

МОЛЕКУЛЯРНЫЙ ЛАНДШАФТ ЗЛОКАЧЕСТВЕННЫХ ОПУХОЛЕЙ ТОЛСТОЙ КИШКИ: ИНТЕГРАТИВНЫЙ АНАЛИЗ ИЗМЕНЕНИЙ

Кутилин Д. С., Худына Ю. Е., Маслов А. А., Колесников Е. Н., Легенько Н. Н.,
Фомин А. А., Анисимов А. Е., Аракелова А. Ю., Чалхакян Л. Х., Дашков А. В.

Федеральное государственное бюджетное учреждение «Национальный медицинский исследовательский центр онкологии» Министерства здравоохранения Российской Федерации, Ростов-на-Дону, Российская Федерация, e-mail: k.denees@yandex.ru

Колоректальный рак является одним из наиболее распространенных злокачественных новообразований в мире, занимая третье место по заболеваемости и второе по смертности среди всех онкологических заболеваний. Для понимания основ и деталей патогенеза, а также учитывая актуальность поиска новых высокоспецифичных молекулярных маркеров для ранней диагностики и прогноза, необходим комплексный анализ молекулярных данных опухолевой ткани в сравнении с условно здоровой. Поэтому целью настоящего исследования стал интегративный анализ молекулярных изменений при колоректальном раке на основе данных TCGA, включая анализ копийности и экспрессии генов, экспрессии некодирующих РНК, мутационного профиля и эпигенетических аномалий, для выявления потенциальных биомаркеров. Данные были получены из базы данных TCGA. В анализ был включен 451 образец первичной колоректальной аденокарциномы. Использовались следующие типы данных: сегментированные данные копийности, данные секвенирования экзонов (WES), данные экспрессии мРНК (RNA-seq) и данные метилирования ДНК (Illumina Infinium HumanMethylation450). Выявлены изменения копийности, включая амплификации в 20q, 8q, 13q, 7p и 7q, и делеции в 18q, 18p, 17p, 16p и 4q, которые затрагивают ключевые онкогены и супрессоры опухолей. Также идентифицировано 1436 значимо мутированных генов, включая классические онкогены APC, TP53, KRAS, PIK3CA, а также гены ARID1A, RNF43 и FBXW7. Обнаружены сильные корреляции между копийностью и экспрессией генов ($> 0,8$ для 5440 генов), а также между метилированием ДНК и транскрипцией ($< -0,8$ для 25 генов). Выявлено значимое изменение 229 сигнальных путей, включая пути регуляции иммунной системы, клеточной адгезии и ключевые онкогенные сигнальные каскады (Wnt, MAPK, PI3K, TGF- β , p53). Подтверждена реактивация раково-тестикулярных антигенов (MAGEA3, PRAME и др.) в опухолевой ткани, что поддерживает их потенциал как иммунотерапевтических мишеней. Результаты исследования имеют потенциальное значение для разработки персонализированных подходов к лечению колоректального рака, включая таргетную терапию и иммунотерапию.

Ключевые слова: колоректальный рак, метилирование, копийность генов, экспрессия генов, некодирующие РНК, раково-тестикулярные антигены.

MOLECULAR LANDSCAPE OF COLON CANCER: AN INTEGRATIVE ANALYSIS OF CHANGES

Kutulin D. S., Khudyna Yu. E., Maslov A. A., Kolesnikov E. N., Legenko N. N.,
Fomin A. A., Anisimov A. E., Arakelova A. Yu., Chalkhakhyan L. Kh., Dashkov A. V.

Federal State Budgetary Institution "National Medical Research Center for Oncology" of the Ministry of Health of the Russian Federation, Rostov-on-Don, Russian Federation, e-mail: k.denees@yandex.ru

Colorectal cancer is one of the most common malignancies worldwide, ranking third in incidence and second in mortality among all oncological diseases. To understand the fundamentals and details of pathogenesis, and given the relevance of the search for new highly specific molecular markers for early diagnosis and prognosis, a comprehensive analysis of molecular data of tumor tissue compared to conventionally healthy tissue is necessary. Therefore, the aim of this study was an integrative analysis of molecular changes in colorectal cancer based on TCGA data, including analysis of gene copy number and expression, expression of non-coding RNAs, mutational profile, and epigenetic changes. Data were obtained from the TCGA database. The analysis included 451 samples of primary colorectal adenocarcinoma. The following data types were used: segmented copy number data, whole-exome sequencing (WES) data, mRNA expression data (RNA-seq), and DNA methylation data (Illumina Infinium Human Methylation450). Copy number changes were identified, including amplifications at 20q, 8q, 13q, 7p, and 7q, and deletions at 18q, 18p, 17p, 16p, and 4q, affecting key oncogenes and tumor suppressors. A total of 1,436 significantly mutated genes were also identified, including the classical oncogenes APC, TP53, KRAS, PIK3CA, and ARID1A, RNF43, and FBXW7. Strong correlations were found between copy number and gene expression

(>0.8 for 5,440 genes), as well as between DNA methylation and transcription (<-0.8 for 25 genes). Significant changes in 229 signaling pathways were identified, including those regulating the immune system, cell adhesion, and key oncogenic signaling cascades (Wnt, MAPK, PI3K, TGF- β , and p53). Reactivation of cancer-testicular antigens (MAGEA3, PRAME, and others) in tumor tissue was confirmed, supporting their potential as immunotherapeutic targets. The study results have potential implications for the development of personalized approaches to colorectal cancer treatment, including targeted therapy and immunotherapy.

Ключевые слова: colorectal cancer, methylation, gene copy number, gene expression, non-coding RNA, cancer-testicular antigens.

Введение

Колоректальный рак (КРР) является одним из наиболее распространенных злокачественных новообразований в мире, занимая третье место по заболеваемости и второе по смертности среди всех онкологических заболеваний [1]. Ежегодно в мире регистрируется более 1,9 млн новых случаев заболевания, а смертность достигает 935 тыс. случаев в год [2]. Несмотря на значительные успехи в диагностике и лечении, выживаемость пациентов с метастатической формой заболевания остается низкой, составляя менее 15 % за пятилетний период [3].

Молекулярный патогенез КРР характеризуется накоплением генетических и эпигенетических изменений, которые приводят к нарушению нормальной регуляции клеточного цикла, апоптоза, дифференцировки и миграции [4]. Классическая модель развития КРР, предложенная Fearon и Vogelstein, описывает последовательное накопление мутаций в генах *APC*, *KRAS*, *TP53* и др., приводящее к прогрессии от нормальной слизистой к аденоме и далее к аденокарциноме [4]. Особый интерес представляют раково-тестикулярные антигены (СТА) – группа белков, экспрессия которых в норме ограничена клетками зародышевой линии, но реактивируется в различных опухолях, включая КРР. Благодаря высокой иммуногенности и ограниченному профилю экспрессии СТА рассматриваются как перспективные мишени для иммунотерапии [5]. Кроме того, значительный вклад в регуляцию онкогенеза вносят некодирующие РНК – длинные некодирующие РНК (lncRNA) и микроРНК (miRNA), которые модулируют экспрессию генов на транскрипционном и посттранскрипционном уровнях, в том числе участвуя в регуляции экспрессии раково-тестикулярных антигенов [6, 7].

В последние годы были выявлены различные молекулярные подтипы КРР, которые отличаются по клиническим проявлениям, прогнозу и ответу на терапию [8]. Проект TCGA (The Cancer Genome Atlas) предоставил беспрецедентный объем данных о геномных, транскриптомных, эпигеномных и протеомных изменениях при различных типах рака, включая КРР [9]. Для понимания основ и деталей патогенеза, а также учитывая актуальность поиска новых высокоспецифичных молекулярных маркеров для ранней диагностики и прогноза [10], необходим комплексный анализ молекулярных данных опухолевой ткани в сравнении с условно здоровой.

Цель исследования – интегративный анализ молекулярных изменений при КРР на основе данных TCGA, включая анализ копийности и экспрессии генов, экспрессии некодирующих РНК, мутационного профиля и эпигенетических аномалий, для выявления потенциальных биомаркеров.

Материал и методы исследования

Источники данных. Данные были получены из базы данных TCGA (The Cancer Genome Atlas, <https://portal.gdc.cancer.gov>, данные загружены в марте 2024 г., версия данных: GDC Data Release 32.0). В анализ был включен 451 образец первичной колоректальной аденокарциномы (COAD – 321, READ 130). Использованы парные образцы опухоль / нормальная ткань (41 образец нормальной слизистой из TCGA, а также дополнительные нормальные ткани из базы GTEx v8 для валидации). Использовались следующие типы данных: сегментированные данные копийности, данные секвенирования экзонов (WES), данные экспрессии мРНК (RNA-seq) и данные метилирования ДНК (Illumina Infinium HumanMethylation450). Важно отметить, что были использованы предварительно обработанные данные RNA-seq образцов опухолевой и нормальной ткани толстой кишки. Для каждого гена рассчитаны медианные значения экспрессии в группах «опухоль» и «норма», логарифмическое изменение (Log_2FC) и скорректированное p -значение (adj.p). Анализ включал 12 345 транскриптов (уникальных Ensembl ID).

Анализ копийности генов. Для идентификации значимых амплификаций и делеций использовался алгоритм GISTIC 2.0.22 [11]. Параметры анализа: порог для амплификации: 0,1, порог для делеции: 0,1, максимальное количество сегментов на образец: 2000, ширина пика: 0,7, уровень доверия: 0,99. Значимость определялась по $Q\text{-value} \leq 0,25$. Результаты включали: фокальные амплификации и делеции, а также изменения на уровне плеч хромосом.

Анализ экспрессии генов. Ген считался дифференциально экспрессированным (DEG) при $|\text{Log}_2\text{FC}| \geq 1$ и $\text{adj.p} < 0,05$. Для выделения наиболее значимых изменений использован порог $|\text{Log}_2\text{FC}| \geq 2$. Для визуализации использованы библиотеки Plotly. Volcano plot построен для всех генов, отвечающих порогу по adj.p ($p < 0,05$). Столбчатые диаграммы отображают топ-20 генов с наибольшим повышением и понижением экспрессии. Таблицы включают детальную информацию по DEG, CTA, lncRNA и miRNA. Раково-тестикулярные антигены идентифицировали на основе курируемой базы CTDatabase [12] и перекрестной проверки по литературным источникам. Длинные некодирующие РНК (lncRNA) и микроРНК (miRNA) определяли по номенклатуре Gene Symbol (наличие префиксов LINC, MIR, суффиксов -AS, а также ручная аннотация).

Анализ мутаций. Для выявления значимо мутированных генов использовался алгоритм MutSig 2CV v3.1 [13]. Критерии значимости: $Q\text{-value} \leq 0,1$. Анализ включал оценку мутационной нагрузки, кластеризации мутаций и функционального воздействия.

Корреляционный анализ. Корреляции между копийностью генов и экспрессией мРНК оценивали с использованием коэффициента Пирсона. Корреляции между метилированием ДНК и экспрессией генов оценивались с использованием коэффициента Спирмена с применением поправки Бенджамини – Хохберга (FDR) для контроля ложноположительных результатов ($Q\text{-value} \leq 0,05$).

Анализ обогащения сигнальных путей. Для идентификации значимо обогащенных сигнальных путей использовался гипергеометрический тест с поправкой на множественные сравнения (FDR). База данных генных наборов: MSigDB v4.0 (c2.cp).

Внешняя валидация. Для подтверждения ключевых результатов (основные амплификации/делеции, DEG, мутации) использована независимая когорта из базы GEO (GSE143985, 98 образцов KPP) и данные CPTAC (Colon Cancer, 110 образцов).

Статистический анализ. Статистический анализ проводился с использованием программного обеспечения R v4.1.0. Для визуализации данных использовались пакеты ggplot2, pheatmap и ComplexHeatmap.

Результаты исследования и их обсуждение

Изменения копийности генов. Анализ GISTIC выявил 22 статистически значимые фокальные амплификации ($Q\text{-value} \leq 0,25$) и 44 статистически значимые ($Q\text{-value} \leq 0,25$) фокальные делеции (рис. 1). Наиболее частые амплификации наблюдались в хромосомах 20q, 8q, 13q и 17q, что согласуется с предыдущими исследованиями при KPP [14]. Наиболее частые делеции выявлены в хромосомах 16p, 4q, 6q и 18q. Анализ на уровне плеч хромосом выявил 23 значимых результата. Наиболее частые амплификации наблюдались в плечах 20q (70 % образцов), 8q (51 % образцов), 13q (58 % образцов), 7p (55 % образцов) и 7q (50 % образцов). Наиболее частые делеции выявлены в плечах 18q (61 % образцов), 18p (59 % образцов), 17p (57 % образцов), 15q (36 % образцов) и 14q (33 % образцов).

Изменения экспрессии генов. В результате анализа выявлено 2847 дифференциально экспрессирующихся генов (при пороге $|\text{Log2FC}| \geq 1$ и $\text{adj.p} < 0,05$), что составляет приблизительно 14 % от общего числа анализируемых генов. Из них 1523 гена (53,5 %) были гиперэкспрессированы в опухолевой ткани, а 1324 гена (46,5 %) – гипоекспрессированы. При более строгом пороге ($|\text{Log2FC}| \geq 2$, $\text{adj.p} < 0,05$) выявлено 140 дифференциально экспрессирующихся генов, из них 121 с повышенной экспрессией и 19 с пониженной.

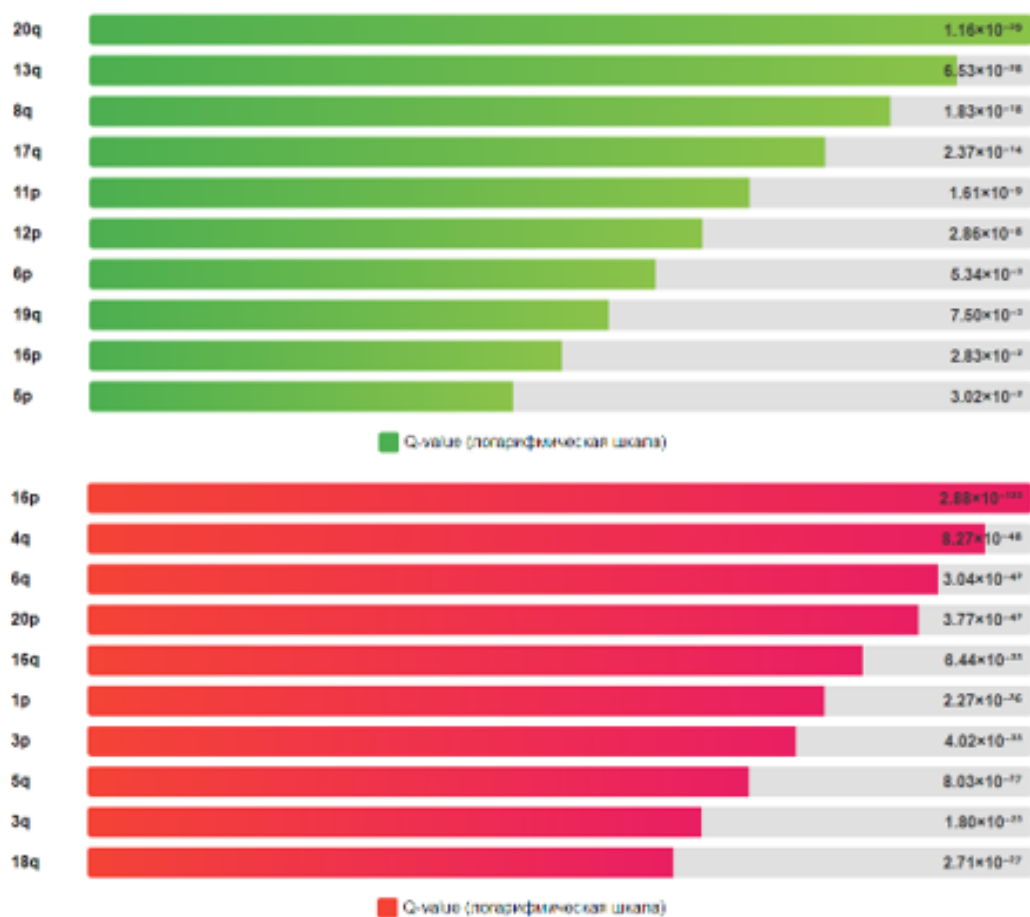


Рис. 1. Распределение амплификаций (зеленый цвет) и делеций (красный цвет) по хромосомам (топ-10).

Примечание: составлен авторами по результатам данного исследования

Наиболее значительное увеличение экспрессии ($\text{Log}_2\text{FC} > 4$) демонстрируют гены, вовлеченные в клеточную адгезию (*CLDN1*, *CLDN2*, *CLDN4*), пролиферацию (*UBE2C*, *RRM2*, *TK1*) и воспалительный ответ (*CXCL8*, *CCL20*) (рис. 2).

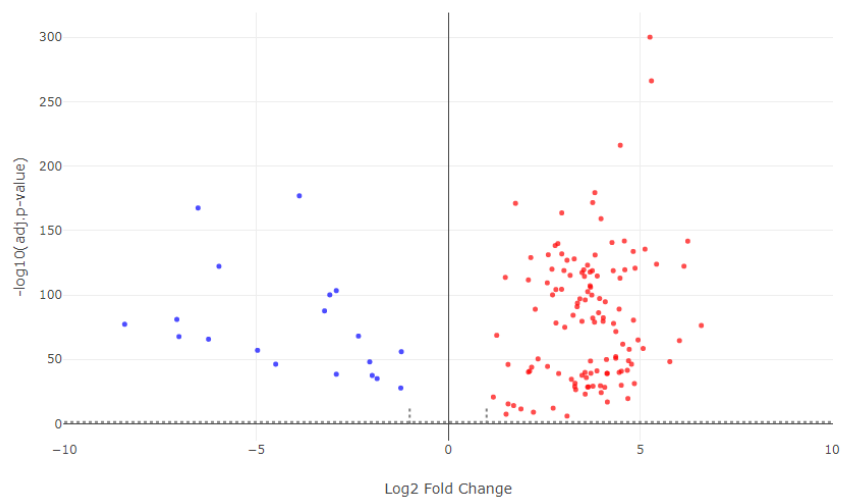


Рис. 2. Volcano plot дифференциальной экспрессии генов при КРР.

По оси X – $\text{Log}_2(\text{Fold Change})$ опухоль/норма, по оси Y – $\log_{10}(\text{adj.p-value})$.

Красные точки – значимо повышенная экспрессия ($\text{adj.p} < 0,05$, $\text{Log}_2\text{FC} > 1$),

синие – значимо пониженная ($\text{adj.p} < 0,05$, $\text{Log}_2\text{FC} < -1$).

Серые точки – незначимые изменения. Горизонтальная линия

соответствует $\text{adj.p} = 0,05$, вертикальные – $\text{Log}_2\text{FC} = \pm 1$.

Примечание: составлен авторами по результатам данного исследования

Среди генов с выраженным снижением ($\text{Log}_2\text{FC} < -4$) – компоненты мышечного сокращения (*ACTG2*, *MYH11*, *CNN1*, *DES*) и гены, ассоциированные с нормальной функцией стромы (*ADAM33*, *DPT*) (рис. 2 и 3).

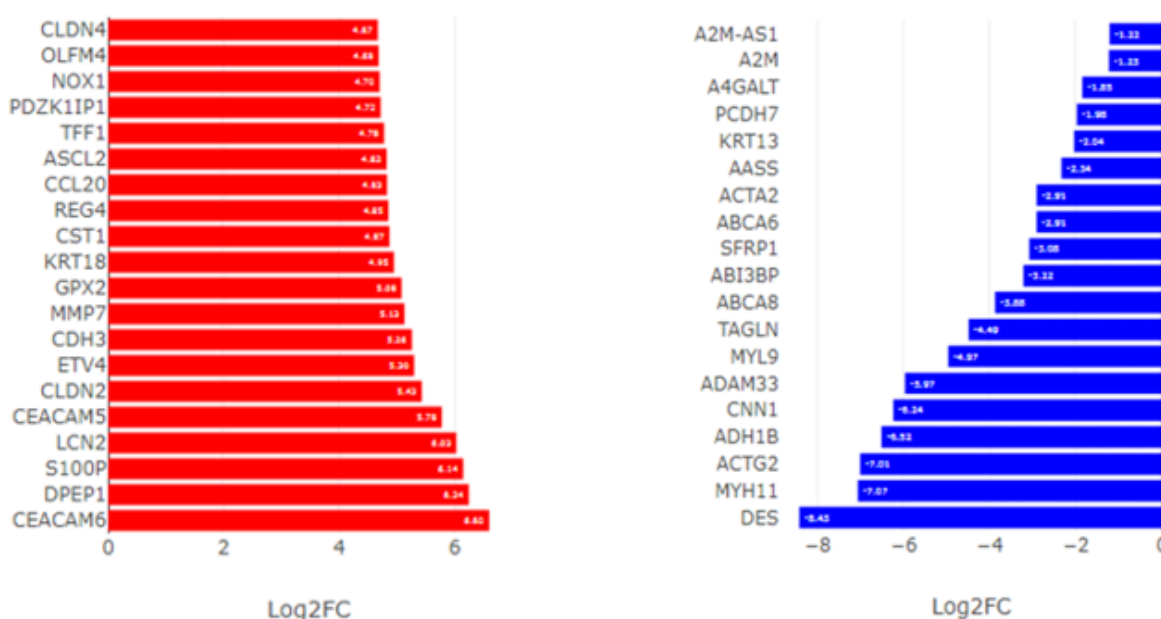


Рис. 3. Топ-20 генов с наиболее выраженным повышением (слева) и понижением (справа) экспрессии в опухоли относительно нормальной ткани.

Значения представлены как $\text{Log}_2(\text{Fold Change})$.

Примечание: составлен авторами по результатам данного исследования

В анализируемом наборе данных идентифицировано 3 гена, относящихся к семейству раково-тестикулярных антигенов (РТА). Большинство из них показали повышенную экспрессию в опухолевых образцах (табл. 1).

Среди наиболее значимых – *MAGEA3*, *MAGEA6*, *PRAME*. Эти данные согласуются с предыдущими исследованиями, указывающими на реактивацию РТА при КРР [15].

Таблица 1

РТА, дифференциально экспрессирующиеся при КРР

Ген	log ₂ FC	P _{adj}
<i>MAGEA3</i>	4,512	1,20e-30
<i>MAGEA6</i>	3,987	5,60e-25
<i>PRAME</i>	3,876	8,30e-42

Примечание: составлена авторами на основе полученных данных в ходе исследования

Также выявлено 23 дифференциально экспрессированных микроРНК, включая *miR-21* ($\log_2FC = 2,8$, $adjp = 1,2e-45$), которая известна как онкогенная микроРНК, способствующая пролиферации и инвазии клеток КРР.

Таблица 2

МикроРНК, дифференциально экспрессирующиеся при КРР

МикроРНК	log ₂ FC	P _{adj}	Известные функции при КРР
<i>miR-21</i>	+2,80	1,20e-45	Онкогенная, подавляет <i>PTEN</i>
<i>miR-143</i>	-1,90	3,40e-38	Опухолевый супрессор
<i>miR-145</i>	-2,10	1,80e-42	Опухолевый супрессор
<i>miR-155</i>	+2,30	5,60e-51	Онкогенная, регуляция иммунного ответа

Примечание: составлена авторами на основе полученных данных в ходе исследования

Идентифицировано 156 lncRNA с измененной экспрессией. Среди них *CRNDE* (Colorectal Neoplasia Differentially Expressed) показал $\log_2FC = 2,66$ ($adjp = 1,82e-102$). Эта lncRNA способствует пролиферации клеток КРР через активацию Wnt/ β -катенинового сигнального пути [16].

Таблица 3

lncRNA, дифференциально экспрессирующиеся при КРР

lncRNA	log ₂ FC	P _{adj}	Функциональная роль
<i>CRNDE</i>	+2,66	1,82e-102	Активация Wnt/ β -катенина
<i>HOTAIRM1</i>	-1,82	1,15e-65	Регуляция HOXA генов
<i>PVT1</i>	+1,44	5,21e-42	Онкогенная, регуляция MYC
<i>MALAT1</i>	-2,83	1,23e-64	Регуляция сплайсинга

Примечание: составлена авторами на основе полученных данных в ходе исследования

Корреляция копийности генов, экспрессии мРНК и эпигенетических изменений.

Анализ выявил 5440 генов с положительной корреляцией между показателями копийности и экспрессией мРНК (90-й перцентиль). Наиболее сильные корреляции ($> 0,8$) наблюдались для генов на хромосоме 20 (рис. 4).

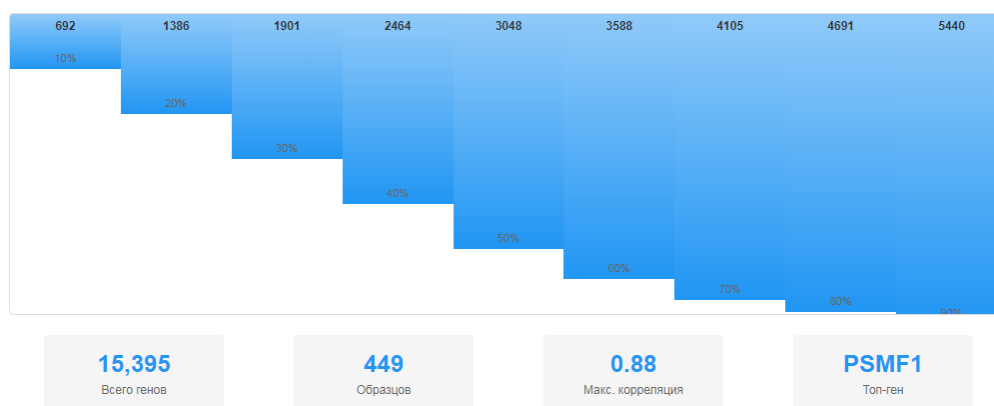


Рис. 4. Распределение коэффициентов корреляции копийности и экспрессии генов.

Примечание: составлен авторами по результатам данного исследования

Анализ выявил сильную отрицательную корреляцию между метилированием ДНК и экспрессией генов. Наиболее сильные отрицательные корреляции ($< -0,8$) наблюдались для генов, представленных на рис. 5.

Мутационный профиль. Анализ MutSig 2CV выявил 1436 значимо мутированных генов ($q \leq 0,1$ в GDAC-отчете). Биологическая значимость выявленных мутаций требует дополнительной фильтрации. Наиболее часто мутированные гены представлены на рис. 6.

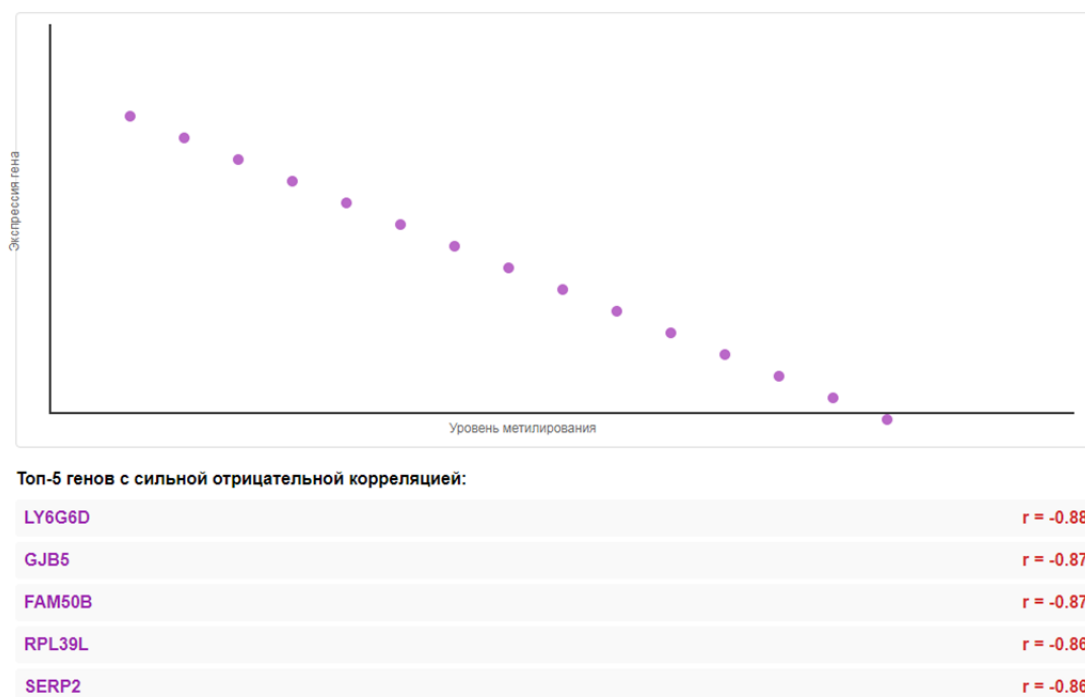


Рис. 5. Отрицательная корреляция метилирования промоторов и экспрессии генов.

Примечание: составлен авторами по результатам данного исследования

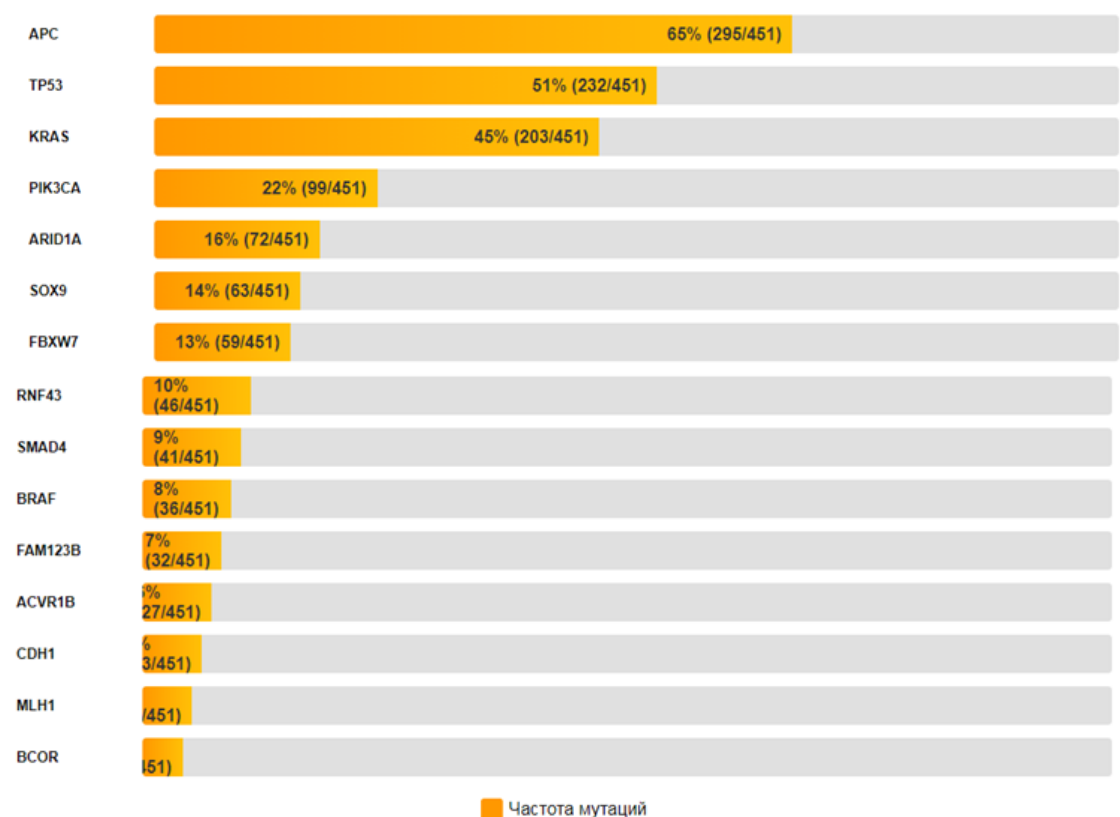


Рис. 6. Частота мутаций в топ-15 генах при KPP.

Примечание: составлен авторами по результатам данного исследования

Обогащение сигнальных путей. Анализ обогащения выявил 229 значимо измененных сигнальных путей ($q \leq 0,01$). Наиболее значимые пути представлены на рис. 7.

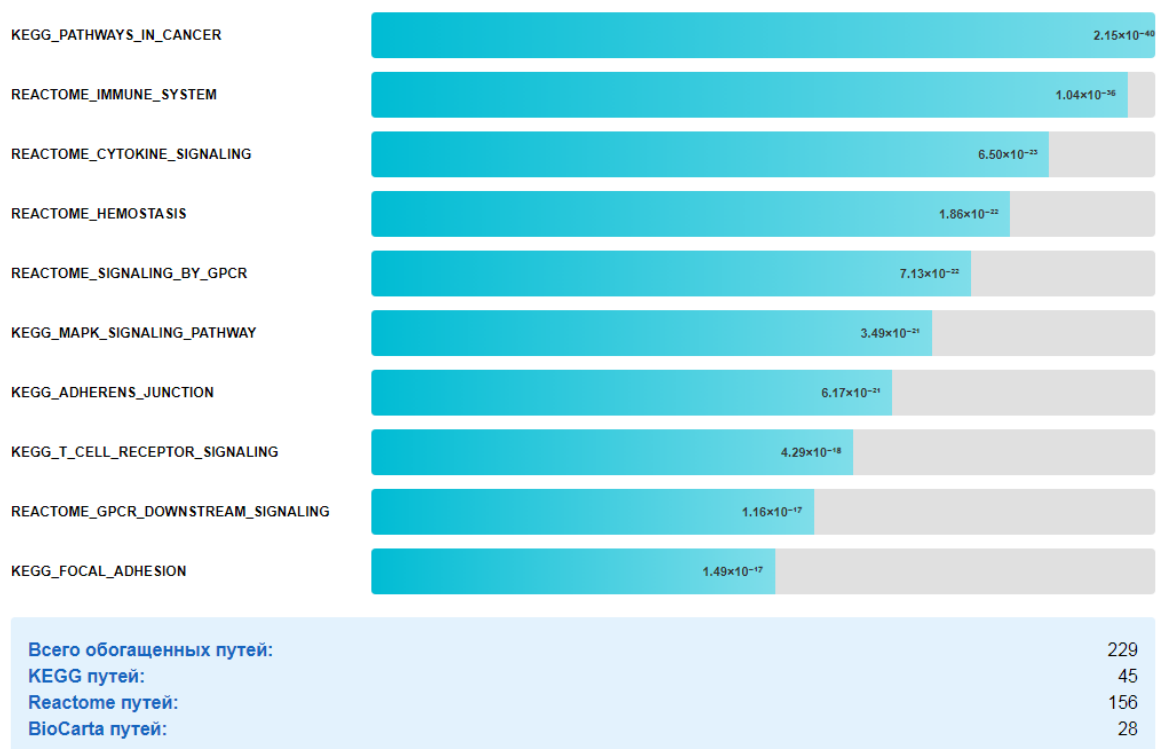


Рис. 7. Топ-10 обогащенных сигнальных путей при KPP.

Примечание: составлен авторами по результатам данного исследования

Изменения копийности генов. Выявленные амплификации в 20q, 8q и 13q согласуются с предыдущими исследованиями КРР [17]. Хромосома 20q содержит ряд онкогенов, включая *AURKA*, *TPX2*, *BCL2L1* и *MYBL2*, которые играют ключевую роль в регуляции клеточного цикла, апоптоза и хромосомной сегрегации [18]. Амплификация 8q24.21 включает ген *MYC*, который является ключевым регулятором пролиферации, метаболизма и апоптоза опухолевых клеток [19]. Амплификация 13q12.13 затрагивает ген *USP12*, который кодирует дубиквитирующую протеазу и может способствовать стабилизации онкогенных белков [20]. Делеции в 16p, 4q и 6q затрагивают потенциальные супрессоры опухолей. Регион 16p13.3 содержит ген *CREBBP*, который функционирует как коактиватор транскрипции и регулятор ацетилирования гистонов [21]. Делеции в 18q включают гены *SMAD2*, *SMAD4* и *DCC*, которые являются ключевыми компонентами сигнального пути TGF- β и регуляторами клеточной адгезии [22]. Делеция 4q22.1 затрагивает ген *CCAR1* (холецистокининовый рецептор), который может играть роль в регуляции пролиферации эпителиальных клеток [23]. Интересно, что амплификации в 7p и 7q наблюдались в 55 и 50 % образцов соответственно. Хромосома 7 содержит гены *EGFR* (7p11.2) и *BRAF* (7q34), которые являются ключевыми компонентами сигнального пути MAPK и потенциальными мишенями для таргетной терапии [24].

Транскриптомный профиль. Проведенный анализ подтвердил значительные изменения транскриптомного профиля в опухолях толстой кишки. Среди генов с максимальным повышением экспрессии – *CLDN1*, *CLDN2*, *CLDN4*, кодирующие белки плотных контактов; их гиперэкспрессия характерна для КРР и ассоциирована с инвазивностью [25]. Гены *UBE2C*, *RRM2*, *TK1* отражают высокую пролиферативную активность опухолевых клеток. Повышение *CXCL8*, *CCL20* указывает на провоспалительное микроокружение, способствующее опухолевой прогрессии [26]. Снижение экспрессии генов гладкомышечного сокращения (*ACTG2*, *MYH11*, *CNN1*, *DES*) является отражением ремоделирования стромы и потери мышечного слоя при инвазии опухоли. Гены *CAV1*, *CAV2*, кодирующие caveолины, также значительно подавлены, что может влиять на сигнальные пути TGF- β и Wnt [27].

Выявленные раково-тестикулярные антигены (*MAGEA3*, *MAGEA6*, *PRAME*) представляют особый интерес для разработки вакцин и адоптивной иммунотерапии КРР, поскольку их экспрессия практически отсутствует в нормальных соматических тканях взрослого человека [28]. Полученные данные подтверждают их значительную активацию в опухолях толстой кишки.

Роль некодирующих РНК в патогенезе КРР. МикроРНК относятся к коротким некодирующим РНК, которые регулируют экспрессию генов, катализируя разрушение мРНК,

либо ингибируя трансляцию мРНК в белок [29]. МикроРНК вносят значительный вклад в инициацию и развитие различных молекулярных событий, включая инициацию онкогенеза, прогрессирование и метастазирование опухолей, что делает микроРНК потенциальными биомаркерами для оценки прогрессирования и прогноза КРР. Хотя микроРНК регулируют экспрессию генов, кодирующих белки, главным образом посредством деградации или сайленсинга мРНК, появляется все больше свидетельств того, что микроРНК могут взаимодействовать с lncRNA, что, в свою очередь, также обеспечивает регуляцию экспрессии генов-мишеней [30]. Выявленные микроРНК и lncRNA участвуют в ключевых сигнальных путях канцерогенеза. Так, miR-21, наиболее изученная онкогенная микроРНК при КРР, подавляет экспрессию опухолевых супрессоров *PTEN* и *PDCD4*, способствуя пролиферации и выживанию опухолевых клеток [31]. miR-143/miR-145 кластер – пониженная экспрессия этих опухолевых супрессорных микроРНК характерна для КРР. Они регулируют экспрессию *KRAS*, *MYC* и других онкогенов [32]. lncRNA *CRNDE* действует как «спонж» для miR-181a-5p, модулируя Wnt/ β -катениновый сигнальный путь. Высокая экспрессия *CRNDE* коррелирует с плохим прогнозом [16]. Пониженная экспрессия lncRNA *HOTAIRM1* ($\log_2FC = -1,82$) ассоциирована с прогрессией КРР через регуляцию генов *HOXA* кластера [33].

Мутационный профиль. Высокая частота мутаций в генах *APC* (65 % пациентов) и *TP53* (51 % пациентов) подтверждает их ключевую роль в патогенезе КРР [34]. Мутации в гене *APC* приводят к конститутивной активации сигнального пути Wnt/ β -catenin, что является ранним событием в развитии КРР. Мутации в гене *TP53* нарушают функции опухолевого супрессора p53, включая контроль клеточного цикла, апоптоз и репарацию ДНК [35]. Мутации в гене *KRAS* (45 % пациентов) являются важным прогностическим и предиктивным фактором. Мутации *KRAS* ассоциированы с резистентностью к анти-EGFR терапии (цетуксимаб, панитумумаб) [36]. Мутации в гене *PIK3CA* (22 % пациентов) активируют сигнальный путь PI3K/AKT/mTOR, что способствует пролиферации и выживанию опухолевых клеток [37]. Мутации в гене *ARID1A* (16 % пациентов) указывают на нарушение комплекса SWI/SNF, что согласуется с данными о его роли в хроматиновом ремоделировании и регуляции транскрипции [38]. Мутации в гене *RNF43* (10 % пациентов) указывают на нарушение регуляции сигнального пути Wnt через убиквитинирование рецепторов франтиза [39]. Мутации в гене *MLH1* (4 % пациентов) указывают на дефекты системы репарации неспаренных оснований (MMR), что характерно для микросателлитно-нестабильных (MSI) опухолей [40]. Однако частота мутаций в гене *MLH1* в проведенном исследовании ниже, чем ожидаемая для спорадических КРР с дефицитом MMR, что может указывать на эпигенетические механизмы инактивации этого гена.

Корреляция копийности и экспрессии генов. Сильные положительные корреляции между копийностью и экспрессией генов ($> 0,8$) наблюдались для генов на хромосоме 20 (*PSMF1*, *PTPRA*, *YTHDF1*, *RTFDC1*, *TRPC4AP*). Это указывает на то, что амплификации в этих регионах функционально значимы и приводят к увеличению экспрессии соответствующих генов [41]. Ген *PSMF1* кодирует протеасомный ингибитор, который может способствовать выживанию опухолевых клеток через ингибирование протеасомной деградации онкогенных белков [42]. Ген *PTPRA* кодирует протеин-тирозинфосфатазу, которая может регулировать сигнальные пути, включая Src и MAPK [43]. Ген *YTHDF1* кодирует белок, связывающийся с м6А-модифицированной мРНК и регулирующий трансляцию, что может способствовать экспрессии онкогенов [44].

Эпигенетические изменения. Сильная отрицательная корреляция между метилированием промоторов и экспрессией генов подтверждает важную роль эпигенетических механизмов в регуляции транскрипции при КРР. Гиперметилирование промоторов генов-супрессоров опухолей является одним из ключевых механизмов их инактивации [45]. Ген *LY6G6D*, показавший наиболее сильную отрицательную корреляцию ($r = -0,88$), кодирует белок, связанный с GPI-якорем, функция которого при КРР не до конца изучена [46]. Ген *GJB5* кодирует коннексин 31.1, который участвует в формировании щелевых контактов и может регулировать межклеточную коммуникацию [47]. Ген *FAM50B* экспрессируется преимущественно в яйцках, и его роль при КРР требует дальнейшего изучения [48].

Сигнальные пути. Значимое обогащение путей, связанных с канцерогенезом, иммунной системой и сигнальными каскадами, указывает на комплексность молекулярных изменений при КРР. Обогащение путей адгезии (адгеренс-юнкшн, фокальная адгезия) согласуется с ролью нарушений клеточной адгезии в инвазии и метастазировании [35]. Обогащение путей, связанных с иммунной системой, указывает на важную роль иммунного микроокружения в прогрессии КРР. Это согласуется с данными о том, что иммунная инфильтрация является важным прогностическим фактором при КРР [35]. На рис. 8 представлена итоговая схема молекулярных изменений при КРР.

Ключевые результаты (основные амплификации 20q, 8q и делеции 18q, мутации *APC/KRAS*, повышение *MAGEA3*) были подтверждены в независимой когорте GSE143985 (98 образцов) и данных CPTAC (Colon Cancer). Коэффициенты корреляции копийность – экспрессия для генов-кандидатов (*PSMF1*, *YTHDF1*) составили 0,75–0,82 ($p < 0,001$).

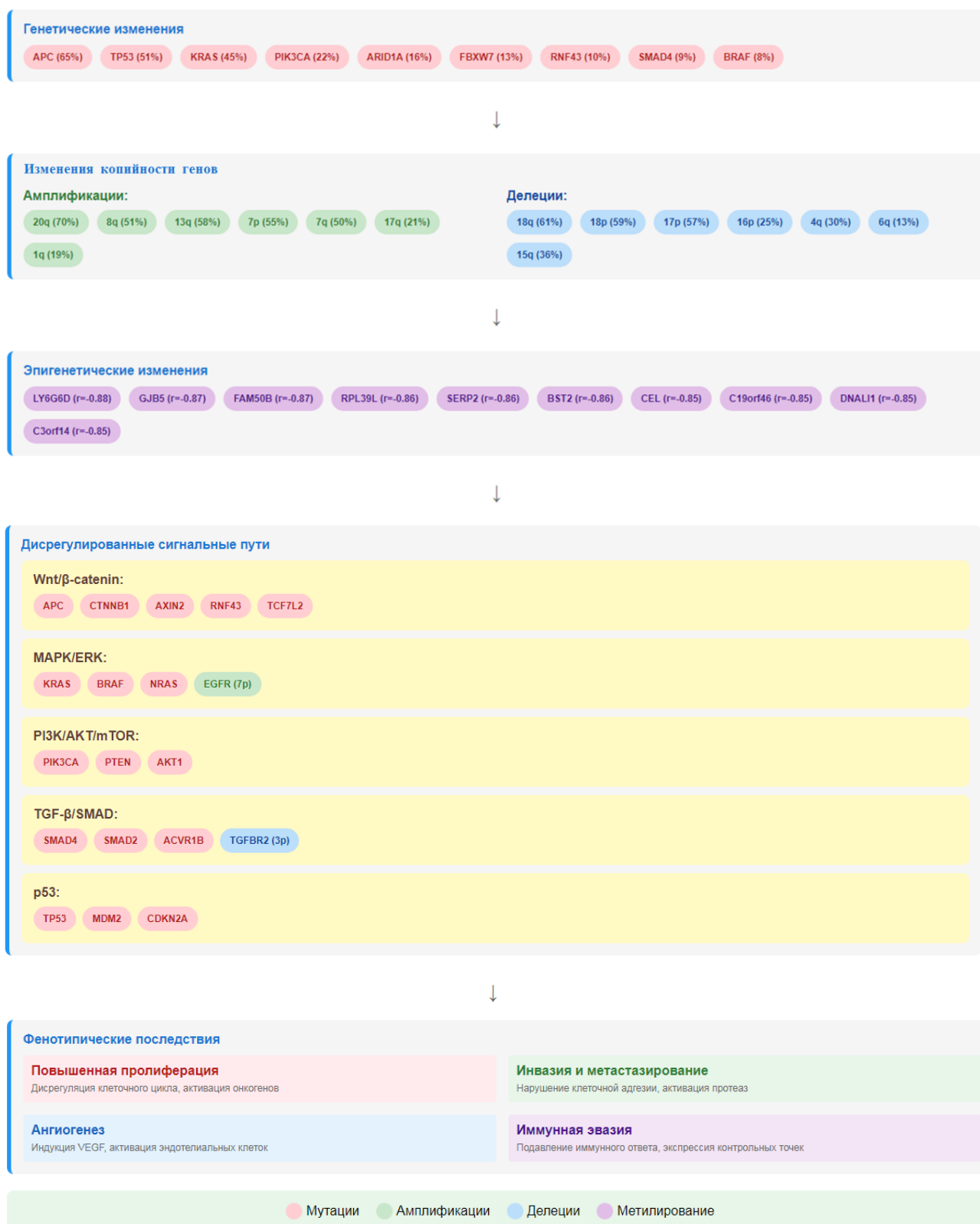


Рис. 8. Интегративная схема молекулярных изменений при КРР.

Примечание: составлен авторами по результатам данного исследования

Ограничения исследования. Несмотря на масштаб интегрированного анализа, исследование имеет ряд ограничений: проведен ретроспективный анализ данных TCGA, а это предполагает возможные систематические смещения (batch effects) при малой

представленности нормальных тканей (относительно опухолевых) и учитывая обработку образцов в разных центрах. Также нами не проводилась функциональная валидация (необходимы эксперименты *in vitro/in vivo*), не учитывалось детально влияние молекулярных подтипов (MSI, CIMP, CIN) опухоли и неоднородность COAD/READ. При этом статистический список 1436 мутированных генов не эквивалентен клинически значимым драйверным мутациям; требуется дополнительная фильтрация.

Заключение

Таким образом, выявлены изменения копийности, включая амплификации в 20q (70 %), 8q (51 %), 13q (58 %), 7p (55 %) и 7q (50 %), и делеции в 18q (61 %), 18p (59 %), 17p (57 %), 16p (25 %) и 4q (30 %), которые затрагивают ключевые онкогены и супрессоры опухолей. Также идентифицировано 1436 значимо мутированных генов, включая классические онкогены *APC* – 65 %, *TP53* – 51 %, *KRAS* – 45 %, *PIK3CA* – 22 %, а также гены *ARID1A* – 16 %, *RNF43* – 10 % и *FBXW7* – 13 %. Обнаружены сильные корреляции между копийностью и экспрессией генов (> 0,8 для 5440 генов), а также между метилированием ДНК и транскрипцией (< -0,8 для 25 генов).

Выявлено значимое изменение 229 сигнальных путей, включая пути канцерогенеза, иммунной системы, клеточной адгезии и ключевые онкогенные сигнальные каскады (Wnt, MAPK, PI3K, TGF- β , p53). Обнаружено повышение экспрессии РТА (*MAGEA3*, *PRAME* и др.) в опухолевой ткани, что поддерживает их потенциал как иммунотерапевтических мишеней. Результаты исследования имеют потенциальное значение для разработки персонализированных подходов к лечению КРР, включая таргетную терапию и иммунотерапию. Полученные данные не предназначены для непосредственного применения в клинической практике без дополнительной валидации.

Список литературы

1. Bray F., Ferlay J., Soerjomataram I., Siegel R. L., Torre L. A., Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries // *CA Cancer J Clin.* 2018. Vol. 68 (6). P. 394–424. DOI: 10.3322/caac.21492.
2. Sung H., Ferlay J., Siegel R. L., Laversanne M., Soerjomataram I., Jemal A., Bray F. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries // *CA Cancer J Clin.* 2021. Vol. 71 (3). P. 209–249. DOI: 10.3322/caac.21660. URL: <https://pubmed.ncbi.nlm.nih.gov/33538338> (дата обращения: 10.05.2026).

3. Siegel R. L., Miller K. D., Fuchs H. E., Jemal A. Cancer Statistics, 2022 // *CA Cancer J Clin.* 2022. Vol. 72 (1). P. 7–33. DOI: 10.3322/caac.21708.
4. Fearon E. R., Vogelstein B. A genetic model for colorectal tumorigenesis // *Cell.* 1990. Vol. 61. P. 759–767. URL: <https://pubmed.ncbi.nlm.nih.gov/2188735> (дата обращения: 10.05.2026). DOI: 10.1016/0092-8674(90)90186-I.
5. Кутилин Д. С., Кит О. И. Зависимость выживаемости и метастазирования у больных колоректальным раком от транскрипционной активности РТ-генов // *Сибирский онкологический журнал.* 2022. Т. 21 (1). С. 37–46. DOI: 10.21294/1814-4861-2022-21-1-37-46.
6. Anastasiadou E., Jacob L. S., Slack F. J. Non-coding RNA networks in cancer // *Nat Rev Cancer.* 2018. Vol. 18 (1). P. 5–18. URL: <https://pubmed.ncbi.nlm.nih.gov/29170536> (дата обращения: 10.05.2026).
7. Кутилин Д. С. Регуляция экспрессии генов раково-тестикулярных антигенов у больных колоректальным раком // *Молекулярная биология.* 2020. Т. 54. № 4. С. 580–595. URL: <https://www.elibrary.ru/item.asp?id=42963750> (дата обращения: 10.05.2026). DOI: 10.31857/S0026898420040096.
8. Guinney J., Dienstmann R., Wang X., de Reyniès A., Schlicker A., Soneson C., Marisa L., Roepman P., Nyamundanda G., Angelino P., Bot B. M., Morris J. S., Simon I. M., Gerster S., Fessler E., De Sousa E., Melo F., Missiaglia E., Ramay H., Barras D., Homicsko K., Maru D., Manyam G. C., Broom B., Boige V., Perez-Villamil B., Laderas T., Salazar R., Gray J. W., Hanahan D., Tabernero J., Bernards R., Friend S. H., Laurent-Puig P., Medema J. P., Sadanandam A., Wessels L., Delorenzi M., Kopetz S., Vermeulen L., Tejpar S. The consensus molecular subtypes of colorectal cancer // *Nat Med.* 2015. Vol. 21 (11). P. 1350–1356. DOI: 10.1038/nm.3967.
9. Cancer Genome Atlas Network. Comprehensive molecular characterization of human colon and rectal cancer // *Nature.* 2012. Vol. 487 (7407). P. 330–337. URL: <https://www.nature.com/articles/nature11252> (дата обращения: 10.05.2026). DOI: 10.1038/nature11252.
10. Кутилин Д. С., Худына Ю. Е., Маслов А. А., Колесников Е. Н., Кожушко М. А., Анисимов А. Е., Озеркова Е. А., Камлык Д. В., Легенько Н. Н. Транскрипционный профиль генов раково-тестикулярных антигенов у больных с локализованным и местно-распространенным колоректальным раком // *Современные проблемы науки и образования.* 2026. № 1. URL: <https://science-education.ru/ru/article/view?id = 34445> (дата обращения: 01.06.2026).
11. Mermel C. H., Schumacher S. E., Hill B., Meyerson M. L., Beroukhi R., Getz G. GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers // *Genome Biol.* 2011. Vol. 12 (4). P. R41. URL:

<https://pubmed.ncbi.nlm.nih.gov/21527027> (дата обращения: 10.05.2026). DOI: 10.1186/gb-2011-12-4-r41.

12. CTDatabase. [Электронный ресурс]. URL: <http://www.cta.lncc.br> (дата обращения: 20.05.2026).

13. Lawrence M. S., Stojanov P., Polak P., Kryukov G. V., Cibulskis K., Sivachenko A., Carter S. L., Stewart C., Mermel C. H., Roberts S. A., Kiezun A., Hammerman P. S., McKenna A., Drier Y., Zou L., Ramos A. H., Pugh T. J., Stransky N., Helman E., Kim J., Sougnez C., Ambrogio L., Nickerson E., Shefler E., Cortés M. L., Auclair D., Saksena G., Voet D., Noble M., DiCara D., Lin P., Lichtenstein L., Heiman D. I., Fennell T., Imielinski M., Hernandez B., Hodis E., Baca S., Dulak A. M., Lohr J., Landau D. A., Wu C. J., Melendez-Zajgla J., Hidalgo-Miranda A., Koren A., McCarroll S. A., Mora J., Crompton B., Onofrio R., Parkin M., Winckler W., Ardlie K., Gabriel S. B., Roberts C. W. M., Biegel J. A., Stegmaier K., Bass A. J., Garraway L. A., Meyerson M., Golub T. R., Gordenin D. A., Sunyaev S., Lander E. S., Getz G. Mutational heterogeneity in cancer and the search for new cancer-associated genes // *Nature*. 2013. Vol. 499 (7457). P. 214–218. DOI: 10.1038/nature12213.

14. Douglas E. J., Fiegler H., Rowan A., Halford S., Bicknell D. C., Bodmer W., Tomlinson I. P., Carter N. P. Array comparative genomic hybridization analysis of colorectal cancer cell lines and primary carcinomas // *Cancer Res*. 2004. Vol. 64 (14). P. 4817–25. DOI: 10.1158/0008-5472.CAN-04-0328. URL: <https://pubmed.ncbi.nlm.nih.gov/15256451> (дата обращения: 10.05.2026).

15. Tarnowski M., Czerewaty M., Deskur A., Safranow K., Marlicz W., Urasińska E., Ratajczak M. Z., Starzyńska T. Expression of Cancer Testis Antigens in Colorectal Cancer: New Prognostic and Therapeutic Implications // *Dis Markers*. 2016. Vol. 2016. P. 1987505. URL: <https://pubmed.ncbi.nlm.nih.gov/27635108> (дата обращения: 10.05.2026). DOI: 10.1155/2016/1987505.

16. Li J., Ji Y., Chen N., Wang H., Fang C., Yin X., Jiang Z., Dong Z., Zhu D., Fu J., Zhou W., Jiang R., He L., Hantao Z., Shi G., Cheng L., Su X., Dai L., Deng H. A specific upregulated long noncoding RNA in colorectal cancer promotes cancer progression // *JCI Insight*. 2022. Vol. 7 (15). P. e158855. DOI: 10.1172/jci.insight.158855.

17. Carvalho B., Postma C., Mongera S., Hopmans E., Diskin S., van de Wiel M. A., van Criekinge W., Thas O., Matthäi A., Cuesta M. A., Terhaar Sive Droste J. S., Craanen M., Schröck E., Ylstra B., Meijer G. A. Multiple putative oncogenes at the chromosome 20q amplicon contribute to colorectal adenoma to carcinoma progression // *Gut*. 2009. Vol. 58 (1). P. 79–89. URL: <https://pubmed.ncbi.nlm.nih.gov/18829976> (дата обращения: 10.05.2026). DOI: 10.1136/gut.2007.143065.

18. Ptashkin R. N., Pagan C., Yaeger R., Middha S., Shia J., O'Rourke K. P., Berger M. F., Wang L., Cimera R., Wang J., Klimstra D. S., Saltz L., Ladanyi M., Zehir A., Hechtman J. F. Chromosome 20q Amplification Defines a Subtype of Microsatellite Stable, Left-Sided Colon Cancers with Wild-type RAS/RAF and Better Overall Survival // *Mol Cancer Res.* 2017. Vol. 15 (6). P. 708–713. URL: <https://pmc.ncbi.nlm.nih.gov/articles/PMC5588907> (дата обращения: 10.05.2026). DOI: 10.1158/1541-7786.MCR-16-0352.
19. Iyer N. G., Ozdag H., Caldas C. p300/CBP and cancer // *Oncogene.* 2004. Vol. 23 (24). P. 4225–4231. URL: <https://pubmed.ncbi.nlm.nih.gov/15156177> (дата обращения: 10.05.2026). DOI: 10.1038/sj.onc.1207118.
20. Markowitz S. D., Roberts A. B. Tumor suppressor activity of the TGF-beta pathway in human cancers // *Cytokine Growth Factor Rev.* 1996. Vol. 7 (2). P. 93–102. URL: <https://pubmed.ncbi.nlm.nih.gov/8864357> (дата обращения: 10.05.2026). DOI: 10.1016/1359-6101(96)00001-9.
21. Kinzler K. W., Vogelstein B. Lessons from hereditary colorectal cancer // *Cell.* 1996. Vol. 87 (2). P. 159–170. DOI: 10.1016/s0092-8674(00)81333-1.
22. Levine A. J., Oren M. The first 30 years of p53: growing ever more complex // *Nat Rev Cancer.* 2009. Vol. 9 (10). P. 749–758. DOI: 10.1038/nrc2723.
23. Douillard J. Y., Oliner K. S., Siena S., Tabernero J., Burkes R., Barugel M., Humblet Y., Bodoky G., Cunningham D., Jassem J., Rivera F., Kocáková I., Ruff P., Błasińska-Morawiec M., Šmakal M., Canon J. L., Rother M., Williams R., Rong A., Wietzorek J., Sidhu R., Patterson S. D. Panitumumab-FOLFOX4 treatment and RAS mutations in colorectal cancer // *N. Engl J. Med.* 2013. Vol. 369 (11). P. 1023–1034. DOI: 10.1056/NEJMoa1305275.
24. Ogino S., Goel A. Molecular classification and correlates in colorectal cancer // *J. Mol Diagn.* 2008. Vol. 10 (1). P. 13–27. URL: <https://pubmed.ncbi.nlm.nih.gov/18165277> (дата обращения: 10.05.2026).
25. Dhawan P., Singh A. B., Deane N. G., No Y., Shiou S. R., Schmidt C., Neff J., Washington M. K., Beauchamp R. D. Claudin-1 regulates cellular transformation and metastatic behavior in colon cancer // *J. Clin Invest.* 2005. Vol. 115 (7). P. 1765–1776. DOI: 10.1172/JCI24543.
26. Wang D., Dubois R. N., Richmond A. The role of chemokines in intestinal inflammation and cancer // *Curr Opin Pharmacol.* 2009. Vol. 9 (6). P. 688–696. DOI: 10.1016/j.coph.2009.08.003.
27. Bender F. C., Raymond M. A., Bron C., Quest A. F. Caveolin-1 levels are down-regulated in human colon tumors, and ectopic expression of caveolin-1 in colon carcinoma cell lines reduces cell tumorigenicity // *Cancer Res.* 2000. Vol. 60 (20). P. 5870–5878. URL: <https://pubmed.ncbi.nlm.nih.gov/11059785> (дата обращения: 10.05.2026).

28. Almeida L. G., Sakabe N. J., de Oliveira A. R., Silva M. C., Mundstein A. S., Cohen T., Chen Y. T., Chua R., Gurung S., Gnjatic S., Jungbluth A. A., Caballero O. L., Bairoch A., Kiesler E., White S. L., Simpson A. J., Old L. J., Camargo A. A., Vasconcelos A. T. CTdatabase: a knowledge-base of high-throughput and curated data on cancer-testis antigens // *Nucleic Acids Res.* 2009. Vol. 37 (Database issue). P. D816-9. DOI: 10.1093/nar/gkn673.
29. Димитриади Т. А., Бурцев Д. В., Дженкова Е. А., Кутилин Д. С. МикроРНК как маркеры прогрессирования предраковых заболеваний в рак шейки матки // *Современные проблемы науки и образования.* 2020. №. 1. URL: <http://science-education.ru/ru/article/view?id=29529> (дата обращения: 10.05.2026).
30. Новикова И. А., Тимошкина Н. Н., Кутилин Д. С. Дифференциальная экспрессия микроРНК в опухолевых и нормальных тканях толстой кишки // *Якутский медицинский журнал.* 2020. Т. 4. С. 74–82. DOI: 10.25789/YMJ.2020.72.19.
31. Masuda T., Hayashi N., Kuroda Y., Ito S., Eguchi H., Mimori K. MicroRNAs as Biomarkers in Colorectal Cancer // *Cancers (Basel).* 2017. Vol. 9 (9). P. 124. DOI: 10.3390/cancers9090124.
32. Carter J. V., Galbraith N. J., Yang D., Burton J. F., Walker S. P., Galandiuk S. Blood-based microRNAs as biomarkers for the diagnosis of colorectal cancer: a systematic review and meta-analysis // *Br J. Cancer.* 2017. Vol. 116 (6). P. 762–774. DOI: 10.1038/bjc.2017.12.
33. Wu Y., Xu X. Long non-coding RNA signature in colorectal cancer: research progression and clinical application // *Cancer Cell Int.* 2023. Vol. 23 (1). P. 28. DOI: 10.1186/s12935-023-02867-0.
34. Giannakis M., Hodis E., Jasmine Mu X., Yamauchi M., Rosenbluh J., Cibulskis K., Saksena G., Lawrence M. S., Qian Z. R., Nishihara R., Van Allen E. M., Hahn W. C., Gabriel S. B., Lander E. S., Getz G., Ogino S., Fuchs C. S., Garraway L. A. RNF43 is frequently mutated in colorectal and endometrial cancers // *Nat Genet.* 2014. Vol. 46 (12). P. 1264–1266. DOI: 10.1038/ng.3127.
35. Li Q., Geng S., Luo H., Wang W., Mo Y. Q., Luo Q., Wang L., Song G. B., Sheng J. P., Xu B. Signaling pathways involved in colorectal cancer: pathogenesis and targeted therapy // *Signal Transduct Target Ther.* 2024. Vol. 9 (1). P. 266. DOI: 10.1038/s41392-024-01953-7.
36. Siddiqui A. D., Piperdi B. KRAS mutation in colon cancer: a marker of resistance to EGFR-I therapy // *Ann Surg Oncol.* 2010. Vol. 17 (4). P. 1168–1176. DOI: 10.1245/s10434-009-0811-z.
37. Wang Y., Rozen V., Zhao Y., Wang Z. Oncogenic activation of *PIK3CA* in cancers: Emerging targeted therapies in precision oncology // *Genes Dis.* 2024. Vol. 12 (2). P. 101430. URL: <https://pubmed.ncbi.nlm.nih.gov/39717717> (дата обращения: 10.05.2026). DOI: 10.1016/j.gendis.2024.101430.
38. Wu J. N., Roberts C. W. ARID1A mutations in cancer: another epigenetic tumor suppressor? // *Cancer Discov.* 2013. Vol. 3 (1). P. 35–43. DOI: 10.1158/2159-8290.CD-12-0361.

39. Huang K., Ding S., Chen K., Guo C., Wu M., Zhang R., Wang L., Fu Y., Xiao J., Liu X. *RNF43* mutation as a predictor of immunotherapeutic efficacy in colorectal cancer // *Am J. Cancer Res.* 2023. Vol. 13 (11). P. 5549–5558. URL: <https://pubmed.ncbi.nlm.nih.gov/38058823> (дата обращения: 10.05.2026).
40. Chen L., Chen G., Zheng X., Chen Y. Expression status of four mismatch repair proteins in patients with colorectal cancer: clinical significance in 1238 cases // *Int J. Clin Exp Pathol.* 2019. Vol. 12 (10). P. 3685–3699. URL: <https://pubmed.ncbi.nlm.nih.gov/31933757> (дата обращения: 10.05.2026).
41. Shao X., Lv N., Liao J., Long J., Xue R., Ai N., Xu D., Fan X. Copy number variation is highly correlated with differential gene expression: a pan-cancer study // *BMC Med Genet.* 2019. Vol. 20 (1). P. 175. DOI: 10.1186/s12881-019-0909-5.
42. Yang S., Tang D., Zhao Y. C., Liu H., Luo S., Stinchcombe T. E., Glass C., Su L., Shen S., Christiani D. C., Wang Q., Wei Q. Potentially functional variants of ERAP1, PSMF1 and NCF2 in the MHC-I-related pathway predict non-small cell lung cancer survival // *Cancer Immunol Immunother.* 2021. Vol. 70 (10). P. 2819–2833. DOI: 10.1007/s00262-021-02877-9.
43. Yadav L., Pietilä E., Öhman T., Liu X., Mahato A. K., Sidorova Y., Lehti K., Saarma M., Varjosalo M. PTPRA Phosphatase Regulates GDNF-Dependent RET Signaling and Inhibits the RET Mutant MEN2A Oncogenic Potential // *iScience.* 2020. Vol. 23 (2). P. 100871. URL: <https://pubmed.ncbi.nlm.nih.gov/32062451> (дата обращения: 10.05.2026). DOI: 10.1016/j.isci.2020.100871.
44. Chen Z., Zhong X., Xia M., Zhong J. The roles and mechanisms of the m6A reader protein YTHDF1 in tumor biology and human diseases // *Mol Ther Nucleic Acids.* 2021. Vol. 26. P. 1270–1279. DOI: 10.1016/j.omtn.2021.10.023.
45. Ng J. M., Yu J. Promoter hypermethylation of tumour suppressor genes as potential biomarkers in colorectal cancer // *Int J. Mol Sci.* 2015. Vol. 16 (2). P. 2472–2496. DOI: 10.3390/ijms16022472.
46. Corrales L., Hipp S., Martin K., Sabarth N., Tirapu I., Fuchs K., Thaler B., Walterskirchen C., Bauer K., Fabits M., Bergmann M., Binder C., Chetta P. M., Vogt A. B., Adam P. J. LY6G6D is a selectively expressed colorectal cancer antigen that can be used for targeting a therapeutic T-cell response by a T-cell engager // *Front Immunol.* 2022. Vol. 13. P. 1008764. DOI: 10.3389/fimmu.2022.1008764.
47. Kibschull M., Colaco K., Matysiak-Zablocki E., Winterhager E., Lye S. J. Connexin31.1 (Gjb5) deficiency blocks trophoblast stem cell differentiation and delays placental development // *Stem Cells Dev.* 2014. Vol. 23 (21). P. 2649–2660. DOI: 10.1089/scd.2014.0013.

48. Wu S., Chen S., Sun X., Chen X. Exploration of signature-related FAM genes and correlation between FAM50A expression and the pathogenesis and prognosis of hepatocellular carcinoma // Transl Cancer Res. 2025. Vol. 14 (8). P. 4720–4747. DOI: 10.21037/tcr-2025-171.

Конфликт интересов: Авторы заявляют об отсутствии конфликта интересов.

Conflict of interest: The authors declare that there is no conflict of interest.

Финансирование: Авторы заявляют об отсутствии внешнего финансирования.

Financing: The research was performed without external funding.