

## ЭКСПЕРИМЕНТАЛЬНЫЕ ИССЛЕДОВАНИЯ

© Коллектив авторов

### МАСС-СПЕКТРОМЕТРИЧЕСКИЙ MRM АНАЛИЗ FDA-ВЕРИФИЦИРОВАННЫХ БЕЛКОВ В ПЛАЗМЕ КРОВИ ЗДОРОВЫХ ДОБРОВОЛЬЦЕВ

*С.Е. Новикова<sup>1\*</sup>, Т.Е. Фарафонова<sup>1</sup>, О.В. Тихонова<sup>1</sup>, Н.А. Шушкова<sup>1</sup>, М.А. Пятницкий<sup>1</sup>,  
В.Г. Згода<sup>1</sup>, Е.А. Пономаренко<sup>1</sup>, А.В. Лисица<sup>1</sup>, А.И. Григорьев<sup>2</sup>, В.А. Тутельян<sup>3</sup>, А.И. Арчаков<sup>1</sup>*

<sup>1</sup>НИИ биомедицинской химии имени В.Н. Ореховича,  
119121, Москва, ул. Погодинская, 10; \*эл. почта: novikova.s.e3101@gmail.com

<sup>2</sup>Институт медико-биологических проблем РАН, 123007, Москва, Хорошевское шоссе, 76а

<sup>3</sup>Федеральный исследовательский центр питания и биотехнологии, 109240, Москва, Устьинский проезд, 2/14

Протеомный состав служит важнейшей характеристикой биологического объекта, позволяющей оценить его состояние. Направленный масс-спектрометрический анализ, а именно мониторинг множественных реакций (multiple reaction monitoring, MRM) с использованием синтетических изотопно-меченых пептидов в качестве внутренних стандартов (SIS) является основной альтернативой методу ИФА для анализа диагностически значимых белков. По результатам работ с применением MRM создан прототип тест-системы, использующей целевой масс-спектрометрический метод для мультиплексного количественного анализа белков, утвержденных Управлением по санитарному надзору за качеством пищевых продуктов и медикаментов (Food and Drug Administration, FDA) в качестве биомаркеров, в цельной плазме крови. С использованием данного подхода удалось измерить содержание 42 белков в 31 образце в диапазоне концентраций, охватывающем 5 порядков величин. При этом, межиндивидуальная вариабельность для 30 из 42 зарегистрированных белков оказалась менее 40%. Наибольший разброс наблюдался при измерении содержания гаптоглобина (68%), константного участка тяжелой цепи иммуноглобулина класса D (90%), ангиотензина (72%), связывающего половые стероиды глобулина SHBG (100%) и липопротеин-(а) (136%). Полученные результаты по концентрации белков коррелируют с литературными данными (Hortin et al., 2008, *Clinical Chemistry*, **54**, 1608) с  $R^2=0,84$ . Разработанный прототип тест-системы, основанной на масс-спектрометрическом направленном анализе белков, можно рассматривать в качестве альтернативы методикам, использующим моноклональные антитела.

**Ключевые слова:** MRM анализ; изотопно-меченые внутренние пептидные стандарты; FDA-верифицированные белки  
**DOI:** 10.18097/PBMC20206604294

## ВВЕДЕНИЕ

Определение белковых маркеров представляет собой основной инструмент клинической химии и лежит в основе диагностики многих патологических состояний, в том числе таких социально-значимых заболеваний как инфаркт миокарда, нарушение функций печени и почек, диабет, гормональные нарушения, воспаление и онкологические заболевания.

На сегодняшний день “золотым стандартом” в данной области является определение белков в плазме или сыворотке крови с помощью различных типов иммуноферментного анализа (ИФА), основанного на использовании антител: иммунотурбидиметрии, иммунохемилюминесценции и иммунофлуоресценции. Использование антител обеспечивает высокую чувствительность, специфичность и точность анализа; при этом недостатками метода остаются кросс-реактивность антител, низкая адаптируемость к мультиплексному анализу или одновременному анализу множества белковых аналитов, а также высокая стоимость разработки новых тестов.

В то же время всё больше исследований указывают на значимость анализа не отдельно взятого биомаркера, а панели белков, часто включающих десятки аналитов, в то время как для целей прецизионной медицины хорошо иметь представление

сразу о всех диагностически значимых белках [1, 2]. Важно отметить различие между истинно мультиплексными тестами, которые позволяют измерить несколько белков в одном и том же образце одновременно и с применением одного и того же метода, и комплексными тестами (bundled tests), которые представляют собой независимые анализы, результаты которых рассматриваются вместе. Явление кросс-реактивности мешает технической реализации истинно мультиплексного теста для анализа десятков и сотен белков с использованием антител. В то же время мультиплексный анализ позволит повысить его чувствительность и специфичность для диагностики [3].

Масс-спектрометрические методы, бурно развивающиеся в последние годы, рассматривают в качестве альтернативы методикам с использованием антител. Особое значение имеют количественные подходы, поскольку для диагностических целей важно оценить содержание белковых аналитов в биологическом материале. Панорамный количественный анализ высокого разрешения без использования стабильных изотопных меток применяют для исследования белков плазмы крови как наиболее распространенного и малоинвазивного биоматериала. Данный подход позволил провести количественный анализ 345 белков *in silico* на основании расчётного параметра — нормализованной

интенсивности пептидного иона (LFQ intensity) [4]. При этом внутрииндивидуальная вариабельность ( $n=15$ ) оказалась ниже 20% для 55% всех белков, в то время как только 19% белков показали межиндивидуальную вариабельность ( $n=10$ ) ниже 20% [4].

Направленная или таргетная масс-спектрометрия, благодаря её высокой чувствительности, селективности, мультиплексному характеру анализа и возможности количественной оценки белковых аналитов, представляет особенный интерес для применения в диагностике. В основе мониторинга множественных реакций (multiple reaction monitoring, MRM) лежит регистрация сигнала для протеотипических пептидов, которые можно рассматривать как уникальные “зонды” для каждого диагностически значимого белка. Высокая чувствительность метода позволяет проводить анализ диагностически значимых белков в цельной плазме крови, где динамический диапазон концентраций белков превышает 10 порядков [5]. Помимо хорошей чувствительности и точности, метод MRM обеспечивает частичное секвенирование аминокислотной последовательности пептида, что определяет более высокую специфичность по сравнению с методами, основанными на использовании антител. Низкая стоимость и простота синтеза новых пептидных стандартов SIS, по сравнению с получением новых моноклональных антител, позволяет гибко модифицировать и расширять уже существующую панель аналитов, добавляя новые белки и перегруппировывая целевые белки. Такой модульный характер количественного MRM метода делает возможным создание тест-систем с учётом индивидуальных потребностей каждого человека (т.н. custom-tailored assay), как болезнь-ориентированных, так и ориентированных на мониторинг и оценку персонального физиологического статуса обследуемого. Кроме того количественная информация о белках, дополненная геномными и метаболомными данными, может обеспечить системный взгляд на молекулярный “портрет” индивида, что соответствует целям прецизионной медицины.

На настоящий момент с помощью направленной масс-спектрометрии получены внушительные объёмы данных, однако большая их часть представляет собой результат анализа биологического материала, полученного от пациентов с различными заболеваниями. В то же время представление о количественном и качественном составе протеома здорового человека остается неполным. При этом физиологически обусловленную разницу в содержании белков, или своего рода физиологический “шум”, необходимо учитывать при разработке и валидации методов определения белковых маркеров с помощью масс-спектрометрии.

Таким образом, нами была поставлена цель определить биологическую вариабельность белков, утверждённых Управлением по санитарному надзору за качеством пищевых продуктов и медикаментов (Food and Drug Administration, FDA) в качестве

биомаркеров, в образцах, полученных от здоровых добровольцев, с помощью метода MRM с использованием изотопно-меченых пептидных стандартов для количественной оценки.

## МЕТОДИКА

### *Биологический материал*

В работе использовали образцы плазмы крови человека, полученной от 23 здоровых добровольцев (10 мужчин и 13 женщин), в возрасте от 23 до 42 лет (медиана 30 лет), которые в течение 2 дней находились под медицинским наблюдением в клинике лечебного питания Федерального исследовательского центра питания и биотехнологии. Для данной группы добровольцев были выполнены общий анализ крови и мочи, а также биохимический анализ крови. Плазма крови от 8 здоровых мужчин, кандидатов в космонавты, была предоставлена Институтом медико-биологических проблем РАН.

Венозную кровь отбирали в вакуумные пробирки с  $K_2$  ЭДТА; для получения плазмы кровь центрифугировали сразу после забора при 1300 g в течение 10 мин. Визуально оценивали наличие гемолиза. Полученную плазму аликвотили по 200 мкл и хранили при  $-80^{\circ}\text{C}$  до анализа.

### *Выбор белков, протеотипических пептидов и синтез изотопно-меченых внутренних стандартов (SIS)*

В работе использованы белки, одобренные FDA для использования в качестве клинко-диагностических маркеров. За основу был взят список из 109 уникальных белковых аналитов, опубликованных Anderson и соавт. [7], для которых были приведены в соответствие 139 уникальных записей базы данных Uniprot (<https://www.uniprot.org>).

Протеотипические пептиды выбирали в базе данных NextProt (<https://www.nextprot.org>), проверку уникальности аминокислотной последовательности пептида в пределах биологического вида проводили с помощью ресурса Peptide unicity checker (<https://www.nextprot.org/viewers/unicity-checker/app/index.html>). Условиями выбора кандидатных пептидов для синтеза были преимущественная длина — 9 и более аминокислотных остатков, обязательное отсутствие аминокислотных остатков глутамина (Q) и глутаминовой кислоты (E) на N-конце аминокислотной последовательности пептида, обязательное отсутствие метионина (M) и преимущественное отсутствие цистеина (C). В результате были выбраны 246 пептидов, картируемых на 120 уникальных записей базы данных Uniprot.

Твердофазный пептидный синтез осуществляли на автоматическом пептидном синтезаторе Overture (“Protein Technologies”, Великобритания), как описано ранее [6]. При синтезе изотопно-меченых пептидов вместо обычного лизина или аргинина использовали изотопно-меченые аминокислоты Fmoc-Lys-OH- $^{13}\text{C}_6,^{15}\text{N}$ , или Fmoc-Arg-OH- $^{13}\text{C}_6,^{15}\text{N}$  (“Cambridge Isotope Laboratories”, США).

*Пробоподготовка образцов  
к масс-спектрометрическому анализу*

Содержание общего белка в образцах плазмы крови определяли колориметрическим методом с помощью коммерческого набора Pierce™ BCA Protein Assay Kit (“Pierce”, США) в соответствии с рекомендациями производителя. Для анализа использовали аликвоты образцов плазмы, содержащие 170 мкг белка.

Для восстановления и алкилирования дисульфидных связей аликвоту плазмы крови, содержащую 170 мкг общего белка, инкубировали в присутствии 30 мМ трис(2-карбоксиэтил)фосфина (TCEP), 50 мМ 2-хлорацетамида (CAA) и 0,2% сурфактанта ProteaseMax при температуре 80°C в течение 40 мин. Затем реакционную смесь разбавляли в 6 раз 50 мМ триэтиламоний бикарбонатным буфером (ТЭАБ, pH 8,5) и добавляли сурфактант ProteaseMax до конечной концентрации 0,015%, а также раствор трипсина с концентрацией 0,5 мкг/мкл в соотношении масса трипсина к массе общего белка 1:70, затем пробу инкубировали в течение ночи. По окончании инкубации к реакционной смеси добавляли муравьиную кислоту до конечной концентрации 5% с последующим центрифугированием при 14000 g в течение 5 мин. Надосадочную жидкость высушивали и растворяли в 42 мкл раствора, содержащего 0,1% муравьиную кислоту и 148 стандартов SIS, содержание которых варьировало от 14 фмоль/мкг до 170 фмоль/мкг общего белка.

*Масс-спектрометрический анализ в режиме SRM*

Каждый экспериментальный образец анализировали в трёх технических повторах.

Хроматографическое разделение осуществлялось с помощью системы Agilent 1200 series (“Agilent Technologies”, США), соединённой с тройным квадрупольным масс-анализатором TSQ Quantiva (“Thermo Scientific”, США). Образец объёмом 3,5 мкл, содержащий 15 мкг нативных пептидов и стандарты SIS, разделяли с использованием аналитической колонки ZORBAX SB-C18 (150×0,5 мм, диаметр частиц 5 мкм, “Agilent Technologies”) в градиенте ацетонитрила со скоростью потока 20 мкл/мин. В начале колонку уравнивали 5% раствором В (80% ацетонитрил в 0,1% муравьиной кислоте) и 95% раствором А (0,1% муравьиная кислота) в течение 5 мин, затем концентрацию раствора В линейно увеличивали до 60% за 30 мин, после чего концентрацию раствора В увеличивали до 99% за 1 мин и в течение 5 мин промывали колонку 99% раствором В, затем в течение 1 мин концентрация возвращалась к начальным условиям, в которых колонка уравнивалась в течение 9 мин. Масс-спектрометрический анализ осуществляли в режиме динамического мониторинга множественных реакций (dMRM), используя следующие настройки МС детектора: напряжение на капилляре составляло 4000 В, скорость осушающего газа (азот) составляла 7 л в мин, скорость аксиального газа (азот) составляла 5 л в мин, температура капилляра — 350°C, окно изоляции для первого и третьего квадрупольных — 0,7 Да, время цикла сканирования — 1,2 с, давление газа

(аргон) в ячейке соударения — 1,5 мТорр. Окно времени удерживания на колонке обращённой фазы составило 1,8 мин для каждого прекурсорного иона. В ходе анализа мониторовались переходы или транзиции, приведённые в таблице 1 дополнительных материалов, лист “Таблица MRM транзиций”. Для MRM анализа целевых белков применяли по 2-4 транзиции на прекурсор. Для количественной оценки в ПО Skyline использовались все мониторируемые транзиции.

Перед постановкой серии образцов была осуществлена калибровка тройного квадрупольного масс-анализатора TSQ Quantiva. После калибровки и перед началом анализа был выполнен 1 вкол пустой пробы и 1 вкол смеси пептидных стандартов в 0,1% муравьиной кислоте. Измерения всех образцов были проведены одной серией. Между каждыми тремя техническими повторами для экспериментальных образцов, осуществлялся вкол пустой пробы, представляющей собой раствор 0,1% муравьиной кислоты в воде, в таком же объёме и при таких же хромато-масс-спектрометрических условиях, как и для экспериментального образца. По окончании анализа серии образцов был выполнен вкол смеси пептидных стандартов в 0,1% муравьиной кислоте.

*Биоинформатический анализ*

Результаты были проверены с использованием программного обеспечения Skyline MacCoss Lab Software (версия 4.1.0) для сравнения хроматографических профилей эндогенного пептида и соответствующего стандарта SIS. Соотношение площадей под пиком для эндогенного пептида и соответствующего стандарта SIS рассчитывалось в Skyline в автоматическом режиме. Для определения количества белка соотношение, рассчитанное в Skyline, умножали на известное содержание каждого стандарта SIS. За измерение каждого FDA-одобренного белка принимали среднее арифметическое значение содержания, рассчитанное на основании результатов MRM анализа в трёх технических повторах для образца плазмы крови после гидролитического расщепления трипсином, выполненного в одном повторе.

Статистический анализ и отображение полученных результатов было выполнено с применением открытого языка программирования R версии 3.6.0 [www.r-project.org], использовались библиотеки ComplexHeatmap, plyr и stringr.

## РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

*Оценка технической, внутри- и межиндивидуальной  
вариабельности измерения пептидов  
масс-спектрометрическим MRM методом*

Для разработки MRM метода мы выбрали 246 протеотипических пептидов, для которых синтезировали и провели очистку пептидных стандартов SIS. По результатам пилотного MRM анализа были отобраны 148 пептидных стандартов SIS, картируемых на аминокислотные

последовательности, которым соответствовали 111 уникальных записей базы данных Uniprot. Поскольку данные пептиды отличались наиболее интенсивным масс-спектрометрическим сигналом, их использовали для измерения содержания нативных пептидов в экспериментальных образцах. В таблице приведены название белка, его уникальный идентификатор в базе данных Uniprot, клиническая значимость, а также выбранные пептиды. В таблице 1 дополнительных материалов (лист “Таблица MRM транзиций”) приведены транзиции, представляющие собой пару прекурсорный ион/фрагментарный ион, и нормализованную энергию соударения, используемые для MRM анализа.

В результате из 148 пептидов, картируемых на 111 FDA-верифицированных белков, 70 пептидов, картируемых на 52 белка, были зарегистрированы хотя бы в 1 из 31 экспериментальном образце, 59 пептидов, картируемых на 44 белка, были зарегистрированы во всех экспериментальных образцах ( $n=31$ ). Результаты измерений приведены в таблице 1 дополнительных материалов (лист “Все измерения”).

На рисунке 1 схематично показан осуществлённый эксперимент, в котором оценивали следующие основные расчётные параметры: (1) техническая вариабельность (TechV), представляющая собой коэффициент вариации концентрации для каждого пептида, измеренной в результате трёх технических повторов хромато-масс-спектрометрического эксперимента в режиме MRM (LC-MRM); (2) среднее значение концентрации, измеренной в результате трёх технических повторов LC-MRM эксперимента, данную величину считали измерением для каждого пептида и использовали далее для вычисления внутри- и межиндивидуальной вариабельности; (3) внутрииндивидуальная вариабельность (IntraV), представляющая собой коэффициент вариации между измерениями ( $\text{TechV} < 20\%$ ), выполненными в 10 образцах гидролизата, полученного в результате ферментативного расщепления одного объединённого образца плазмы крови; (4) межиндивидуальная (биологическая) вариабельность (Inter(bio)), представляющая собой коэффициент вариации между измерениями ( $\text{TechV} < 20\%$ ,  $\text{IntraV} < 30\%$ ), выполненными в 31 индивидуальном образце.

Поскольку ферментативное расщепление трипсином и масс-спектрометрический анализ могут вносить вклад в погрешность измерения, на первом этапе оценивали техническую воспроизводимость измерений и пробоподготовки (рис. 2 и 3). В первом случае мы оценивали техническую вариабельность (TechV), во втором — внутрииндивидуальную вариабельность (IntraV).

При проведении масс-спектрометрических измерений техническую вариабельность в пределах 20% считают вполне приемлемой [4, 7]. Рисунок 2 показывает распределение технической вариабельности и долю измерений, для которых она составила менее 20%, а также её зависимость от содержания пептида. Как следует из рисунка 2А,

для 1466 (80,1%), 1758 (96,1%) и 1827 (99,9%) измерений техническая вариабельность составила менее 5%, 10% и 20% соответственно. Таким образом, подавляющее большинство измерений удовлетворяет допустимой технической вариабельности. При этом, как показывает рисунок 2Б, начиная с микромолярных концентраций и выше, техническая вариабельность оказалась менее 10%, в то время, как в наномолярных концентрациях и ниже техническая вариабельность увеличивалась пропорционально содержанию пептида, приближаясь к пороговому значению в 20%. В нижней области концентрационного диапазона соотношение сигнала к шуму (S/N) для пептидных ионов снижается, что приводит к увеличению ошибки измерений и диктует потребность оптимизации метода в отношении чувствительности и точности. Для увеличения чувствительности может быть применено иммуноаффинное обогащение низкопредставленных белков плазмы крови или двумерная хроматография, объединяющая обогащение в щелочных условиях с последующим аналитическим разделением в кислых условиях [8, 9].

Внутрииндивидуальная вариабельность (IntraV) отражает воспроизводимость ферментативного расщепления трипсином. Для оценки данного параметра проводили MRM анализ объединённого образца плазмы крови, для которого гидролитическое расщепление было выполнено в 10 повторах. На рисунке 3 показано распределение технической (TechV), внутри-(IntraV) и межиндивидуальной (InterV (BioV)) вариабельностей для 59 пептидов, картируемых на 44 белка, зарегистрированных во всех экспериментальных образцах. Рисунок 3А демонстрирует для какой доли пептидов в каком процентном диапазоне находится вариабельность на каждом уровне (TechV, IntraV, InterV (BioV)). Как следует из рисунка 3А, минимальное значение внутрииндивидуальной вариабельности (IntraV) оказалось больше 5%. При этом для 30 (50,8%), 38 (64,4%), и 55 (93,2%) пептидов внутрииндивидуальная вариабельность (IntraV) пептидов, отражающая воспроизводимость гидролиза, составила менее 15%, 20%, и 30% соответственно (рис. 3А и 3Б). Для 4 (6,8%) пептидов внутрииндивидуальная вариабельность оказалась более 30%, поэтому они были исключены из дальнейшей оценки межиндивидуальной (биологической) вариабельности (рис. 3А и 3Б). Рисунок 3В показывает, что средняя концентрация пептидов, измеренная в 10 образцах гидролизата, полученных из одного объединённого образца, и концентрация пептидов, измеренная в 31 индивидуальном образце, коррелируют с  $R^2=0,99$ , что позволит в будущем использовать объединённый образец для стандартизации метода, в том числе определения порога детекции и нижней границы определяемых концентраций, линейного диапазона измерений и т.д.

Эффективность гидролитического расщепления во многом зависит от доступности аминокислотных остатков, к которым трипсин проявляет специфичность, а именно, аргинина (R) и лизина (K).

## MRM АНАЛИЗ FDA-ВЕРИФИЦИРОВАННЫХ БЕЛКОВ

Таблица. Белки, одобренные FDA для клинического применения, их клиническая значимость и протеотипические пептиды измеренные методом MRM. Uniprot ID — уникальный идентификатор базы данных Uniprot

Uniprot ID	Название белка	Клиническая значимость	Аминокислотная последовательность пептида
P24298	Аланинаминотрансфераза (АЛТ)	Оценка функции печени	ALELEQELR
P24298	Аланинаминотрансфераза (АЛТ)	Оценка функции печени	LLVAGEGHTR
P04075	Альдолаза А	Диагностика миодистрофий	GILAADESTGSIK
P05062	Альдолаза В	Диагностика миодистрофий	LDQGGAPLAGTNK
P09972	Альдолаза С	Диагностика миодистрофий	AEVNLAAQGK
			ALQASALNAWR
P08697	Альфа 2-антиплазмин	Комплексная оценка состояния плазминовой системы	LGNQEPGGQTALK
			FDPSLTQR
P01009	Альфа-1-антитрипсин	Диагностика причин эмфиземы, выявление причин нарушений функции печени	LSITGTYYDLK
			LQHLENELTHDIITK
P02763	Альфа-1-кислый гликопротеин (AGP, орозомукоид)	Диагностика коллагенозов, злокачественных опухолей, гемолиза, воспалительных процессов	YVGGQEHFAHLLILR
P01023	Альфа-2-макроглобулин	Диагностика фиброза печени, диагностика нефротического синдрома	NEDSLVFVQTDK
			AIGYLNTGYQR
P02771	Альфа-фетопротеин (альфа-ФП)	Пренатальная диагностика патологий развития плода, диагностика низкодифференцированных опухолей	DFNQFSSGEK
P04746	Амилаза панкреатическая (Р-тип)	Диагностика панкреатита	LTGLLDLAEK
P04745	Амилаза слюнная (S-тип)	Диагностика панкреатита	LSGLLDLALGK
P01019	Ангиотензиноген	Диагностика артериальной гипертензии	ALQDQLVLVAAK
P12821	Ангиотензин-превращающий фермент сыворотки	Диагностика саркоидоза	FDVNQLQNTTIK
			VTNDTESDINYLLK
P01008	Антитромбин 3	Диагностика нарушения свертываемости крови	ANRPFLVFIR
			DDLTVSDAFHK
P04114	Аполипопротеин В	Оценка риска развития атеросклероза	VPSYTLILPSLELPVLHVPR
			TGISPLALIK
P02647	Аполипопротеин А1	Оценка риска развития сердечно-сосудистых заболеваний, расчёт коэффициента аполипопротеин В / аполипопротеин А1	ATEHLSTLSEK
			LLDNWDSVTSTFSK
P02656	Аполипопротеин С3	Оценка риска развития атеросклероза	GWVTDGFSSLK
P17174	Аспартатаминотрансфераза (АСТ)	Оценка функции печени	VNLGVGAYR
P08833	Белок 1, связывающий инсулиноподобный фактор роста (IGFBP-1)	Оценка развития плода	IPGSPEIR
P17936	Белок 3, связывающий инсулиноподобный фактор роста (IGFBP-3)	Диагностика нарушений роста	AGASSAGLGPVVR
P61769	Бета-2-микроглобулин	Диагностика миеломы, лимфомы и воспалительных процессов. Оценка функции почек.	VEHSDLFSK
P02775	Бета-тромбоглобулин	Оценка активации тромбоцитов	TTSGIHPK
			NIQSLEVIGK
P43251	Биотинидаза	Диагностика недостаточности биотинидазы	VDLITFDTPFAGR
			LSSGLVTAALYGR
P36269	Гамма-глутамилтранспептидаза (гамма-ГТ)	Оценка функции печени	DLLGETLAQLIR

Таблица (продолжение). Белки, одобренные FDA для клинического применения, их клиническая значимость и протеотипические пептиды измеренные методом MRM. Uniprot ID — уникальный идентификатор базы данных Uniprot

Uniprot ID	Название белка	Клиническая значимость	Аминокислотная последовательность пептида
P00738	Гаптоглобин	Диагностика внутрисосудистого гемолиза, оценка функций печени, выявление реакции острой фазы: инфекции, воспаления, опухоли, ожога, обморожения, аутоиммунных заболеваний.	TEGDGVYTLNNEK
P02790	Гемопексин	Оценка метаболизма железа и функции печени	SGAQATWTELPWPHEK
P04278	Глобулин, связывающий половые гормоны (ГСПГ)	Оценка статуса половых гормонов	LDVDQALNR IALGGLLPASNLR
P06744	Глюкозо-6-фосфат-изомераза	Диагностика различных типов рака	TFTTQETITNAETAK
P48735	Изоцитратдегидрогеназа, митохондриальный тип IDH2	Диагностика и оценка ответа на лечение при глиомах	YFDLGLPNR
O75874	Изоцитратдегидрогеназа, цитоплазматический тип IDH1	Диагностика и оценка ответа на лечение при глиомах	TVEAEAAHGTVTR GQETSTNPIASIFAWTR
P01880	Иммуноглобулин D	Оценка гуморального иммунитета, диагностика иммунодефицитных состояний	DAHLTWEVAGK VPTGGVEEGLLER
P01854	Иммуноглобулин E	Диагностика аллергических заболеваний	NGTLTVTSTLPVGTR
P01876	Иммуноглобулин IgA-1	Оценка гуморального иммунитета, диагностика иммунодефицитных состояний	TPLTATLSK
P01877	Иммуноглобулин IgA-2	Оценка гуморального иммунитета, диагностика иммунодефицитных состояний	DASGATFTWTPSSGK
P01857	Иммуноглобулин IgG-1	Оценка гуморального иммунитета, диагностика иммунодефицитных состояний	GPSVFPLAPSSK GPSVFPLAPCSR
P01871	Иммуноглобулин IgM	Оценка гуморального иммунитета, диагностика иммунодефицитных состояний	VSVFVPPR YAATSQVLLPSK
P05111	Ингибин A	Оценка функции яичников, диагностика опухолей яичников	GLAQEAEEGLFR ALFLDALGPPAVTR
P08476	Ингибин B	Оценка функции яичников. Оценка функции яичек и сперматогенеза	VGNGYVEIEDDIGR
P05121	Ингибитор активатора плазминогена (SERPINE1)	Выявление генетической предрасположенности к инфаркту миокарда, ишемической болезни сердца, преэклампсии, атеросклерозу, инсулинорезистентности и ожирению	VFQQAQASK DEISTTDAIFVQR
P05019	Инсулиноподобный фактор роста	Исследование функции гипофиза	GFYFNKPTGYGSSSR
P01834	Каппа цепь иммуноглобулинов, константный участок	Диагностика моноклональных гаммапатий	DSTYSLSSTLTLSK
P24666	Кислая фосфатаза	Выявление карциномы простаты	LVTQDNISENWR SPIAEAVFR
P12277	Креатинкиназа MB, B тип	Диагностика острого и рецидивирующего инфаркта	VLTPELYAELR LAVEALSSLDGDLAGR
P06732	Креатинкиназа MB, M тип	Диагностика острого и рецидивирующего инфаркта	TDLNHNENLK
P07195	Лактатдегидрогеназа цепь H	Диагностика инфаркта миокарда и других видов повреждения сердечной мышцы	SLADELALVDVLEDK IVVVTAGVR

## MRM АНАЛИЗ FDA-ВЕРИФИЦИРОВАННЫХ БЕЛКОВ

Таблица (продолжение). Белки, одобренные FDA для клинического применения, их клиническая значимость и протеотипические пептиды измеренные методом MRM. Uniprot ID — уникальный идентификатор базы данных Uniprot

Uniprot ID	Название белка	Клиническая значимость	Аминокислотная последовательность пептида
P00338	Лактатдегидрогеназа цепь M	Диагностика инфаркта миокарда и других видов повреждения сердечной мышцы	VTLTSEEEAR
P02788	Лактоферрин	Диагностика неспецифических воспалительных заболеваний и бактериальных кишечных инфекций	DGAGDVAFIR GGSFQLNELQGLK
P61626	Лизоцим	Оценка функции почек и иммунной системы	ATNYNAGDR
P16233	Липаза панкреатическая	Диагностика панкреатита	IVETNVGK FYLDTGDAENFAR
P08519	Липопротеин (а)	Оценка риска развития болезней сердца и сосудов, в том числе ишемической болезни сердца (ИБС), цереброваскулярной болезни сердца (ЦВБ) и заболеваний периферических сосудов	GTYSTTVTGR
Q13093	Липопротеин-ассоциированная фосфолипаза A2	Диагностика метаболического синдрома	LDTLWIPNK
P06858	Липопротеинлипаза	Оценка обмена холестерина	GLGDVDQLVK
P0CF74	Лямбда цепь иммуноглобулинов 6, константный участок	Диагностика моноклональных гаммапатий	ADGSPVNTGVETTTSPK
P05164	Миелопероксидаза	Оценка функции иммунной системы	IANVFTNAFR
P02144	Миоглобин	Диагностика повреждения сердечной мышцы, оценка функции почек при травме	HGATVLTALGGILK
P15941	Муцин 1 (CA 15-3)	Диагностика рака молочной железы	YVPPSSDTR
Q8WX17	Муцин 16 (CA-125)	Диагностика рака яичника	NGTQLQNFTLDR NIEDALNQLFR
P14618	Опухолевый маркер 2 (ТМ 2) — пируваткиназа	Диагностика рака толстой кишки	GSGTAEVELK
P01270	Паратиреоидный гормон (ПТГ)	Оценка функции паращитовидных желёз, оценка обмена кальция и фосфора	LQDVHNFVALGAPLAPR
P00747	Плазминоген	Оценка функционального состояния фибринолитической системы плазмы крови	HSIFTPETNPR
P02766	Преальбумин (транстретин)	Оценка функции печени, оценка и коррекция состояния питания у больных с неотложными состояниями	AADDTWEPFASGK
P01236	Пролактин	Исследование функции гипофиза, диагностика бесплодия и нарушения половой функции	AVEIEEQTK SWNEPLYHLVTEVR
P07288	Простатспецифический антиген	Диагностика аденомы и воспалительных заболеваний предстательной железы. Диагностика рака предстательной железы.	LSEPAELTDAVK
P04070	Протеин С	Диагностика тромбозов и тромботических осложнений	GDSPWQVVLLDSK
P07225	Протеин S	Оценка генетической недостаточности протеина S и связанного с ней риска развития тромбозов	IETISHEDLQR VYFAGFPR
P06731	Раковый эмбриональный антиген (РЭА)	Диагностика и мониторинг лечения злокачественной опухоли	TLTLFNVTR

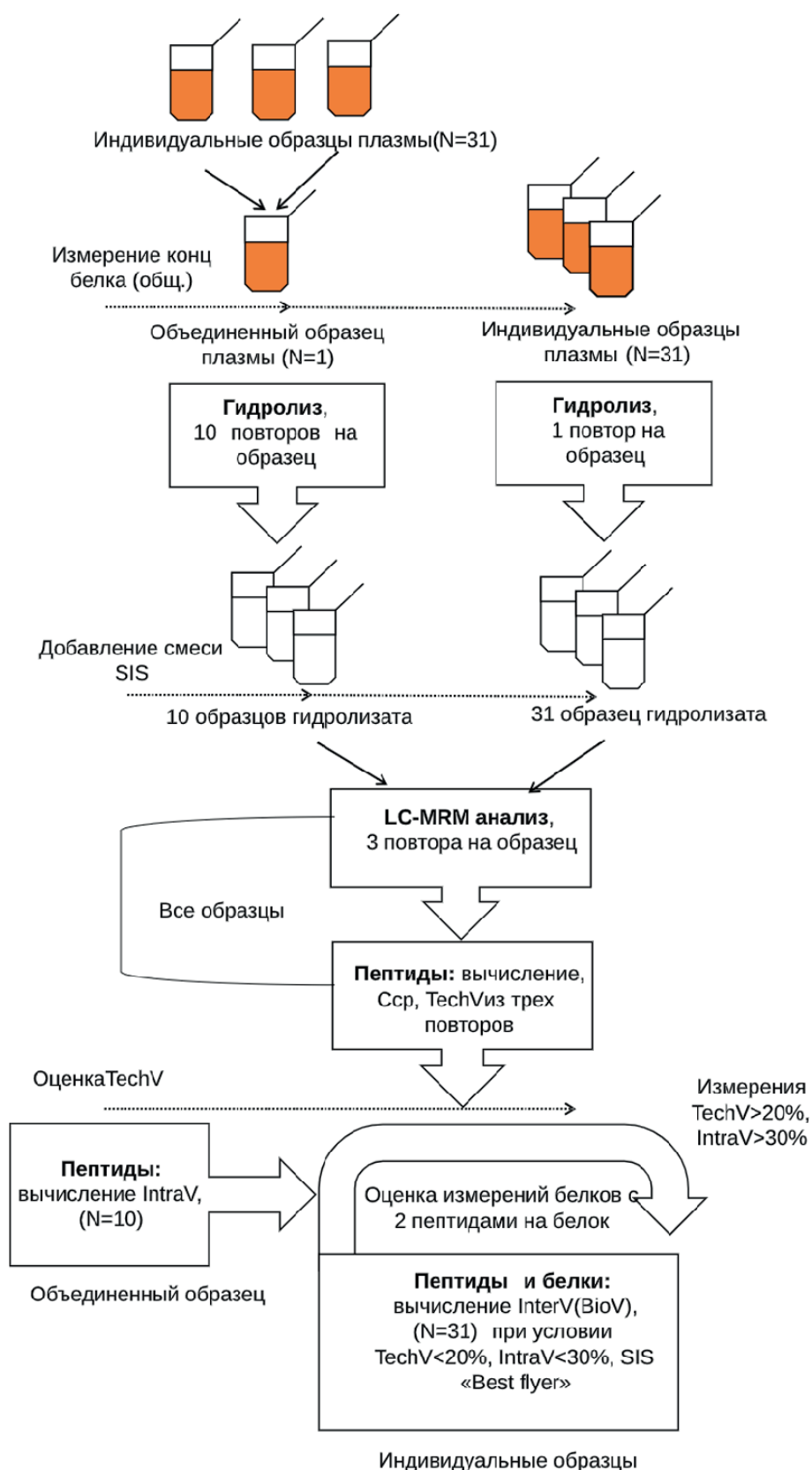
Таблица (продолжение). Белки, одобренные FDA для клинического применения, их клиническая значимость и протеотипические пептиды измеренные методом MRM. Uniprot ID — уникальный идентификатор базы данных Uniprot

Uniprot ID	Название белка	Клиническая значимость	Аминокислотная последовательность пептида
P00797	Ренин	Выявление причин артериальной гипертензии, скрининг первичного гиперальдостеронизма — синдрома Конна	VVFDTGSSNVWVPSSK
P09455	Ретинол-связывающий белок 1	Оценка обмена витамина А	IANLLKPKDK
P02753	Ретинол-связывающий белок 4 (плазменный ретинол-связывающий белок)	Диагностика инсулинорезистентности, оценка уровня витамина А	FSGTWYAMAK
			YWGVASFLQK
P04626	Рецептор HER2/neu	Определения наличия HER-2 (HER2/neu) – рецепторов при раке молочной железы и желудка	NNQLALTLDITNR
P02786	Рецептор к трансферину	Оценка содержания железа и активности эритропоэза	VSASPLLYTLIEK
P14784	Рецептор человеческого интерлейкина 2, субъединица бета	Диагностика гематологических злокачественных новообразований	LTTVDIVTLR
P31785	Рецептор человеческого интерлейкина 2, субъединица гамма	Диагностика гематологических злокачественных новообразований	LSESQLELNWNNR
P05155	C1-ингибитор	Оценка активности системы комплемента, диагностика аутоиммунных заболеваний	LLDSLPSDTR
P02745	C1q компонент комплемента субъединица А	Оценка классического пути запуска системы комплемента, диагностика аутоиммунных заболеваний	VGYPGPGSGLGAR
P02746	C1q компонент комплемента субъединица В	Оценка классического пути запуска системы комплемента, диагностика аутоиммунных заболеваний	LEQGENVFLQATDK
P02747	C1q компонент комплемента субъединица С	Оценка классического пути запуска системы комплемента, диагностика аутоиммунных заболеваний	TNQVNSGGVLLR
			FQSVFTVTR
P09871	C1s компонент комплемента	Оценка классического пути запуска системы комплемента, диагностика аутоиммунных заболеваний	SNALDIIFQTDLTGQK
			TNFDNDIALVR
P01024	C3 компонент комплемента	Оценка классического и альтернативного пути запуска системы комплемента, диагностика аутоиммунных заболеваний	LVAYYTLIGASGQR
P01031	C5 компонент комплемента	Оценка классического и альтернативного пути запуска системы комплемента, диагностика аутоиммунных заболеваний	GIYGTISR
			LQGTLPVEAR
P01241	Соматотропный гормон	Исследование функции гипофиза	LHQLAFDITYQEFEEAYIPK
Q00796	Сорбитолдегидрогеназа	Оценка функции печени	ALEAFETFK
P01266	Тиреоглобулин	Контроль лечения и выявление рецидива папиллярной и фолликулярной аденокарциномы щитовидной железы	LGVNVTWK
			GGADVASIHLTLAR
P05543	Тироксин-связывающий глобулин (TBG)	Оценка уровня тиреоид-связывающих белков в крови	NALALFVLPK
P00750	Тканевой активатор плазминогена	Диагностика тромбоза глубоких вен, инфаркта миокарда, ишемической болезни сердца, инсульта	HEALSPFYSER
P02787	Трансферин	Оценка метаболизма железа, диагностика анемий, оценка функции печени	SASDLTWDNLK

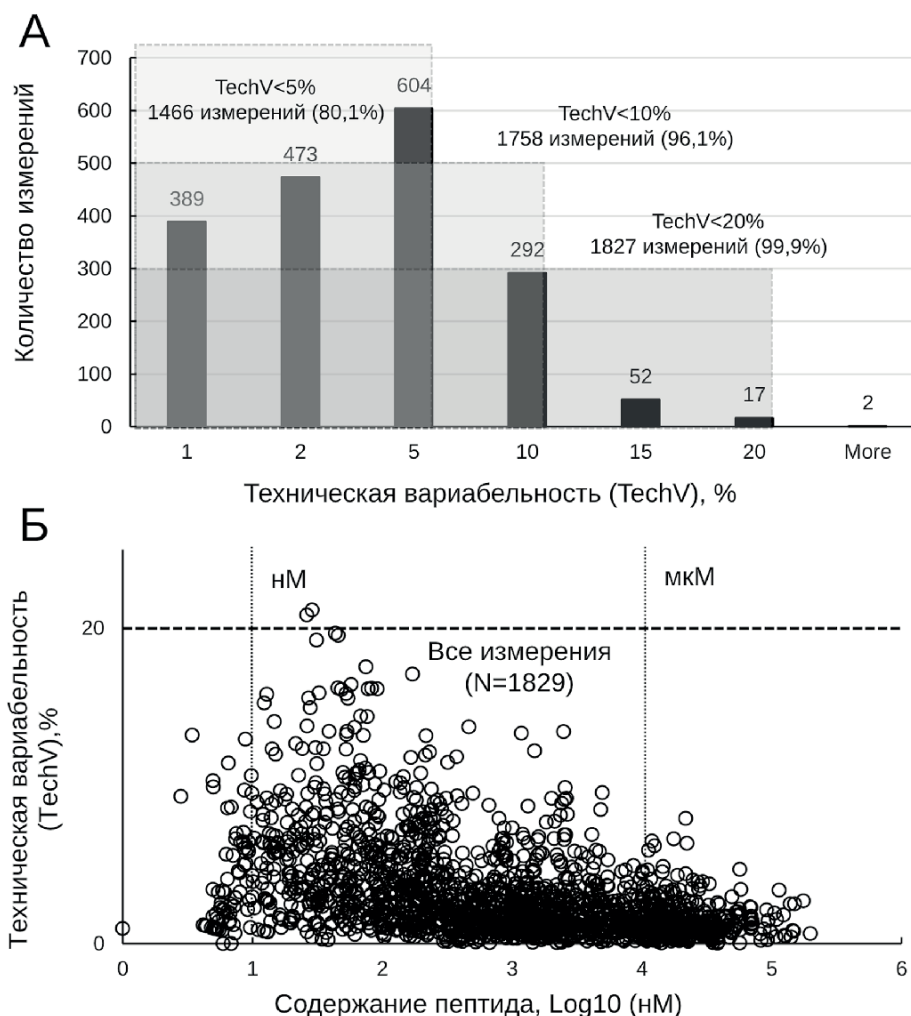
# MRM АНАЛИЗ FDA-ВЕРИФИЦИРОВАННЫХ БЕЛКОВ

Таблица (продолжение). Белки, одобренные FDA для клинического применения, их клиническая значимость и протеотипические пептиды измеренные методом MRM. Uniprot ID — уникальный идентификатор базы данных Uniprot

Uniprot ID	Название белка	Клиническая значимость	Аминокислотная последовательность пептида
P07477	Трипсин	Диагностика панкреатита	VSTISLPTAPPATGTK
P19429	Тропонин I	Диагностика инфаркта миокарда	AYATEPHAK
P45378	Тропонин T, быстрая скелетная мускулатура	Диагностика инфаркта миокарда	VDFDDIQK
P13805	Тропонин T, медленная скелетная мускулатура	Диагностика инфаркта миокарда	AQELSDWIHQLESEK
P45379	Тропонин T, сердечная мышца	Диагностика инфаркта миокарда	DLNELQALIEAHFENR
P00749	Урокиназный активатор плазминогена	Оценка фибринолитической ферментной системы. Диагностика опухоли мочеполовой системы или желудочно-кишечного тракта	DYSADTLAHHNDIALLK
P00751	Фактор В комплемента	Оценка альтернативной системы комплемента, диагностика заболеваний иммунной системы	STGSWSTLK VSEADSSNADWVTK
P04275	Фактор Виллебранда	Диагностика коагулопатии	IGEADFNR YLSHDSFLVSQGDR
P00488	Фактор свёртываемости крови 13, субъединица A1	Выявление генетической предрасположенности к инфаркту миокарда, тромбозам.	LSIQSSPK
P05160	Фактор свёртываемости крови 13, субъединица B	Оценка свёртываемости крови	LIENGYFHPVK
P00740	Фактор свёртывания крови IX (фактор Кристмаса)	Диагностика гемофилии B, оценка функций печени	SALVLQYLR
P02792	Ферритин	Оценка запасов железа в организме	LNQALLDLHALGSAR
P02765	Фетуин A (AHSG)	Оценка обмена кальция и остеогенеза	FSVVYAK
P02675	Фибриноген, цепь b	Оценка способности организма к тромбообразованию	DNENVVNEYSSELEK
P02679	Фибриноген, цепь g	Оценка способности организма к тромбообразованию	YEASILTHDSSIR
P02671	Фибриноген, цепь a	Оценка способности организма к тромбообразованию	GSESGIFTNTK GLIDEVNQDFTNR
P02751	Фибронектин	Оценка эффективности проводимой терапии при многих инфекционных и соматических состояниях	VPGTSTSATLTGLTR
P15309	Фосфатаза кислая простатическая	Выявление карциномы простаты	DFIATLGK SPIDTFPTDPIK
P09923	Фосфатаза щелочная, изофермент кишечного типа	Выявление заболеваний печени или костей, диагностика закупорки желчевыводящих путей	ANFQTIGLSAAAR GFYLFVEGGR
P05186	Фосфатаза щелочная, ткань-неспецифичный изофермент	Выявление заболеваний печени или костей, диагностика закупорки желчевыводящих путей	LDGLDLVDTWK ANEGTVGVSAATER
P06276	Холинэстераза	Диагностика отравления пестицидами, оценка функционального статуса холинэстеразы	YLTNTESTR IFFPGVSEFGK
P00450	Церулоплазмин	Обмен меди, диагностика заболевания Вильсона–Коновалова	GAYPLSIEPIGVR
P01034	Цистатин C	Определение исходного функционального состояния почек	ALDFAVGEYNK
P16860	NT-proBNP — мозговой натрийуретический гормон	Диагностика сердечной недостаточности	SHPLGSPGSASDLETSGLQEQR



**Рисунок 1.** Схема эксперимента. Измерение общего белка (Конц. белка общ.) для всех образцов выполнялась с помощью коммерческого набора Pierce™ BCA Protein Assay Kit (“Pierce”). SIS — изотопно-меченные пептидные стандарты (Stable Isotope-labeled Standard). SIS “Best Flyer” — изотопно-меченый пептидный стандарт с наилучшей ионизационной способностью. Сср — средняя концентрация белка в нМ, TechV — техническая вариабельность; IntraV — внутрииндивидуальная вариабельность; Inter (BioV) — межиндивидуальная (биологическая) вариабельность.

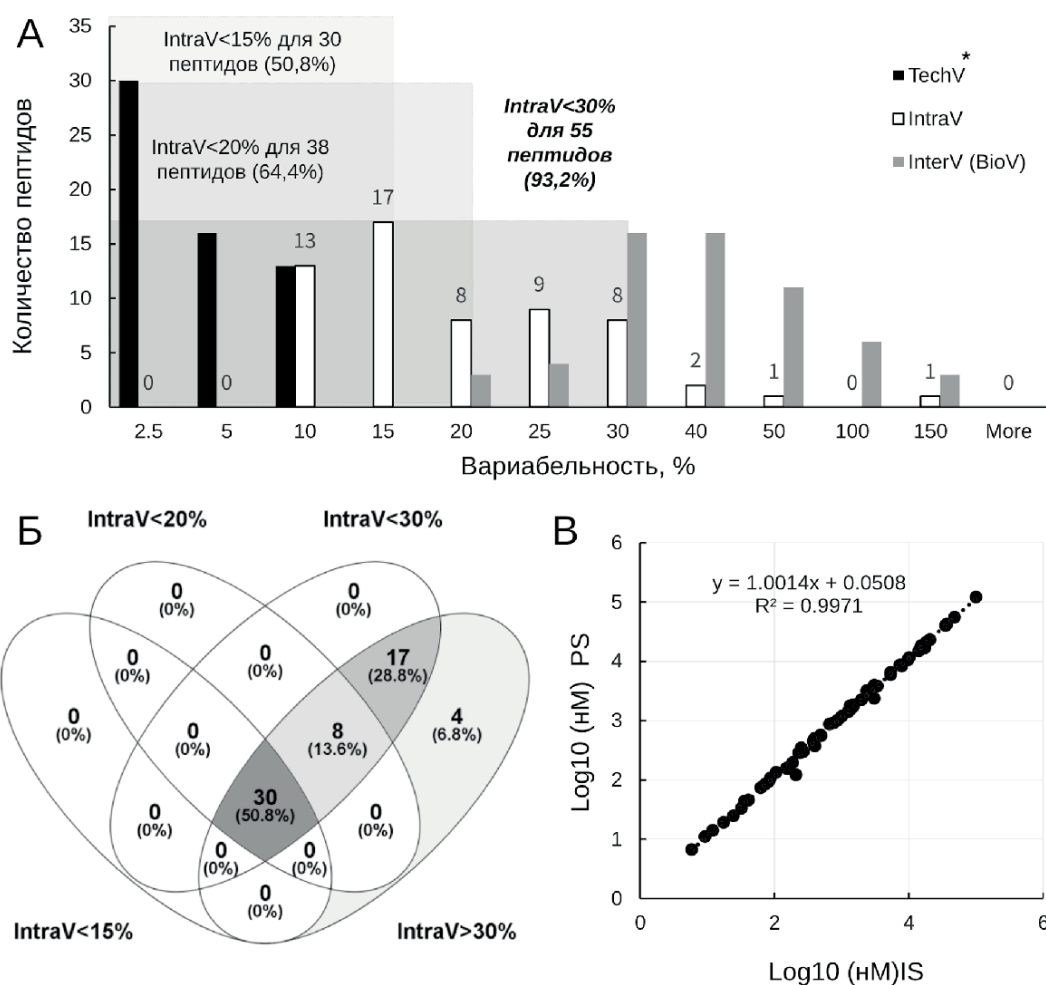


**Рисунок 2.** Все выполненные измерения ( $n=1829$ ) для 59 пептидов, картируемых на 44 белка, в 31 образце. Техническая вариабельность (TechV) представляет собой коэффициент вариации концентрации для каждого пептида, измеренной в результате 3 технических повторов LC-MRM эксперимента. **А.** Распределение технической вариабельности (TechV). **Б.** Зависимость технической вариабельности (TechV) от содержания пептида. Результаты представлены в виде десятичного логарифма концентрации белка, измеренной в цельной плазме крови в нМ (Log10 нМ).

Стабильные посттрансляционные модификации (ПТМ) этих остатков, включающие ацетилирование, убиквитинилирование и сумоилирование, а также гликозилирование, характерное для белков плазмы крови [10], снижают эффективность действия трипсина. Так, в случае ацетилирования нормальной работе фермента мешают стерические эффекты, а также потеря положительного заряда на  $\epsilon$ -аминогруппе, в результате чего остаток аспартата в каталитическом центре трипсина не может стабилизировать данную аминокислотную последовательность белка, поскольку это может отражаться на эффективности расщепления белка трипсином и влиять, в конечном счёте, на детекцию его пептидов. Такая задача требует междисциплинарного подхода с привлечением биоинформатических методов. С другой стороны, наличие ПТМ и уровень модификации белкового продукта, часто связывают

с возникновением и развитием болезни. Например, N-гликозилирование белков плазмы крови, таких как альфа-1-кислый гликопротеин, альфа-1-антитрипсин, альфа-2-макроглобулин, аполипопротеин В и др., связывают с различными заболеваниями [10].

Рисунок 4 демонстрирует на каком уровне была техническая (TechV\*), внутри- (IntraV) и межиндивидуальная (биологическая) (InterV (BioV)) вариабельность для индивидуальных пептидов. Как видно из рисунка 4, для каждого отдельно взятого белка соотношение внутри- и межиндивидуальной вариабельности оказывается уникальным. В ряде случаев, для пептидов таких белков, как фактор фон Виллебранда (VWF) (YLSHDSFLVSQGDR), антитромбин III (ANT3) (ANRPFLVFIR), аполипопротеин В (APOB) (VPSYTLILPSLELPVLHVPR), ингибитор протеаз C1 (IC1) (LLDSLPSDTR), субъединица С компонента комплемента C1q (C1QC) (FQSVFTVTR), гемоплексин (HEMO) (SGAQATWTELPWPHEK) внутрииндивидуальная вариабельность очень близка

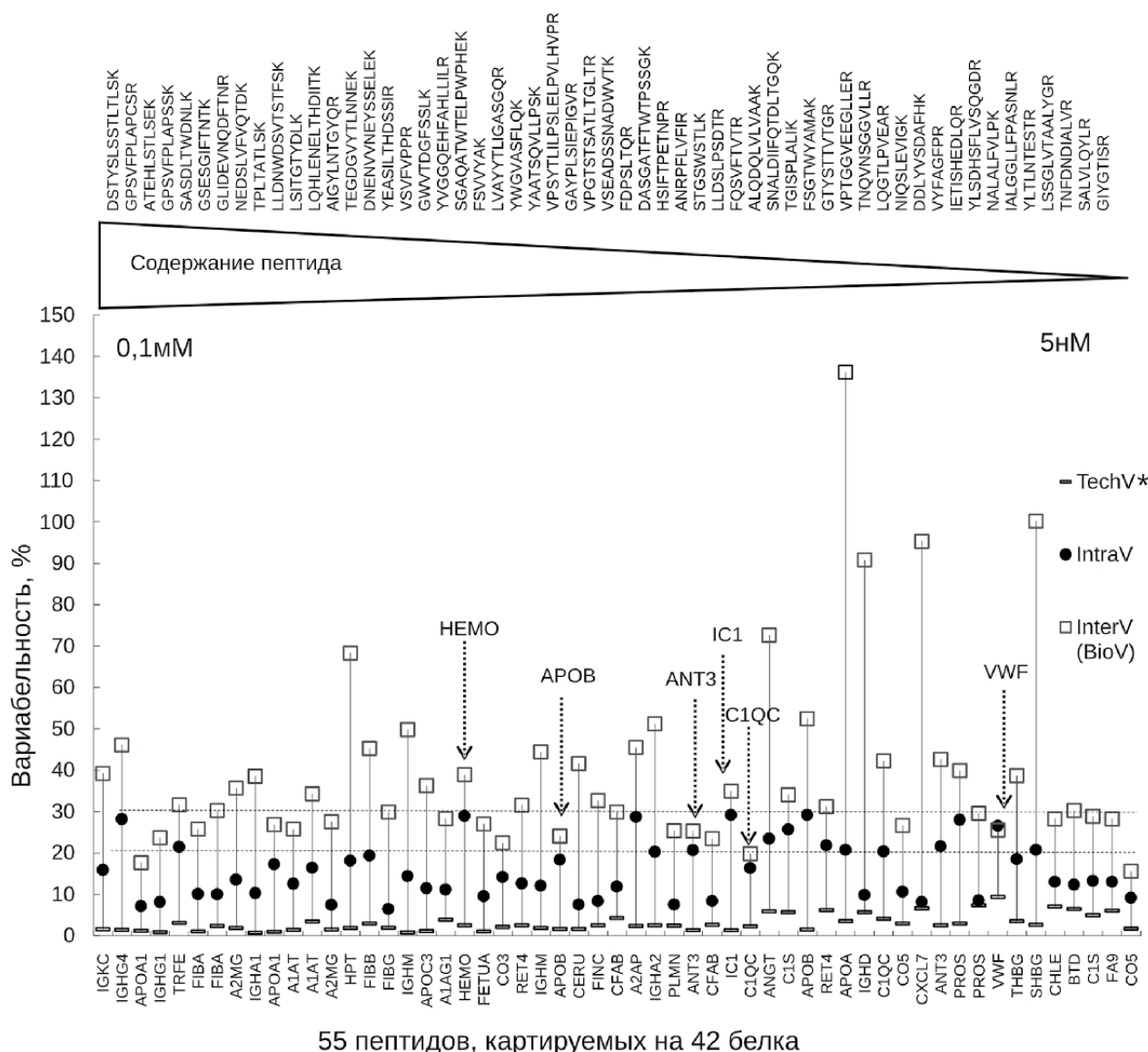


**Рисунок 3.** А. Распределение технической (TechV), внутри- (IntraV) и межиндивидуальной (InterV (BioV)) вариальности для 59 пептидов, картируемых на 44 белка, зарегистрированных во всех экспериментальных образцах (n=31). TechV\* — средняя техническая вариальность; IntraV — внутрииндивидуальная вариальность; Inter (BioV) — межиндивидуальная (биологическая) вариальность. Б. Диаграмма Венна, демонстрирующая пересечение пептидов, для которых IntraV оказалась менее 15%, 20% и 30%, а также более 30%. В. Корреляция среднего содержания пептидов, измеренных в индивидуальных образцах и объединенном образце. Log10 (нМ) PS — десятичный логарифм средней концентрации белка, измеренной в объединенном образце плазмы крови в нМ (Pooled Sample) (n=10); Log10 (нМ) IS — десятичный логарифм средней концентрации белка, измеренной в индивидуальных образцах плазмы крови в нМ (Individual Sample) (n=31).

к межиндивидуальной (биологической). Эффективность гидролиза сильно зависит от аминокислотной последовательности и этот факт необходимо учитывать при отборе кандидатных протеотипических пептидов для мультиплексного количественного анализа.

Согласно руководству по анализу и интерпретации масс-спектрометрических данных HPP (Human Proteome Project), достоверность идентификации белка повышается, если на него картируется два и более уникальных пептидов [11]. Для 13 FDA-верифицированных белков MRM анализ осуществлялся по двум уникальным протеотипическим пептидам. Для таких белков вычислили отношение концентрации пептида I (ПептI) и пептида II (ПептII), измеренной в плазме крови в нМ (рис. 5А), а также содержание каждого пептида от общего количества ПептI и ПептII в процентах (рис 5Б). Как видно из рисунка 5, для белков PROS, FIBA и A1AT содержание ПептI и ПептII было приблизительно

эквимолярным, в то время как для 5 белков содержание ПептI и ПептII отличалось от 2 до 2,5 раз, и ещё для 5 белков из 13 содержание ПептI и ПептII отличалось более чем в 5 раз. В то же время для большинства белков, кроме антитромбина (ANT3), аполипопротеина В (APOB) и компонента комплемента C1q (C1QC) отношение содержания двух протеотипических пептидов сохранялось стабильным между индивидуумами (относительное стандартное отклонение <30%). Эти результаты могут указывать на существование стабильных и повторяющихся у индивидуумов по локализации ПТМ, влияющих на эффективность гидролитического расщепления трипсином. Для дальнейшей оценки межиндивидуальной вариальности белков использовались протеотипические пептиды, в большем количестве образующиеся в результате ферментативного расщепления, и для которых измеренная в плазме крови концентрация была выше.



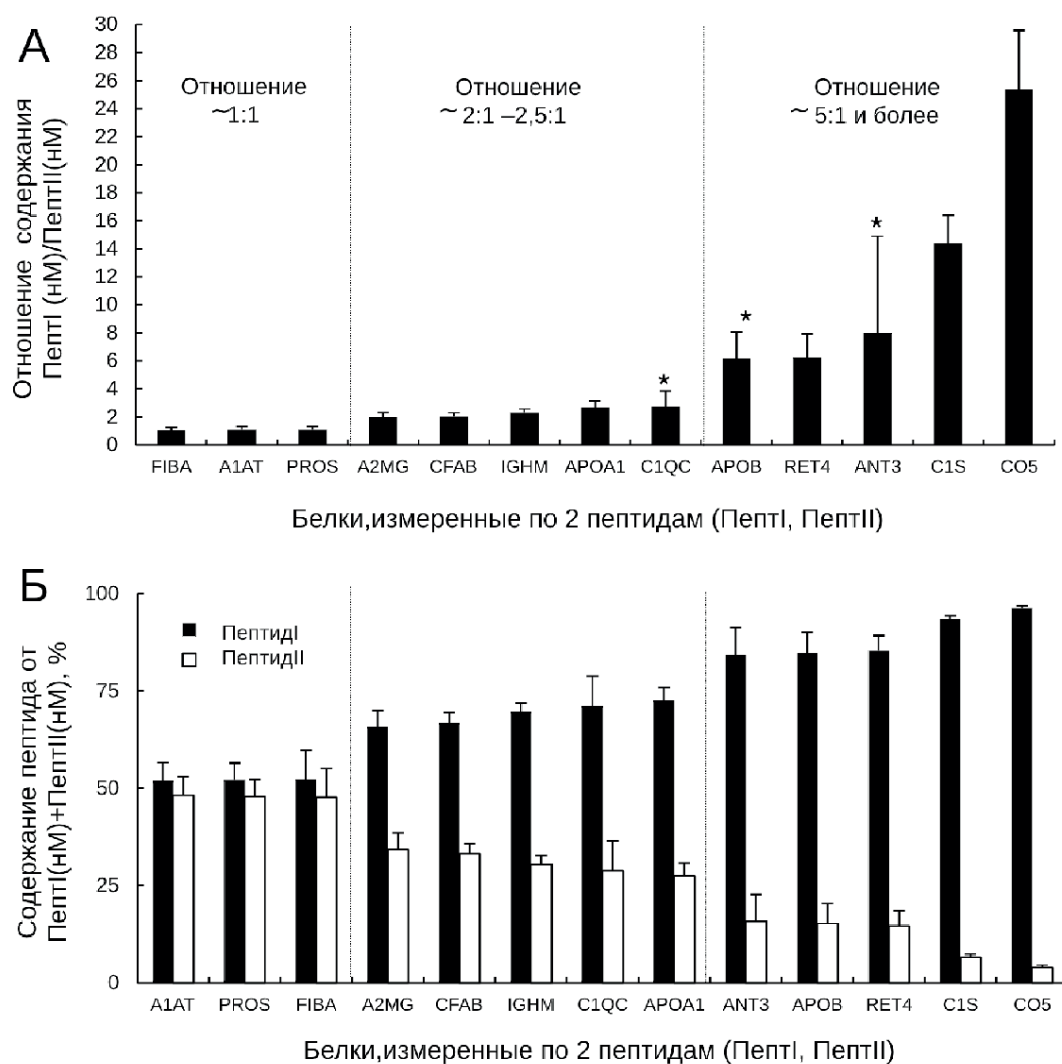
**Рисунок 4.** Индивидуальная техническая (TechV\*), внутри- (IntraV) и межиндивидуальная (биологическая) (InterV (BioV)) вариабельность в процентах для 55 пептидов, картируемых на 42 белка, зарегистрированных во всех экспериментальных образцах (N=31), TechV<20%, IntraV<30%. TechV\* — средняя техническая вариабельность; IntraV — внутрииндивидуальная вариабельность; InterV (BioV) — межиндивидуальная (биологическая) вариабельность. Пептиды показаны в порядке уменьшения содержания. Пунктирными линиями показано отсечение вариабельности в 20% и 30%.

*Протеомное ядро плазмы крови человека:  
оценка межиндивидуальной вариабельности  
FDA-верифицированных белков  
у здоровых добровольцев*

Во всех экспериментальных образцах (n=31) было зарегистрировано 42 белка по протеотипическим пептидам, для которых TechV<20%, IntraV<30%. Данный набор белков можно рассматривать как протеомное ядро плазмы крови здорового человека.

Диаграмма на рисунке 6 отражает разброс содержания 42 белков протеомного ядра плазмы крови, выраженного в виде десятичного логарифма концентрации белка, измеренной в цельной плазме крови в нМ (Log10 (нМ)). Как показывает рисунок 6,

диапазон молярной концентрации охватывал примерно 5 порядков. Минимальная и максимальная концентрации были определены для пептидов, соответствующих фактору свёртывания крови IX (FA9) ( $12 \pm 3,4$  нМ) и константному участку лёгких цепей иммуноглобулинов каппа (IGKC) ( $0,1 \pm 0,04$  мкМ). На графике также виден разброс значений для отдельных белков в выборке. Для ряда белков, а именно цепей фибриногена а, b, g (FIBA, FIBB, FIBG), и для гаптоглобина (HPT), результаты измерений в отдельных пробах сильно отличались от остальной выборки. В случае цепей фибриногена а, b, g их содержание было снижено в одном и том же образце, что может быть результатом замены плазмы на сыворотку или запуска каскада свёртывания

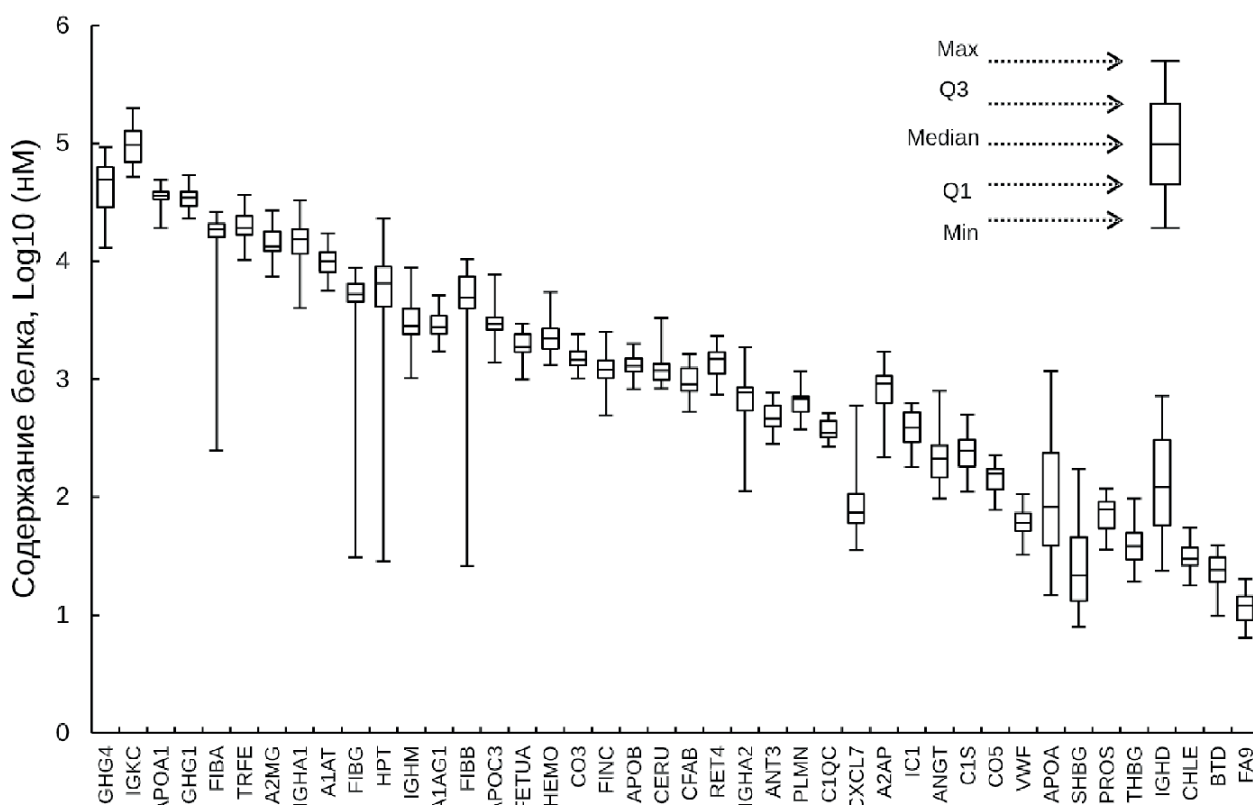


**Рисунок 5.** Содержание пептидов, картируемых на один белок. Показаны 26 пептидов, картируемых на 13 белков; на каждый белок картируется по 2 пептида — пептид I (ПептI) и пептид II (ПептII). А. Отношение концентрации пептидов ПептI и ПептII, измеренной в цельной плазме крови в нМ. Б. Содержание каждого пептида от общего количества ПептI и ПептII (ПептI(нМ)+ПептII(нМ)) в процентах. Показаны средние значения, вычисленные для 31 образца плазмы крови.\* — относительное стандартное отклонение >30%.

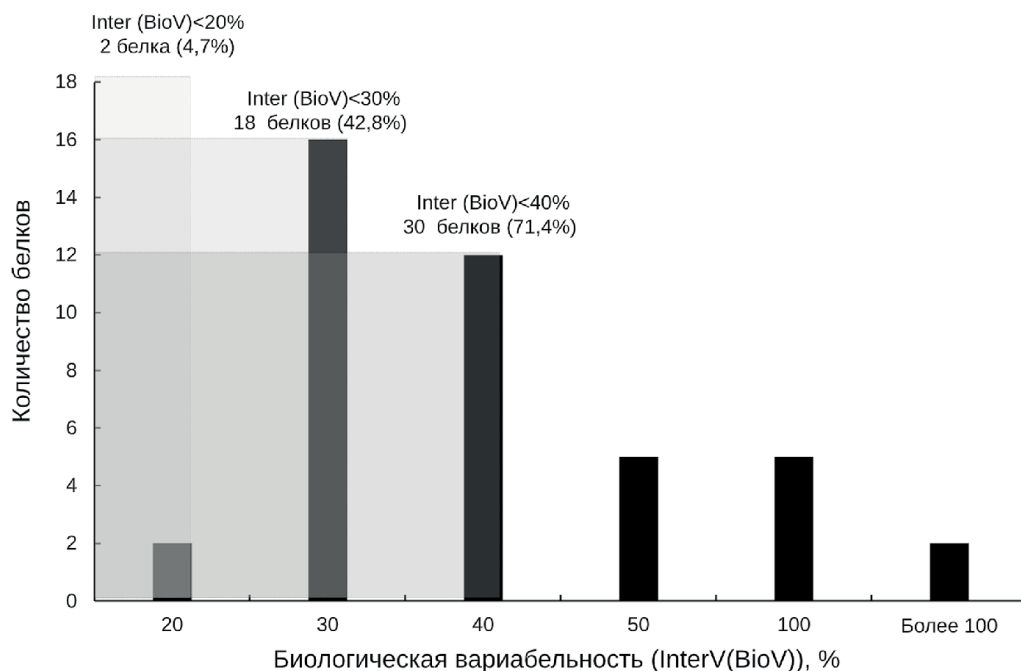
в результате нарушения процедуры пробоподготовки плазмы из цельной крови. В случае гаптоглобина причиной снижения содержания белка в одном образце может быть аминокислотная замена на данном участке белка или его сильная модификация. Используемый для анализа протеотипический пептид гаптоглобина (TEGDGVYTLNNEK) картируется на область белка с 119 по 131 аминокислотный остаток. По литературным данным, в положении 129 может происходить замена аспарагина на аспарагиновую кислоту ( $N \rightarrow D$ ,  $\Delta M_r = 0,985$  Да), а в положении 129 может происходить замена глутаминовой кислоты на лизин ( $E \rightarrow K$ ,  $\Delta M_r = 0,942$  Да) [12, 13]. Во втором случае аминокислотная замена приводит к появлению дополнительного сайта для расщепления трипсина, что может приводить к образованию укороченного пептида с последовательностью TEGDGVYTLNNEK. Более того, протеотипический пептид гаптоглобина (TEGDGVYTLNNEK) картируется на альфа-цепь, образующуюся при созревании белка в результате

гидролитического расщепления. Именно альфа-цепь может существовать в двух аллельных формах: альфа 1 (83 аминокислотных остатка, укороченная форма) и альфа 2 (142 аминокислотных остатка, полноразмерная форма). Таким образом, одной из причин сниженного содержания белка может быть аллельный полиморфизм, который ассоциируют с развитием различных заболеваний различных органов, в том числе сердечно-сосудистой системы [14, 15]. Метод MRM позволяет отслеживать и принимать во внимание как артефакты пробоподготовки, так и структурные различия аминокислотной последовательности, обусловленные точечными мутациями или посттрансляционными модификациями.

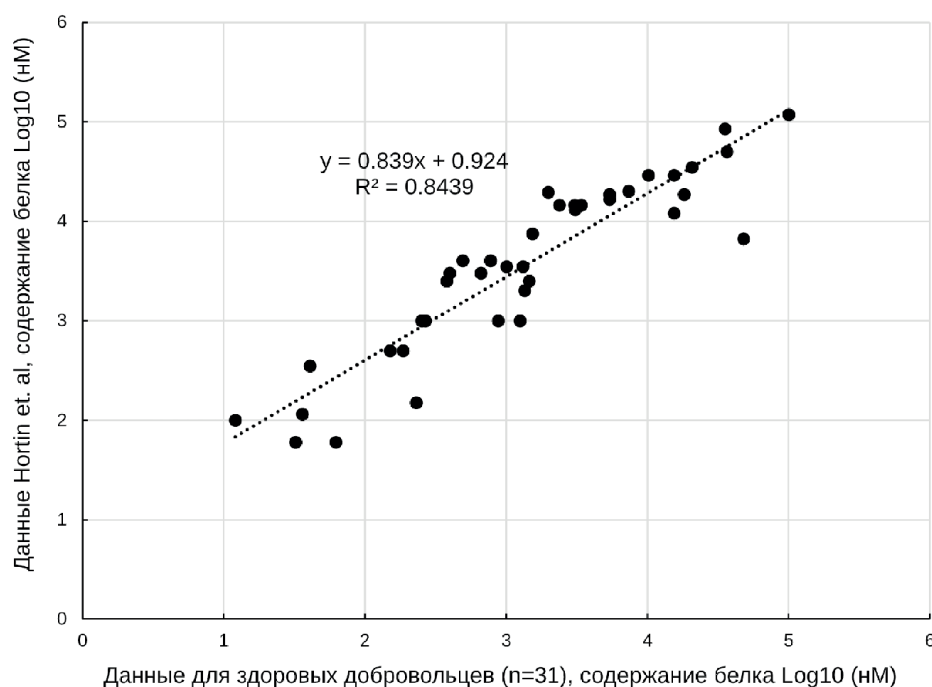
На рисунке 7 показано распределение биологической варируемости для 42 диагностически значимых белков: для 30 (71,4%) белков, 18 (42,8%) белков и 2 (4,7%) белков межиндивидуальная варируемость (Inter (BioV)) оказалась менее 40%, 30% и 20% соответственно.



**Рисунок 6.** Диаграмма размаха (диаграмма “ящик с усами”) для содержания 42 белков протеомного ядра плазмы крови. Результаты показаны в виде десятичного логарифма концентрации белка, измеренной в цельной плазме крови в нМ (Log10 (нМ)). Границы “ящика” демонстрируют 25-й (Q1) и 75-й (Q3) процентиля соответственно, линия в середине “ящика” соответствует медиане (50-й процентиль) (Median), планки погрешностей отражают максимальные (Max) и минимальные (Min) значения.



**Рисунок 7.** Распределение биологической вариабельности (Inter (BioV)) для 42 белков протеомного ядра плазмы крови, измеренных во всех экспериментальных образцах (n=31), TechV<20%, IntraV<30%, для белков с двумя протеотипическими пептидами выбран пептид, для которого измеренная в плазме крови концентрация была выше.



**Рисунок 8.** Корреляция содержания 39 белков протеомного ядра плазмы крови, измеренного во всех образцах цельной плазмы крови здоровых добровольцев, и содержания этих же белков по данным литературного обзора Hortic и соавт. [18]. Содержание показано в виде десятичного логарифма концентрации белка, измеренной в цельной плазме крови в нМ (Log10 (нМ)). Для белков PROS, BTG, CXCL7, содержание которых было измерено во всех образцах от здоровых добровольцев, в литературном обзоре не было данных о количественном содержании.

Наибольшая межиндивидуальная вариабельность была обнаружена для гаптоглобина (68%), константного участка тяжелой цепи иммуноглобулина класса D (90%), связывающего половые стероиды глобулина SHBG (100%) и липопротеин(а) (136%). Примеры MRM сигналов приведены на рисунке 1 дополнительных материалов. Уровень содержания липопротеина(а) коррелирует с риском развития ишемической болезни сердца [16]. Иммуноглобулин D является важным иммуномодулятором, поддерживающим баланс между иммунным ответом и воспалением на уровне врожденного иммунитета [17].

Экспериментальные количественные данные, полученные в результате MRM анализа цельной плазмы крови, сравнили с данными литературного обзора Hortic и соавт. [18], в котором была собрана информация о концентрациях наиболее высоко представленных белков в плазме крови здоровых людей, полученная с помощью различных физико-химических методов, в том числе масс-спектрометрии, гель-электрофореза и ИФА. График на рисунке 8 демонстрирует корреляцию содержания 39 белков протеомного ядра плазмы крови, измеренного в экспериментальных образцах и известного по литературным данным. Содержание выражено в виде десятичного логарифма концентрации белка, измеренной в цельной плазме крови в нМ (Log10 (нМ)). Как видно из рисунка 8, экспериментальные количественные данные коррелируют с литературными данными с  $R^2=0,84$ , что подчеркивает применимость метода направленной масс-спектрометрии для адекватной количественной оценки диагностически значимых белков.

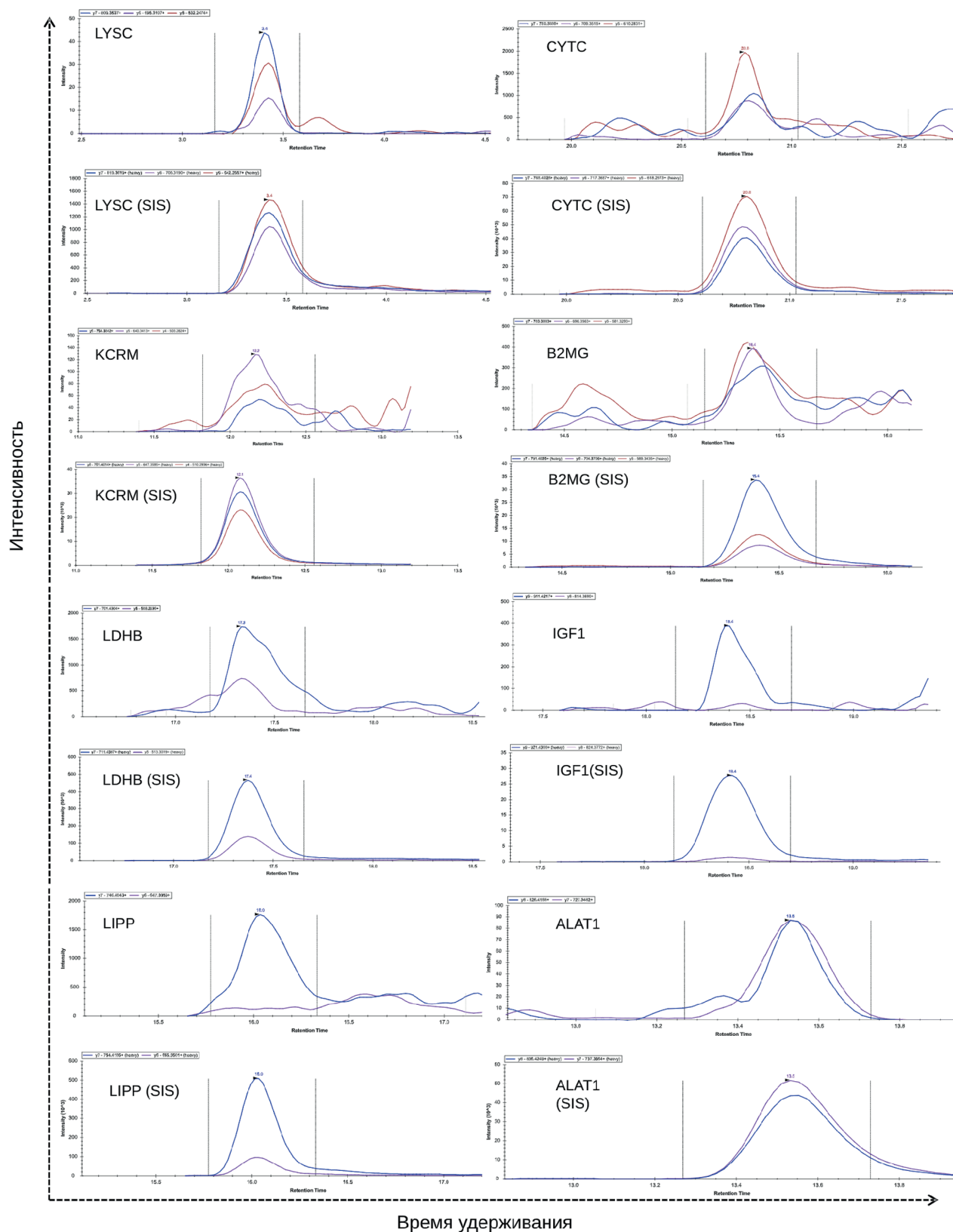
*Непостоянно регистрируемые и никогда не регистрируемые MRM у здоровых добровольцев белки*

Пептиды для 8 белков непостоянно регистрировались в экспериментальных образцах. Примеры MRM сигналов приведены на рисунке 9.

Из редко регистрируемых белков лизоцим (LYSC, ATNYNAGDR) отмечен у 25 из 31 здоровых добровольцев (11 женщин и 14 мужчин, проходивших обследование в клинике лечебного питания). Протеотипический пептид для лизоцима отличается высокой степенью гидрофильности, элюируется с колонки обращенной фазы в самом начале аналитического градиента, в результате чего хромато-масс-спектрометрический сигнал отличается высокой нестабильностью как с точки зрения времени удерживания, так и с точки зрения интенсивности сигнала.

Цистатин С (CYTC, ALDFAVGEYNK), уровень которого коррелирует с риском развития сердечно-сосудистых заболеваний и позволяет оценить функцию почек, обнаруживался у 17 из 31 здоровых добровольцев (9 женщин и 8 мужчин (1 кандидат в космонавты и 7 мужчин, проходивших обследование в клинике лечебного питания)).

Субъединица креатинкиназы М-типа (KCRM, TDLNHENLK), которая в комбинации с такой же субъединицей М-типа преимущественно находится в скелетных мышцах, а в комбинации с субъединицей В-типа — в сердечной мышце, встречалась у 8 из 31 здоровых добровольцев (4 женщины и 4 мужчин, проходивших обследование в клинике лечебного питания).



**Рисунок 9.** Результаты MRM анализа для лизоцима (LYSC, ATNYNAGDR), цистатина С (CYTC, ALDFAVGEYNK), субъединицы креатинкиназы М-типа (KCRM, TDLNHNENLK), бета-2-микроглобулина (B2MG, VEHSDFSFSK), бета цепи лактатдегидрогеназы (LDHB, IVVVTAGVR), инсулиноподобного фактора роста 1 (IGF1, GFYFNKPTGYGSSSR), панкреатической липазы (LIPP, IIVETNVGK), аланинаминотрансферазы (ALAT1, LLVAGEGHTR); сигнал от природного пептида и синтетического стандарта для пептидов (по оси x — время удерживания, по оси y — интенсивность сигнала) (визуализация в ПО Skyline). В лучшем качестве результаты MRM анализа для каждого белка показаны на рисунках 3-10 файла с дополнительными материалами.

Бета-2-микроглобулин (B2MG, VEHSDFSFSK), экспрессирующийся на поверхности многих клеток, в особенности лимфоцитов, является маркером активности иммунной системы, начиная с воспаления и заканчивая онкогематологическими заболеваниями, кроме того высокий уровень бета-2-микроглобулина коррелирует со снижением почечного клиренса. Данный белок обнаруживался у 7 из 31 здоровых добровольцев (5 женщин и 2 мужчин, проходивших обследование в клинике лечебного питания).

Бета-цепь лактатдегидрогеназы (LDHB, IVVVTAGVR), обнаруживаемая преимущественно в сердечной мышце и являющаяся маркером её поражения, детектирована у 6 из 31 здоровых добровольцев, при этом 5 образцов были получены от кандидатов в космонавты (мужчины) и только 1 образец от здорового добровольца, проходившего обследование в клинике лечебного питания (женщина).

Инсулиноподобный фактор роста 1 (IGF1, GFYFNKPTGYGSSSR), уровень содержания которого коррелирует с количеством гормона роста, обнаружен у 5 здоровых добровольцев (4 мужчин (3 кандидата в космонавты и мужчина, проходивший обследование в клинике лечебного питания) и 1 женщина).

Панкреатическая липаза (LIPP, IVETNVGK), отражающая функции поджелудочной железы, обнаружена только в образце одного здорового добровольца (мужчина, кандидат в космонавты).

Аланинаминотрансфераза (ALAT1, LLVAGEGHTR), отражающая функции печени, обнаружена только в образце одного здорового добровольца (женщина, проходившая обследование в клинике лечебного питания).

Для 55 белков масс-спектрометрический сигнал отсутствовал во всех образцах плазмы крови. В их числе онкомаркеры CAEM5, ERBB2, альфа-фетопротеин (АФП) и простатспецифический антиген (ПСА).

Возможными причинами отсутствия масс-спектрометрического сигнала для редко регистрируемых белков могут быть высокая гидрофильность протеотипического пептида, интерференция и/или низкая интенсивность сигнала вследствие малого содержания целевого пептида в образце.

Для белков, непостоянно регистрируемых в экспериментальных образцах, актуальна оптимизация метода с целью улучшения качества и интенсивности сигнала для пептидных ионов. Для повышения чувствительности может быть применено иммуноаффинное обогащение низкопредставленных белков плазмы крови или двумерная хроматография, объединяющая обогащение в щелочных условиях с последующим аналитическим разделением в кислых условиях [8, 9, 19]. Для белков, никогда не регистрируемых у здоровых добровольцев, кроме технического совершенствования метода возможен поиск данных белков с помощью MRM метода у пациентов с различными патологиями.

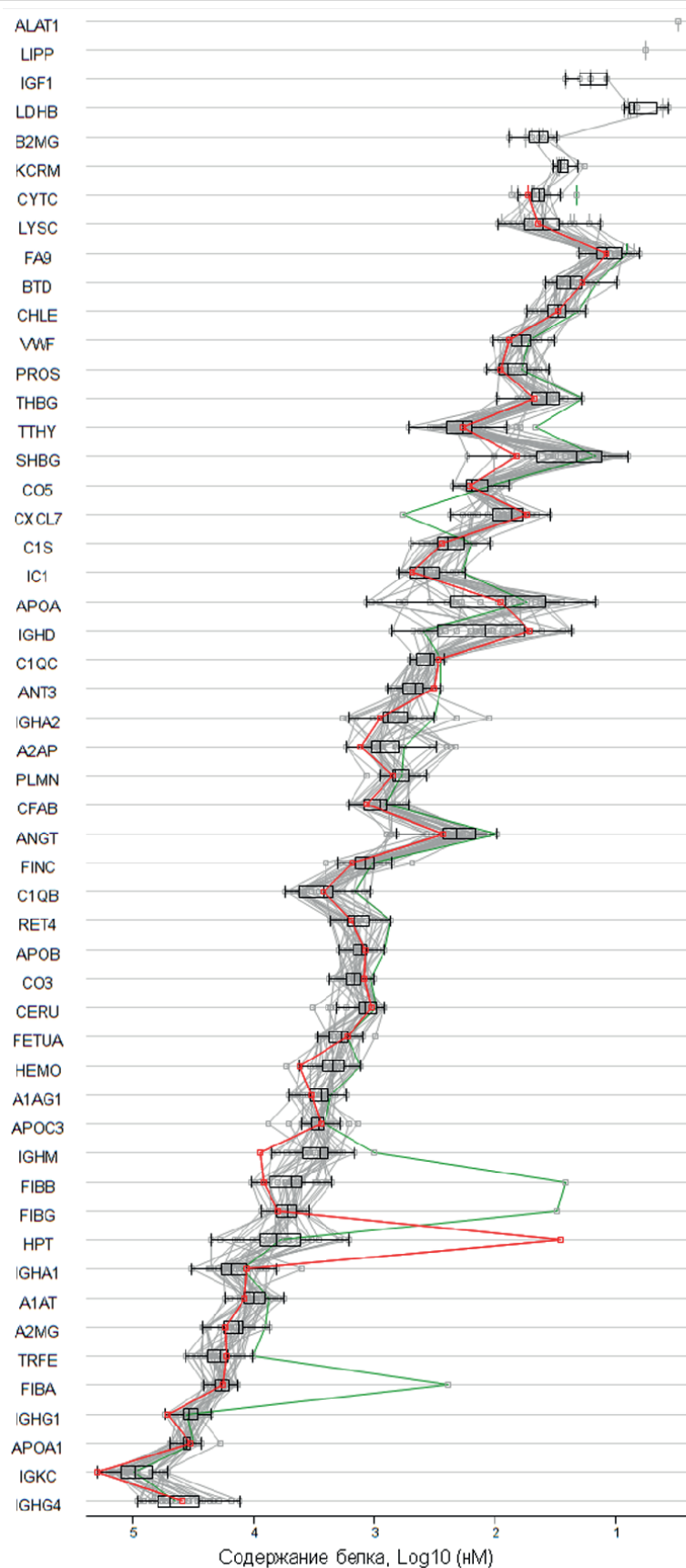
*Смена парадигмы: переход от независимых и комплексных тестов к мультиплексным тестам для создания цифрового образа протеома плазмы человека*

Независимые и комплексные тесты с использованием антител основаны на использовании различных физико-химических методов измерения содержания белков (иммунотурбидиметрия, иммунохемилюминесценция, иммунофлуоресценция, УФ-кинетический тест и т.д.). В результате измеренные величины, даже в случае комплексных тестов, не всегда можно привести к одному виду, и каждый белковый анализ необходимо рассматривать по отдельности в сравнении с его референсными нормами. Одним из главных преимуществ мультиплексных тестов является возможность статистической визуализации данных в виде тепловых карт и профилей экспрессии для всех измеренных аналитов. На рисунках 10-12 показана визуализация результатов измерения FDA-верифицированных белков в образцах плазмы крови, полученных от 31 здорового добровольца. Рисунок 10 показывает профили для каждого здорового добровольца, включающие 44 белка, которые были детектированы в каждом образце плазмы крови здоровых пациентов, и 8 белков, которые непостоянно регистрировались в экспериментальных образцах. Подобное представление данных хорошо демонстрирует образцы, в которых для отдельных белков были характерны выбросы, в виде “сигнала” на фоне “шума” межиндивидуальной вариабельности. Данный формат представления данных может быть полезен при скрининге большой выборки образцов, поскольку обращает внимание на отдельные пробы, содержание белков в которых сильно отличается, что позволяет заподозрить наличие патологического процесса.

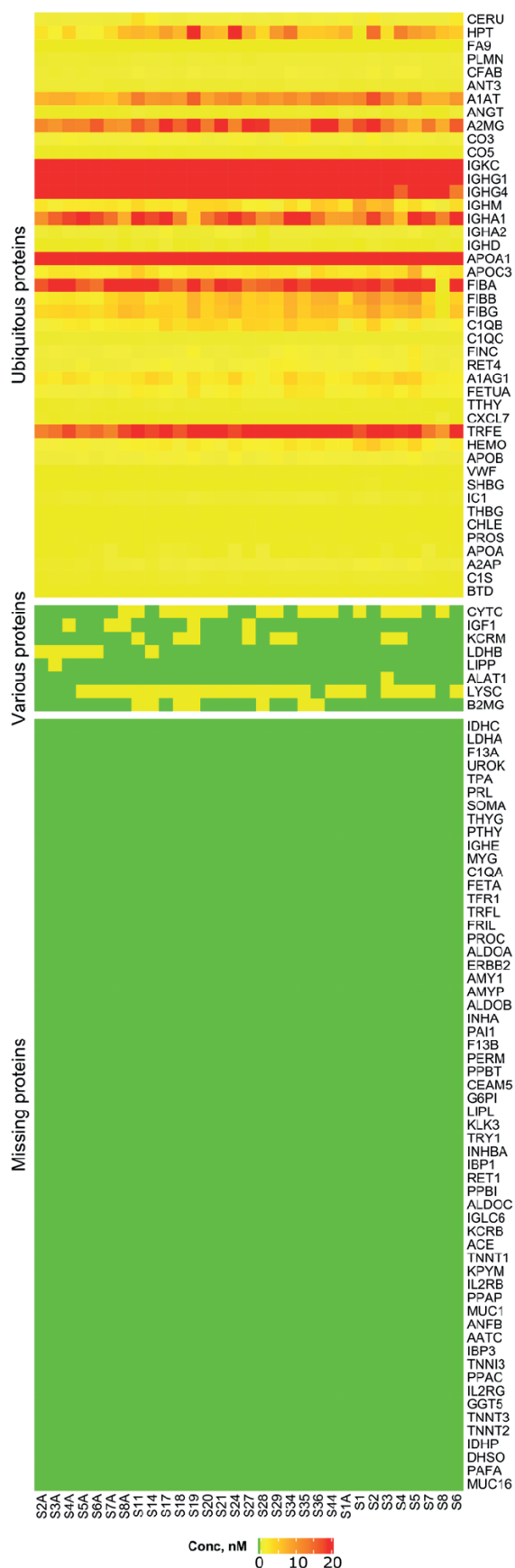
Рисунок 11 показывает результаты анализа паттерна детекции для 111 FDA-верифицированных белков-маркеров в виде тепловой карты. Аналогичный рисунок был построен для паттерна присутствия/отсутствия белков.

Исходя из результатов анализа, все белки были разделены на три группы. В первую группу (Ubiquitous proteins) входили белки, которые были детектированы в каждом образце плазмы крови здоровых пациентов. Вторую группу (Various proteins) составили белки, которые детектировались только у части пациентов. Наконец, третью группу (Missing proteins) составили белки из списка FDA, которые не были зарегистрированы ни в одном из проанализированных образцов. Количество белков группы Ubiquitous proteins составило n=44, количество белков группы Various proteins составило n=8, количество белков группы Missing proteins составило n=59 (Списки белков, составляющих каждую группу приведены в таблицах 1-3 дополнительных материалов).

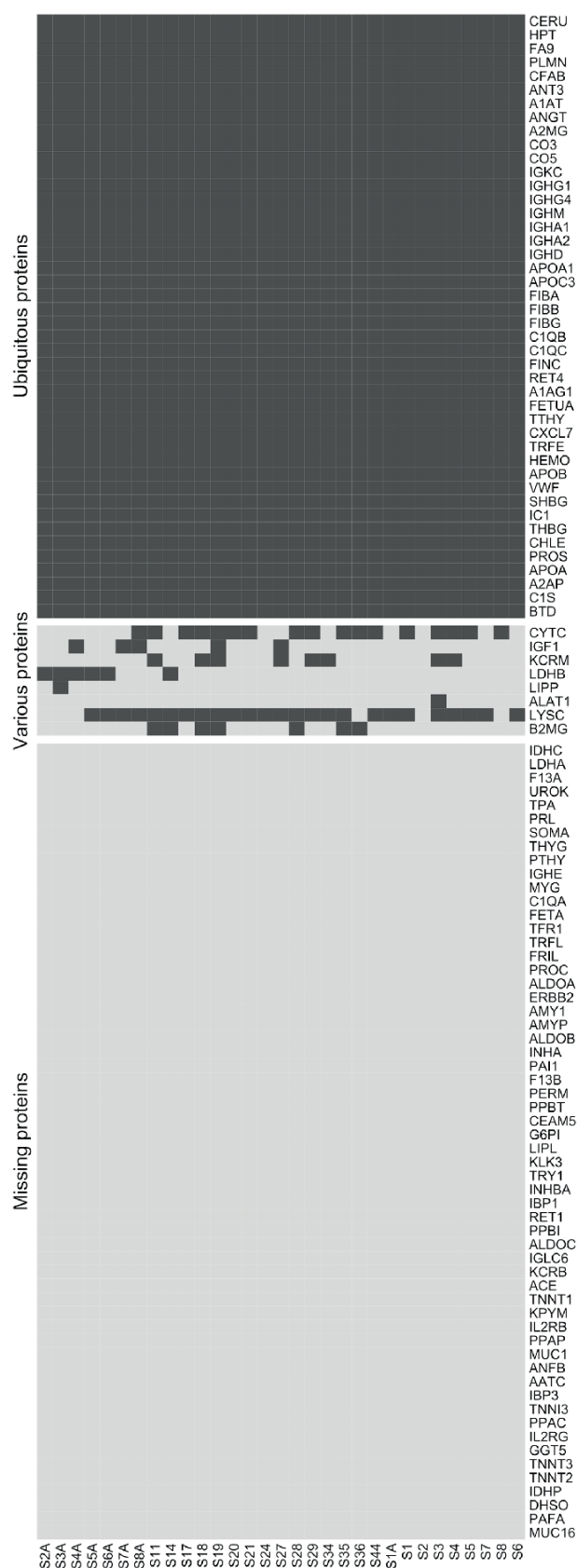
Рисунок 12 показывает в виде тепловой карты кластеризацию образцов на основании профилей содержания для 42 белков, составляющих протеомное ядро плазмы крови, и 8 белков, непостоянно регистрируемых в экспериментальных образцах.



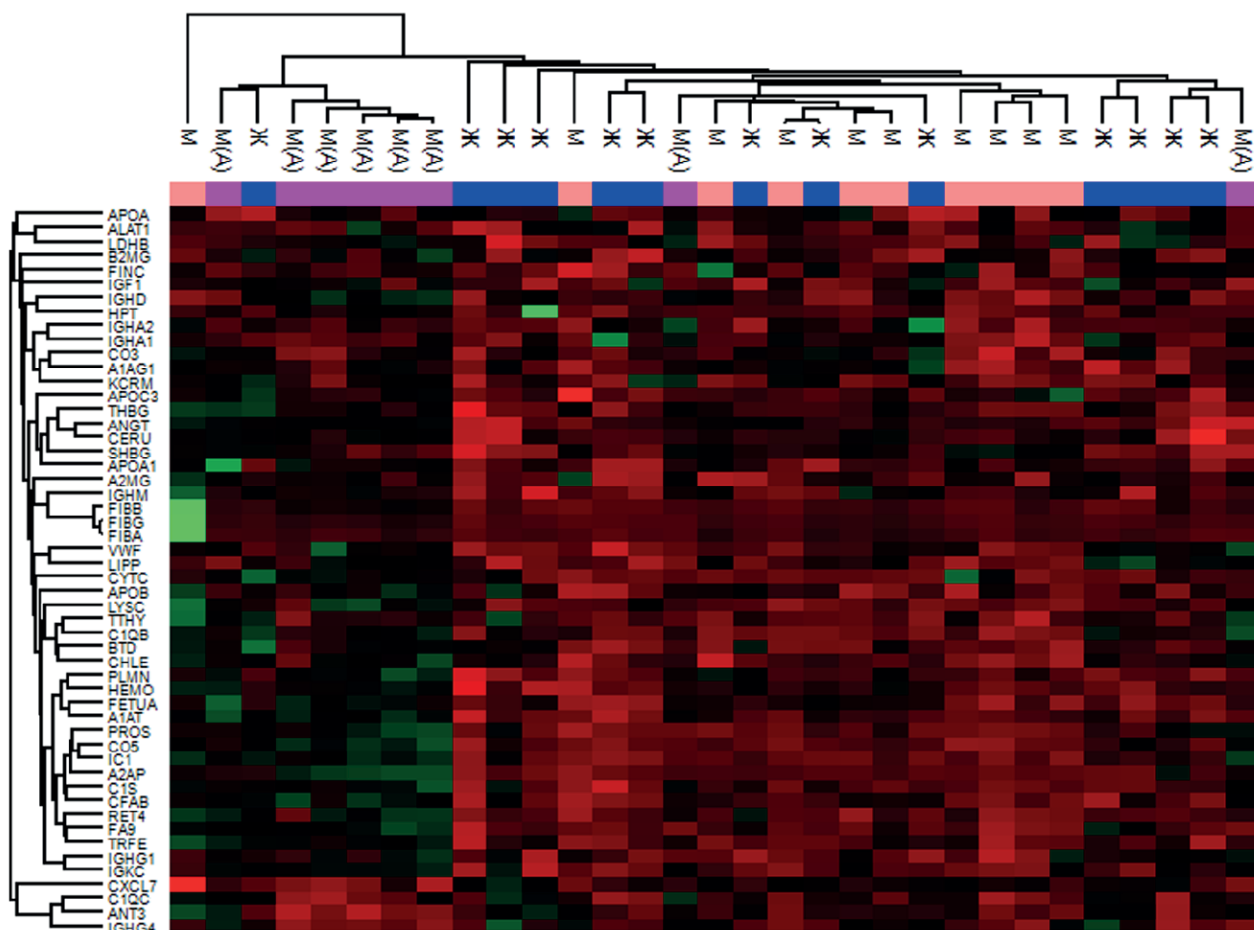
**Рисунок 10.** Профили для образцов плазмы крови, полученной от 31 здорового добровольца, включающие 44 белка, которые были детектированы в каждом образце плазмы крови здоровых пациентов, и 8 белков, которые непостоянно регистрировались в экспериментальных образцах. По оси у отложены значения  $\text{Log}_2$  nM. Визуализация была выполнена в программном обеспечении Perseus 1.6.0.7. Выделены пробы, для отдельных белков которых были получены значения, сильно отличающиеся от остальной выборки.



**Рисунок 11.** Тепловая карта измерения концентрации для 111 белков, утверждённых FDA в качестве биомаркеров (FDA-верифицированные белки).



**Рисунок 12.** Тепловая карта паттерна детекции 111 белков утверждённых FDA в качестве биомаркеров.



**Рисунок 13.** Тепловая карта распределения концентраций для 42 белков, составляющих протеомное ядро плазмы крови, и 8 белков, которые непостоянно регистрировались в экспериментальных образцах, и кластеризация экспериментальных образцов (n=31). Визуализация была выполнена в программном обеспечении Perseus 1.6.0.7.

Возможность статистической визуализации обеспечивает задел для создания цифрового образа протеома плазмы человека.

Как видно из рисунка 13, 6 из 8 образцов, полученных от кандидатов в космонавты, группируются в отдельный кластер, что может быть обусловлено биологическими причинами, поскольку данная группа здоровых добровольцев испытывает повышенные и специфические, такие как состояние невесомости, физические нагрузки в ходе тренировок. Применение дополнительной статистической визуализации позволяет выделить группы с похожим профилем для белков протеомного ядра плазмы крови здорового человека.

## ЗАКЛЮЧЕНИЕ

Анализ белковых маркеров для диагностики заболеваний человека остается важной задачей для современной медицины. Технология ИФА, использующая антитела, до сих пор является “золотым стандартом” биохимической диагностики биологических жидкостей, в первую очередь плазмы крови. Несмотря на очевидные достоинства метода, такие как точность и чувствительность, современная медицина нуждается в новых подходах

к количественной и качественной оценке белковых аналитов. Все больше исследований указывают на важность одновременного анализа множества белковых маркеров или “панели” белков, для оценки состояния здоровья человека с позиции прецизионной медицины [20]. Для мультиплексного измерения белков был разработан масс-спектрометрический направленный метод MRM, техническая вариабельность которого для большинства измерений оказалась менее 10%. Погрешность, которую вносит процедура пробоподготовки отличается от одного протеотипического пептида к другому, но не превышает 30%. Для определения количества белка использовали протеотипический пептид, выход которого в результате ферментативного расщепления был максимальным. На данный момент удалось измерить содержание 42 белков в 31 образце в диапазоне концентраций, охватывающем примерно 5 порядков. Межиндивидуальная вариабельность для 30 из 42 белков оказалась менее 40%. Наибольший разброс наблюдался для гаптоглобина (68%), константного участка тяжелой цепи иммуноглобулина класса D (90%), ангиотензина (72%) связывающего половые стероиды глобулина SHBG (100%) и липопротеин-(а) (136%). Полученные количественные данные коррелируют с литературными данными [18] с  $R^2=0,84$ .

По результатам работ нами был создан прототип тест-системы для мультиплексного, количественного анализа FDA-верифицированных белков в цельной плазме крови, отличающейся высокой специфичностью и чувствительностью. Данная тест-система пластична, её можно модифицировать путём добавления дополнительных белковых аналитов и/или перегруппировки компонентов для создания персонализированных панелей аналитов с учётом индивидуальных потребностей каждого человека. Возможность статистической визуализации данных в виде тепловых карт и профилей экспрессии для всех измеренных аналитов открывает новые возможности для интерпретации результатов, полученных на больших когортах добровольцев. Разработанный прототип тест-системы можно рассматривать в качестве альтернативы методикам, использующим моноклональные антитела.

## БЛАГОДАРНОСТИ

Работы были выполнены с помощью оборудования ЦКП ИБМХ “Протеом человека”, протоколы исследований можно найти на сайте <http://proteocenter.ibmc.msk.ru>.

## ФИНАНСИРОВАНИЕ

Работа выполнена при поддержке гранта РНФ №20-15-00410.

## СОБЛЮДЕНИЕ ЭТИЧЕСКИХ СТАНДАРТОВ

У всех добровольцев было получено письменное информированное согласие.

## КОНФЛИКТ ИНТЕРЕСОВ

Авторы заявляют об отсутствии конфликта интересов.

Дополнительные материалы доступны в электронной версии статьи на сайте журнала ([pbmc.ibmc.msk.ru](http://pbmc.ibmc.msk.ru)).

## ЛИТЕРАТУРА

- Lomnytska M., Pinto R., Becker S., Engström U., Gustafsson S., Björklund C., Templin M., Bergstrand J., Xu L., Widengren J., Epstein E., Franzén B., Auer G. (2018) *Biomark. Res.*, **6**(1), 2. DOI: 10.1186/s40364-018-0118-y
- Bhardwaj M., Gies A., Werner S., Schrotz-King P., Brenner H. (2017) *Clin. Transl. Gastroenterol.*, **8**(11), e128.
- Niu L., Song X., Wang N., Xue L., Song X., Xie L. (2019) *Cancer Sci.*, **110**(1), 433-442.
- Geyer P.E., Kulak N.A., Pichler G., Holdt L.M., Teupser D., Mann M. (2016) *Cell Syst.*, **2**(3), 185-195.
- Anderson N.L., Anderson N.G. (2002) *Mol. Cell. Proteomics*, **1**(11), 845-867.
- Kopylov A.T., Ponomarenko E.A., Ilgisonis E.V., Pyatnitskiy M.A., Lisitsa A.V., Poverennaya E.V., Kiseleva O.I., Farafonova T.E., Tikhonova O.V., Zavialova M.G., Novikova S.E., Moshkovskii S.A., Radko S.P., Morukov B.V., Grigoriev A.I., Paik Y.K., Salekdeh G.H., Urbani A., Zgoda V.G., Archakov A.I. (2018) *J. Proteome Res.*, **18**(1), 120-129.
- Yeh C.Y., Adusumilli R., Kullolli M., Mallick P., John E.M., Pitteri S.J. (2017) *Biomark Res.*, **5**, 30. DOI: 10.1186/s40364-017-0110-y. eCollection 2017.
- Kopylov A.T., Ilgisonis E.V., Moysa A.A., Tikhonova O.V., Zavialova M.G., Novikova S.E., Lisitsa A.V., Ponomarenko E.A., Moshkovskii S.A., Markin A.A., Grigoriev A.I., Zgoda V.G., Archakov A.I. (2016) *J. Proteome Res.*, **15**(11), 4039-4046.
- An B., Zhang M., Pu J., Shen S., Qu Y., Chen Y.J., Huo S., Wang X., Polli J.R., Balthasar J.P., Herzog D., Ferrari L., Staack R.F., Richter W.F., Otteneder M.B., Benincosa L.J., Zhou S., Vazvaei F., Qu J. (2018) *Anal. Chem.*, **90**(3), 1870-1880.
- Clerc F., Reiding K.R., Jansen B.C., Kammeijer G.S., Bondt A., Wührer M. (2016) *Glycoconjugate J.*, **33**(3), 309-343.
- Deutsch E.W., Lane L., Overall C.M., Bandeira N., Baker M.S., Pineau C., Moritz R.L., Corrales F., Orchard S., van Eyk J.E., Paik Y., Weintraub S.T., Vandenbrouck Y., Omenn G.S. (2019) *J. Proteome Res.*, **18**(12), 4108-4116.
- van der Straten A., Herzog A., Cabezyn T., Bollen A. (1984) *FEBS Lett.*, **168**(1), 103-107.
- Bensi G., Raugei G., Klefenz H., Cortese R. (1985) *EMBO J.*, **4**(1), 119-126.
- Sadrzadeh S.M.H., Bozorgmehr J. (2004) *Am. J. Clin. Path.*, **121**, 97-104.
- Hamdy G., Hendy O., Mahmoud H., El A., Ali S., Khalaf F. (2014) *Egypt. J. Med. Hum. Genet.*, **15**(3), 257-264.
- Nordestgaard B.G., Chapman M.J., Ray K., Borén J., Andreotti F., Watts G.F., Ginsberg H., Amarenco P., Catapano A., Descamps O.S., Fisher E., Kovanen P.T., Kuivenhoven J.A., Lesnik P., Masana L., Reiner Z., Taskinen M.R., Tokgözoğlu L., Tybjaerg-Hansen A. (2010) *Eur. Heart J.*, **31**(23), 2844-2853.
- Chen K., Cerutti A. (2010) *Immunol. Revs.*, **237**(1), 160-179.
- Hortin G.L., Sviridov D., Anderson N.L. (2008) *Clin. Chem.*, **54**(10), 1608-1616.
- Ilgisonis E.V., Kopylov A.T., Ponomarenko E.A., Poverennaya E.V., Tikhonova O.V., Farafonova T.E., Novikova S., Lisitsa A.V., Zgoda V.G., Archakov A.I. (2018) *J. Proteome Res.*, **17**(12), 4258-4266.
- Rosser C.J., Dai Y., Miyake M., Zhang G., Goodison S. (2014) *BMC Biotechnol.*, **14**(1), 24. DOI: 10.1186/1472-6750-14-24.

Поступила в редакцию: 21. 05. 2020.  
После доработки: 07. 07. 2020.  
Принята к печати: 07. 07. 2020.

MASS-SPECTROMETRIC MRM ANALYSIS OF FDA-VERIFIED PROTEINS  
IN THE BLOOD PLASMA OF HEALTHY VOLUNTEERS

*S.E. Novikova<sup>1\*</sup>, T.E. Farafonova<sup>1</sup>, O.V. Tikhonova<sup>1</sup>, N.A. Shushkova<sup>1</sup>, M.A. Pyatnitsky<sup>1</sup>,  
V.G. Zgoda<sup>1</sup>, E.A. Ponomarenko<sup>1</sup>, A.V. Lisitsa<sup>1</sup>, A.I. Grigoryev<sup>2</sup>, V.A. Tutelyan<sup>3</sup>, A.I. Archakov<sup>1</sup>*

<sup>1</sup>Institute of Biomedical Chemistry,  
10 Pogodinskaya str., Moscow, 119121 Russia; \*e-mail: novikova.s.e3101@gmail.com

<sup>2</sup>Institute of Biomedical Problems, Russian Academy of Sciences,  
76a Khoroshevskoye shosse, Moscow, 123007 Russia

<sup>3</sup>Institute of Nutrition, 2/14 Ustinsky proezd, Moscow, 109240 Russia

The proteomic composition of a biological sample serves as the most important feature of a biological object, and it allows discriminating normal and pathological conditions. Targeted mass spectrometric analysis, namely, multiple reaction monitoring (MRM) using synthetic isotopically-labeled internal standard (SIS), is the main alternative to the ELISA method for the analysis of diagnostically significant proteins. Based on the MRM results, a prototype test system has been developed; it employs the targeted mass spectrometric method for multiplex, quantitative analysis of FDA-verified proteins in whole blood plasma. Using this approach, it was possible to measure the content of 42 proteins in 31 samples in a concentration range spanning five orders of magnitude. The interindividual variability for 30 of the 42 registered proteins was less than 40%. The largest scatter was observed for haptoglobin (68%), immunoglobulin heavy constant delta IGHD (90%), angiotensin (72%), sex hormone-binding globulin SHBG (100%) and lipoprotein-(a) (136%). The obtained results on the concentration of proteins correlate with published data (Hortin et al., 2008, *Clinical Chemistry*, **54**, 1608) with  $R^2=0.84$ . The developed prototype test system based on targeted mass spectrometric analysis of proteins can be considered as an alternative to methods using monoclonal antibodies.

**Key words:** SRM analysis; isotope-labeled internal peptide standards; FDA-verified proteins

**Funding.** This work was supported by the Russian Science Foundation (project No. 20-15-00410).

Received: 21.05.2020, revised: 07.07.2020, accepted: 07.07.2020.