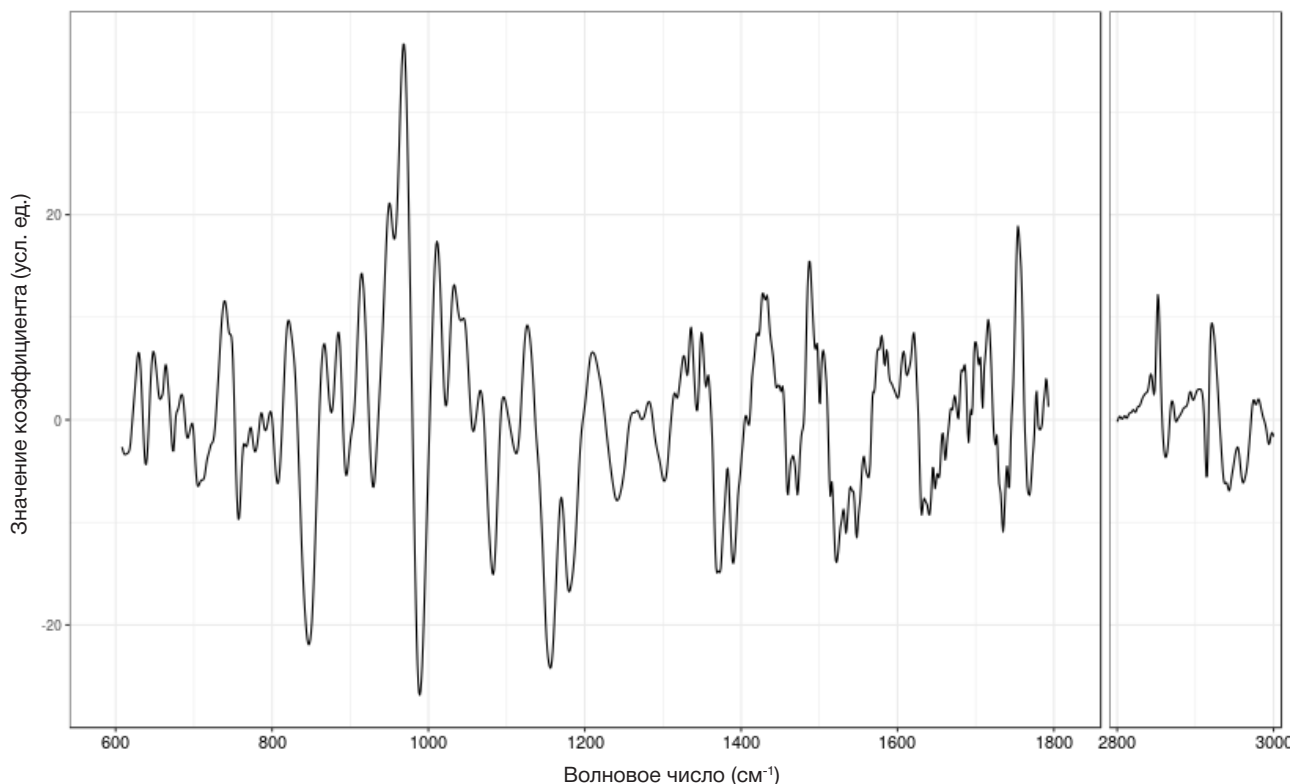


Рис. 4. Коэффициенты линейного дискриминанта, полученного с помощью метода PCA-LDA. Проекция на рис. 3 были получены с помощью этих коэффициентов. Коэффициенты можно интерпретировать как значимость участков со знаком, указывающим на направление влияния. Например, высокие значения исходного спектра в областях, где коэффициенты отрицательны, увеличивают вероятность того, что изолят не представляет собой целевой патоген

разрешающей способности различных методов колебательной спектроскопии: FTIR, спектроскопии Рамановского рассеяния и спектроскопии Рамановского рассеяния с поверхностным усилением (SERS). Авторы делают вывод о том, что FTIR и спектроскопия Рамановского рассеяния дают достоверные и примерно одинаковые результаты, но уступают по чувствительности методу SERS. Метод SERS, напротив, обладает существенно более высокой чувствительностью, но не обеспечивает воспроизводимости результатов идентификации.

В последнее время также появляется много работ, посвященных использованию FTIR в медицинской микробиологии [25–28]. В первой из них на примере ряда патогенных и непатогенных бактерий *P. aeruginosa*, *P. putida*, *E. coli*, *E. faecium*, *Streptomyces lividans*, *B. subtilis*, *B. cereus* и пекарских дрожжей *Saccharomyces cerevisiae* сравнивается разрешающая способность методов колебательной спектроскопии с применением современной приборной базы: SERS (точность 74,9%), спектроскопии Рамановского рассеяния (97,8%) и FTIR (96,2%). В последней работе описывается метод экспресс-идентификации бактерий в форме микроколоний диаметром 50–300 мкм с помощью новейшего Фурье-спектрометра IR-BioType (Bruker): колонии автоматически переносятся с агаризованной среды на подложки из CaF₂, а для анализа спектров методом главных компонент используется искусственная нейронная сеть (ANN), доступная через сервер Bruker.

Результаты цитированных работ свидетельствуют, что FTIR-спектры достаточно полно характеризуют химический состав клеток (включая биополимеры, являющиеся строительным материалом стенок клетки и мембран, а также внутриклеточные ДНК, фосфолипиды, сахара и др.), чтобы обеспечить возможность: (а) надежной межвидовой дискриминации патогенных бактерий; (б) точной идентификации микроорганизмов до вида; (в) выявления с помощью соответствующих цифровых библиотек спектров принадлежности вновь полученного изолята к таксономической группе внутри вида. Решение этих задач посредством экспресс-анализа, основанного исключительно на данных ИК-спектроскопии, позволило бы оптимизировать процедуру антибиотикотерапии для лечения конкретного больного, избегая назначения препаратов, заведомо неэффективных для борьбы с выявленным у него этиологическим агентом. Однако проблемой на пути внедрения этого подхода в практику клинической микробиологии является отсутствие алгоритма, позволяющего в автоматическом режиме анализировать FTIR-спектры культур микроорганизмов. Такой алгоритм должен выявлять те компоненты спектра, которые зависят только от генетических особенностей штамма и не зависят от условий его культивирования: состава среды, фазы роста и степени деградации культуры и т. д. Особенностью же всех процитированных выше работ является стремление авторов стандартизировать условия культивирования микроорганизмов, что трудно достижимо в условиях клинической диагностики и увеличивает время проведения анализа.

В рамках нашей работы была решена задача по разработке алгоритма, позволяющего идентифицировать видовую принадлежность бактерий вне зависимости от фазы роста и среды культивирования. Для этого наборы клинических изолятов бактерий видов *S. aureus*, *E. faecalis*, *E. faecium*, *K. pneumoniae*, *E. coli*, *S. marcescens*, *E. cloacae*, *A. baumannii*, *P. aeruginosa*, *S. epidermidis* и *C. albicans* культивировали в различных условиях: на нескольких

средах в течение различного времени. Используя метод главных компонент, вычленили наиболее информативную часть FTIR-спектра бактериальных культур. Результат анализа фиксировали в виде системы коэффициентов, которая позволяла за короткое время идентифицировать видовую принадлежность новых изолятов по их FTIR-спектрам. Точность нового метода была оценена в слепом тесте на модели изолятов *S. aureus* и их смесей с *P. aeruginosa*, *E. coli* и *K. pneumoniae*.

Результаты показывают, что предложенный алгоритм анализа FTIR-спектров после обучения на выборке из 11 патогенов различной таксономической принадлежности (бактерии и дрожжи-аскомицеты) достоверно выявляет наличие в культуре *S. aureus* вне зависимости от длительности культивирования (24 или 48 ч). Все образцы

Таблица. Предсказанная вероятность присутствия в образце *S. aureus*

№ образца	Смесь	Предсказанная вероятность присутствия в образце <i>S. aureus</i>
1	MSSA + <i>K. pneumoniae</i>	96,8%
1	MSSA + <i>K. pneumoniae</i>	95,1%
2	MSSA + <i>E. coli</i>	40,0%
2	MSSA + <i>E. coli</i>	41,5%
3	MRSA + <i>K. pneumoniae</i>	96,4%
3	MRSA + <i>K. pneumoniae</i>	97,0%
4	MRSA + <i>P. aeruginosa</i>	90,6%
4	MRSA + <i>P. aeruginosa</i>	87,7%
5	MRSA + <i>E. coli</i>	82,4%
5	MRSA + <i>E. coli</i>	73,0%

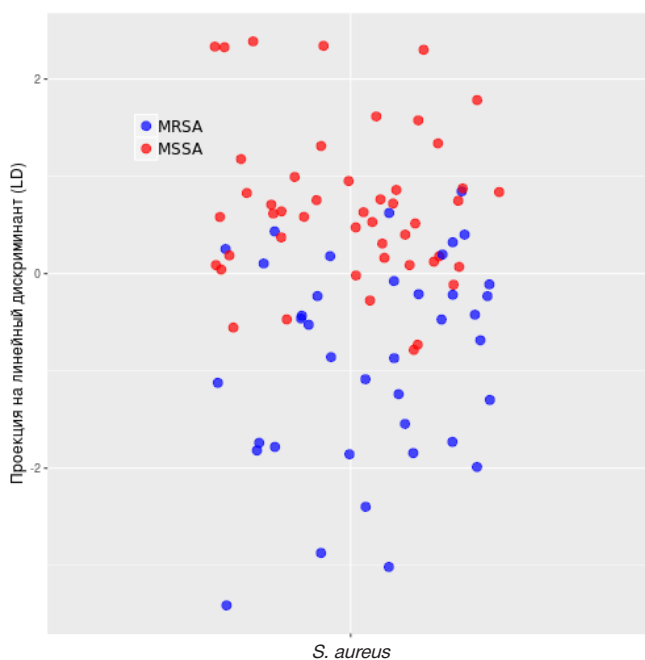


Рис. 5. Проекция на линейный дискриминант, разделяющий штаммы *S. aureus* с фенотипами MSSA и MRSA. Красным отмечены проекции спектров штаммов MSSA, синим MRSA. Группы не разделяются со 100% точностью, однако 80% точек разделяются правильно по условию LD > 0

до снятия спектра инактивировались 70% этанолом, что облегчает организацию работы с вирулентными штаммами и хранение образцов для повторного анализа. Присутствие в образце цельной крови и примесей посторонних микроорганизмов (грамотрицательных палочек *E. coli*, *K. pneumoniae* и *P. aeruginosa*) в концентрациях, примерно равных концентрации *S. aureus*, не снижает способность алгоритма выявлять целевой патоген. Модель с точностью до 80% может предсказывать наличие у изолятов *S. aureus* фенотипа лекарственной устойчивости (MRSA/MSSA). Результаты тестирования позволяют надеяться, что описанный алгоритм после необходимого расширения «обучающей» базы данных будет способен выявлять любые другие патогены, культивируемые на средах любого удобного для работы состава.

ВЫВОДЫ

Описан способ формирования баз данных и программы сравнения, пригодных для идентификации патогенных микроорганизмов с помощью инфракрасной Фурье-спектроскопии вне зависимости от стадии роста культуры и состава среды. Способ может быть успешно реализован на базе стандартного и доступного по стоимости серийного прибора Spectrum Two (Perkin-Elmer; США). Проведена апробация способа на модели клинических изолятов *S. aureus*, которые достоверно дискриминировались от других распространенных возбудителей гнойных и внутрибольничных инфекций: *E. faecalis*, *E. faecium*, *K. pneumoniae*, *E. coli*, *S. marcescens*, *E. cloacae*, *A. baumannii*, *P. aeruginosa* и *C. albicans* в виде чистых культур и парных смесей.

Литература

1. Попов Д. А. Послеоперационные инфекционные осложнения в кардиохирургии. *Анналы хирургии*. 2013; (5): 15–21.
2. Припутневич Т. В., Любасовская Л. А., Дубоделов Д. В., Мелкумян А. Р., Игонина Е. П., Акимкин В. Г., Дегтярев Д. Н., Сухих Г. Т. Эффективная профилактика и лечение ИСМП в родовспомогательных учреждениях Российской Федерации: нерешенные вопросы организации и контроля. *Вестник Росздравнадзора*. 2017; (4): 34–41.
3. Попов Д. А., Надточей Е. А. Алгоритм диагностики бактериемии у кардиохирургических больных в орит. *Анестезиология и реаниматология*. 2017; 62 (5): 382–7.
4. Shlens J. A tutorial on Principal Components Analysis. April 7, 2014; Version 3.02.
5. Fisher RA. The Use of Multiple Measurements in Taxonomic Problems. *Annals of Eugenics*. 1936; 7 (2): 179–188.
6. Hastie T, Tibshirani R, Friedman J. 4.3 Linear Discriminant Analysis. *The Elements of Statistical Learning* (2 ed.). New York: Springer; 2009. 763 p.
7. Hastie T, Tibshirani R, Friedman J. 7.10 Cross-Validation. *The Elements of Statistical Learning* (2 ed.). New York: Springer; 2009.
8. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, 2016. Vienna, Austria. URL <https://www.R-project.org/>
9. Kuhn M. Contributions from Jed Wing, Steve Weston, Andre Williams, Chris Keefer, Allan Engelhardt, Tony Cooper, Zachary Mayer, Brenton Kenkel, the R Core Team, Michael Benesty, Reynald Lescarbeau, Andrew Ziem, Luca Scrucca, Yuan Tang and Can Candan. *Caret: Classification and Regression Training*. R package version 6.0-71. 2016. <https://CRAN.R-project.org/package=caret>
10. Venables WN, Ripley BD. *Modern Applied Statistics with S*. New York: Springer; 2002. 495 p.
11. Wickham H. *ggplot2: Elegant Graphics for Data Analysis*. New York: Springer; 2009.
12. Naumann D, Helm D, Labischinski H. Microbiological characterizations by FT-IR spectroscopy. *Nature*. 1991; 351 (6321): 81–2.
13. Oust A, Mørretrø T, Kirschner C, Narvhus JA, Kohler A. Evaluation of the robustness of FT-IR spectra of lactobacilli towards changes in the bacterial growth conditions. *FEMS Microbiol Lett*. 2004 Oct 1; 239 (1): 111–6.
14. Wenning M, Theilmann V, Scherer S. Rapid analysis of two food-borne microbial communities at the species level by Fourier-transform infrared microspectroscopy. *Environ Microbiol*. 2006 May; 8 (5): 848–57.
15. Miguel Gómez MA, Bratos Pérez MA, Martín Gil FJ, Dueñas Díez A, Martín Rodríguez JF, Gutiérrez Rodríguez P, Orduña Domingo A, Rodríguez Torres A. Identification of species of *Brucella* using Fourier transform infrared spectroscopy. *J Microbiol Methods*. 2003 Oct; 55 (1): 121–31.
16. Rebuffo-Scheer CA, Kirschner C, Staemmler M, Naumann D. Rapid species and strain differentiation of non-tuberculous mycobacteria by Fourier-Transform Infrared microspectroscopy. *J Microbiol Methods*. 2007 Feb; 68 (2): 282–90. Epub 2006 Oct 19.
17. Rebuffo-Scheer CA, Dietrich J, Wenning M, Scherer S. Identification of five *Listeria* species based on infrared spectra (FTIR) using macrosamples is superior to a microsample approach. *Anal Bioanal Chem*. 2008 Mar; 390 (6): 1629–35.
18. Schäwe R, Fetzer I, Tönniges A, Härtig C, Geyer W, Harms H, Chatzinotas A. Evaluation of FT-IR spectroscopy as a tool to quantify bacteria in binary mixed cultures. *J Microbiol Methods*. 2011 Aug; 86 (2): 182–7.
19. Wenning M, Scherer S. Identification of microorganisms by FTIR spectroscopy: perspectives and limitations of the method. *Appl Microbiol Biotechnol*. 2013 Aug; 97(16): 7111–20.
20. Wenning M, Breitenwieser F, Konrad R, Huber I, Busch U, Scherer S. Identification and differentiation of food-related bacteria: A comparison of FTIR spectroscopy and MALDI-TOF mass spectrometry. *J Microbiol Methods*. 2014 Aug; 103: 44–52.
21. Rygula A, Jekiel K, Szostak-Kot J, Wrobel TP, Baranska M. Application of FT-Raman spectroscopy for in situ detection of microorganisms on the surface of textiles. *J Environ Monit*. 2011 Nov; 13 (11): 2983–7.
22. Driver T, Bajhaiya AK, Allwood JW, Goodacre R, Pittman JK, Dean AP. Metabolic responses of eukaryotic microalgae to environmental stress limit the ability of FT-IR spectroscopy for species identification. *Algal Res*. 2015 Sep; 11: 148–155.
23. Alvarez-Ordóñez A, Mouwen DJ, López M, Prieto M. Fourier transform infrared spectroscopy as a tool to characterize molecular composition and stress response in foodborne pathogenic bacteria. *J Microbiol Methods*. 2011 Mar; 84 (3): 369–78.
24. Maquelin K, Kirschner C, Choo-Smith LP, Ngo-Thi NA, van Vreeswijk T, Stämmler M, Endtz HP, Bruining HA, Naumann D, Puppels GJ. Prospective study of the performance of vibrational spectroscopies for rapid identification of bacterial and fungal pathogens recovered from blood cultures. *J Clin Microbiol*. 2003 Jan; 41 (1): 324–9.
25. Muhamadali H, Subaihi A, Mohammadtaheri M, Xu Y, Ellis DI, Ramanathan R, Bansal V, Goodacre R. Rapid, accurate, and comparative differentiation of clinically and industrially relevant microorganisms via multiple vibrational spectroscopic fingerprinting. *Analyst*. 2016 Aug 15; 141 (17): 5127–36.
26. Lasch P, Stämmler M, Zhang M, Baranska M, Bosch A, Majzner K. FT-IR Hyperspectral Imaging and Artificial Neural Network Analysis for Identification of Pathogenic Bacteria. *Anal Chem*. 2018 Jul 11.
27. Bishop CM. *Pattern Recognition and Machine Learning*. New York: Springer; 2006. ISBN-10: 0-387-31073-8.
28. Beleites C, Sergo V. *HyperSpec: a package to handle hyperspectral data sets in R*, R package version 0.98-20150304. <http://hyperspec.r-forge.r-project.org/>

References

1. Popov DA. Postoperative infectious complications in cardiac surgery. *Annaly hirurgii*. 2013; (5): 15–21.
2. Priputnevich TV, Ljubasovskaja LA, Dubodelov DV, Melkumjan AR, Igonina EP, Akimkin VG, Degtjarev DN, Suhih GT. Jeffektivnaja profilaktika i lechenie ISMP v rodovspomogatel'nyh uchrezhdenijah Rossijskoj Federacii: nereshennye voprosy organizacii i kontrolja. *Vestnik Roszdravnadzora*. 2017; (4): 34–41.
3. Popov DA, Nadtochej EA. Algorithm of bacteremia diagnostic in cardiosurgical patients in ICU. *Anesteziol Reanimatol*. 2017; 62 (5): 382–7.
4. Shlens J. A tutorial on Principal Components Analysis. April 7, 2014; Version 3.02.
5. Fisher RA. The Use of Multiple Measurements in Taxonomic Problems. *Annals of Eugenics*. 1936; 7 (2): 179–188.
6. Hastie T, Tibshirani R, Friedman J. 4.3 Linear Discriminant Analysis. *The Elements of Statistical Learning* (2 ed.). New York: Springer; 2009. 763 p.
7. Hastie T, Tibshirani R, Friedman J. 7.10 Cross-Validation. *The Elements of Statistical Learning* (2 ed.). New York: Springer; 2009.
8. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, 2016. Vienna, Austria. URL <https://www.R-project.org/>
9. Kuhn M. Contributions from Jed Wing, Steve Weston, Andre Williams, Chris Keefer, Allan Engelhardt, Tony Cooper, Zachary Mayer, Brenton Kenkel, the R Core Team, Michael Benesty, Reynald Lescarbeau, Andrew Ziem, Luca Scrucca, Yuan Tang and Can Candan. *Caret: Classification and Regression Training*. R package version 6.0-71. 2016. <https://CRAN.R-project.org/package=caret>
10. Venables WN, Ripley BD. *Modern Applied Statistics with S*. New York: Springer; 2002. 495 p.
11. Wickham H. *ggplot2: Elegant Graphics for Data Analysis*. New York: Springer; 2009.
12. Naumann D, Helm D, Labischinski H. Microbiological characterizations by FT-IR spectroscopy. *Nature*. 1991; 351 (6321): 81–2.
13. Oust A, Møretø T, Kirschner C, Narvhus JA, Kohler A. Evaluation of the robustness of FT-IR spectra of lactobacilli towards changes in the bacterial growth conditions. *FEMS Microbiol Lett*. 2004 Oct 1; 239 (1): 111–6.
14. Wenning M, Theilmann V, Scherer S. Rapid analysis of two food-borne microbial communities at the species level by Fourier-transform infrared microspectroscopy. *Environ Microbiol*. 2006 May; 8 (5): 848–57.
15. Miguel Gómez MA, Bratos Pérez MA, Martín Gil FJ, Dueñas Díez A, Martín Rodríguez JF, Gutiérrez Rodríguez P, Orduña Domingo A, Rodríguez Torres A. Identification of species of *Brucella* using Fourier transform infrared spectroscopy. *J Microbiol Methods*. 2003 Oct; 55 (1): 121–31.
16. Rebuffo-Scheer CA, Kirschner C, Staemmler M, Naumann D. Rapid species and strain differentiation of non-tuberculous mycobacteria by Fourier-Transform Infrared microspectroscopy. *J Microbiol Methods*. 2007 Feb; 68 (2): 282–90. Epub 2006 Oct 19.
17. Rebuffo-Scheer CA, Dietrich J, Wenning M, Scherer S. Identification of five *Listeria* species based on infrared spectra (FTIR) using macrosamples is superior to a microsample approach. *Anal Bioanal Chem*. 2008 Mar; 390 (6): 1629–35.
18. Schawe R, Fetzer I, Tönniges A, Härtig C, Geyer W, Harms H, Chatzinotas A. Evaluation of FT-IR spectroscopy as a tool to quantify bacteria in binary mixed cultures. *J Microbiol Methods*. 2011 Aug; 86 (2): 182–7.
19. Wenning M, Scherer S. Identification of microorganisms by FTIR spectroscopy: perspectives and limitations of the method. *Appl Microbiol Biotechnol*. 2013 Aug; 97(16): 7111–20.
20. Wenning M, Breitenwieser F, Konrad R, Huber I, Busch U, Scherer S. Identification and differentiation of food-related bacteria: A comparison of FTIR spectroscopy and MALDI-TOF mass spectrometry. *J Microbiol Methods*. 2014 Aug; 103: 44–52.
21. Rygula A, Jekiel K, Szostak-Kot J, Wrobel TP, Baranska M. Application of FT-Raman spectroscopy for in situ detection of microorganisms on the surface of textiles. *J Environ Monit*. 2011 Nov; 13 (11): 2983–7.
22. Driver T, Bajhaiya AK, Allwood JW, Goodacre R, Pittman JK, Dean AP. Metabolic responses of eukaryotic microalgae to environmental stress limit the ability of FT-IR spectroscopy for species identification. *Algal Res*. 2015 Sep; 11: 148–155.
23. Alvarez-Ordóñez A, Mouwen DJ, López M, Prieto M. Fourier transform infrared spectroscopy as a tool to characterize molecular composition and stress response in foodborne pathogenic bacteria. *J Microbiol Methods*. 2011 Mar; 84 (3): 369–78.
24. Maquelin K, Kirschner C, Choo-Smith LP, Ngo-Thi NA, van Vreeswijk T, Stämmler M, Endtz HP, Bruining HA, Naumann D, Puppels GJ. Prospective study of the performance of vibrational spectroscopies for rapid identification of bacterial and fungal pathogens recovered from blood cultures. *J Clin Microbiol*. 2003 Jan; 41 (1): 324–9.
25. Muhamadali H, Subaihi A, Mohammadtaheri M, Xu Y, Ellis DI, Ramanathan R, Bansal V, Goodacre R. Rapid, accurate, and comparative differentiation of clinically and industrially relevant microorganisms via multiple vibrational spectroscopic fingerprinting. *Analyst*. 2016 Aug 15; 141 (17): 5127–36.
26. Lasch P, Stämmler M, Zhang M, Baranska M, Bosch A, Majzner K. FT-IR Hyperspectral Imaging and Artificial Neural Network Analysis for Identification of Pathogenic Bacteria. *Anal Chem*. 2018 Jul 11.
27. Bishop CM. *Pattern Recognition and Machine Learning*. New York: Springer; 2006. ISBN-10: 0-387-31073-8.
28. Beleites C, Sergio V. *HyperSpec: a package to handle hyperspectral data sets in R*, R package version 0.98-20150304. <http://hyperspec.r-forge.r-project.org/>