

Федеральное государственное бюджетное учреждение науки
ИНСТИТУТ БИООРГАНИЧЕСКОЙ ХИМИИ
ИМ. АКАДЕМИКОВ М.М. ШЕМЯКИНА И Ю.А. ОВЧИННИКОВА
РОССИЙСКОЙ АКАДЕМИИ НАУК

На правах рукописи

Арапиди Георгий Павлович

ПОИСК И ИДЕНТИФИКАЦИЯ ПОТЕНЦИАЛЬНЫХ БИОМАРКЕРОВ
РАКА ЯИЧНИКОВ В СЫВОРОТКЕ КРОВИ ЧЕЛОВЕКА

03.01.04 – биохимия

Диссертация на соискание ученой степени
кандидата биологических наук

Научный руководитель:
кандидат химических наук,
Зиганшин Рустам Хусманович

Москва 2015 год

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ.....	6
Актуальность темы исследования	6
Цели и задачи исследования	7
Научная новизна работы.....	8
Теоретическая и практическая значимость работы.....	8
Апробация работы.....	9
Личный вклад автора	9
ЛИТЕРАТУРНЫЙ ОБЗОР	10
1. Рак яичников.....	10
1.1. Злокачественные новообразования	10
1.2.Рак яичников.....	11
1.3. Современные методы диагностики рака яичников	13
2. Протеомные технологии для поиска новых маркеров онкологических заболеваний.....	17
2.1. Протеомика и пептидомика	17
2.2. Масс-спектрометрия – основной инструментальный метод протеомики	21
2.2.1. Методы ионизации биологических макромолекул	21
2.2.2. Основные типы масс-анализаторов	23
2.2.3. Tandemная масс-спектрометрия	25
2.2.4. LC-MS/MS – tandemная масс-спектрометрия, сопряженная с высокоэффективной жидкостной хроматографией.....	26
2.2.5. Методы количественного масс-спектрометрического анализа	27
2.3. Биоинформатические подходы в протеомике.....	29
2.3.1 Построение математических моделей классификации.....	30
2.3.2 Идентификация пептидов и валидация результатов идентификации	31
2.4 Кровь, как объект биомаркерных исследований	33

2.4.1 Плазма и сыворотка крови	34
2.4.2. Гипотеза образования пептидных биомаркеров	35
2.4.3. Основные сложности протеомных и пептидомных исследований производных крови	38
2.4.4. Методы выделения пептидов из плазмы/сыворотки крови.....	39
2.4.4.1. Ультрафильтрация	40
2.4.4.2. Преципитация белков органическими растворителями	40
2.4.4.3. Твердофазная экстракция.....	41
2.4.4.4. Выделение пептидов на хроматографических сорбентах с ограниченным доступом к сорбционным центрам	44
2.4.4.5 Деплеция основных белков плазмы/сыворотки крови.....	45
2.4.5 Методы десорбции пептидов с поверхности основных белков плазмы крови	45
2.4.6. Использование масс-спектрометрии для поиска в плазме и сыворотке крови потенциальных маркеров онкологических заболеваний.....	46
2.4.7. Современные достижения в области поиска белковых/пептидных биомаркеров рака яичников	48
МАТЕРИАЛЫ И МЕТОДЫ	52
1. Образцы сыворотки крови.....	52
2. МАЛДИ МС профилирование сыворотки крови.....	54
2.1 Фракционирование образцов сыворотки крови.....	54
2.1.1 Протокол фракционирования с использованием MB-WCX.....	55
2.1.2 Протокол фракционирования с использованием MB-HIC 8 (18) ...	55
2.1.3 Протокол фракционирования с использованием MB-IMAC Cu.....	56
2.2. Стадия прогревания элюатов образцов сыворотки крови	57
2.3. МАЛДИ масс-спектрометрическое оборудование	57
2.4. Анализ масс-спектрометрических данных	58
3. Идентификация потенциальных пептидных биомаркеров.....	60

3.1 Фракционирование образцов сыворотки крови.....	60
3.1.1 Протокол фракционирования с использованием MB-WCX.....	60
3.2 Стадия прогрева элюатов образцов сыворотки крови	60
3.3 Обессоливание на микроколонках с обращено фазовой поверхностью .	61
3.3.1 Протокол фракционирования на микроколонках с обращено фазовой поверхностью	61
3.4 Масс-спектрометрическое оборудование.....	61
3.5 Идентификация пептидов и биоинформатический анализ.....	62
4. Количественный анализ содержания пептидов в сыворотке крови методом SWATH.....	63
4.1 Масс-спектрометрическое оборудование.....	64
4.2 Биоинформатический анализ данных SWATH.....	64
РЕЗУЛЬТАТЫ.....	66
1. Метод выделения пептидов из сыворотки крови человека	66
1.1 Сравнение пептидного и белкового состава элюатов до и после прогрева	66
1.2 Прогревание белка Apolipoprotein A-I	69
2. Построение классификационных моделей	70
2.1 Фракционирование образцов сыворотки крови с использованием различных типов магнитных микрочастиц	70
2.2 Построение классификационных моделей на основе профилирования на магнитных микрочастицах MB-WCX.....	73
2.3 Построение классификационных моделей сравнения групп «рак яичников» и «номинальный контроль»	78
2.4 Влияние длительности инкубации крови до получения сыворотки на последующий масс-спектрометрический профиль	80
3. Идентификация потенциальных биомаркеров рака яичников	83
4. Анализ потенциальных биомаркеров количественным безметочным масс- спектрометрическим методом SWATH	89

5. Взаимосвязь результатов МАЛДИ масс-спектрометрического профилирования и количественного анализа на основе метода SWATH.....	92
ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ	94
1. Преимущества, которые дает разработанный метод выделения пептидов из сыворотки крови человека	94
2. Кислотный гидролиз на стадии прогрева элюатов.....	95
3. Результаты МАЛДИ масс-спектрометрического профилирования образцов сыворотки крови.....	98
4. Идентификация пептидов в сыворотке крови человека.....	100
5. Количественный анализ пептидов в сыворотке крови.....	101
ЗАКЛЮЧЕНИЕ	106
ВЫВОДЫ	107
БЛАГОДАРНОСТИ.....	109
СПИСОК СОКРАЩЕНИЙ И УСЛОВНЫХ ОБОЗНАЧЕНИЙ	110
СЛОВАРЬ ТЕРМИНОВ.....	112
СПИСОК ЛИТЕРАТУРЫ	113
ПРИЛОЖЕНИЕ 1.....	127

ВВЕДЕНИЕ

Актуальность темы исследования

Онкологические заболевания из года в год занимают ведущие позиции в структуре заболеваемости и смертности среди социально-значимых заболеваний. Последние несколько лет появилась тенденция к уменьшению среднего возраста пациентов, которым впервые в жизни был поставлен диагноз рак. Рак яичников называют «тихим убийцей», поскольку первые симптомы у многих больных появляются, когда рак находится на поздней стадии развития. По статистике, шансы вылечиться у таких пациентов низкие. В этой ситуации именно от правильной и своевременной диагностики зависит больше, чем от лечения. На сегодня разработано и используется несколько методов диагностики рака яичников, но ни один не обладает необходимыми значениями чувствительности и специфичности. Разработка эффективных методов диагностики является важным направлением в современной онкологии рака яичников.

Протеомика является наукой пост геномной эры, направленной на детальный анализ всех белков организма. Протеомика осуществляет сравнительный анализ больших групп белков в клетках, тканях или целом организме для нахождения взаимосвязей между белками и другими макромолекулами организма в различных состояниях и внешних условиях. Современная протеомика находится на стыке таких наук как геномика, биохимия, биоинформатика, системная биология. В основе этой области исследований лежат базы данных сиквенсов геномов, масс-спектрометрия, как основной инструментальный метод анализа белков и пептидов и биоинформатика, коррелирующая результаты масс-спектрометрического анализа со структурами белков, транслированными из геномных баз данных. Многие исследователи связывают перспективу разработки новых эффективных диагностик социально-значимых заболеваний, в том числе рака яичников, именно с успехами развития протеомики и её подходов. Несмотря на это, за последние 10 лет не было

предложено ни одного нового метода диагностики рака яичников, разработанного протеомными методами.

Пептидом является наименее изученной частью протеома. Долгое время пептидом рассматривали исключительно как “биологический мусор”, “биологический шум”, нестабильный, неспособный быть биологически значимым. Однако, давно известно, что пептиды играют важную роль в организме в качестве гормонов, факторов роста, цитокинов или защитных агентов. Пептиды участвуют в процессах передачи сигналов в дыхательной, сердечно-сосудистой, эндокринной, воспалительной и нервной системе. Сейчас все больше групп начинают изучать пептидом как потенциальный источник диагностической информации, отражающий биологические процессы внутри организма.

Цели и задачи исследования

Цель настоящей работы состояла в выявлении и идентификации в сыворотки крови пациенток с раком яичников диагностически значимых для этого заболевания пептидов.

Для достижения этой цели были поставлены следующие задачи:

1. Разработать высокопроизводительный и воспроизводимый метод выделения пептидов из сыворотки крови человека.
2. Используя разработанный метод выделения пептидов, провести сравнительное профилирование сывороток крови практически здоровых доноров и пациенток с раком яичников времяпролетной МАЛДИ масс-спектрометрией (MALDI-TOF MS) с целью выявления пептидов - потенциальных маркеров рака яичников.
3. Провести выделение и идентификацию пептидов из пулированных образцов сыворотки крови практически здоровых доноров и пациенток с раком яичников методом тандемной масс-спектрометрии, сопряжённой с высокоэффективной жидкостной хроматографией (LC-MS/MS).
4. Провести соотнесение значений m/z потенциальных маркеров рака яичников, выявленных в результате сравнительного MALDI-TOF MS

профилирования образцов сывороток крови практически здоровых доноров и пациенток с раком яичников, с идентифицированными методом LC-MS/MS пептидами.

5. Используя количественный безметочный масс-спектрометрический метод анализа (SWATH) пептидных фракций пулированных образцов сыворотки крови практически здоровых доноров и пациенток с раком яичников, выявить пептиды, содержание которых в исследованных группах образцов различается в 2 или более раза.

6. Выявить корреляции между количественными данными, полученными в результате анализа пептидных фракций сывороток крови методом SWATH и MALDI-TOF MS профилирования для пептидов, с совпадающими значениями m/z .

Научная новизна работы

В диссертации разработан и описан метод выделения пептидов из сыворотки крови человека. С использованием разработанного метода было проведено МАЛДИ масс-спектрометрическое профилирование образцов сыворотки крови практически здоровых доноров, пациенток с подтвержденным диагнозом рак яичников, а также пациенток с доброкачественными гинекологическими заболеваниями. По результатам LC-MS/MS анализа удалось идентифицировать почти восемь тысяч уникальных пептидов. Большое количество идентифицированных пептидов позволило провести детальные маркерные сравнения на количественном уровне, с использованием количественного безметочного масс-спектрометрического метода анализа SWATH. По глубине анализа наши результаты превосходят результаты ведущих групп в области пептидомики.

Теоретическая и практическая значимость работы

Настоящей работой мы внесли вклад в развитие клинической протеомики в вопросах разработки перспективной диагностики рака яичников. Предложенный нами метод выделения пептидов из сыворотки крови позволяет разрушать

комплексы пептидов с основными белками плазмы крови и осаждать высокомолекулярные белки. В работе показано, что при использовании разработанного метода выделения пептидов увеличивается количество регистрируемых МАЛДИ масс-спектрометрических пиков, улучшается качество и воспроизводимость масс-спектрометрических профилей. На основе разработанного метода пробоподготовки нам удалось выявить отличия между группами пациентов как на уровне МАЛДИ масс-спектрометрического профиля, так и на уровне количественного анализа пептидов методом SWATH. Данные по идентификации 7996 уникальных пептидов, а также по их количественному анализу представляют практический интерес для возможных дальнейших работ по выявлению биомаркеров.

Апробация работы

Основные результаты исследования были представлены на российских и международных конференциях: “51ая научная конференция МФТИ” (Москва, 2008), “4ый съезд ВМСО Масс-спектрометрия и её прикладные проблемы” (Москва, 2009), “4ый российский симпозиум Белки и пептиды” (Новосибирск, 2009), “Научная конференция Фундаментальная Наука для Биотехнологии и Медицины 2010” (Долгопрудный, Россия, 2010), “The 31st European Peptide Symposium (Copenhagen, Denmark, 2010), “The 10th HUPO World Congress” (Geneva, Switzerland, 2011) и “The 38th FEBS Congress” (Санкт-Петербург, 2013).

По теме диссертации опубликовано 16 печатных работ – 5 статей в реферируемых научных журналах и 11 в трудах конференций.

Личный вклад автора

Все экспериментальные научные результаты, изложенные в диссертации, их анализ, обобщение литературных данных и оценка результатов выполнены лично автором под руководством Зиганшина Р.Х.

ЛИТЕРАТУРНЫЙ ОБЗОР

1. Рак яичников

1.1. Злокачественные новообразования

Тело взрослого человека состоит примерно из 100 триллионов клеток [1]. В норме клетки растут, делятся и умирают в зависимости от потребностей организма. Организм контролирует замещение погибших клеток новыми поделившимися клетками. Скорость деления клеток различна в разных органах и тканях. Если клетки какой-то ткани изменяются под действием определенных факторов, эти клетки могут начать бесконтрольное деление и рост – образуется опухоль [2].

Опухолевые клетки обладают рядом свойств, не присущих нормальным клеткам организма: утрата специфической структуры и функции, изменение антигенного состава, агрессивный рост с разрушением окружающих тканей. Клинически опухоли представляют собой очаги роста патологической (анормальной) ткани в различных органах и структурах организма [2]. Объединяя различные критерии классификации опухолей, можно выделить две большие группы:

- Доброкачественные опухоли. Клетки доброкачественных опухолей в процессе опухолевой трансформации утрачивают способность контроля клеточного деления, но частично или почти полностью сохраняют способность к дифференцировке. По своей структуре доброкачественные опухоли напоминают ткань, из которой они происходят. Характерно также и частичное сохранение специфической функции ткани. Клинически доброкачественные опухоли проявляются как медленно растущие новообразования различной локализации. Доброкачественные опухоли постепенно сдавливают прилежащие структуры и ткани, но никогда не проникают в них. Доброкачественные опухоли, как правило, хорошо поддаются хирургическому лечению и редко рецидивируют, однако часто перерождаются в злокачественные опухоли.

- Злокачественные опухоли. Клетки злокачественных опухолей претерпевают значительные изменения, ведущие к полной утрате контроля над делением и дифференцировкой. Порой, определить источник опухоли довольно трудно из-за высокой степени атипизма. Клинически злокачественные опухоли проявляются весьма разнообразно. Им свойственен как очаговый рост, так и быстрый, агрессивный рост с проникновением в окружающие ткани и органы, кровеносные и лимфатические сосуды с образованием метастаз. Злокачественные опухоли, как правило, трудно поддаются лечению и часто рецидивируют. Именно злокачественным опухолям приписывают высокую смертность среди всех опухолевых патологий [2, 3].

Злокачественные опухоли делятся в зависимости от вида ткани, из которой они возникли [2]. Рак (карцинома) — злокачественная опухоль из эпителиальной ткани. Аденокарцинома – злокачественная опухоль из эпителия железистых тканей.

1.2. Рак яичников

Яичники – парные женские половые железы, расположенные в полости малого таза, выполняют генеративную функцию (являются местом, где развиваются и созревают женские половые клетки), а также являются железами внутренней секреции и вырабатывают половые гормоны – эндокринная функция [1].

Рак яичников – злокачественные опухоли яичников эпителиального происхождения. Рак яичников может возникать как из неизменной эпителиальной ткани (первичный), так и в результате трансформации доброкачественного процесса (вторичный). Возможно развитие рака вследствие поражении эндометрия яичников при уже имеющемся раке молочной железы или желудочно-кишечного тракта (метастатический) [2].

Рак яичников занимает седьмое место по заболеваемости и шестое место в списке причин смертности среди онкологических заболеваний в России (Таблица 1).

Таблица 1. Статистика за 2006 год по заболеваемости и смертности от злокачественных новообразований среди женского населения России [3]

Заболеваемость		Смертность	
Рак молочной железы	19,8%	Рак молочной железы	17,1%
Немеланомные новообразования	13,3%	Рак желудка	12,4%
Рак желудка	7,5%	Рак ободочной кишки	9,1%
Рак ободочной кишки	7,0%	Рак прямой кишки	6,5%
Рак тела матки	6,8%	Рак легкого	6,4%
Рак шейки матки	5,2%	Рак яичников	5,9%
Рак яичников	4,9%	Рак поджелудочной железы	5,3%

В 2006 году в России было диагностировано 12.556 случаев рака яичников, что на 2,8% больше, чем в 2005 году (12.214) и на 12,2% больше, чем в 1996 году (11.186). Число умерших от рака яичников в 2006 году составило 7.666 человек, что на 1,9% больше, чем в 2005 году (7.520) и на 21,7% больше, чем в 1999 году (6.299) [3].

Рак яичников называют «тихим убийцей», поскольку на ранних стадиях он протекает бессимптомно. Клинически у многих больных первые проявления рака яичников связаны с распространением опухоли за пределы яичника, а иногда и за пределы малого таза. По данным Противоракового Общества России среди всех диагностированных случаев рака яичников в 2002 году 64,2% составили III-IV стадии [4]. По данным Международной Федерации Акушерства и Гинекологии (FIGO) 5-летняя выживаемость среди пациентов, получивших лечение по поводу рака яичников I стадии, составляет 70%, II стадии – 46%, III стадии – 20%, а при IV стадии этот показатель снижается до 5% [5]. На сегодняшний день не существует методов своевременной диагностики рака яичников. Поиск новых методов диагностики является важным направлением в современной науке и медицине [6].

1.3. Современные методы диагностики рака яичников

Течение заболевания на ранних стадиях рака яичников проходит бессимптомно, поэтому злокачественную опухоль на I-II стадии удастся обнаружить либо при проведении периодического осмотра (к примеру, при генетической предрасположенности к раку яичников или при обнаружении доброкачественной опухоли в яичниках), либо случайно, при осмотре на какое-то другое заболевание органов брюшины, при операции в области малого таза [2].

Несмотря на высокую смертность, рак яичников достаточно редкое заболевание с частотой встречаемости около 40 случаев на 100 тысяч [7]. Поэтому, помимо чувствительности диагностики, необходимо контролировать её специфичность, чтобы травмы или смерть, связанные с хирургическим вмешательством, в случае неверно диагностированного рака, не перекрывали случаи успешной диагностики и своевременного лечения. Большинство клиницистов высказывают мнение, что на каждый случай правильно диагностированного рака яичников допустимо ошибочно выносить подобный диагноз не более чем в 9 случаях. Таким образом, подсчет показывает, что минимально допустимая специфичность метода диагностики равна 99,6% [6, 7]. Таблица 2 демонстрирует насколько существующие методы далеки от минимальных требований по специфичности и от решения проблемы ранней диагностики.

Таблица 2: Существующие методы диагностики рака яичников, а также их значения чувствительности и специфичности [6]

Метод диагностики	Стадия заболевания	Чувствительность	Специфичность
СА-125 [8]	ранние стадии	50-62%	95%
	поздние стадии	90%	95%
HE4 [9]	все стадии	95%	72,9%
УЗИ [10]	все стадии	98,1%	80,8%
УЗИ с эффектом Доплера [11]	все стадии	98%	87%
СА-125 с УЗИ [12]	все стадии	78%	99,9%

Основным методом диагностики рака яичников до сих пор остается гистологическое исследование биоптата, которое может дать точный ответ о характере и структуре опухоли [13, 14]. Однако, генетическая гетерогенность и неоднородность развития рака яичников приводит к многообразию морфологических форм, что затрудняет диагностирование стадии злокачественного процесса [13, 14]. Так в своей работе Kobel с соавторами использовали систему классификации, основанную на морфологии опухоли и показали, что она значительно отличается от системы, предлагаемой FIGO (Fédération Internationale de Gynécologie et d'Obstétrique, Международная Федерация Гинекологии и Акушерства); последняя считается международным стандартом и используется повсеместно [13]. Кроме того, любое гистологическое исследование внутренних органов инвазивно, поэтому данная диагностика не подходит для рутинного применения.

Трансвагинальное, трансректальное или классическое ультразвуковое исследование (УЗИ) позволяет детально изучить структуру яичников, чтобы различить морфологические изменения, причиной которых могут быть злокачественные процессы [6]. УЗИ позволяет детектировать кисты в яичниках, ненормальности в структуре ткани или уплотнения, факторы часто сопутствующие образованию рака яичников [15]. Преимуществом этих методов исследования является их неинвазивность, что позволяет использовать их для проведения периодического осмотра. С другой стороны, в одном из исследований было показано, что у 56% женщин в постменопаузе, которые умерли от причин отличных от гинекологических или онкологических, имелись доброкачественные новообразования размером менее 5 см [16]. Эти данные могут объяснить, почему УЗИ имеет низкую специфичность по отношению к раку яичников (Таблица 2). В другом исследовании было продемонстрировано, что повторное УЗИ через 4-6 недель увеличивает специфичность [15]. Многие современные протоколы УЗИ учитывают несколько показателей, такие как объем яичников, их форма, образование папиллом, кист, структуру стенок, толщину перегородок и прочее

[10]. Кроме того, предложены улучшения метода, использующие красители и эффект Доплера [10, 11].

Большие надежды врачи и исследователи, работающие в области лечения рака яичников, связывают с разработкой новых методов диагностики, основанных на измерении уровня биомаркеров [6]. К настоящему моменту одобрение FDA (Food and Drug Administration, управление по продовольствию и медикаментам) в США получили только два маркера: СА-125 [17] (cancer antigen 125, раковый антиген 125) и HE4 [18] (Human epididymis protein 4, Белок 4 эпидидимиса человека), а также два мультипараметрических теста: OVA1 и ROMA. Единичные маркеры СА-125 и HE4 одобрены для мониторинга рецидива или прогрессирования рака яичников, но не для скрининга [9]. Тесты OVA1 и ROMA были одобрены для применения в случаях, когда образование в яичниках уже выявлено, но не ясно является ли оно злокачественным и как срочно необходима операция [19, 20].

СА-125 является хорошо изученным [21-24] и широко используемым в клинической диагностике [6, 8] сывороточным биомаркером рака яичников. Уровень СА-125 менее 35 Ед/мл принят за норму. Примерно у 83% пациентов с диагнозом рак яичников уровень СА-125 выше 35 Ед/мл. Если разграничивать различные стадии заболевания, то на поздних стадиях уровень СА-125 превышает норму в 90% случаев, а на ранних – только в 50% (Таблица 2). Другие исследования показали, что в 25% случаев наблюдается повышенный уровень СА-125 ещё за 5 лет до диагностирования рака яичников [25]. К сожалению, СА-125 демонстрирует недостаточную специфичность, поскольку повышение уровня маркера в крови может сопутствовать некоторым другим заболеваниям, включая рак (поджелудочной железы, молочной железы, мочевого пузыря, печени, легких), доброкачественные новообразования (эндометриоз, миома матки), цирроз печени и даже при менструации и беременности [6, 26]. Для увеличения чувствительности и специфичности диагностического подхода предлагают использовать СА-125 в сочетании с другими маркерами [27, 28] или с другими методами диагностики, к

примеру с УЗИ [12]. Разработка мультипараметрических тестов является популярным и многообещающим трендом в области диагностики [29].

Маркер HE4 в 2009 году был одобрен FDA для мониторинга прогрессирования рака яичников или возможного рецидива [18]. Показана явная корреляция HE4 с определенными подтипами рака яичников: 100% специфичности при эндометриоидной форме и 93% при серозной [30]. Использование только HE4 позволяет достичь 95% чувствительности и 72,9% специфичности [9]. В комбинации с СА-125 панель из двух маркеров показала 93,8% чувствительности и 74,9% специфичности [31]. HE4 обладает большей специфичностью по отношению к доброкачественным новообразованиям в сравнении с СА-125: 71% специфичности у HE4 и 64% у СА-125 [9]. Для HE4 было показано, что он потенциально может быть использован в качестве маркера рака эндометрия [32].

Тест OVA1 основан на измерении содержания СА-125, транстиретина, аполипопротеина AI, трансферрина и β 2 микроглобулина. Чувствительность теста, по разным оценкам, составляет более 91%, а вот специфичность для общей группы – 43% [20]. Поэтому тест одобрен для использования в случаях, когда образование в яичниках уже выявлено, но не ясно является ли оно злокачественным. Специфичность при таком использовании может достигать 95,4% [20]. Тест ROMA (The Risk of Ovarian Malignancy Algorithm; алгоритм диагностики риска развития злокачественной опухоли в яичниках) использует данные о содержании маркеров СА-125 и HE4 и о климактерическом статусе. Тест ROMA в 2011 году получил разрешение FDA на использование для диагностики рака яичников в случае, когда образование в яичниках уже выявлено [20]. Чувствительность ROMA для таких случаев составляет 80%, а специфичность – 85,4% [33].

2. Протеомные технологии для поиска новых маркеров онкологических заболеваний

После завершения проекта генома человека у исследователей появилась надежда, что удастся открыть генетические причины многих социально-значимых заболеваний и научиться их эффективно предсказывать. Для достижения этой цели предполагалось максимально полно изучить генетические полиморфизмы человека и найти те, которые связаны с предрасположенностью к заболеваниям [34]. Эти надежды окрепли благодаря открытию ряда генов, связанных с такими опасными и распространенными заболеваниями, как рак молочной железы, рак яичников, ВИЧ-инфекция [34, 35]. Однако полученные результаты не воспроизводились: в различных исследованиях находили непересекающиеся группы связанных с болезнью SNP (Single nucleotide polymorphism – однонуклеотидный полиморфизм) [36-39]. Накопленная за годы исследований статистика позволила в 2006 году создать такой проект как “The Cancer Genome Atlas” (атлас геномов рака), описывающий все исследованные SNP при различных видах рака [40, 41].

Последние 10-20 лет начали появляться технологии, направленные на детальный анализ конкретных групп молекул. Так зародились новые направления в науке: протеомика, метаболомика, липидомика и др. Эти науки, их также принято называть «омики», направлены на развитие успеха проекта генома человека [37, 38, 42, 43].

2.1. Протеомика и пептидомика

Протеомика является областью научных исследований постгеномной эры, направленной на анализ перехода между генетической информацией и фенотипом [36, 38, 43-45]. Термин «протеом» был введен в 1995 году для описания полного набора белков, кодируемых геномом [44]. Позднее в 1997 году был предложен термин «протеомика» по аналогии с термином «геномика» (наука о генах) [36]. Само слово «протеом» произошло от сочетания слов «**протеин**» и «**геном**». Основным предметом изучения протеомики являются структуры, пост-трансляционные модификации и взаимодействия белков в живых организмах.

Задачей протеомики является идентификация и количественный анализ экспрессии белков в клетках, тканях или целом организме в зависимости от их состояния и внешних условий [45, 46]. Исследователи проводят сравнительный анализ больших групп белков: от белков, вовлеченных в тот или иной биологический процесс, до полного протеома [47]. В основе этой области исследований лежат базы данных сиквенсов геномов, масс-спектрометрия, как основной инструментальный метод анализа белков и пептидов и биоинформатика, коррелирующая результаты масс-спектрометрического анализа со структурами белков, транслированными из геномных баз данных.

Данные геномных исследований не позволяют однозначно судить о транскрибируемости генов, а наличие транскрипта не означает, что он будет транслироваться. Кроме того, транскрипт не позволяет однозначно говорить о структуре белка, его модификациях, сплайсинге и локализации [36, 45]. Для ответа на эти вопросы и необходим арсенал современной протеомики.

Традиционно изучение белков являлось одним из разделов биохимии, но после определения структуры всей геномной ДНК человека и ряда других организмов у исследователей появились новые возможности. В частности, появились исчерпывающие базы данных о структуре белков многих организмов. Это позволило, к примеру, идентифицировать белки по молекулярной массе их протеолитических фрагментов. Импульс к развитию протеомики дает совершенствование методов масс-спектрометрического анализа и биоинформатики [38].

Термин «пептидомика» был впервые предложен в феврале 2000 года группой немецких ученых под руководством Schulz-Knappe и обозначал сокращение от термина «пептидная протеомика», качественный и количественный анализ всех пептидов и коротких белков в заданной биологической системе в определенный момент времени [42, 48]. Не существует четкого деления между белками и пептидами, но большинство авторов называют пептидами полиаминокислоты с молекулярной массой до 10 кДа [39, 42, 49]. Пептиды содержатся во всех

биологических жидкостях организма человека, в клетках и тканях, где они играют важную роль в качестве гормонов, факторов роста, цитокинов или защитных агентов [39, 50]. Кроме того, биоактивные пептиды участвуют в процессах передачи сигналов в дыхательной, сердечно-сосудистой, эндокринной, воспалительной и нервной системе [51].

Помимо регуляторных пептидов существует большой пул пептидов, которые являются продуктами протеолитической деградации белков – деградом – их называют “биологическим мусором”, “биологическим шумом”, нестабильным, неспособным быть биологически значимым. Долгое время весь пептидом рассматривали исключительно с этой точки зрения [52-54]. Сейчас все больше групп начинают изучать деградом как потенциальный источник диагностической информации [55-57]. Содержащиеся в биологических жидкостях и тканях продукты деградации белков, могут отражать биологические процессы внутри организма [56]. Истинную роль пептидома ещё предстоит изучить [58].

Пептидом может содержать важную диагностическую информацию. По аминокислотной последовательности пептидов можно установить белки предшественники и вероятные протеазы, которые привели к образованию наблюдаемого пептидного пула [56, 59]. Кроме того, важную информацию несет количество того или иного пептида в организме. В циркулирующих жидкостях, как правило, пептиды находятся в ассоциированном состоянии с большими белками переносчиками. Такие комплексы, вероятно, достаточно стабильны и образуются вблизи мест непосредственного выхода пептидов в кровоток. Таким образом, место синтеза белков переносчиков является ценной диагностической информацией [56, 60].

Пептиды являются более удобным объектом исследования, чем белки. Во-первых, пептиды по сравнению с белками свободнее мигрируют из тканей, где они образовались, в кровь и другие биологические жидкости [56]. Для исследований и диагностики биологические жидкости организма, такие как сыворотка или плазма крови, доступнее, чем пораженные ткани. Во-вторых, для исследования пептидома

современными масс-спектрометрическими подходами не требуется проводить протеолиз, так как пептиды имеют подходящую для масс-спектрометрического анализа длину аминокислотной последовательности [61]. В-третьих, профилирование пептидома (методом SELDI/MALDI-MS) производительнее, чем профилирование протеома (к примеру, 2D-PAGE) [58]. Поэтому, существует мнение, что перспективнее изучать изменения на уровне пептидов, чем белков [42].

Для изучения пептидома доступно большинство хорошо отработанных протеомных методов и подходов [58]. В частности, tandemная масс-спектрометрия является сегодня превалирующим методом идентификации пептидов. Однако, идентификация эндогенных пептидов по их MS/MS спектрам является гораздо более сложной задачей, чем идентификация белков по tandemным масс-спектрам их триптических фрагментов [62]. Основная сложность идентификации эндогенных пептидов заключается в том, что, они образуются в результате расщепления белков-предшественников протеазами с заранее неизвестной специфичностью. Поэтому при поиске в базах данных белковых последовательностей нельзя ограничить выборку рассматриваемых потенциальных кандидатов специфичностью конкретной протеазы, как это делается в случае расщепления белков трипсином при протеомных исследованиях. Это может на 2-3 порядка увеличить объем вычислений и повысить требования к качеству интерпретируемых MS/MS спектров. В протеомных работах ставится задача идентифицировать белок по триптическим фрагментам [58, 62]. При этом не важно, если некоторые из пептидов слишком большие или маленькие для анализа, поскольку все равно будет достаточное количество пептидов подходящего размера и аминокислотной последовательности. Более того, достаточным для идентификации белка считается обнаружение всего двух его триптических пептидов. В пептидомных работах ставится задача охарактеризовать эндогенную форму пептида, включая его пост трансляционные модификации [62].

2.2. Масс-спектрометрия – основной инструментальный метод протеомики

Масс-спектрометрия – это физико-химический метод анализа, заключающийся в переводе молекул образца в газообразное состояние и ионизации молекул с последующим разделением и регистрацией образующихся при этом ионов. Современные модели масс-спектрометров обладают высоким разрешением, высокой чувствительностью, широким массовым и динамическим диапазоном. В зависимости от решаемых задач используется определенная модель масс-спектрометра, которая может превосходить остальные по каким-либо показателям и уступать другим моделям по прочим характеристикам [38]. Сегодня масс-спектрометрия используется для анализа различных биологических объектов, к примеру: ДНК [63], комплексов белков [64], вирусов [65]. В современных протеомных работах по поиску и идентификации потенциальных биомаркеров метод масс-спектрометрии выполняет две задачи: идентифицировать наличие пептида/белка в образце (качественный анализ) и определить концентрацию (количественный анализ) [38].

Масс-спектрометрия изначально была областью физики. Существовавшие тогда методы масс-спектрометрического исследования не позволяли ионизовать и перевести в газообразное состояние биологические молекулы целиком [66]. Эта проблема была решена в 80-ых годах XX века, когда были предложены так называемые «мягкие» методы ионизации: группы Tanaka и Karas предложили метод МАЛДИ [67, 68], а группа Fenn предложила метод ионизации электро распылением [69]. В 2002 году за разработку методов масс-спектрометрического анализа биологических макромолекул Koichi Tanaka и John Fenn вместе с Kurt Wuthrich (за работу в области ЯМР спектроскопии) получили нобелевскую премию по химии [70].

2.2.1. Методы ионизации биологических макромолекул

MALDI – Matrix-Assisted Laser Desorption/Ionization (Матрично-Активированная Лазерная Десорбция/Ионизация, МАЛДИ) – десорбционный

метод «мягкой» ионизации, обусловленной воздействием импульсами лазерного излучения на матрицу с анализируемым веществом. В качестве матрицы используются специальные вещества, свойства которых обуславливают понижение деструктивных свойств лазерного излучения, перевод в газообразное состояние и ионизацию анализируемого вещества [71]. В основном в качестве матрицы используют низкомолекулярные органические кислоты: α -Циано-4-Гидроксикоричная кислота, 2,5-Дигидроксibenзойная кислота, 2-(4-Гидроксифенилазо)-бензойная кислота, Синапиновая кислота.

Масс-спектрометры, использующие метод ионизации MALDI, являются наиболее чувствительным (способны детектировать вещества содержанием менее 1 фмоль) и позволяют детектировать в основном однозарядные ионы, что облегчает исследование сложных смесей. Кроме того, MALDI толерантен к примесям (до определенных концентраций) таким как соли или детергенты [72]. MALDI чаще всего объединяют с времяпролетными масс-анализаторами, поскольку такая комплектация наилучшим образом отвечает периодическому/импульсному режиму работы лазера [66].

ESI (Electrospray ionization) – электроспрей или ионизация распылением в электрическом поле – это метод «мягкой» ионизации для получения ионов в газовой фазе из раствора. В источник ионизации растворитель с анализируемым веществом вводится через проводящий капилляр, на который подается высокое напряжение (около 3 кВ) относительно электрода, образующего входное отверстие в масс-анализатор [73]. Растворитель рассеивается в виде аэрозоли, мелких заряженных капель, которые начинают быстро испаряться до тех пор, пока для них не достигается предела Рэлея, и они не разрываются на капли меньшего размера [73]. Процесс повторяется до полного испарения растворителя и образования ионов в газовой фазе.

Одна из основных особенностей электроспрея это образование многозарядных ионов [73]. Это позволяет анализировать белки до 10 кДа внутри относительно небольшого массового диапазона (менее 2500 m/z). Кроме того,

фрагментация многозарядных ионов приводит к образованию спектров с большим числом дочерних ионов, что увеличивает шанс достоверной идентификации.

2.2.2. Основные типы масс-анализаторов

Анализатор масс-спектрометра измеряет величину m/z ионов, разделяя их или во времени, или в пространстве. В настоящий момент для идентификации пептидов наиболее часто используются квадрупольный, времяпролетный масс-анализаторы, ионная ловушка, анализатор ионно-циклотронного резонанса и орбитальная ионная ловушка [66]. Каждый масс-анализатор характеризуется своей скоростью, разрешением, массовым диапазоном, точностью определения m/z и возможностью использования тандемной масс-спектрометрии.

Времяпролетный масс-анализатор (от английского time-of-flight, TOF) был предложен группой Stephens в 1946 году [74]. Наиболее простой вариант – это линейный времяпролетный масс-анализатор. В высоком вакууме бесполевого пространства анализатора ионы дрейфуют с постоянной скоростью, практически не испытывая столкновений. Поскольку начальная скорость ионов обратно пропорциональна квадратному корню величины m/z , более легкие ионы достигают детектор раньше более тяжелых. Разрешение времяпролетного масс-анализатора прямо пропорционально длине камеры дрейфа, однако при увеличении её длины возникают трудности с фокусировкой пучка ионов. Для увеличения разрешения времяпролетного масс анализатора на конце камеры дрейфа помещают рефлектор, «электростатическое зеркало» [75]. Отражаясь в рефлектроне, ионы регистрируются на противоположном конце анализатора, проходя суммарно удвоенную длину анализатора. Кроме того, рефлектор позволяет измерять тандемные масс-спектры, поскольку разделяет ионы с разной кинетической энергией.

В 1953 году Paul и Steinwedel описали метод удержания и анализа ионов с помощью двух- и трехмерного электростатического поля квадруполя [76, 77]. В 1989 году «за разработку метода удержания одиночных ионов» Wolfgang Paul получил Нобелевскую премию по физике совместно с Hans Georg Dehmelt.

Линейный квадрупольный масс-фильтр состоит из четырех параллельных гиперболических стержней [77]. Совокупность постоянного и переменного электрического поля позволяет поддерживать и фокусировать поток ионов в заданном диапазоне m/z . Изменяя электрическое поле, сканируют непрерывный поток ионов. Квадруполи часто используют в качестве «ионных ворот», «ионных проводников» и столкновительных ячеек для фрагментации [78]. Линейный квадруполь анализирует непрерывный поток ионов, поэтому его обычно используют совместно с источником ионизации ESI. Минусом прибора является узкий массовый диапазон (примерно до 4000 Да) и точность определения m/z иона (ошибка 0,2 – 0,3 Да) [79].

Работа ионной ловушки основана на тех же физических принципах, что и работа квадруполя, но в ловушке ионы накапливаются и, затем, последовательно измеряются [76, 77]. Для анализа ионов используют либо режим «нестабильности» (от англ. *instability scan method*), либо режим «резонансного возбуждения» (от англ. *resonance excitation method*) [80]. Высокая скорость ионной ловушки хорошо сочетается с методом ионизации ESI. Ионная ловушка обычно используется для получения тандемных масс-спектров. Для этого в ловушке удерживаются только ионы заданного диапазона m/z , методом CID или ETD проводят фрагментацию ионов и анализируют фрагменты. Ионная ловушка позволяет повторять процесс несколько раз, анализируя фрагменты фрагментов [80].

Comisarow и Marshall предложили масс-анализатор ионно-циклотронного резонанса (ИЦР) в 1970 году. ИЦР на сегодня имеет самые высокие характеристики по точности определения m/z и разрешению [81]. Ионы анализируемого вещества прецессируют в магнитном поле анализатора. Частота прецессии, зависит от m/z этих ионов. Для регистрации частоты прецессии измеряют наведенный ток. Суммарный наведенный ток раскладывается с помощью преобразований Фурье до частот прецессий [82]. ИЦР позволяет последовательно проводить несколько стадий фрагментации. Кроме того, конструкция анализатора позволяет использовать для фрагментации различные методы: CID, ETD и прочие.

Орбитальная ионная ловушка, является аналогом ИЦР, однако вместо магнитного поля для удержания ионов используется электрическое поле [83]. Особенностью прибора является веретенообразная форма электрода, которая позволяет создать поле нужной конфигурации. Так же, как и ИЦР, орбитальная ионная ловушка обладает высокой точностью определения m/z , высоким разрешением и позволяет проводить фрагментацию различными методами. Преимуществом ловушки над ИЦР является более высокая чувствительность и более низкая стоимость [83].

Каждый масс анализатор обладает особенностями, которые определяют область его применения. Комбинация из нескольких последовательно расположенных масс анализаторов позволяет использовать сильные стороны каждого из них на нужном этапе анализа. Так комбинация последовательно расположенных двух квадрупольных и одного времяпролетного масс-анализатора нашла широкое применение для идентификации, а так же количественного анализа сложных белковых смесей [66]. Сильными сторонами таких приборов является их способность точно отбирать ионы для фрагментации, проводить быстрый анализ исследуемой смеси и получать тандемные масс-спектры с высокой точностью определения m/z родительских и дочерних ионов.

2.2.3. Тандемная масс-спектрометрия

Деградация по Эдману традиционно применяется для определения аминокислотной последовательности пептидов и белков. Масс-спектрометрия способна определять структуры пептидов по их спектрам фрагментации, используя меньшие количества образца [66].

Для получения спектров фрагментации, или MS/MS спектров, были предложены различные методы. Наиболее часто используемый метод фрагментации это CID/CAD (от английского, Collision Induced/Activated Dissociation), основанный на столкновении ионов анализируемого вещества с инертным газом (к примеру: Ar, He, Ne или N₂) [84]. Различают метод CID при низких (10 – 100 эВ) и высоких (более 1 кэВ) энергиях молекул, в зависимости от

значения кинетической энергии, до которой разгоняются ионы анализируемого вещества до соударения. При столкновении часть кинетической энергии переходит во внутреннюю энергию молекул, что приводит к разрыву связей и образованию фрагментов. Заряд может остаться на С- или N-концевом фрагменте. Разработана номенклатура ионов фрагментов в зависимости от места разрыва полипептидной цепи и фрагмента, на котором остался заряд [85, 86].

Помимо CID/CAD были предложены и другие методы фрагментации. Все они основаны на повышении внутренней энергии ионов исследуемого вещества. Дополнительная энергия может быть передана посредством инфракрасного излучения (IRMPD, infrared multiphoton dissociation) [87], лазерного излучения (LID, laser-induced dissociation), захвата электрона (ECD, electron capture dissociation) [88], отделения электрона (EDD, electron-detachment dissociation) [89], переноса электрона (ETD, electron transfer dissociation) [90] и прочее. Многие методы комплементарны друг другу, что позволяет получать более представленные спектры фрагментации при использовании нескольких методов. Так ионная ловушка и ИЦР позволяют параллельно использовать методы CID и ETD, что увеличивает эффективность и достоверность идентификаций [91].

2.2.4. LC-MS/MS – tandemная масс-спектрометрия, сопряженная с высокоэффективной жидкостной хроматографией

Прямой масс-спектрометрический анализ сложной пептидной смеси затруднен по ряду причин [72, 92, 93]. Хроматографическое разделение пептидов повышает эффективность анализа и позволяет получить большее количество идентификаций. Метод ионизации ESI позволяет напрямую соединить масс-спектрометр с хроматографом, что увеличивает чувствительность и динамический диапазон анализа. Однако это сопряжение потребовало ряд дополнительных доработок: уменьшение диаметра используемых хроматографических колонок и миниатюризация всей хроматографической системы [94]. При потоках жидкости в 100-200 нл/мин удалось достичь 100% эффективности ионизации

электрораспылением, так что интенсивность сигнала стала пропорциональна содержанию анализируемого компонента в образце [95].

При LC-MALDI анализе отдельные пептидные фракции наносятся на масс-спектрометрическую мишень [96]. Достоинством данного подхода является отсутствие временных ограничений для проведения фрагментации, что позволяет глубже проанализировать смесь, а при необходимости повторить или дополнить анализ позднее. Однако измерение каждого компонента смеси требует относительно большого времени и для анализа сложных смесей такой подход не используется.

2.2.5. Методы количественного масс-спектрометрического анализа

Одной из задач протеомики является количественная оценка изменений уровня экспрессии белков в клетках, тканях или целом организме при воздействии на них различных внешних факторов [45, 46]. Из-за технических ограничений масс-спектрометрия, основной инструментальный подход современной протеомики, не может быть использована для непосредственного сравнения содержания разных компонент в одном образце [97]. В то же время, сравнение содержания одного и того же компонента в разных образцах может быть успешно проведено методами масс-спектрометрии. Использование стандарта известной концентрации в качестве одного из образцов, позволяет перейти от относительного количественного измерения к абсолютному.

Для реализации этой задачи первоначально были разработаны методы количественной протеомики с использованием стабильных тяжелых изотопов кислорода, углерода, азота и дейтерия [97]. Эти тяжелые атомы включаются в состав пептидов или белков химически после выделения (к примеру, ICAT (Isotope-Coded Affinity Tags – изотопно-меченые аффинные метки) [98], iTRAQ (isobaric Tag for Relative and Absolute Quantitation – изотопные метки для относительного и абсолютного количественного анализа) [99]) или естественным путем в процессе развития культуры или организма (метод SILAC (Stable Isotope Labeling by Amino acids in Cell culture – мечение клеточной культуры аминокислотами со

стабильными изотопами) [100-102]), либо замещают легкие аналоги (захват атомов кислорода ^{18}O из водного раствора в процессе гидролиза белка [103]), и на масс-спектре появляются дополнительные компоненты, соответствующие m/z более тяжелых пептидов. Таким образом, помечают пептиды одного из сравниваемых образцов, смешивают образцы и, сравнивая суммарный ионный сигнал от легкого и тяжелого пептида, получают информацию об относительном изменении содержания пептида между образцами. Методы количественного анализа с использованием стабильных изотопов имеют ряд технических недостатков, связанных с ограниченной специфичностью меток, необходимостью учитывать наложение сигналов меченных пептидов на естественное изотопное распределение и ограничением числа одновременно сравниваемых образцов [97].

Количественные безметочные масс-спектрометрические подходы используют зависимость между содержанием измеряемого пептида и уровнем масс-спектрометрического сигнала. По принципу работы масс-спектрометра безметочные подходы можно разделить на DDA, DIA (Data-dependent/independent acquisitions – результат-зависимые/независимые измерения) [104, 105] и направленные протеомные подходы (SRM или MRM – Single/Multiple Reaction Monitoring [106, 107], измерение одной/нескольких реакций).

При DDA процесс работы масс-спектрометра представляет собой циклы, включающие этап измерения общего спектра всех компонент исследуемой смеси, отбор ионов прекурсоров для фрагментации на основании интенсивностей и последовательное проведение MS/MS анализа выбранных компонент [104]. При этом параллельно с информацией об интенсивности родительских ионов собираются MS/MS данные для идентификации пептидов. В качестве количественной оценки могут быть использованы суммарный ионный ток как самого иона прекурсора, так и дочерних ионов из MS^2 спектра, или даже количество tandemных спектров, измеренных для конкретного пептида. Основной проблемой DDA на первых порах было недостаточное количество учитываемых точек за время элюции пептида с колонки, которое определяется временем DDA

цикла. На современных быстрых масс-спектрометрах данная проблема практически отсутствует. Сегодня основная сложность это корректный анализ спектров (выравнивание, устранение шума) и экстрагирование хроматограмм индивидуальных компонент [93].

Направленные протеомные подходы (SRM и MRM) используются для количественного измерения ограниченный набор пептидов, которые должны быть известны ещё до начала измерения [106]. Данные подходы избавлены от ошибок квантификации, связанных с возможным перекрытием ионов прекурсоров (как по m/z , так и по времени хроматографического выхода), поскольку используют информацию о дочерних ионах из MS^2 спектра. Минусом подходов SRM и MRM является ограниченное количество компонент, которые удастся измерить за один масс-спектрометрический эксперимент [105].

При DIA после измерения общего спектра всех компонент исследуемой смеси, не происходит отбор единичных родительских ионов на MS/MS анализа, а проводится либо тотальная фрагментация всех компонент смеси (к примеру, методом MS^E) [108], либо ионы прекурсоры пропускаются большими окнами (к примеру, в 25 Да в методе SWATH, Sequential Window Acquisition of all Theoretical fragment-ion spectra – измерение спектров всех теоретических ионов-фрагментов при использовании последовательных окон пропускания) [105]. После проведения LC-MS/MS эксперимента с помощью специальных программ проводится соотнесение между родительскими и дочерними ионами. DIA предъявляет чрезвычайно высокие требования к точности и скорости работы прибора и реализован только на самых современных машинах.

2.3. Биоинформатические подходы в протеомике

Масс-спектрометрические методы анализа превратили протеомику в высокопроизводительную область исследования. Огромное количество накопленных масс-спектрометрических данных, требует адекватного и тщательного анализа, который в настоящее время считается основным сдерживающим фактором развития протеомики [93].

2.3.1 Построение математических моделей классификации

Масс-спектрометрическое профилирование образцов биологических жидкостей (сыворотки, плазмы и пр.) позволяет обнаруживать различия между группами доноров, у которых забирали исследуемый биологический материал. Биоинформатический анализ способен выявить те компоненты спектра, которые статистически наилучшим образом отличают экспериментальные группы. Выбранные масс-спектрометрические пики, а также их весовые коэффициенты составляют математическую модель классификации, которая может быть использована в дальнейшем для диагностирования новых пациентов.

Для построения математических моделей классификации метод простого перебора всех возможных комбинаций детектируемых масс-спектрометрических пиков не подходит, поскольку число возможных комбинаций велико. Существует ряд математических алгоритмов, позволяющих определить оптимальную комбинацию масс-спектрометрических пиков, к примеру: алгоритм Быстрой Классификации, Генетический Алгоритм (ГА), Обучаемая Нейронная Сеть (ОНС), алгоритм Опорных Векторов [109].

Концепция ГА была разработана Holland [110]. Принцип алгоритма базируется на идее эволюции, в которой наиболее приспособленные индивидуумы имеют наибольший шанс для выживания. Отбор осуществляется при помощи оптимизации функции стоимости, направленной на наибольшее расхождение между сравниваемыми классами. Процедура скрещивания объединяет случайно выбранные комбинации пиков с образованием дочерних комбинаций, которые замещают родительские. Для предупреждения преждевременного схождения эволюционирующих комбинаций в локальные оптимумы проводят случайные замены пиков. ГА позволяет получить лишь оптимальное решение и не может гарантировать нахождение наилучшей комбинации пиков, так как не были просчитаны все комбинации. Преимущество ГА заключается в малых затратах вычислительного времени, обеспечивающего, достаточно хороший результат.

Алгоритм ОНС представляет собой модификацию алгоритма Hammer [111]. Принцип работы основан на определении оптимальной границы, разделяющей два множества данных. При сравнении набора спектров исследуемых групп пациентов, ОНС строит характеристические спектры для сравниваемых групп – прототипы. При получении спектра для диагностики модель соотносит новый спектр с прототипами исходных классов. Преимуществом ОНС является производительность для больших массивов данных.

2.3.2 Идентификация пептидов и валидация результатов идентификации

Интерпретация результатов MS/MS анализа комплексной пептидной смеси является сложной вычислительной задачей [66, 112]. Тандемный масс-спектр представляет собой запись фрагментов полипептидной цепи, при этом, часть компонент отсутствует, а часть по интенсивности лишь незначительно перекрывает шум. Кроме того, современные LC-MS/MS масс-спектрометры способны проводить несколько десятков MS/MS в секунду – в результате необходимо проанализировать несколько сотен тысяч спектров фрагментации от одного эксперимента.

Для идентификации пептидов по MS/MS спектрам подходят те же программные продукты, что были разработаны для идентификации белков [58, 66]. Однако, при оценке достоверности идентификации, а также при валидации результатов идентификации не следует оценивать достоверность идентификации белков или использовать критерии, связанные с обнаружением нескольких пептидов, относящихся к одному белку. Эти подходы не работают, так как факт наличия целого белка в исходной смеси находится под сомнением. Можно выделить два основных решения для идентификации пептидов по MS/MS спектрам [66, 113, 114].

Наиболее популярный подход – это использование баз данных (БД) всех белковых последовательностей исследуемого организма [112]. Белки из БД виртуально щепятся до пептидов. Затем, спектр фрагментации сопоставляется с теми из полученных пептидов, для которых величина m/z равна измеренной для

иона прекурсора, в пределах известной ошибки масс-спектрометра. Для оценки корреляции пептида со спектром фрагментации различные программы используют разные алгоритмы и критерии. Идентифицируют пептид с наибольшей корреляцией. Данный подход может приводить к ложно положительным идентификациям и чувствителен к количеству потенциально возможных пептидов, сопоставляемых с каждым MS/MS спектром. Поскольку в образовании пептидома участвовали протеазы различной специфичности, точно указать правило щепления белков невозможно. Чтобы скомпенсировать возрастающее число потенциально возможных пептидов, накладываются дополнительные требования на точность масс-спектрометров. Для реализации данного подхода важно грамотно выбрать БД белковых структур и подобрать набор потенциально возможных пост-трансляционных модификаций, чтобы эффективно проверить все возможные варианты и огородить себя от ложно положительных идентификаций. Предложено несколько десятков поисковых программ, и каждый год появляются новые [112]. Наиболее известными являются Mascot [115], SEQUEST [116], X!Tandem [117]. Кроме поиска по БД белковых структур предложены программы, осуществляющие поиск по спектральным библиотекам, реальных спектров, которые предварительно были проанализированы и интерпретированы. Данный подход использует информацию об относительной интенсивности пиков MS/MS спектров, но ограничен небольшим размером созданных спектральных библиотек. Наиболее популярными реализациями подхода являются SpectraST [118], NIST MS [119], LIBQUEST [120], X!Hunter [121].

Метод *de novo* секвенирования позволяет восстановить аминокислотную последовательность пептида по расстоянию между соседними пиками в MS/MS спектре. Подход требует точного определения не только m/z прекурсора, но и дочерних пиков. Он используется, когда нет подходящей БД (к примеру, для организмов, у которых не расшифрован геном или для идентификации новой изоформы белка). Во многих случаях не удастся обнаружить на спектре переходы для каждой аминокислоты в сиквенсе пептида, но хорошо просматривается часть

сиквенса – тег сиквенса. Поиск таких тегов по БД белковых последовательностей производительнее оригинального *de novo* секвенирования. Если по БД не удастся обнаружить пептид, содержащий тег и коррелирующий с остальной частью MS/MS спектра, есть возможность идентифицировать высоко гомологичный пептид. Кроме того, некоторые поисковые программы позволяют анализировать возможные мутации и SNP [114]. PEAKS [122], PepNovo [123] и Lutefisk [124] наиболее часто используемые программы для *de novo* секвенирования.

Каждый метод идентификации может давать ложно положительные результаты. Для уменьшения количества ошибок при идентификации и для повышения количества истинных идентификаций применяют согласованный подход, одновременно используя несколько поисковых программ/алгоритмов [114]. Поскольку каждый алгоритм использует свои критерии корреляции аминокислотной последовательности и MS/MS спектра, необходимо приводить все результаты к универсальной оценке корреляции [125]. Согласованный подход реализован в таких программных продуктах как Scaffold и Trans Proteomic Pipeline (TPP) [126, 127]. Кроме того, данные программы позволяют валидировать результаты идентификации, оценивая количество ложно положительных идентификаций при помощи поиска по БД последовательностей несуществующих белков, а так же при помощи алгоритма PeptideProphet [128, 129].

2.4 Кровь, как объект биомаркерных исследований

Чтобы максимально эффективно получать ответы на интересующие медицинские вопросы, нужно, прежде всего, правильно выбрать объект исследования: биологическая жидкость, ткань и т.д. В данном вопросе существует много критериев для оценки: доступность материала в медицинских центрах, простота получения, стабильность после получения, состав образца, близость к очагу заболевания и прочее [130]. Кроме того, следует сразу учесть факторы, которые будут иметь значение при потенциальной диагностике: инвазивность метода получения, этические вопросы и прочее.

2.4.1 Плазма и сыворотка крови

В работах, посвященных поиску потенциальных биомаркеров, используются различные биологические жидкости (плазма и сыворотка крови, моча, слюна, спинномозговая жидкость, бронхоальвеолярная промывная жидкость, слеза, амниотическая жидкость и прочие), а также тканевые или клеточные экстракты [131-145]. Наиболее часто в подобных работах исследуют плазму, сыворотку крови, мочу или слюну. Прочие биологические жидкости выбирают, как правило, при изучении локализованных заболеваний. Их использование при исследовании системных заболеваний не рекомендовано, ввиду сложности последующего использования в потенциальной диагностике, а также риска и дискомфорта для пациента.

В работах разных групп исследователей подробно рассматриваются преимущества и недостатки каждой из перечисленных биологических жидкостей как объектов биомаркерных исследований, изучается влияние различных стадий пробоподготовки на последующие результаты, обсуждаются вопросы стандартизации протоколов забора и работы с этими жидкостями [49, 54, 146-149]. Условия получения, хранения и работы с плазмой и сывороткой крови разработаны в рамках инициатив международной организации протеома человека (HUPO, Human Proteome Organization) [137].

HUPO рекомендует предпочтительно использовать плазму при анализе циркулирующих в крови пептидов и низкомолекулярных белков [49, 150]. Считается, что плазма ближе по составу к крови, чем сыворотка, поскольку при получении плазмы ингибируют каскад коагуляции [49, 137]. Согласно протоколу получения плазмы крови рекомендуется немедленное добавление EDTA к крови, после отбора [137, 151]. Поскольку пептидный и белковый состав плазмы изменчив, следует как можно быстрее замораживать аликвоты плазмы крови до -80°C [137].

Преимуществом использования сыворотки крови является относительная простота и низкая инвазивность получения. Сыворотка является одним из наиболее

часто архивируемых биологических объектов. Использование этих готовых коллекций при проработанной методике пробоподготовки позволяет проводить на основе сыворотки крови биомаркерные исследования за относительно короткие сроки [137]. Кроме того, пептидный и белковый составы сыворотки после отделения от форменных тел крови считаются стабильными [54, 150]. К недостаткам сыворотки крови как объекта исследования относят процесс образование пептидов при свертывании крови, который зависит от многих трудно контролируемых факторов [49, 137]. Известно, что к основным факторам, влияющим на пептидный состав сыворотки крови, относятся тип используемых при заборе крови вакутейнеров, уровень гемолиза эритроцитов и время инкубации крови до отделения сыворотки от сгустка форменных элементов центрифугированием [54, 146]. Тем не менее, большинство исследователей продолжает использовать в своих работах сыворотку крови. Рекомендуемые протоколы получения сыворотки крови включают инкубацию при комнатной температуре в течение 60 минут [146]. После формирования сгустка и до его отделения от сыворотки центрифугированием возможна инкубация или транспортировка на льду длительностью до 3 часов [54]. Аликвоты сыворотки крови рекомендуют хранить при -80°C . Среди исследователей нет единого мнения по вопросу сохраняется ли стабильность белкового профиля сыворотки крови при многократном повторении циклов заморозки и последующей разморозки образцов [54, 146].

2.4.2. Гипотеза образования пептидных биомаркеров

Термин «биомаркер» был определен следующим образом: параметр, который может быть измерен, чтобы отличить нормальный биологический процесс от патологического процесса или определить фармакологический ответ на введение лекарства (NIH Biomarker Definition Working Group) [152].

Опухоль следует рассматривать в условиях её микроокружения [56, 60]. Взаимодействия опухолевых клеток со здоровыми эпителиальными и стромальными клетками, с сосудами капилляров и с иммунной системой

регулируются различными ферментами, цитокинами, молекулами внеклеточного матрикса и факторами роста [60]. Все это верно и для клеток здоровой ткани. Но микроокружение опухоли реагирует на процесс постоянного клеточного деления и роста, клеточной инвазии и измененной функции иммунной системы, оказывает постоянное воздействие на опухолевые клетки через действие ферментов (например, киназ или фосфатаз) и активность протеиназ (например, матриксных металлопротеиназ). Это порождает несбалансированный или измененный состав молекул внутри опухоли по сравнению с нормальной тканью [56].

Гипотеза происхождения пептидома объясняет, как многие белки и пептиды из микроокружения опухоли попадают в кровоток, в том числе функционально активные белки, такие как протеазы (Рисунок 1). Если белок слишком большой, чтобы проникнуть через стенку кровеносного капилляра, в кровотоке он все равно будет представлен своими протеолитическими фрагментами. Кроме того, повышенное гидростатическое давление вблизи постоянно растущей опухоли облегчает выход молекул из микроокружения опухоли в кровоток. Таким образом, кровь содержит в себе, в качестве маркеров утечки, все белки, или хотя бы их фрагменты, которые в данный момент продуцируются микроокружением опухоли. Комбинация пептидных маркеров, которые отражают особенности взаимодействия опухоли и её микроокружения могут позволить создать эффективный диагностикум [56].

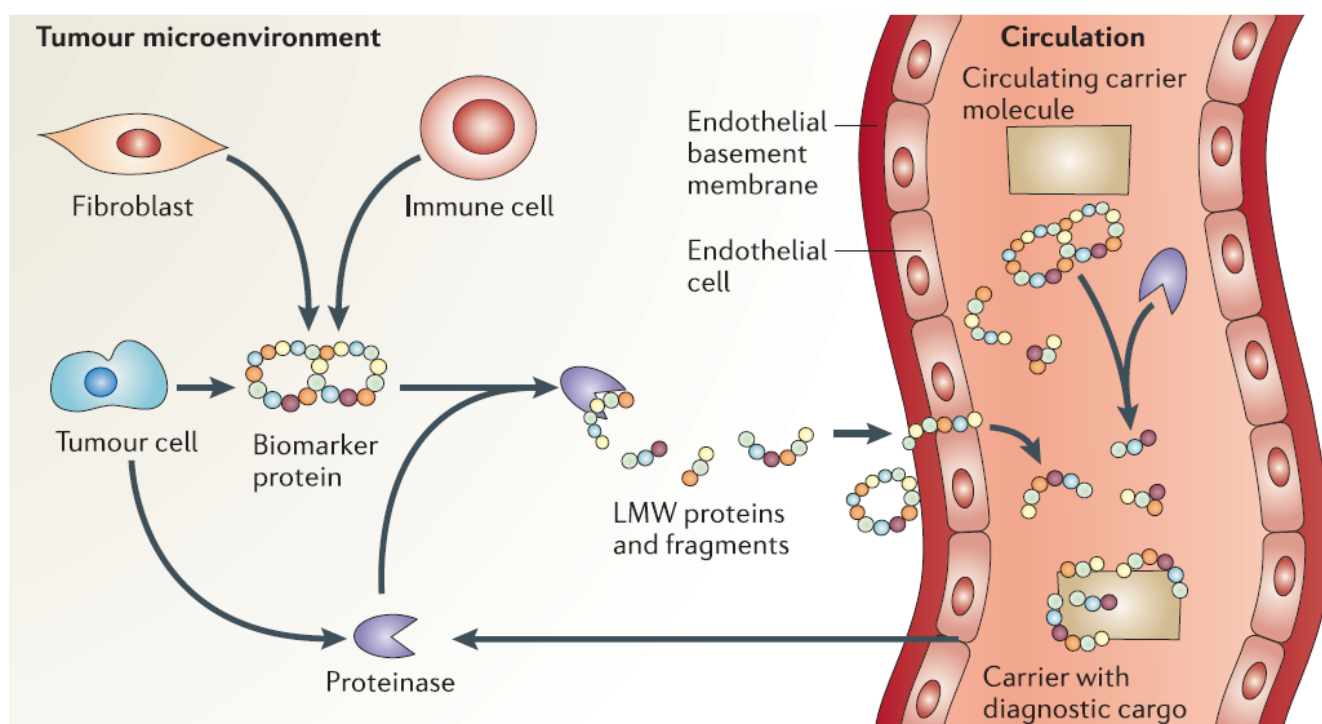


Рисунок 1: Гипотеза происхождения пептидома. Пептиды и фрагменты белков, попадающие в тканевое микроокружение из различных типов клеток. Протеолитические ферменты в тканях создают фрагменты, которые попадают в циркулирующие биологические жидкости. Идентификация пептидов и паттерна их протеолитического щепления дает два вида диагностической информации. Рисунок взят из работы Petricoin и соавторов [56].

Описанная выше гипотеза происхождения «истинных» биомаркеров основывается на предположении, что сама опухоль имеет достаточные размеры, чтобы продуцируемые ими маркеры могли существенно перекрыть уровень индивидуальной изменчивости пептидома человека. Поскольку до сих пор не было обнаружено истинно опухолевых биомаркеров, можно предположить, что концентрация подобного рода маркеров крайне низкая и/или они выводятся из организма [56]. Работы Villanueva и соавторов по поиску потенциальных биомаркеров в сыворотке и плазме крови показали, что найденные характерные пептиды, соответствуют фрагментам основных белков/пептидов крови (фибриноген α , комплемент C3f, брадикинин) [59]. Однако содержание этих фрагментов в крови онкологических больных значительно отличается от среднего значения для здоровых доноров. Согласно новой выдвинутой гипотезе

биомаркеры, которые удалось идентифицировать, не являются высокоспецифичными молекулами, синтез которых связан непосредственно с развитием опухоли. В процессе трансформации клетки злокачественной опухоли претерпевают существенные изменения в строении и выполняемых функциях, что находит отражение в экспрессии специфических ферментов. Попадая в межклеточное окружение или кровоток, ферменты сохраняют активность и расщепляют белки, которые в норме содержатся в этих жидкостях. При этом высокая активность ферментов позволяет им создать концентрацию специфических фрагментов большую, чем концентрация самих ферментов. Эта гипотеза «суррогатных маркеров» очень хорошо согласуется с последними работами в области идентификации потенциальных биомаркеров [59].

2.4.3. Основные сложности протеомных и пептидомных исследований производных крови

Кровь и её производные (плазма и сыворотка) используются как основной объект для исследований во многом потому, что кровь является соединительной тканью, которая контактирует практически со всеми клетками организма. Кровь содержит следы метаболической активности, в том числе маркеры изменений в метаболизме, вызванные патологическими процессами [56]. Ещё одним аргументом в пользу использования крови в качестве основного объекта исследования является техническая простота и низкий уровень инвазивности получения и подготовки плазмы или сыворотки крови [146, 150].

Помимо вышеперечисленных преимуществ имеются существенные недостатки использования крови в качестве объекта для протеомных исследований. Во-первых, кровь имеет очень сложный состав, теоретически включающий в себя в следовых количествах все белки организма [53]. Оценивается, что различных гликоформ и сплайс вариантов только белков плазмы крови может насчитываться до 40 тысяч [92]. Кровь содержит в себе белки многих других тканей организма. С учетом различных сплайс вариантов и посттрансляционных модификаций общее количество белков в плазме крови может быть около 500 тысяч. А если учесть ещё

и разнообразные формы иммуноглобулинов, которые присутствуют в крови, то оцениваемое число белков плазмы крови превышает 10 млн [92].

Во-вторых, кровь содержит белки в очень широком динамическом диапазоне концентраций [53], который намного превосходит аналитические возможности современных масс-спектрометрических методов. Динамический диапазон содержания белков в плазме крови, измеряемых на сегодняшний день иммуноферментными методами, составляет 11 порядков. Аналитические возможности современных протеомных подходов оцениваются в 4-6 порядков. Для исследования потенциальных маркеров необходимы подходы, имеющие динамический диапазон концентраций 5-9 порядков [53].

В то же время, основное количество белков крови приходится на несколько десятков белков. Известно, что на 10 наиболее представленных белков приходится 90% от общего количества белка по массе, а на 20 наиболее представленных белков – 98% от общего количества белка по массе [92]. Основные белки плазмы крови могут скрывать сигнал от многих низко представленных белков и пептидов.

Так же следует упомянуть, что пептидный состав плазмы и сыворотки крови существенно различаются. Так более чем 40% пептидов, детектированных в сыворотке крови, не обнаруживаются в плазме [49, 137]. Эти данные не позволяют использовать сыворотку и плазму крови совместно для сравнения обнаруженных пептидов.

Ещё одним важным с точки зрения масс-спектрометрии недостатком является высокое содержание в крови примесей: солей и липидов. Эти примеси уменьшают эффективность ионизации исследуемых молекул и чувствительность, а при высоком содержании могут и вовсе привести к отсутствию целевого сигнала [72].

2.4.4. Методы выделения пептидов из плазмы/сыворотки крови

Первым шагом при анализе пептидного пула является его выделение из сложной многокомпонентной смеси, такой как сыворотка или плазма крови [153]. Ключевой задачей является устранение из смеси высоко представленных белков и

обогащение состава низко представленных пептидов. Каждый метод имеет свои достоинства. Многие методы могут быть использованы последовательно [154].

2.4.4.1. Ультрафильтрация

Главное отличие пептидов от белков, с точки зрения физических свойств, это их молекулярный вес. Ультрафильтрация с использованием пористых мембран с отсечкой по молекулярной массе это наиболее простой и широко распространенный метод отделения пептидов от высокомолекулярных белков. Zheng с соавторами использовали фильтры с мембраной на 10 кДа, чтобы выделить пептиды сыворотки крови для прямого анализа на масс-спектрометре с линейной ионной ловушкой и Фурье преобразованием (LTQ-FT MS). Им удалось идентифицировать, с точностью определения массы иона прекурсора 2 ppm, более чем 300 уникальных пептидов с каждого образца [62]. Поскольку концентрация пептидов в биологических жидкостях низкая, для концентрирования приходится использовать большие объемы жидкости, что ведет к существенному увеличению времени фильтрации. Кроме того, другие низкомолекулярные компоненты, такие как соли, так же концентрируются в результате ультрафильтрации.

Ну с соавторами использовали ультрафильтрацию до 10 кДа с последующей гель-фильтрацией [155]. Из низкомолекулярной фракции им удалось идентифицировать 1181 уникальный пептид, относящийся к 371 белку предшественнику. Прямой LC-MS/MS анализ высокомолекулярной пептидной фракции привел к идентификации только 25 пептидов. После трипсинолиза высокомолекулярной фракции им удалось идентифицировать 250 белков по их триптическим фрагментам.

2.4.4.2. Преципитация белков органическими растворителями

Органический растворитель уменьшает диэлектрическую константу раствора белков, что приводит к электростатическому притяжению заряженных молекул. Гидрофобные взаимодействия белков нарушаются в присутствии органического растворителя, а электростатические взаимодействия становятся главными и

приводят к агрегации и преципитации белков. В это же время низкомолекулярные белки и пептиды остаются в растворе.

Alpert с соавторами впервые использовали органический растворитель для удаления высокомолекулярных белков [156]. Они показали, что белки массой более 20 кДа преципитируют при добавлении двух объемов ацетонитрила. В своих экспериментах они наблюдали агрегацию шести наиболее представленных белков, на которые приходилось более 90% от общего количества белка по массе. Кроме того, этот метод не допускает потери пептидов, связанные с неспецифической сорбцией на мембране как в случае с ультрафильтрацией.

В дальнейшем было показано, что энергичное перемешивание с постепенным увеличением концентрации органического растворителя, увеличивает воспроизводимость и эффективность выделения пептидов и очистки от белков. Это позволило использовать данный подход для сравнительного анализа экспрессии белков в сыворотке крови пациентов с раком поджелудочной железы и здоровых доноров [157]. Chertov и соавторы обнаружили, что добавление к ацетонитрилу 0,1% трифторуксусной кислоты способствует диссоциации комплексов пептидов и основных белковых плазмы крови [158]. Это повышает эффективность экстракции пептидов. Кроме того, кислотные группы образуют нерастворимые соли с положительно заряженными аминокислотными группами белков при pH ниже изоэлектрической точки.

2.4.4.3. Твердофазная экстракция

Твердофазная экстракция широко используется для очистки образцов перед анализом. Villanueva с соавторами отработал методику MALDI масс-спектрометрического профилирования пептидов сыворотки крови на основе автоматизированной твердофазной экстракции на базе магнитных частиц [147]. В своей работе они сравнили эффективность выделения пептидов с использованием C1, C2, C3, C8 и C18 гидрофобных функциональных групп, и обнаружили, что при использовании магнитных частиц с группами C8 масс-спектрометрически детектируется наибольшее количество пептидных компонент. Baumann и соавторы

стандартизировали протокол для повышения воспроизводимости MALDI масс-спектрометрических профилей сыворотки крови. Они отработали протоколы для магнитных частиц с гидрофобной, катионообменной и метало аффинной поверхностью [149].

Помимо функционализированных магнитных частиц обычные хроматографические колонки так же используются для выделения и фракционирования пептидов. Группой Yates [159] был впервые предложен метод фракционирования, основанный на последовательном объединении сильно катионообменной и обращено фазовой хроматографии. После сорбции пептидов на катионообменнике, отдельные пептидные фракции элюируются на обращено фазовую колонку при помощи последовательных ступенек концентрации соли. Затем, после промывки от солей, с обращено фазовой колонки пептиды элюируются для анализа в масс-спектрометр. Преимуществом данного метода является то, что нет необходимости в отдельном механизме для совмещения колонок. Это снижает используемые на каждой стадии объемы растворов для элюции и повышает концентрации элюируемых компонентов. Двумерная хроматография стала популярным методом фракционирования в протеомике и пептидомике. Данный подход был использован в работе по изучению пептидома плазмы крови человека: группе Tian удалось идентифицировать 988 уникальных пептидов [160].

Для увеличения эффективности двумерной хроматографии при сохранении ортогональности подходов разделения были предложены обращено фазовые сорбенты стабильные при высоких pH. В отличие от многомерной хроматографии при новом подходе первая стадия разделения при высоком значении pH проводится отдельно, в режиме офлайн [161]. Ступеньки элюции, затем, анализируются стандартным одномерным LC-MS/MS методом при помощи обращено фазовой хроматографии при низком pH. Ортогональность методов разделения достигается за счет изменения заряда кислых и основных аминокислот при разных pH. Группе

Li при помощи данного подхода удалось идентифицировать 56 нейропептидов крысы, включая 17 ранее необнаруженных пептидов [162].

Благодаря высокой специфичности своей поверхности, наноматериалы различных сорбционных свойств так же используются для очистки пептидов. К примеру, Li с соавторами провели анализ эндогенных пептидов плазмы крови, используя многослойные углеродные нанотрубки [163]. Им удалось идентифицировать с высокой достоверностью 374 уникальных пептида при помощи 2D LC-MS/MS подхода. Позднее были разработаны нанопористые поверхности и частицы на основе диоксида кремния. Группе Tian удалось использовать высокоупорядоченные мезопористые материалы, что повысило селективность выделения пептидов из плазмы крови и очистку от белков по механизму напоминающему ультрафильтрацию [160]. В своей работе они сравнили мезопористые частицы с размером пор 2, 8 и 12 нм и обнаружили, что материал MCM-41 с размером пор 2 нм наиболее эффективен для выделения пептидов в диапазоне от 1 до 12 кДа и очистки от остальных белков плазмы крови с массой более 12 кДа. Данный способ выделения пептидов объединяет в себе адсорбционные методы и ультрафильтрацию. Для более полного изучения пептидома было предложено химически модифицировать нанопористые материалы закрепляя на них катионо- и анионообменные группы. Одновременное использование сорбентов с катионо- или анионообменными группами, а так же с не модифицированными поверхностями (их специфичности близка к гидрофобной обращенной фазе) увеличивает покрытие исследуемого пептидома [164].

Сегодня научные группы все чаще исследуют пептиды, обладающие пост трансляционными модификациями. Как правило, для выделения модифицированных пептидов используют сорбенты с иммобилизованными металл аффинными хроматографическими поверхностями. В своей работе Ни и соавторы использовали материал MCM-41 модифицированный функционализированными группами фосфоната титана. Это позволило им одновременно очистить пептиды от белков и селективно выделить фосфопептиды для сравнительного MALDI-TOF

профилирования сыворотки крови пациентов с гепатоцеллюлярной карциномой и здоровых доноров [165].

2.4.4.4. Выделение пептидов на хроматографических сорбентах с ограниченным доступом к сорбционным центрам

Все вышеперечисленные методы выделения и фракционирования пептидов используются обычно в качестве отдельной стадии пробоподготовки. Для того, чтобы иметь возможность напрямую масс-спектрометрически анализировать сложные смеси, был предложен специальный материал, толерантный к многократному введению биологических жидкостей [166]. Материал называли RAM (restricted-access material), что можно перевести как материал с ограниченным доступом к сорбционным центрам. Внешняя поверхность частиц материала не обладает сорбционными свойствами. Внутри частиц расположены функционально активные группы, но доступ во внутреннее пространство частиц ограничен размерами пор. RAM сорбент выполняет сразу две функции во время хроматографии. Во-первых, хроматографическая колонка с RAM сорбентом беспрепятственно пропускает высокомолекулярные соединения, такие как белки. Эти вещества смываются с колонки во время нанесения и не мешают дальнейшему анализу. Во-вторых, на внутренней поверхности частиц проходит непосредственно хроматографирование пептидного пула образцов в зависимости от сорбционной специфичности активных центров. Выпускают RAM сорбенты с обращено фазовой, ионнообменной или аффинной поверхностью.

Ну с соавторами использовали колонку с RAM сорбентом для идентификации пептидов методом LC-MS/MS [167]. RAM сорбент, который они использовали, обладал сильно катионо обменными хроматографическими свойствами и пропускным порогом 15 кДа. При помощи только одной колонки с RAM сорбентом им удалось идентифицировать около 400 пептидов, введя напрямую на колонку 2 мкл сыворотки крови. Причем, среди любых трех повторов воспроизводимость была не менее 60%. Еще одним преимуществом, которое было показано в работе Ну и соавторов, является то, что колонки с RAM сорбентом можно комбинировать

с обычными нанопоточными хроматографическими колонками. Используя в качестве ортогонального метода разделения материал C18, авторам удалось идентифицировать 1286 пептидов из 20 мкл сыворотки крови [167].

2.4.4.5 Деплеция основных белков плазмы/сыворотки крови

Ещё один способ выравнивания концентрации между высоко представленными белками и низко представленными полипептидами предложил Soloviev [168, 169]. Для деплеции были разработаны специальные комбинаторные библиотеки пептидных лигандов «Protein Equalizer Technology», короткие аминокислотные последовательности, на которые сорбируются белки и пептиды. Высокое разнообразие аминокислотных последовательностей при конечной емкости колонки приводит к выравниванию концентрации между различными компонентами смеси, существенно уменьшив содержание высоко представленных белков и сконцентрировав остальные низко представленные компоненты.

2.4.5 Методы десорбции пептидов с поверхности основных белков плазмы крови

Альбумин, иммуноглобулины и другие высоко представленные белки крови сорбируют на своей поверхности и переносят различные низкомолекулярные компоненты крови, в том числе, небольшие белки и пептиды [170-175]. Согласно теоретическим оценкам, эти низкомолекулярные компоненты протеома крови практически полностью ассоциированы на поверхности переносчиков [171, 172]. Комплекс с большим белком способен защитить пептиды от выведения из кровотока в почках и способствует их накоплению в крови в количествах пригодных для анализа (Рисунок 1).

Предложено несколько способов разрушения пептид-белковых комплексов [171-178]. Один из способов предполагает выделение на афинных сорбентах интересующих белков вместе с ассоциированными на нём низкомолекулярными белками и пептидами. Проводят десорбцию пептидной фракции и исследуют её состав [171-173]. Так в своей работе группа Gundry исследовала десорбированные

с альбумина пептиды [170]. Комбинирую различные методы разделения, им удалось идентифицировать 35 белков, ассоциированных с альбумином.

Многие авторы не выделяют отдельные фракции пептидов, ассоциированные с какими-либо конкретными белками, а проводят тотальную десорбцию пептид-белковых комплексов, и затем выделяют пептиды [147, 174, 175, 178-180]. Так в своей работе Villanueva и соавторы исследовали влияние различных стадий на качество пептидных MALDI масс-спектрометрических профилей и для разрушения комплексов пептидов с основными белками плазмы крови использовали 8 М мочевины, 8 М мочевины в 20 mM DTT или 0,2% н-октилглюкозид. Они показали, что при использовании 0,2% н-октилглюкозида детектируется втрое большее количество масс-спектрометрических пиков в области от 0,8 до 15 кДа, чем при использовании мочевины или мочевины с DTT [147].

Ещё одной методикой, при которой происходит десоциация пептид-белковых комплексов, является капиллярный электрофорез [179, 180]. В денатурирующих условиях электрофореза белки теряют свою третичную структуру и пептиды естественным образом десорбируются с их поверхности. Затем, электрофоретически низкомолекулярные белки и пептиды отделяются от высокомолекулярных белков.

2.4.6. Использование масс-спектрометрии для поиска в плазме и сыворотке крови потенциальных маркеров онкологических заболеваний

В 2002 году группа американских ученых во главе с Petricoin опубликовала свои результаты сравнительного анализа сыворотки крови пациентов больных раком яичников и здоровых доноров методом SELDI-TOF-MS [134]. Согласно их данным, для отбора масс-спектрометрических пиков, позволяющих создать специфичную и чувствительную диагностику по образцам сыворотки крови, нет необходимости идентифицировать пептиды, образующие эти пики. В результате анализа двух или более групп масс-спектров находят «пептидный паттерн», набор масс-спектрометрических пиков, наилучшим образом отличающий данные группы

друг от друга. Данный паттерн в дальнейшем может быть использован для диагностирования заболевания по масс-спектру.

В работах Petricoin и соавторы использовали масс-спектрометры с методом ионизации SELDI (Surface-enhanced laser desorption/ionization – ионизация с усиливаемой лазерной десорбцией-ионизацией поверхностью) [130, 181]. Суть этого метода заключается в использовании в качестве масс-спектрометрической мишени поверхностей, избирательно связывающих наборы белков на основе адсорбционных, электростатических или специфических аффинных взаимодействий. Фракционирование сложных биологических смесей происходит прямо на поверхности масс-спектрометрической мишени, которую после процедуры связывания целевых веществ с поверхностью и отмывки от примесей помещают в масс-спектрометр для анализа. MALDI обладает рядом преимуществ перед SELDI [182]. Во-первых, для уменьшения компонентного состава сложных биологических образцов перед их анализом методом MALDI масс-спектрометрии используют различные методы фракционирования в объеме, которые чувствительнее, чем фракционирование на модифицированной поверхности мишени SELDI. Во-вторых, разрешение и чувствительность самих масс-спектрометров SELDI уступает современным MALDI масс-спектрометрам.

MALDI/SELDI масс-спектрометрическое профилирование имеет ряд сложностей, связанных с технической стороной работы, а не с принципами масс-спектрометрического анализа «пептидных паттернов» [52, 138]. В качестве основных недостатков подхода чаще всего отмечают отсутствие данных о структуре пептидов, факторы, занижающие чувствительность, и возможный эффект «переучивания» во время сравнительного биоинформатического анализа. Тем не менее, данный подход широко применяется в ряде групп и сегодня, поскольку он прост с технической стороны и не требует идентификации выявленных компонент [54, 165]. Некоторые работы уже вышли на различные стадии масштабных исследований, к примеру, тест OvaCheck разрабатываемый

докторами Petricoin и Liotta в Национальном Институте Рака США (NCI, National Cancer Institute) [183].

2.4.7. Современные достижения в области поиска белковых/пептидных биомаркеров рака яичников

К настоящему моменту одобрение FDA в США получили только два маркера: СА-125 [17] и HE4 [18], а так же два мультипараметрических теста: OVA1 и ROMA. Единичные маркеры СА-125 и HE4 одобрены для мониторинга рецидива или прогрессирования рака яичников, но не для скрининга [9]. Тесты OVA1 и ROMA были одобрены для применения в случаях, когда образование в яичниках уже выявлено, но не ясно является ли оно злокачественным и как срочно необходима операция [19]. На сегодня маркер СА-125 и УЗИ являются единственными используемыми методами периодического обследования для выявления рака яичников. Однако низкая чувствительность для ранних стадий и низкая специфичность этих методов диагностики (Таблица 2) подталкивает исследователей искать новые маркеры данного заболевания.

За последние десять лет было предложено множество потенциальных белковых маркеров на рак яичников. В своих работах авторы используют биологический материал, полученный от пациентов с уже выявленным заболеванием. Различия между протеомным составом у здоровых доноров и больных раком рассматриваются как потенциальные биомаркеры. Выявленные потенциальные маркеры проверяют на дополнительных группах образцов и рассчитывают специфичность и чувствительность маркеров. В таблице 3 приводятся данные последних работ по поиску и валидации потенциальных белковых маркеров рака яичников.

Таблица 3: Потенциальные биомаркеры рака яичников, наиболее часто упоминаемые в литературе с приведенными значениями чувствительностей и специфичностей

Название белка	Чувствительность	Специфичность
Mesothelin [184]	60 – 72%	85 – 98%
TTR (Transthyretin) [9]	96% (вместе с СА-125, АроА1 и transferrin)	98% (вместе с СА-125, АроА1 и transferrin)
АроА1 (Apolipoprotein AI) [9]	93,9% (ранние стадии, вместе с СА-125 и TTR)	95% (вместе с СА-125 и TTR)
VCAM-1 (Vascular cell adhesion molecule 1) [9]	86% (ранние стадии, вместе с СА-125) 93% (поздние стадии, вместе с СА-125)	98% (вместе с СА-125)
IL-6 и IL-8 (Interleukin) [9]	65,5% (ранние стадии) 94,1% (вместе с СА-125 и SAA)	98% 93,1% (вместе с СА-125 и SAA)
B7-H4 [9, 184]	97% (ранние стадии, вместе с СА-125)	65% (вместе с СА-125)
SAA (Serum amyloid A) [9]	95,2% (вместе с СА-125)	–
Transferrin [9]	89% (вместе с СА-125, TTR и АроА1)	92% (вместе с СА-125, TTR и АроА1)
OPN (Osteopontin) [9, 184]	81 – 93,8% 96% (вместе с leptin, prolactin и IGF)	33,7% 94% (вместе с leptin, prolactin и IGF)
KLK (Kallikreins) [9]	70% (для KLK11)	95% (для KLK11)
OVX1 (ovarian cancer antigenX1) [9]	88% (вместе с М-CSF, СА19–19 и СА72–74)	92,5% (вместе с М-CSF, СА19–19 и СА72–74)
VEGF [184] [9]	77% (вместе с СА-125)	87% (вместе с СА-125)
Prostasin [184]	92%	94%
Macrophage colony-stimulating factor [184]	90%	98%
HP- α (Haptoglobin- α) [185]	64%	90%

Несмотря на высокие значения специфичности и чувствительности опубликованных диагностических моделей, единичные потенциальные маркеры вышли на клинический уровень. Одно из возможных объяснений этому несоответствию кроется в принципе формирования групп пациентов для сравнения. Во всех вышеперечисленных работах по масс-спектрометрическому профилированию проводили сравнение между двумя патофизиологически гомогенными группами пациентов: практически здоровые доноры (в большинстве случаев) и пациенты, страдающие заболеванием, которое исследуется. При этом специфичность предлагаемых диагностических моделей по отношению к другим заболеваниям не исследуется, что негативно сказывается на реальном диагностическом потенциале этих моделей. Многие из обнаруживаемых потенциальных биомаркеров ассоциированы с общими воспалительными процессами и могут коррелировать с различными инфекционными заболеваниями [186].

Другой причиной отсутствия реальных маркеров рака яичников может являться недостаток знаний о возможных белковых формах. В работе Tang с соавторами показали как важно отличать различные изоформы одного белка [187]. Используя LTQ Orbitrap XL им удалось идентифицировать уникальные участки четырех изоформ тропомиозина, считающегося потенциальным маркером рака яичников. При помощи 5500 QTRAP от ABSciex им удалось количественно (методом MRM) измерить концентрации каждой изоформы по уникальным фрагментам. Авторы показали, что согласно количественным данным содержание трех изоформ тропомиозина существенно повышено в образцах сыворотки пациенток с раком яичников по сравнению с контрольными образцами. Данная работа указывает на необходимость использовать исчерпывающие данные об изоформах белков при поиске и количественном анализе потенциальных биомаркеров.

Работ, посвященных поиску и идентификации потенциальных пептидных маркеров, проводится значительно меньше, чем аналогичных поисковых работ по

белковым маркерам. К примеру, запрос в PubMed «ovar* cancer biomarker* peptide*» возвращает только 23 статьи, а запрос «ovar* cancer biomarker* protein*» – 116 статьи. Пептиды реже используются, в качестве потенциальных биомаркеров, потому что с ними сложнее работать, и поскольку основную их массу принято рассматривать как «биологический мусор» [56-58]. В последнее время интерес к исследованию пептидов возрастает на фоне открытия новых пептидных регуляторов у животных и растений [188-190], а так же с развитием методов анализа пептидов [58].

МАТЕРИАЛЫ И МЕТОДЫ

1. Образцы сыворотки крови

Сыворотки крови пациенток с диагнозом рак яичников получены из Российского Научного Центра Рентгенорадиологии Минздрава РФ, Москва. Всего в работе использованы образцы от 34 пациенток, средний возраст которых составил 52 года (далее, группа «рак яичников»). По стадиям образцы в группе «рак яичников» распределены следующим образом: 4 образца от пациенток на I стадии, 4 – от пациенток на II стадии, 22 – III стадия и 4 – IV стадия. Стадию заболевания определяли цитоморфологическими исследованиями биоптатов. Данную коллекцию использовали в работе по МАЛДИ масс-спектрометрическому профилированию сыворотки крови.

В работе по идентификации и количественному анализу потенциальных биомаркеров рака яичника была использована коллекция образцов сыворотки крови пациенток с диагнозом рак яичников из Российского онкологического научного центра им. Н. Н. Блохина РАМН, Москва. Всего для работы были взяты образцы от 60 пациенток, средний возраст которых составил 50 лет (далее так же, группа «рак яичников»). Заболевание определяли цитоморфологическими исследованиями биоптатов. Для нивелирования отличий, связанных с индивидуальной изменчивостью пациенток, образцы пулировали по 10 штук, и таким образом получили 6 пулов.

В качестве контроля были использованы образцы сыворотки крови 60 практически здоровых женщин (далее, группа «здоровые доноры»), полученные из клинико-диагностической лаборатории ООО НПФ «Литех», Москва. Средний возраст женщин контрольной группы 49 лет. В работе по МАЛДИ масс-спектрометрическому профилированию сыворотки крови были использованы только 33 образца из данной коллекции, чтобы статистически уравновесить выборку образцов в двух сравниваемых группах. Для идентификации и

количественного анализа потенциальных биомаркеров рака яичника все 60 образцов были пулированы по 10 штук, и таким образом было получено 6 пулов.

В качестве дополнительного контроля в работе по МАЛДИ масс-спектрометрическому профилированию были использованы образцы сыворотки крови пациенток с доброкачественными гинекологическими новообразованиями: 35 пациенток с аденомиозом (далее, группа «аденомиоз»), 20 пациенток больных одновременно аденомиозом и гистеромиомой (далее, группа «аденомиоз с гистеромиомой»), 13 пациенток с гиперпластическими процессами эндометрия (далее, группа «гиперпластические процессы эндометрия»), 12 пациенток с ректоцервикальным эндометриозом (далее, группа «ректоцервикальный эндометриоз»), 27 пациенток с эндометриоидными кистами яичников (далее, группа «эндометриоидные кисты яичников») и 29 пациенток с синдромом поликистозных яичников (далее, группа «синдром поликистозных яичников») (Таблица 4). Все образцы были получены из клинической больницы Российского Университета Дружбы Народов. Поставленные диагнозы проверяли цитоморфологическими исследованиями биоптатов.

Таблица 4: Информация по количеству и среднему возрасту пациенток для каждой экспериментальной группы, использованной в работе по МАЛДИ масс-спектрометрическому профилированию сыворотки крови

Группы пациенток	Средний возраст пациенток	Число образцов, для построения мат. моделей	Число образцов, для внешней валидации мат. моделей
Здоровые доноры	49	17	16
Рак яичников	52	17	17
Аденомиоз	41	18	17
Аденомиоз с гистеромиомой	42	10	10
Гиперпластические процессы эндометрия	41	7	6
Ректоцервикальный эндометриоз	43	6	6
Эндометриоидные кисты яичников	40	14	13
Синдром поликистозных яичников	42	15	14

Кровь отбирали из локтевой вены при помощи вакутейнеров с активатором свертывания (REF11030, Vacutest Kima, Italy). Сыворотку получали после инкубации крови в течении 1 часа при комнатной температуре. Сыворотку отделяли от форменных тел крови центрифугированием (RCF 670 g, 15 мин, при комнатной температуре). После центрифугирования сыворотку хранили в отдельных микропробирках при температуре -80°C.

Для изучения влияния возможных отклонений от протокола получения сыворотки крови на воспроизводимость масс-спектрометрических профилей этих сывороток, были использованы образцы крови двух здоровых мужчин 23 и 49 лет. Сыворотки из этих образцов крови получали по различным протоколам, отличающимся длительностью инкубации крови при комнатной температуре (1, 2, 4, 8 и 24 часа) до отделения сыворотки от сгустка центрифугированием (RCF 670 g, 15 мин, при комнатной температуре). Аналогичные образцы сыворотки крови были получены также и после предварительного выдерживания крови на льду в течение 1, 2, 4 и 8 часов.

2. МАЛДИ МС профилирование сыворотки крови

2.1 Фракционирование образцов сыворотки крови

Фракционирование образцов сыворотки крови проводили на автоматизированной платформе ClinProt Robot (Bruker Daltonics) по нижеприведенным протоколам. Для фракционирования использовали стандартные наборы для профилирования, содержащие магнитные микрочастицы с функционализированной поверхностью MB-HIC 8, MB-HIC 18, MB-WCX и MB-IMAC Cu производства компании Bruker Daltonics.

Образцы сыворотки крови различных групп наносили в 96-луночный планшет для фракционирования на платформе ClinProt Robot в порядке, сформированном функцией RANDBETWEEN программы MSExcel. Рандомизация применялась, чтобы устранить влияние относительного расположения образцов при фракционировании на результаты профилирования. Кроме того, в каждую 96-луночный планшет в позиции A:1, D:3, G:5, B:8, E:10 и H:12 наносили один и тот

же контрольный образец сыворотки крови, который никогда не использовался для построения или валидации классификационных моделей. Спектры этих образцов использовались исключительно для контроля воспроизводимости профилирования как внутри каждого 96-луночного планшета, так и между ними.

2.1.1 Протокол фракционирования с использованием MB-WCX

Перед использованием магнитные микрочастицы тщательно перемешать в течение 1 минуты. Затем, 10 мкл частиц смешать с 10 мкл раствора MB-WCX Binding Solution в каждой лунке 96-луночного планшета. После добавления 5 мкл сыворотки крови в каждую лунку, планшет инкубировать в течение 5 минут при комнатной температуре. Планшет поместить в специальный магнитный держатель. При этом в течение 1 минуты магнитные частицы осядут на поверхности стенок лунок, и раствор станет прозрачным. После этого 25 мкл раствора отобрать из каждой лунки. Для отмывки микрочастиц 100 мкл раствора MB-WCX Washing Solution добавить в каждую лунку планшета. Планшет переставить в держателе 10 раз, меняя положение относительно магнита. После инкубации в держателе в течении 1 минуты, 100 мкл раствора отобрать из каждой лунки. Промывку магнитных частиц повторить дважды. Для элюции 10 мкл раствора MB-WCX Elution Solution добавить в каждую лунку. Планшет инкубировать в держателе в течение 2 минут. Затем, перенести в отдельный 96-луночный планшет 10 мкл раствора из каждой лунки. 10 мкл раствора MB-WCX Stabilization Solution добавить к элюату в каждой лунке.

2.1.2 Протокол фракционирования с использованием MB-HIC 8 (18)

Перед использованием магнитные микрочастицы перемешать, переворачивая баночку с частицами вверх дном и обратно 20 раз. Затем, 5 мкл сыворотки крови смешать с 10 мкл раствора MB-HIC Binding Solution в каждой лунке 96-луночного планшета. После добавления 10 мкл раствора с микрочастицами в каждую лунку, планшет инкубировать в течение 1 минут при комнатной температуре. Поместить планшет в специальный магнитный держатель. В течение 20 секунд магнитные частицы осядут на поверхности стенок лунок, и раствор станет прозрачным. После

этого 25 мкл раствора отобрать из каждой лунки. Для отмывки микрочастиц 100 мкл раствора MB-HIC Washing Solution добавить в каждую лунку. Планшет переставить в держателе 20 раз, меняя положение относительно магнита. После 20 секунд инкубации в держателе, 100 мкл раствора отобрать из каждой лунки. Промывку магнитных частиц повторить дважды. Для элюции 10 мкл 50% ацетонитрила в деионизованной воде добавить в каждую лунку. Планшет инкубировать вне держателя в течение 1 минут, а затем поставить в магнитный держатель ещё на 30 секунд. Наконец, 20 мкл элюата из каждой лунки перенести в отдельный 96-луночный планшет.

2.1.3 Протокол фракционирования с использованием MB-IMAC Cu

Перед использованием магнитные микрочастицы тщательно перемешать в течение 1 минуты. Для отмывки микрочастиц 50 мкл раствора MB-WCX Binding Solution (BS) добавить к 5 мкл раствора частиц в каждой лунке 96-луночного планшета. Планшет переставить в специальном магнитном держателе 10 раз, меняя положение относительно магнита. После 20 секунд инкубации в держателе, 55 мкл раствора отобрать из каждой лунки. Промывку магнитных частиц повторить дважды. Затем, последовательно добавить в каждую лунку 20 мкл BS и 5 мкл сыворотки крови. Планшет инкубировать вне магнитного держателя в течение 5 минут. После инкубации в держателе в течение 20 секунд, 25 мкл раствора отобрать из каждой лунки. Для отмывки частиц 100 мкл раствора MB-WCX Washing Solution добавить в каждую лунку, планшет переставить в держателе 10 раз и инкубировать в держателе в течение 20 секунд. Затем, 100 мкл раствора отобрать из каждой лунки. Промывку магнитных частиц повторить дважды. Для элюции 10 мкл раствора MB-WCX Elution Solution добавить в каждую лунку. Планшет инкубировать вне держателя в течение 5 минут, а затем поместить в магнитный держатель ещё на 20 секунд. 20 мкл элюата перенести из каждой лунки в отдельный 96-луночный планшет.

2.2. Стадия прогревания элюатов образцов сыворотки крови

Стадия прогревания элюатов, полученных в результате фракционирования образцов сыворотки крови, проводили следующим образом. Планшет извлекали из робота, лунки плотно закрывали крышками (low profile thermo-strip, Thermo Scientific) и прогревали в течение 15 мин при 98°C на водяном термостате, после чего охлаждали до комнатной температуры. При помощи центрифугирования осаждали образовавшийся конденсат и агрегировавшие белки (Eppendorf Centrifuge 5804 R, ротор A-2-DWP, RCF 2250 g, 15 мин).

2.3. МАЛДИ масс-спектрометрическое оборудование

После прогревания или после фракционирования (если стадия прогревания не проводилась) элюаты смешивали с раствором матрицы (смесь 2,5-дигидроксibenзойной кислоты (0,24 мг/мл) и альфа-циано-гидроксикоричной кислоты (0,3 мг/мл) в ацетонитрил/метанол 1:1) в соотношении 1:15 и наносили на масс-спектрометрическую мишень AnchorChip 600/384 (производства Bruker Daltonics) в четырех повторах по 0,8 мкл. Эти операции проводили с использованием автоматизированной системы пробоподготовки ClinProt Robot (Bruker Daltonics).

Масс-спектры получали с использованием времяпролетного масс-спектрометра Ultraflex™ TOF/TOF фирмы Bruker Daltonics. Десорбцию образцов осуществляли иррадиацией азотным лазером (длина волны 337 нм), работающим при частоте 25 Гц. Для удаления пиков матрицы использовали максимальный уровень подавления сигнала вплоть до 500 Да. Спектры регистрировали в линейном режиме положительно заряженных ионов в диапазоне масс от 0,6 до 12 кДа. Для калибровки использовали калибровочную смесь, содержащую пептиды и белки в диапазоне масс 1-17 кДа. Для увеличения чувствительности детекции избыток матрицы удаляли 8 импульсами лазера при мощности 60% с последующей аккумуляцией данных при мощности лазера 45%. Для каждого спектра суммировали результаты 720 лазерных импульсов (по 60 импульсов с 12 различных

точек пятна). Суммировали спектры с отношением сигнал/шум > 5 и разрешением > 300 .

Для автоматического получения МАЛДИ масс-спектров с каждой точки мишени создавали AutoXecute лист очередности измерений при помощи написанной нами на С# программы, которая позволяла рандомизировать порядок измерений. Рандомизация применялась, чтобы устранить влияние очередности измерений на результаты профилирования.

2.4. Анализ масс-спектрометрических данных

Масс-спектрометрические данные анализировали с использованием компьютерной программы ClinProTools 2.1 (Bruker Daltonics). Для удаления базовой линии использовали алгоритм Top Hat Baseline с параметром минимальной ширины 50%. Для перекалибровки спектров допускали отклонение до 1000 ppm, если хотя бы 30% масс-спектрометрических пиков совпадало с реперными пиками. Анализ вариаций площадей масс-спектрометрических пиков между группами «рак яичников» и «здоровые доноры» проводили на основе статистического теста Welch's t-test (который является упрощенной формой Student's t-test) в случае нормального распределения вариации площадей или при помощи теста Wilcoxon rank-sum, если нормального распределения не наблюдалось. Для определения, является ли распределение вариации площадей нормальным, использовали тест Anderson-Darling (модификация теста Kolmogorov-Smirnov). Если значение p-value теста Anderson-Darling для вариации площади конкретного масс-спектрометрического пика было больше 0,05, то распределение считали нормальным и использовали данные теста Welch's t-test. В противном случае использовали данные теста Wilcoxon rank-sum. Если значение p-value теста Welch's t-test или Wilcoxon rank-sum не превышало 0,05 для вариации площади какого-либо масс-спектрометрического пика, то данный пик считали статистически значимым для отличия групп «рак яичников» и «здоровые доноры».

Математические модели для классификации масс-спектров, полученных после профилирования сыворотки крови, строили на основе Генетического

Алгоритма (ГА) и алгоритма Обучаемой Нейронной Сети (ОНС). При построении классификационных моделей использовали следующие параметры. Для ГА: количество пиков, входящих в создаваемую модель классификации – не более 7; число поколений – не более 500; начальное число комбинаций пиков выбиралось автоматически; уровень мутаций – 0.2; уровень кроссинговера – 0,5. Количество пиков, используемых для построения классификационной модели на основе ОНС, выбиралось алгоритмом автоматически (но не более 25); максимальное число итераций – 99х100; число прототипов выбиралось автоматически. Для внутренней валидации использовали алгоритм Leave one out.

Массивы спектров каждой из групп «рак яичников» и «здоровые доноры» разделяли пополам. На основании одной половины строили классификационные модели, а оставшиеся использовали для внешней валидации построенных моделей (Таблица 4).

Для построения математических моделей классификации, которые способны отличить пациенток группы «рак яичников» от пациенток с доброкачественными гинекологическими заболеваниями, мы сформировали объединенные группы. В группу «номинальный контроль 1» входили масс-спектрометрические профили образцов групп «аденомиоз» и «эндометриоидные кисты яичников». В группу «номинальный контроль 2» входили масс-спектрометрические профили образцов групп «аденомиоз», «эндометриоидные кисты яичников» и «ректоцервикальный эндометриоз». В группу «номинальный контроль 3» входили масс-спектрометрические профили образцов групп «аденомиоз», «эндометриоидные кисты яичников», «ректоцервикальный эндометриоз», «гиперпластические процессы эндометрия» и «здоровые доноры». Массивы спектров каждой из групп «рак яичников» и «номинальный контроль №» разделяли пополам. На основании одной половины строили классификационные модели, а оставшиеся использовали для внешней валидации построенных моделей.

Классификационные модели называли согласно экспериментальным группам, которые использовали для их построения. Первым в названии модели стоит название целевой группы патологии («рак яичников»).

3. Идентификация потенциальных пептидных биомаркеров

3.1 Фракционирование образцов сыворотки крови

Фракционирование пулированных образцов сыворотки крови проводили вручную на магнитных микрочастицах MB-WCX (производства Bruker Daltonics) по нижеприведенному протоколу.

3.1.1 Протокол фракционирования с использованием MB-WCX

Перед использованием магнитные микрочастицы тщательно перемешать в течение 1 минуты. Затем, 200 мкл частиц смешать с 200 мкл раствора MB-WCX Binding Solution. После добавления 100 мкл сыворотки крови, раствор инкубировать в течение 5 минут при комнатной температуре. Микропробирку поместить в специальный магнитный держатель. При этом в течение 1 минуты магнитные частицы осядут на поверхности стенок пробирки, и раствор станет прозрачным. После этого 500 мкл раствора отобрать из пробирки. Для отмывки микрочастиц 1000 мкл раствора MB-WCX Washing Solution добавить к частицам. Микропробирку переставить в держателе 10 раз, меняя положение относительно магнита. После инкубации в держателе в течении 1 минуты, 1000 мкл раствора отобрать. Промывку магнитных частиц повторить дважды. Для элюции 200 мкл раствора MB-WCX Elution Solution добавить к частицам. Смесь инкубировать в держателе в течение 2 минут. Затем, перенести в отдельную микропробирку 200 мкл раствора. 200 мкл раствора MB-WCX Stabilization Solution добавить к элюату.

3.2 Стадия прогревания элюатов образцов сыворотки крови

Полученные в результате фракционирования образцов сыворотки крови, элюаты прогревали в течение 15 мин при 98⁰С на водяном термостате, после чего образцы охлаждали до комнатной температуры. При помощи центрифугирования осаждали образовавшийся конденсат и агрегировавшие белки (Eppendorf Centrifuge 5415 D, RCF 16100 g, 15 мин).

3.3 Обессоливание на микроколонках с обращено фазовой поверхностью

Обессоливание элюатов после прогревания проводили на микроколонках с обращено фазовой поверхностью по нижеприведенному протоколу. Микроколонки подготавливали самостоятельно по протоколам рекомендованных в статье Rappsilber с соавторами [191].

После обессоливания элюаты концентрировали на вакуумном центрифужном концентраторе SpeedVac (SVC-100H, Savant) до объема 5 мкл, добавляли к элюату 10 мкл раствора 5% ацетонитрила в 0,1% муравьиной кислоте и анализировали методом LC-MS/MS. Для масс-спектрометрического анализа использовали по 5 мкл образца на каждый технический повтор.

3.3.1 Протокол фракционирования на микроколонках с обращено фазовой поверхностью

Для фракционирования 400 мкл элюата подготовить 4 микроколонки по 3 слоя мембраны в каждой. Дальнейшие действия выполнять параллельно на 4 микроколонках. Для промывки микроколонки пропустить через неё последовательно 20 мкл метанола и 20 мкл 0,1% трифторуксусной кислоты. После этого пропустить через микроколонку 100 мкл элюата. Для промывки микроколонки от солей пропустить через неё 20 мкл 0,1% трифторуксусной кислоты. Для элюции пропустить через микроколонку 30 мкл 50% ацетонитрила в 0,1% трифторуксусной кислоте.

3.4 Масс-спектрометрическое оборудование

LC-MS/MS анализ проводили при помощи масс-спектрометров TripleTOF5600+ с источником ионизации NanoSpray III фирмы ABSciex, объединенном с нано-HPLC системой NanoLC Ultra 2D+ фирмы Eksigent. Буфер для загрузки образцов и буфер А: 98,9% вода, 1% метанол, 0,1% муравьиная кислота (объемные доли). Буфер Б: 99,9% ацетонитрил, 0,1% муравьиная кислота (объемные доли). Образцы загружали в петлю Chrom XP C18 3 мкм 120 Å 350 мкм*0.5 мм (Eksigent) при скорости потока 3 мкл/мин в течении 10 минут и

элюировали через отдельную колонку 3C18-CL-120 (3 мкм 120 Å) 75 мкм*150 мм (Eksigent) при скорости потока 300 нл/мин при градиенте от 5 до 40% буфера Б в течении 120 минут. Колонку и петлю промывали между измерениями 95% буфера Б в течении 7 минут, а затем 5% буфера Б в течении 25 минут.

Для экспериментов по идентификации масс-спектрометр TripleTOF5600+ (ABSciex) функционировал в режиме IDA. Данный режим предполагает циклическую работу прибора. Каждый цикл включает измерение обзорного MS спектра с последующим получением спектров фрагментации наиболее интенсивных родительских ионов. Обзорный спектр измеряли в диапазоне m/z 300-1250 при времени накопления сигнала 250 мс. Спектры дочерних ионов измеряли для 50 наиболее интенсивных родительских ионов с интенсивностью выше 400 сигналов в секунду и зарядовым состоянием от 2 до 5. Спектры дочерних ионов измеряли при следующих параметрах: ширина пропускания квадрупольного масс-фильтра 0,7 Да; диапазон m/z 200-1800; время накопления сигнала 50 мс; энергия столкновений линейно поднималась от 25 до 55 эВ за время накопления сигнала. Допустимая погрешность определения m/z родительского иона 20 ppm, а дочерних – 50 ppm. Динамическое исключение было настроено на добавление родительских ионов во временный лист исключений на 15 с после каждого их MS/MS анализа, чтобы измерить спектр фрагментации в точке максимально близкой к вершине пика (минимальная ширина пика была около 30 с).

3.5 Идентификация пептидов и биоинформатический анализ

Полученные в результате LC-MS/MS анализа масс-спектрометрические данные первоначально были проанализированы при помощи программы ProteinPilot (version 4.5) с использованием алгоритма поиска Paragon против базы данных белковых последовательностей UniProtKB по таксону HUMAN, скачанная с официального сайта <http://www.uniprot.org/> 14 Марта 2013 года. Данный анализ позволяет уменьшить ошибку определения m/z ионов прекурсоров за счет калибровки при условии максимизации числа получаемых идентификаций. При

помощи программы ProteinPilot были получены файлы масс-листов в формате .mgf для дальнейшего анализа.

Для идентификации пептидов и белков по MS/MS спектрам применили согласованный подход, основанный на использовании двух поисковых алгоритмов: Mascot и X!Tandem. Параметры для всех поисковых программ выставляли одинаковыми: неспецифический гидролиз, отсутствие постоянных или переменных модификаций, точность определения m/z иона прекурсора – 20 ppm, точность определения дочерних ионов – 50 ppm (0,04 Da для Mascot). При анализе была использована актуальная на тот момент база данных белковых последовательностей UniProtKB по таксону HUMAN, скачанная с официального сайта <http://www.uniprot.org/> 14 Марта 2013 года.

Результаты идентификации объединялись и валидировались при помощи программного продукта Scaffold (Proteome Software). Валидацию проводили только по пептидам, устанавливая допустимое пороговое значение FDR на уровне 95%.

Основная задача работы и сама методика пробоподготовки не подразумевала идентификации целых белков, а только их пептидных фрагментов (или, в исключительном случае, низкомолекулярных белков). Идентификация отдельных фрагментов белков в некоторых случаях (гомогенные белки или изоформы одного белка) не позволяла соотнести пептид только с одним конкретным белком. Поэтому, если один или несколько пептидов одновременно относились к нескольким белкам, то такие белки объединяли в белковые группы.

4. Количественный анализ содержания пептидов в сыворотке крови методом SWATH

Результаты идентификаций пептидов на основе MS/MS анализа были использованы для построения библиотеки переходов для SWATH анализа. Для анализа были использованы по 3 пулированных образца из групп «здоровые доноры» и «рак яичников», которые подготавливали точно так же, как для

идентификации пептидов. Для всех SWATH LC-MS/MS экспериментов проводили 3 технических повтора.

4.1 Масс-спектрометрическое оборудование

Для SWATH LC-MS/MS анализ использовали те же параметры хроматографической системы, что и для LC-MS/MS экспериментов по идентификации пептидов.

Параметры масс-спектрометрической системы были следующие: диапазон m/z родительских ионов 300-1250, спектры дочерних ионов снимались для окон пропускания первого квадруполь в 27,9 а.е.м., время накопления сигнала для каждого окна 0,1 сек, суммарное время цикла на весь диапазон масс 3,36 сек, что в случае средней ширины пика в 30 секунд позволяет снимать около 9 точек на пик. Диапазон m/z измерения дочерних ионов 200-1800. Энергия столкновений в течение времени измерения дочерних спектров (0,1 с) линейно изменяется от 32,5 до 47,5 В.

4.2 Биоинформатический анализ данных SWATH

Библиотека переходов была создана с помощью плагина MS/MS(ALL) with SWATH™ Acquisition MicroApp (версия 1.0.0.653) программы PeakView Software (версия 1.2.0.3, фирмы ABSciex) с использованием только наиболее достоверно идентифицированных пептидов (достоверность пептидов по расчетам ProteinPilot 95%). Для поиска соответствий между SWATH LC-MS/MS экспериментами и библиотекой переходов использовали окна времени удерживания в 5 мин и допустимую погрешность m/z в 50 ppm. Для количественного анализа для каждого из пептидов в библиотеке были использованы 10 его наиболее интенсивных фрагментов.

Анализ SWATH результатов проводили в программе MarkerView (версия 1.2.1.1, ABSciex). Перед началом анализа для всех сравниваемых экспериментов была проведена нормализация площадей пиков, исходя из суммарной интенсивности для всех пептидов. Были рассчитаны относительные изменения концентрации пептидов между группами «здоровые доноры» и «рак яичников».

Значимым считалось увеличение/уменьшение содержания пептида в два и более раз. Для пептидов, содержание которых отличалось в десять и более раз, была проведена ручная валидация корректности соотнесения библиотечных спектров со спектрами, автоматически ассоциированными с данными пептидами программой PeakView Software.

Далее был проведен анализ на основе t-критерия Стьюдента для поиска наиболее отличающихся компонент между образцами групп «здоровые доноры» и «рак яичников». Значение параметра p-value указывает на вероятность того, что выявленные различия являются случайными. Высокие значения параметра p могут наблюдаться, например, при низком соотношении сигнал/шум в образце для данных переходов или в случае некорректного соотнесения списка переходов из библиотеки с реальными экспериментальными SWATH спектрами (например, при ошибке в калибровке или при сильном сдвиге времени удерживания пептида на колонке). В качестве порогового значения для параметра p-value мы использовали 0,05.

РЕЗУЛЬТАТЫ

1. Метод выделения пептидов из сыворотки крови человека

Нашей группой разработан новый метод выделения пептидов из сыворотки крови человека. Данный метод прост в проведении и по сравнению со стандартными методиками выделения пептидов не требует дополнительных стадий с добавлением к смеси каких-либо химических соединений [175, 178, 192-198].

1.1 Сравнение пептидного и белкового состава элюатов до и после прогрева

Для оценки преимуществ, которые дает введение дополнительной стадии прогрева, мы провели фракционирование пяти образцов из группы «здоровые доноры» на различных типах магнитных микрочастиц. МАЛДИ масс-спектры элюатов были измерены до и после прогрева (Рисунок 2).

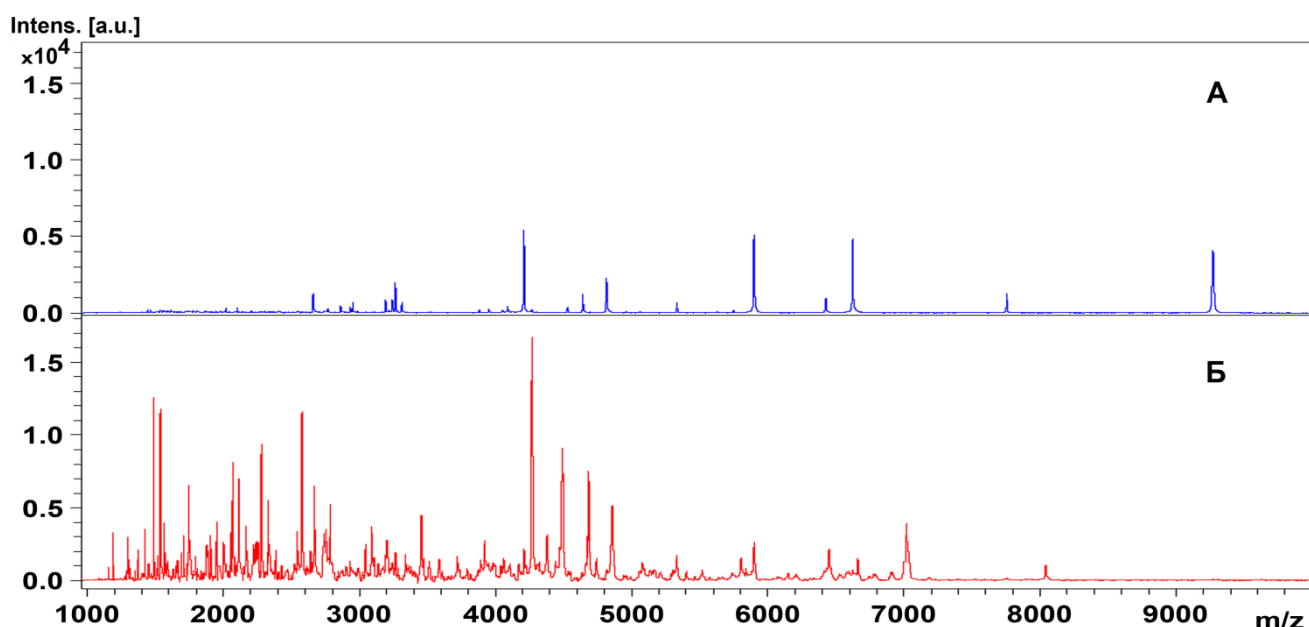


Рисунок 2: МАЛДИ масс-спектры фракционированного на магнитных микрочастицах MB-WCX образца сыворотки крови практически здорового донора до (А) и после (Б) прогрева элюата. По оси ординат отложена интенсивность спектра, по оси абсцисс – отношение массы к заряду в Да.

Из рисунка видно, что после стадии прогрева увеличивается как число пиков в спектре, так и общая интенсивность спектра. Точное число детектируемых масс-спектрометрических пиков в элюатах, полученных после фракционирования на различных типах магнитных микрочастиц было оценено по результатам анализа сыворотки крови пяти практически здоровых доноров (Таблица 5).

Таблица 5: Количество детектируемых масс-спектрометрических пиков во фракциях сыворотки крови пяти практически здоровых доноров, полученных с использованием магнитных микрочастиц с функционализированной поверхностью с прогреванием элюатов и без (среднее значение $\pm$ стандартное отклонение)

Протокол фракционирования	Тип использованных магнитных микрочастиц		
	MB-WCX	MB-NIC 8	MB-IMAC Cu
Без прогрева элюатов	23 $\pm$ 6	30 $\pm$ 7	22 $\pm$ 5
С прогреванием элюатов	376 $\pm$ 34	365 $\pm$ 32	257 $\pm$ 24

В фракционированных образцах сыворотки крови после прогрева­ния улучшается воспроизводимость масс-спектрометрических профилей элюатов (Рисунок 3).

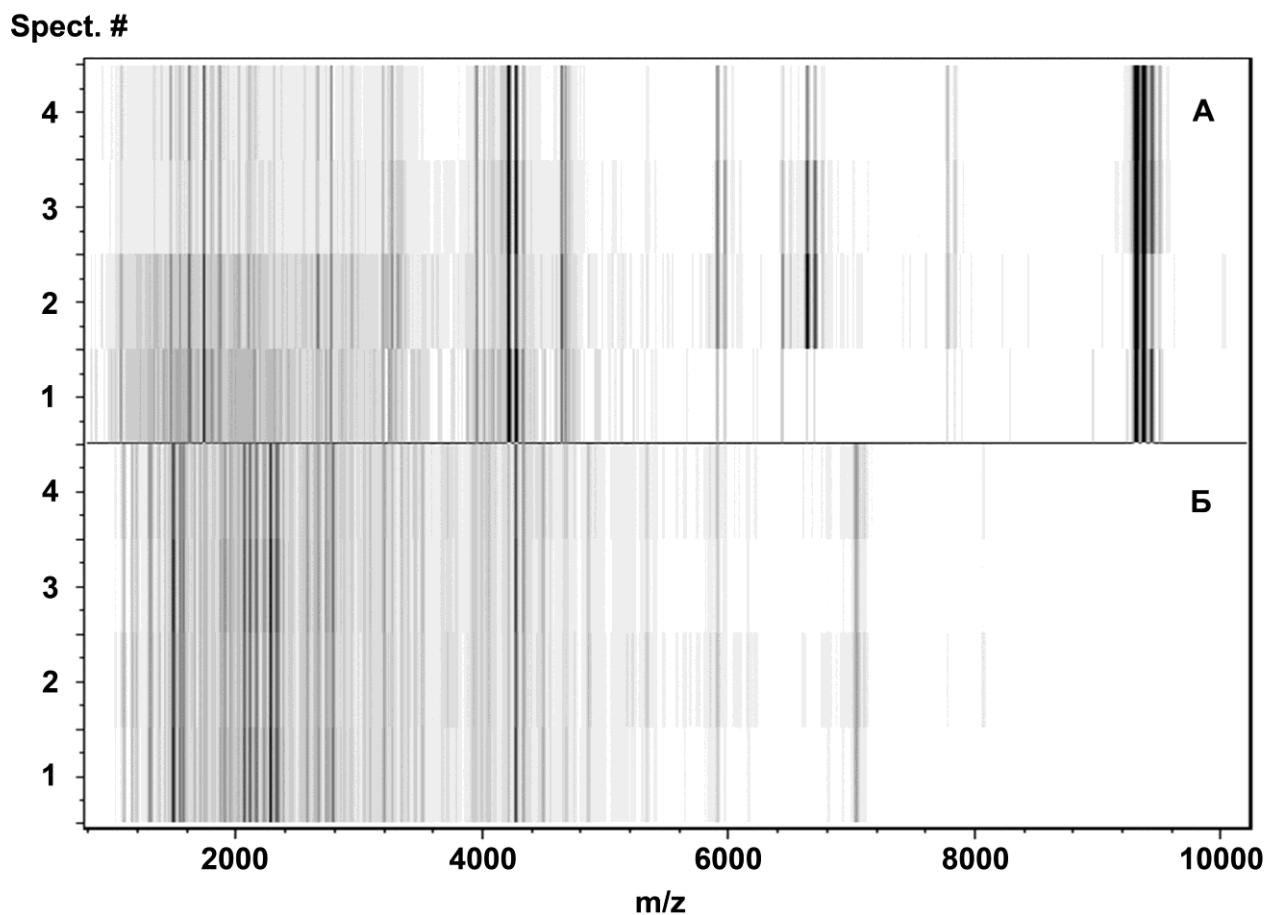


Рисунок 3: Представленные в виде псевдо-геля МАЛДИ масс-спектрометрические профили образца сыворотки крови практически здорового донора, полученные после его фракционирования на магнитных микрочастицах MB-WCX с последующим прогреванием (А) и без (Б). Каждое измерение проводили в четырех повторах. По оси ординат отложены номера спектров, по оси абсцисс – отношение массы к заряду в Да. Интенсивность окрашивания полос соответствует интенсивности пика в спектре, нормированном по общему ионному току.

SDS-PAGE анализ образцов сыворотки крови, фракционированных с использованием магнитных микрочастиц с функционализированной поверхностью, показал значительные различия в белковом профиле образцов до и после стадии прогревания элюатов (Рисунок 4).

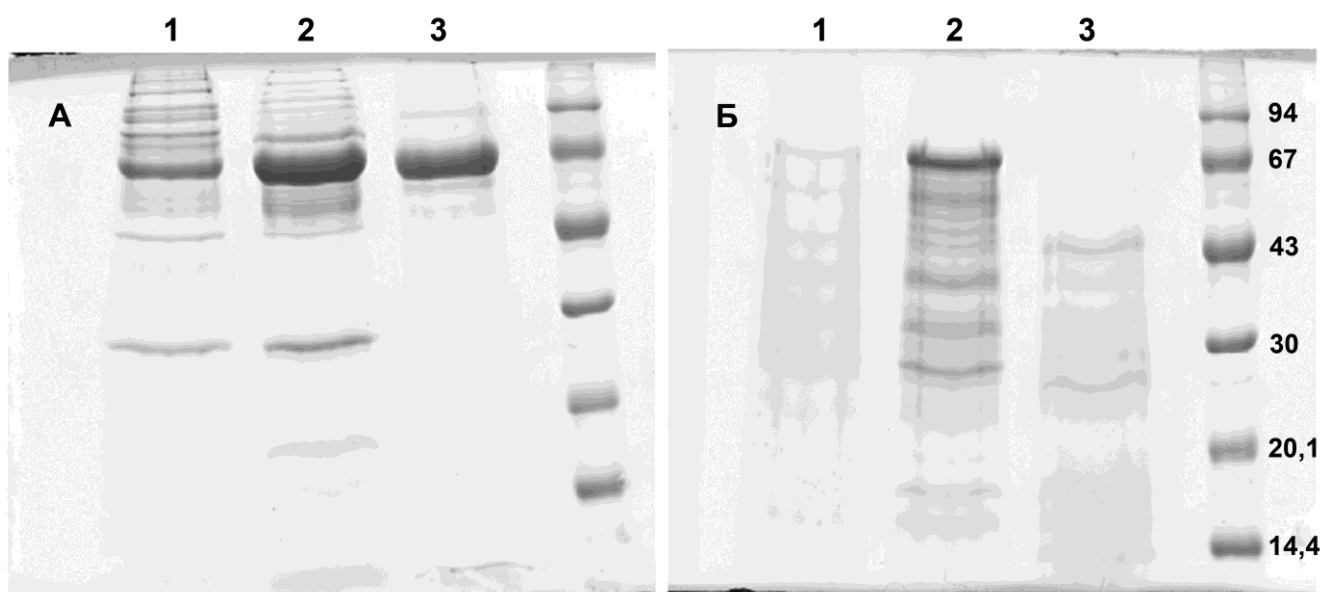


Рисунок 4: 12% Т SDS-PAGE образцов сыворотки крови практически здоровых доноров, фракционированных на магнитных микрочастицах MB-WCX (1), MB-HIC8 (2), MB-IMAC Cu (3) до (А) и после (Б) стадии прогревания.

1.2 Прогревание белка Apolipoprotein A-I

МАЛДИ масс-спектрометрический анализ очищенного белка Apolipoprotein A-I (sp|P02647|APOA1_HUMAN, Sigma A0722-1MG, Apolipoprotein A-I from human plasma, $\geq 85\%$ SDS-PAGE, buffered aqueous solution) растворенного в воде качества milli-Q не показал существенных изменений в пептидном профиле образца до и после прогревания (Рисунок 5А и Рисунок 5Б). Однако прогревание Apolipoprotein A-I в буфере для элюции MB-WCX (MB-WCX Elution Solution и MB-WCX Stabilization Solution 1:1, pH = 2÷3) привело к существенному изменению МАЛДИ спектра (Рисунок 5В).

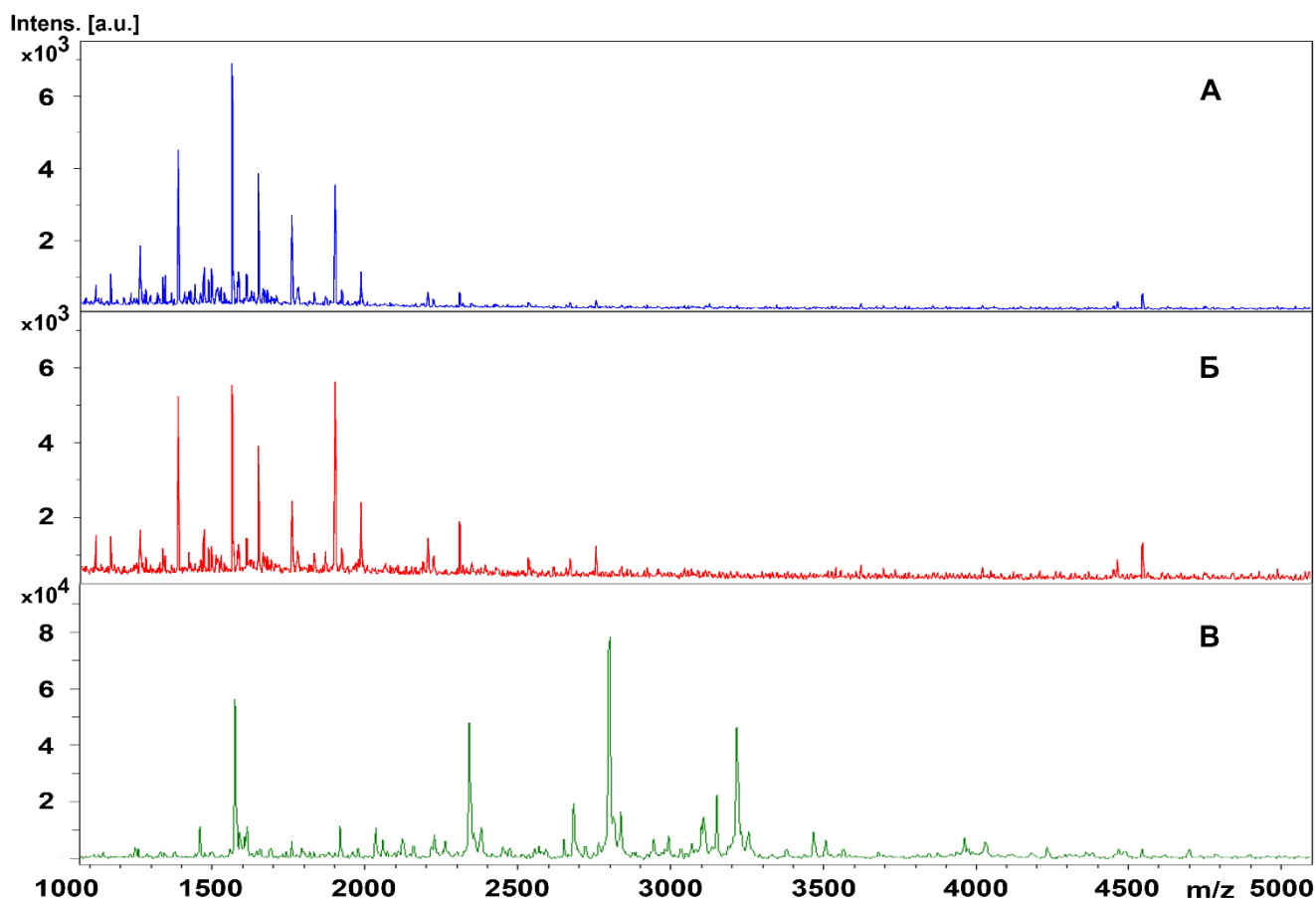


Рисунок 5: МАЛДИ масс-спектры образца очищенного белка Apolipoprotein A-I до (А) и после (Б) прогревания в воде качества milli-Q, а также после прогревания в буфере для элюции MB-WCX (В). По оси ординат отложена интенсивность спектра, по оси абсцисс – отношение массы к заряду в Да.

Все приводимые ниже результаты, если отдельно не оговаривается, получены с использованием протокола включающего стадию прогревания элюатов [175].

2. Построение классификационных моделей

2.1 Фракционирование образцов сыворотки крови с использованием различных типов магнитных микрочастиц

МАЛДИ масс-спектрометрические профили элюатов, полученных после фракционирования одних и тех же образцов сыворотки крови на различных типах магнитных микрочастиц, различались между собой значительно (Рисунок 6) [175, 199]. Наиболее однородную картину представлял спектр элюата, полученный после фракционирования на магнитных микрочастицах MB-WCX. Количество детектируемых в масс-спектрах пиков см. выше (Таблица 5, нижняя строка).

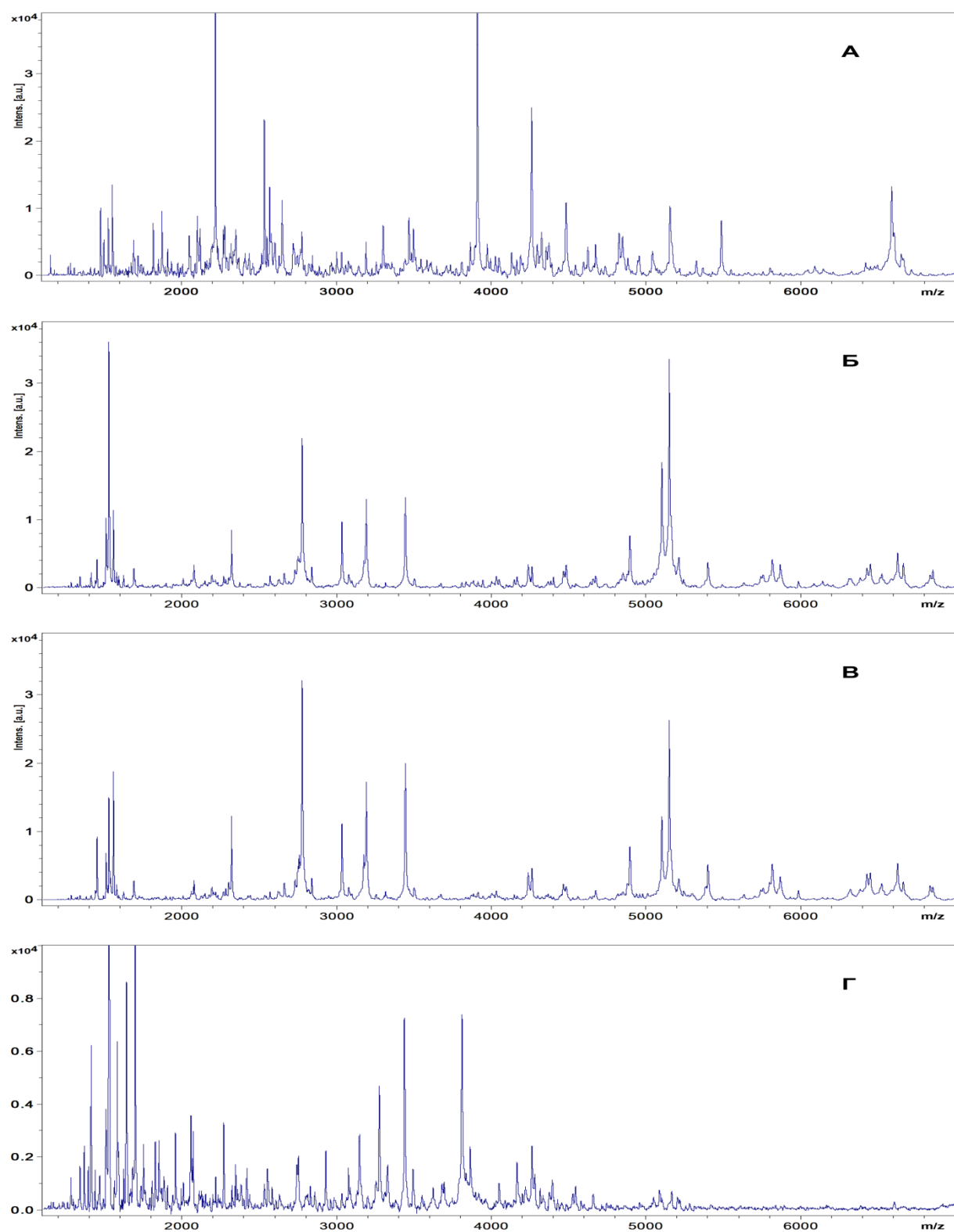


Рисунок 6: Характерные МАЛДИ масс-спектрометрические профили образцов сыворотки крови, полученные после фракционирования на магнитных микрочастицах с функционализированной поверхностью: MB-WCX (А), MB-HIC 8 (Б), MB-HIC 18 (В) и MB-IMAC Cu (Г).

На основе Генетического Алгоритма (ГА) и алгоритма Обучаемой Нейронной Сети (ОНС) были построены классификационные модели для групп «рак яичников» и «здоровые доноры», определены значения чувствительности¹ и специфичности² построенных моделей (Таблица 6) [175, 199]. Из таблицы видно, что пептидные наборы, полученные при использовании частиц MB-WCX, обеспечивают создание на их основе классификационных моделей, обладающих самыми высокими значениями чувствительности и специфичности. Математические модели «рак яичников / здоровые доноры», построенные на основе сравнения профилей образцов сыворотки крови после фракционирования на магнитных микрочастицах MB-WCX, показали близкие в 100% значения чувствительности и специфичности (Таблица 6, Рисунок 8) [196, 200].

Таблица 6: Значения чувствительности и специфичности классификационных моделей, построенных на основе ГА и алгоритма ОНС для групп «рак яичников» и «здоровые доноры», фракционированных на всех использованных типах магнитных микрочастиц

Тип магнитных микрочастиц	Генетический Алгоритм		Обучаемая Нейронная Сеть	
	Чувствительность	Специфичность	Чувствительность	Специфичность
MB-HIC 8	88%	96%	88%	98%
MB-HIC 18	98%	83,3%	81,3%	100%
MB-IMAC Cu	97,2%	88%	100%	90%
MB-WCX	100%	97,1%	93,3%	100%

¹ Чувствительностью (sensitivity) называют процент правильно диагностированных больных раком к их общему числу

² Специфичностью (specificity) называют процент правильно определенных здоровых людей по отношению к их общему числу

Мы приняли решение проводить статистический анализ, последующие идентификацию и количественный анализ потенциальных биомаркеров рака яичников только с использованием магнитных микрочастиц MB-WCX.

2.2 Построение классификационных моделей на основе профилирования на магнитных микрочастицах MB-WCX

Анализ вариации площадей масс-спектрометрических пиков между группами «рак яичников» и «здоровые доноры» позволил выявить 18 пиков статистически значимых для отличия этих групп (Таблица 7, Рисунок 7). Из рисунка видно, что для большинства масс-спектрометрических пиков области характерного распределения площадей (между 25% и 75% квантилями) между группами образцов не пересекаются. Не являясь высоко специфичными биомаркерами, эти пики в совокупности друг с другом и с менее значимыми пиками позволяют отличать масс-спектры образцов сыворотки крови групп «рак яичников» и «здоровые доноры». Классификационные модели используют комбинации из нескольких пиков для получения статистически более значимых отличий между масс-спектрами групп образцов.

Таблица 7: Масс-спектрометрические пики, являющиеся статистически значимыми для отличия групп «рак яичников» и «здоровые доноры» по данным тестов Welch's t-test и Wilcoxon rank-sum.

МС пик		Среднее значение площади под МС пиками		Стандартное отклонение для вариации площади под МС пиками		p-value статистических тестов		
Номер	m/z	Группа «рак яичников»	Группа «здоровые доноры»	Группа «рак яичников»	Группа «здоровые доноры»	Welch's t-test	Wilcoxon rank-sum	Anderson-Darling
118	3524,24	8,93	31,2	7,34	13,07	0,000989	0,000475	0,0000748
125	3905,16	13,49	7,83	6,57	1,75	0,0503	0,00175	0,000195
66	2082,05	68,43	136,18	44,55	39,51	0,00667	0,00256	0,0342
55	1885,87	50,59	28,76	20,8	8,64	0,0246	0,00662	0,00916
83	2422,3	22,81	44,48	16,31	18,14	0,0278	0,00662	0,00165
109	3263,44	26,96	12,21	10,87	5,08	0,00667	0,00147	0,0713
192	6180,52	3,98	8,17	1,87	3,18	0,00773	0,00279	0,123
73	2210,6	27,48	56,53	23,78	37,66	0,107	0,0103	0,0000248
105	3137,77	15,7	28,06	8,92	7,9	0,0121	0,00225	0,613
119	3535,91	6,51	12,28	4,81	3,62	0,0246	0,0054	0,521
74	2215,8	25,01	35,02	6,52	8,41	0,0246	0,00662	0,674
103	3094,81	16,46	26,7	7,89	8,12	0,0255	0,00662	0,305
57	1945,11	38,44	103,84	44,55	54,33	0,0246	0,0306	0,0000748
88	2671,32	77,74	101,58	20	19,32	0,0351	0,0351	0,902
212	6797,65	2,65	1,38	1,32	0,53	0,0357	0,0295	0,111
94	2787,01	114,27	173,78	42,62	58,01	0,0443	0,0503	0,351
63	2050,35	19,69	26,17	5,34	6,01	0,0454	0,0306	0,6
80	2331,1	135,66	177,57	40,73	33,17	0,0473	0,0386	0,562

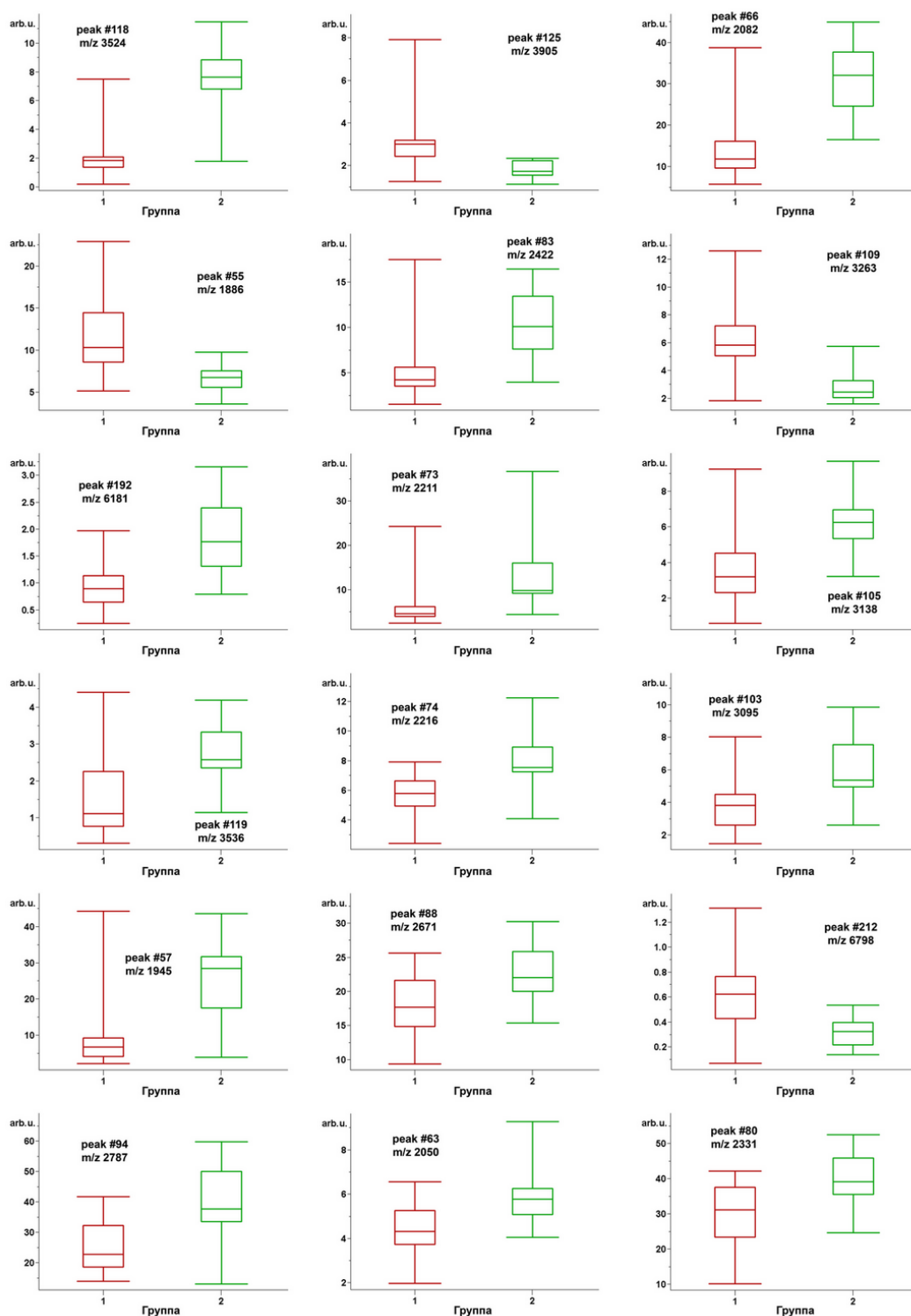


Рисунок 7: Статистические диаграммы вариации площадей масс-спектрометрических пиков в спектрах образцов групп «рак яичников» и «здоровые доноры». Приведены диаграммы вариации площадей только для статистически значимых МС пиков (Таблица 7). По оси ординат отложены нормированные площади пика, по оси абсцисс – группы спектров: 1 – «рак яичников», 2 – «здоровые доноры». Диаграмма состоит из столбцов, отсекающих 25% и 75% квантили, медиан столбцов, максимальных и минимальных значений площадей пиков.

Математические модели, построенные на основе сравнения МАЛДИ масс-спектрометрических профилей образцов сыворотки крови групп «рак яичников» и «здоровые доноры» после фракционирования на магнитных микрочастицах MB-WCX, показали близкие в 100% значения чувствительности и специфичности (Таблица 6, Рисунок 8) [175, 196, 199, 200]. Среди масс-спектрометрических пиков статистически значимых для отличия групп «рак яичников» и «здоровые доноры» (Таблица 7) 7 из 18 были выбраны при построении моделей на основе ГА (Таблица 9) и/или ОНС (Таблица 8).

Таблица 8: Масс-спектрометрические пики, выбранные алгоритмом Обучаемой Нейронной Сети при построении модели. Указана область m/z , соответствующая расположению пика, величина m/z вершины пика. Статистический вес определяет важность масс-спектрометрического пика для построенной модели.

Номер	m/z	Начало m/z	Конец m/z	Статистический вес
38	1540,03	1534,94	1545,93	0,0551
40	1569,78	1565,86	1577,23	0,0000
47	1710,05	1705,05	1717,86	0,0423
51	1753,49	1750,49	1758,31	0,0295
53	1793,98	1788,62	1801,75	0,0106
57	1945,11	1936,97	1949,33	0,0162
65	2068,19	2061,28	2077,15	0,1355
68	2113,51	2106,76	2117,86	0,0003
86	2641,07	2633,02	2650,15	0,0045
91	2753,85	2745,85	2756,62	0,0975
94	2787,01	2780,29	2795,73	0,0034
103	3094,81	3091,99	3101,9	0,1033
105	3137,77	3129,84	3143,79	0,0004
109	3263,44	3255,23	3269,99	0,0000
116	3455,06	3440,56	3467,71	0,0044

Номер	m/z	Начало m/z	Конец m/z	Статистический вес
118	3524,24	3519,34	3531,43	0,2970
119	3535,91	3531,43	3544,22	0,0000
127	3933,28	3929,45	3941,08	0,0932
132	4013,78	4008,04	4023,86	0,0036
145	4308,32	4299,04	4318,59	0,0037
154	4475,45	4462,8	4482,26	0,0327
168	5083,06	5073,06	5091,53	0,0448

Таблица 9: Масс-спектрометрические пики, выбранные Генетическим Алгоритмом при построении модели. Указана область m/z, соответствующая расположению пика, величина m/z вершины пика. Статистический вес определяет важность масс-спектрометрического пика для построенной модели.

Номер	m/z	Начало m/z	Конец m/z	Статистический вес
40	1569,78	1565,86	1577,23	0,75
65	2068,19	2061,28	2077,15	0,06
68	2113,51	2106,76	2117,86	0,16
105	3137,77	3129,84	3143,79	1,09
109	3263,44	3255,23	3269,99	1,22
118	3524,24	3519,34	3531,43	1,52
119	3535,91	3531,43	3544,22	0,98

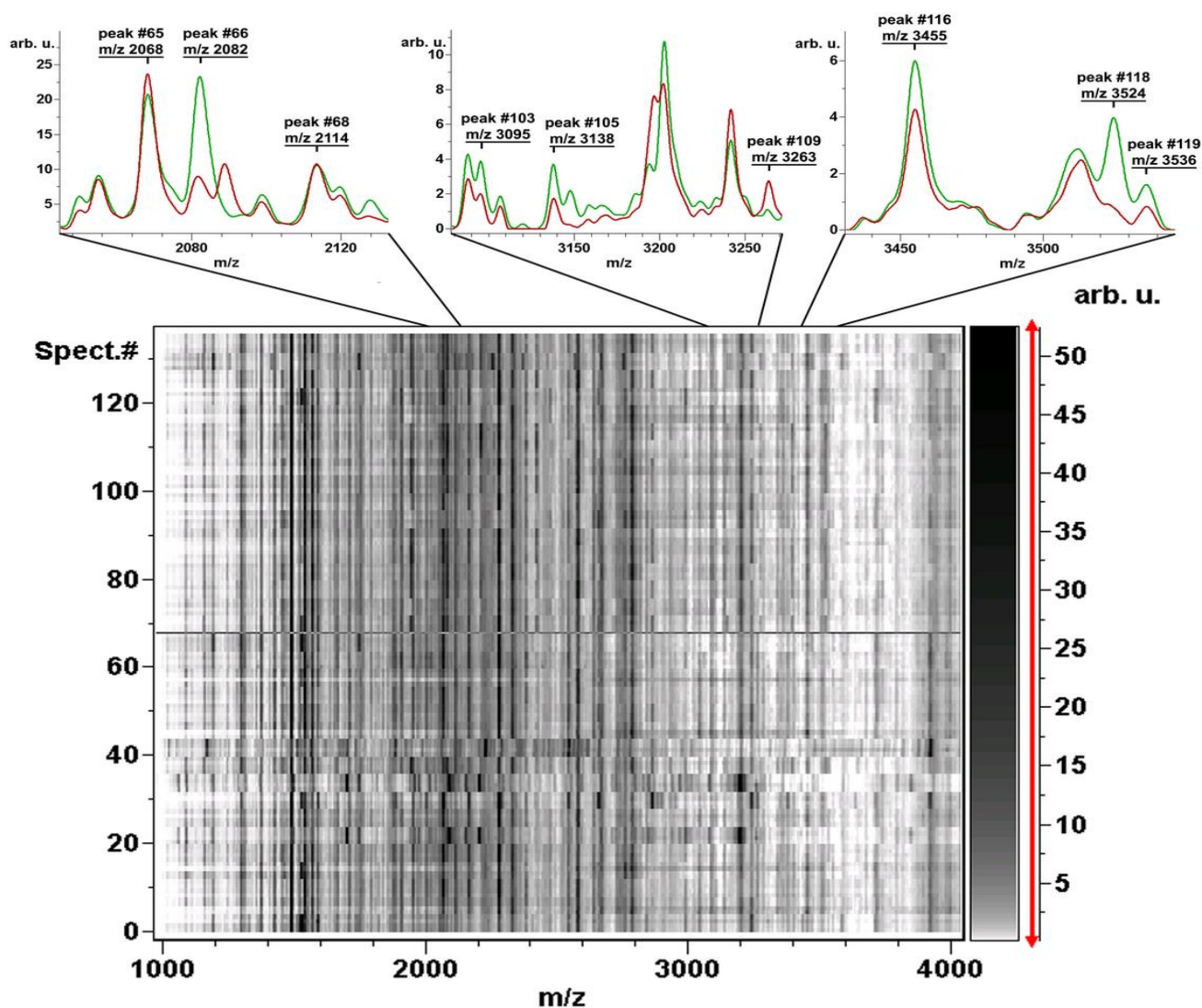


Рисунок 8: Итоговый анализ масс-спектрометрических профилей образцов сывороток крови, фракционированных с использованием микрочастиц MB-WCX. В виде псевдо-геля представлены масс-спектрометры образцов групп «рак яичников» (снизу) и «здоровые доноры» (сверху). В верхней части рисунка показаны отдельные участки усредненных масс-спектров групп «рак яичников» (красный цвет) и «здоровые доноры» (зеленый цвет).

2.3 Построение классификационных моделей сравнения групп «рак яичников» и «номинальный контроль»

Математические модели, построенные на основе сравнения МАЛДИ масс-спектрометрических профилей образцов сыворотки крови групп «рак яичников» и «здоровые доноры» после фракционирования на магнитных микрочастицах MB-

WCX, показали низкие значения специфичности по отношению к доброкачественным гинекологическим заболеваниям (Таблица 10).

Таблица 10: Специфичность модели «рак яичников / здоровые доноры», построенной на основании ГА, относительно доброкачественных гинекологических заболеваний

Группы пациенток	Специфичность
Аденомиоз	26,1%
Аденомиоз с гистеромиомой	25,0%
Гиперпластические процессы эндометрия	46,2%
Ректоцервикальный эндометриоз	25,0%
Эндометриоидные кисты яичников	48,1%
Синдром поликистозных яичников	44,8%

Для построения математических моделей классификации, которые способны отличить пациенток группы «рак яичников» от пациенток с доброкачественными гинекологическими заболеваниями, мы сформировали смешанные группы, состоящие из образцов группы «здоровые доноры», а также групп пациенток с доброкачественными гинекологическими заболеваниями. Группа «номинальный контроль 1» включала образцы групп «аденомиоз» и «эндометриоидные кисты яичников». Группа «номинальный контроль 2» состояла из образцов групп «аденомиоз», «эндометриоидные кисты яичников» и «ректоцервикальный эндометриоз». В группу «номинальный контроль 3» входили образцы групп «здоровые доноры», «аденомиоз», «эндометриоидные кисты яичников», «ректоцервикальный эндометриоз» и «гиперпластические процессы эндометрия».

Классификационные модели, построенные для групп «рак яичников» и «номинальный контроль» показали высокие значения чувствительности и специфичности (Таблица 11), но низкие значения кросс-специфичности по отношению к тем патологиям, которые не входили в группы сравнения (Таблица 12).

Таблица 11: Чувствительность и специфичность классификационных моделей, способных различать рак яичников от доброкачественных гинекологических заболеваний

Классификационная модель	Чувствительность	Специфичность
рак яичников / номинальный контроль 1	93,3%	100%
рак яичников / номинальный контроль 2	81,3%	97,0%
рак яичников / номинальный контроль 3	87,5%	98,1%

Таблица 12: Специфичность классификационных моделей, способных различать рак яичников от доброкачественных гинекологических заболеваний, по отношению к различным группам пациенток. Курсивным начертанием обозначены значения специфичности по отношению к тем группам пациенток, образцы которых использовались при построении модели.

Группы пациенток	Специфичности классификационных моделей		
	рак яичников / номинальный контроль 1	рак яичников / номинальный контроль 2	рак яичников / номинальный контроль 3
Здоровые доноры	43,8%	56,3%	<i>100%</i>
Аденомиоз	<i>100%</i>	<i>100%</i>	<i>100%</i>
Аденомиоз с гистеромиомой	83,3%	66,7%	57,1%
Гиперпластические процессы эндометрия	100%	100%	<i>100%</i>
Ректоцервикальный эндометриоз	91,7%	91,7%	<i>100%</i>
Эндометриоидные кисты яичников	<i>100%</i>	<i>100%</i>	96,3%
Синдром поликистозных яичников	78,6%	100%	100%

2.4 Влияние длительности инкубации крови до получения сыворотки на последующий масс-спектрометрический профиль

Для выполнения данной работы были взяты образцы сыворотки крови пациенток трех различных медицинских центров и одной частной диагностической фирмы. Отличия в протоколах получения сыворотки в этих учреждениях в основном сводились к разнице во времени инкубации крови до отделения сыворотки от сгустка (от одного до нескольких часов). Кроме того, в ООО НПФ «Литех» (поставщик образцов группы «здоровые доноры») сыворотку могли

получать после того, как в течение нескольких часов пробирки с кровью везли в лабораторию на льду.

Для выяснения влияния длительности инкубации крови до получения сыворотки на последующий масс-спектрометрический профиль сыворотки, мы провели забор крови у двух практически здоровых мужчин 23 и 49 лет. Сыворотки из этих образцов крови получали по различным протоколам, отличающимся длительностью инкубации крови при комнатной температуре (1, 2, 4, 8 и 24 часа) до отделения сыворотки от сгустка центрифугированием. Аналогичные образцы сыворотки крови были получены также и после предварительного выдерживания крови на льду в течение 1, 2, 4 и 8 часов [175].

Масс-спектрометрические профили образцов сыворотки крови после их фракционирования на магнитных микрочастицах MB-WCX (без стадии прогревания элюатов) показали неудовлетворительную воспроизводимость, как внутри группы из четырех повторов, так и между группами (Рисунок 9).

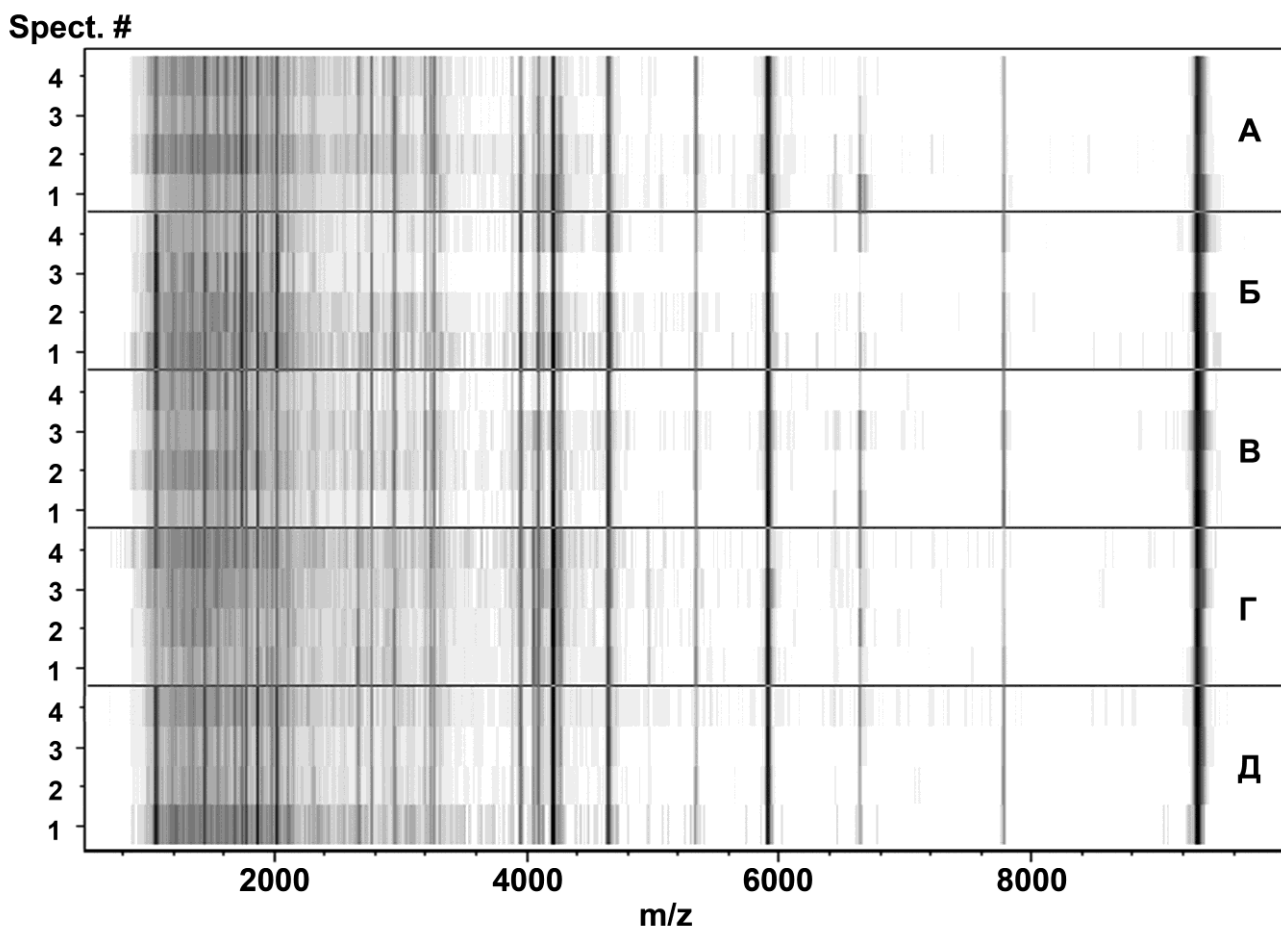


Рисунок 9: Представленные в виде псевдо-геля масс-спектрометрические профили фракционированных с использованием магнитных микрочастиц MB-WCX образцов сыворотки крови, полученных после 1 (А), 2 (Б), 4 (В), 8 (Г) или 24 (Д) часов инкубации крови до отделения сыворотки от сгустка центрифугированием. Стадия прогревания элюатов не проводилась. Каждое измерение проводили в четырех повторях.

Прогревание элюатов приводит к значительному увеличению воспроизводимыми вне зависимости от условий получения сыворотки (Рисунок 10). Кроме того, ещё раз заметим, что возрастает количество детектируемых пиков, особенно в низкомолекулярной области спектра.

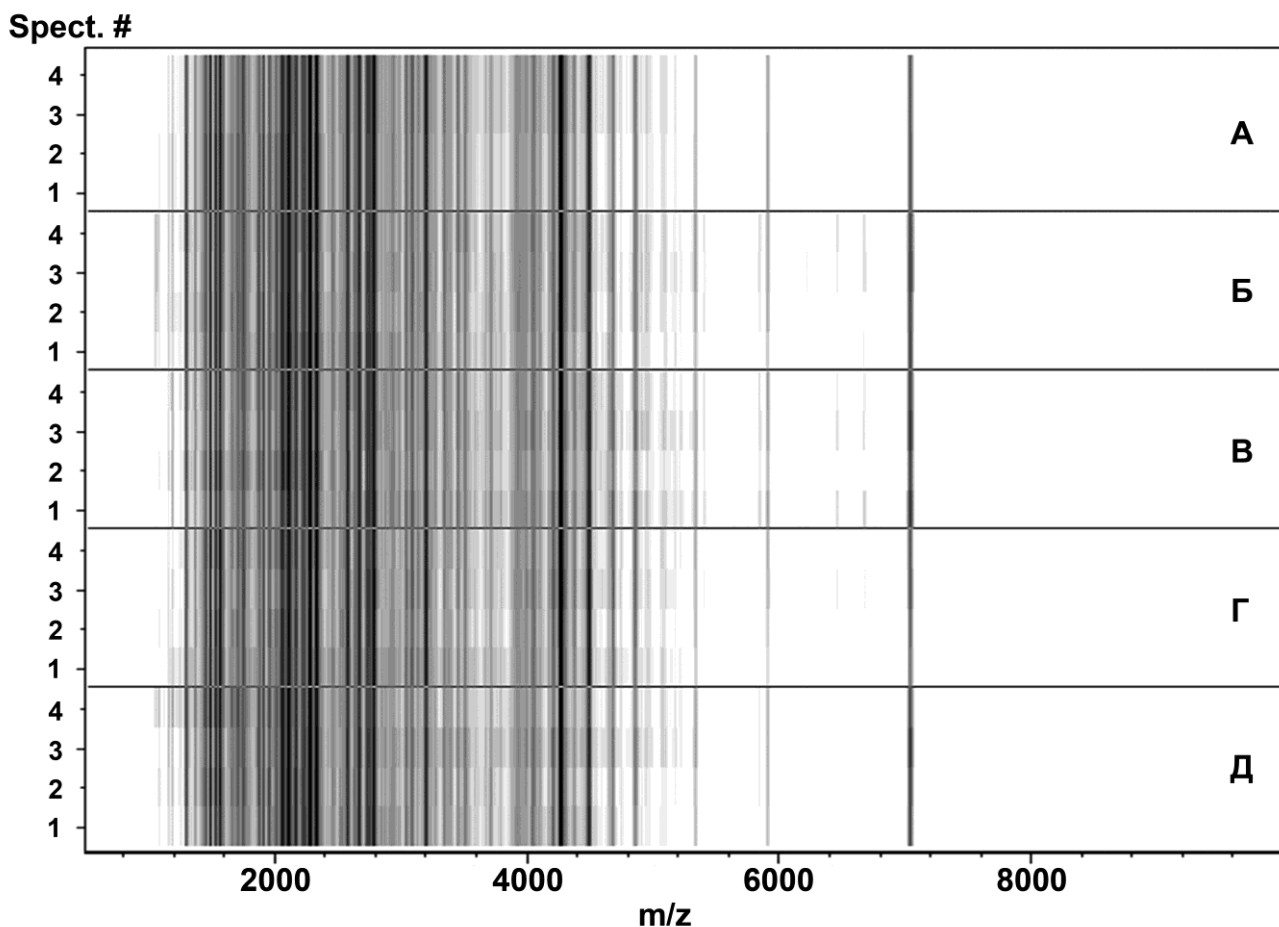


Рисунок 10: Представленные в виде псевдо-геля масс-спектрометрические профили фракционированных с использованием магнитных микрочастиц MB-WCX образцов сыворотки крови, полученных после 1 (А), 2 (Б), 4 (В), 8 (Г) или 24 (Д) часов инкубации крови до отделения сыворотки от сгустка центрифугированием. Стадия прогревания проводилась. Каждое измерение проводили в четырех повторах.

Аналогичное увеличение количества детектируемых низкомолекулярных пептидов и улучшение воспроизводимости мы наблюдали и при анализе образцов, полученных после выдерживания крови на льду в течение 1, 2, 4 и 8 часов [175].

3. Идентификация потенциальных биомаркеров рака яичников

Методами масс-спектрометрии был проведен поиск и валидация пептидных биомаркеров, специфичных для образцов сывороток крови пациенток с раком яичников. На первом этапе работы образцы сыворотки крови практически здоровых доноров и пациенток с диагнозом рак яичников были проанализированы

с точки зрения получения максимального числа идентифицируемых в них пептидов [175]. Всего таким способом было проанализировано 6 пулов по 10 сывороток из каждой группы образцов: «рак яичников» и «здоровые доноры».

Всего было идентифицировано 7996 уникальных пептидов, являющихся фрагментами 2068 уникальных белков (Рисунок 11 и Рисунок 12). В пулах группы «рак яичников» было идентифицировано 2058 уникальных пептидов, не идентифицированных в пулах группы «здоровые доноры». Из них 790 пептидов относятся к белкам, фрагменты которых не были идентифицированы для группы «здоровые доноры» (Рисунок 11).

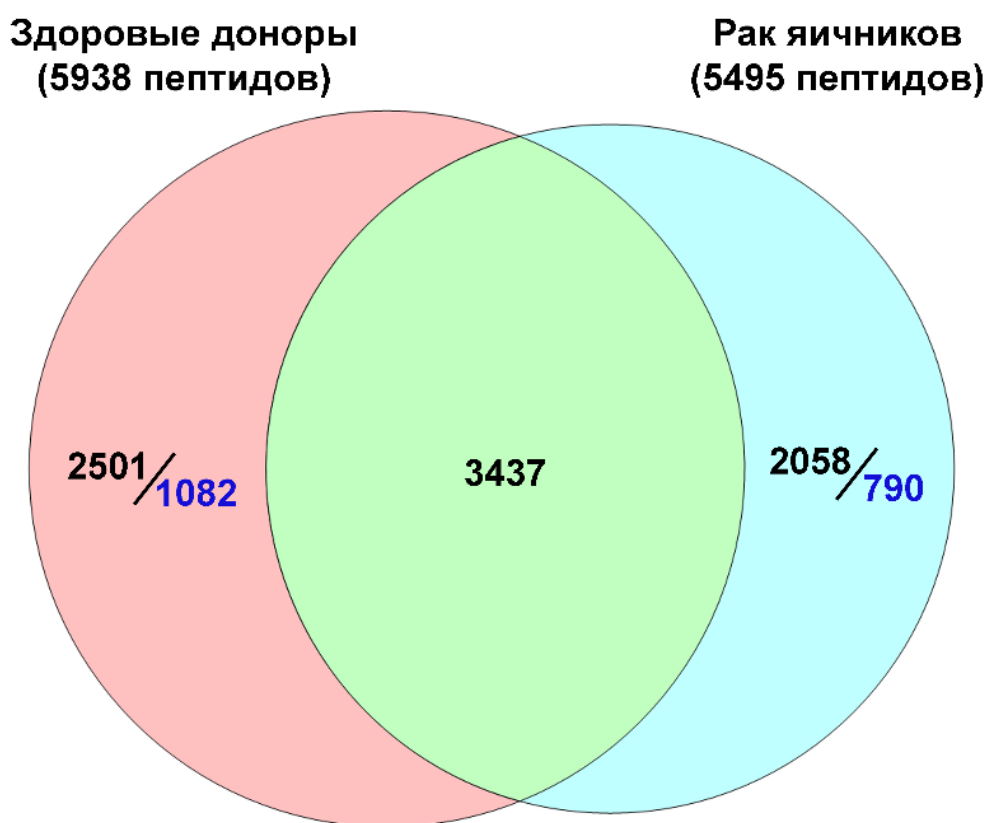


Рисунок 11: Диаграмма распределения идентифицированных пептидов между экспериментальными группами. Синим указано количество пептидов, из числа идентифицированных только в данной экспериментальной группе, относящихся к уникальным для данной группы белкам.

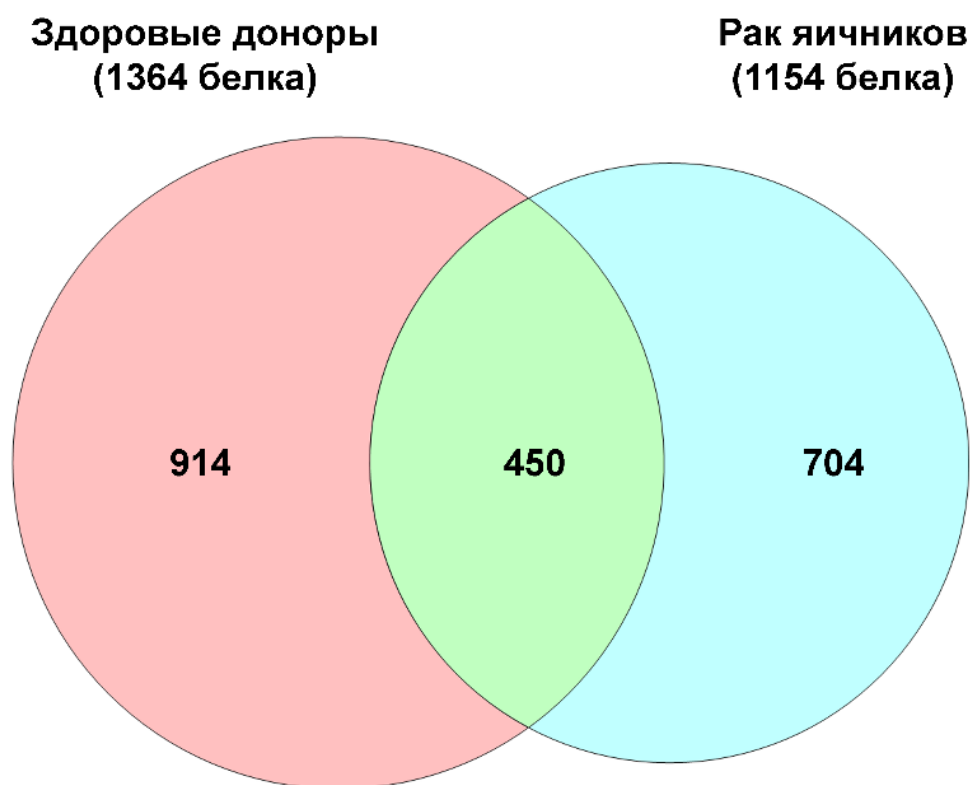


Рисунок 12: Диаграмма распределения белков, предшественников идентифицированных пептидов, между экспериментальными группами.

Количество пептидов, являющихся фрагментами одного белка, достигало 851 как в случае с белком Apolipoprotein A-I (Таблица 13). Наложив аминокислотные последовательности этих пептидов на сиквенс белка, мы увидели, так называемые, «пептидные лестницы» [59] (Рисунок 13).

Таблица 13: 20 белков, представленных наибольшим количеством уникальных идентифицированных пептидов

#	Уникальный номер белка	Название белка	Молекулярная масса белка, кДа	Количество уникальных пептидов, идентифицированных как фрагменты данного белка	
				для группы «здоровые доноры»	для группы «рак яичников»
1	P02647	Apolipoprotein A-I	31	851	819
2	P06396	Gelsolin	86	258	202
3	P02765	Alpha-2-HS-glycoprotein	39	232	222
4	P02671	Fibrinogen alpha chain	95	176	248
5	P02768	Serum albumin	69	197	184
6	Q14624	Inter-alpha-trypsin inhibitor heavy chain H4	103	202	169
7	P01024	Complement C3	187	194	174
8	P04004	Vitronectin	54	166	169
9	P10909	Clusterin	52	126	114
10	P06727	Apolipoprotein A-IV	45	123	74
11	P01042	Kininogen-1	72	96	100
12	P0DJI8	Serum amyloid A-1 protein	14	53	127
13	P02656	Apolipoprotein C-III	11	77	85
14	P0C0L5	Complement C4-B	193	88	74
15	A8K008	cDNA FLJ78387	52	76	62
16	P00734	Prothrombin	70	72	64
17	B4E1Z4	Complement factor B	141	67	58
18	P04196	Histidine-rich glycoprotein	60	57	66
19	B4E1B2	cDNA FLJ53691, highly similar to Serotransferrin	75	59	54
20	P02652	Apolipoprotein A-II	11	63	44

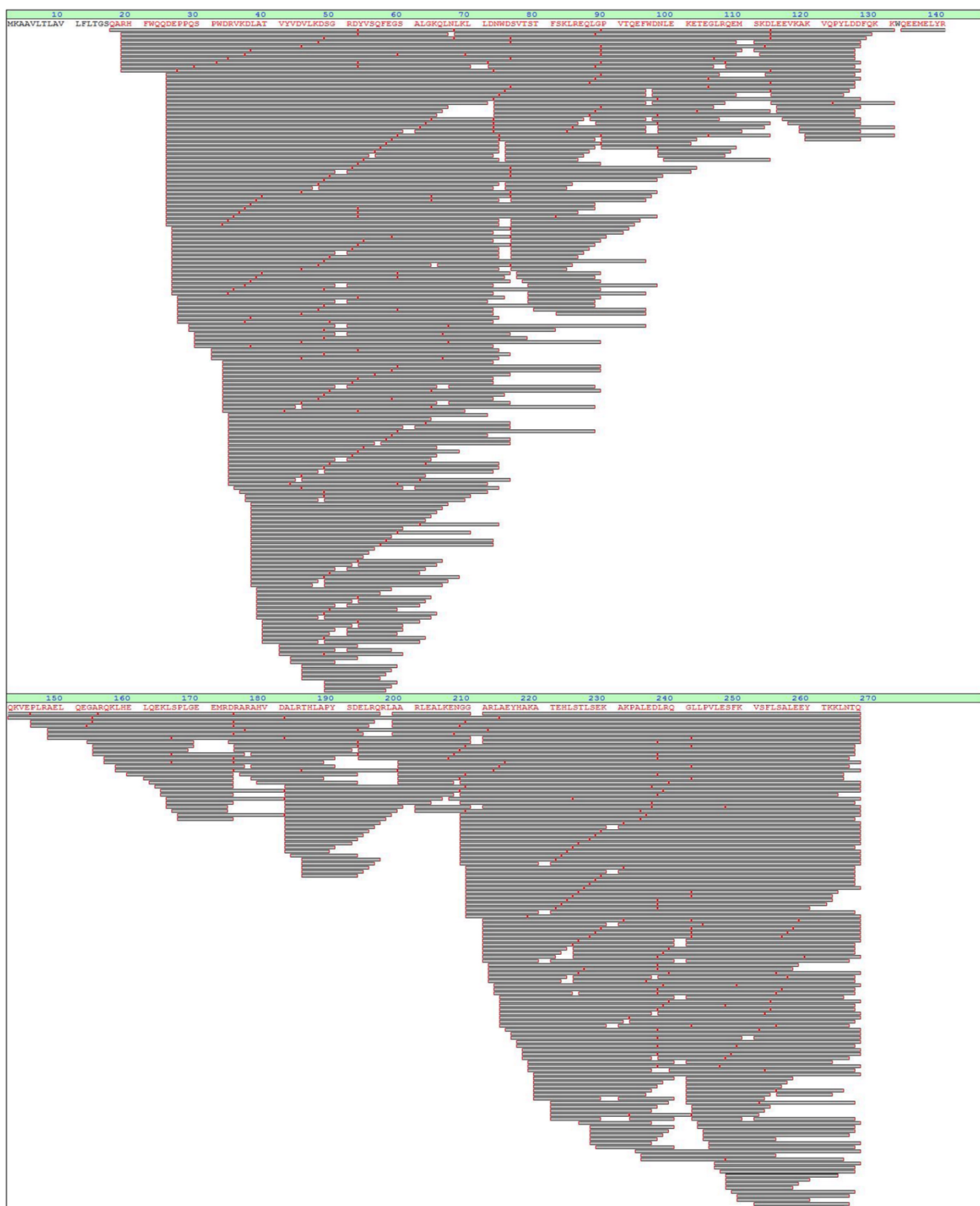


Рисунок 13: Наложение аминокислотных последовательности уникальных идентифицированных пептидов, фрагментов белка Apolipoprotein A-I на сиквенс белка.

Для 7996 идентифицированных уникальных пептидов мы подсчитали количество аминокислот, которые остаются на N- и C-конце при выщеплении пептидов из белков предшественников (Рисунок 14). Мы так же учли аминокислоты в последовательностях белков, предшествующие и следующие за идентифицированными пептидами. Наибольшее отклонения по количеству C-концевых групп приходится на аспарагиновую кислоту, а по количеству N-концевых групп – на пролин и серин.

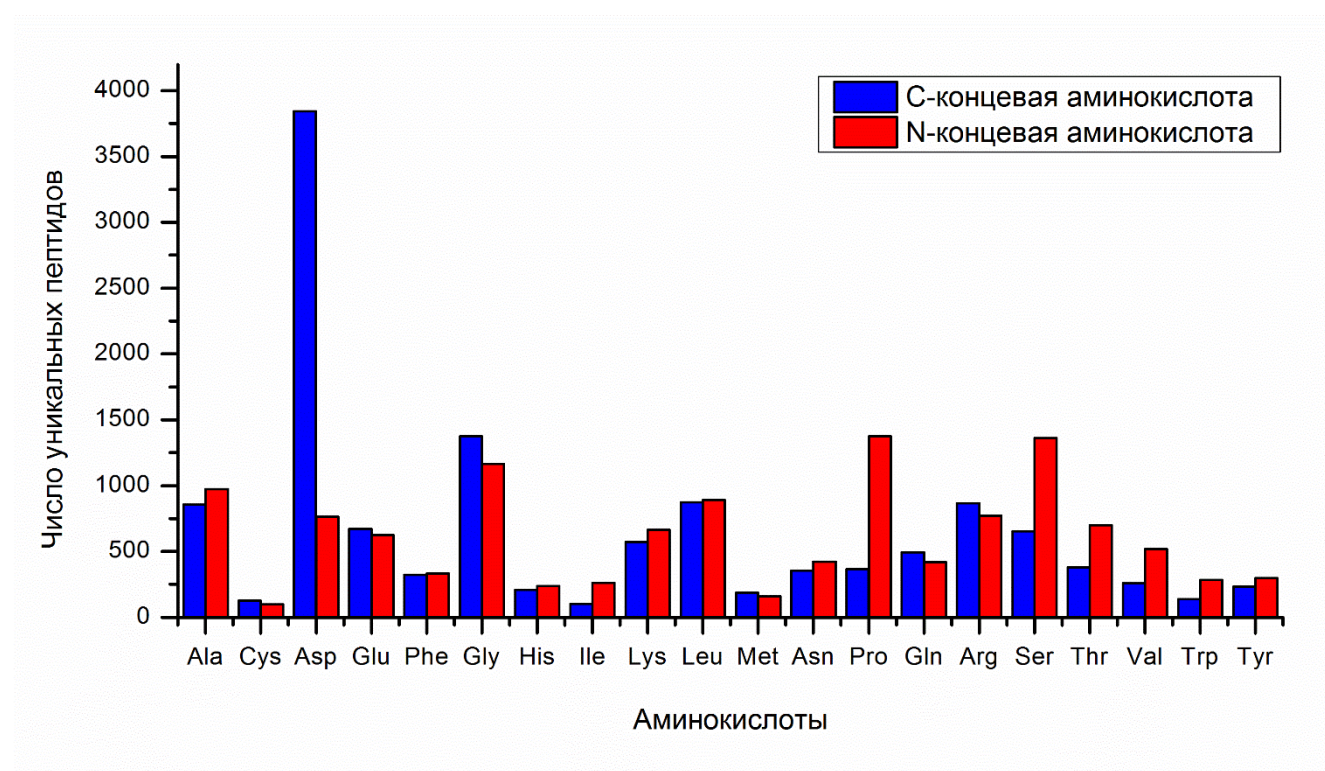


Рисунок 14: График распределения C- и N-концевых аминокислот по частоте встречаемости на концах идентифицированных пептидов. Для расчетов использовались последовательности пептидов, а также аминокислоты, предшествующие и следующие за пептидами (всего 12325 сайтов гидролиза).

4. Анализ потенциальных биомаркеров количественным безметочным масс-спектрометрическим методом SWATH

Пулированные образцы сыворотки крови практически здоровых доноров и пациенток с диагнозом рак яичников были проанализированы при помощи количественного безметочного масс-спектрометрического метода SWATH (см. Приложение 1). Для выявления пептидов, содержание которых отличается между сравниваемыми группами, рассчитали логарифм по основанию 2 от отношения содержания пептида в образцах сыворотки крови пациентов групп «рак яичников» и «здоровые доноры», соответственно. Кроме того, для выявления статистически значимых отличий был применен t-критерий Стьюдента. Для отбора использовались значения модуля логарифма большие или равные 1, а также значения p-value t-критерия Стьюдента меньшие или равные 0,05.

В пулах группы «рак яичников» было обнаружено 36 уникальных пептидов, концентрация которых в 2 или более раз выше, чем в пулах группы «здоровые доноры», а в пулах группы «здоровые доноры» было обнаружено 28 уникальных пептидов, концентрация которых в 2 или более раз выше, чем в пулах группы «рак яичников» (см. Приложение 1 и Рисунок 15). В пулах группы «рак яичников» было обнаружено 11 уникальных пептидов, концентрация которых в 10 или более раз выше, чем в пулах группы «здоровые доноры», а в пулах группы «здоровые доноры» было обнаружено 2 уникальных пептида, концентрация которых в 10 или более раз выше, чем в пулах группы «рак яичников» (Таблица 14).

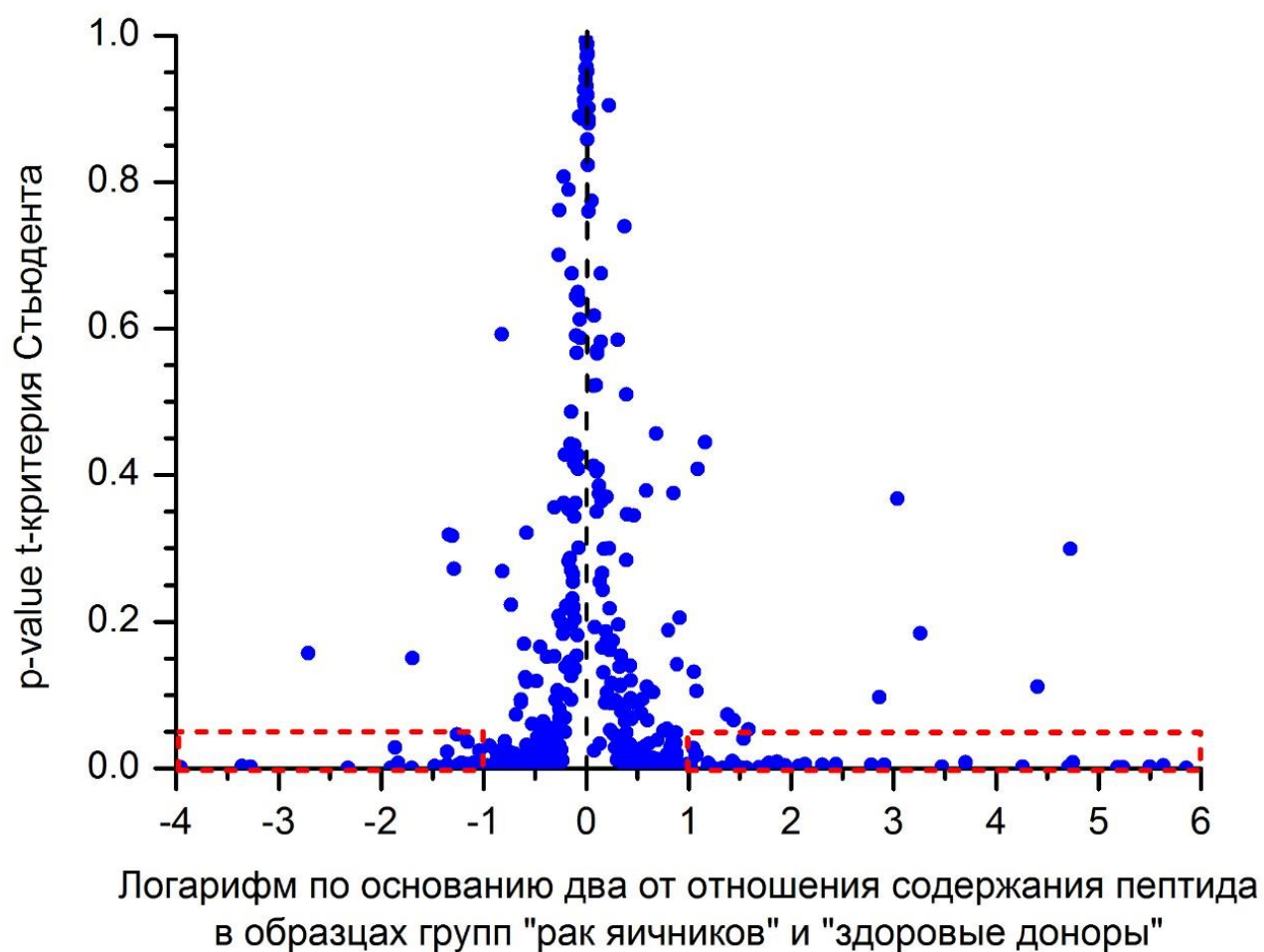


Рисунок 15: График распределения пептидов, измеренных количественным безметочным масс-спектрометрическим методом SWATH, в зависимости от их содержания в образцах групп «рак яичников» и «контроль», а также в зависимости от значимости пептидов при сравнении групп.

Красным выделены пептиды, содержание которых значимо отличается между группами.

Таблица 14: Список пептидов, имеющих значимое отличие в образцах сыворотки крови пациентов групп «рак яичников» и «здоровые доноры» по t-критерию Стьюдента, а также содержание которых в образцах этих групп пациентов отличается в 10 и более раз.

#	Аминокислотная последовательность пептида	p-value t-критерия Стьюдента	Содержание пептида в образцах сыворотки крови пациентов		Десятичный логарифм отношения содержания пептида в образцах групп «рак яичников» и «здоровые доноры», соответственно
			для группы «рак яичников»	для группы «здоровые доноры»	
1	RSFFSFLGEAFD	3.29E-04	2.78E+07	4.80E+05	1.76
2	VTTVASHTSDSDVPSG VTEVVVKLFDSD	3.30E-03	3.18E+06	6.40E+04	1.70
3	SFFSFLGEAFDGARD	2.33E-03	6.26E+06	1.38E+05	1.66
4	RSFFSFLGEAFDGARD	8.06E-04	1.07E+07	2.84E+05	1.58
5	RSFFSFLGEAFDGAR	1.35E-03	9.83E+05	2.70E+04	1.56
6	SLADQAANEWGRSGK D	8.20E-03	9.95E+05	3.70E+04	1.43
7	PNHFRPAGLPEKY	1.91E-03	1.28E+07	4.90E+05	1.42
8	FFSFLGEAFDGARD	1.91E-03	5.32E+06	2.04E+05	1.42
9	RSFFSFLGEAFDGARD MWR	1.94E-03	1.07E+05	5.56E+03	1.28
10	DRVKDLATVYVDVLK DSG	1.08E-03	2.20E+04	3.41E+05	-1.19
11	ADSGEGDFLAEGGGV R	7.62E-03	9.06E+05	6.95E+04	1.11
12	SGEGDFLAEGGGVR	1.94E-03	5.51E+05	4.96E+04	1.05
13	PETEEAALVAIDYINQ NLPWGYKHTL	2.76E-03	3.15E+04	3.23E+05	-1.01

5. Взаимосвязь результатов МАЛДИ масс-спектрометрического профилирования и количественного анализа на основе метода SWATH

Для всех МАЛДИ масс-спектрометрических (МС) пиков, вошедших в модели, построенные на основании ГА и ОНС (Таблица 9, Таблица 8), а также для МС пиков, являющихся статистически значимыми для отличия групп «рак яичников» и «здоровые доноры» (Таблица 7), – всего 33 МС пика – мы искали соответствия в списке идентифицированных пептидов, отобранных для количественного анализа методом SWATH. Значения m/z для идентифицированных пептидов брались с учетом протонирования, а также были оценены максимально допустимые отклонения m/z МАЛДИ МС пика от среднего (0,5 Да) и максимально возможная ошибка определения массы масс-спектрометром (200 ppm для Bruker UltraFlex™ и 20 ppm для ABSciex TripleTOF 5600+).

Для 16 МАЛДИ МС пиков были найдены соответствия между m/z МС пиков и массами пептидов, отобранных для анализа методом SWATH (Таблица 15). Для МАЛДИ МС пика m/z 2671.32 было найдено 2 пептида со схожими массами. Для выявления значимых корреляций оценили изменение содержания пептидов при переходе между группами «здоровые доноры» и «рак яичников». Для идентифицированных пептидов мы рассчитали логарифм по основанию два от отношения содержания пептида в образцах групп «рак яичников» и «здоровые доноры», соответственно. Для МАЛДИ МС пиков мы рассчитали логарифм по основанию два от отношения средних значений площадей под МАЛДИ МС пиками для образцов групп «рак яичников» и «здоровые доноры», соответственно. При сопоставленной корреляции результатов МАЛДИ профилирования и идентификации пептидов знаки обоих логарифмов должны совпадать. Сопоставленные корреляции выделены полужирным шрифтом: см. Таблица 15 – всего 15 пептидов из 16 проанализированных.

Таблица 15: Идентифицированные пептиды, отобранные для количественного анализа методом SWATH, чьи массы перекрываются с МАЛДИ масс-спектрометрическими (МС) пиками, вошедших в модели, построенные на основании ГА и ОНС, а также для МС пиков, являющихся статистически значимыми для отличия групп «рак яичников» и «здоровые доноры». Полутолстым шрифтом выделены сонаправленные корреляции результатов МАЛДИ профилирования и анализа методом SWATH.

#	Аминокислотная последовательность пептида	Масса, Да	Логарифм отношения содержания пептидов в группах «рак яичников» и «здоровые доноры», соответственно (по данным SWATH)	m/z МАЛДИ МС	Логарифм отношения площадей под МАЛДИ МС пиками между группами «рак яичников» и «здоровые доноры», соответственно
1	DAHKSEVAHRFKD	1539.67	0.87	1540.03	0.08
2	DEPPQSPWDRVKD	1568.66	0.83	1569.78	0.45
3	DAHKSEVAHRFKDLG	1709.88	1.57	1710.05	-0.05
4	GFDGIPDNVDAALALPAH	1792.97	0.84	1793.98	0.17
5	AATLHTSTAMAAQHGMDDD	1944.08	0.57	1945.11	1.43
6	GVYTLNNEKQWINKAVGD	2049.27	1.65	2050.35	0.41
7	APRIKKIVQKKLAGDESAD	2067.42	1.23	2068.19	0.00
8	GEQRISLPESLKRIPIED	2080.37	0.95	2082.05	0.99
9	GTGQKQIWRIEGSNKVPVD	2112.37	0.57	2113.51	0.06
10	DEPPQSPWDRVKDLATVYVD	2330.54	1.06	2331.10	0.39
11	DEPPQSPWDRVKDLATVYVDVLK	2671.00	1.81	2671.32	0.39
12	EPPQSPWDRVKDLATVYVDVLKD	2671.00	1.51	2671.32	0.39
13	SGRDYVSQFEGSALGKQLNLKLLD N	2753.06	1.24	2753.85	0.34
14	DEPPQSPWDRVKDLATVYVDVLK D	2786.09	0.89	2787.01	0.60
15	RVEPYGENFNKALVQQMEQLRQ KLGPAGD	3453.88	0.39	3455.06	0.42
16	KAKPALEDLRQGLLPVLESFKVSFL SALEEYTKKLNTQ	4306.03	0.36	4308.32	0.33

ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

1. Преимущества, которые дает разработанный метод выделения пептидов из сыворотки крови человека

Усовершенствование стандартного протокола фракционирования образцов сыворотки крови с использованием магнитных микрочастиц с функционализированной поверхностью для последующего масс-спектрометрического анализа образцов было проведено по двум основным причинам. Во-первых, нами было показано, что масс-спектрометрические профили образцов сыворотки крови, получаемые по стандартному протоколу, имеют малое количество пиков и неудовлетворительную воспроизводимость (Таблица 5 и Рисунок 9). Во-вторых, в последние годы появились как теоретические расчеты [171], так и экспериментальные подтверждения [172] того, что часть низкомолекулярных белков и пептидов плазмы и сыворотки крови находятся в ассоциированном состоянии с основными белками плазмы крови.

Одним из наиболее простых, на наш взгляд, способов разрушения межмолекулярных связей, удерживающих пептиды на поверхности высоко представленных белков крови, является нагревание. Предварительные исследования показали, что данный способ не вносит изменений или модификации в аминокислотный состав пептидов и совместим с последующим масс-спектрометрическим анализом. Кроме того, подобранное время прогревания образца в течение 15 минут, на наш взгляд, не является критичным для общего времени анализа.

Сравнение масс-спектрометрических профилей образцов сыворотки крови, полученных по модифицированному и стандартному протоколам, показало существенные преимущества первого. После стадии прогревания элюатов на порядок возрастает число детектируемых пиков, увеличивается общая интенсивность спектра (Таблица 5 и Рисунок 2). Одновременно улучшается воспроизводимость масс-спектрометрических профилей образцов (Рисунок 10).

Мы предположили, что эти улучшения положительным образом скажутся на выявлении наборов масс-спектрометрических пиков, характерных для конкретной патологии. Такое предположение основано на том, что разрушение пептидно-белковых комплексов не просто увеличивает выборку, в которой можно производить поиск потенциальных маркеров, а увеличивает содержание именно тех пептидов, которые в течение относительно длительного времени накапливались в кровотоке на белках-носителях и могут отражать процессы, происходящие в организме.

Увеличение количества детектируемых масс-спектрометрических пиков также связано с уменьшением содержания высокомолекулярных белков в элюатах после прогревания (Рисунок 4). Падение концентрации высокомолекулярных белков после нагревания образцов происходит, вероятно, в результате их денатурации и последующей необратимой в условиях эксперимента агрегации и/или сорбции белков на стенках пробирок. Мы предполагаем, что стабильность пептидного профиля прогретых образцов обусловлена, как уменьшением сродства пептидов к денатурированным белкам-носителям (особенно в присутствии органических растворителей, содержащихся в элюирующем растворе), так и уменьшением (в результате агрегации и/или сорбции белков) концентрации белков-носителей в растворе.

Еще одним достоинством разработанного метода выделения пептидов является то, что он позволяет нивелировать негативное влияние различных условий получения сыворотки крови на воспроизводимость её пептидного состава (сравните Рисунок 9 и Рисунок 10). Это свойство метода открывает возможности для использования в биомаркерных исследованиях образцов, накопленных в биобанках, создававшихся в течение многих десятков лет.

2. Кислотный гидролиз на стадии прогревания элюатов

Особенностью предложенного метода выделения пептидов сыворотки крови является сопутствующий кислотный гидролиз в буфере для элюции MB-WCX при pH равном 2÷3. При изучении кинетики этой протеолитической химической

реакции было показано, что катализируемый кислотой гидролиз проходит в 100 раз быстрее для остатка аспарагиновой кислоты, чем для любого другого остатка [201]. Для механизма реакции на первой стадии необходимо наличие кислых условий (рН менее 2,1), при этом не важно какая конкретно кислота используется [202, 203]. Было показано, что протеолиз существенно ускоряется при повышении температуры более 108⁰С [203]. В последних работах по данной тематике сообщается о преимуществах использования микроволн для увеличения скорости реакции [204, 205]. В работе Swatkoski с соавторами предлагается использовать данный механизм гидролиза белков наравне с трипсинолизом для изучения протеомов различных организмов, и сообщается о включении энзиматической реакции кислотного гидролиза в перечень параметров для анализа на поисковом алгоритме MASCOT [204].

При идентификации пептидов мы обнаружили, что наибольшее количество фрагментов соответствует белку Apolipoprotein A-I – 851 (Таблица 13, Рисунок 13). Прогреванию Apolipoprotein A-I в буфере для элюции MB-WCX привело к существенному изменению МАЛДИ спектра (Рисунок 5). Для идентифицированных пептидов мы обнаружили большое количество С-концевых аспарагиновых кислот (Рисунок 14), что можно объяснить кислотным гидролизом. С другой стороны, в работе Карр с соавторами показали, что при MS/MS наибольшей интенсивностью обладают фрагменты, образованные при фрагментации по С-концевой аспарагиновой кислоте, а также N-концевому пролину, глицину и серину [206]. Эти данные объясняются как лабильностью аминокислотной связи при соответствующих группах, так и различной способностью удерживать протон. В наших данных, помимо С-концевых аспарагиновых кислот в составе идентифицированных пептидов, мы обнаружили большое количество N-концевых пролинов, серинов и глицинов (Рисунок 14).

На естественные процессы образования пептидов в организме сказывается лабильность связи, образованной определенными аминокислотными остатками. В процессе прогрева в кислом буфере мы способствуем ускорению кислотного

гидролиза, особенно по аспарагиновой кислоте. Мы не можем точно сказать, какое происхождение имеют пептиды, идентифицированные нами в ходе этой работы: какая их часть является нативной, а какие образованы в результате кислотного гидролиза при прогревании. Поскольку наблюдаемый нами паттерн гидролиза существенно шире гидролиза по аспарагиновой кислоте (количество пептидов, образованных в результате гидролиза по аспарагиновой кислоте не превышает 30%; см. Рисунок 14) мы считаем, что часть идентифицированных нами пептидов являются нативными.

Проведенная нами работа использует в качестве ключевой идеи сравнение образцов сыворотки крови от двух групп пациенток: практически здоровых доноров и больных раком яичников. Мы разработали системы рандомизации коллекций образцов и последующих элюатов, систему мониторинга качества масс-спектров, мы проводили все эксперименты параллельно с контрольной и целевой группой образцов. Таким образом, разработанный нами протокол пробоподготовки и последующего анализа позволяет выявлять значимые отличия между пептидными пулами сыворотки крови групп пациентов вне зависимости от происхождения этих пептидов.

В своих работах Petricoin и Liotta подчеркивали, что в ответ на экстенсивный рост и потерю дифференцировки опухолью организм, вероятно, продуцирует специальные белки, или же сама опухоль может продуцировать белки с измененной структурой [56, 60]. Уникальные фрагменты таких белков могут быть обнаружены в кровотоке и применяться в качестве биомаркеров (Рисунок 1). Обнаруженные нами уникальные для опухоли пептиды, могут являться фрагментами белков, экспрессия которых меняется при возникновении и/или развитии опухоли, образованными либо в результате гидролиза в самом организме, либо при кислотном гидролизе на стадии пробоподготовки.

3. Результаты МАЛДИ масс-спектрометрического профилирования образцов сыворотки крови

Разработанный протокол выделения пептидов из сыворотки крови мы использовали для поиска потенциальных биомаркеров рак яичников. На основе Генетического Алгоритма и алгоритма Обучаемой Нейронной Сети нам удалось построить классификационные модели (использующие найденные характерные паттерны пиков) с показателями чувствительности и специфичности близкими к 100% (Таблица 6).

Детальное изучение кросс-специфичности построенных математических моделей (Таблица 10) показало, что они обладают низкой специфичностью по отношению к образцам групп доброкачественных гинекологических патологий, которые не использовались при построении моделей. В таблице 10 представлены данные о количестве образцов той или иной группы пациенток, классифицированном как «здоровые доноры» моделью «рак яичников / здоровые доноры». К примеру, модель распознала, как относящихся к группе «здоровые доноры» 26,1% образцов группы «аденомиз» и по 25% образцов групп «аденомиоз с гистеромиомой» и «ректоцервикальный эндометриоз». Иначе говоря, модель распознала, как «рак яичников» 73,9% образцов группы «аденомиз» и по 75% образцов групп «аденомиоз с гистеромиомой» и «ректоцервикальный эндометриоз». Таким образом, модель, построенная на основании сравнения двух групп пациентов («рак яичников» и «здоровые доноры»), не позволяет проводить эффективную диагностику образцов из патофизиологически гетерогенных групп пациентов.

На первый взгляд, полученные результаты ставят под вопрос использование МАЛДИ масс-спектрометрического профилирования для поиска биомаркеров, специфичных для заболевания. Однако в онкологии диагностик с абсолютной специфичностью не встречается, да и большой необходимости в них сейчас нет. Сегодня гораздо важнее уметь отличать целевое заболевание от нескольких патологий, которые могут его скрывать [6]. Типичным примером является рак и

доброкачественные опухоли той же локализации. Так, к примеру, было показано, что эндометриоз имеет схожий с раком яичников патогенетический механизм и, возможно, общую этиологию. По статистике у 10-15% пациенток с раком яичников так же обнаруживается эндометриоз [207, 208], а в 1-5% эндометриоз перерождается в раковую опухоль [209, 210]. Эти данные показывают необходимость разработки диагностики, способной отличать рак яичников от эндометриоза.

В большинстве случаев добавление дополнительной группы к «номинальному контролю» увеличивает специфичность модели по отношению к этой группе. К примеру, модель «рак яичников / номинальный контроль 3», для построения которой использовали образцы групп «здоровые доноры», «аденомиоз», «эндометриозидные кисты яичников», «ректоцервикальный эндометриоз» и «гиперпластические процессы эндометрия», показала низкую специфичность только по отношению к группе «аденомиоз с гистеромиомой», которая не использовалась при построении модели (Таблица 12). Однако бывают исключения. Так модели «рак яичников / номинальный контроль 2» и «рак яичников / номинальный контроль 3» показали высокую специфичность по отношению к группе «синдром поликистозных яичников», которая не использовалась при построении данных моделей (Таблица 12).

Улучшение специфичности диагностической модели, путем введения дополнительных патологий в группу «номинальный контроль», отрицательно сказывается на чувствительности модели. Чувствительность заметно падает на 6-7% в ряду моделей «рак яичников / здоровые доноры», «рак яичников / номинальный контроль 1», «рак яичников / номинальный контроль 3» (сравните Таблица 6 и Таблица 11). Таким образом, МАЛДИ масс-спектрометрическое профилирование подходит для сравнения ограниченной выборки патологий. К примеру, МАЛДИ профилирование способно отличить рака от доброкачественных опухолей той же локализации. Вероятно, по аналогичной причине единственные к настоящему моменту одобренные FDA в США мультипараметрические тесты,

OVA1 и ROMA, применимы только в случаях, когда образование в яичниках уже выявлено, но не ясно является ли оно злокачественным и как срочно необходима операция [19, 20].

4. Идентификация пептидов в сыворотке крови человека

По результатам LC-MS/MS анализа нам удалось идентифицировать почти восемь тысяч уникальных пептидов (Рисунок 11 и Рисунок 12). По глубине анализа наши результаты находятся на одном уровне с результатами ведущих групп в области протеомики [211-213]. Большое количество идентифицированных пептидов, особенно уникальных для какой-то из групп сравнения, позволяет проводить более детальные маркерные сравнения, как на количественном уровне, так и путем сопоставления данных с МАЛДИ масс-спектрометрическим профилированием.

Для идентификации пептидов мы использовали две поисковые программы (MASCOT и X!Tandem). Более того, идентифицированные пептиды дополнительно валидировали при помощи алгоритмов статистического анализа, заложенных в программе Scaffold. Такой многостадийный аккуратный анализ необходим, поскольку в отличие от протеомных работ, в которых достоверность идентификации пептидов дополнительно подтверждается идентификацией белка предшественника, при пептидном анализе нет каких-либо предположений о необходимости идентифицировать несколько пептидов на белок. Поскольку при пробоподготовке к пептидному анализу стадия гидролиза не используется, идентификация пептидов, относящихся к какому-либо белку, не означает идентификацию белка предшественника. Таким образом, биоинформатический анализ пептидома требует большего количества стадий валидации и по возможности нескольких независимых алгоритмов идентификации.

В некоторых работах исследователи пептидома сыворотки/плазмы крови отмечают идентификацию так называемых «пептидных лестниц», которые могут быть продуктами активности концевых пептидаз [59]. Обнаруженные в нашей работе «пептидные лестницы» для многих белков предшественников являются, с

одной стороны, указанием на достоверность результатов идентификации. Подобно идентификации нескольких пептидов на белок, пептиды, идентифицированные в составе «пептидных лестниц» усиливают достоверности друг друга. С другой стороны, наличие пептидных лестниц может указывать на нативный механизм образования пептидов (Рисунок 13). В своих работах Petricoin исходит из гипотезы, что формирование естественного пептидного пула плазмы крови проходит в два этапа. Вначале белки предшественники расщепляются на относительно крупные пептиды, которые затем подвергаются гидролизу под действием концевых пептидаз [56].

Для многих белков предшественников нам удалось идентифицировать значительное количество уникальных фрагментов (до 851 в случае Apolipoprotein A-I; Таблица 13). Существуют методы количественного анализа белков, применяемый в протеомике, основанные на сравнении числа идентифицированных пептидов [214]. На двадцать наиболее представленных белков плазмы крови приходится 98% от общего количества белка по массе [92]. Если бы основной вклад в формирование идентифицированного нами пептидома вносил кислотный гидролиз, мы бы наблюдали значительную часть этих белков в списке двадцати, представленных наибольшим количеством уникальных идентифицированных пептидов (Таблица 13). Однако в последний список входит только 7 белков из числа 20 мажорных белков плазмы крови.

5. Количественный анализ пептидов в сыворотке крови

Методом количественного безметочного анализа SWATH нам удалось определить относительное содержания 384 уникальных пептидов (191 достоверно по *p*-value *t*-критерия Стьюдента) между образцами групп «здоровые доноры» и «рак яичников» (см. Приложение 1 и Рисунок 15). В пулах группы «рак яичников» было обнаружено 36 пептидов, концентрация которых в 2 или более раз выше, чем в пулах группы «здоровые доноры», и 28 пептидов, концентрация которых в 2 или более раз ниже, чем в пулах группы «здоровые доноры».

Соотнесение данных МАЛДИ масс-спектрометрического профилирования и результатов количественного анализа позволяет выявить пептиды, МС сигнал которых привел к появлению значимых отличий в интенсивности МАЛДИ МС пиков. В работах, посвященных построению классификационных моделей на основе МАЛДИ профилирования, отмечается, что вклад в выявленные пики отличия, скорее всего, вносят сразу несколько пептидных компонент [59]. Нам удалось идентифицировать 16 уникальных пептидов, значение массы которых соотносится со значением m/z МАЛДИ МС пиков, для 15 из них при переходе между группами «рак яичников» и «здоровые доноры» содержание по данным SWATH меняется сонаправлено с аналогичным изменением площади под МАЛДИ МС пиками. Для МС пика m/z 2671,32 было найдено 2 пептидных сиквенса со схожими массами (Таблица 15).

По данным SWATH ни для одного из 16 пептидов содержание не меняется более чем в 10 раз (см. Таблица 15), тогда как среди всех пептидов, измеренных методом SWATH, в пулах группы «рак яичников» было обнаружено 11 пептидов, концентрация которых в 10 или более раз выше, чем в пулах группы «здоровые доноры», и 2 пептида, концентрация которых в 10 или более раз ниже, чем в пулах группы «здоровые доноры» (Таблица 14). Такое большое расхождение, скорее всего, говорит о низком абсолютном содержании большинства пептидов из таблицы 15, раз их большое относительное изменение не позволило сформировать МАЛДИ МС пик, значимый для классификации.

Среди 15 пептидов, которые были характеризованы как специфичные для группы «рак яичников» на основе согласованных данных двух различных подходов, SWATH и МАЛДИ МС профилирования, представлены фрагменты 8 различных белков (Таблица 16). Большая часть белков предшественников относится к белкам, входящим в 20 представленных наибольшим количеством уникальных идентифицированных пептидов (Таблица 13). К ним относятся такие белки, как Аполипопротеины А-I и А-IV, сывороточный альбумин, белки, относящиеся к системе комплемента, и некоторые другие. Многие из белков

предшественников входят в 20 наиболее представленных белков плазмы крови [92]. Обнаружение полипептидных фрагментов мажорных белков плазмы хорошо согласуется с гипотезой «суррогатных маркеров» и последними работами в области идентификации потенциальных пептидных биомаркеров [59].

Таблица 16: Идентифицированные пептиды, содержание которых при переходе между группами «рак яичников» и «здоровые доноры» меняется сонаправлено по данным SWATH и МАЛДИ МС профилирования.

#	Аминокислотная последовательность пептида	Название вероятного белка предшественника	Логарифм отношения содержания пептидов в группах «рак яичников» и «здоровые доноры», соответственно (по данным SWATH)
1	DAHKSEVAHRFKD	Serum albumin	0,42
2	DEPPQSPWDRVKD	Apolipoprotein A-I	0,84
3	GFDGIPDNVDAALALPAH	Vitronectin	2,08
4	AATLHTSTAMAAQHGMDDD	Gelsolin	-7,96
5	GVYTLNNEKQWINKAVGD	Haptoglobin (Fragment)	-0,05
6	APRIKKIVQKKLAGDESAD	Platelet basic protein	0,44
7	GEQRISLPESLKRIPIED	Complement C3	0,76
8	GTGQKQIWRIEGSNKVPVD	Gelsolin	3,88
9	DEPPQSPWDRVKDLATVYVD	Apolipoprotein A-I	1,13
10	DEPPQSPWDRVKDLATVYVDV LK	Apolipoprotein A-I	0,15
11	EPPQSPWDRVKDLATVYVDVL KD	Apolipoprotein A-I	0,50
12	SGRDYVSQFEGSALGKQLNLK LLDN	Apolipoprotein A-I	1,17
13	DEPPQSPWDRVKDLATVYVDV LKD	Apolipoprotein A-I	0,78
14	RVEPYGENFNKALVQQMEQLR QKLGPAGD	Apolipoprotein A-IV	0,06
15	KAKPALEDLRQGLLPVLESFKV SFLSALEEYTKKLNTQ	Apolipoprotein A-I	-1,37

Среди тех белков, которые нельзя отнести к 20 наиболее представленным по данным проекта протеома плазмы крови [92], остаются Витронектин (Vitronectin, уникальный номер по БД UniProtKB: P04004), Гельзолин (Gelsolin, P06396), Platelet basic protein (P02775), а так же Apolipoprotein A-IV (P06727).

Витронектин представляет собой гликопротеин общей массой около 75 кДа [215]. Он принимает участие в фибринолизе, опосредует клеточную адгезию и миграцию, ингибирует мембраноатакующий цитолитический комплекс системы комплемента и связывает некоторые серпины. Также витронектин играет роль в стимуляции пролиферации и роста клеток в таких процессах, как рост и метастазирование злокачественных опухолей, заживление ран и резорбция кости. Синтез витронектина происходит главным образом в гепатоцитах печени. RGD-последовательность (Arg45 – Asp47) в составе витронектина позволяет ему взаимодействовать с интегриновыми рецепторами на поверхности клеток [216]. Таким образом витронектин способствует адгезии клеток к внеклеточному матриксу и их миграции [217], участвуя тем самым в заживлении ран, росте злокачественных новообразований и их метастазировании.

Гельзолин – это актин-связывающий белок, который является ключевым регулятором сборки актиновой нити. Кроме того, гельзолин способен разрывать актиновые нити со 100% эффективностью [218]. Внутри клеток гельзолин локализован в цитозоле и митохондриях. Вне клеток гельзолин обнаружен в плазме крови [219]. Регулируя сборку актинового волокна, гельзолин играет важную роль в формировании подосом, цилиндрических актиновых структур на внешней поверхности плазматической мембраны [220]. Подосомы важны для миграции клеток, а также для инвазии (проникновении клеток одной ткани в структуру другой ткани). Некоторые клетки активно используют особенности данной структуры, к примеру раковые клетки или остеокласты [221]. Гельзолин также способен ингибировать апоптоз путем стабилизации митохондрий [219]. В Таблице 16 указаны два пептида белка гельзолин: Ala424 – Asp442 и Gly443 – Asp461. Важно заметить, что они непосредственно граничат в структуре белка, но их

изменение при переходе между группами «рак яичников» и «здоровые доноры» меняется противоположно направлено. Такое расхождение можно объяснить, разницей в активности белка в опухолевых клетках и в нормальных клетках (к примеру, второй пептид полностью лежит в актин-связывающем, Ca^{2+} -чувствительном домене (Ala434 – Ala782), и данная часть белка лучше защищена от гидролиза, когда белок правильно фолдирован).

Platelet basic protein или C-X-C motif chemokine 7 (CXCL7), или Beta-Thromboglobulin является небольшим цитокином массой 13,9 кДа и относится к СХС семейству хемокинов. Этот белок выделяется в больших количествах из тромбоцитов после их активации [222] и стимулирует различные процессы, включая митогенеза, синтез внеклеточного матрикса, метаболизма глюкозы и синтеза активатора плазминогена.

ЗАКЛЮЧЕНИЕ

В ходе данной работы разработан метод выделения пептидов из сыворотки крови человека. При помощи разработанного метода с использованием магнитных микрочастиц с катионообменной поверхностью проведено фракционирование образцов сыворотки крови практически здоровых доноров, пациенток с подтвержденным диагнозом рак яичников, а также пациенток с доброкачественными гинекологическими заболеваниями. МАЛДИ масс-спектрометрическое профилирование позволило построить классификационные модели со значениями чувствительности и специфичности близкими к 100%. Однако было показано, что МАЛДИ масс-спектрометрическое профилирование подходит для дифференцирования ограниченной выборки патологий и может быть использовано, к примеру, для отличия рака от доброкачественных опухолей той же локализации.

Для идентификации пептидов, потенциальных маркеров рака яичников, а также для их количественного анализа методом SWATH провели LC-MS/MS анализ с использованием высокопроизводительных масс-спектрометров TripleTOF5600+. Данные по идентификации 7996 уникальных пептидов, а также по их количественному анализу представляют практический интерес для возможных дальнейших работ по выявлению биомаркеров на существенно больших группах пациентов. При этом в пулах образцов сыворотки от пациенток больных раком яичников идентифицировано 2058 уникальных пептидов, не идентифицированных в пулах образцов от практически здоровых доноров.

ВЫВОДЫ

1. Разработанные протоколы выделения из сыворотки крови человека пептидно-белковых фракций позволяют на порядок увеличить количество регистрируемых во фракциях сыворотки крови МАЛДИ масс-спектрометрических пиков в области значений m/z менее 6 кДа, а также обеспечивают значительное улучшение качества и воспроизводимости масс-спектрометрических профилей получаемых фракций, по сравнению со стандартными протоколами фракционирования.

2. МАЛДИ масс-спектрометрическое профилирование сывороток крови практически здоровых доноров, пациенток с подтвержденным диагнозом рак яичников, а также пациенток с доброкачественными гинекологическими заболеваниями позволило построить классификационные модели со значениями чувствительности и специфичности близкими к 100%.

3. Показано, что МАЛДИ масс-спектрометрическое профилирование подходит для дифференцирования ограниченной выборки патологий и может быть использовано для отличия рака от доброкачественных опухолей той же локализации.

4. На основании разработанного метода выделения пептидов, в пулах сыворотки крови практически здоровых доноров и пациенток с диагнозом рак яичников идентифицировано 7996 уникальных пептидов, являющихся фрагментами 2068 белков. В пулах образцов сыворотки от пациенток больных раком яичников идентифицировано 2058 уникальных пептидов, не идентифицированных в пулах образцов от практически здоровых доноров. Из них 790 пептидов относятся к белкам, фрагменты которых не идентифицированы в пулах сыворотки от практически здоровых доноров.

5. Количественным безметочным масс-спектрометрическим методом SWATH в пулах образцов сыворотки крови пациенток с подтвержденным диагнозом рак яичников обнаружено 36 пептидов, содержание которых в 2 или более раз выше, чем в пулах сыворотки крови практически здоровых доноров и 28

пептидов, содержание которых в 2 или более раз ниже, чем в пулах образцов от практически здоровых доноров.

6. Среди пептидов, измеренных методом SWATH, выявлено 16 уникальных последовательностей, значение массы которых соответствует значениям m/z МАЛДИ масс-спектрометрических пиков. Для 15 пептидов при переходе между группами «рак яичников» и «здоровые доноры» содержание по данным SWATH меняется сонаправлено с данными МАЛДИ масс-спектрометрического профилирования.

БЛАГОДАРНОСТИ

Я благодарю моего научного руководителя Зиганшина Рустама Хусмановича за все знания и навыки, которые я получил, работая в лаборатории, за подробные консультации по тематике работы, за доверие и поддержку, которые он мне оказал. Я благодарю научных сотрудников лаборатории протеомики Института Биоорганической Химии (ИБХ) РАН и моих коллег за консультации и помощь в работе, за прекрасную атмосферу, которая безусловно способствовала моей работе.

Я благодарю всех моих учителей и преподавателей в Московском Физико-Техническом Институте (Государственный Университет), сотрудников и преподавателей учебно-научного центра ИБХ РАН за все знания и навыки, которые я получил в процессе обучения.

СПИСОК СОКРАЩЕНИЙ И УСЛОВНЫХ ОБОЗНАЧЕНИЙ

RCF	relative centrifugal force (относительная центробежная сила)
MB-HIC	Magnetic Beads based Hydrophobic Interaction Chromatography (магнитные микрочастицы с обращено фазовой поверхностью)
MB-IMAC Cu	Magnetic Beads bases on Immobilized Metal ion Affinity Chromatography on Cuprum (магнитные микрочастицы с иммобилизированными ионами медь на поверхности)
MB-WCX	Magnetic Beads based Weak Cation Exchange Chromatography (магнитные микрочастицы со слабой катионообменной поверхностью)
MB-HIC 8 (18)	магнитные микрочастицы с обращено фазовой поверхностью
MB-IMAC Cu	магнитные микрочастицы с иммобилизированными ионами меди
nano-LC-ESI-MS/MS	тандемная масс-спектрометрия с ионизацией электрораспылением, сопряженная с нанопоточной высокоэффективной жидкостной хроматографией

Q-TOF	квадруполь-времяпролетный масс-спектрометр
SWATH	Sequential Window Acquisition of all Theoretical fragments (Последовательная периодическая регистрация всех теоретических фрагментов)
IDA	Information-Dependent Acquisition (Информационно-зависимая регистрация)
БД	база данных
ВЭЖХ	высокоэффективная жидкостная хроматография
МАЛДИ	матрично-активируемая лазерная десорбция, ионизация

СЛОВАРЬ ТЕРМИНОВ

Чувствительность (sensitivity): процент правильно диагностированных больных раком к их общему числу

Специфичность (specificity): процент правильно определенных здоровых людей по отношению к их общему числу

Заболеваемость: число больных с впервые в жизни установленным диагнозом

СПИСОК ЛИТЕРАТУРЫ

1. Roza G (2006) *Inside the Human Body: Using Scientific and Exponential Notation* (Rosen Publishing Group's PowerKids Press).
2. Блохин НН & Петерсон БЕ (1979) *Клиническая онкология* (Медицина, Москва) p 648.
3. Давыдов МИ & Аксель ЕМ (2008) *Вестник РОНЦ им. Н.Н. Блохина РАМН*.
4. Сайт Противоракового Общества России. (13.06 2015): <http://www.pror.ru/pror.shtml>.
5. Сайт Международной Федерации Акушерства и Гинекологии (FIGO). (14.06 2015): <http://www.figo.org/>.
6. Nossov V, Amneus M, Su F, *et al.* The early detection of ovarian cancer: from traditional methods to proteomics. Can we really do better than serum CA-125? (2008) *Am J Obstet Gynecol* **199**(3):215-223.
7. Jacobs IJ & Menon U Progress and challenges in screening for early detection of ovarian cancer. (2004) *Mol Cell Proteomics* **3**(4):355-366.
8. Visintin I, Feng Z, Longton G, *et al.* Diagnostic markers for early detection of ovarian cancer. (2008) *Clin Cancer Res* **14**(4):1065-1072.
9. Nguyen L, Cardenas-Goicoechea SJ, Gordon P, *et al.* Biomarkers for early detection of ovarian cancer. (2013) *Women's health* **9**(2):171-185; quiz 186-177.
10. Ueland FR, DePriest PD, Pavlik EJ, *et al.* Preoperative differentiation of malignant from benign ovarian tumors: the efficacy of morphology indexing and Doppler flow sonography. (2003) *Gynecol Oncol* **91**(1):46-50.
11. Kupesic S & Plavsic BM Early ovarian cancer: 3-D power Doppler. (2006) *Abdominal imaging* **31**(5):613-619.
12. Jacobs IJ, Skates SJ, MacDonald N, *et al.* Screening for ovarian cancer: a pilot randomised controlled trial. (1999) *Lancet* **353**(9160):1207-1210.
13. Kobel M, Kalloger SE, Boyd N, *et al.* Ovarian carcinoma subtypes are different diseases: implications for biomarker studies. (2008) *PLoS medicine* **5**(12):e232.
14. Muramatsu T, Mukai M, Sato S, *et al.* Clinical usefulness of serum and immunohistochemical markers in patients with stage Ia and Ic ovarian cancer. (2005) *Oncol Rep* **14**(4):861-865.
15. Bailey CL, Ueland FR, Land GL, *et al.* The malignant potential of small cystic ovarian tumors in women over 50 years of age. (1998) *Gynecol Oncol* **69**(1):3-7.
16. Valentin L, Skoog L, & Epstein E Frequency and type of adnexal lesions in autopsy material from postmenopausal women: ultrasound study with histological correlation. (2003) *Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology* **22**(3):284-289.
17. Bast RC, Jr., Klug TL, St John E, *et al.* A radioimmunoassay using a monoclonal antibody to monitor the course of epithelial ovarian cancer. (1983) *N Engl J Med* **309**(15):883-887.
18. Anastasi E, Marchei GG, Viggiani V, *et al.* HE4: a new potential early biomarker for the recurrence of ovarian cancer. (2010) *Tumour Biol* **31**(2):113-119.

19. Fung ET A recipe for proteomics diagnostic test development: the OVA1 test, from biomarker discovery to FDA clearance. (2010) *Clin Chem* **56**(2):327-329.
20. Leung F, Diamandis EP, & Kulasingam V From bench to bedside: discovery of ovarian cancer biomarkers using high-throughput technologies in the past decade. (2012) *Biomarkers in medicine* **6**(5):613-625.
21. Kabawat SE, Bast RC, Jr., Bhan AK, *et al.* Tissue distribution of a coelomic-epithelium-related antigen recognized by the monoclonal antibody OC125. (1983) *International journal of gynecological pathology : official journal of the International Society of Gynecological Pathologists* **2**(3):275-285.
22. Patankar MS, Jing Y, Morrison JC, *et al.* Potent suppression of natural killer cell response mediated by the ovarian tumor marker CA125. (2005) *Gynecol Oncol* **99**(3):704-713.
23. Gubbels JA, Belisle J, Onda M, *et al.* Mesothelin-MUC16 binding is a high affinity, N-glycan dependent interaction that facilitates peritoneal metastasis of ovarian tumors. (2006) *Molecular cancer* **5**(1):50.
24. Theriault C, Pinard M, Comamala M, *et al.* MUC16 (CA125) regulates epithelial ovarian cancer cell growth, tumorigenesis and metastasis. (2011) *Gynecol Oncol* **121**(3):434-443.
25. Zurawski VR, Jr., Orjaseter H, Andersen A, *et al.* Elevated serum CA 125 levels prior to diagnosis of ovarian neoplasia: relevance for early detection of ovarian cancer. (1988) *Int J Cancer* **42**(5):677-680.
26. Skates SJ, Menon U, MacDonald N, *et al.* Calculation of the risk of ovarian cancer from serial CA-125 values for preclinical detection in postmenopausal women. (2003) *J Clin Oncol* **21**(10 Suppl):206s-210s.
27. Andersen MR, Goff BA, Lowe KA, *et al.* Use of a Symptom Index, CA125, and HE4 to predict ovarian cancer. (2010) *Gynecol Oncol* **116**(3):378-383.
28. Escudero JM, Auge JM, Filella X, *et al.* Comparison of serum human epididymis protein 4 with cancer antigen 125 as a tumor marker in patients with malignant and nonmalignant diseases. (2011) *Clin Chem* **57**(11):1534-1544.
29. Autelitano DJ, Raineri L, Knight K, *et al.* Performance of a multianalyte test as an aid for the diagnosis of ovarian cancer in symptomatic women. (2012) *Journal of translational medicine* **10**:45.
30. Drapkin R, von Horsten HH, Lin Y, *et al.* Human epididymis protein 4 (HE4) is a secreted glycoprotein that is overexpressed by serous and endometrioid ovarian carcinomas. (2005) *Cancer Res* **65**(6):2162-2169.
31. Moore RG, Miller MC, Disilvestro P, *et al.* Evaluation of the diagnostic accuracy of the risk of ovarian malignancy algorithm in women with a pelvic mass. (2011) *Obstet Gynecol* **118**(2 Pt 1):280-288.
32. Huhtinen K, Suvitie P, Hiissa J, *et al.* Serum HE4 concentration differentiates malignant ovarian tumours from ovarian endometriotic cysts. (2009) *Br J Cancer* **100**(8):1315-1319.

33. Jacob F, Meier M, Caduff R, *et al.* No benefit from combining HE4 and CA125 as ovarian tumor markers in a clinical setting. (2011) *Gynecol Oncol* **121**(3):487-491.
34. Black DM & Solomon E The search for the familial breast/ovarian cancer gene. (1993) *Trends Genet* **9**(1):22-26.
35. Poeschla E, Corbeau P, & Wong-Staal F Development of HIV vectors for anti-HIV gene therapy. (1996) *Proc Natl Acad Sci U S A* **93**(21):11395-11399.
36. James P Protein identification in the post-genome era: the rapid rise of proteomics. (1997) *Q Rev Biophys* **30**(4):279-331.
37. Moore RE, Kirwan J, Doherty MK, *et al.* Biomarker discovery in animal health and disease: the application of post-genomic technologies. (2007) *Biomark Insights* **2**:185-196.
38. Aebersold R & Mann M Mass spectrometry-based proteomics. (2003) *Nature* **422**(6928):198-207.
39. Baggerman G, Verleyen P, Clynen E, *et al.* Peptidomics. (2004) *J Chromatogr B Analyt Technol Biomed Life Sci* **803**(1):3-16.
40. Chin L, Hahn WC, Getz G, *et al.* Making sense of cancer genomic data. (2011) *Genes Dev* **25**(6):534-555.
41. Chin L, Andersen JN, & Futreal PA Cancer genomics: from discovery science to personalized medicine. (2011) *Nat Med* **17**(3):297-303.
42. Schulz-Knappe P, Schrader M, & Zucht HD The peptidomics concept. (2005) *Comb Chem High Throughput Screen* **8**(8):697-704.
43. Anderson NL The roles of multiple proteomic platforms in a pipeline for new diagnostics. (2005) *Mol Cell Proteomics* **4**(10):1441-1444.
44. Wilkins MR, Sanchez JC, Gooley AA, *et al.* Progress with proteome projects: why all proteins expressed by a genome should be identified and how to do it. (1996) *Biotechnology & genetic engineering reviews* **13**:19-50.
45. Anderson NL & Anderson NG Proteome and proteomics: new technologies, new concepts, and new words. (1998) *Electrophoresis* **19**(11):1853-1861.
46. Soloviev M & Finch P Peptidomics, current status. (2005) *J Chromatogr B Analyt Technol Biomed Life Sci* **815**(1-2):11-24.
47. Engholm-Keller K & Larsen MR Technologies and challenges in large-scale phosphoproteomics. (2013) *Proteomics* **13**(6):910-931.
48. Schrader M & Schulz-Knappe P Peptidomics technologies for human body fluids. (2001) *Trends Biotechnol* **19**(10 Suppl):S55-60.
49. Tammen H, Schulte I, Hess R, *et al.* Peptidomic analysis of human blood specimens: comparison between plasma specimens and serum by differential peptide display. (2005) *Proteomics* **5**(13):3414-3422.
50. Adermann K, John H, Standker L, *et al.* Exploiting natural peptide diversity: novel research tools and drug leads. (2004) *Curr Opin Biotechnol* **15**(6):599-606.
51. Falth M, Skold K, Norrman M, *et al.* SwePep, a database designed for endogenous peptides and mass spectrometry. (2006) *Mol Cell Proteomics* **5**(6):998-1005.

52. Diamandis EP Point: Proteomic patterns in biological fluids: do they represent the future of cancer diagnostics? (2003) *Clin Chem* **49**(8):1272-1275.
53. Anderson NL & Anderson NG The human plasma proteome: history, character, and diagnostic prospects. (2002) *Mol Cell Proteomics* **1**(11):845-867.
54. Di Girolamo F, Alessandroni J, Somma P, *et al.* Pre-analytical operating procedures for serum Low Molecular Weight protein profiling. (2010) *J Proteomics* **73**(3):667-677.
55. Rai AJ & Vitzthum F Effects of preanalytical variables on peptide and protein measurements in human serum and plasma: implications for clinical proteomics. (2006) *Expert Rev Proteomics* **3**(4):409-426.
56. Petricoin EF, Belluco C, Araujo RP, *et al.* The blood peptidome: a higher dimension of information content for cancer biomarker discovery. (2006) *Nat Rev Cancer* **6**(12):961-967.
57. Liotta LA & Petricoin EF Serum peptidome for cancer detection: spinning biologic trash into diagnostic gold. (2006) *J Clin Invest* **116**(1):26-30.
58. Hu L, Ye M, & Zou H Recent advances in mass spectrometry-based peptidome analysis. (2009) *Expert Rev Proteomics* **6**(4):433-447.
59. Villanueva J, Shaffer DR, Philip J, *et al.* Differential exoprotease activities confer tumor-specific serum peptidome patterns. (2006) *J Clin Invest* **116**(1):271-284.
60. Liotta LA & Kohn EC The microenvironment of the tumour-host interface. (2001) *Nature* **411**(6835):375-379.
61. Soloviev M & Finch P Peptidomics: bridging the gap between proteome and metabolome. (2006) *Proteomics* **6**(3):744-747.
62. Zheng X, Baker H, & Hancock WS Analysis of the low molecular weight serum peptidome using ultrafiltration and a hybrid ion trap-Fourier transform mass spectrometer. (2006) *J Chromatogr A* **1120**(1-2):173-184.
63. Murray KK DNA sequencing by mass spectrometry. (1996) *J Mass Spectrom* **31**(11):1203-1215.
64. Kriwacki RW & Siuzdak G Probing protein-protein interactions with mass spectrometry. (2000) *Methods Mol Biol* **146**:223-238.
65. Fuerstenau SD, Benner WH, Thomas JJ, *et al.* Mass Spectrometry of an Intact Virus. (2001) *Angew Chem Int Ed Engl* **40**(3):541-544.
66. Boonen K, Landuyt B, Baggerman G, *et al.* Peptidomics: the integrated approach of MS, hyphenated techniques and bioinformatics for neuropeptide analysis. (2008) *J Sep Sci* **31**(3):427-445.
67. M. Karas, D. Bachmann, U. Bahr, *et al.* Matrix-Assisted Ultraviolet Laser Desorption of Non-Volatile Compounds. (1987) *International Journal of Mass Spectrometry and Ion Processes* **78**:53-68.
68. Koichi Tanaka, Hiroaki Waki, Yutaka Ido, *et al.* Protein and Polymer Analyses up to m/z 100 000 by Laser Ionization Time-of-flight Mass Spectrometry (1988) *Rapid Commun. Mass Spectrom.* **2**(8):151-153.
69. Fenn JB, Mann M, Meng CK, *et al.* Electrospray ionization for mass spectrometry of large biomolecules. (1989) *Science* **246**(4926):64-71.

70. Cho A & Normile D Nobel Prize in Chemistry. Mastering macromolecules. (2002) *Science* **298**(5593):527-528.
71. Kussmann M & Roepstorff P Sample preparation techniques for peptides and proteins analyzed by MALDI-MS. (2000) *Methods Mol Biol* **146**:405-424.
72. Cohen SL & Chait BT Influence of matrix solution conditions on the MALDI-MS analysis of peptides and proteins. (1996) *Anal Chem* **68**(1):31-37.
73. Kebarle P & Ho Y (1997) *Electrospray Ionization Mass Spectrometry: Fundamentals, Instrumentation and Applications* (Wiley, New York) p 577.
74. Stephens WE A Pulsed Mass Spectrometer with Time Dispersion. (1946) *Phys. Rev.* **69**:691.
75. Cotter RJ (1997) *Time-of-flight mass spectrometry: instrumentation and applications in biological research* (American Chemical Society, Washington) p 326.
76. Paul W. & H. S Ein neues Massenspektrometer ohne Magnetfeld. (1953) *Zeitschrift Naturforschung Teil A* **8**.
77. Paul Electromagnetic Traps for Charged and Neutral Particles. (1990) *Reviews of Modern Physics* **63**(3):9.
78. Chris M. Lock & Dyer E Characterisation of High Pressure Quadrupole Collision Cells Possessing Direct Current Axial Fields. (1990) *Rapid Communications in Mass Spectrometry* **13**(5):432–448.
79. Dawson PH (1997) *Quadrupole Mass Spectrometry and Its Applications (AVS Classics in Vacuum Science and Technology)* (American Institute of Physics, Maryland) p 376.
80. March F Quadrupole ion trap mass spectrometry: theory, simulation, recent developments and applications. (1998) *Rapid Communications in Mass Spectrometry* **12**(20):1543–1554.
81. Nikolaev EN, Boldin IA, Jertz R, *et al.* Initial experimental characterization of a new ultra-high resolution FTICR cell with dynamic harmonization. (2011) *J Am Soc Mass Spectrom* **22**(7):1125-1133.
82. Comisarow MB & Marshall AG The early development of Fourier transform ion cyclotron resonance (FT-ICR) spectroscopy. (1996) *J Mass Spectrom* **31**(6):581-585.
83. Zubarev RA & Makarov AA Orbitrap Mass Spectrometry. (2013) *Anal Chem*.
84. Hunt DF, Buko AM, Ballard JM, *et al.* Sequence analysis of polypeptides by collision activated dissociation on a triple quadrupole mass spectrometer. (1981) *Biomed Mass Spectrom* **8**(9):397-408.
85. Roepstorff P & Fohlman J Proposal for a common nomenclature for sequence ions in mass spectra of peptides. (1984) *Biomed Mass Spectrom* **11**(11):601.
86. Biemann K Appendix 5. Nomenclature for peptide fragment ions (positive ions). (1990) *Methods Enzymol* **193**:886-887.
87. Little DP, Speir JP, Senko MW, *et al.* Infrared multiphoton dissociation of large multiply charged ions for biomolecule sequencing. (1994) *Anal Chem* **66**(18):2809-2815.

88. Zubarev RA, Horn DM, Fridriksson EK, *et al.* Electron capture dissociation for structural characterization of multiply charged protein cations. (2000) *Anal Chem* **72**(3):563-573.
89. Budnik BA, Haselmann KF, & Zubarev RA Electron detachment dissociation of peptide di-anions: an electron-hole recombination phenomenon. (2001) *Chemical Physics Letters* **342**:299-302.
90. Syka JE, Coon JJ, Schroeder MJ, *et al.* Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. (2004) *Proc Natl Acad Sci U S A* **101**(26):9528-9533.
91. Zubarev RA, Zubarev AR, & Savitski MM Electron capture/transfer versus collisionally activated/induced dissociations: solo or duet? (2008) *J Am Soc Mass Spectrom* **19**(6):753-761.
92. The Clinical Potential of the Human Plasma Proteome. (14.06 2015):
<http://www.plasmaproteome.org/Presentations/PPI%20slides%20for%20the%20website%20Jan%202003.pdf>.
93. Cappadona S, Baker PR, Cutillas PR, *et al.* Current challenges in software solutions for mass spectrometry-based quantitative proteomics. (2012) *Amino acids* **43**(3):1087-1108.
94. Williamsa JD & Burinsky DJ Mass spectrometric analysis of complex mixtures then and now: the impact of linking liquid chromatography and mass spectrometry. (2001) *Int. J. Mass Spectrom.* **212**(1-3):111-133.
95. Thomas R Recent Developments in LC-MS-MS for the Identification and Measurement of Nanoscale Amounts of Proteins and Peptides. (2001) *Spectroscopy* **16**(1):28 - 37.
96. Ericson C, Phung QT, Horn DM, *et al.* An automated noncontact deposition interface for liquid chromatography matrix-assisted laser desorption/ionization mass spectrometry. (2003) *Anal Chem* **75**(10):2309-2315.
97. Ong SE & Mann M Mass spectrometry-based proteomics turns quantitative. (2005) *Nat Chem Biol* **1**(5):252-262.
98. Flory MR, Griffin TJ, Martin D, *et al.* Advances in quantitative proteomics using stable isotope tags. (2002) *Trends Biotechnol* **20**(12 Suppl):S23-29.
99. Ross PL, Huang YN, Marchese JN, *et al.* Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. (2004) *Mol Cell Proteomics* **3**(12):1154-1169.
100. Foster LJ, De Hoog CL, & Mann M Unbiased quantitative proteomics of lipid rafts reveals high specificity for signaling factors. (2003) *Proc Natl Acad Sci U S A* **100**(10):5813-5818.
101. Schlosser A, Amanchy R, & Otto H Identification of tyrosine-phosphorylation sites in the nuclear membrane protein emerlin. (2006) *The FEBS journal* **273**(14):3204-3215.
102. Geiger T, Cox J, Ostasiewicz P, *et al.* Super-SILAC mix for quantitative proteomics of human tumor tissue. (2010) *Nat Methods* **7**(5):383-385.

103. Stewart, II, Thomson T, & Figeys D 18O labeling: a tool for proteomics. (2001) *Rapid Commun Mass Spectrom* **15**(24):2456-2465.
104. Domon B & Aebersold R Mass spectrometry and protein analysis. (2006) *Science* **312**(5771):212-217.
105. Gillet LC, Navarro P, Tate S, *et al.* Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis. (2012) *Mol Cell Proteomics* **11**(6):O111 016717.
106. Lange V, Picotti P, Domon B, *et al.* Selected reaction monitoring for quantitative proteomics: a tutorial. (2008) *Molecular systems biology* **4**:222.
107. Anderson L & Hunter CL Quantitative mass spectrometric multiple reaction monitoring assays for major plasma proteins. (2006) *Mol Cell Proteomics* **5**(4):573-588.
108. Plumb RS, Johnson KA, Rainville P, *et al.* UPLC/MS(E); a new approach for generating molecular fragment information for biomarker structure elucidation. (2006) *Rapid Commun Mass Spectrom* **20**(13):1989-1994.
109. Vapnik VN (1998) *Statistical Learning Theory (Adaptive and Learning Systems for Signal Processing, Communications and Control Series)*.
110. Holland JH (1992) *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence* (The MIT Press).
111. Hammer B, Strickert M, & Villmann T Supervised Neural Gas with General Similarity Measure. (2005) *Neural Processing Letters* **21**:21-44.
112. Kapp E & Schutz F Overview of tandem mass spectrometry (MS/MS) database search algorithms. (2007) *Current protocols in protein science / editorial board, John E. Coligan ... [et al.] Chapter 25*:Unit25 22.
113. Shadforth I, Crowther D, & Bessant C Protein and peptide identification algorithms using MS for use in high-throughput, automated pipelines. (2005) *Proteomics* **5**(16):4082-4095.
114. Xu C & Ma B Software for computational peptide identification from MS-MS data. (2006) *Drug discovery today* **11**(13-14):595-600.
115. Perkins DN, Pappin DJ, Creasy DM, *et al.* Probability-based protein identification by searching sequence databases using mass spectrometry data. (1999) *Electrophoresis* **20**(18):3551-3567.
116. Eng JK, McCormack AL, & Yates JR An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. (1994) *J Am Soc Mass Spectrom* **5**(11):976-989.
117. Sadygov RG & Yates JR, 3rd A hypergeometric probability model for protein identification and validation using tandem mass spectral data and protein sequence databases. (2003) *Anal Chem* **75**(15):3792-3798.
118. Lam H, Deutsch EW, Edde JS, *et al.* Development and validation of a spectral library searching method for peptide identification from MS/MS. (2007) *Proteomics* **7**(5):655-667.

119. Stein SE Estimating probabilities of correct identification from results of mass spectral library searches. (1994) *J Am Soc Mass Spectrom* 5(4):316-323.
120. Yates JR, 3rd, Morgan SF, Gatlin CL, *et al.* Method to compare collision-induced dissociation spectra of peptides: potential for library searching and subtractive analysis. (1998) *Anal Chem* 70(17):3557-3565.
121. Craig R, Cortens JC, Fenyo D, *et al.* Using annotated peptide mass spectrum libraries for protein identification. (2006) *J Proteome Res* 5(8):1843-1849.
122. Ma B, Zhang K, Hendrie C, *et al.* PEAKS: powerful software for peptide de novo sequencing by tandem mass spectrometry. (2003) *Rapid Commun Mass Spectrom* 17(20):2337-2342.
123. Frank A & Pevzner P PepNovo: de novo peptide sequencing via probabilistic network modeling. (2005) *Anal Chem* 77(4):964-973.
124. Taylor JA & Johnson RS Sequence database searches via de novo peptide sequencing by tandem mass spectrometry. (1997) *Rapid Commun Mass Spectrom* 11(9):1067-1075.
125. Keller A, Purvine S, Nesvizhskii AI, *et al.* Experimental protein mixture for validating tandem mass spectral analysis. (2002) *OMICS* 6(2):207-212.
126. Keller A, Eng J, Zhang N, *et al.* A uniform proteomics MS/MS analysis platform utilizing open XML file formats. (2005) *Molecular systems biology* 1:2005 0017.
127. Deutsch EW, Mendoza L, Shteynberg D, *et al.* A guided tour of the Trans-Proteomic Pipeline. (2010) *Proteomics* 10(6):1150-1159.
128. Keller A, Nesvizhskii AI, Kolker E, *et al.* Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. (2002) *Anal Chem* 74(20):5383-5392.
129. Nesvizhskii AI, Vitek O, & Aebersold R Analysis and validation of proteomic data generated by tandem mass spectrometry. (2007) *Nat Methods* 4(10):787-797.
130. De Bock M, de Seny D, Meuwis MA, *et al.* Challenges for biomarker discovery in body fluids using SELDI-TOF-MS. (2010) *J Biomed Biotechnol* 2010:906082.
131. Liotta LA, Ferrari M, & Petricoin E Clinical proteomics: written in blood. (2003) *Nature* 425(6961):905.
132. Schipper R, Loof A, de Groot J, *et al.* Salivary protein/peptide profiling with SELDI-TOF-MS. (2007) *Ann N Y Acad Sci* 1098:498-503.
133. Schaub S, Wilkins J, Weiler T, *et al.* Urine protein profiling with surface-enhanced laser-desorption/ionization time-of-flight mass spectrometry. (2004) *Kidney Int* 65(1):323-332.
134. Petricoin EF, Ardekani AM, Hitt BA, *et al.* Use of proteomic patterns in serum to identify ovarian cancer. (2002) *Lancet* 359(9306):572-577.
135. Papale M, Pedicillo MC, Thatcher BJ, *et al.* Urine profiling by SELDI-TOF/MS: monitoring of the critical steps in sample collection, handling and analysis. (2007) *J Chromatogr B Analyt Technol Biomed Life Sci* 856(1-2):205-213.
136. Papale M, Pedicillo MC, Di Paolo S, *et al.* Saliva analysis by surface-enhanced laser desorption/ionization time-of-flight mass spectrometry (SELDI-TOF/MS): from sample collection to data analysis. (2008) *Clin Chem Lab Med* 46(1):89-99.

137. Luque-Garcia JL & Neubert TA Sample preparation for serum/plasma profiling and biomarker identification by mass spectrometry. (2007) *J Chromatogr A* **1153**(1-2):259-276.
138. Diamandis EP Analysis of serum proteomic patterns for early cancer diagnosis: drawing attention to potential problems. (2004) *J Natl Cancer Inst* **96**(5):353-356.
139. Battino M, Ferreiro MS, Gallardo I, *et al.* The antioxidant capacity of saliva. (2002) *J Clin Periodontol* **29**(3):189-194.
140. Banez LL, Prasanna P, Sun L, *et al.* Diagnostic potential of serum proteomic patterns in prostate cancer. (2003) *J Urol* **170**(2 Pt 1):442-446.
141. Bertucci F, Birnbaum D, & Goncalves A Proteomics of breast cancer: principles and potential clinical applications. (2006) *Mol Cell Proteomics* **5**(10):1772-1786.
142. Grus FH, Podust VN, Bruns K, *et al.* SELDI-TOF-MS ProteinChip array profiling of tears from patients with dry eye. (2005) *Invest Ophthalmol Vis Sci* **46**(3):863-876.
143. Hu S, Loo JA, & Wong DT Human body fluid proteome analysis. (2006) *Proteomics* **6**(23):6326-6353.
144. Noble JL, Dua RS, Coulton GR, *et al.* A comparative proteinomic analysis of nipple aspiration fluid from healthy women and women with breast cancer. (2007) *Eur J Cancer* **43**(16):2315-2320.
145. Shores KS & Knapp DR Assessment approach for evaluating high abundance protein depletion methods for cerebrospinal fluid (CSF) proteomic analysis. (2007) *J Proteome Res* **6**(9):3739-3751.
146. Villanueva J, Philip J, Chaparro CA, *et al.* Correcting common errors in identifying cancer-specific serum peptide signatures. (2005) *J Proteome Res* **4**(4):1060-1072.
147. Villanueva J, Philip J, Entenberg D, *et al.* Serum peptide profiling by magnetic particle-assisted, automated sample processing and MALDI-TOF mass spectrometry. (2004) *Anal Chem* **76**(6):1560-1570.
148. Hsieh SY, Chen RK, Pan YH, *et al.* Systematical evaluation of the effects of sample collection procedures on low-molecular-weight serum/plasma proteome profiling. (2006) *Proteomics* **6**(10):3189-3198.
149. Baumann S, Ceglarek U, Fiedler GM, *et al.* Standardized approach to proteome profiling of human serum based on magnetic bead separation and matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. (2005) *Clin Chem* **51**(6):973-980.
150. Rai AJ, Stemmer PM, Zhang Z, *et al.* Analysis of Human Proteome Organization Plasma Proteome Project (HUPO PPP) reference specimens using surface enhanced laser desorption/ionization-time of flight (SELDI-TOF) mass spectrometry: multi-institution correlation of spectra and identification of biomarkers. (2005) *Proteomics* **5**(13):3467-3474.
151. Yi J, Kim C, & Gelfand CA Inhibition of intrinsic proteolytic activities moderates preanalytical variability and instability of human plasma. (2007) *J Proteome Res* **6**(5):1768-1781.

152. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. (2001) *Clin Pharmacol Ther* **69**(3):89-95.
153. Jiang X, Ye M, & Zou H Technologies and methods for sample pretreatment in efficient proteome and peptidome analysis. (2008) *Proteomics* **8**(4):686-705.
154. Chertov O, Simpson JT, Biragyn A, *et al.* Enrichment of low-molecular-weight proteins from biofluids for biomarker discovery. (2005) *Expert Rev Proteomics* **2**(1):139-145.
155. Hu L, Li X, Jiang X, *et al.* Comprehensive peptidome analysis of mouse livers by size exclusion chromatography prefractionation and nanoLC-MS/MS identification. (2007) *J Proteome Res* **6**(2):801-808.
156. Alpert AJ & Shukla AK Precipitation of Large, High-Abundance Proteins from Serum with Organic Solvents. (2003) *The 8th Annual Meeting of the Association for Biomolecular Resource Facilities (ABRF 2003) Denver, CO, USA (2003) (Poster P111-W).*
157. Chen J, Anderson M, Misek DE, *et al.* Characterization of apolipoprotein and apolipoprotein precursors in pancreatic cancer serum samples via two-dimensional liquid chromatography and mass spectrometry. (2007) *J Chromatogr A* **1162**(2):117-125.
158. Chertov O, Biragyn A, Kwak LW, *et al.* Organic solvent extraction of proteins and peptides from serum as an effective sample preparation for detection and identification of biomarkers by mass spectrometry. (2004) *Proteomics* **4**(4):1195-1203.
159. Washburn MP, Wolters D, & Yates JR, 3rd Large-scale analysis of the yeast proteome by multidimensional protein identification technology. (2001) *Nat Biotechnol* **19**(3):242-247.
160. Tian R, Zhang H, Ye M, *et al.* Selective extraction of peptides from human plasma by highly ordered mesoporous silica particles for peptidome analysis. (2007) *Angew Chem Int Ed Engl* **46**(6):962-965.
161. Gilar M, Olivova P, Daly AE, *et al.* Two-dimensional separation of peptides using RP-RP-HPLC system with different pH in first and second separation dimensions. (2005) *J Sep Sci* **28**(14):1694-1703.
162. Dowell JA, Heyden WV, & Li L Rat neuropeptidomics by LC-MS/MS and MALDI-FTMS: Enhanced dissection and extraction techniques coupled with 2D RP-RP HPLC. (2006) *J Proteome Res* **5**(12):3368-3375.
163. Li X, Xu S, Pan C, *et al.* Enrichment of peptides from plasma for peptidome analysis using multiwalled carbon nanotubes. (2007) *J Sep Sci* **30**(6):930-943.
164. Tian R, Ren L, Ma H, *et al.* Selective enrichment of endogenous peptides by chemically modified porous nanoparticles for peptidome analysis. (2009) *J Chromatogr A* **1216**(8):1270-1278.
165. Hu L, Zhou H, Li Y, *et al.* Profiling of endogenous serum phosphorylated peptides by titanium (IV) immobilized mesoporous silica particles enrichment and MALDI-TOFMS detection. (2009) *Anal Chem* **81**(1):94-104.

166. Boos K-S & Grimm C-H High-performance liquid chromatography integrated solid-phase extraction in bioanalysis using restricted access precolumn packings. (1999) *TrAC Trends in Analytical Chemistry* **18**(3):175-180.
167. Hu L, Boos KS, Ye M, *et al.* Selective on-line serum peptide extraction and multidimensional separation by coupling a restricted-access material-based capillary trap column with nanoliquid chromatography-tandem mass spectrometry. (2009) *J Chromatogr A* **1216**(28):5377-5384.
168. Soloviev M, Barry R, Scrivener E, *et al.* Combinatorial peptidomics: a generic approach for protein expression profiling. (2003) *J Nanobiotechnology* **1**(1):4.
169. Scrivener E, Barry R, Platt A, *et al.* Peptidomics: A new approach to affinity protein microarrays. (2003) *Proteomics* **3**(2):122-128.
170. Gundry RL, Fu Q, Jelinek CA, *et al.* Investigation of an albumin-enriched fraction of human serum and its albuminome. (2007) *Proteomics Clin Appl* **1**(1):73-88.
171. Geho DH, Liotta LA, Petricoin EF, *et al.* The amplified peptidome: the new treasure chest of candidate biomarkers. (2006) *Curr Opin Chem Biol* **10**(1):50-55.
172. Lowenthal MS, Mehta AI, Frogale K, *et al.* Analysis of albumin-associated peptides and proteins from ovarian cancer patients. (2005) *Clin Chem* **51**(10):1933-1945.
173. Kikuchi S, Honda K, Handa Y, *et al.* Serum albumin-associated peptides of patients with uterine endometrial cancer. (2007) *Cancer Sci* **98**(6):822-829.
174. Hortin GL, Shen RF, Martin BM, *et al.* Diverse range of small peptides associated with high-density lipoprotein. (2006) *Biochem Biophys Res Commun* **340**(3):909-915.
175. Ziganshin R, Arapidi G, Azarkin I, *et al.* New method for peptide desorption from abundant blood proteins for plasma/serum peptidome analyses by mass spectrometry. (2011) *J Proteomics* **74**(5):595-606.
176. Zhou M, Lucas DA, Chan KC, *et al.* An investigation into the human serum "interactome". (2004) *Electrophoresis* **25**(9):1289-1298.
177. Mehta AI, Ross S, Lowenthal MS, *et al.* Biomarker amplification by serum carrier protein binding. (2003) *Dis Markers* **19**(1):1-10.
178. Ziganshin R, Arapidi GP, Azarkin IV, *et al.* [Proteomic technologies for identification of serum potential biomarkers of autoimmune demyelinating polyneuropathies]. (2011) *Bioorg Khim* **37**(1):36-44.
179. Camerini S, Polci ML, Liotta LA, *et al.* A method for the selective isolation and enrichment of carrier protein-bound low-molecular weight proteins and peptides in the blood. (2007) *Proteomics Clin Appl* **1**(2):176-184.
180. Ly L & Wasinger VC Peptide enrichment and protein fractionation using selective electrophoresis. (2008) *Proteomics* **8**(20):4197-4208.
181. Hutchens TW Erratum to: 10th asilomat conference on mass spectrometry: Time of flight mass spectrometry. (1993) *J Am Soc Mass Spectrom* **4**(7):612.

182. Merchant M & Weinberger SR Recent advancements in surface-enhanced laser desorption/ionization-time of flight-mass spectrometry. (2000) *Electrophoresis* **21**(6):1164-1177.
183. Villanueva J & Tempst P OvaCheck: let's not dismiss the concept. (2004) *Nature* **430**(7000):611.
184. Anderson GL, McIntosh M, Wu L, *et al.* Assessing lead time of selected ovarian cancer biomarkers: a nested case-control study. (2010) *J Natl Cancer Inst* **102**(1):26-38.
185. Kobayashi E, Ueda Y, Matsuzaki S, *et al.* Biomarkers for screening, diagnosis, and monitoring of ovarian cancer. (2012) *Cancer Epidemiol Biomarkers Prev* **21**(11):1902-1912.
186. Zhang Y, Guo B, & Bi R Ovarian cancer: biomarker proteomic diagnosis in progress. (2012) *Applied biochemistry and biotechnology* **168**(4):910-916.
187. Tang HY, Beer LA, Tanyi JL, *et al.* Protein isoform-specific validation defines multiple chloride intracellular channel and tropomyosin isoforms as serological biomarkers of ovarian cancer. (2013) *J Proteomics* **89**:165-178.
188. Andrews SJ & Rothnagel JA Emerging evidence for functional peptides encoded by short open reading frames. (2014) *Nature reviews. Genetics* **15**(3):193-204.
189. Hanada K, Higuchi-Takeuchi M, Okamoto M, *et al.* Small open reading frames associated with morphogenesis are hidden in plant genomes. (2013) *Proc Natl Acad Sci U S A* **110**(6):2395-2400.
190. Ma J, Ward CC, Jungreis I, *et al.* Discovery of Human sORF-Encoded Polypeptides (SEPs) in Cell Lines and Tissue. (2014) *J Proteome Res.*
191. Rappsilber J, Mann M, & Ishihama Y Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. (2007) *Nat Protoc* **2**(8):1896-1906.
192. Радзинский ВЕ, Сорокина АВ, Зиганшин РХ, *et al.* Аденомиоз – болезнь загадок и предположений. Перспективы постгеномных исследований. (2011) *Доктор Ру, Часть 2, Эндокринология* **9**(68):28-31.
193. Сорокина А, Радзинский В, Тотчиев Г, *et al.* Потенциальные протеомные маркеры аденомиоза в сыворотке крови. (2010) *Врач* **1**:61-63.
194. Сорокина АВ, Радзинский ВЕ, Зиганшин РХ, *et al.* Новый подход к диагностике аденомиоза с использованием протеомного профилирования сыворотки крови. (2011) *Доктор Ру, Часть 1, Гинекология* **9**(68):5-8.
195. Сорокина АВ, Радзинский ВЕ, Зиганшин РХ, *et al.* Алгоритм диагностики аденомиоза с использованием неинвазивных методов исследования. (2011) *Вестник национального медико-хирургического центра им. Н.И.Пирогова* **1**:124-128.
196. Сорокина АВ, Радзинский ВЕ, Зиганшин РХ, *et al.* Поиск пептидных маркеров гинекологических заболеваний в сыворотке крови с использованием МАЛДИ масс-спектрометрии. (2011) *Вестник РУДН, Серия Медицина, Акушерство и гинекология* **6**:122-130.

197. Сорокина АВ, Радзинский ВЕ, Мустафина ЕА, *et al.* Масс-спектрометрия – новый подход в диагностике аденомиоза и рака тела матки. (2011) *Опухоли женской репродуктивной системы* **2**:65-72.
198. Сорокина АВ, Радзинский ВЕ, Сохова ЗМ, *et al.* Потенциальные протеомные маркеры доброкачественных заболеваний матки в сыворотке крови. (2011) *НПЖ Акушерство и Гинекология* **3**:47-51.
199. Ziganshin R, Alekseev DG, Arapidi GP, *et al.* [Serum proteome profiling for ovarion cancer diagnosis using ClinProt magnetic bead technique and MALDI-TOF-mass-spectrometry]. (2008) *Biomed Khim* **54**(4):408-419.
200. Сорокина А, Зиганшин Р, Арапиди Г, *et al.* Поиск специфичных маркеров рака яичников в сыворотке крови женщин с использованием МАЛДИ масс-спектрометрии. (2012) *Врач* **1**:39-42.
201. Schultz J (1967) Cleavage at Asparfic Acid. *Methods in Enzymology*, ed Hirs CHW), Vol 11, pp 255-263.
202. Li A, Sowder RC, Henderson LE, *et al.* Chemical cleavage at aspartyl residues for protein identification. (2001) *Anal Chem* **73**(22):5395-5402.
203. Inglis AS Cleavage at aspartic acid. (1983) *Methods Enzymol* **91**:324-332.
204. Swatkoski S, Gutierrez P, Wynne C, *et al.* Evaluation of microwave-accelerated residue-specific acid cleavage for proteomic applications. (2008) *J Proteome Res* **7**(2):579-586.
205. Basile F & Hauser N Rapid online nonenzymatic protein digestion combining microwave heating acid hydrolysis and electrochemical oxidation. (2011) *Anal Chem* **83**(1):359-367.
206. Kapp EA, Schutz F, Reid GE, *et al.* Mining a tandem mass spectrometry database to determine the trends and global factors influencing peptide fragmentation. (2003) *Anal Chem* **75**(22):6251-6264.
207. Munksgaard PS & Blaakaer J The association between endometriosis and ovarian cancer: a review of histological, genetic and molecular alterations. (2012) *Gynecol Oncol* **124**(1):164-169.
208. Nezhat F, Datta MS, Hanson V, *et al.* The relationship of endometriosis and ovarian malignancy: a review. (2008) *Fertility and sterility* **90**(5):1559-1570.
209. Kokcu A Relationship between endometriosis and cancer from current perspective. (2011) *Archives of gynecology and obstetrics* **284**(6):1473-1479.
210. Sayasneh A, Tsivos D, & Crawford R Endometriosis and ovarian cancer: a systematic review. (2011) *ISRN obstetrics and gynecology* **2011**:140310.
211. Wisniewski JR, Ostasiewicz P, Dus K, *et al.* Extensive quantitative remodeling of the proteome between normal colon tissue and adenocarcinoma. (2012) *Molecular systems biology* **8**:611.
212. Nagaraj N, Wisniewski JR, Geiger T, *et al.* Deep proteome and transcriptome mapping of a human cancer cell line. (2011) *Molecular systems biology* **7**:548.
213. Beck M, Schmidt A, Malmstroem J, *et al.* The quantitative proteome of a human cell line. (2011) *Molecular systems biology* **7**:549.

214. Ishihama Y, Oda Y, Tabata T, *et al.* Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. (2005) *Mol Cell Proteomics* **4**(9):1265-1272.
215. Schwartz I, Seger D, & Shaltiel S Vitronectin. (1999) *The international journal of biochemistry & cell biology* **31**(5):539-544.
216. D'Souza SE, Ginsberg MH, & Plow EF Arginyl-glycyl-aspartic acid (RGD): a cell adhesion motif. (1991) *Trends in biochemical sciences* **16**(7):246-250.
217. Stefansson S & Lawrence DA The serpin PAI-1 inhibits cell migration by blocking integrin alpha V beta 3 binding to vitronectin. (1996) *Nature* **383**(6599):441-443.
218. Sun HQ, Yamamoto M, Mejillano M, *et al.* Gelsolin, a multifunctional actin regulatory protein. (1999) *J Biol Chem* **274**(47):33179-33182.
219. Koya RC, Fujita H, Shimizu S, *et al.* Gelsolin inhibits apoptosis by blocking mitochondrial membrane potential loss and cytochrome c release. (2000) *J Biol Chem* **275**(20):15343-15349.
220. Varon C, Tatin F, Moreau V, *et al.* Transforming growth factor beta induces rosettes of podosomes in primary aortic endothelial cells. (2006) *Mol Cell Biol* **26**(9):3582-3594.
221. Gimona M, Buccione R, Courtneidge SA, *et al.* Assembly and biological role of podosomes and invadopodia. (2008) *Current opinion in cell biology* **20**(2):235-241.
222. Majumdar S, Gonder D, Koutsis B, *et al.* Characterization of the human beta-thromboglobulin gene. Comparison with the gene for platelet factor 4. (1991) *J Biol Chem* **266**(9):5785-5789.

ПРИЛОЖЕНИЕ 1

Результаты анализа количественным безметочным масс-спектрометрическим методом SWATH. В перечне белков, которые могут быть предшественниками данного идентифицированного пептида, приводятся уникальные номера белков согласно базе данных UniProtKB, скачанной с официального сайта <http://www.uniprot.org/> 14 Марта 2013 года.

Аминокислотная последовательность пептида	Перечень белков, которые могут быть предшественниками данного пептида	Значение p-value t-критерия Стьюдента	Содержание пептида в образцах сыворотки крови		Отношение содержания пептида в образцах групп «рак яичников» и «здоровые доноры», соответственно
			для группы "здоровые доноры"	для группы "рак яичников"	
RSFFSFLGEAFD	P0DJI8	3.29E-04	4.80E+05	2.78E+07	5.81E+01
VTTVASHTSDSDVPSGVTEVVVKLFDS	P10909-2	3.30E-03	6.40E+04	3.18E+06	4.97E+01
SFFSFLGEAFDGARD	P0DJI8	2.33E-03	1.38E+05	6.26E+06	4.53E+01
RSFFSFLGEAFDGARD	P0DJI8	8.06E-04	2.84E+05	1.07E+07	3.78E+01
RSFFSFLGEAFDGAR	P0DJI8	1.35E-03	2.70E+04	9.83E+05	3.64E+01
SLADQAANEWGRSGKD	P0DJI8	8.20E-03	3.70E+04	9.95E+05	2.69E+01
RSFFSFLGEAFDGARDM	P0DJI8	2.99E-01	1.61E+05	4.27E+06	2.65E+01
PNHFRPAGLPEKY	P0DJI8	1.91E-03	4.90E+05	1.28E+07	2.61E+01
FFSFLGEAFDGARD	P0DJI8	1.91E-03	2.04E+05	5.32E+06	2.61E+01
SPWDRVKDLATVYVDVLKDSGRD	P02647	1.11E-01	4.74E+05	1.01E+07	2.12E+01
RSFFSFLGEAFDGARDMWR	P0DJI8	1.94E-03	5.56E+03	1.07E+05	1.92E+01
DRVKDLATVYVDVLKDSG	P02647	1.08E-03	3.41E+05	2.20E+04	6.45E-02
ADSGEGDFLAEGGGVR	P02671	7.62E-03	6.95E+04	9.06E+05	1.30E+01
SGEGDFLAEGGGVR	P02671	1.94E-03	4.96E+04	5.51E+05	1.11E+01
PETEEAALVAIDYINQNLPGYKHTL	C9JV77	2.76E-03	3.23E+05	3.15E+04	9.74E-02
PAFISKTNNHIMKLTKGLIKDALENID	Q8IVC0	1.74E-03	1.63E+05	1.68E+04	1.04E-01
SEAEDASLLSFMQGYMKHAT	B0YIW2	1.84E-01	6.07E+04	5.82E+05	9.58E+00

Аминокислотная последовательность пептида	Перечень белков, которые могут быть предшественниками данного пептида	Значение p-value t-критерия Стьюдента	Содержание пептида в образцах сыворотки крови		Отношение содержания пептида в образцах групп «рак яичников» и «здоровые доноры», соответственно
			для группы "здоровые доноры"	для группы "рак яичников"	
APRIKKIVQKKLAGD	P02775	3.67E-01	7.90E+04	6.48E+05	8.21E+00
DSGEGDFLAEGGGVR	P02671	4.37E-03	1.09E+06	8.19E+06	7.51E+00
ALSSVQESQVAQQARG	B0YIW2	9.67E-02	2.07E+05	1.51E+06	7.29E+00
EGDFLAEGGGVR	P02671	4.67E-03	7.70E+05	5.31E+06	6.89E+00
RAHYDLRHTFMGVVSLGSPSGEVSHPRKT	C9JV77	1.57E-01	2.17E+06	3.31E+05	1.53E-01
DSGEGDFLAEGGGV	P02671	5.29E-03	3.79E+05	2.05E+06	5.42E+00
PETEEAALVAIDYINQNLPGWGYKHTLNQIDEVK	C9JV77	3.81E-04	3.82E+05	7.64E+04	2.00E-01
SPWDRVKDLATVYVDVLKD	P02647	3.11E-03	6.31E+04	3.14E+05	4.97E+00
DLAEGGGVR	P02671	4.60E-03	6.39E+05	3.16E+06	4.94E+00
GDFLAEGGGVR	P02671	4.87E-03	3.78E+05	1.66E+06	4.40E+00
SSSYSKQFTSSTSYNRGDSTFESKSY	P02671	2.37E-03	1.23E+05	5.18E+05	4.22E+00
ALSSVQESQVAQQARGWVTDFGSSLKDY	B0YIW2	3.31E-03	2.88E+04	1.11E+05	3.84E+00
PAYLHYYDPAGAED	P08567	2.38E-04	3.68E+05	9.82E+04	2.67E-01
SETESRGSESGIFTNTKESSSHHPGIAEFPSRG	P02671	2.79E-02	3.09E+06	8.47E+05	2.74E-01
SSSYSKQFTSSTSYNRGD	P02671	8.29E-03	5.47E+05	1.99E+06	3.64E+00
RVEPYGENFNKALVQQMEQLRQKL	P06727	7.39E-03	8.75E+04	2.45E+04	2.81E-01
RHFWQQDEPPQSPWDRVKD	P02647	7.25E-03	4.79E+04	1.65E+05	3.44E+00
MNFRPGVLSSRQLGLPGPPDVPDHA	Q14624-4	9.59E-05	9.81E+05	3.02E+05	3.08E-01
DLRHTFMGVVSLGSPSGEVSHPRKT	C9JV77	1.50E-01	3.02E+05	9.34E+04	3.10E-01
SSSYSKQFTSSTSYNRG	P02671	8.07E-04	1.24E+05	4.01E+05	3.22E+00
EDLRQGGLLPVLESFKVSFLSALEEYTKKLNTQ	P02647	5.25E-02	2.32E+05	6.98E+05	3.01E+00

Аминокислотная последовательность пептида	Перечень белков, которые могут быть предшественниками данного пептида	Значение p-value t-критерия Стьюдента	Содержание пептида в образцах сыворотки крови		Отношение содержания пептида в образцах групп «рак яичников» и «здоровые доноры», соответственно
			для группы "здоровые доноры"	для группы "рак яичников"	
GFSSLKDYWSTVKDKFSEF	B0YIW2	4.33E-04	1.42E+05	4.20E+05	2.96E+00
GLVEKVQAAVGTSAAPVPSDNH	P02649	3.99E-02	1.13E+05	3.27E+05	2.90E+00
AGLVEKVQAAVGTSAAPVPSDNH	P02649	3.48E-04	1.97E+05	5.69E+05	2.89E+00
KAKPALEDLRQGLLPVLESFKVSFLSALEEYTK KLNTQ	P02647	2.73E-03	1.40E+06	5.03E+05	3.59E-01
ALSSVQESQVAQQARGWVTD	B0YIW2	5.12E-03	6.34E+05	1.73E+06	2.74E+00
ALSSVQESQVAQQAR	B0YIW2	7.34E-03	4.16E+04	1.13E+05	2.71E+00
LYSKSTAAMSTYTGIFTDQVLSVLKGEE	K7ER74	6.58E-02	2.51E+05	6.79E+05	2.71E+00
SSSYSKQFTSSTS YNRGDSTFES	P02671	9.84E-03	3.02E+05	8.13E+05	2.69E+00
PDLQVSAMQHQT VLELTETGVEAAAAS AISVA R	B4E1F0	1.47E-03	5.67E+06	2.13E+06	3.77E-01
LRQGLLPVLESFKVSFL	P02647	7.28E-02	2.67E+04	6.95E+04	2.60E+00
EFVSETESRGSESGIFTNTKESSSHHPGIAEFPSR G	P02671	4.04E-03	1.11E+05	4.33E+04	3.90E-01
RVEPYGENFNKALVQQMEQLRQKLGP HAGD	P06727	2.26E-02	5.18E+05	2.02E+05	3.91E-01
RGSESGIFTNTKESSSHHPGIAEFPSRG	P02671	3.18E-01	7.09E+05	2.81E+05	3.96E-01
SEAEDASLLSFMQG	B0YIW2	2.29E-04	4.35E+05	1.09E+06	2.51E+00
IFTNTKESSSHHPGIAEFPSRG	P02671	3.17E-01	6.95E+05	2.81E+05	4.05E-01
PALEDLRQGLLPVLESFKVSFLSALEEYTKKLNT Q	P02647	2.72E-01	1.79E+06	7.31E+05	4.09E-01
DEPPQSPWDRVKDLATVYVDVLKDS	P02647	4.56E-02	4.15E+05	1.73E+05	4.18E-01

Аминокислотная последовательность пептида	Перечень белков, которые могут быть предшественниками данного пептида	Значение p-value t-критерия Стьюдента	Содержание пептида в образцах сыворотки крови		Отношение содержания пептида в образцах групп «рак яичников» и «здоровые доноры», соответственно
			для группы "здоровые доноры"	для группы "рак яичников"	
YPKQTQVSVLPEGGETPLFKQFFKNWRD	P06396	4.62E-03	1.16E+06	4.89E+05	4.22E-01
SEAEDASLLSFMQ	B0YIW2	2.74E-04	7.36E+04	1.74E+05	2.36E+00
GNTEGLQKSLAELGGHLDQQVEEF	P06727	6.65E-03	2.02E+06	8.69E+05	4.30E-01
PLPSKETIEQEKQAGES	Q0P5N8	3.11E-04	4.16E+05	1.80E+05	4.34E-01
AAVGTSAAPVPSDNH	P02649	6.76E-03	8.91E+04	2.04E+05	2.29E+00
FITKMDYPKQTQVSVLPEGGETPLFKQFFKNWRD	P06396	6.17E-03	4.56E+05	2.00E+05	4.39E-01
TFPGFFSPMLGEFVSETESRGSESGIFTNTKESSS HHPGIAEFPSRG	P02671	4.78E-05	2.34E+06	5.33E+06	2.28E+00
ASLLSFMQGYMKHATKTAKD	B0YIW2	4.45E-01	4.32E+05	9.68E+05	2.24E+00
YWSVDPLDRAMAELA	P06396	1.48E-04	5.53E+05	2.47E+05	4.46E-01
GLGLSYLSSHIANVERVPF	P06396	3.56E-02	6.28E+04	2.82E+04	4.48E-01
DGTGQKQIWRIEGSNKVPVD	P06396	1.22E-03	5.88E+04	2.73E+04	4.64E-01
LDDDLQGHVLD	B4E1C2	7.38E-03	2.01E+05	9.38E+04	4.67E-01
GLGLSYLSSHIANVERVPFD	P06396	4.64E-03	1.96E+06	9.16E+05	4.68E-01
PETEEAALVAIDYINQNLPGYKHTLNQID	C9JV77	1.23E-04	1.43E+07	6.71E+06	4.70E-01
QNTKSPLFMGKVVNPTQK	E9KL23	4.08E-01	1.65E+05	3.51E+05	2.13E+00
SEAEDASLLSFMQGYMKHATKTAKDA	B0YIW2	1.05E-01	2.71E+05	5.72E+05	2.11E+00
AVGTSAAPVPSDNH	P02649	1.90E-02	1.78E+05	3.76E+05	2.11E+00
VEKVQAAVGTSAAPVPSDNH	P02649	1.71E-02	1.52E+05	3.19E+05	2.10E+00
PITVIDEIRDLLYIGKDRKNPRED	E7EVA3	7.22E-03	1.90E+05	9.09E+04	4.79E-01

Аминокислотная последовательность пептида	Перечень белков, которые могут быть предшественниками данного пептида	Значение p-value t-критерия Стьюдента	Содержание пептида в образцах сыворотки крови		Отношение содержания пептида в образцах групп «рак яичников» и «здоровые доноры», соответственно
			для группы "здоровые доноры"	для группы "рак яичников"	
PETEEAALVAIDYINQDLPWGYKHTLNQID	B7Z556	1.59E-04	1.44E+07	6.89E+06	4.79E-01
DLRQGLLPVLESFKVSFLSALEEYTKKLNTQ	P02647	1.32E-01	1.34E+05	2.79E+05	2.08E+00
NALQSGNSQESVTEQD	Q7Z3Y4	3.21E-03	1.47E+05	7.08E+04	4.83E-01
GIFTDQVLSVLKGEE	K7ER74	2.73E-02	5.25E+04	1.09E+05	2.07E+00
GLGLSYLSSHIANVERVPFDAATLHTSTAMAAQ HGMDDD	P06396	2.41E-02	3.16E+05	1.53E+05	4.85E-01
FDTASTGKTFFPGFFSPMLGEFVSETESRGSESGIF TNTKESSSHHPGIAEFPSRG	P02671	2.43E-03	1.30E+06	2.64E+06	2.04E+00
VAYYLAWIREHTVS	Q8IZZ5	3.54E-05	4.98E+05	2.49E+05	4.99E-01
YWSVDPLDRAMAELAA	P06396	2.96E-04	6.01E+05	3.02E+05	5.02E-01
KVEQAVETEPEPE	P02649	5.61E-03	9.86E+04	1.95E+05	1.98E+00
KVEQAVETEPEPEL	P02649	4.83E-03	1.25E+05	2.45E+05	1.96E+00
PDQTDGLGLSYLSSHIANVERVPFDAATLHTST AMAAQHGMDDD	P06396	3.74E-03	3.04E+05	1.56E+05	5.12E-01
PKKGHIYQGSEADSVFSGFLIPSA	P02745	3.05E-02	2.37E+05	1.23E+05	5.19E-01
IFTDQVLSVLKGEE	K7ER74	2.51E-03	5.17E+05	9.94E+05	1.92E+00
PPQSPWDRVKDLATVYVDVLKD	P02647	2.05E-01	1.68E+05	3.17E+05	1.88E+00
LATVYVDVLKDSGRDYVSQFEGSALGKQLNLK LLD	P02647	1.33E-03	3.74E+05	6.96E+05	1.86E+00
SEAEDASLLSFMQGYMKHATKTAKDALSSVQE	B0YIW2	1.87E-04	6.33E+05	1.17E+06	1.85E+00
DEPPQSPWDRVKDLATVYVDVLKDSGR	P02647	1.42E-01	7.65E+05	1.41E+06	1.85E+00

Аминокислотная последовательность пептида	Перечень белков, которые могут быть предшественниками данного пептида	Значение p-value t-критерия Стьюдента	Содержание пептида в образцах сыворотки крови		Отношение содержания пептида в образцах групп «рак яичников» и «здоровые доноры», соответственно
			для группы "здоровые доноры"	для группы "рак яичников"	
HSGSTFFKYYLQGAKIPKPEASFSPR	B2RMS9	1.97E-02	3.27E+05	6.03E+05	1.84E+00
GGETPLFKQFFKNWRD	P06396	2.34E-02	8.43E+04	4.58E+04	5.43E-01
EEAGARVQQNVPSGTD	A5PL32	4.88E-02	5.41E+05	9.94E+05	1.84E+00
DAHKSEVAHRFKDLGEENFKALVL	P02768	3.43E-02	6.93E+04	1.27E+05	1.83E+00
GSEMVVAGKLQDRG	Q14624-4	1.12E-02	9.10E+04	4.98E+04	5.47E-01
TPDVSSALDKLKEFGN	B2R526	9.03E-03	2.38E+04	4.35E+04	1.83E+00
TPDVSSALDKLKEFGNTLE	B2R526	1.08E-02	1.26E+05	2.27E+05	1.81E+00
DEPPQSPWDRVKDLATVYVDVLK	P02647	3.75E-01	1.24E+06	2.24E+06	1.81E+00
EEAGARVQQNVPSGTD	A5PL32	2.72E-02	1.42E+05	2.55E+05	1.80E+00
VDSGNDVTDIADD	P00738	1.18E-02	2.37E+06	4.26E+06	1.80E+00
VEGHLSFLEKDLRD	P06727	5.98E-04	2.54E+05	1.42E+05	5.57E-01
GLDEDSRAIEGR	P00734	3.59E-02	4.09E+05	7.30E+05	1.79E+00
LDPEVRPTSAVAA	B0YIW2	1.74E-02	1.72E+05	3.06E+05	1.78E+00
DRVKDLATVYVDVLKDSGRD	P02647	5.92E-01	3.82E+05	2.16E+05	5.65E-01
SEAEDASLLSFMQGYMKHATKTAKDALSSVQE SQVAQQARGWVTD	B0YIW2	2.84E-02	2.68E+06	4.74E+06	1.77E+00
LAIATPLPPTSAHGNVAEGETKPD	P02790	2.69E-01	4.79E+06	2.72E+06	5.67E-01
AATLHTSTAMAAQHGMDDD	P06396	2.15E-03	5.58E+05	3.16E+05	5.67E-01
GTGQKQIWRIEGSNKVPVD	P06396	2.74E-03	9.35E+05	5.30E+05	5.67E-01
SEAEDASLLSFMQGYMKHATKTAKDALSSVQE SQVAQQARG	B0YIW2	1.88E-01	8.22E+05	1.43E+06	1.74E+00

Аминокислотная последовательность пептида	Перечень белков, которые могут быть предшественниками данного пептида	Значение p-value t-критерия Стьюдента	Содержание пептида в образцах сыворотки крови		Отношение содержания пептида в образцах групп «рак яичников» и «здоровые доноры», соответственно
			для группы "здоровые доноры"	для группы "рак яичников"	
AVTL DAGLVYDAYLAPNNLKPVVAEFYGSKE D	B4E1B2	3.68E-02	2.28E+05	1.31E+05	5.76E-01
LRHTFMGVVSLGSPSGEVSH P	C9JV77	3.80E-03	3.16E+04	5.46E+04	1.73E+00
GLLPVLESFKVSFLSALEEYTKKLNTQ	P02647	5.32E-02	2.85E+06	4.93E+06	1.73E+00
AGSAFAVHDLEED	P00738	1.84E-02	3.24E+05	5.58E+05	1.72E+00
MVGELRDLSPDD	Q15828	6.95E-03	4.39E+05	2.56E+05	5.82E-01
ENPFAQGALKSED	P06396	1.50E-03	9.42E+05	5.59E+05	5.93E-01
HPNSPLDEENLTQENQD	G3V5I3	5.07E-02	1.09E+05	1.84E+05	1.68E+00
IALCLKSSPAVITD	B2R7F8	2.20E-02	3.75E+05	2.23E+05	5.94E-01
VLKDSGRDYVSQFEGSALGKQLNLKLLDNWD	P02647	2.23E-01	1.04E+05	6.28E+04	6.02E-01
GVYTLNNEKQWINKAVGD	P00738	1.65E-04	2.09E+04	3.44E+04	1.65E+00
PDAPSPPLGAPGLPPAGSPPD SHVLLAAPP G	C9JV77	6.49E-03	1.29E+06	7.90E+05	6.14E-01
GSESGIFTNTKESSSHHPGIAEFPSRG	P02671	1.94E-02	5.45E+05	3.36E+05	6.16E-01
SPGSGNARPNNPDWGT FEEVSGNVSPGT	P02671	3.63E-03	8.58E+04	5.29E+04	6.17E-01
VLESFKVSFLSALEEYTKKLNTQ	P02647	3.82E-02	4.71E+05	7.64E+05	1.62E+00
PDQTDGLGLSYLSSHIANVERVPFD	P06396	2.24E-04	2.78E+06	1.72E+06	6.20E-01
IAVEWESNGQPENNYKTPPVLDSD	A8K008	2.06E-04	1.65E+05	1.02E+05	6.21E-01
TPDVSSALDKLKEFGNTLED	B2R526	1.07E-02	2.18E+06	3.50E+06	1.61E+00
LATVYVDVLKDSGRDYVSQFEGSALGKQLNLK LLDN	P02647	4.56E-01	4.93E+05	7.91E+05	1.60E+00
TPQEVAEAFSSLTETIEGVDAEDGHGPGEQQK RKIVLD	E7EVA3	7.30E-02	8.48E+05	5.28E+05	6.23E-01

Аминокислотная последовательность пептида	Перечень белков, которые могут быть предшественниками данного пептида	Значение p-value t-критерия Стьюдента	Содержание пептида в образцах сыворотки крови		Отношение содержания пептида в образцах групп «рак яичников» и «здоровые доноры», соответственно
			для группы "здоровые доноры"	для группы "рак яичников"	
IQKVIDQFGE	P00734	1.26E-02	4.49E+05	7.19E+05	1.60E+00
TVVQPSVGAA	C9JV77	6.20E-03	9.08E+05	5.69E+05	6.27E-01
ATEHLSTLSEKAKPALEDLRQGLLPVLESFKVSF LSALEEYTKKLNTQ	P02647	7.35E-03	2.80E+06	4.44E+06	1.59E+00
EPLDDYVNTQGASLF	B2R7F8	1.73E-02	5.52E+04	3.48E+04	6.30E-01
TEHLSTLSEKAKPALEDLRQGLLPVLESFKVSFL SALEEYTKKLNTQ	P02647	1.11E-02	3.27E+06	5.17E+06	1.58E+00
DAHKSEVAHRFKDLG	P02768	1.04E-01	1.66E+05	2.62E+05	1.57E+00
LLPVLESFKVSFLSALEEYTKKLNTQ	P02647	1.34E-02	4.66E+06	7.33E+06	1.57E+00
GFPQVSMFFTHTFPK	P04196	4.02E-05	1.64E+05	1.05E+05	6.41E-01
GLPGPPDVPDHAAYHPF	Q14624-4	8.96E-02	9.67E+05	6.22E+05	6.43E-01
PETEEAALVAIDYINQNLPGWGYKHT	C9JV77	7.22E-03	6.22E+06	4.01E+06	6.44E-01
DEAGSEADHEGTHSTKRGHAKSRPV	P02671	9.35E-02	6.21E+05	4.00E+05	6.45E-01
GRTSAGTRQPQFISRD	D9ZGG2	6.22E-04	2.87E+04	4.39E+04	1.53E+00
NTKSPLFMGKVVNPTQK	E9KL23	1.30E-02	1.01E+05	1.55E+05	1.53E+00
TFMGVVSLGSPSGEVSHPRKT	C9JV77	6.33E-03	2.83E+05	1.85E+05	6.54E-01
AEDGHGPGEQQKRKIVLD	E7EVA3	3.35E-02	4.47E+04	6.81E+04	1.52E+00
PDAPRIKKIVQKKLAGDESA	P02775	1.69E-01	3.33E+05	2.19E+05	6.57E-01
DAHKSEVAHRFKDLGEENFKALVLIIFAQY	P02768	5.10E-04	2.30E+06	1.51E+06	6.58E-01
ADEAGSEADHEGTHSTKRGHAKSRPV	P02671	1.24E-01	4.25E+05	2.81E+05	6.62E-01
TATSEYQTFFNPRTFG	P00734	6.55E-02	1.07E+05	1.62E+05	1.51E+00

Аминокислотная последовательность пептида	Перечень белков, которые могут быть предшественниками данного пептида	Значение p-value t-критерия Стьюдента	Содержание пептида в образцах сыворотки крови		Отношение содержания пептида в образцах групп «рак яичников» и «здоровые доноры», соответственно
			для группы "здоровые доноры"	для группы "рак яичников"	
FPQVSMFFTHTFPK	P04196	4.93E-03	2.43E+05	1.61E+05	6.62E-01
SQEEKTEALTSAKRYIETD	P06396	1.12E-02	1.69E+06	1.12E+06	6.62E-01
EPPQSPWDRVKDLATVYVDVLKD	P02647	1.11E-01	2.07E+06	3.13E+06	1.51E+00
PVLESFKVSFLSALEEYTKKLNT	P02647	6.36E-03	1.78E+05	2.69E+05	1.51E+00
GATYDLEGHPQYLLD	P0C0L5	7.37E-03	2.69E+05	1.79E+05	6.64E-01
TPDVSSALDKLKEFG	B2R526	9.60E-05	8.46E+05	1.27E+06	1.50E+00
TDTGALLFIGKILD	P36955	3.79E-01	6.81E+04	1.02E+05	1.50E+00
SYKMADEAGSEADHEGTHSTKRGHAKSRPV	P02671	3.21E-01	2.67E+05	1.78E+05	6.67E-01
PDAPPSPPLGAPGLPPAGSPPD	C9JV77	3.17E-02	1.81E+07	1.21E+07	6.69E-01
GDKLFGPDLKLVPPMEEDYPQFGSPK	P08697	1.18E-01	1.12E+05	7.46E+04	6.69E-01
EPLDDYVNTQGASLFSVTKKQLGAG	B2R7F8	2.25E-02	8.21E+05	5.51E+05	6.72E-01
LRQGLLPVLESFKVSF	P02647	3.01E-02	1.14E+05	1.70E+05	1.49E+00
RGPLVKYSSDYFQAPSD	Q08380	4.45E-03	2.25E+05	1.51E+05	6.74E-01
EPLDDYVNTQGASLFSVTKKQLG	B2R7F8	3.99E-03	3.43E+05	2.32E+05	6.76E-01
ESFKVSFLSALEEYTKKLNTQ	P02647	1.33E-03	1.24E+06	1.83E+06	1.48E+00
PNWVVLNQGREIVQTMNSD	Q53FR6	7.20E-03	3.35E+05	2.28E+05	6.80E-01
RVVEESELARTAGYGINILGMD	P02748	9.33E-03	7.61E+05	1.12E+06	1.47E+00
SEKAKPALEDLRQGLLPVLESFKVSFLSALEEYTKKLNTQ	P02647	9.45E-02	3.15E+06	4.61E+06	1.46E+00
PELTESSGSASHID	P02748	7.44E-02	6.07E+05	8.81E+05	1.45E+00
APPSPLGAPGLPPAGSPPD	C9JV77	7.38E-03	1.78E+07	1.24E+07	6.94E-01

Аминокислотная последовательность пептида	Перечень белков, которые могут быть предшественниками данного пептида	Значение p-value t-критерия Стьюдента	Содержание пептида в образцах сыворотки крови		Отношение содержания пептида в образцах групп «рак яичников» и «здоровые доноры», соответственно
			для группы "здоровые доноры"	для группы "рак яичников"	
PAKLDQIESVITATSANTQLVLETLAQMDDLQD FLSKDKHAD	P19823	5.99E-02	2.68E+06	1.86E+06	6.94E-01
FVELGTQPATQ	P02652	1.73E-02	3.02E+05	4.35E+05	1.44E+00
AGLVYDAYLAPNNLKPVVAEFYGSKE	B4E1B2	1.44E-02	1.20E+05	8.38E+04	6.96E-01
SEAEDASLLSFMQGYMKH	B0YIW2	1.84E-03	7.97E+04	1.14E+05	1.43E+00
VSFLSALEEYTKKLNTQ	P02647	2.57E-02	1.64E+06	2.35E+06	1.43E+00
LPPTSAHGNVAEGETKPD	P02790	2.38E-03	3.68E+06	2.59E+06	7.03E-01
DVSSALDKLKEFGNTLED	B2R526	2.58E-02	7.28E+05	1.04E+06	1.42E+00
GSEMVVAGKLQDRGPD	Q14624-4	1.35E-03	9.87E+05	6.95E+05	7.04E-01
TPDVSSALDKLK	B2R526	1.53E-04	4.75E+04	6.73E+04	1.42E+00
SEAEDASLLSFMQGYMKHATKTAKD	B0YIW2	4.23E-03	4.42E+06	6.26E+06	1.42E+00
PPTSAHGNVAEGETKPD	P02790	3.53E-03	3.58E+05	2.54E+05	7.08E-01
SRQLGLPGPPDVPDHAAYHPF	Q14624-4	2.06E-02	1.31E+06	9.30E+05	7.10E-01
TQNIFFMSKVTNPQA	G3V5I3	2.12E-02	5.28E+05	7.39E+05	1.40E+00
PAKLDQIESVITATSANTQLVLETLAQMDDLQD FLSKDKHA	P19823	1.19E-01	3.01E+05	2.16E+05	7.15E-01
EPLDDYVNTQGASLFSVTKKQL	B2R7F8	4.16E-02	8.22E+04	5.90E+04	7.18E-01
PPQSPWDRVKDLATVYVDVLKDSGRD	P02647	8.90E-02	8.81E+04	1.22E+05	1.39E+00
GLVYDAYLAPNNLKPVVAEFYGSKE	B4E1B2	5.63E-03	3.27E+05	2.37E+05	7.24E-01
SYKMADEAGSEADHEGTHSTKRGHA	P02671	3.45E-01	2.55E+04	3.51E+04	1.38E+00
LRHTFMGVVSLGSPSGEVSHPRKT	C9JV77	2.61E-02	2.14E+07	1.55E+07	7.25E-01

Аминокислотная последовательность пептида	Перечень белков, которые могут быть предшественниками данного пептида	Значение p-value t-критерия Стьюдента	Содержание пептида в образцах сыворотки крови		Отношение содержания пептида в образцах групп «рак яичников» и «здоровые доноры», соответственно
			для группы "здоровые доноры"	для группы "рак яичников"	
SGRDYVSQFEGSA	P02647	2.21E-02	1.38E+05	1.90E+05	1.38E+00
MGVVSLGSPSGEVSHPRKT	C9JV77	1.26E-02	4.42E+06	3.22E+06	7.28E-01
GYLFQLLRIAD	P04196	2.51E-02	3.98E+06	2.90E+06	7.30E-01
APNNLKPVVAEFYGSKE	B4E1B2	1.66E-01	1.61E+05	1.18E+05	7.33E-01
GQPENNYKTTPPVLD	A8K008	3.72E-02	2.12E+05	1.56E+05	7.36E-01
SDVPSGVTEVVVKLFD	P10909-2	6.71E-02	1.00E+06	1.36E+06	1.36E+00
MNFRPGVLSSRQLGLPGPDVDPDHAAYHPF	Q14624-4	1.97E-02	1.75E+07	2.38E+07	1.36E+00
GYLFQLLRIA	P04196	6.14E-02	4.84E+05	3.57E+05	7.38E-01
EVQLVESGGGLVKPGG	A8K008	1.07E-02	8.31E+04	6.14E+04	7.38E-01
SALEEYTKKLNTQ	P02647	9.52E-02	7.79E+05	1.05E+06	1.35E+00
LEEVKAKVQPYLD	P02647	1.20E-01	3.50E+05	4.73E+05	1.35E+00
SVMPGLKMTMDKTGLLLSD	Q14624-4	6.86E-03	4.94E+04	3.66E+04	7.41E-01
PDAPLQPVTPLQLFEG	P0C0L5	2.17E-02	4.72E+05	3.50E+05	7.42E-01
KVEQAVETEPEPELRQQTEWQSG	P02649	1.40E-01	5.03E+04	6.77E+04	1.35E+00
NSPLDEENLTQENQD	G3V5I3	3.26E-02	7.82E+04	1.05E+05	1.35E+00
FRPGVLSSRQLGLPGPDVDPDHA	Q14624-4	1.22E-02	3.49E+05	2.60E+05	7.44E-01
TATSEYQTFFNPRTFGSGEAD	P00734	7.00E-02	6.54E+04	8.79E+04	1.34E+00
PATYGQFYGGDSYIILYNYRHGGRQGQIIYN	P06396	5.38E-02	1.64E+05	1.22E+05	7.45E-01
AFVLKTPSAAYL	P06396	7.19E-03	6.34E+05	4.74E+05	7.48E-01
NLKPVVAEFYGSKE	B4E1B2	6.41E-02	1.84E+05	1.38E+05	7.50E-01
GNTEGLQKSLAELGGHLD	P06727	3.34E-02	3.55E+05	2.67E+05	7.53E-01

Аминокислотная последовательность пептида	Перечень белков, которые могут быть предшественниками данного пептида	Значение p-value t-критерия Стьюдента	Содержание пептида в образцах сыворотки крови		Отношение содержания пептида в образцах групп «рак яичников» и «здоровые доноры», соответственно
			для группы "здоровые доноры"	для группы "рак яичников"	
AGLVYDAYLAPNNLKPVVAEFYGSKE	B4E1B2	1.14E-02	4.66E+06	3.53E+06	7.58E-01
SFMYDTPQEVAEAFSSLTETIEGVDAEDGHGP GEQQKRKIVLD	E7EVA3	8.81E-02	4.49E+05	5.91E+05	1.32E+00
ATWLSLFSSEESNLGANNYDDYRMD	D9ZGG2	3.47E-01	1.85E+05	2.44E+05	1.32E+00
TYHYLPFSLPHRRPHFFPKSRIV	P10909-2	5.10E-01	1.11E+05	1.46E+05	1.31E+00
PAPEVGASKPEGID	D9ZGG2	5.70E-02	3.91E+05	2.97E+05	7.61E-01
VFLGMFLYEYARRHPD	P02768	7.37E-03	3.74E+05	2.85E+05	7.62E-01
YVSQFEGSALGK	P02647	4.83E-02	1.81E+05	2.37E+05	1.31E+00
MADEAGSEADHEGTHSTKRGHAKSRPV	P02671	2.84E-01	1.22E+05	1.60E+05	1.31E+00
TKSPLFMGKVVNPTQK	E9KL23	4.17E-03	2.31E+05	3.02E+05	1.31E+00
LRHTFMGVVSLGSPSGEVSHPR	C9JV77	5.69E-02	2.18E+06	1.67E+06	7.65E-01
SPLFMGKVVNPTQK	E9KL23	4.73E-02	5.14E+05	6.72E+05	1.31E+00
MNFRPGVLSSRQLGLPGPPDVPDHAA	Q14624-4	4.06E-03	4.21E+05	3.22E+05	7.65E-01
ESFKVSFLSALEEYTKKLNT	P02647	4.31E-02	6.76E+04	8.80E+04	1.30E+00
GQPENNYKTPPVLDSD	A8K008	1.52E-01	7.10E+04	5.46E+04	7.68E-01
SFLSALEEYTKKLNTQ	P02647	1.14E-02	8.36E+05	1.09E+06	1.30E+00
GVLSSRQLGLPGPPDVPDHAAAYHPF	Q14624-4	6.41E-02	6.84E+06	8.88E+06	1.30E+00
SPAYRVPVAVQGEDTVQSLTQGD	P01024	1.98E-03	2.19E+05	2.84E+05	1.30E+00
HTFMGVVSLGSPSGEVSHPRKT	C9JV77	7.39E-01	3.52E+05	4.55E+05	1.29E+00
LVYDAYLAPNNLKPVVAEFYGSKE	B4E1B2	3.00E-02	4.01E+05	3.11E+05	7.75E-01
VFTMPEDYEYTVYD	D9ZGG2	2.66E-02	4.67E+05	3.64E+05	7.79E-01

Аминокислотная последовательность пептида	Перечень белков, которые могут быть предшественниками данного пептида	Значение p-value t-критерия Стьюдента	Содержание пептида в образцах сыворотки крови		Отношение содержания пептида в образцах групп «рак яичников» и «здоровые доноры», соответственно
			для группы "здоровые доноры"	для группы "рак яичников"	
SEAEDASLLSFMQGYMKHATKTAKDAL	B0YIW2	1.79E-02	8.60E+04	1.10E+05	1.28E+00
ALRTHLAPYSDELQRRLA	P02647	3.58E-02	1.12E+05	1.43E+05	1.28E+00
GDVFTMPEDYTYVYDD	D9ZGG2	1.19E-02	1.67E+06	2.12E+06	1.27E+00
SEAEDASLLSFMQGYMKHATKTAK	B0YIW2	1.28E-02	2.32E+05	2.95E+05	1.27E+00
PQVSMFFTHTFPK	P04196	5.11E-02	5.12E+04	4.04E+04	7.88E-01
KVSFLSALEEYTKKLNTQ	P02647	7.76E-02	6.50E+05	8.24E+05	1.27E+00
ALRTHLAPYSD	P02647	1.14E-01	1.32E+06	1.67E+06	1.26E+00
GSPAYRVPVAVQGEDTVQSLTQGD	P01024	2.29E-02	5.71E+05	7.20E+05	1.26E+00
LATVYVDVLKDSGRDYVSQFEGSAL	P02647	1.53E-01	4.33E+04	5.47E+04	1.26E+00
SALGKQLNLKLLD	P02647	1.12E-01	1.43E+05	1.80E+05	1.26E+00
PPYPRSIAQYWL	D9ZGG2	1.49E-02	1.20E+06	9.59E+05	7.97E-01
SFKVSFLSALEEYTKKLNT	P02647	1.38E-01	1.17E+05	1.47E+05	1.25E+00
AYLAPNNLKPVVAEFYGSKE	B4E1B2	4.58E-03	3.64E+06	2.91E+06	8.00E-01
TLEIPGNSDPNMIPDGD	P0C0L5	3.35E-02	4.84E+05	3.87E+05	8.01E-01
SGRDYVSQFEGSALGKQLNLKLLDN	P02647	1.96E-01	1.55E+06	1.94E+06	1.24E+00
SEAEDASLLSFMQGYMKHA	B0YIW2	1.45E-02	2.27E+05	2.83E+05	1.24E+00
PVVAEFYGSKE	B4E1B2	5.43E-02	2.42E+05	1.95E+05	8.05E-01
APRIKKIVQKKLAGDESA	P02775	3.56E-01	3.81E+05	3.07E+05	8.06E-01
SLMPFSPYEPLNFHAM	P10909-2	4.12E-02	5.58E+05	4.50E+05	8.06E-01
NVHSGSTFFKYQLQAKIPKPEA	B2RMS9	5.85E-01	3.11E+05	3.85E+05	1.24E+00
DAHKSEVAHRFKDLGEENFKALVLIFAQ	P02768	1.53E-01	5.82E+04	4.70E+04	8.08E-01

Аминокислотная последовательность пептида	Перечень белков, которые могут быть предшественниками данного пептида	Значение p-value t-критерия Стьюдента	Содержание пептида в образцах сыворотки крови		Отношение содержания пептида в образцах групп «рак яичников» и «здоровые доноры», соответственно
			для группы "здоровые доноры"	для группы "рак яичников"	
APRIKKIVQKKLAGDESAD	P02775	4.21E-02	1.74E+06	2.15E+06	1.23E+00
IHFHSPAFQHPPTEFIREGD	P10909-2	1.14E-02	6.93E+04	8.55E+04	1.23E+00
GPRRYTIAALLSPYSYSTTAVVTNPKE	A6XGL1	9.34E-02	5.81E+05	4.73E+05	8.14E-01
SVTSTFSKLREQLG	P02647	8.80E-02	7.28E+05	8.93E+05	1.23E+00
PQTFYYAVAVVKK	B4E1B2	1.45E-02	3.56E+05	2.91E+05	8.17E-01
GNTEGLQKSLAELG	P06727	2.72E-02	8.60E+04	7.05E+04	8.20E-01
RVKDLATVYVDVLKD	P02647	9.26E-02	1.13E+06	1.38E+06	1.22E+00
GDVFTMPEDYTYVD	D9ZGG2	2.80E-02	6.19E+05	7.53E+05	1.22E+00
PRRYTIAALLSPYSYSTTAVVTNPKE	A6XGL1	1.06E-01	7.14E+05	5.89E+05	8.25E-01
PERLGREGVQKED	P01024	2.08E-01	5.35E+04	4.44E+04	8.30E-01
SYKMADEAGSEADHEGTHST	P02671	1.97E-02	9.67E+04	8.04E+04	8.32E-01
PYTVSHYAVGDE	Q6MZM7	9.16E-02	1.09E+05	1.31E+05	1.20E+00
DEPPQSPWDRVKD	P02647	7.00E-01	2.83E+06	2.35E+06	8.32E-01
FKVSFLSALEEYTKKLNTQ	P02647	4.70E-02	1.21E+06	1.46E+06	1.20E+00
LAIATPLPPTSAHGNVAEGETKPPDP	P02790	3.04E-02	1.51E+05	1.26E+05	8.33E-01
PITVTVPVEV	P10909-2	6.04E-03	2.00E+05	1.67E+05	8.34E-01
DQTVSDNELQEMSNQGSKYVVK	P10909-2	1.74E-01	3.12E+04	3.73E+04	1.20E+00
LEDLRQGLLPVLESFKVSFLSALEEYTKKLNTQ	P02647	7.61E-01	2.96E+05	2.47E+05	8.35E-01
GFDGIPDNVDAALALPAH	D9ZGG2	8.06E-02	1.25E+05	1.04E+05	8.36E-01
GFPRGDKLFGPDLKLVPPMEEDYPQFGSPK	P08697	6.76E-02	4.44E+05	3.71E+05	8.36E-01
NALQSGNSQESVTEQDSKD	Q7Z3Y4	5.00E-02	1.12E+05	9.43E+04	8.39E-01

Аминокислотная последовательность пептида	Перечень белков, которые могут быть предшественниками данного пептида	Значение p-value t-критерия Стьюдента	Содержание пептида в образцах сыворотки крови		Отношение содержания пептида в образцах групп «рак яичников» и «здоровые доноры», соответственно
			для группы "здоровые доноры"	для группы "рак яичников"	
LATVYVDVLKDSGRDYVSQFEG	P02647	1.16E-01	4.22E+05	5.00E+05	1.19E+00
SDVPSGVTEVVVKLFDSD	P10909-2	8.82E-02	9.15E+06	1.08E+07	1.19E+00
PQTFYYAVAVVKKDSG	B4E1B2	2.49E-02	1.79E+05	1.51E+05	8.45E-01
SPMYSITPNILRLESEETMVLEAHD	P01024	1.98E-01	2.83E+06	2.40E+06	8.46E-01
LATVYVDVLKDSGRDYVSQFEGSALG	P02647	5.21E-02	8.74E+04	1.03E+05	1.18E+00
KLFGPDLKLVPMEEDYPQFGSPK	P08697	7.07E-02	6.48E+05	5.50E+05	8.50E-01
DAHKSEVAHRFKDLGEENFKA	P02768	2.18E-01	1.04E+05	1.23E+05	1.17E+00
DQTVSDNELQEMSNQG	P10909-2	1.61E-01	5.34E+05	6.26E+05	1.17E+00
ISETFRRFQMIPLD	Q59E99	1.02E-02	3.35E+05	2.85E+05	8.53E-01
MLGEFVSETESRGSESGIFTNTKESSSHHPGIAEF PSRG	P02671	1.83E-01	7.16E+05	6.12E+05	8.54E-01
THRIHWESASLL	P01024	3.00E-01	6.32E+04	7.39E+04	1.17E+00
SGRDYVSQFEGSALGKQ	P02647	9.05E-01	4.33E+05	5.05E+05	1.17E+00
KVSFLSALEEYTKKLNT	P02647	3.62E-01	3.69E+04	3.17E+04	8.60E-01
LKMTMDKTGLLLSD	Q14624-4	8.07E-01	2.45E+04	2.11E+04	8.61E-01
PRIKKIVQKKLAGDESAD	P02775	9.75E-02	6.97E+05	8.06E+05	1.16E+00
PQTFYYAVAVVKKD	B4E1B2	6.92E-02	3.93E+06	3.40E+06	8.65E-01
LPSLAADFVESKD	P02768	4.96E-02	6.58E+06	5.70E+06	8.66E-01
FQLFKNGVAQEPVHLD	P04217	1.03E-01	1.92E+05	2.21E+05	1.15E+00
GNPEQTPVLKPEEEAPAPEVGASKPEGID	D9ZGG2	1.74E-01	3.25E+05	3.75E+05	1.15E+00
DAHKSEVAHRFKD	P02768	4.28E-01	1.52E+06	1.32E+06	8.68E-01

Аминокислотная последовательность пептида	Перечень белков, которые могут быть предшественниками данного пептида	Значение p-value t-критерия Стьюдента	Содержание пептида в образцах сыворотки крови		Отношение содержания пептида в образцах групп «рак яичников» и «здоровые доноры», соответственно
			для группы "здоровые доноры"	для группы "рак яичников"	
VFTMPEDEYTVYDD	D9ZGG2	1.38E-01	1.52E+06	1.32E+06	8.70E-01
DAHKSEVAHRFKDLGEENFKALVLIA	P02768	3.71E-01	4.43E+05	5.09E+05	1.15E+00
PEQGVTEVTGQYEREKA	B2RMS9	1.01E-01	1.28E+05	1.12E+05	8.70E-01
VPSGVTEVVVKLF	P10909-2	2.21E-01	1.79E+05	1.56E+05	8.72E-01
LATVYVDVLKD	P02647	1.87E-01	5.36E+06	6.12E+06	1.14E+00
ETYVPKEFNAETFTFHAD	P02768	8.90E-02	1.05E+06	1.20E+06	1.14E+00
SYFVELGTQPATQ	P02652	2.99E-01	1.47E+05	1.66E+05	1.13E+00
SGPRRYTIAALLSPYSYSTTAVVTNPKE	A6XGL1	2.82E-01	4.11E+06	3.65E+06	8.88E-01
DEPPQSPWDRVKDLATVYVDVLKD	P02647	7.89E-01	1.93E+07	1.72E+07	8.88E-01
SGRDYVSQFEGSALGKQLNLKLLDNWD	P02647	1.31E-01	1.34E+06	1.51E+06	1.12E+00
PRHRQGPVNLLSD	B2RMS9	3.54E-01	2.94E+06	2.62E+06	8.90E-01
PSSPAFRIEDANLIPPVPDD	Q59E99	1.45E-01	6.03E+05	5.38E+05	8.92E-01
HEFLEVPEAQED	P04217	2.44E-01	4.37E+04	4.89E+04	1.12E+00
PPGAPHTEEEGAED	P41271-2	2.86E-01	5.13E+05	4.59E+05	8.95E-01
DAHKSEVAHRFKDLGE	P02768	1.65E-01	2.50E+05	2.78E+05	1.11E+00
YVSQFEGSALGKQ	P02647	2.66E-01	9.47E+04	1.05E+05	1.11E+00
PGKQLYNVEATSYALLALLQLKD	P01024	4.42E-01	7.39E+04	6.64E+04	8.98E-01
SYKMADEAGSEAD	P02671	3.64E-01	1.79E+05	1.99E+05	1.11E+00
TGALLFIGKILD	P36955	4.86E-01	3.45E+05	3.11E+05	9.02E-01
LLSPYSYSTTAVVTNPKE	A6XGL1	1.96E-01	3.06E+05	2.77E+05	9.03E-01

Аминокислотная последовательность пептида	Перечень белков, которые могут быть предшественниками данного пептида	Значение p-value t-критерия Стьюдента	Содержание пептида в образцах сыворотки крови		Отношение содержания пептида в образцах групп «рак яичников» и «здоровые доноры», соответственно
			для группы "здоровые доноры"	для группы "рак яичников"	
LQAQSKGNPEQTPVLKPEEEAPAPEVGASKPEGID	D9ZGG2	6.75E-01	1.55E+05	1.72E+05	1.11E+00
SPMYSIITPNILRLESEE	P01024	1.26E-01	1.44E+05	1.30E+05	9.04E-01
GPSLTSDLQAQSKGNPEQTPVLKPEEEAPAPEVGASKPEGID	D9ZGG2	9.32E-02	1.48E+05	1.34E+05	9.05E-01
SSLSAVFEHFAMMQRD	D9ZGG2	2.70E-01	5.70E+04	5.16E+04	9.05E-01
DAHKSEVAHRFKDLGEENFKALVLI AF	P02768	6.75E-01	4.29E+05	3.89E+05	9.06E-01
NFRPGVLSSRQLGLPGPPDVPDHAAYHPF	Q14624-4	5.82E-01	2.11E+06	2.33E+06	1.10E+00
PSLTSDLQAQSKGNPEQTPVLKPEEEAPAPEVGASKPEGID	D9ZGG2	1.33E-01	1.27E+05	1.15E+05	9.07E-01
GSKGPLDQLEKGGETAQSAD	Q8IVC0	2.14E-01	2.16E+06	1.96E+06	9.09E-01
VPSGVTEVVVKLFDSD	P10909-2	2.54E-01	2.02E+07	2.22E+07	1.10E+00
SVTSTFSKLREQLGPVTQEFWD	P02647	3.36E-02	9.66E+05	1.06E+06	1.10E+00
PEQTPVLKPEEEAPAPEVGASKPEGID	D9ZGG2	2.31E-01	1.20E+06	1.09E+06	9.13E-01
SLMPFSPYEPLNFHA	P10909-2	2.65E-01	1.04E+06	9.53E+05	9.14E-01
VPSGVTEVVVKLFDSD	P10909-2	3.86E-01	1.64E+06	1.79E+06	1.09E+00
PGVLSSRQLGLPGPPD	Q14624-4	2.54E-01	1.40E+06	1.28E+06	9.15E-01
YSYSTTAVVTNPKE	A6XGL1	3.74E-01	1.43E+05	1.56E+05	1.09E+00
PYTVSHYAVGD	Q6MZM7	2.20E-01	1.36E+05	1.25E+05	9.18E-01
KKLVAASQAALGL	P02768	4.40E-01	3.84E+06	3.53E+06	9.21E-01
GDEELLRFSN	P02760	3.43E-01	4.75E+06	4.38E+06	9.21E-01

Аминокислотная последовательность пептида	Перечень белков, которые могут быть предшественниками данного пептида	Значение p-value t-критерия Стьюдента	Содержание пептида в образцах сыворотки крови		Отношение содержания пептида в образцах групп «рак яичников» и «здоровые доноры», соответственно
			для группы "здоровые доноры"	для группы "рак яичников"	
SPASSKPISINYRTEID	Q6MZM7	4.16E-01	5.94E+04	5.47E+04	9.22E-01
PYSYSTTAVVTNPKE	A6XGL1	4.08E-01	4.48E+05	4.84E+05	1.08E+00
DAHKSEVAHRFKDLGEENFKALVLIAFAQYLQ	P02768	5.66E-01	1.53E+05	1.66E+05	1.08E+00
KLVAASQAALGL	P02768	2.03E-01	5.03E+06	4.66E+06	9.27E-01
PPTSAHGNVAEGETKPDVDTERCS	P02790	1.36E-01	1.08E+05	9.97E+04	9.27E-01
YVSQFEGSALGKQLNLKLLD	P02647	5.70E-01	9.20E+05	9.90E+05	1.08E+00
PDSTGALVEED	P36955	4.05E-01	8.57E+05	9.20E+05	1.07E+00
DRVKDLATVYVDVLK	P02647	3.62E-01	6.16E+05	5.74E+05	9.32E-01
SPMYSIITPNILRLE	P01024	3.50E-01	6.51E+05	6.98E+05	1.07E+00
YVSQFEGSALGKQLNLKLLDNWD	P02647	5.90E-01	1.10E+06	1.02E+06	9.35E-01
PDAPRIKKIVQKKLAGD	P02775	6.44E-01	3.59E+05	3.36E+05	9.35E-01
SLMPFSPYEPL	P10909-2	1.54E-01	6.79E+05	6.37E+05	9.37E-01
GKKLVAASQAALGL	P02768	5.67E-01	2.84E+06	2.66E+06	9.37E-01
EEGKKLVAASQAALGL	P02768	5.23E-01	7.29E+05	7.78E+05	1.07E+00
PSSPAFRIEDANLIPPVPDDKFQD	Q59E99	1.81E-01	3.27E+05	3.07E+05	9.41E-01
DEPPQSPWDRVKDLATVYVD	P02647	1.93E-01	1.17E+07	1.24E+07	1.06E+00
SARLNSQRLVFNRPFILMFIVDNNILFLGKVNRP	P05154	6.17E-01	3.33E+05	3.52E+05	1.06E+00
GEQRISLPESLKRIPIED	P01024	4.27E-01	7.91E+05	7.49E+05	9.46E-01
SPMYSIITPNILRLESEETMVLEAHDAQGD	P01024	4.08E-01	2.99E+06	2.83E+06	9.47E-01
GVTEVVVKLFDS	P10909-2	6.50E-01	3.08E+05	2.92E+05	9.48E-01

Аминокислотная последовательность пептида	Перечень белков, которые могут быть предшественниками данного пептида	Значение p-value t-критерия Стьюдента	Содержание пептида в образцах сыворотки крови		Отношение содержания пептида в образцах групп «рак яичников» и «здоровые доноры», соответственно
			для группы "здоровые доноры"	для группы "рак яичников"	
SPMYSITPNILRLESEETMVLEAHDAQGDVPVT VTVHD	P01024	2.43E-02	2.06E+06	2.17E+06	1.06E+00
PDAPRIKKIVQKKLAGDESAD	P02775	4.13E-01	4.10E+06	4.31E+06	1.05E+00
NPEQTPVLKPEEEAPAEVVGASKPEGID	D9ZGG2	5.22E-01	4.06E+05	4.27E+05	1.05E+00
AVLPQEEEGSGGGQLVTEVTKKED	P02760	3.01E-01	5.80E+04	5.52E+04	9.51E-01
YVSQFEGSALGKQLNLKLLDN	P02647	8.89E-01	5.62E+05	5.36E+05	9.53E-01
YSTTAVVTNPKE	A6XGL1	6.39E-01	2.28E+05	2.17E+05	9.55E-01
NSVFDIFELTGAARKGSGRRLVKGPD	Q59E99	6.12E-01	4.33E+04	4.15E+04	9.58E-01
SHVLLAAPPQHQLHRAHYD	C9JV77	5.87E-01	7.03E+05	6.77E+05	9.63E-01
VPSGVTEVVVKLFD	P10909-2	7.74E-01	2.36E+06	2.45E+06	1.04E+00
YVSQFEGSALGKQLNLKLL	P02647	8.86E-01	1.75E+05	1.71E+05	9.75E-01
MLGEFVSETESRG	P02671	8.91E-01	9.90E+04	9.71E+04	9.81E-01
LPGPPDVPDHAAYHPF	Q14624-4	9.01E-01	3.54E+05	3.59E+05	1.02E+00
DEPPQSPWDRVKDLATVYVDVLKDSGRD	P02647	8.86E-01	2.46E+07	2.50E+07	1.02E+00
SPYSYSTTAVVTNPKE	A6XGL1	8.80E-01	6.39E+05	6.48E+05	1.01E+00
PDAPLQPVTPLQLFEGR	P0C0L5	7.60E-01	1.07E+05	1.09E+05	1.01E+00
EGKKLVAASQAALGL	P02768	9.11E-01	7.24E+05	7.15E+05	9.87E-01
SHTSDSDVPSGVTEVVVKLFD	P10909-2	9.27E-01	4.26E+05	4.21E+05	9.87E-01
PITVTVPVEVSRKNPKFMETVAEKALQEYRKKH REE	P10909-2	9.75E-01	1.21E+06	1.23E+06	1.01E+00
LSPYSYSTTAVVTNPKE	A6XGL1	9.04E-01	2.70E+05	2.67E+05	9.89E-01

Аминокислотная последовательность пептида	Перечень белков, которые могут быть предшественниками данного пептида	Значение p-value t-критерия Стьюдента	Содержание пептида в образцах сыворотки крови		Отношение содержания пептида в образцах групп «рак яичников» и «здоровые доноры», соответственно
			для группы "здоровые доноры"	для группы "рак яичников"	
QLGLPGPPDVPDHAAYHPF	Q14624-4	8.23E-01	4.58E+06	4.63E+06	1.01E+00
NRIPESGGDNSVFD	Q59E99	9.03E-01	5.22E+05	5.27E+05	1.01E+00
PEVGASKPEGID	D9ZGG2	9.19E-01	2.88E+05	2.90E+05	1.01E+00
SYSTTAVVTNPKE	A6XGL1	8.99E-01	6.96E+05	7.02E+05	1.01E+00
PGVLSSRQLGLPGPPDVPDHAAYHPF	Q14624-4	8.58E-01	8.29E+06	8.36E+06	1.01E+00
NRIPESGGDNSVFDIFELTGAARKGSGRRLVKGP D	Q59E99	9.51E-01	4.03E+06	4.06E+06	1.01E+00
TYHYLPFSLPHR	P10909-2	9.11E-01	1.86E+05	1.85E+05	9.93E-01
TPVLKPEEEAPAEVGASKPEGID	D9ZGG2	9.18E-01	2.55E+05	2.57E+05	1.01E+00
LESFKVSFLSALEEYTKKLNTQ	P02647	9.88E-01	1.04E+06	1.05E+06	1.01E+00
GSYAAWLSRD	P0C0L5	9.41E-01	1.02E+05	1.01E+05	9.94E-01
EEAPAEVGASKPEGID	D9ZGG2	9.54E-01	3.86E+04	3.84E+04	9.94E-01
TFGSGEADCGLRPLFEKKSLEDKTERELLESYID GR	P00734	9.54E-01	2.95E+05	2.96E+05	1.00E+00
ISELVYGAQLQPLD	P01008	9.31E-01	3.14E+05	3.16E+05	1.00E+00
ALLSPYSYSTTAVVTNPKE	A6XGL1	9.71E-01	9.55E+04	9.58E+04	1.00E+00
VAFVKHQTVPQNTGGKNPD	B4E1B2	9.57E-01	1.03E+05	1.03E+05	1.00E+00
DAHKSEVAHRFKDLGEENFKALVLIIFAQYL	P02768	9.84E-01	3.95E+05	3.96E+05	1.00E+00
SGRDYVSQFEGSALGKQLNLKLLD	P02647	9.93E-01	1.35E+06	1.35E+06	9.99E-01