

УДК 616-006.6

©Коллектив авторов

ПОИСК ПОТЕНЦИАЛЬНЫХ БИОМАРКЕРОВ РАКА ЖЕЛУДКА МАСС-СПЕКТРОМЕТРИЧЕСКИМ ПРОФИЛИРОВАНИЕМ НИЗКОМОЛЕКУЛЯРНОГО ПРОТЕОМА ПЛАЗМЫ КРОВИ

В.Е. Шевченко, Н.Е. Арноцкая, Е.В. Огородникова, М.М. Давыдов,
М.А. Ибраев, И.Н. Туркин, М.И. Давыдов*

Российский онкологический научный центр имени Н.Н. Блохина
Российской академии медицинских наук,
115478 Москва, Каширское шоссе, д. 24, тел.: (499)3235611;
факс: (499)3242924; эл. почта: vshev@nm.ru

В настоящее время отсутствуют надежные биомаркеры сыворотки/плазмы крови для раннего обнаружения рака желудка (РЖ) – одной из наиболее распространённых злокачественных опухолей. В данном исследовании разработана, унифицирована и апробирована новая методология поиска биомаркеров РЖ – профилирование низкомолекулярного протеома (НМП) плазмы крови (1-17 кДа). Данный подход включал три основных компонента: предварительную подготовку образцов, времяпролетную масс-спектрометрию с матрично-активированной лазерной десорбцией/ионизацией (MALDI-TOF-MS), обработку данных с помощью биоинформационного пакета программ.

Продемонстрированы возможности и перспективность разработанного подхода для обнаружения потенциальных маркеров РЖ при анализе НМП 69 образцов плазмы крови больных РЖ (I-IV стадии) и 238 контрольных образцов. В результате проведенного исследования обнаружены пептиды/полипептиды, которые могут в перспективе использоваться для детектирования этой патологии.

Ключевые слова: биомаркеры, рак желудка, протеом, плазма крови, масс-спектрометрия.

ВВЕДЕНИЕ

Рак желудка (РЖ) занимает четвертое место в мире среди злокачественных новообразований и второе в структуре онкологической смертности после рака лёгких. Метастазы возникают у 80-90% больных раком желудка [1]. Как правило, на ранних стадиях РЖ имеет невыраженные клинические проявления и неспецифические симптомы, что затрудняет точную постановку диагноза и назначение соответствующей терапии. Хотя ряд инструментальных методов, таких как компьютерная томография, эндоскопическое ультразвуковое сканирование, магнитно-резонансная томография и позитронная эмиссионная томография используются для диагностики РЖ, их возможности ограничены. Все эти методы точны

только при средней и распространенной стадии рака желудка. В настоящее время хирургия или эндоскопическая резекция слизистой оболочки желудка с последующим гистологическим анализом биоматериала считаются наиболее точными диагностическими методами для рака желудка. Если РЖ диагностируется на ранней стадии, 5-летняя выживаемость составляет приблизительно 90% [2], тогда как у больных с поздней стадией – менее 10% [3]. Таким образом, ранняя диагностика увеличивает возможность сэкономить время для активной терапии и улучшить уровень выживаемости пациентов. К сожалению, биомаркеры РЖ, такие как раковоэмбриональный антиген (СЕА), карбогидратные антигены СА19–9, СА72–4 и СА50 для сыворотки крови и СЕА, СА19–9 и эмбриональный сульфогликопротеин

* - адресат для переписки

для желудочного сока имеют достаточно низкую специфичность и чувствительность и не используются для ранней диагностики [4].

Таким образом, необходимы новые маркеры для скрининга и ранней диагностики рака желудка. Учитывая гетерогенность РЖ, сложный характер молекулярной, клеточной и клинической информации, необходимой для лучшей диагностики РЖ и его терапии, единичные биомаркеры не могут удовлетворять всем критериям. Необходимо проводить поиск комплексов биомаркеров, которые могли бы использоваться в качестве диагностического инструмента. В связи с этим наблюдается повышенный интерес к применению протеомики для исследований в области клинической диагностики и прогностической медицины. Одна из основных целей протеомики – поиск и идентификация белков, дифференциально экспрессированных (ДЭ) в биологических жидкостях больных и контрольной группы.

Сыворотка/плазма крови человека содержит тысячи пептидов, большинство из которых, как считают, является фрагментами больших белков, частично разрушенных эндогенными протеолитическими ферментами [5, 6]. Сложный низкомолекулярный протеом (НМП), который они создают, может обеспечить новый и надежный отпечаток биологических событий во всем организме.

Масс-спектрометрия (MS) – новая технология, которая привела к быстрому прогрессу в области протеомных исследований. С помощью времяпролетной масс-спектрометрии с матрично-активированной лазерной десорбцией/ионизацией (MALDI-TOF-MS) можно получить профили экспрессии белков в биологических образцах человека и обнаружить потенциальные биомаркеры рака. Недавно показана эффективность этой технологии для открытия белковых маркеров в сыворотке/плазме крови человека [7-11].

В настоящей работе представлены данные по поиску потенциальных маркеров опухолевого роста, присутствующих в плазме крови больных раком желудка. Мы описываем технологическую платформу для одновременного измерения больших количеств пептидов/полипептидов НМП плазмы, включающую предварительную обработку образцов с помощью магнитных микрочастиц с катионообменной поверхностью и MALDI-TOF-MS анализ. В своем исследовании мы исходили из того,

что маркеры РЖ могут секретироваться в кровоток самой опухолью, окружающими её тканями или другими органами и тканями в ответ на опухолевый рост. Проведено сравнительное MALDI-TOF-MS профилирование НМП (МВ<17кДа) образцов плазмы крови больных РЖ и индивидов группы контроля. Детектировано 255 полипептидов, из которых 143 были дифференциально экспрессированными в группах случай/контроль. Используя разработанный подход, мы показали, что плазму крови от больных I-II и III-IV стадии РЖ можно отличить с достаточно высокой чувствительностью и специфичностью от образцов контрольной группы с помощью наборов дискриминационных пептидов/полипептидов.

МЕТОДИКА

Образцы крови

Образцы крови больных РЖ получали из торако-абдоминального отдела НИИ клинической онкологии ФГБУ “РОНЦ имени Н.Н. Блохина” РАМН. В качестве контрольной группы (КГ) использовали образцы от практически здоровых индивидов, полученных в отделении переливания крови НИИ клинической онкологии ФГБУ “РОНЦ имени Н.Н. Блохина” РАМН. От всех больных и здоровых доноров получены информационные согласия на проведение исследований.

Протокол получения и хранения плазмы крови

Забор крови из вены (около 4 мл) проводили в пробирки Vacuta (с антикоагулянтом ЭДТА) натошак у больных до хирургического вмешательства или до начала применения противоопухолевой терапии. Кровь центрифугировали в течение 10 мин при 3000 об/мин (1600 g) при комнатной температуре, полученную плазму аликвотировали по 100 мкл и хранили при -80°C. Перед использованием образцы размораживали при комнатной температуре в течение 15 мин и немедленно обрабатывали.

Клиническая характеристика больных

Клинико-морфологические характеристики больных представлены в таблице 1. Для исследования использовали плазму от 69 больных (47 мужчин, 22 женщины)

Таблица 1. Клинико-морфологическая характеристика больных раком желудка.

Кол-во больных	Пол		Возраст					Глубина опухолевой инвазии (Т)	
	Мужчины	Женщины	Менее 45 лет	45-60 лет	Более 60 лет	Медиана (годы)	Диапазон (годы)	T ₁₋₂	T ₃₋₄
n	47	22	5	27	37	60	19-82	13	56
%	68,1	31,9	7,2	39,1	53,7			18,8	81,2
Кол-во больных	Количество поражённых регионарных лимфатических узлов (N)		Стадия болезни						
	N ⁻	N ⁺	IA	IB	II	IIIA	IIIB	IV (M0)	IV (M1)
n	25	44	7	3	18	18	4	12	7
%	36,2	63,8	10,1	4,3	26,1	26,1	5,8	17,4	10,2
Кол-во больных	Локализация опухоли								
	Тело желудка	Проксимальный отдел	Проксимальный отдел + тело желудка	Антральный отдел	Антральный отдел + тело желудка	Весь желудок	Субтотальный + тотальный рак желудка	Другие	
n	9	16	16	10	2	5	8	3	
%	13,0	23,2	23,2	14,5	2,9	7,3	11,6	4,3	

с гистологически подтверждённым диагнозом рак желудка (I стадия – 10, II стадия – 18, III стадия – 22, IV стадия – 19 больных) и 238 практически здоровых лиц (165 мужчин, 73 женщины). Возраст больных РЖ варьировал от 19 до 82 лет, лиц КГ – от 18 до 57 лет. Образцы больных РЖ и КГ произвольно делили на подгруппы (35 и 120) для построения классификационных моделей (процесс обучения) с помощью биоинформационного пакета ClinProTools™ 2.1 (“Bruker Daltonics”, Германия) и внешней проверки их дискриминационной способности (34 и 118).

Фракционирование образцов плазмы крови

Для предварительного фракционирования плазмы крови использовали набор для профилирования НМП сыворотки/плазмы крови Profiling Kit 100 MB-WCX (“Bruker Daltonics”), содержащий магнитные микрочастицы со слабой катионообменной поверхностью и соответствующий протокол, рекомендованный производителем. Часть полученной фракции пептидов/полипептидов (2 мкл) наносили на стальную полированную 384-луночную MALDI-мишень (Bruker MTP 384 target plate polished steel TF) в 2 повторах, после высушивания на воздухе образец покрывали раствором матрицы (2 мкл). В качестве матрицы применяли смесь

2,5-дигидроксibenзойной (3 мг/мл) и α-циано-гидрокси-коричной (2,4 мг/мл) кислот в смеси метанол–ацетонитрил–вода в соотношении 5:4:1 [7].

Времяпролетная масс-спектрометрия с матрично-активированной лазерной десорбцией/ионизацией

Протеомный анализ проводили на времяпролётном масс-спектрометре Ultraflex II (“Bruker Daltonics”) по методике, разработанной нами ранее [11]. В результате был получен сигнал суммарной интенсивности не менее чем 10⁶.

Анализ масс-спектрометрических данных

Для процессирования и генерации масс-фингерпринта с помощью биоинформационного пакета ClinProTools™ 2.1 обрабатывали все спектры с отношением сигнал/шум > 3 и разрешением 800 в диапазоне *m/z* 1000-17000 Да. При построении классификационных моделей (процесс обучения) с помощью ClinProTools™ 2.1 и внешней проверки их дискриминационной способности образцы больных РЖ и контроля делили пополам, при этом одну часть использовали при построении моделей (обучение), а вторую — при их внешней валидации.

ПОИСК ПОТЕНЦИАЛЬНЫХ МАРКЕРОВ РАКА ЖЕЛУДКА

В результате получали серию нормализованных площадей пиков в каждом спектре, которые использовались ClinProTools™ 2.1 при формировании средних масс-спектров каждого класса и построении моделей, отличающих образцы больных РЖ и контрольной группы.

Методы классической статистики применялись ClinProTools™ 2.1 в комбинации с различными алгоритмами построения классификационных моделей, такими как генетический алгоритм (ГА), метод опорных векторов (МОВ), управляемая нейронная сеть (УНС), алгоритм быстрой классификации (АБК). Разделительную способность модели оценивали перекрёстной и внешней валидацией.

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

В данном исследовании мы проверили и подтвердили гипотезу о том, что НМП плазмы крови больных РЖ отличается от протеома индивидов без рака, и это различие можно использовать для поиска потенциальных маркеров опухолевого роста. Сравнение масс-спектральных профилей случая и контроля, полученных ClinProTools™ 2.1 из серии данных в обучающей фазе эксперимента показало их подобие, однако, ряд ионов, отвечающих молекулам пептидов/полипептидов, имели

различные интенсивности для двух популяций (табл. 2).

Анализ показал, что пики масс-спектра представляли, в основном, однозарядные ионы протонированных пептидов/полипептидов. В обучающем тесте воспроизводимо детектировали 255 различных пиков на спектр при отношении сигнал/шум >3 в диапазоне масс от 1000 до 17000 Да (203 для 110 кДа, 52 для 1017 кДа), что давало матрицу из 255×155 интенсивностей (площадей) пиков для двух классов. Данные статистической обработки, часть из которых представлена в таблице 2, выявили 143 пика с достоверно различающимися уровнями сигналов ($p < 0,01$) в масс-спектрах образцов плазмы больных I-IV стадии РЖ и КГ. Анализ усредненных масс-спектров показал, что интенсивности 102 пиков увеличивались у больных РЖ, а 41 пика – уменьшались ($p < 0,01$). Все эти ионы относились к ДЭ пептидам/полипептидам, которые можно рассматривать в качестве потенциальных маркеров рака желудка.

Наибольший интерес вызывали пептиды/полипептиды, представленные в масс-спектре протонированными ионами, интенсивности которых с наибольшей статистической достоверностью отличались между двумя классами (I-IV стадия РЖ/КГ): ***m/z* 2395** (6,7 vs 12,57), **2342** (19,04 vs 38,04), **2787** (14,67 vs 23,67), **3199** (18,17 vs 27,97), **2820** (54,61 vs 87,72), **8598,96** (2,08 vs 0,93)

Таблица 2. Некоторые результаты статистического анализа усреднённых масс-спектров низкомолекулярного протеома плазмы крови индивидов двух групп (I-IV стадия рака желудка/контроль) после обработки программой ClinProTools™ 2.1.

Mass (<i>m/z</i>), Da	DAve	PTTA	Ave1	Ave2	StdDev1	StdDev2	AUC
2395,47	5,87	7,4E-35	6,7	12,57	2,33	3,73	0,92
2341,79	19,00	5,5E-28	19,04	38,04	8,04	17,23	0,84
2787,36	9,00	1,2E-17	14,67	23,67	6,12	6,18	0,85
3199,14	9,80	1,9E-17	18,17	27,97	6,73	7,14	0,84
2409,82	7,53	1,5E-16	12,06	19,59	5,44	6,21	0,82
2820,94	33,11	2,8E-14	54,61	87,72	26,66	23,44	0,84
1553,40	7,10	4,3E-12	20,59	27,69	6,43	5,45	0,83
2833,79	18,67	1,2E-10	29,58	48,24	18,61	14,52	0,82
1887,73	6,65	6,6E-08	20,33	26,98	7,89	8,25	0,80
8598,96	1,15	1,2E-07	2,08	0,93	1,41	1,36	0,88
1896,57	65,86	2,8E-05	89,37	23,5	111,9	47,33	0,85
1741,32	3,61	3,2E-05	5,22	1,61	6,29	1,17	0,80
8563,22	0,56	0,001	0,95	0,39	1,3	0,27	0,86
6629,49	58,94	0,044	64,16	5,22	231,97	6,26	0,81
6431,41	17,44	0,083	19,13	1,7	79,91	1,91	0,82

Примечание: Mass - *m/z* иона, Да; DAve - разница средних значений площадей пиков для двух классов; PTTA - *p*-величина для *t*-теста; AveN - среднее значение для площади пика класса N; StdDevN - стандартное отклонение значения площади пика для класса N; AUC - площадь под ROC-кривой.

и некоторые другие. Как видно из таблицы 2, эти пики имели высокие значения площади ($AUC \geq 0,80$) под ROC-кривой (*receiver operating characteristic, операционная характеристика приёмника*), характеризующей их дискриминационную способность. Распределение числа пиков масс-спектра в зависимости от диапазона AUC иллюстрирует таблица 3. Анализ полученных данных показал, что из 255 пиков масс-спектра 163 проявляли дискриминационную способность ($AUC \geq 0,60$). Наивысшее значение $AUC=0,92$ относилось к пептиду, которому отвечал ион с m/z 2395.

Таблица 3. Распределение числа пиков масс-спектра в зависимости от диапазона изменения площади под ROC-кривой.

Диапазон изменения площади под ROC-кривой (AUC)	Число пиков масс-спектра
$AUC \geq 0,80$	15
$0,79 \geq AUC \geq 0,75$	27
$0,74 \geq AUC \geq 0,70$	32
$0,69 \geq AUC \geq 0,60$	89

Для иллюстрации дискриминационной способности комбинации из двух таких пиков масс-спектра на рисунке 1 представлено двумерное распределение их площадей (в произвольных единицах) для каждого образца (РЖ/КГ), используемого в процессе обучения ClinProTools™ 2.1. Порядковые номера пиков в спектре и их величины m/z указаны над x- и y- осями. По оси x отложены значения площадей для пика с m/z 2395, по оси y – для пика с m/z 2342. Эллипсами выделены стандартные отклонения этих значений для каждого класса. Как можно видеть из приведенного графика, большая часть значений площадей этих пиков увеличена в образцах КГ (обозначенные кружками) по сравнению с образцами больных РЖ (обозначенные крестиками).

Генерацию моделей проводили в диапазоне масс-спектров от 1000 до 17000 Да. Для предварительной оценки их дискриминационной способности использовали протокол перекрестной проверки, в котором программа ClinProTools™ 2.1 случайно выбирала 2/3 данных в качестве набора обучения и 1/3 как испытательный набор,

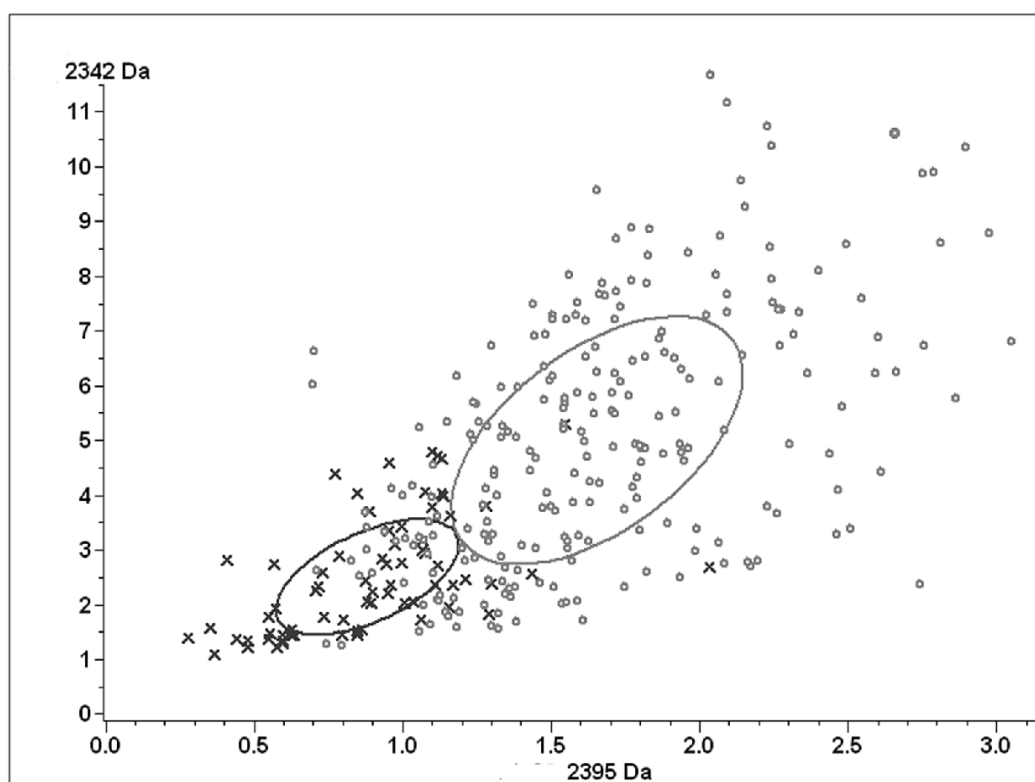


Рисунок 1. Статистическое распределение образцов плазмы крови (I-IV стадия РЖ/контроль), основанное на значениях площадей пиков с m/z 2342 и 2395 Да (в произвольных единицах) в процессе обучения биоинформационного пакета ClinProTools 2.1. (X) - распределение площадей пиков, принадлежащих образцам плазмы крови больных раком желудка; (o) - распределение площадей пиков, принадлежащих образцам плазмы крови индивидов контрольной группы.

ПОИСК ПОТЕНЦИАЛЬНЫХ МАРКЕРОВ РАКА ЖЕЛУДКА

чтобы определить, насколько правильно данная модель могла предсказать рак желудка. В таблице 4 представлены результаты перекрестной проверки на обучающей серии образцов, которые демонстрируют высокую дискриминационную способность моделей, построенных методом ГА, МОВ, УНС (>93%).

Полученные при анализе обучающей серии образцов наборы пиков, отличающие образцы плазмы крови больных РЖ от КГ, использовали для внешней проверки в независимой серии из 152 образцов. Основываясь на результатах испытательной серии, вычисляли чувствительность и специфичность построенных дискриминационных моделей. Как видно из таблицы 4, все модели могли классифицировать образцы случая и контроля с достаточно высокой чувствительностью и специфичностью. Наивысшие значения этих параметров имели наборы пиков, сформированные методом УНС. Анализ качественного состава модели УНС, представленного в таблице 5, показал, что она состояла из 21 пика (в диапазоне m/z от 1332 до 7449 Да), 14 из которых имели достоверно различающиеся уровни сигнала ($p < 0,01$) для двух классов, а 9 пиков имели значения $AUC > 0,7$, подтверждающую их высокую дискриминационную способность.

Из 14 пиков пять отвечали протонированным пептидам/полипептидам, экспрессия которых увеличивалась при РЖ в сравнении с КГ, а девять – с пониженной экспрессией (см. табл. 5). Данная модель разделяла образцы плазмы крови больных I-IV стадии РЖ и контроля с чувствительностью 83,9% и специфичностью 91,5% (табл. 4). Для сравнения в таблице 6 приведены подобные параметры для модели, построенной с помощью метода ГА, состоящей из 5 пиков (в диапазоне m/z от 1179 до 10715 Да) с достоверно различающимися уровнями сигнала ($p < 0,01$) для двух классов, три из которых имели значения $AUC > 0,7$. Эта дискриминационная модель разделяла образцы плазмы крови больных I-IV стадии РЖ и контроля с чувствительностью 69,1% и специфичностью 91,5% (табл. 4). Снижение чувствительности по сравнению с предыдущей моделью объясняется, по-видимому, присутствием меньшего числа дискриминирующих пиков.

Далее полученные модели, изучали на способность разделять образцы плазмы крови от больных с различными стадиями РЖ и контроля. Для валидации использовали два набора, состоящих из комбинации образцов КГ ($n=118$) с образцами от больных I-II стадии ($n=14$) и III-IV стадии ($n=20$) рака желудка. Как видно из таблицы 4,

Таблица 4. Чувствительность и специфичность построенных дискриминационных моделей (данные обучающей и тестовой серии образцов).

Классификационные алгоритмы	Чувствительность, %	Специфичность, %
Данные обучающей серии образцов (I-IV стадия РЖ/КГ)		
ГА	80,6	96,9
МОВ	82,1	99,3
УНС	83,8	93
АБК	84,2	85,3
Данные тестовой серии образцов (I-IV стадия РЖ/КГ)		
ГА	69,1	91,5
МОВ	75,0	97,9
УНС	83,9	91,5
АБК	77,4	82,9
Данные тестовой серии образцов (I-II стадия РЖ/КГ)		
ГА	65,7	91,5
МОВ	75,9	97,9
УНС	75,9	91,5
АБК	77,1	82,9
Данные тестовой серии образцов (III-IV стадия РЖ/КГ)		
ГА	75,9	91,5
МОВ	77,1	97,9
УНС	91,4	91,5
АБК	79,3	82,9

Таблица 5. Качественный состав дискриминационной модели, построенной с помощью метода управляемых нейронных сетей.

Масса (m/z), Да	Области интегрирования, используемые для классификации		Вес пика	AUC	Изменение сигнала РЖ/КТ ($p<0,01$)
	Стартовая масса (m/z), Да	Конечная масса (m/z), Да			
1847,92	1842,55	1853,78	0,16	0,63	↓
7448,99	7443,72	7458,75	0,11	0,56	↓
2295,86	2292,09	2304,26	0,09	0,53	
2395,47	2391,44	2401,13	0,09	0,92	↓
4611,66	4608,13	4620,83	0,08	0,72	↓
6100,9	6077,76	6133	0,05	0,65	↓
3279,72	3273,65	3288,63	0,04	0,71	↓
5097,93	5086,66	5107,35	0,04	0,59	
2901,31	2894,67	2909,86	0,04	0,67	↓
5536,97	5526,63	5550,98	0,04	0,75	↑
1741,36	1736,64	1744,42	0,03	0,79	↑
2022,6	2016,77	2030,78	0,02	0,73	↑
3868,77	3858,69	3878,44	0,02	0,58	
5683,39	5678,21	5696,54	0,02	0,56	
4477,05	4467,88	4480,38	0,02	0,63	↓
1897,57	1892,46	1903,59	0,02	0,85	↑
2820,94	2812,58	2827,85	0,01	0,84	↓
3316,76	3310,41	3326,76	0,01	0,72	↑
1887,72	1882,41	1892,46	0,01	0,81	↓
4824,59	4820,52	4835,97	0,01	0,52	
1331,9	1327,86	1336,16	0,01	0,55	

Таблица 6. Качественный состав дискриминационной модели, построенной с помощью генетического алгоритма.

Масса (m/z), Да	Области интегрирования, используемые для классификации		Вес пика	AUC	Изменение сигнала РЖ/КТ ($p<0,01$)
	Стартовая масса (m/z), Да	Конечная масса (m/z), Да			
3199,14	3196	3207,84	1,00	0,84	↓
2395,47	2391,44	2401,13	1,34	0,92	↓
1178,53	1175,46	1181,99	0,38	0,65	↑
10714,84	10707,35	10730,82	0,49	0,72	↑
2802,49	2794,69	2812,38	0,51	0,69	↓

наблюдалось увеличение чувствительности всех моделей, созданных различными алгоритмами, при переходе от I-II к III-IV стадии рака желудка. В частности, набор пептидов/полипептидов, сформированный УНС, разделял образцы плазмы крови больных I-II стадии РЖ от образцов КГ с чувствительностью 75,9%, тогда как образцы III-IV стадии с чувствительностью 91,4%.

Рак желудка, протекающий вначале практически бессимптомно, в дальнейшем имеет неблагоприятный прогноз и

сопровождается метастазами и рецидивами [12, 13]. Ранняя диагностика РЖ может увеличить эффективность противоопухолевой терапии и уменьшить вероятность прогрессии болезни, её распространения и смерти больного. Однако в настоящее время не существует надёжных диагностических биомаркеров и нет никакой удовлетворительной стратегии для скрининга и ранней диагностики рака желудка. Поэтому в последние годы проводятся активные исследования в этой

области с применением самых современных аналитических методов. Считается, что 20–25% всех белков, кодируемых геномом, должны секретироваться клетками [14] и попадать в сыворотку/плазму крови человека, формируя один из самых информативных и важных с клинической точки зрения протеомов. Так как возникновение и развитие рака приводит к характерным изменениям в протеоме сыворотки/плазмы крови больного, можно использовать протеомные методы для поиска потенциальных биомаркеров РЖ в этих биологических жидкостях [15–17]. Особый интерес представляет мало изученная низкомолекулярная фракция их протеома, в которой, как полагают, могут находиться потенциальные биомаркеры опухолевого роста и мишени для таргетной терапии, обладающие повышенной сосудистой проницаемостью [18–21], в отличие от белков с большим молекулярным весом.

Масс-спектрометрия с усиленной поверхностью лазерной десорбцией/ионизацией (SELDI-MS) была одной из первых технологий, применявшихся при профилировании клинических образцов сыворотки/плазмы крови для обнаружения потенциальных биомаркеров рака желудка. Liang Y. и др. [22] использовали SELDI-MS для фингерпринта сыворотки крови больных РЖ, гастритом и индивидов КГ и обнаружили в масс-спектре ДЭ пики с m/z 5910, 5084, 6640 и 8691 Да. В этой работе впервые продемонстрированы возможности SELDI-MS для скрининга повреждений желудка. Su Y. и др. [23] изучили образцы сыворотки крови 127 больных РЖ и 100 индивидов КГ. Авторы идентифицировали пики с m/z 1468, 3935 и 7560 Да, которые использовали для диагностики рака желудка. Дискриминационная модель на их основе показала чувствительность 95,6% и специфичность 92,0%. Другие исследователи также использовали SELDI-MS для диагностики рака желудка. В работе [24] с использованием чипа CM10 были проанализированы 100 образцов плазмы крови (60 от больных РЖ и 40 от КГ). Алгоритм дерева решений построил модель из 17 пиков, валидация которой слепым методом показала чувствительность 93,3% и специфичность 90%. Zhonghua Y.F. и др. [25] применили SELDI-MS с чипом Q10 для сравнения профилей сыворотки крови от 36 больных РЖ и 46 пациентов с поверхностным гастритом. Чувствительность и специфичность построенной модели

составила 80,0 и 88,9%, соответственно. В настоящее время SELDI-MS продолжает использоваться в исследованиях по раку желудка. В работе [26] авторы применили эту платформу для ранней диагностики РЖ, исследуя образцы сыворотки 169 больных I-й стадией РЖ и 83 индивидов КГ, аннотированных по полу и возрасту. Для диагностической модели были отобраны пики с m/z 2873, 3163, 4526, 5762, 6121 и 7778 Да. Дискриминационная модель эффективно отличала образцы больных РЖ от контрольных образцов с чувствительностью и специфичностью 93,5 и 91,6%, соответственно. Кроме того, были идентифицированы 3 из 6 пиков с m/z 2873, 6121 и 7778 Да.

До настоящего времени опубликовано небольшое количество работ с использованием MALDI-TOF-MS для обнаружения в сыворотке/плазме крови биомаркеров рака желудка. Ebert M.P.A. и др. [27] получили масс-спектры НМП сыворотки крови 14 больных РЖ и 14 здоровых индивидов с помощью MALDI-TOF-MS и идентифицировали пик с m/z 1466 Да (фибринопептид А), который значительно возрастал у больных РЖ. Однако увеличение его уровня в сыворотке не является специфичным для РЖ и может наблюдаться также при раке печени, яичников и уротелиальном раке [28–33]. В 2008 году Chang и др. [34] изучили пептидные профили образцов желудочного сока от 34 больных РЖ и 106 здоровых субъектов [34] и идентифицировали в масс-спектре пять ДЭ пиков. Пики с m/z 2187, 2387 и 3572 Да отвечали протонированным пептидам, уровни которых снижались при раке желудка (с чувствительностью 85, 88 и 79%), в то время как пики с m/z 2753 и 4132 Да – возрастали, диагностируя РЖ с меньшей чувствительностью 47 и 68%, соответственно [34].

Несколько предыдущих исследований также использовали MALDI-TOF-MS для идентификации пептидов и определения их уровней экспрессии [35–37]. Изучая сыворотку от 28 больных РЖ и 16 здоровых пациентов, Gao X. и др. сообщили, что семь пептидов, которым отвечали пики с m/z 4268, 56363, 5248, 5754, 2933, 1450 и 1349 Да, могут использоваться в качестве потенциальных серологических биомаркеров для рака желудка [36].

В отличие от этих исследований для сравнительного профилирования НМП плазмы крови больных РЖ и КГ

мы использовали MALDI-TOF-MS технологию после предварительного фракционирования образцов на магнитных частицах MB-WCX ("Bruker Daltonics"). Эта технологическая платформа значительно увеличивает чувствительность и воспроизводимость анализа, что подтверждается рядом исследований по профилированию образцов сыворотки от онкологических больных [28-30]. Кроме того, MALDI-TOF-MS анализ в комбинации с современными биоинформационными программами (например, ClinProTools™ 2.1, используемой в этом исследовании) может применяться для поиска кандидатных маркеров заболеваний человека и увеличивать потенциал клинических протеомных исследований. В своей работе мы подтвердили это мнение.

При обработке полученных данных с помощью ClinProTools™ 2.1 были обнаружены 143 иона, которые отвечали ДЭ протонированным пептидам/полипептидам. Некоторые дискриминационные пики имели повышенную интенсивность сигнала у больных РЖ в сравнении с КГ (например, с m/z 1898; 3317 и др.), в то время как другие – пониженную (например, с m/z 2901; 4612; 6101 и др.). Модели, созданные всеми четырьмя алгоритмами (ГА, МОВ, НСУ, АБК), показывали достаточно высокую чувствительность и специфичность как в процессе обучения, так и при внешней валидации. В фазе обучения была получена средняя чувствительность 82,7% и специфичность 93,6%, а при слепом тесте для образцов:

I-IV стадия РЖ vs КГ -76,4 и 91,0%; I-II стадия РЖ vs КГ – 73,7 и 91,0%; III-IV стадия РЖ vs КГ – 81,0 и 91,0%, соответственно. Полученные данные демонстрируют способность моделей использоваться при ранней диагностике рака желудка.

В задачи этого исследования не входила идентификация обнаруженных пептидов/полипептидов, однако проведённый анализ литературы показал, что масса 30 из 255 ионов усреднённых масс-спектров совпадали с точностью ± 1 Да с массами ионов протонированных пептидов/полипептидов, ранее идентифицированных другими авторами (см. табл. 7). Следует отметить, что 20 из них имели достаточно высокую дискриминационную способность ($AUC > 0,60$) и статистически достоверное различие сигналов между группами случая и контроля ($p < 0,05$). К ним относились фрагменты фибринопептида А (1078, 1350), комплемента С3f (1212, 1450, 2021), комплемента С4а (1499, 1628, 1741, 1898, 3200), ИТН4 (1531, 2358, 2879), фибриногена А (2660, 2931, 5335), аполипопротеинов А4, Е, С1 (1928, 2410, 6630), С1 ингибитора (4153). Для иллюстрации на рисунке 2 приведены усреднённые спектральные характеристики четырёх ионов, отвечающих протонированным фрагментам комплемента С4а (1741, 1898), фибриногена А (2661), аполипопротеина С1 (6629). Дополнительно показаны ROC-кривые для выбранных пиков вместе с их площадью под кривой (AUC), характеризующие их высокую дискриминационную способность.

Таблица 7. Список идентифицированных пептидов/полипептидов низкомолекулярного протеома сыворотки/плазмы крови человека.

m/z иона (Да)	Идентифицированный пептид/протеин	Ссылка	m/z иона (Да)	Идентифицированный пептид/протеин	Ссылка
1078	Фибринопептид А, фрагмент	[38]	2358	ИТН4, фрагмент	[39, 41]
1212	Комплемент С3f, фрагмент	[39]	2410	Аполипопротеин Е, фрагмент	[39]
1350	Фибринопептид А, фрагмент	[39]	2545	Фибриноген А, фрагмент	[40]
1450	Комплемент С3f, фрагмент	[39]	2660	Фибриноген А, фрагмент	[39, 40]
1465	Фибринопептид А, фрагмент	[27, 39]	2862	Фибриноген А, фрагмент	[40]
1499	Комплемент С4а, фрагмент	[39]	2879	ИТН4, фрагмент	[41]
1531	ИТН4, фрагмент	[40]	2931	Фибриноген А, фрагмент	[39, 40]
1628	Комплемент С4а, фрагмент	[39]	3142	ИТН4, фрагмент	[41]
1741	Комплемент С4а, фрагмент	[39]	3200	Комплемент С4а, фрагмент	[39]
1866	Фибриноген А, фрагмент	[40]	3240	Фибриноген А, фрагмент	[39, 40]
1897	Комплемент С4а, фрагмент	[39, 40]	4153	С1 ингибитор, фрагмент	[42]
1928	Аполипопротеин А4, фрагмент	[39]	5335	Фибриноген А, фрагмент	[40]
1945	Кининоген, фрагмент	[39]	5903	Фибриноген А, фрагмент	[40, 42, 43]
2022	Комплемент С3f, фрагмент	[39, 40]	6630	Аполипопротеин С1	[42, 44]
2127	Кининоген, фрагмент	[39]	7764	Фактор тромбоцитов 4	[42, 45]

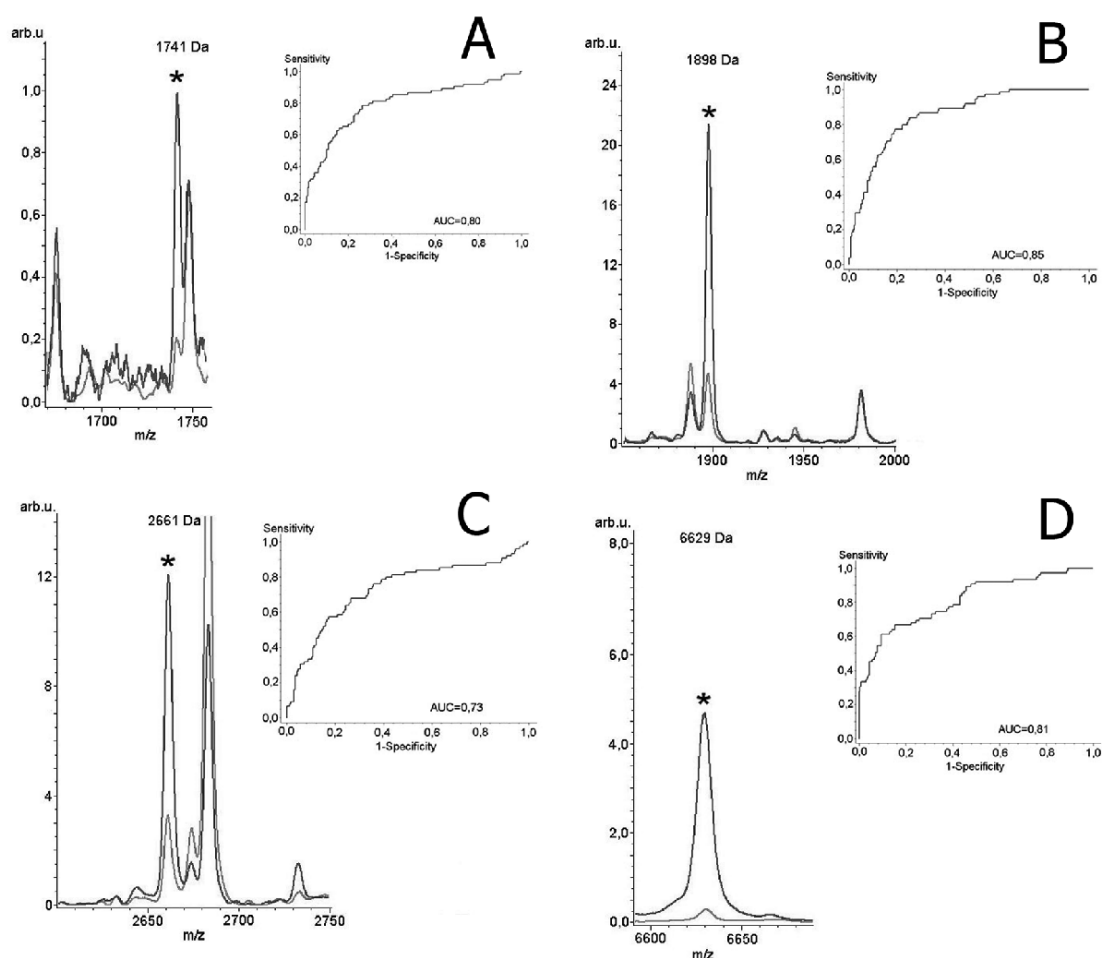


Рисунок 2. Описание усредненных спектральных характеристик четырех ионов (отмеченных звёздочкой), отвечающих протонированным фрагментам компонента C4a (1741, 1898), фибриногена А (2661), аполипопротеина С1 (6629). Дополнительно показаны ROC-кривые для выбранных пиков вместе с их площадью под кривой (AUC)

По нашему мнению, обнаруженные потенциальные маркеры РЖ представляют собой циркулирующие в крови пептиды/полипептиды или их фрагменты, уровни концентраций которых модулируются опухолевым процессом. Можно утверждать с большой долей вероятности, что часть отобранных маркеров обладает высокой дискриминационной способностью. Тем не менее, число проанализированных нами образцов плазмы крови больных РЖ относительно невелико, что может сказываться на проверке качества представленных моделей. В этой связи необходимы дальнейшие исследования для подтверждения и уточнения полученных результатов на большем количестве образцов. После идентификации и синтеза ключевых пептидов/полипептидов могут быть

созданы антитела для разработки высокопроизводительных тест-систем. В дальнейшем эти наборы могут быть включены в проспективные испытания для оценки диагностической эффективности этих белков в распознавании рака желудка.

ЗАКЛЮЧЕНИЕ И ВЫВОДЫ

В настоящем исследовании разработана и апробирована новая методология поиска маркеров рака желудка при сравнительном профилировании низкомолекулярного протеома (1-17 кДа) плазмы крови больных РЖ и индивидов контрольной группы. Данный подход включал: предварительное фракционирование образцов магнитными микрочастицами с катионообменным покрытием для

специфического захвата биомолекул, MALDI-TOF масс-спектрометрию, обработку данных с помощью биоинформационного пакета программ. Данная платформа успешно использовалась для подготовки клинических образцов и проведения MALDI-TOF-MS профилирования НМП плазмы крови. Мы изучили возможность использования различных классификационных алгоритмов для генерации диагностически значимых наборов пептидов/полипептидов из масс-спектрометрических данных для дискриминации образцов от больных с различными стадиями РЖ и контрольной группы. Продемонстрированы возможности и перспективность разработанного подхода для обнаружения потенциальных маркеров РЖ при скрининге 69 образцов плазмы крови больных РЖ и 238 образцов контрольной группы. В результате проведенного исследования обнаружены пептиды/полипептиды, которые могут в перспективе использоваться для детектирования этой патологии.

ЛИТЕРАТУРА

1. *Parkin D.M., Bray F., Ferlay J., Pisani P.* (2005) *Cancer J. Clin.*, **55**, 74–108.
2. *Ohta H., Noguchi Y., Takagi K., Nishi Y., Kajitani T., Kati Y.* (1987) *Cancer*, **60**, 1099–1106.
3. *Peddanna N., Holt S., Verma R.S.* (1995) *Anticancer Res.*, **15**, 2055–2064.
4. *Ebert M.P., Rocken C.* (2006) *Eur. J. Gastroenterol. Hepatol.*, **18**, 847–853.
5. *Tirumalai R.S., Chan K.C., Prieto D.A., Issaq H.J., Conrads T.P., Veenstra T.D.* (2003) *Mol. Cell Proteomics*, **2**, 1096–1103.
6. *Anderson N.L., Anderson N.G.* (2002) *Mol. Cell Proteomics*, **1**, 845–867.
7. *Villanueva J., Philip J., Entenberg D., Chaparro C.A., Tanwar M.K., Holland E.C., Tempst P.* (2004) *Anal. Chem.*, **76**, 1560–1570.
8. Шевченко В.Е., Арноцкая Н.Е., Трифонова О.П., Дашкевич А.С., Юрченко В.А., Заридзе Д.Г. (2007) *Масс-спектрометрия*, **4**, 245–254.
9. *Shevchenko V.E., Arnotskaya N.E., Zaridze D.G.* (2010) *Eur. J. Mass Spectrom.*, **16**, 539–549.
10. *Pyatnitskiy M.A., Lisitsa A.V., Moshkovskii S.A., Arnotskaya N.E., Akhmedov B.B., Zaridze D.G., Polotskii B.E., Shevchenko V.E.* (2011) *J. Analyt. Chem.*, **66**, 1369–1375.
11. Шевченко В.Е., Мадина Х., Поддубная И.В., Арноцкая Н.Е., Таупов М.А., Юрченко В.А., Заридзе Д.Г. (2011) *Вестник РОНЦ им. Н.Н. Блохина РАМН*, **22**(3), 27–33.
12. *Wang J., Yu J.C., Kang W.M., Ma Z.Q.* (2012) *Surg. Oncol.*, **21**, 119–123.
13. *Saka M., Morita S., Fukagawa T., Katai H.* (2011) *Jpn. J. Clin. Oncol.*, **41**, 307–313.
14. *Chen Y., Yu P., Luo J., Jiang Y.* (2003) *Mamm. Genome*, **14**, 859–865.
15. *Caffrey R.E.* (2010) *Methods Mol. Biol.*, **641**, 167–183.
16. *Cho W.C.* (2007) *Genomics, Proteomics, Bioinformatics*, **5**, 77–85.
17. *Cho W.C.* (2007) *Mol. Cancer*, **6**, 25.
18. *Jain R.K.* (1987) *Cancer Metastasis Rev.*, **6**(4), 559–593.
19. *Dvorak H.F., Nagy J.A., Dvorak J.T., Dvorak A.M.* (1988) *Am. J. Pathol.*, **133**, 95–109.
20. *Yuan F., Dellian M., Fukumura D., Leunig M., Berk D.A., Torchilin V.P., Jain R.K.* (1995) *Cancer Res.*, **55**(17), 3752–3756.
21. *Dellian M., Yuan F., Trubetskoy V.S., Torchilin V.P., Jain R.K.* (2000) *Br. J. Cancer*, **82**, 1513–1518.
22. *Liang Y., Fang M., Li J., Liu C.B., Rudd J.A., Kung H.F., Yew D.T.* (2006) *Exp. Mol. Pathol.*, **81**, 176–180.
23. *Su Y., Shen J., Qian H., Ma H., Ji J., Ma H., Ma L., Zhang W., Meng L., Li Z., Wu J., Jin G., Zhang J., Shou C.* (2007) *Cancer Sci.*, **98**, 37–43.
24. *Lim J.Y., Cho J.Y., Paik Y.H., Chang Y.S., Kim H.G.* (2007) *Int. J. Biol. Markers*, **22**, 281–286.
25. *Wang Y.S., Shen J., Yu J.K., Zhang L., Ma J.L., You W.C., Pan K.F.* (2008) *Zhonghua Yu Fang Yi Xue Za Zhi*, **42**, 677–680.
26. *Li P., Zhang D., Guo C.* (2012) *Mol. Med. Report*, **5**, 1531–1535.
27. *Ebert M.P.A., Niemeyer D., Deininger S.O., Wex T., Knippig C., Hoffmann J., Sauer J., Albrecht W., Knippig P., Rolcken C.* (2006) *J. Proteome Res.*, **5**, 2152–2158.
28. *Fiedler G.M., Baumann S., Leichtle A., Oltmann A., Kase J., Thiery J., Ceglarek U.* (2007) *Clin. Chem.*, **53**, 421–428.
29. *Schaub N.P., Jones K.J., Nyalwidhe J.O., Cazares L.H., Karbassi I.D., Semmes O.J., Feliberti E.C., Perry R.R., Drake R.R.* (2009) *J. Am. Coll. Surg.*, **208**, 970–978.
30. *Schwamborn K., Krieg R.C., Grosse J., Reulen N., Weiskirchen R., Knuechel R., Jakse G., Henkel C.* (2009) *Eur. Urol.*, **56**, 989–999.
31. *Theodorescu D., Wittke S., Ross M.M., Walden M., Conaway M., Just I., Mischak H., Frierson H.F.* (2006) *Lancet Oncol.*, **7**, 230–240.
32. *Bergen H.R., Vasmatazis G., Cliby W.A., Johnson K.L., Oberg A.L., Muddiman D.C.* (2003) *Dis. Markers*, **19**, 239–249.
33. *Orvisky E., Drake S.K., Martin B.M., Abdel-Hamid M., Ransom H., Varghese R.S., An Y., Saha D., Hortin G.L., Loffredo C.A., Goldman R.* (2006) *Proteomics*, **6**, 2895–2902.
34. *Chang W.C., Hsu P.I., Chen Y.Y., Hsiao M., Lu P.J.* (2008) *Proteomics Clin. Appl.*, **2**, 55–62.
35. *Liu W.T., Gao X., Cai Q., Li J.F., Zhu Z.L., Li C., Yao X.X., Yang Q.M., Xiang M., Yan M., Zhu Z.G.* (2010) *Front. Biosci.*, **2**, 961–971.

ПОИСК ПОТЕНЦИАЛЬНЫХ МАРКЕРОВ РАКА ЖЕЛУДКА

36. Gao X., Liu W.T., Yang Q.M., Liu B.Y., Cai Li., Li J.F., Zhu Z.L., Xiang M., Yan M., Zhu Z.G. (2010) *Chin. J. Clin. Oncol.*, **16**, 933–936.
37. Li T., Shen H.M. (2009) *Modern Oncol.*, **17**, 1513–1517.
38. Marshall J., Kupchak P., Zhu W., Yantha J., Vrees T., Furesz S., Jacks K., Smith C., Kireeva I., Zhang R., Takahashi M., Stanton E., Jackowski G. (2003) *J. Proteome Res.*, **2**, 361–372.
39. Villanueva J., Shaffer D.R., Philip J., Chaparro C.A., Erdjument-Bromage H., Olshen A.B., Fleisher M., Lilja H., Brogi E., Boyd J., Sanchez-Carbayo M., Holland E.C., Cordon-Cardo C., Scher H.I., Tempst P. (2006) *J. Clin. Invest.*, **116**, 271–284.
40. Peng J., Stanley A.J., Cairns D., Selby P.J., Banks R.E. (2009) *Proteomics*, **9**, 492–498.
41. Fung E.T., Yip T.T., Lomas L., Wang Z., Yip C., Meng X.Y., Lin S., Zhang F., Zhang Z., Chan D.W., Weinberger S.R. (2005) *Int. J. Cancer*, **115**, 783–789.
42. Timms J.F., Arslan-Low E., Gentry-Maharaj A., Luo Z., T'Jampens D., Podust V.N., Ford J., Fung E.T., Gammernan A., Jacobs I., Menon U. (2007) *Clin. Chem.*, **53**, 645–656.
43. Villanueva J., Martorella A.J., Lawlor K., Philip J., Fleisher M., Robbins R.J., Tempst P. (2006) *Mol. Cell Proteomics*, **5**, 1840–1852.
44. Ward D.G., Suggett N., Cheng Y., Wei W., Johnson H., Billingham L.J., Ismail T., Wakelam M.J., Johnson P.J., Martin A. (2006) *Br. J. Cancer*, **94**, 1898–1905.
45. Vermeulen R., Lan Q., Zhang L., Gunn L., McCarthy D., Woodbury R.L., McGuire M., Podust V.N., Li G., Chatterjee N., Mu R., Yin S., Rothman N., Smith M.T. (2005) *Proc. Natl. Acad. Sci. USA*, **102**, 17041–17046.

Поступила: 10. 02. 2013.

SEARCH OF POTENTIAL GASTRIC CANCER BIOMARKERS USING LOW MOLECULAR WEIGHT BLOOD PLASMA PROTEOME PROFILING BY MASS SPECTROMETRY

**V.E. Shevchenko, N.E. Arnotskaya, E.V. Ogorodnikova, M.M. Davidov, M.A. Ibraev,
I.N. Turkin, M.I. Davidov**

N.N. Blokhin Russian Cancer Research Center of Russian Academy of Medical Sciences,
Kashirskoye sh., 24, Moscow, 115478 Russia; tel.: (499)3235611; fax: (499)3242924;
e-mail: vshev@nm.ru

Gastric cancer, one of the most widespread malignant tumors, still lacks reliable serum/plasma biomarkers of its early detection. In this study we have developed, unified, and tested a new methodology for search of gastric cancer biomarkers based on profiling of low molecular weight proteome (LMWP) (1–17 kDa). This approach included three main components: sample pre-fractionation, matrix-assisted laser desorption ionization time of flight mass spectrometry (MALDI-TOF-MS), data analysis by a bioinformatics software package. Applicability and perspectives of the developed approach for detection of potential gastric cancer markers during LMWP analysis have been demonstrated using 69 plasma samples from patients with gastric cancer (stages I–IV) and 238 control samples. The study revealed peptides/polypeptides, which may be potentially used for detection of this pathology.

Key words: biomarkers, gastric cancer, proteome, blood plasma, mass spectrometry.