

(19)



**Евразийское  
патентное  
ведомство**

(11) **045105**

(13) **B1**

**(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОМУ ПАТЕНТУ**

**(45)** Дата публикации и выдачи патента  
**2023.10.30**

**(51)** Int. Cl. *C12Q 1/6886* (2018.01)

**(21)** Номер заявки  
**202293350**

**(22)** Дата подачи заявки  
**2022.11.15**

---

**(54) СПОСОБ РАННЕЙ ДИАГНОСТИКИ КОЛОРЕКТАЛЬНОГО РАКА В  
КАЗАХСТАНСКОЙ ПОПУЛЯЦИИ НА ОСНОВЕ ИСПОЛЬЗОВАНИЯ ПАНЕЛИ  
МУЛЬТИГЕННОГО ТЕСТИРОВАНИЯ**

---

**(31)** 2022/0693.1

**(32)** 2022.11.04

**(33)** KZ

**(43)** 2023.10.25

**(96)** KZ2022/063 (KZ) 2022.11.15

**(71)(73)** Заявитель и патентовладелец:  
**АКЦИОНЕРНОЕ ОБЩЕСТВО  
"КАЗАХСКИЙ НАУЧНО-  
ИССЛЕДОВАТЕЛЬСКИЙ  
ИНСТИТУТ ОНКОЛОГИИ И  
РАДИОЛОГИИ" (KZ)**

**(56)** ZHUNUSSOVA G., et al. Mutation Spectrum of Cancer-Associated Genes in Patient With Early Onset of Colorectal Cancer. FRONTIERS OF ONCOLOGY, 2019, Vol. 9, Art. 673, Разделы "RESULTS", "DISCUSSION", Табл. 2, 3, фиг. 4, 5

PARSA F.G., et al. Fanconi Anemia Pathway in Colorectal Cancer: A Novel Opportunity for Diagnosis, Prognosis and Therapy. J. PERS MED., 2022, Vol. 12(3), 396. Раздел "3.1 Germline Monoallelic Mutations", последний абзац, табл. 1  
WO-2021016089

**(72)** Изобретатель:  
**Афонин Георгий Алексеевич,  
Кайдарова Диляра Радиковна,  
Балтаев Нурлан Анаркулович (KZ)**

---

**(57)** Изобретение относится к медицине, а именно к онкологии, и может быть использовано для улучшения ранней диагностики колоректального рака (КРР) в казахстанской популяции. Важным этапом диагностики является молекулярно-генетическое тестирование на "панелях генов", представляющих собой диагностические системы для анализа мутаций генов, ассоциированных с различными вариантами КРР. Обнаруженные в исследовании мутации генов можно классифицировать как причинные изменения, приводящие к более раннему, по сравнению с общепопуляционным, развитию КРР и использовать как дополнительный маркер в ранней диагностике КРР. По результатам исследования были идентифицированы патогенные мутации, не имеющиеся ни в одной популяционной базе данных и традиционно не связанные с КРР - BRCA2, FANCI, FANCC, DICER1, которые рекомендуются для добавления в панели мультигенного тестирования, ассоциированные с КРР. Анализ мутаций генов пациентов молодого возраста на основе секвенирования с дополненной "панелью генов" позволит определить уровень риска для конкретного пациента и его кровных родственников, модифицировать диагностические и терапевтические подходы при КРР.

---

**B1**

**045105**

**045105**

**B1**

Изобретение относится к медицине, а именно к онкологии, и может быть использовано для улучшения ранней диагностики колоректального рака в казахстанской популяции.

В Республике Казахстан отмечается ежегодный рост заболеваемости колоректальным раком (КРР), который находится на 3 месте по частоте встречаемости в общей структуре онкологической патологии. Также в Республике Казахстан наблюдается ежегодный рост числа случаев КРР у лиц в молодом возрасте. Анализ генов пациентов молодого возраста на основе секвенирования нового поколения (NGS) позволяет определить уровень риска для конкретного пациента и его кровных родственников, модифицировать диагностические и терапевтические подходы. Важным этапом диагностики лиц с наследственно отягощенным анамнезом является молекулярно-генетическое тестирование, основанное на выборе так называемых "панелей генов", представляющих собой диагностические панели для сайт-специфического анализа генов, ассоциированных с полипозами и различными вариантами КРР.

К настоящему времени результаты по разработке диагностических панелей для обнаружения КРР в Республике Казахстан представлены в единичных публикациях. Увеличение заболеваемости КРР у лиц молодого возраста, отсутствие ранней диагностики при семейных и наследственных формах данного вида рака, высокие показатели запущенности заболевания и низкие показатели выживаемости при КРР (общая 5-летняя выживаемость при III и IV стадиях составляет 59,4 и 39,5% соответственно), обуславливают необходимость совершенствования ранней диагностики с помощью внедрения генетического скрининга.

Известен способ молекулярно-генетической диагностики - "Панель ColoNext для выявления пациентов с наследственным риском колоректального рака и/или полипов", разработанный компанией Ambry Genetics (США) ([www.ambrygen.com/providers/genetic-testing/6/oncology/colonext/](http://www.ambrygen.com/providers/genetic-testing/6/oncology/colonext/)). Панель включает 20 генов, ассоциированных с КРР: APC, AXIN2, BMPR1A, CDH1, CHEK2, EPCAM, GREM1, MLH1, MSH2, MSH3, MSH6, MUTYH, NTHL1, PMS2, POLD1, POLE, PTEN, SMAD4, STK11, TP53. Эти гены оцениваются с помощью секвенирования нового поколения (NGS) или секвенирования по Сэнгеру всех кодирующих доменов, а также фланкирующих 5'- и 3'-концов всех интронов и нетранслируемых областей.

Однако недостатком данного способа является то, что набор генов (панель) сформирован для этногенетически отличных от казахстанского населения когорт пациентов из США, Европы, Юго-Восточной Азии и Китая. Известно, что некоторые патогенные варианты мутаций проявляют этническую и расовую специфичность в модификации риска КРР. Кроме того, указанная панель включает не все локусы, значимые для молекулярно-генетической диагностики КРР.

Известен также способ диагностики колоректального рака, использующий панель мультигенного тестирования для пациентов с высоким риском семейного и наследственного колоректального рака, рекомендованную NCCN (The National Comprehensive Cancer Network) - Национальной комплексной онкологической сетью (NCCN Clinical Practice Guidelines in Oncology. Genetic/Familial High Risk Assessment: Colorectal. Version 3.2019. - P. 8-87. [www.henryford.com/-/media/files/henry-ford/hcp/pathology/nccn-genetics-colon.pdf](http://www.henryford.com/-/media/files/henry-ford/hcp/pathology/nccn-genetics-colon.pdf)).

Панель "Наследственный и семейный рак толстой кишки" направлена на обнаружение патогенных мутаций в генах, ассоциированных с наследственными формами колоректального рака и наследственными опухолевыми синдромами, связанными с повышенным риском развития КРР. Данная панель включает 22 гена: APC, ATM, AXIN2, BLM, BMPR1A, CHEK2, EPCAM, GALNT12, GREM1, MLH1, MSH2, MSH6, MUTYH biallelic pathogenic variants, MUTYH heterozygotes, NTHL1, PMS2, POLD1, POLE, PTEN, SMAD4, STK11, TP53.

Недостатком данного способа является то, что в панель не включены локусы генов, мутации в которых также ассоциированы с колоректальным раком населения Республики Казахстан.

Прототипом данного способа является метод молекулярно-генетической диагностики колоректального рака на основе NGS. Были получены данные по частотному распределению аллельных вариантов и мутаций ключевых генов колоректального канцерогенеза (APC, TP53, DCC, MLH1, MSH2, K-ras, GSTT1, GSTM1) в казахстанских популяциях. Впервые для казахстанской популяции определен спектр мутаций критических районов генов (APC, TP53, MLH1, MSH2). Определены индивидуальные генотипы, достоверно ассоциированные с высоким риском развития спорадических вариантов колоректального рака с учетом этнической принадлежности (Жунусова Г.С. "Разработка панелей генетических маркеров для скрининга семейных и спорадических случаев колоректального рака в казахстанских популяциях", дисс. на соискание степени доктора философии, (2014)).

Недостатком данного способа является отсутствие исследований других генов, также ассоциированных с КРР.

Поэтому задачей исследования является разработка способа ранней диагностики КРР с использованием панели мультигенного тестирования для профилактики колоректального рака в Республике Казахстан.

Техническая задача способа осуществляется путем добавления в панель мультигенного тестирования, ассоциированную с КРР, новых локусов генов: BRCA2, FANCI, FANCC, DICER1.

Способ осуществляется следующим образом.

Материалом для молекулярно-генетического исследования являются ДНК, извлеченные из лимфоцитов периферической крови пациентов в возрасте от 17 до 50 лет.

Производят забор 5 мл крови из периферической вены пациента, или из внутривенного катетера, или из внутривенного порта длительного доступа в стерильных условиях с использованием вакуумной системы VenoSafe с этилендиаминтетрауксусной кислотой (ЭДТА) в качестве антикоагулянта.

Геномная ДНК экстрагируется из лимфоцитов периферической крови согласно протоколу выделения ДНК стандартным методом с применением специальных наборов реактивов: GeneJet Genomic DNA Purification Kit (ThermoScientific, США). Качественная и количественная оценка образцов ДНК проводится с использованием методов спектрофотометрии на оборудовании Eppendorf BioPhotometerplus (Eppendorf, Германия) или флуорометрии на оборудовании Quantus™ Fluorometer (Promega, США) с применением набора реактивов QuantiFluor® dsDNA System (Promega, США). Чистота образца (наличие РНК и белка) считается удовлетворительной при абсорбции света молекулами ДНК (при длине волны 260-280 нм). Согласно протоколу исследования, количество выделенной двухцепочечной ядерной ДНК составляет не менее 50 нг. Выделенные образцы ДНК хранятся при температуре -20°C до включения в эксперимент.

Массивное параллельное секвенирование осуществляется на генетическом анализаторе нового поколения MiSeq (Illumina, США). Используются наборы реактивов TruSight Rapid Capture Kit (Illumina) и TruSight Cancer Sequencing Panel (Illumina) - системы для анализа кодирующих последовательностей и фланкирующих некодирующих регионов генов, более чем 1700 экзонов 94 генов (255 килобаз ядерного генома) и 284 замен единичных пар оснований, ассоциированных с различными видами рака ([http://products.illumina.com/products/trusight\\_cancer.html](http://products.illumina.com/products/trusight_cancer.html)).

Для подготовки библиотек 50 нг двухцепочечной ДНК каждого образца энзиматически фрагментируются с использованием Nextera транспосом, на концах фрагментов прикрепляются адаптерные последовательности (tag). Проводится очистка tag-ментированных фрагментов ДНК от Nextera транспосом, добавляются общие адаптеры, требуемые для генерации кластеров и амплификации, tag-ментированные ДНК амплифицируются ПЦР с последующей очисткой амплифицированных фрагментов. Очистка библиотек ДНК проводится для удаления ненужных продуктов амплификации с использованием магнитных частиц. Полные библиотеки количественно и качественно оцениваются с помощью метода флуорометрии на приборе Quantus™ Fluorometer (Promega, США) и с применением набора реактивов Agilent DNA High Sensitivity Chip на биоанализаторе Agilent Technologies 2100 Bioanalyzer (AgilentTechnologies, США). Затем по 500 нг каждой ДНК-библиотеки объединяются в один пул, с последующей двойной гибридизацией для захвата меченых зондов, специфичных для целевых областей генов. Негибридизованный материал удаляется с помощью промывания (Enrichment Wash Solution). Перед ПЦР-амплификацией захваченную библиотеку очищают с использованием магнитных частиц для удаления негибридизованного материала, а затем обогащенную библиотеку амплифицируют и снова очищают с использованием магнитных частиц. Затем всю библиотеку количественно оценивают на приборе Quantus™ Fluorometer, для оценки качества библиотеки используют набор реагентов Agilent High Sensitivity и биоанализатор AgilentTechnologies 2100. Библиотеки ДНК с молярностью 4 мМ подвергаются кластерной генерации в проточных ячейках. Проводится 300 циклов секвенирования в соответствии с парными концами фрагментов в разных направлениях с использованием картриджа MiSeq (MiSeq Reagent Kit v.2 и MiSeq Reagent Kit v.3) на платформе MiSeq или другого секвинатора. Первичные данные секвенирования анализируются с использованием программного обеспечения (ПО) MiSeq Reporter v.2.4, в результате обработки создаются файлы в форматах FASTQ, BAM и VCF.

Для оценки качества секвенирования используют показатель Phred (Q score), который оценивает вероятность ошибочного распознавания нуклеотида в процессе секвенирования. В качестве допустимого порогового значения качества секвенирования (Q score) выбирают показатель 30, что соответствует уровню частоты ошибок 1:1000. Выровненные и упорядоченные секвенированные последовательности нуклеотидов сопоставляются с референсной последовательностью генома человека (GRCH37.p5/hg19) с использованием геномной базы данных (БД) Burrows-WheelerAligner. Далее проводится поиск вариантов специфических районов генома, для этого используют ПО Genome Analysis Toolkit (GATK, Broad Institute, Cambridge, США). Дальнейший анализ файлов в формате FASTQ проводится с помощью программы NextGENe v.2.3.4.3. (Softgenetics, State College, США), которая использует алгоритм поиска специфических последовательностей в геномной БД Burrows-Wheeler Transform. Для работы в программе NextGENe первичные MiSeq FASTQ-файлы конвертируются в FASTA-файлы, которые выравниваются с референсной последовательностью генома человека GRCH37.p5/hg19 с использованием метода двойных чтений.

Биоинформационный анализ проводится с использованием двух методических подходов. При первом подходе данные секвенирования анализируются с использованием ПО MiSeq Reporter v.3.0. Данная программа позволяет проводить выравнивание и упорядочивание секвенированных последовательностей, которые сопоставляются с референсной последовательностью генома с использованием алгоритма Burrows-Wheeler Aligner. Поиск и детекция вариантов для специфических районов генома проводятся с

использованием GATK. Во втором подходе процесс оптимизирован с помощью пакетов ПО. Картирование и выравнивание чтений относительно референсных последовательностей генома (версия GRCH37.p5/hg19) проводятся с помощью алгоритма Bowtie 2.

Оценка качества чтения последовательностей проводится с использованием ПО FastQC и MultiQC. Конвертация SAM-файлов в BAM-файлы, сортировка картированных прочтений и их индексирование, удаление промежуточных файлов, объединение BAM-файлов, а также идентификация повторов в них осуществляются с использованием пакетов ПО Picard Tools и SAM Tools/BCF Tools. Для запуска программ устанавливаются и используются ПО Java Runtime Environment и R Bioconductor. Фильтрация и детекция геномных вариантов проводятся с помощью стратегии GATK Haplotype caller.

Все этапы биоинформационного анализа и вычислительные операции выполняются с использованием рабочей станции HPZ440 (Intel® Xeon® CPU E5-1603 v3@ 2.80GHz 2.80Gb, 64Gb RAM) и платформы MiSeq Illumina (Intel® Cores™ i7-2710QE CPU@ 2.10GHz, 16Gb RAM) или другого сервера и компьютера, удовлетворяющего указанным требованиям. Результаты секвенирования, представленные выровненными последовательностями в виде VCF-файлов, анализируются путем ПО Variant Studio Data Analysis v.2.2. (Illumina). Для анализа используются специфические районы генов с покрытием более 10 нуклеотидов. Определенные однонуклеотидные варианты и небольшие инсерции/делеции аннотируются и фильтруются для экзонных несинонимичных вариантов или вариантов с альтернативным сплайсингом и сдвигом рамки считывания.

Генетические варианты аннотируются в соответствии с номенклатурой Международной базы вариантов секвенирования Human Genome Variation Society (HGVS, [www.vamomen.hgvs.org](http://www.vamomen.hgvs.org)). Для интерпретации используются геномные БД: база данных однонуклеотидных замен (Single Nucleotide Polymorphism Database, <http://www.ncbi.nlm.nih.gov/projects/SNP/>) и база клинических генетических вариантов (ClinVar, <http://www.ncbi.nlm.nih.gov/clinvar/>) Национального центра биотехнологической информации США, Лейденская открытая база геномных вариантов (Leiden Open Variation Database, LOVD, <http://www.lovd.nl/3.0/home>), база данных Международного общества гастроинтестинальных наследственных опухолей (International Society for Gastrointestinal Hereditary Tumours, InSiGHT, <https://www.insight-group.org/>) и каталог соматических мутаций при раке (Catalogue of Somatic Mutations in Cancer, COSMIC, <http://cancer.sanger.ac.uk/cancergenome/projects/cosmic/>). Для визуализации вариантов было применено ПО Integrative Genomics Viewer.

Идентифицированные варианты классифицируются в соответствии с рекомендациями Американской коллегии медицинской генетики и геномики (American College of Medical Genetics and Genomics) на патогенные, вероятно патогенные, варианты с неопределенным функциональным значением, вероятно доброкачественные и доброкачественные ([https://www.acmg.net/ACMG/Medical-Genetics-Practice-Resources/ACT\\_Sheets\\_and\\_Algorithms/](https://www.acmg.net/ACMG/Medical-Genetics-Practice-Resources/ACT_Sheets_and_Algorithms/)).

Пример исполнения.

В исследование включено 185 пациентов с КРР в возрасте 17-50 лет (экспериментальная группа для молекулярно-генетического анализа), прошедших обследование и лечение в Казахском научно-исследовательском институте онкологии и радиологии, региональных онкологических центрах и онкологических центрах городов Алматы, Астана, Шымкент. Исследование проводилось в рамках НТП BR11065390 "Разработка и развитие инновационных технологий ранней диагностики и лечения злокачественных заболеваний с учетом современных подходов геномики" (ПЦФ МЗ РК), а также результаты исследования вошли в диссертационное исследование на соискание степени доктора философии Афонина Г.А. на тему "Молекулярно-генетический анализ развития колоректального рака у больных в возрасте до 50 лет" по специальности 6D110100 -Медицина.

Количество мужчин составило 98 (53%), женщин - 87 (47%). Средний возраст пациентов составил  $41,1 \pm 0,52$  лет. Наибольшее количество пациентов относилось к возрасту 47 лет (9,2%), 49 лет (7,5%), 43 и 45 лет (по 7%), 40 лет (5,4%) и 36 лет (4,3%).

Разработка способа ранней диагностики КРР с использованием панели мультигенного тестирования проводилась методом, указанным выше (в разделе "Способ осуществляется следующим образом").

В ходе проведенного анализа методом секвенирования были идентифицированы патогенные мутации при разных фенотипах КРР в казахстанской популяции, не имеющиеся ни в одной популяционной базе данных и традиционно не связанные с КРР - в генах BRCA2, FANCI, FANCC, DICER1.

Данные мутации можно классифицировать как причинные изменения, приводящие к более раннему, по сравнению с общепопуляционным, развитию колоректального рака, и их наличие можно использовать как дополнительный маркер в ранней диагностике КРР.

Молекулярно-генетический анализ на основе секвенирования нового поколения позволяет идентифицировать случаи КРР синдромальной и спорадической природы. Это дает возможность определять уровень риска для пациента и его родственников, осуществлять активное раннее выявление предопухолевых состояний и рака и проводить своевременное лечение.

ФОРМУЛА ИЗОБРЕТЕНИЯ

Способ ранней диагностики колоректального рака в казахстанской популяции с использованием панели мультигенного тестирования, включающей гены, достоверно ассоциированные с высоким риском развития спорадических и наследственных вариантов колоректального рака, отличающийся тем, что при секвенировании добавляются отсутствующие в панелях мультигенного тестирования КРР гены BRCA2, FANCI, FANCC, DICER1, частота мутаций которых у пациентов казахстанской популяции достоверно высока и является дополнительным маркером для диагностики колоректального рака.

