

РАК МОЛОЧНОЙ ЖЕЛЕЗЫ: АНАЛИЗ СПЕКТРА СОМАТИЧЕСКИХ ДРАЙВЕРНЫХ МУТАЦИЙ С ПРИМЕНЕНИЕМ ВЫСОКОПРОИЗВОДИТЕЛЬНОГО СЕКВЕНИРОВАНИЯ

К. Ю. Цуканов¹, А. Ю. Красненко¹, Д. О. Коростин¹, А. В. Чуров², И. Ф. Стеценко¹, Н. А. Плотноков¹, С. А. Зарубина³, В. А. Белова³, А. В. Ковыршина³, И. К. Воротников⁴, А. А. Мещеряков⁴, В. В. Ильинский^{5,3,1} ✉

¹ ООО «Генотек», Москва

² Карельский научный центр Российской академии наук, Петрозаводск

³ Институт общей генетики имени Н. И. Вавилова РАН, Москва

⁴ Национальный медицинский исследовательский центр онкологии имени Н. Н. Блохина, Москва

⁵ Научно-исследовательский институт биомедицинской химии имени В. Н. Ореховича, Москва

Рак молочной железы (РМЖ) представляет собой одну из наиболее распространенных форм злокачественных опухолей. Развитие новых подходов к скринингу генетических изменений у больных с опухолями молочной железы поможет значительно снизить общую высокую смертность от рака этого типа и повысить эффективность противоопухолевой терапии. Целью настоящей работы являлось выявление методом высокопроизводительного секвенирования спектра мутаций в составе ключевых онкогенов при РМЖ оценка их патогенности с применением ранее разработанного биоинформатического алгоритма, а также оценка взаимосвязи некоторых мутаций с особенностями клинического проявления заболевания. В исследовании приняла участие 16 пациенток с РМЖ (средний возраст — 50,7 ± 11,3 года). Было обнаружено 58 мутаций в онкогенах *BRCA1*, *BRCA2*, *ATM*, *CDH1*, *CHEK2* и *TP53*. Среди выявленных генетических вариантов с применением биоинформатических подходов найдено 14 мутаций, оказывающих влияние на последовательность кодируемого белка. Большая часть патогенных мутаций идентифицирована в генах *BRCA1/2*, *ATM* и *TP53*.

Ключевые слова: рак молочной железы, высокопроизводительное секвенирование, соматическая мутация, онкогены, *BRCA1*, *BRCA2*, *TP53*

Финансирование: работа выполнена при поддержке Министерства образования и науки РФ (идентификатор соглашения RFMEFI60716X0152).

✉ **Для корреспонденции:** Ильинский Валерий Владимирович
Наставнический пер., д. 17, стр. 1, под. 14, 15, г. Москва, 105120; info@genotek.ru

Статья получена: 14.12.2017 **Статья принята к печати:** 22.12.2017

BREAST CANCER: ANALYSIS OF DRIVER SOMATIC MUTATIONS DETECTED BY NEXT-GENERATION SEQUENCING

Tsukanov KYu¹, Krasnenko AYU¹, Korostin DO¹, Churov AV², Stetsenko IF¹, Plotnikov NA¹, Zarubina SA³, Belova VA³, Kovyrshina AV³, Vоротnikov IK⁴, Mescheryakov AA⁴, Ilyinsky VV^{5,3,1} ✉

¹ Genotek Ltd., Moscow, Russia

² Karelian Research Centre of the Russian Academy of Sciences, Petrozavodsk, Russia

³ Vavilov Institute of General Genetics of the Russian Academy of Sciences, Moscow, Russia

⁴ N. N. Blokhin Russian Cancer Research Center, Moscow, Russia

⁵ Institute of Biomedical Chemistry, Moscow, Russia

Breast cancer (BC) is one of the most common malignancies. There is a need for novel approaches to screening for genetic mutations in patients with BC that will help to reduce high mortality rates caused by this disease and improve treatment outcomes. In this study we employed next generation sequencing to screen a few key genes associated with the risk of breast cancer for mutations. We also evaluated their pathogenicity using the previously proposed bioinformatics-based algorithm and analyzed the associations between some of the detected mutations and the clinical manifestations of the disease. Our study recruited 16 female patients with BC (mean age was 50.7 ± 11.3 years). A total of 58 mutations were detected in the oncogenes *BRCA1*, *BRCA2*, *ATM*, *CDH1*, *CHEK2* and *TP53*. Bioinformatic analysis of the sequencing data revealed 14 mutations that affect the sequence of the encoded proteins. Most deleterious mutations were harbored by the genes *BRCA1/2*, *ATM* and *TP53*.

Keywords: breast cancer, next-generation sequencing, somatic mutation, oncogenes, *BRCA1*, *BRCA2*, *TP53*

Funding: this work was supported by the Ministry of Education and Science of the Russian Federation (Project ID RFMEFI60716X0152).

✉ **Correspondence should be addressed:** Valery Ilyinsky
Nastavnicheskyy per., d.17, str. 1, pod. 14, 15, Moscow, 105120; info@genotek.ru

Received: 14.12.2017 **Accepted:** 22.12.2017

Рак молочной железы (РМЖ) является вторым по распространенности типом рака в мире, занимает первое место в структуре онкологической заболеваемости и второе — в структуре смертности среди женщин [1]. Заболеваемость РМЖ увеличивается с возрастом и особенно велика среди женщин 60–65 лет, при этом патология зачастую диагностируется на поздних стадиях, что и определяет высокий уровень смертности от нее. Наибольшую опасность представляет собой метастазирующий РМЖ, который крайне трудно поддается лечению, несмотря на применение комбинированных схем, включающих химио- и гормонотерапию, а также использование таргетных противоопухолевых препаратов. Общий уровень 5-летней выживаемости пациентов при РМЖ составляет 55 %. В этой связи представляется актуальной разработка подходов для более эффективного скрининга РМЖ и выбора средств для таргетной терапии, учитывающих молекулярно-генетические особенности опухолей.

Стремительное развитие технологий высокопроизводительного секвенирования (next-generation sequencing, NGS) привело к получению большого количества данных о генетических вариантах [2]. Идентифицировано множество мутаций, связанных с развитием РМЖ, в том числе соматические и герминальные мутации в генах *PIK3CA*, *STK11/LKB1*, *CDH1*, *ATM*, *CHEK2*, *BRIP1*, *PALB2*, а также мутации в высокопенетрантных генах *TP53*, *PTEN*, *MLH1*, *BRCA1*, *BRCA2*, ассоциированных с наследственными формами РМЖ [3].

Большинство мутаций, возникающих в опухолях, являются соматическими и играют важную роль в патогенезе онкологических заболеваний и развитии *de novo* резистентности к лекарственным препаратам. Поэтому многие исследования направлены на профилирование вариантов в образцах опухолей с применением NGS. В результате было идентифицировано значительное количество вариантов с неизвестной функцией. Для их описания необходимы математические алгоритмы, позволяющие обрабатывать большие массивы данных в автоматическом режиме, предсказывать потенциально патогенные мутации и отличать их от нейтральных изменений генома опухолевых клеток. Результаты подобных исследований могут быть использованы для разработки скрининговых и диагностических инструментов (включая технологии жидкостной биопсии), а также инструментов подбора таргетных противоопухолевых препаратов.

В настоящей работе представлены результаты анализа спектра мутаций в составе ключевых онкогенов при РМЖ с применением метода высокопроизводительного секвенирования и ранее разработанного биоинформатического алгоритма для функциональной аннотации мутаций и оценки их патогенности.

ПАЦИЕНТЫ И МЕТОДЫ

Для проведения исследования были получены образцы опухолей от 16 пациенток со злокачественными новообразованиями молочной железы в возрасте от 27 до 76 лет (средний возраст — $50,7 \pm 11,3$ года), проходивших комплексное обследование и получавших комбинированное лечение в Национальном медицинском исследовательском центре онкологии имени Н. Н. Блохина (Москва). Критериями включения в исследование были: возраст 18–70 лет, женский пол, наличие гистологически и цитологически подтвержденного диагноза «рак молочной же-

лезы». Критериями исключения являлись наличие других форм новообразований в анамнезе и беременность.

Стадию опухолевого процесса устанавливали согласно классификации TNM [4]. В исследование были включены пациентки со стадиями T1–3N0–3M0–1.

Работа была выполнена с соблюдением принципов добровольности и конфиденциальности. От всех пациентов были получены письменные информированные согласия на проведение исследования. Клиническая характеристика больных представлена в табл. 1.

Выделение ДНК и контроль качества

Выделение ДНК из образцов опухолевой ткани проводили с применением набора DNeasy Blood and Tissue Kit (Qiagen, США). Измельченные образцы опухолевой ткани после добавления буфера ATL обрабатывали протеиназой K и инкубировали при 56 °C до полного лизиса. Затем обрабатывали ферментом PNHазой A. Далее последовательно добавляли 200 мкл буфера AL и 96 % этанола. Полученную смесь переносили на спин-колонки и центрифугировали при 8 000 g в течение 1 мин. Затем образцы промывали с применением буферов AW1 и AW2 для удаления солей (гуанидина и SDS). Для элюции ДНК колонки обрабатывали буфером 1x Low-TE (дважды, по 30 мкл), затем инкубировали и центрифугировали согласно протоколу производителя. Контроль качества полученной ДНК проводили на флуориметре Qubit 3.0 (Thermo Fisher Scientific, США), а также с помощью электрофореза в 1 % агарозном геле с добавлением бромистого этидия.

Секвенирование таргетной панели онкогенов

Из образцов ДНК опухолей молочной железы готовили библиотеки с применением набора реагентов NEBNext Ultra DNA Library Prep Kit for Illumina (New England Biolabs, США). Двойное баркодирование библиотек проводили с помощью ПЦП с применением наборов реагентов NEBNext Ultra DNA Library Prep Kit for Illumina и NEBNext Multiplex Oligos for Illumina (Dual Index Primers Set 1, New England Biolabs). Контроль качества полученных библиотек фрагментов ДНК проводили на приборе Agilent Bioanalyzer 2100 (Agilent Technologies, США) с помощью набора реагентов этого же производителя High Sensitivity Kit в соответствии с протоколом производителя.

Для таргетного обогащения кодирующих регионов геномов опухолей использовали набор MYbaits Onconome KL v1.5 Panel (MYcroarray, США). Анализ проводили с применением системы высокопроизводительного секвенирования HiSeq 2500 (Illumina, США) методом парных прочтений длиной 100 нуклеотидов. Подготовку образцов и запуск секвенирования осуществляли согласно стандартным протоколам Illumina.

Биоинформатическая обработка данных

Для биоинформатической обработки полученных при секвенировании NGS-данных применяли ранее разработанный алгоритм [5]. На начальном этапе проводили оценку качества прочтений, полученных при секвенировании ДНК опухолей: последовательности, имевшие качество прочтения ниже 10, удаляли с помощью программного обеспечения Cutadapt [6]. После этого прочтения картировали на

Таблица 1. Клиническая характеристика больных раком молочной железы (n = 16)

Характеристика	Значение, абс. (%)
Возраст, лет	50,7 ± 11,3
Операция:	
проводилась	16 (100)
не проводилась	0 (0)
T-стадия опухоли (согласно TNM):	
T1	10 (62,5)
T2	5 (31,3)
T3	1 (6,2)
Наличие метастазов в лимфоузлах:	
без метастазов, M0	10 (62,5)
с метастазами, M1	6 (37,5)
Экспрессия рецепторов эстрогена (ER):	
ER+	11 (68,8)
ER-	5 (31,2)
Экспрессия рецепторов прогестерона (PR):	
PR+	10 (62,5)
PR-	6 (37,5)
Экспрессия HER2/neu:	
Her2+	9 (56,3)
Her2-	7 (43,7)
Экспрессия Ki-67:	
низкая (< 14 %)	13 (81,3)
высокая (≥ 14 %)	3 (18,7)
Адьювантная химиотерапия:	
проводилась	10 (62,5)
не проводилась	6 (37,5)
Адьювантная гормональная терапия:	
проводилась	9 (56,3)
не проводилась	7 (43,7)
Лучевая терапия:	
проводилась	0 (0)
не проводилась	16 (100)

референсный геном hg19 (GRCh37.p13) с помощью алгоритма BWA (Burrows–Wheeler Aligner) [7]. ПЦП-дубликаты удаляли специализированной командой `mdup` в составе программной платформы SAMtools [8].

Для поиска мутаций применяли инструмент MuTect [9]. В качестве значимых вариантов, рассматривали последовательности ДНК, число покрытий которых в результате секвенирования составило не менее 12.

Для оценки функционального эффекта обнаруженных мутаций проводили их аннотирование и предсказание их влияния на кодируемый белок на основе анализа геномных координат фрагментов с помощью программы SnpEff [10].

РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ

С применением метода высокопроизводительного секвенирования на платформе Illumina мы провели скрининг образцов 16 опухолей молочной железы на наличие мутаций. Среди онкогенов в составе таргетной панели были исследованы *BRCA1*, *BRCA2*, *ATM*, *CDH1*, *CHEK2*, *MRE11A*, *NBN*, *PALB2*, *PTEN*, *RAD50*, *RAD51C*, *TP53*, *SEC23B*. В

результате применения биоинформатического алгоритма для анализа данных были обнаружены 58 точечных мутаций в генах *BRCA1*, *BRCA2*, *ATM*, *CDH1*, *CHEK2* и *TP53*, среди них — 19 гомозигот и 39 гетерозигот. Список уникальных мутаций представлен в табл. 2.

На рисунке представлено распределение частот встречаемости мутаций в генах, в которых их было обнаружено больше всего: *ATM*, *TP53* и *BRCA1*. Наиболее часто встречающиеся мутации — с.376-283T>C (*TP53*), с.3994-193T>C и с.8010+186C>T (*ATM*), а также с.5215+66G>A (*BRCA1*).

Среди обнаруженных нами генетических вариантов по результатам биоинформатической обработки данных и аннотирования мутаций с применением баз данных были отобраны мутации с потенциально значимым влиянием на регуляторную или белковую последовательность. Для оценки патогенности и консервативности выявляемых генетических вариантов использовали данные, которые извлекали из таких источников, как COSMIC (Catalogue of Somatic Mutations In Cancer) [11] и dbNSFP [12]. Дополнительно для предсказания патогенности обнаруженных вариантов, оценки их эффекта на функцию кодируемого белка применяли утилиты SIFT (Sorting Intolerant From

Таблица 2. Однонуклеотидные варианты (мутации), идентифицированные у пациенток с раком молочной железы (n = 16)

Образец	Представленность мутаций в образце, %	Мутация	Ген	Геномная локализация
1	18,7	c.4828G>A	<i>BRCA2</i>	Chr13:32913320
		c.3994-193T>C	<i>ATM</i>	Chr11:108158134
		c.-73G>A	<i>CHEK2</i>	Chr22:29137870
2	25,5	c.5215+66G>A	<i>BRCA1</i>	Chr17:41215825
		c.3994-193T>C	<i>ATM</i>	Chr11:108158134
		c.8010+186C>T	<i>ATM</i>	Chr11:108204881
		c.524G>A	<i>TP53</i>	Chr17:7578406
		c.376-283T>C	<i>TP53</i>	Chr17:7578837
3	29,4	c.8755-272A>G	<i>BRCA2</i>	Chr13:32953182
		c.376-283T>C	<i>TP53</i>	Chr17:7578837
4	26,5	c.3994-193T>C	<i>ATM</i>	Chr11:108158134
		c.8010+186C>T	<i>ATM</i>	Chr11:108204881
		c.376-283T>C	<i>TP53</i>	Chr17:7578837
5	26,5	c.5215+66G>A	<i>BRCA1</i>	Chr17:41215825
		c.3994-193T>C	<i>ATM</i>	Chr11:108158134
		c.8010+186C>T	<i>ATM</i>	Chr11:108204881
		c.1289C>T	<i>CHEK2</i>	Chr22:29091797
		c.376-283T>C	<i>TP53</i>	Chr17:7578837
6	20,6	c.5215+66G>A	<i>BRCA1</i>	Chr17:41215825
		c.3994-193T>C	<i>ATM</i>	Chr11:108158134
		c.8010+186C>T	<i>ATM</i>	Chr11:108204881
7	25,5	c.146C>G	<i>ATM</i>	Chr11:108098576
		c.376-283T>C	<i>TP53</i>	Chr17:7578837
8	25,5	c.3994-193T>C	<i>ATM</i>	Chr11:108158134
		c.790C>T	<i>CDH1</i>	Chr16:68844202
		c.376-283T>C	<i>TP53</i>	Chr17:7578837
9	28,5	c.5070A>C	<i>BRCA2</i>	Chr13:32913562
		c.-73G>A	<i>CHEK2</i>	Chr22:29137870
		c.469G>T	<i>TP53</i>	Chr17:7578461
10	37,3	c.5215+66G>A	<i>BRCA1</i>	Chr17:41215825
		c.4258C>T	<i>ATM</i>	Chr11:108160350
		c.376-283T>C	<i>TP53</i>	Chr17:7578837
11	26,5	c.5215+66G>A	<i>BRCA1</i>	Chr17:41215825
		c.1865C>T	<i>BRCA1</i>	Chr17:41245683
		c.384G>A	<i>BRCA1</i>	Chr17:41256196
		c.54G>T	<i>BRCA1</i>	Chr17:41276060
		c.3994-193T>C	<i>ATM</i>	Chr11:108158134
		c.-73G>A	<i>CHEK2</i>	Chr22:29137870
		c.743G>A	<i>TP53</i>	Chr17:7577538
		c.376-283T>C	<i>TP53</i>	Chr17:7578837
12	28,5	c.5215+66G>A	<i>BRCA1</i>	Chr17:41215825
		c.3994-193T>C	<i>ATM</i>	Chr11:108158134
		c.8010+186C>T	<i>ATM</i>	Chr11:108204881
		c.376-283T>C	<i>TP53</i>	Chr17:7578837
13	22,6	c.5215+66G>A	<i>BRCA1</i>	Chr17:41215825
		c.1192G>C	<i>ATM</i>	Chr11:108119786
		c.3994-193T>C	<i>ATM</i>	Chr11:108158134
		c.8010+186C>T	<i>ATM</i>	Chr11:108204881
		c.1342C>T	<i>CDH1</i>	Chr16:68849439
		c.376-283T>C	<i>TP53</i>	Chr17:7578837
14	22,6	c.5215+66G>A	<i>BRCA1</i>	Chr17:41215825
		c.376-283T>C	<i>TP53</i>	Chr17:7578837
15	26,5	c.5215+66G>A	<i>BRCA1</i>	Chr17:41215825
		c.560-1G>C	<i>TP53</i>	Chr17:7578290
		c.376-283T>C	<i>TP53</i>	Chr17:7578837
16	28,5	c.3994-193T>C	<i>ATM</i>	Chr11:108158134
		c.8010+186C>T	<i>ATM</i>	Chr11:108204881
		c.376-283T>C	<i>TP53</i>	Chr17:7578837

Tolerant) и PolyPhen2 [13, 14]. Информацию о частоте мутаций получали на основе данных проекта 1000 Genomes и консорциума ExAC [15, 16].

Всего в результате проведенного анализа отобрали 14 мутаций, влияющих на белковую последовательность: *BRCA2* — с.4828G>A (p.Val1610Met), с.5070A>C (p.Lys1690Asn); *TP53* — с.524G>A (p.Arg175His), с.469G>T (p.Val157Phe); *CHEK2* — с.1289C>T (p.Thr430Ile); *ATM* — с.146C>G (p.Ser49Cys), с.4258C>T (p.Leu1420Phe), с.1192G>C (p.Asp398His); *CDH1* — с.790C>T (p.Gln264), с.1342C>T (p.Gln448); *BRCA1* — с.1865C>T (p.Ala622Val), с.384G>A (p.Met128Ile), с.54G>T (p.Met18Ile).

ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

Для детекции известных мутаций в генах, связанных с риском развития РМЖ, в России получили наибольшее распространение методы, основанные на полимеразной цепной реакции. Однако на современном этапе развития методов генетической диагностики одной из наиболее перспективных технологий выявления генетических вариантов в клетках злокачественных опухолей является высокопроизводительное секвенирование. При этом наиболее актуально его применение для изучения вариативности исследуемых участков геномов опухолей, характеризующихся высокой генетической гетерогенностью. В данной работе нами был исследован методом высокопроизводительного секвенирования спектр мутаций ключевых онкогенов при РМЖ и проведено тестирование разработанного ранее алгоритма для биоинформатической обработки NGS-данных.

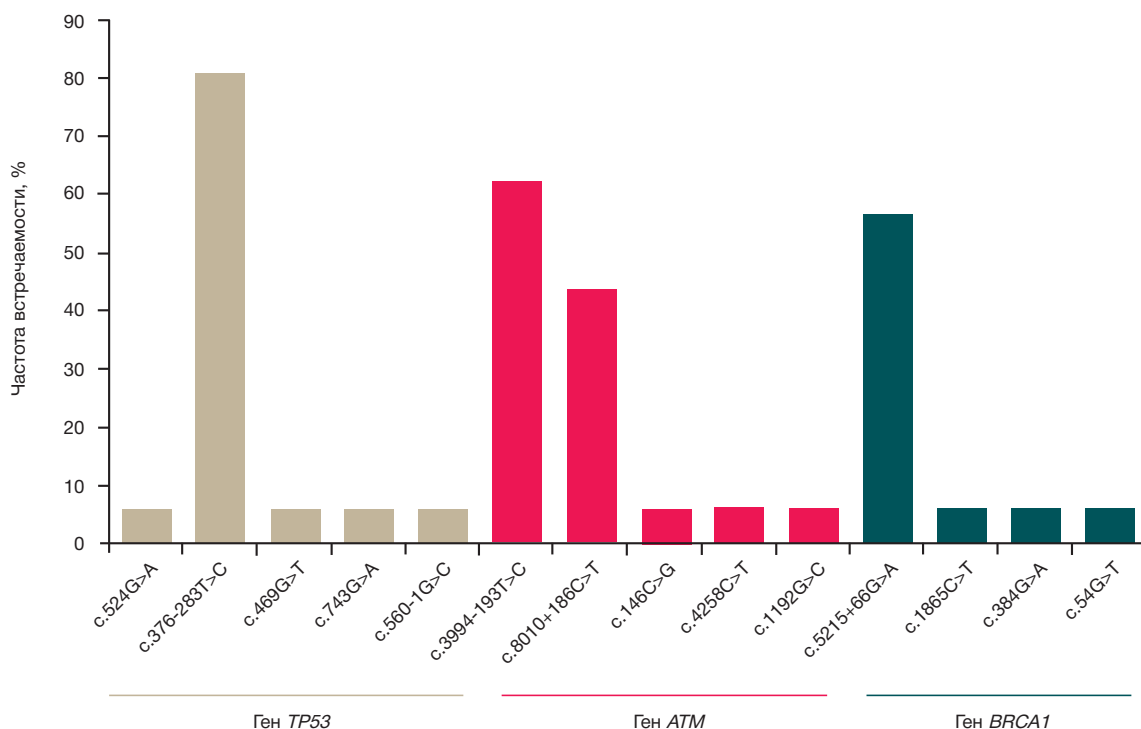
Одним из наиболее детально изученных генов, играющих важную роль в патогенезе РМЖ, является *TP53*, регулирующий клеточный цикл, процессы апоптоза и репарации ДНК. Мутации в этом гене приводят к нарушению вышеуказанных процессов и способствуют развитию злокачественных новообразований. *TP53* представляет

собой онкосупрессор, мутации в нем обнаруживаются примерно в половине всех случаев онкологических заболеваний и более чем в 30 % случаев РМЖ. Что касается спорадического РМЖ, то при нем частота мутаций в гене *TP53* варьирует от 25 до 86 % — в зависимости от стадии опухолевого процесса и методов детекции. Прогностическая значимость мутаций в *TP53* при РМЖ была подробно изучена [17]. Среди мутаций, выявленных нами, чаще всего встречалась мутация с.376-283T>C: у 13 из 16 пациентов (81 %).

При РМЖ отмечена относительно высокая частота возникновения мутаций в генах *BRCA1* и *BRCA2*, в особенности для некоторых форм РМЖ. Эти гены участвуют в регуляции многих клеточных процессов, обеспечивая прежде всего поддержание стабильности генома, гомологичную рекомбинацию при репарации разрывов двуцепочечной ДНК. Мутации в этих генах часто приводят к нарушению нормального протекания вышеуказанных процессов и являются этиологическим фактором развития наследственного РМЖ, значительно увеличивают индивидуальный риск развития рака. Примерно четверть всех случаев наследственного РМЖ ассоциирована с мутациями в *BRCA1/2* [17].

В России распространены мутации в *BRCA1*, составляющие порядка 80 % всех мутаций в этих двух генах при РМЖ. К примеру, одной из наиболее часто встречающихся является мутация 5382insC (*rs80357906*), приводящая к сдвигу рамки считывания и потере функции кодируемым белком. Значительную долю обнаруженных нами вариантов также составили мутации в генах *BRCA1* и *BRCA2*, при этом чаще всего встречалась мутация с.5215+66G>A (*rs3092994*) в гене *BRCA1*, обнаруженная у 9 из 16 пациентов (52,9 %).

Результаты общего анализа спектра мутаций в генах *ATM*, *TP53* и *BRCA1* в целом согласуются с данными литературы, в соответствии с которыми при РМЖ к наиболее часто встречающимся мутациям относятся изменения в гене *TP53* [17]. Сопоставимы с известными данными и



Распределение частот встречаемости мутаций в генах *ATM*, *TP53* и *BRCA1* у пациенток с раком молочной железы

результаты оценки разнообразия генетических вариантов в генах *BRCA1/2*. Важно, что мутации в этих генах ассоциированы с неблагоприятным прогнозом и развитием инфильтративного протокового РМЖ. Наличие таких мутаций также учитывается при оценке объема хирургического вмешательства [17]. По нашим данным, из 12 пациенток с РМЖ и мутациями в генах *BRCA1* и *BRCA2* у 8 был диагностирован именно инфильтративный протоковый РМЖ. При этом наиболее часто (у 6 пациенток из 8) встречалась мутация с.5215+66G>A (*BRCA1*).

В результате биоинформатического анализа данных, полученных в ходе высокопроизводительного секвенирования, нами было обнаружено множество мутаций-драйверов в образцах злокачественных опухолей молочной железы. С применением различных баз данных были отобраны и аннотированы мутации со значимым функциональным эффектом. Всего нами было обнаружено 14 патогенных мутаций, оказывающих влияние на аминокислотную последовательность кодируемых белков. В каждом из исследованных образцов была обнаружена минимум одна такая мутация. Разработанный нами протокол для биоинформатического анализа NGS-данных позволяет обрабатывать их в автоматическом режиме.

Литература

- DeSantis CE, Ma J, Goding Sauer A, Newman LA, Jemal A. Breast Cancer Statistics, 2017, Racial Disparity in Mortality by State. *CA Cancer J Clin.* 2017 Nov; 67 (6): 439–48. DOI: 10.3322/caac.21412.
- Yohe S, Thyagarajan B. Review of Clinical Next-Generation Sequencing. *Arch Pathol Lab Med.* 2017 Nov; 141 (11): 1544–57. DOI: 10.5858/arpa.2016-0501-RA.
- Melchor L, Benítez J. The complex genetic landscape of familial breast cancer. *Hum Genet.* 2013 Aug; 132 (8): 845–63. DOI: 10.1007/s00439-013-1299-y.
- Giuliano AE, Connolly JL, Edge SB, Mittendorf EA, Rugo HS, Solin LJ et al. Breast Cancer-Major changes in the American Joint Committee on Cancer eighth edition cancer staging manual. *CA Cancer J Clin.* 2017 Jul 8; 67 (4): 290–303. DOI: 10.3322/caac.21393.
- Цуканов К. Ю., Красненко А. Ю., Плахина Д. А., Коростин Д. О., Чуров А. В., Дружиловская О. С. и др. Биоинформатический протокол для обработки NGS-данных и идентификации мутаций в солидных опухолях человека. *Биомедицинская химия.* 2017; 63 (5): 413–7. DOI: 10.18097/PBMC20176305413.
- Martin M. Cutadapt Removes Adapter Sequences from High-Throughput Sequencing Reads. *EMBnet.journal.* 2011; 17: 10–2.
- Li H, Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics.* 2009 Jul 15; 25 (14): 1754–60. DOI: 10.1093/bioinformatics/btp324.
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N et al. 1000 Genome Project Data Processing Subgroup. The Sequence Alignment/Map format and SAMtools. *Bioinformatics.* 2009 Aug 15; 25 (16): 2078–9. DOI: 10.1093/bioinformatics/btp352.
- Cibulskis K, Lawrence MS, Carter SL, Sivachenko A, Jaffe D, Sougnez C et al. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat Biotechnol.*

References

- DeSantis CE, Ma J, Goding Sauer A, Newman LA, Jemal A. Breast Cancer Statistics, 2017, Racial Disparity in Mortality by State. *CA Cancer J Clin.* 2017 Nov; 67 (6): 439–48. DOI: 10.3322/caac.21412.
- Yohe S, Thyagarajan B. Review of Clinical Next-Generation

Выводы

Высокопроизводительное секвенирование в комбинации с современными алгоритмами для биоинформатической обработки данных позволяет проводить поиск генетических вариантов и оценку функционального эффекта выявляемых мутаций в опухолях, что представляет интерес для применения данных технологий в медицинских целях. На сегодняшний день уровень развития методов NGS позволяет проводить молекулярную классификацию опухолей при РМЖ, определять подтипы опухолей в зависимости от спектра выявляемых мутаций и профиля экспрессии генов, подбирать на основе этих данных наиболее подходящие противоопухолевые препараты. Одной из главных задач в сфере онкогенетики является поиск и разработка инструментов для идентификации биомаркеров РМЖ, удобных для клиницистов при постановке диагноза и назначении адекватной терапии. Следующим этапом в развитии данной области исследований должно стать совершенствование биоинформатических подходов, переход к системам автоматического анализа генетического профиля опухолей и внедрение технологий NGS в клиническую практику.

- 2013 Mar; 31 (3) :213–9. DOI: 10.1038/nbt.2514.
- Cingolani P, Platts A, Wang le L, Coon M, Nguyen T, Wang L et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly (Austin).* 2012 Apr-Jun; 6 (2): 80–92. DOI: 10.4161/fly.19695.
- Forbes SA, Beare D, Bindal N, Bamford S, Ward S, Cole CG et al. COSMIC: High-Resolution Cancer Genetics Using the Catalogue of Somatic Mutations in Cancer. *Curr Protoc Hum Genet.* 2016 Oct 11; 91: 10.11.1-10.11.37. DOI: 10.1002/cphg.21.
- Liu X, Wu C, Li C, Boerwinkle E. dbNSFP v3.0: A One-Stop Database of Functional Predictions and Annotations for Human Nonsynonymous and Splice-Site SNVs. *Hum Mutat.* 2016 Mar; 37 (3): 235–41. DOI: 10.1002/humu.22932.
- Ng PC, Henikoff S. SIFT: Predicting amino acid changes that affect protein function. *Nucleic Acids Res.* 2003 Jul 1; 31 (13): 3812–4.
- Adzhubei I, Jordan DM, Sunyaev SR. Predicting functional effect of human missense mutations using PolyPhen-2. *Curr Protoc Hum Genet.* 2013 Jan; Chapter 7: Unit7.20. DOI: 10.1002/0471142905.hg0720s76.
- 1000 Genomes Project Consortium, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, Korbel JO et al. A global reference for human genetic variation. *Nature.* 2015 Oct 1; 526 (7571): 68–74. DOI: 10.1038/nature15393.
- Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature.* 2016 Aug 18; 536 (7616): 285–91. DOI: 10.1038/nature19057.
- Гришина К. А., Музаффарова Т. А., Хайленко В. А., Карпунин А. В. Молекулярно-генетические маркеры молочной железы. *Опухоли женской репродуктивной системы.* 2016; 12 (3): 36–42. DOI: 10.17650/1994-4098-2016-12-3-26-42.

- Sequencing. *Arch Pathol Lab Med.* 2017 Nov; 141 (11): 1544–57. DOI: 10.5858/arpa.2016-0501-RA.
- Melchor L, Benítez J. The complex genetic landscape of familial breast cancer. *Hum Genet.* 2013 Aug; 132 (8): 845–63. DOI: 10.1007/s00439-013-1299-y.

4. Giuliano AE, Connolly JL, Edge SB, Mittendorf EA, Rugo HS, Solin LJ et al. Breast Cancer-Major changes in the American Joint Committee on Cancer eighth edition cancer staging manual. *CA Cancer J Clin*. 2017 Jul 8; 67 (4): 290–303. DOI: 10.3322/caac.21393.
5. Tsukanov KYu, Krasnenko AYu, Plakhina DA, Korostin DO, Churov AV, Druzhilovskaya OS et al. A bioinformatic pipeline for NGS data analysis and mutation calling in human solid tumors. *Biomed Khim*. 2017; 63 (5): 413–7.
6. Martin M. Cutadapt Removes Adapter Sequences from High-Throughput Sequencing Reads. *EMBnet.journal*. 2011; 17: 10–2.
7. Li H, Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*. 2009 Jul 15; 25 (14): 1754–60. DOI: 10.1093/bioinformatics/btp324.
8. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N et al. 1000 Genome Project Data Processing Subgroup. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 2009 Aug 15; 25 (16): 2078–9. DOI: 10.1093/bioinformatics/btp352.
9. Cibulskis K, Lawrence MS, Carter SL, Sivachenko A, Jaffe D, Sougnez C et al. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat Biotechnol*. 2013 Mar; 31 (3) :213–9. DOI: 10.1038/nbt.2514.
10. Cingolani P, Platts A, Wang le L, Coon M, Nguyen T, Wang L et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly (Austin)*. 2012 Apr-Jun; 6 (2): 80–92. DOI: 10.4161/fly.19695.
11. Forbes SA, Beare D, Bindal N, Bamford S, Ward S, Cole CG et al. COSMIC: High-Resolution Cancer Genetics Using the Catalogue of Somatic Mutations in Cancer. *Curr Protoc Hum Genet*. 2016 Oct 11; 91: 10.11.1-10.11.37. DOI: 10.1002/cphg.21.
12. Liu X, Wu C, Li C, Boerwinkle E. dbNSFP v3.0: A One-Stop Database of Functional Predictions and Annotations for Human Nonsynonymous and Splice-Site SNVs. *Hum Mutat*. 2016 Mar; 37 (3): 235–41. DOI: 10.1002/humu.22932.
13. Ng PC, Henikoff S. SIFT: Predicting amino acid changes that affect protein function. *Nucleic Acids Res*. 2003 Jul 1; 31 (13): 3812–4.
14. Adzhubei I, Jordan DM, Sunyaev SR. Predicting functional effect of human missense mutations using PolyPhen-2. *Curr Protoc Hum Genet*. 2013 Jan; Chapter 7: Unit7.20. DOI: 10.1002/0471142905.hg0720s76.
15. 1000 Genomes Project Consortium, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, Korbel JO et al. A global reference for human genetic variation. *Nature*. 2015 Oct 1; 526 (7571): 68–74. DOI: 10.1038/nature15393.
16. Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature*. 2016 Aug 18; 536 (7616): 285–91. DOI: 10.1038/nature19057.
17. Grishina KA, Muzaffarova TA, Khaylenko VA, Karpukhin AV. [Molecular genetic markers of breast cancer]. *Tumors of female reproductive system*. 2016; 12 (3): 36–42. DOI: 10.17650/1994-4098-2016-12-3-26-42. Russian.