



Оценка возможности применения нанопорового секвенирования ампликонов для определения однонуклеотидных полиморфизмов на примере образцов ДНК В-клеточных лимфом

^{1,2}Е.С. Жданова Е.С., ²А.М. Морозова А.М., ^{1,2}А.М. Ермаков*

¹ ООО «Нанопорус», Российская Федерация, 142201, Московская обл., г. Серпухов, ул. Сиреневая д.8.

² ФГБУН Институт теоретической и экспериментальной биофизики РАН,
Российская Федерация, 142290, Московская обл., г. Пущино, ул. Институтская, д.3.

*E-mail: ao_ermakov@rambler.ru

Резюме

Введение. Высокопроизводительное секвенирование ДНК в настоящее время является стандартной частью лечения онкологических заболеваний, как для диагностики лечения, так и на ранних стадиях скрининга. Среди методов секвенирования доминирует NGS, но он имеет ряд существенных недостатков. Развивающаяся технология нанопорового секвенирования, преодолевает все вышеперечисленные ограничения NGS. Точность такого секвенирования приблизилась к NGS, а сама технология более доступна в плане бюджета. Нанопоровое секвенирование уже нашло широкое применение в различных приложениях и имеется значительная перспектива его применения в анализе полиморфизмов для выявления генетических вариаций для диагностических и прогностических задач.

Цель исследования: на основе разработанных ампликонных панелей для детекции однонуклеотидных полиморфизмов на примере образцов ДНК лимфом провести исследование по оптимизации процесса амплификации, нанопорового секвенирования и анализа данных с применением отечественной реагентной базы и отечественного нанопорового секвенатора «Нанопорус Про».

Материалы и методы. Выделенная и очищенная с помощью спин-колонок ДНК из парафиновых срезов была амплифицирована при помощи подобранных праймеров (47 пар) в условиях 4-х сетового мультиплекса. Полученные библиотеки ДНК были секвенированы при помощи нанопорового секвенатора «Нанопорус Про».

Результаты и заключение. Были исследованы и оптимизированы условия мультиплексной ПЦР на основе нескольких вариантов отечественных ДНК полимераз. В ходе анализа полученных данных секвенирования было определено 45 из 47 референсных мишеней и получены высокоточные данные о наличии однонуклеотидных полиморфизмов в исследуемых точках вариаций. Таким образом, была показана высокая перспективность применения нанопорового секвенирования для мультиплексного анализа клинически-важных мутаций в участках генома раковых опухолей с использованием отечественной реагентики и оборудования.

Ключевые слова: нанопоровое секвенирование, NGS, однонуклеотидные полиморфизмы, лимфомы

Цитирование: Жданова Е.С., Морозова А.М., Ермаков А.М. Оценка возможности применения нанопорового секвенирования ампликонов для определения однонуклеотидных полиморфизмов на примере образцов ДНК В-клеточных лимфом. *Biomics*. 2025. 17(3). 229-237. DOI: 10.31301/2221-6197.bmcs.2025-19

© **Авторы,** Жданова Е.С., Морозова А.М., Ермаков А.М., 2025

Evaluation of the possibility of using nanopore sequencing of amplicons for the determination of single nucleotide polymorphisms using the example of DNA samples from B-cell lymphomas

Zhdanova E.S.^{1,2}, Morozova A.M.², Ermakov AM^{1,2*}

¹Nanoporus LLC, Russian Federation, 142201, Moscow region, Serpukhov, Sirenevaya str., 8.

²Institute of Theoretical and Experimental Biophysics of the Russian Academy of Sciences, Russian Federation, 142290, Moscow region, Pushchino, Institutskaya str., 3

*E-mail: ao_ermakovy@rambler.ru

Resume

Introduction. High-throughput DNA sequencing is currently a standard part of cancer treatment, both for treatment diagnosis and early-stage screening. NGS dominates among the sequencing methods, and it has a number of significant drawbacks. The evolving nanopore sequencing technology overcomes all of the above limitations of NGS. The accuracy of such sequencing is closer to NGS, and the technology itself is more affordable in terms of budget. Nanopore sequencing has already found wide application in various applications and there is a significant prospect of its application in polymorphism analysis to identify genetic variations for diagnostic and prognostic tasks.

The purpose of the study: based on the developed amplicon panels for the detection of single nucleotide polymorphisms on the example of DNA lymphoma samples, to conduct a study on optimizing the amplification process, nanopore sequencing and data analysis using a domestic reagent database and a domestic nanopore sequencer "Nanoporus Pro".

Materials and methods. The DNA isolated and purified using spin columns from paraffin sections was amplified using selected primers (47 pairs) in a 4-set multiplex. The resulting DNA libraries were sequenced using the nanopore sequencer Nanoporus Pro.

Results and conclusion. Multiplex PCR conditions based on several variants of domestic DNA polymerases were investigated and optimized. During the analysis of the obtained sequencing data, 45 out of 47 reference targets were identified and highly accurate data on the presence of single nucleotide polymorphisms at the studied points of variation were obtained. Thus, the high prospects of using nanopore sequencing for multiplex analysis of clinically important mutations in cancer genome regions using domestic reagents and equipment have been shown.

Keywords: nanopore sequencing, NGS, single nucleotide polymorphisms, lymphomas

Citation: Zhdanova E.S., Morozova A.M., Ermakov AM. Evaluation of the possibility of using nanopore sequencing of amplicons for the determination of single nucleotide polymorphisms using the example of DNA samples from B-cell lymphomas. *Biomics*. 2025. 17(3). 229-237. DOI: 10.31301/2221-6197.bmcs.2025-19

© **The Authors**, Zhdanova E.S., Morozova A.M., Ermakov AM., 2025

Введение

Высокопроизводительное секвенирование ДНК уже стало стандартной частью лечения онкологических заболеваний, при этом множество лабораторно разработанных тестов дают информацию о прогнозе для пациента [Dowsett et al., 2013], выборе терапии [Simon et al., 2013; Steuer et al., 2018] и минимальной остаточной болезни [Tie et al., 2016]. Кроме того, секвенирование ДНК изучается как метод, позволяющий проводить ранний скрининг рака у бессимптомных групп населения [Chabon et al., 2020]. В настоящее время в клиническом высокопроизводительном секвенировании и

исследованиях доминируют платформы на основе секвенирования путем синтеза (секвенирование следующего поколения (NGS)) – это секвенаторы Illumina и MGI/BGI.

Несмотря на то, что NGS обладает множеством преимуществ, включая высокую производительность, высокую чувствительность, надежность и воспроизводимость, у NGS также есть несколько существенных ограничений: это - длина прочтений NGS ограничена примерно 300-400 (при парноконцевых прочтениях) нуклеотидами, что делает его менее подходящим и чувствительным к более протяженным изменениям ДНК, таким как

вариации числа копий и хромосомные транслокации [Cohen et al., 2018]. Также NGS требует много времени для проведения секвенирования, так как там происходит синтез ДНК и весь процесс может занимать несколько дней без возможности проводить оценку и анализ данных в режиме реального времени. Пробоподготовка образцов для этого типа секвенирования длительна и трудоемка, а биоинформатическая интерпретация данных требует немалых вычислительных ресурсов [Satam et al., 2023]. И самое главное - NGS требует высоких капиталовложений, порядка 1 млн долларов США достигает стоимость для высокопроизводительного инструмента (например, Novaseq), а более доступные платформы NGS, такие как MiSeq, приводят к примерно 10-кратному увеличению затрат на секвенирование за прочтение.

Развивающееся сегодня нанопоровое секвенирование (называемое также секвенированием 3-го поколения) преодолевает все вышеперечисленные ограничения NGS. Данная технология основывается на прямом прочтении цепочки ДНК посредством нанопоры, и длина единичного прочтения может быть от 20 до 1 млн нуклеотидов [Baslan et al., 2021]. В ранних версиях реагентов и проточных ячеек точность такого секвенирования не была высокоточной - достигала порядка 95% (Q10-15), за что эта технология и подвергалась критике [Sanderson et al., 2023]. Дальнейшее развитие технологии компанией Oxford Nanopore привело к созданию высокоточной химии и нейросетевых моделей бейзколинга, которые сейчас позволяют получить точность единичных прочтений на уровне 99.75% (Q26), что уже сравнимо с ранее считавшимся эталонным качеством секвенирования на приборах NGS [Hall et al., 2024]. И самое главное преимущество – нанопоровые секвенаторы портативны, пробоподготовка может занимать от 20 мин и не требует сложно оборудованной лаборатории, стоимость приборов и расходных материалов значительно ниже, чем в NGS. При этом анализ получаемых данных секвенирования можно проводить уже через 5 мин после запуска в режиме реального времени без использования больших вычислительных ресурсов [Magi et al., 2018]. Этот тип секвенирования уже нашел свое широкое применение в анализе геномов, секвенировании 16S рРНК, ампликонном секвенировании вирусов и др. [MacKenzie et al., 2023]. Поэтому перспективно использовать данную технологию также и для анализа полиморфизмов в ампликонном секвенировании для выявления генетических

вариаций для диагностических и прогностических задач.

В настоящем исследовании на основе разработанных ампликонных панелей для детекции однонуклеотидных полиморфизмов на примере образцов ДНК лимфом было проведено исследование по оптимизации процесса амплификации, нанопорового секвенирования и анализа данных с применением отечественной реагентной базы и отечественного нанопорового секвенатора «Нанопорус Про».

Материалы и методы

Выделение ДНК из парафиновых срезов образцов тканей опухолей.

В качестве образцов для проверки метода были выбраны образцы срезов из парафиновых блоков опухолей лимфом. ДНК из парафиновых блоков выделяли с помощью набора SKY Amp FFPE DNA (Скайджин, Россия) согласно прилагаемому протоколу. Концентрацию полученной ДНК измеряли с помощью прибора Qubit 4 (Thermo FS, США) набором QuDye HS (Люмипроб, Россия). Качество выделенной ДНК проверяли с помощью горизонтального агарозного гель-электрофореза. *Мультиплексная ПЦР, получение ампликонов.*

Для амплификации целевых участков геномной ДНК каждого исследуемого образца, содержащей необходимые SNP использовали 4 сета праймеров. Каждый из них состоял из 11-12 пар праймеров, что обеспечивало мультиплицирование от 11 до 12 мишеней в виде SNP. Мишени выбирались, исходя из литературных данных по молекулярному профилированию В-клеточных лимфом [Morin et al., 2022; Alduaij et al., 2023].

Праймеры к целевым участкам подбирали с помощью онлайн-сервиса Primer Blast NCBI (NCBI,). Размер предполагаемых ампликонов составлял от 400 до 900 пар нуклеотидов. Для постановки ПЦР реакции использовали несколько наборов - Genta Taq-nAB ДНК-полимераза и Genta TaqF ДНК-полимераза (Гентерра, Россия), БиоМастер LR HS-ПЦР (2x) (Биолабмикс, Россия), а также QIAGEN Multiplex PCR Kit (QIAGEN, Германия). Реакцию проводили в объеме 50 мкл, в каждую реакцию добавляли 30 нг ДНК. Схема циклов была следующая: Предварительная денатурация: 94°C - 3 мин, далее 35 циклов: 94°C - 30 сек, 60°C - 1 мин, 68°C - 4 мин и терминальная элонгация 68°C - 10 мин.

Продукты амплификации проверяли с помощью горизонтального агарозного гель-электрофореза.

Таблица 1.

Список генов и полиморфизмов, анализируемых в исследовании, а также распределение по ПЦР сетам.

Table 1. List of genes and polymorphisms analyzed in the study, as well as distribution by PCR sets.

сет 1 / set 1	сет 2 / set 2	сет 3 / set 3	сет 4 / set 4
TCF3_rs1860661	ID3_rs2742946	FBXO11_rs2537742	CREBBP_rs3025684
TCF3_rs2074888	DDX3X_rs953114	FBXO11_rs41295278	CREBBP_rs2239317
TCF3_rs55677929	CCND3_rs3218086	FBXO11_rs55740729	CREBBP_rs11644593
TCF3_rs879255271	FOXO1_rs17592236	FBXO11_rs2710163	CREBBP_rs12599143
TCF3_rs1052918	TFAP4_rs251732	FBXO11_rs3732191	CCND3_rs33966734
TCF3_rs41275834	TNFRSF14_rs4870	FBXO11_rs6713506	ARID1A_rs34502618
TCF3_rs72618599	TNFRSF14_rs2234163	KMT2D_rs12580349	ARID1A_rs12685
CCND3_rs9529	TNFRSF14_rs2234167	KMT2D_rs7975791	SMARCA4_rs11880865
CCND3_rs1051130	TNFRSF14_rs2281852	KMT2D_rs11168831	SMARCA4_rs11879293
CCND3_rs1410492	CREBBP_rs130005	CREBBP_rs129963	SMARCA4_rs11085754
CCND3_rs2479717	CREBBP_rs130021	CREBBP_rs129984	SMARCA4_rs3786725
	KMT2D_rs11614738	CCND3_rs3218092	SMARCA4_rs2072382

Подготовка библиотеки ампликонов для нанопорового секвенирования.

Библиотеку для секвенирования образцов ДНК ампликонов готовили с использованием набора SQK-NBD114.96 (Oxford Nanopore, Англия) согласно прилагаемым протоколам.

В ходе первичной обработки образцов из каждого экспериментального образца, содержащего предоставленные ампликоны, готовили общий пул, путем смешения 4-х пулов по 10 мкл каждого ампликона в общую пробирку. Далее производили очистку ампликонов с помощью магнитных частиц SynMag (Синтол, Россия). Для этого в пул ампликонов добавляли равный объем магнитных частиц, пипетировали и вращали в ротаторе 5 мин. Далее осаждали частицы на магнитном штативе, отбирали супернатант и промывали частицы 2 раза по 200 мкл 80% этанолом не снимая с магнита. После последней промывки отбирали этанол, и частицы в пробирке просушивали 30 сек. Далее производили элюирование ампликонной ДНК деионизированной водой (10 мкл).

Для приготовления библиотеки использовали порядка 300 нг ДНК ампликонов из каждого анализируемого образца. Процедуру проводили согласно рекомендуемым протоколам, прилагаемым к набору SQK-NBD114.96. Для этого необходимое количество мкл каждого образца ДНК доводили до 12,5 мкл воды качества MilliQ и смешивали с 1,75 мкл Ultra II End-prep reaction buffer и 0,75 мкл Ultra II End-prep enzyme mix (NEBNext Ultra II End Repair/dA-Tailing Module, NEB, Англия). Реакцию проводили в режиме 20°C - 5 мин и 65°C - 5 мин в амплификаторе C1000 (BioRad, США) и после производили очистку реакционной смеси магнитными частицами AMPure

XP (Beckman, США). ДНК элюировали в воде качества MilliQ в объеме 10 мкл при комнатной температуре в течение 5 мин.

Далее производили лигирование ампликонной ДНК каждого образца с баркодом. Для этого смешивали 3,75 мкл очищенной после первого этапа пробоподготовки ДНК с 1,25 мкл нативного баркода (входящий в состав набора SQK-NBD114.96) и 5 мкл ферментной смеси Blunt/TA Ligase Master Mix (NEB, Англия) и инкубировали смесь в течение 20 мин при комнатной температуре и затем добавляли 1 мкл 0,5 М раствора ЭДТА для остановки реакции. После смешивали все образцы после реакции лигирования баркодов в один пул и очищали его магнитными частицами AMPure XP в соотношении 1 к 0,4. Далее производили элюирование ДНК в воде качества MilliQ в объеме 31 мкл.

Далее смешивали 30 мкл подготовленной ДНК после первого этапа, Native Adapter (NA) (входит в состав набора SQK-NBD114.96) – 5 мкл, NEBNext Quick Ligation Reaction Buffer (5X) (NEB, США) – 10 мкл и 5 мкл NEBNext Quick T4 DNA Ligase (NEB). Реакцию лигирования адаптеров проводили при комнатной температуре в течение 20 мин. Реакционную смесь очищали магнитными частицами AMPure XP, при этом промывку частиц осуществляли с помощью SFB буфера (входит в состав набора SQK-NBD114.96). Библиотеку ДНК элюировали в 33 мкл буфера EB (входит в состав набора SQK-NBD114.96) в течение 15 мин при температуре 37°C. На всех промежуточных этапах подготовки библиотеки ДНК осуществляли контроль количества и качества ДНК путем измерения

концентрации (на приборе Qubit 4 (Thermo FS, США) с помощью набора QuDye HS (Люмипроб, Россия).

Подготовленные библиотеки ДНК секвенировали с помощью нанопорового секвенатора «Нанопорус Про» (ООО «Нанопорус», Россия) с ячейкой FLO-PRO114M использованием базового программного обеспечения MinKNOW (Oxford Nanopore, Англия). Использовали 300 нг библиотеки ДНК, полученную смешиванием баркодированных библиотек для мультиплексирования реакции. Библиотеку готовили путем смешивания 100 мкл буфера Sequencing Buffer (SB), 68 мкл раствора Library Beads (LIB) и 32 мкл подготовленной библиотеки ДНК. Полученную смесь загружали в ячейку и запускали процедуру секвенирования. Бейзколлинг и демультиплексирование полученных данных проводили с помощью программы Dorado v1.1.0 (Oxford Nanopore), с использованием модели dna_r10.4.1_e8.2_400bps_sup@v5.2.0 на базе ПО Linux. *Первичная биоинформатическая и статистическая обработка данных секвенирования.*

Первичный статистический анализ длины, качества ридов, представленности ридов каждого образца производили с помощью стандартных отчетов, генерируемых программой MinKNOW (Oxford Nanopore).

Анализ данных ампликонного секвенирования. Анализ данных ампликонного секвенирования (выравнивание

ридов на референсные последовательности ампликонов в формате bam, сгенерированы консенсусные последовательности для каждого вида ампликонов в формате fasta и отчет в виде html файла с информацией для каждого обнаруженного вида ампликонов о покрытии, SNP, SNV) осуществляли с помощью конвейера wf-amplicon (Oxford Nanopore, Англия, <https://github.com/epi2me-labs/wf-amplicon>). Данный конвейер объединяет рабочие процессы по фильтрам качества, длины ридов, выравнивания на референс и генерации консенсусной последовательности, также позволяет определять SNP и SNV. В качестве референсных последовательностей использовали последовательности интересующих участков генов в виде fasta файла. В настройках первичной обработки ридов конвейера ставили обрезку используемых для анализа ридов по длине от 400 до 900 н.п., число используемых для выравнивания ридов не ограничивали, минимальное качество ридов устанавливали на уровне $Q \geq 10$.

Результаты и обсуждение

В результате амплификации целевых участков в 4-х мультиплексных реакциях с разными полимеразой для каждого образца были получены смеси ампликонов, визуализация продуктов амплификации представлена на рисунке 1.

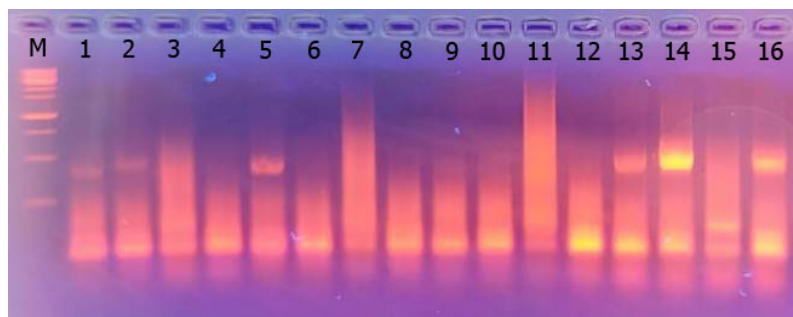


Рисунок 1. Визуализация продуктов амплификации в 4-х праймерных мультиплексных наборах полученных на разных ферментных смесях. М - маркер (250 – 10 000 нп), 1-4 продукты амплификации Genta Taq-nAB полимеразой, 5-8 - продукты амплификации Genta TaqF полимеразой, 9-12 - продукты амплификации BioMaster LR HS-ПЦР, 13-16 - QIAGEN Multiplex PCR Kit.

Figure 1. Visualization of amplification products in 4 primer multiplex sets obtained on different enzyme mixtures. M - marker (250 – 10 000 np), 1-4 - products of Genta Taq-nAB polymerase amplification, 5-8 - products of Genta TaqF polymerase amplification, 9-12 - products of BioMaster LR HS-PCR amplification, 13-16 - QIAGEN Multiplex PCR Kit.

Продукты амплификации визуализируются в виде непрерывных шмеров с выраженными полосами продуктов с более высокой концентрацией. Картина несколько отличается при амплификации образцов разными полимеразой, но в целом можно утверждать, что ПЦР реакция идет во всех 4-х наборах праймеров при использовании всех испытываемых полимераз.

Секвенирование образцов ампликонных библиотек, полученных в ходе смешения продуктов амплификации 4-х наборов каждого из образцов, позволило выявить 45 из 47 референсных последовательностей, которые должны присутствовать в образцах. Качество единичных прочтений составляло в среднем Q18, средняя длина

прочтений 559 нуклеотидов, максимальный выход данных получен на длинах ридов от 400 нуклеотидов (Рисунок 2). Полученные статистические результаты соответствуют ожидаемым, так как большинство

ампликонов представлено размерным рядом порядка 400 – 500 нуклеотидов.

Анализ количества детектируемых ампликонов в каждом из анализируемом представлен в таблице 2.

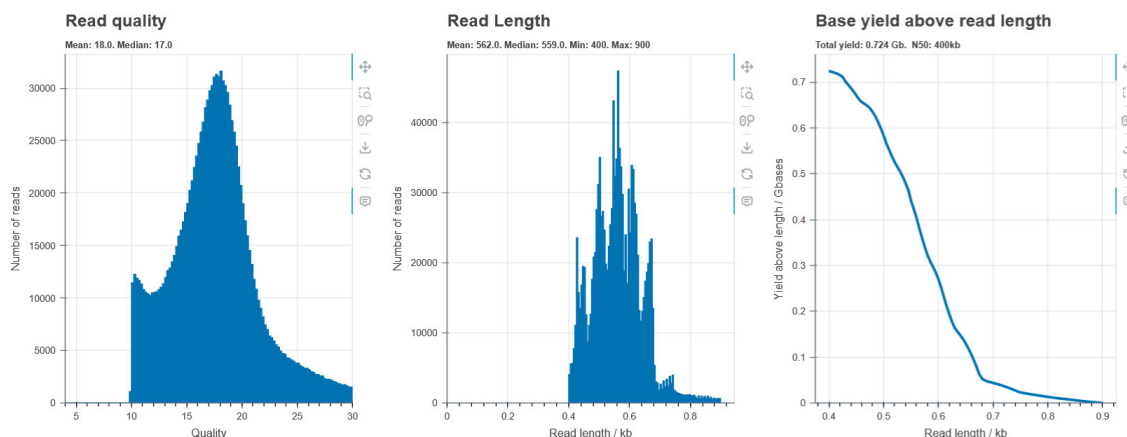


Рисунок 2. Статистические параметры прочтений, полученные в результате секвенирования ампликонов.
Figure 2. Statistical parameters of readings obtained as a result of amplicon sequencing.

Таблица 2.

Суммарные результаты секвенирования по исследуемым образцам.
Table 2. Total sequencing results for the studied samples.

№	1	2	3	4
Образец (тип используемой полимеразы) Sample (type of polymerase used)	Genta Taq-nAB	Genta TaqF	БиоМастер LR HS BioMaster LR HS	QIAGEN Multiplex PCR Kit
Общее число ридов Total number of reads	18.0 k	8.1 k	439.4 k	11.6 k
Число оснований Number of bases	7.4 M	3.5 M	197.1 M	5.1 M
Средняя длина рида Average read length	406	425	450	433
Число видов ампликонов Number of types of amplicons	36	44	45	45
Не маппировано Not mapped	4.3 k (24%)	1.1 k (14%)	38.7 k (9%)	1.4 k (12%)
Полиморфизмы SNP/SNV	57	97	122	109

Как показано в таблице, наиболее представленными были ампликонные библиотеки, полученные с применением ферментных смесей 2 – 3, при этом наибольшее число не маппированных на референсные последовательности ридов наблюдалось в вариантах амплификации 1 и 2.

Был выполнен анализ суммарного числа ридов и оснований для каждого из выявленного в

ходе секвенирования ампликона, а также значение среднего глубины покрытия ридами каждого нуклеотида референсной последовательности (среднее покрытие) и обнаруженное число вариантов в виде однонуклеотидных замен и вставок/делеций (таблица 3).

Таблица 3. Статистические данные ампликонов по отношению к референсу и число вариантов в виде однонуклеотидных замен и вставок/делеций / Table 3. Statistical data for the amplicons in relation to the reference, and the number of variants in the form of single nucleotide polymorphisms and insertions/deletions.

№	Ампликон / Amplicon	Риды Reads	Основания Bases	Среднее покрытие Average cover	Число вариантов Number of variants
1	CCND3_rs1410492	14	5.9 k	41.0	0 (0)
2	KMT2D_rs7975791	15	6.5 k	50.3	10 (3)
3	ARID1A_rs12685	23	11.1 k	53.2	168 (16)
4	FBXO11_rs3732191	34	17.1 k	68.6	8 (4)
5	TNFRSF14_rs2281852	38	19.0 k	49.7	20 (12)
6	KMT2D_rs12580349	40	19.2 k	68.4	3 (0)
7	SMARCA4_rs11880865	41	18.3 k	80.8	13 (6)
8	FBXO11_rs6713506	104	45.7 k	49.6	153 (22)
9	CREBBP_rs12599143	151	75.0 k	57.5	0 (0)
10	FBXO11_rs2710163	169	82.1 k	88.0	4 (0)
11	CREBBP_rs2239317	282	142.7 k	89.7	19 (1)
12	ID3_rs2742946	437	204.8 k	49.7	7 (6)
13	TCF3_rs879255271	504	206.3 k	88.1	10 (4)
14	SMARCA4_rs11879293	572	301.0 k	74.9	11 (5)
15	DDX3X_rs953114	853	436.2 k	65.9	15 (1)
16	CREBBP_rs129984	2.4 k	1.1 M	85.9	0 (0)
17	TNFRSF14_rs2234163	2.5 k	1.4 M	91.4	1 (0)
18	CCND3_rs9529	2.6 k	1.3 M	91.2	9 (0)
19	TFAP4_rs251732	3.2 k	1.6 M	89.5	10 (0)
20	FBXO11_rs41295278	4.1 k	1.4 M	75.8	23 (11)
21	SMARCA4_rs3786725	4.2 k	2.2 M	92.1	2 (0)
22	FBXO11_rs55740729	4.2 k	1.5 M	50.5	15 (4)
23	SMARCA4_rs2072382	4.4 k	2.3 M	93.0	0 (0)
24	KMT2D_rs11614738	5.0 k	2.7 M	86.0	5 (0)
25	CREBBP_rs11644593	5.5 k	2.3 M	80.0	0 (0)
26	TCF3_rs41275834	7.0 k	2.7 M	57.6	12 (3)
27	TCF3_rs2074888	7.5 k	3.5 M	93.6	1 (0)
28	KMT2D_rs11168831	8.7 k	3.7 M	95.3	0 (0)
29	TCF3_rs1860661	11.2 k	4.9 M	95.9	2 (0)
30	CREBBP_rs130021	11.6 k	6.3 M	84.5	6 (0)
31	TCF3_rs55677929	13.8 k	5.5 M	96.7	0 (0)
32	CCND3_rs3218092	14.1 k	6.1 M	95.4	1 (1)
33	CCND3_rs33966734	23.8 k	9.9 M	96.4	0 (0)
34	CREBBP_rs130005	31.1 k	16.9 M	92.9	0 (0)
35	CREBBP_rs129963	32.3 k	13.1 M	96.9	5 (0)
36	CREBBP_rs3025684	41.7 k	19.3 M	96.7	0 (0)
37	CCND3_rs1051130	44.8 k	18.8 M	96.8	5 (0)
38	CCND3_rs2479717	57.7 k	26.9 M	96.6	5 (0)
39	SMARCA4_rs11085754	61.9 k	28.2 M	97.4	0 (0)
40	TNFRSF14_rs2234167	70.9 k	34.1 M	96.7	0 (0)
41	TNFRSF14_rs4870	91.0 k	37.7 M	98.0	0 (0)
42	TCF3_rs1052918	104.2 k	42.0 M	97.7	0 (0)
43	TCF3_rs72618599	105.1 k	42.4 M	97.7	0 (0)
44	CCND3_rs3218086	352.4 k	158.9 M	97.8	0 (0)
45	FOXO1_rs17592236	371.2 k	181.2 M	95.5	0 (0)

Наименее представленными из анализируемых видов ампликонов оказались с мишенями-полиморфизмами CCND3_rs1410492, KMT2D_rs7975791, ARID1A_rs12685, FBXO11_rs3732191, TNFRSF14_rs2281852, KMT2D_rs12580349 и SMARCA4_rs11880865. Очевидно, это может быть связано с неоптимальной структурой праймеров для амплификации данных участков или недостаточной оптимизацией условий проведения ПЦР. Суммарный анализ количества обнаруженных полиморфизмов в каждом консенсусе демонстрирует наличие достаточно большого количества обнаруженных полиморфизмов, причем в участках, где относительно небольшое количество ридов, тогда как на участках с высокой глубиной покрытия обнаруживаемых изменений в последовательности нуклеотидов по сравнению с референсной последовательностью невелико или совсем отсутствует. Это объясняется ошибками секвенирования при невысокой глубине покрытия ридами соответствующей референсной последовательности. Более детальный статистический анализ по отношению к каждому анализируемому образцу после фильтрации недостоверно-определяемых изменений последовательности нуклеотидов каждой мишени, показывает, что статистически-значимые детектируемые однонуклеотидные замены, а также делеции/вставки выявляются уже в единичном количестве и преимущественно именно в месте, где должен располагаться исследуемый полиморфизм.

Заключение

Представленное исследование показывает, что на основе доступной технологии нанопорового секвенирования возможна разработка специализированных таргетных ампликонных панелей, с помощью которых возможно проводить различные диагностические генетические тесты. В нашей экспериментальной панели из 47 мишеней надежно удалось определить 45 – т.е более 95%, из них более 85% надежно детектируются и выравниваются на референсные последовательности более чем на 90%. Полученные результаты указывают на высокую достоверность обнаружения однонуклеотидных полиморфизмов при сравнительно небольшом получаемом объеме данных и затрачиваемом времени на пробоподготовку и секвенирование. В частности в данном исследовании удалось получить приемлемый выход данных в виде числа ридов и оснований уже через 3 часа после запуска процесса секвенирования. Также стоит отметить, что пробоподготовка осуществлялась полностью на отечественной реагентной базе, а секвенирование происходило с использованием отечественного нанопорового секвенатора «Нанопорус Про».

Таким образом, проведенное исследование показывает возможность разработки и внедрения в клиническую практику доступных и бюджетных разнообразных систем детекции мутаций на основе таргетных ампликонных панелей. При этом можно производить секвенирование не на базе NGS, а используя доступную платформу для высокопроизводительного анализа ДНК – нанопоровые секвенаторы. Это позволит сделать доступным проведения подобного рода диагностических исследований для небольших клинико-диагностических лабораторий и в итоге обеспечить большую доступность диагностики для пациентов.

Декларация о конфликте интересов

Авторы заявляют об отсутствии конфликта интересов по представленной статье.

Участие авторов в создании статьи

Ж.Е.С. – проведение исследования, сбор и интерпретация результатов; М.А.М – проведение исследования; Е.А.М. – предоставление материала исследования, интерпретация результатов и осуществление стилистической обработки содержания статьи.

Литература / References

1. Alduaij W, Collinge B, Ben-Neriah S et al. Molecular determinants of clinical outcomes in a real-world diffuse large B-cell lymphoma population. *Blood*. 2023. 141(20). 2493-2507. doi: 10.1182/blood.2022018248
2. Baslan T, Kovaka S, Sedlazeck FJ et al. High resolution copy number inference in cancer using short-molecule nanopore sequencing. *Nucleic Acids Res*. 2021. 49 (21). e124. doi: 10.1093/nar/gkab812
3. Chabon JJ, Hamilton EG, Kurtz DM et al. Integrating genomic features for non-invasive early lung cancer detection. *Nature*. 2020. 580(7802). 245–251. doi: 10.1038/s41586-020-2140-0
4. Cohen JD, Li L, Wang Y et al. Detection and localization of surgically resectable cancers with a multi-analyte blood test. *Science*. 2018. 359(6378). 926–930. doi: 10.1126/science.aar3247
5. Dowsett M, Sestak I, Lopez-Knowles E et al. Comparison of PAM50 risk of recurrence score with onco type DX and IHC4 for predicting risk of distant recurrence after endocrine therapy. *J Clin Pathol*. 2013. 31(22). 2783–2790. doi: 10.1200/JCO.2012.46.1558
6. Hall MB, Wick RR, Judd LM et al. Benchmarking reveals superiority of deep learning variant callers on bacterial nanopore sequence data. *Elife*. 2024. 10 (13). RP98300. doi: 10.7554/eLife.98300

7. Magi A, Semeraro R, Mingrino A et al. Nanopore sequencing data analysis: state of the art, applications and challenges. *Brief Bioinform.* 2018. 19(6). 1256-1272. doi: 10.1093/bib/bbx062
8. MacKenzie M, Argyropoulos C. An Introduction to Nanopore Sequencing: Past, Present, and Future Considerations. *Micromachines (Basel).* 2023. 14(2). 459. doi: 10.3390/mi14020459
9. Morin RD, Arthur SE, Hodson DJ. Molecular profiling in diffuse large B-cell lymphoma: why so many types of subtypes? *Br J Haematol.* 2022. 196(4). 814-829. doi: 10.1111/bjh.17811
10. Sanderson ND, Kapel N, Rodger G et al. Comparison of R9.4.1/Kit10 and R10/Kit12 Oxford Nanopore flowcells and chemistries in bacterial genome reconstruction. *Microb Genom.* 2023. 9(1). mgen000910. doi: 10.1099/mgen.0.000910
11. Satam H, Joshi K, Mangrolia U et al. Next-Generation Sequencing Technology: Current Trends and Advancements. *Biology (Basel).* 2023. 12(7). 997. doi: 10.3390/biology12070997
12. Simon R, Roychowdhury S. Implementing personalized cancer genomics in clinical trials. *Nat Rev Drug Discov.* 2013. 12(5). 358-369. doi: 10.1038/nrd3979
13. Steuer C., Ramalingam S. Tumor mutation burden: leading immunotherapy to the era of precision medicine. *J Clin Oncol.* 2018. 36(7). 631-632. doi: 10.1200/JCO.2017.76.8770
14. Tie J, Wang Y, Tomasetti C et al. Circulating tumor DNA analysis detects minimal residual disease and predicts recurrence in patients with stage II colon cancer. *Sci Transl Med.* 2016. 8(346). 346ra92. doi: 10.1126/scitranslmed.aaf6219