

Идентификация маркеров аденокарциномы желудка на основе биоинформатического поиска и анализа генной экспрессии

В.В. Волкоморов^{1,2}, Е.С. Григорьева^{1,2}, Г.В. Краснов³, Н.В. Литвяков^{1,2}, Н.А. Лисицын³,
М.И. Воевода⁴, А.В. Белковец⁴, С.Г. Афанасьев¹, Н.В. Чердынцева^{1,2}

¹Научно-исследовательский институт онкологии ФГБНУ «Томский национальный исследовательский медицинский центр РАН»;
Россия, 634009 Томск, пер. Кооперативный, 5;

²ФГАОУ ВО «Национальный исследовательский Томский государственный университет»;
Россия, 634050 Томск, пр. Ленина, 36;

³ФГБУН «Институт молекулярной биологии им. В.А. Энгельгардта РАН»;
Россия, 119991 Москва, ул. Вавилова, 32;

⁴ФГБНУ «Научно-исследовательский институт терапии и профилактической медицины»;
Россия, 630089 Новосибирск, ул. Б. Богаткова, 175/1

Контакты: Виктор Владимирович Волкоморов v.volkomorov@gmail.com

Введение. Одна из важных задач современной онкологии — поиск ассоциированных с опухолями молекулярных маркеров, которые могут использоваться для диагностики и прогнозирования рака, оценки степени радикальности операции и последующего лечения, а также раннего выявления рецидивов. Одним из продуктивных вариантов подобного поиска является анализ транскриптомных баз данных с применением методов биоинформатики с последующей валидацией полученных результатов на клиническом материале.

Цель исследования — поиск мембранных белков, которые могут быть использованы для сывороточной диагностики аденокарциномы желудка интестинального гистологического типа.

Материалы и методы. Идентификацию потенциальных маркеров рака желудка (РЖ) проводили с использованием баз данных Gene Ontology и The Cancer Genome Atlas (TCGA). Последующую оценку дифференциальной экспрессии генов выполняли на парных образцах аденокарциномы и нормальной ткани желудка, взятых от 55 пациентов. Экспрессию генов оценивали с помощью полимеразной цепной реакции с обратной транскрипцией в режиме реального времени по методу ΔCq .

Результаты. Сравнительный анализ уровней синтеза матричных РНК (мРНК) нормальных и опухолевых тканей с применением нового алгоритма биоинформатического поиска привел к идентификации 3 наиболее высококопийных транскриптов (SULF1, PMEPA1 и SPARC), внутриклеточное содержание которых заметно повышается при РЖ. При анализе уровня мРНК данных генов в клиническом материале наблюдалось более чем двукратное увеличение уровня экспрессии PMEPA1 и SPARC в 75 % образцов РЖ интестинального гистологического типа. В образцах РЖ диффузного гистологического типа этот показатель составил 25 и 38 % соответственно.

Выводы. Использование оригинального биоинформатического подхода, основанного на анализе данных TCGA, позволило выявить 2 гена (PMEPA1 и SPARC), преимущественно экспрессирующихся в опухолях желудка интестинального типа. Полученные результаты свидетельствуют об актуальности дальнейшего исследования роли этих генов в патогенезе РЖ и оценки клинической значимости уровня их экспрессии в опухолевой ткани.

Ключевые слова: аденокарцинома желудка, гистологический тип, гены-маркеры, генетическая база данных, экспрессия генов, PMEPA1, SPARC, полимеразная цепная реакция в режиме реального времени

DOI: 10.17650/2313-805X-2017-4-1-40-45

Identification of the intestinal type gastric adenocarcinoma transcriptomic markers using bioinformatic and gene expression analysis

V.V. Volkomorov^{1,2}, E.S. Grigor'eva^{1,2}, G.S. Krasnov³, N.V. Litvyakov^{1,2}, N.A. Lisitsyn³,
M.I. Voevoda⁴, A.V. Belkovets⁴, S.G. Afanas'ev¹, N.V. Cherdyntseva^{1,2}

¹Cancer Research Institute, Tomsk National Research Medical Center of the Russian Academy of Sciences;
5 Kooperativnyy Pereulok, Tomsk 634009, Russia;

²National Research Tomsk State University; 36 Lenina Prospekt, Tomsk 634050 Russia;

³V.A. Engelhardt Institute of Molecular Biology of the Russian Academy of Sciences; 32 Vavilova St., Moscow 119991, Russia;

⁴Research Institute of Internal and Preventive Medicine; 175/1 B. Bogatkova St., Novosibirsk 630089, Russia

Introduction. Searching for specific and sensitive molecular tumor markers is one of the important tasks of modern oncology. These markers can be used for early tumor diagnosis and prognosis as well as for prediction of therapeutic response, estimation of tumor volume or to assess disease recurrence through monitoring. Gene expression data base mining followed by experimental validation of results obtained is one of the promising approaches for searching of that kind.

Objective: to identify several membrane proteins which can be used for serum diagnosis of intestinal type of gastric adenocarcinoma.

Materials and methods. We used bioinformatic-driven search using Gene Ontology and The Cancer Genome Atlas (TCGA) data to identify mRNA up-regulated in gastric cancer (GC). Then, the expression levels of the mRNAs in 55 pare clinical specimens were investigated using reverse transcription polymerase chain reaction.

Results. Comparative analysis of the mRNA levels in normal and tumor tissues using a new bioinformatics algorithm allowed to identify 3 high-copy transcripts (*SULF1*, *PMEPA1* and *SPARC*), intracellular content of which markedly increased in GC. Expression analysis of these genes in clinical specimens showed significantly higher mRNA levels of *PMEPA1* and *SPARC* in tumor as compared to normal gastric tissue. Interestingly more than twofold increase in expression level of these genes was observed in 75 % of intestinal-type GC. The same results were found only in 25 and 38 % of diffuse-type GC respectively.

Conclusions. As a result of original bioinformatic analysis using TCGA data base two genes (*PMEPA1* and *SPARC*) were shown to be significantly upregulated in intestinal-type gastric adenocarcinoma. The findings show the importance of further investigation to clarify the clinical value of their expression level in stomach tumors as well as their role in carcinogenesis.

Key words: gastric adenocarcinoma, histological type, genetic markers, gene expression database, gene expression, *PMEPA1*, *SPARC*, reverse transcription polymerase chain reaction assay

Введение

Одна из важных задач современной онкологии — поиск ассоциированных с опухолями молекулярных маркеров, которые могут использоваться для диагностики и прогнозирования рака, оценки степени радикальности операции и последующего лечения, а также раннего выявления рецидивов [1]. Для поиска таких маркеров применяются методы, основанные на сравнительном анализе геномов, транскриптомов и протеомов опухолевых и нормальных клеток в целях обнаружения специфических молекулярно-генетических изменений, возникающих в процессе канцерогенеза. Одним из используемых подходов является идентификация генов, уровень транскрипции которых заметно повышается в опухолях, при этом для применения в клинической практике особенно важны случаи, в которых такое повышение коррелирует с клиническими параметрами. Последующие идентификация и анализ активируемых сигнальных путей, обеспечивающих биологическое поведение опухолевых клеток, позволяют выявить потенциальные мишени для противоопухолевой терапии [2, 3].

Используемые в настоящее время транскриптомные методы поиска молекулярных маркеров рака можно подразделить на 2 группы.

- Сравнительный анализ результатов гибридизации тотальной комплементарной ДНК (кДНК) нормальных и опухолевых тканей с коммерчески распространяемыми микрочипами (результаты таких исследований содержит база данных Oncomine). К недостаткам этого подхода можно отнести невозможность оценки уровня транскрипции низкокопийных генов, а также элиминацию сигналов, получаемых в результате перекрестной гибридизации транскриптов гомологичных генов [4].
- Сравнительный анализ результатов высокопроизводительного секвенирования транскриптома опухолевых и нормальных тканей. Идентификация генов, уровень транскрипции которых наиболее заметно и часто различается

в опухолевой и нормальной тканях, требует применения методов биоинформатики [1, 5]. Для последующей валидации результатов биоинформатического поиска в качестве «золотого стандарта» используется метод количественной полимеразной цепной реакции с обратной транскрипцией в режиме реального времени (ОТ-ПЦР-РВ) [6, 7].

В настоящем исследовании представлены результаты нахождения маркеров, специфичных для аденокарцином желудка интестинального и диффузного гистологических типов, на основе комплексного биоинформатического поиска и валидации дифференциальной экспрессии выявленных генов в опухолевой и нормальной тканях на клиническом материале больных раком желудка (РЖ). Эти маркеры могут использоваться для сывороточной диагностики с применением разработанного нами ранее формата иммуно-ПЦР [8]. Исследований подобного рода в России практически нет, что указывает на актуальность их проведения, которая подчеркивается фактом наличия популяционных особенностей молекулярных маркеров [9, 10].

Материалы и методы

Биоинформатический поиск. Поиск генов для сывороточной диагностики РЖ проводили в 2 этапа.

На 1-м этапе в базе данных Gene Ontology отбирали гены, кодирующие мембранные белки, а также белки, находящиеся в составе эндосом и образующихся из них экзосом. Поиск осуществляли с использованием ключевых слов: (*extracellular region NOT secreted*) OR (*endosome*).

Второй этап отбора заключался в идентификации генов, экспрессия которых повышается в 2 основных формах РЖ согласно данным RNA-Seq, представленным в базе The Cancer Genome Atlas (TCGA). Анализ проведен с помощью разработанной нами программы CrossHub (статья готовится к печати). Нормирование транскриптомных данных выполняли по усеченному среднему значению относительной экспрессии (ТММ) [11, 12].

При анализе дифференциальной экспрессии производилась оценка 2 параметров:

- распределения числа прочтений (ридов) между 2 пулами образцов (норма—опухоль) с использованием *t*-теста для независимых выборок. Общую дисперсию оценивали как сумму наблюдаемой дисперсии и дисперсии для распределения Пуассона;
- числа прочтений в парных образцах норма—опухоль с использованием *t*-теста для зависимых выборок.

По результатам оценки рассчитывали скоринг-фактор *S*, отражающий величину наблюдаемых изменений экспрессии и их достоверность (false discovery rate, FDR):

$$S = \sum_{\substack{i = \text{pools,} \\ \text{pairs}}} (-1) \text{sign}(\log(FC_i)) \log(FC_i) \log(\text{FDR}_i),$$

где FC_{pools} , FC_{pairs} — изменение содержания транскрипта в опухоли по сравнению с нормой при 1-м и 2-м методах расчета соответственно.

Наибольшее значение *S* присваивалось генам, повышенная транскрипция которых выявлялась по обоим критериям оценки: как при сравнении 2 пулов образцов (норма—опухоль), так и при анализе изменений в индивидуальных парных образцах. Достоверность изменений уровней транскрипции (FDR) рассчитывалась с использованием поправки на множественное тестирование Бенджамини—Хохберга [13].

Дополнительными критериями отбора являлись: 1) повышенная транскрипция отобранных генов в других опухолях желудочно-кишечного тракта (раке толстой и прямой кишки); 2) высокое абсолютное значение уровня транскрипции, оцененное по общему числу ридов для опухолевых образцов. Окончательное ранжирование генов *R* проводили исходя из их положения в 4 рейтинг-листах:

$$R = 2P_{\text{stomach}} + P_{\text{rectum}} + P_{\text{colon}} + P_{\text{expr. level?}}$$

где *P* — место гена в соответствующем рейтинге.

Как известно, повышение уровня транскрипции определяется 3 основными факторами: влиянием активирующих транскрипционных факторов, гипометилированием промоторных и энхансерных областей, а также модификацией гистонов. Среди 50 топ-генов с наименьшим значением *R* отбирали те, для которых результаты транскриптомного анализа подтверждались результатами исследования гипометилирования промоторных областей в базе TCGA, поскольку анализ профилей метилирования является одним из наиболее информативных способов обнаружения возможных изменений транскрипции. Для этого с помощью программы CrossHub анализировали данные метилового профилирования 450 тыс. сайтов генома человека, полученные с использованием микрочипов для опухолей желудка и толстой и прямой кишки.

Характеристика пациентов. Для оценки дифференциальной экспрессии генов были использованы парные образцы аденокарциномы и нормальной ткани от 55 (54,5 % мужчин, 45,5 % женщин) больных РЖ (28 — диффузного и 27 — интестинального гистологического типа), получивших лечение в клинике Томского национального исследовательского медицинского центра. Работа проведена с соблюдением принципов добровольности и конфиденциальности в соответствии с «Основными законодательства Российской Федерации об охране здоровья граждан» (Указ Президента Российской Федерации от 24.12.1993 № 2288). Получены разрешение этического комитета института и информированные согласия пациентов. Средний возраст больных составил $58,3 \pm 1,5$ года. Первичная опухоль T1–2 была диагностирована у 40 % пациентов (9 случаев диффузного гистологического типа, 13 — интестинального), опухоль T3–4 — у 60 % (19 случаев диффузного гистологического типа, 14 — интестинального); метастазы

Таблица 1. Последовательности праймеров и зондов для оценки экспрессии генов

Наименование гена	GenBank Accession Number		Последовательность	Температура отжига, °C
<i>ACTB</i> (actin beta)	NM_001101.3	F R Probe	5'-gagaagatgaccagatcatgtt-3' 5'-atagcacagcctggatagcaa-3' FAM 5'-agaccttcaacacccagccat-3' BHQ1	60
<i>SPARC</i>	NM_003118.3	F R Probe	5'-atctccctgtactggcagttc-3' 5'-ctcgggtgggagaggtacc FAM 5'-cagctggaccagcaccattgac-3' BHQ1	60
<i>SULF1</i>	NM_001128205.1	F R UT-метка	5'-gaagtgaccaagttcatgctaattgctgggaagcctctgtt-3' 5'-aggcacaagaataatgttggtc-3' FAM 5'-agcagatcgttcgagcatcgc(dT-BHQ1)gaagtgaccaagttcatgct	60
<i>PMEPA1</i>	NM_020182.4	F R Probe	5'-tgttccagagcatggagatca-3' 5'-gtgcagacagctttagtgg-3' FAM 5'-catcgtggtgtagtggatgatggtgatg-3' BHQ1	60

Примечание. UT-метка (universal primer tag) — универсальный зонд; FAM — флуоресцентный краситель, карбоксифлуоресцеин; BHQ1 — гаситель флуоресценции.

в регионарные лимфатические узлы отмечены у 40 % больных (15 случаев диффузного гистологического типа, 18 – интестинального).

Выделение матричных РНК (мРНК) и синтез кДНК.

Парные образцы опухолей и нормальной ткани желудка получены от пациентов с диагнозом РЖ при оперативном вмешательстве. РНК из замороженного операционного материала выделяли с помощью набора RNeasy Plus mini Kit (Qiagen, Германия). Синтез кДНК (с использованием гексамерных праймеров) проводили с применением набора RevertAid™ (Fermentas, Латвия).

Оценка уровня транскрипции генов. Процедуры выделения мРНК, синтеза кДНК и постановки ОТ-ПЦР-РВ описаны ранее [14]. Для нормализации уровней экспрессии использовали контрольный ген *ACTB* [15]. Последовательности праймеров и зондов (FAM–BHQ) подбирали с помощью программы Vector NTI 11.5 с использованием базы данных GenBank (www.ncbi.nlm.nih.gov) (табл. 1).

Для каждого образца кДНК ПЦР проводили в 2 репликах. Уровень содержания транскрипта оценивали по формуле:

$$R = 2^{Ct(\text{ref}) - Ct(\text{targ})}$$

Статистический анализ уровней транскрипции генов в нормальной и опухолевой тканях проводили с использованием *U*-критерия Манна–Уитни.

Результаты

Алгоритм биоинформатического поиска был направлен на отбор мембранных белков, которые могут быть использованы для сывороточной диагностики РЖ по разработанной нами ранее технологии иммуно-ПЦР [8]. Для этой цели на 1-м этапе биоинформатического поиска отбирали белки, локализованные на внешней стороне плазматической мембраны, а также в составе эндосом. Таким образом, по данным ресурса Gene

Ontology сформирован список из 1983 генов, отвечающих ключевым словам (см. Материалы и методы). Для идентификации генов, уровень транскрипции которых заметно повышается в опухолях желудка по сравнению с нормой, проводили анализ базы данных TCGA, в настоящее время являющейся наиболее информативным ресурсом, объединяющим сведения транскриптомных, экзомных и метиломных исследований более чем для 15 видов рака. Анализ, выполненный с помощью приложения CrossHub, позволил выявить потенциальное повышение экспрессии 130 генов по парным образцам и 155 генов при сравнении пулов образцов нормы и опухоли (FDR < 0,01). Для 39 и 60 генов соответственно уровень экспрессии в опухоли возрастал более чем в 4 раза.

В результате дальнейшего анализа по 4 критериям (расчет *S* для 3 видов рака и оценка абсолютного значения экспрессии при РЖ) был сформирован рейтинг из 50 генов. Дальнейший анализ данных TCGA показал возможное гипометилирование промоторных областей 20 генов: *SULF1*, *COL1A2*, *ESM1*, *COL5A2*, *COL12A1*, *ADAM12*, *COL4A1*, *INHBA*, *VCAN*, *COL3A1*, *KIAA1199*, *COL5A1*, *ACAN*, *PMEPA1*, *SPPI1*, *ADAMTS2*, *SPARC*, *MET*, *COL11A1*, *MMP7*. Гены, кодирующие коллагены, из дальнейшего анализа были исключены, а из оставшихся 13 генов были отобраны 3 (*SULF1*, *PMEPA1* и *SPARC*) с наибольшим внутриклеточным содержанием (число ридов 10, 9 и 45 млн соответственно).

Экспрессия отобранных генов оценена в парных клинических образцах нормальной и опухолевой тканей желудка интестинального и диффузного типов с помощью количественной ОТ-ПЦР-РВ (табл. 2). Для гена *SULF1* статистически значимых различий между уровнями генной экспрессии образцов 2 типов не выявлено. Однако для оставшихся 2 генов (*PMEPA1* и *SPARC*) в опухолевых образцах интестинального типа наблюдалось более чем двукратное повышение уровня транскрипции по сравнению с образцами диффузного типа ($p < 0,001$). Так, для гена *PMEPA1*

Таблица 2. Результаты полимеразной цепной реакции с обратной транскрипцией генов, уровень экспрессии которых повышается преимущественно при раке желудка интестинального типа

Тип	Me (Q25–75), %			
	Опухоль	Норма	%*	<i>p</i>
<i>PMEPA1</i>				
Интестинальный	2,8 (1,9; 5,8)	1,1 (0,6; 1,5)	58,3	0,0005
Диффузный	2,5 (1,6; 3,5)	1,94 (1,0; 3,4)	25,0	0,174
<i>SPARC</i>				
Интестинальный	15,6 (10,2; 31,9)	5,0 (4,0; 7,8)	75,0	0,012
Диффузный	21,2 (7,2; 24,8)	8,7 (7,0; 12,4)	38,5	0,191

Примечание. Me (Q25–75) – медиана и интерквартильный размах уровня экспрессии; *p* – уровень значимости, *U*-критерий Манна–Уитни.
*Процент опухолей с более чем двукратным увеличением уровня экспрессии гена.

более чем двукратное повышение уровня экспрессии наблюдалось в 75 % образцов интестинального типа по сравнению с 25 % образцов диффузного, в то время как для гена *SPARC* это соотношение составляло 75 и 38 % соответственно (см. табл. 2). Таким образом, транскрипты генов *PMEPA1* и *SPARC* являются преимущественно маркерами РЖ интестинального типа, что открывает новые возможности в сывороточной диагностике опухолей желудка.

Обсуждение

Белок *PMEPA1* является трансмембранным, ингибирующим сигнальный путь TGF- β (transforming growth factor beta, трансформирующего фактора роста бета) – полифункционального цитокина, регулирующего процессы клеточной пролиферации, дифференцировки и апоптоза. Ускользание неопластических клеток от ингибирующего действия TGF- β происходит при возникновении большинства эпителиальных опухолей по 2 механизмам. Первый механизм – результат инактивирующих мутаций в генах, кодирующих белки данного пути, а второй – следствие повышения уровня синтеза белков, ингибирующих путь, таких как *PMEPA1*, уровень синтеза которого увеличивается в опухолях легких и молочной железы [10, 16]. Интересно, что после инактивации указанного пути в опухолевых клетках происходит повышение уровня экспрессии гена *TGF- β* . Это может приводить, во-первых, к пролиферации клеток стромы и кровеносных сосудов, во-вторых, к запуску программы эпителиально-мезенхимального перехода, что способствует

образованию мезенхимальных клеток с локомоторным фенотипом, обладающих свойствами стволовых опухолевых клеток и клеток, способных к метастазированию [17, 18].

В данном исследовании также установлена дифференциальная экспрессия гена *SPARC* при интестинальном типе аденокарциномы желудка. Ранее продемонстрировано, что ген *SPARC* кодирует белок внеклеточного матрикса, регулирующий клеточный цикл и смену полярности эпителиальных клеток в процессах инвазии и метастазирования опухолей. Показано, что повышение экспрессии гена *SPARC* в опухолях желудка является прогностическим фактором, указывающим на быстрое прогрессирование заболевания и снижение показателей выживаемости пациентов [3, 19]. В то же время ингибирование активности гена в клеточных линиях РЖ с помощью специфической микроПНК приводит к резкому снижению инвазивного потенциала и активации апоптоза опухолевых клеток [20].

Заключение

Таким образом, в данной работе с использованием оригинального биоинформатического подхода, основанного на анализе данных TCGA, выявлены 2 гена (*PMEPA1* и *SPARC*), преимущественно экспрессирующихся в опухолях желудка интестинального типа. Полученные нами результаты свидетельствуют об актуальности дальнейшего исследования роли этих генов в патогенезе аденокарцином желудка и необходимости проведения оценки их прогностической значимости и возможности использования в качестве мишеней для химиотерапии.

Финансирование

Биоинформатический анализ выполнен при поддержке гранта № 15-04-08731 Российского фонда фундаментальных исследований, экспериментальный анализ – при поддержке гранта № 14-04-31500.

Конфликт интересов

Авторы заявляют об отсутствии конфликта интересов.

ЛИТЕРАТУРА / REFERENCES

1. Diamandis E.P. Towards identification of true cancer biomarkers. *BMC Med* 2014;12(1):156.
2. Hu Y., He K., Wang D. et al. *TMEPA1* regulates EMT in lung cancer cells by modulating the ROS and IRS-1 signaling pathways. *Carcinogenesis* 2013;34(8):1764–72.
3. Jeung H.C., Rha S.Y., Im C.K. et al. A randomized phase 2 study of docetaxel and S-1 versus docetaxel and cisplatin in advanced gastric cancer with an evaluation of *SPARC* expression for personalized. *Cancer* 2011;117(10):2050–7.
4. Ramaswamy S., Tamayo P., Rifkin R. et al. Multiclass cancer diagnosis using tumor gene expression signatures. *Proc Natl Acad Sci U S A* 2001;98(26):15149–54.
5. Букурова Ю.А., Краснов Г.С., Никитина И.Г. и др. Методы поиска маркеров для сывороточной диагностики опухолей. *Молекулярная биология* 2013;47(1):3–11. [Bukurova Yu.A., Krasnov G.S., Nikitina I.G. et al. Serological diagnosis of tumors: methods of marker's search. *Molekulyarnaya biologiya = Molecular Biology* 2013;47(1):3–11. (In Russ.)].
6. Kulasingam V., Diamandis E.P. Strategies for discovering novel cancer biomarkers through utilization of emerging technologies. *Nat Clin Pract Oncol* 2008;5(10):588–99.
7. Kropotova E.S., Zinov'eva O.L., Zyrianova A.F. et al. Expression of genes involved in retinoic acid biosynthesis in human gastric cancer. *Mol Biol* 2013;47(2):280–92.
8. Никитина И.Г., Сабирова Е.Ю., Солопова О.Н. и др. Новый формат иммуно-ПЦР для сывороточной диагностики рака толстой кишки. *Молекулярная биология* 2014;48(1):117–23. [Nikitina I.G., Sabirova E.Yu., Solopova O.N. et al. A new immuno-PCR format for serological diagnosis of colon cancer. *Molekulyarnaya biologiya = Molecular Biology* 2014;48(1):117–23. (In Russ.)].
9. Maconi G., Manes G., Porro G.B. Role of symptoms in diagnosis and outcome of gastric cancer. *World J Gastroenterol* 2008;14(8):1149–55.
10. Watanabe Y., Itoh S., Goto T. et al. *TMEPA1*, a transmembrane TGF- β -inducible protein, sequesters Smad proteins from active participation in TGF- β signaling. *Mol Cell* 2010;37(1):123–34.

11. Anders S., Huber W. Differential expression analysis for sequence count data. *Genome Biol* 2010;11(10):R106.
12. Bullard J.H., Purdom E., Hansen K.D., Dudoit S. Evaluation of statistical methods for normalization and differential expression in mRNA-Seq experiments. *BMC Bioinformatics* 2010;11:94.
13. Benjamini Y., Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J Royal Statistic Society, Series B: Methodol* 1995;57:289–300.
14. Volkomorov V., Grigoryeva E., Krasnov G. et al. Search for potential gastric cancer markers using miRNA databases and their gene expression analysis. *Exp Oncol* 2013; 35(1):1–6.
15. Rajkumar T., Vijayalakshmi N., Gopal G. et al. Identification and validation of genes involved in gastric tumorigenesis. *Cancer Cell Int* 2010;10:45.
16. Sharad S., Ravindranath L., Haffner M.C. et al. Methylation of the PMEPA1 gene, a negative regulator of the androgen receptor in prostate cancer. *Epigenetics* 2014;9(6):918–27.
17. Lamouille S., Xu J., Derynck R. Molecular mechanisms of epithelial-mesenchymal transition. *Nat Rev Mol Cell Biol* 2014;15(3):178–96.
18. Arvelo F., Sojo F., Cotte C. Tumour progression and metastasis. *Ecancermedalscience* 2016;10:617.
19. Sato T., Oshima T., Yamamoto N. et al. Clinical significance of SPARC gene expression in patients with gastric cancer. *J Surg Oncol* 2013;108(6):364–8.
20. Yin J., Chen G., Liu Y. et al. Downregulation of SPARC expression decreases gastric cancer cellular invasion and survival. *J Exp Clin Cancer Res* 2010;29:59.