

УТВЕРЖДАЮ

Проректор по научной работе
ФГАОУ ВО РНИМУ им. Н.И.
Пирогова Минздрава России



Д.В. Ребриков

« 08 » августа 2023 г.

Руководство пользователя
Программное обеспечение «Mutracker»
по ТУ 58.29.32-001-01963278-2023

Москва, 2023

АННОТАЦИЯ

Настоящий документ является Руководством пользователя Программного обеспечения «Mutracker» по ТУ 58.29.32-001-01963278-2023 (далее по тексту – ПО Mutracker), которое работает с данными последовательностей нуклеиновых кислот анализируемых образцов, полученными на Секвенаторе нуклеиновых кислот ГА-2000 (ПУ № РЗН 2022/17851 от 25.07.2022), Секвенаторе генетическом HELICON® G400 и Секвенаторе генетическом HELICON® G50 (ПУ № РЗН 2023/20825 от 16.08.2023).

Документ содержит сведения о назначении и условиях выполнения ПО Mutracker, а также описание последовательности действий пользователя, обеспечивающих запуск, выполнение и завершение работы с ПО Mutracker.

Перечень национальных стандартов, применяемых к ПО Mutracker, представлен в Приложении А данного руководства.

СОДЕРЖАНИЕ

АННОТАЦИЯ	2
СОДЕРЖАНИЕ	3
1. НАЗНАЧЕНИЕ ПРОГРАММЫ	4
1.1 Функции ПО Mutracker	6
1.2 Алгоритм работы ПО Mutracker	8
1.3 Алгоритм автоматизированных вычислений ПО Mutracker	10
1.4 Роли пользователей	12
1.5 Информационная безопасность	13
2. УСЛОВИЯ ВЫПОЛНЕНИЯ ПРОГРАММЫ	14
2.1 Аппаратное обеспечение	14
2.2 Программное обеспечение	16
2.3 Уровень подготовки пользователя	16
2.4 Рекомендации по защите данных	16
3. ИСПОЛЬЗОВАНИЕ ПРОГРАММЫ	17
3.1 Авторизация в ПО Mutracker	17
3.2 Личный кабинет пользователя	18
3.3 Загрузка файлов секвенирования в ПО и запуск анализа	18
3.4 РАБОТА С РЕЗУЛЬТАТАМИ	21
3.4.1 Результат Оценка качества данных	22
3.4.2 Результат Статистика	24
3.4.3 Получение файла вариантов, отличающих образец от референсного генома человека	29
3.4.4 Результат Отчет по носительству	29
3.4.5 Результат Аннотация	31
3.5 РАБОТА С АННОТАЦИЯМИ	32
3.5.1 Загрузка новой аннотации	32
3.5.2 Поиск аннотации в списке уже загруженных	35
3.5.3 Просмотр вариантов	37
3.5.4 Сортировка вариантов	42
3.5.5 Панели генов	42
3.5.5.1 Добавление новой пользовательской панели	43
3.5.5.2 Просмотр пользовательских панелей	44
3.5.5.3 Просмотр панелей PanelApp	45
3.5.6 Фильтрация вариантов	47
3.5.6.1 Фильтрация по панели	47
3.5.6.2 Фильтрация по содержанию аннотации	50
3.5.7 Выбор вариантов для ввода в заключение	53
3.5.7.1 Внесение вариантов в заметки	54
3.5.7.2 Внесение вариантов в заключение	55
3.5.8 Получение файла заключения по результатам фильтрации и интерпретации вариантов	57
4. АДРЕС ДЛЯ ОБРАЩЕНИЯ	68
ПРИЛОЖЕНИЕ А	69
ПЕРЕЧЕНЬ СОКРАЩЕНИЙ	70

1. НАЗНАЧЕНИЕ ПРОГРАММЫ

ПО Mutracker предназначено для обработки и интерпретации результатов, полученных с использованием Секвенатора нуклеиновых кислот ГА-2000 (ПУ № РЗН 2022/17851 от 25.07.2022), Секвенатора генетического HELICON® G400 и Секвенатора генетического HELICON® G50 (ПУ № РЗН 2023/20825 от 16.08.2023) в рамках секвенирования экзона человека.

Высокопроизводительное секвенирование (MPS или NGS) – это набор разных технологических решений, объединенных одной характеристикой: генетические анализаторы (NGS-секвенаторы), относящиеся к этой категории, могут одновременно определять нуклеотидную последовательность очень большого (от нескольких тысяч до нескольких миллиардов) количества коротких фрагментов геномной ДНК, полученной из исходного образца в пределах любого биоматериала (кровь или ткани человека, соскобы и пробы биоматериала из окружающей среды, археологические образцы, пробы грунта и т.д.). Данные о геномной структуре исследуемого образца используются в научных и клинко-диагностических процедурах в областях клинической генетики человека (наследственные заболевания, онкология, инфекции и антибиотикорезистентность, происхождение, установление родства, как важнейший элемент вспомогательных репродуктивных технологий и т.д.), фундаментальной генетики, молекулярной биологии и других направлениях наук о жизни.

Благодаря стремительному развитию технологий NGS, а также вычислительных возможностей компьютеров в XXI веке, произошла тихая революция: NGS из разряда сугубо фундаментального искусства перешел в практическую науку и стал набором рутинных технологий, которые помогают в улучшении качества жизни людей ежедневно. При этом результатом работы NGS-секвенатора являются так называемые сырые файлы – это данные о прочитанных нуклеотидных последова-

тельность. С сырыми данными необходимо провести ряд последовательных вычислительных процедур, называемых биоинформатическим анализом, чтобы привести их в доступный специалисту формат.

В геноме человека есть более 20 000 функциональных компонентов, называемых генами. Каждый из генов может быть представлен множеством разных вариантов или аллелей, немного отличающихся друг от друга, но при этом в норме общий набор генов всегда стабилен. Разнообразие аллелей проявляется в фенотипе – именно благодаря генетике разные люди так отличаются внешне, а близкие родственники в силу законов наследственности имеют больше общих черт. Почти для всех генов человека в каждой клетке имеется по два варианта гена, унаследованных от мамы и папы.

На сегодня описано более 6000 заболеваний человека, относящихся к категории наследственных, то есть вызываемых комбинацией патогенных аллельных вариантов генов. То есть не все аллели генов полезны, есть так называемые мутантные, которые приводят к разнообразным негативным клиническим проявлениям. Большой проблемой наследственных заболеваний является то, что многие из них имеют очень низкую частоту встречаемости (например, 1:1 000 000) и плохо описаны клинически, а также то, что одинаковые клинические проявления могут иметь разную генетическую природу, что делает невозможным корректную диагностику и дальнейшую терапию без эффективного NGS-анализа.

За счет диплоидности негативный эффект патологических аллелей почти всегда компенсируется второй копией гена от другого родителя, несущей корректно работающий ген.

Важно, что у абсолютно каждый человек является носителем как минимум нескольких десятков разных генов с патогенными аллелями. Но генов у человека много, патогенные аллели «размазаны» между ними всеми, и вероятность исхода, при котором у одного человека от папы и мамы достанется по патогенному аллелю в одном и том же гене, достаточно невелика.

Уровень науки на сегодня технически позволяет проводить преконцепционный скрининг пар еще до зачатия ими детей. Благодаря этому подходу можно определить «пары родителей риска» и в рамках медико-генетического консультирования значительно снизить вероятность рождения у них больных детей.

Сейчас стандартом для диагностики наследственных заболеваний или скрининга их носительства является NGS-анализ либо всех генов человека одновременно (экзом), либо панелей генов в зависимости от предположительного типа заболевания, к которому относится пациент. Основное преимущество экзоста – его универсальность, т.к. этот диагностический подход охватывает все 20 000 генов.

Разработанное нами ПО Mutracker представляет собой комплексное решение, с помощью которого можно на современном уровне автоматизации и алгоритмизации вычислений проводить анализ данных NGS-секвенирования – в первую очередь данных экзоста человека – для решения задач по преконцепционному генетическому скринингу, а также поиск наследственных причин клинической картины для пациентов с вероятно генетической причиной особенностей фенотипа (наследственные заболевания).

Область применения: клинико-диагностические и медико-генетические лаборатории медицинских учреждений.

1.1 Функции ПО Mutracker

ПО Mutracker обеспечивает обработку и интерпретацию продуктов анализа результатов, полученных с использованием Секвенатора нуклеиновых кислот ГА-2000 (РУ № РЗН 2022/17851 от 25.07.2022), Секвенатора генетического HELICON® G400 и Секвенатора генетического HELICON® G50 (РУ № РЗН 2023/20825 от 16.08.2023) в рамках секвенирования экзоста человека, которые позволяют создавать лабораторный отчет по статусу носителя наследственных заболеваний для здоровых пациентов (преконцепционный скрининг) или медицинское заключение по

поиску наследственных причин клинической картины для пациентов с вероятно генетической причиной особенностей фенотипа.

В ПО Mutracker реализованы ключевые функции:

- подгрузка файлов данных высокопроизводительного секвенирования (сырые данные);
- оценка качества данных высокопроизводительного секвенирования;
- генерация файла выравнивания сырых данных на референсный геном;
- оценка качества выравнивания сырых данных на референсный геном;
- генерация файла вариантов, отличающих образец от референсного генома;
- генерация файла аннотации вариантов, отличающих образец от референсного генома. В рамках этого процесса проводится классификация, фильтрация и аннотация списка вариантов, отличающих образец от референсного генома;
- сбор статистических параметров качества секвенирования экзома человека (cov_stats);
- подготовка файла со списком вариантов носительства однозначно патогенных вариантов, связанных с развитием наследственных заболеваний;
- формирование данных для лабораторного отчета по результатам обработки процедуры генерации файла со списком вариантов носительства (carriage_gc);
- система фильтрации файла аннотации вариантов (модуль Varsiever в виде веб-клиента) для проведения поиска наследственных причин клинической картины для пациентов с вероятно генетической причиной особенностей фенотипа в ручном режиме экспертом;
- формирование медицинского заключения по результатам фильтрации файла аннотации вариантов.

Схема формата работы ПО Mutracker и типов предоставляемых результатов представлена на Рисунок 1: ПО представляет собой веб-интерфейс, в рамках которого реализована подгрузка файлов высокопроизводительного секвенирования в формате fastq. По итогам первичной обработки данных ПО может предоставлять файлы с метриками качества, лабораторный отчет по результатам поиска патогенных вариантов носительства, аннотацию вариантов. Аннотация вариантов может быть проанализирована с помощью веб-интерфейса Varsiever для подготовки медицинского заключения по результатам фильтрации вариантов в файле аннотации в рамках поиска наследственных причин клинической картины для пациентов с вероятно генетической причиной особенностей фенотипа.

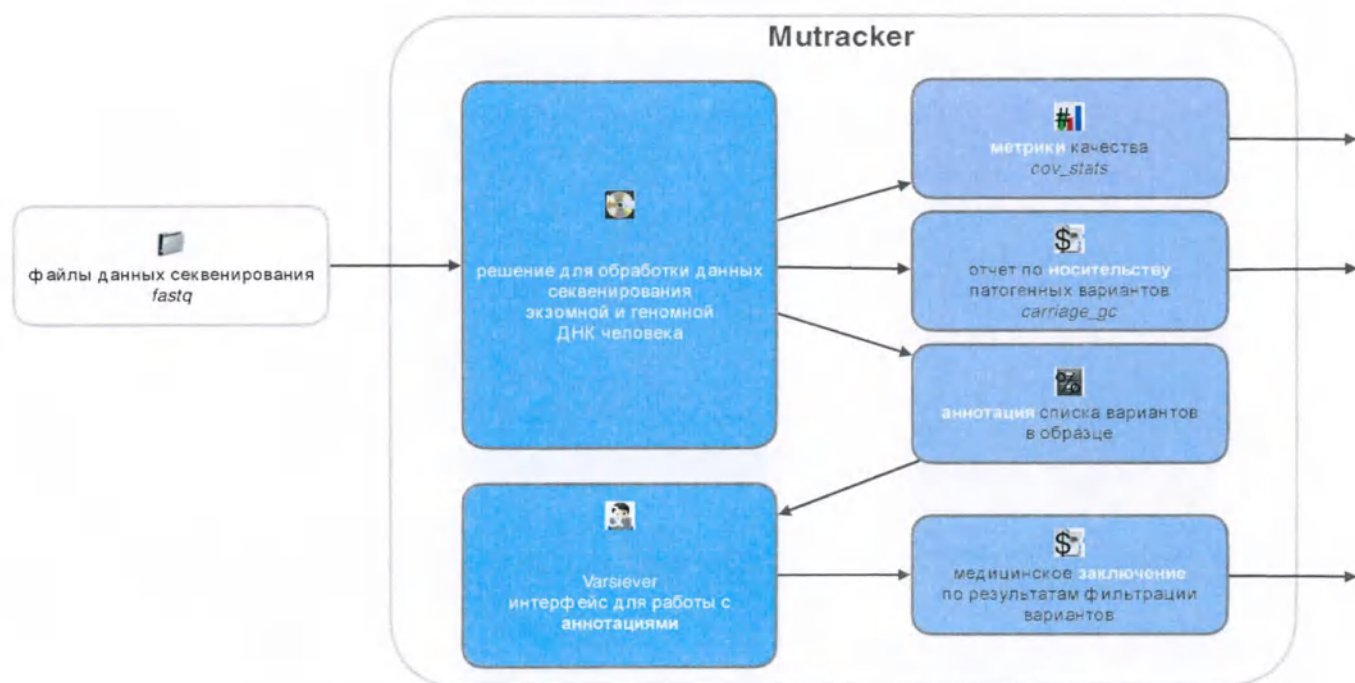


Рисунок 1. Схема формата работы ПО Mutracker.

1.2 Алгоритм работы ПО Mutracker

Алгоритм работы ПО Mutracker представлен на блок-схеме на Рисунок 2.

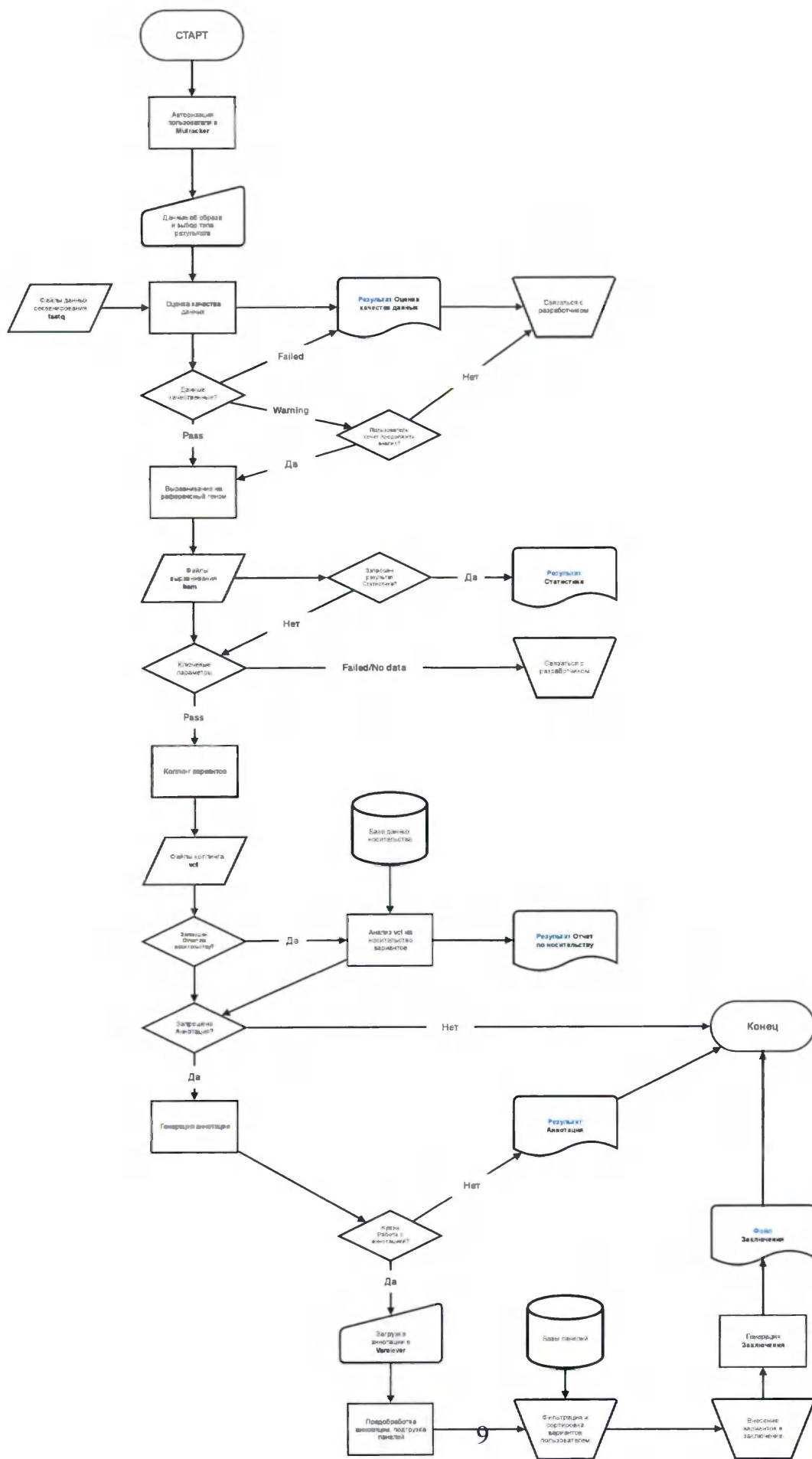


Рисунок 2. Блок-схема работы ПО Mutracker.

1.3 Алгоритм автоматизированных вычислений ПО Mutracker

Важным элементом работы ПО Mutracker является автоматизированная биоинформатическая обработка исходных данных NGS-секвенирования полноэкзомных и полногеномных данных человека, с помощью которой из сырых файлов в формате fastq генерируются необходимые для формирования отчетов и заключений промежуточные файлы. Схема алгоритма автоматизированных вычислений ПО Mutracker представлена на Рисунок 3

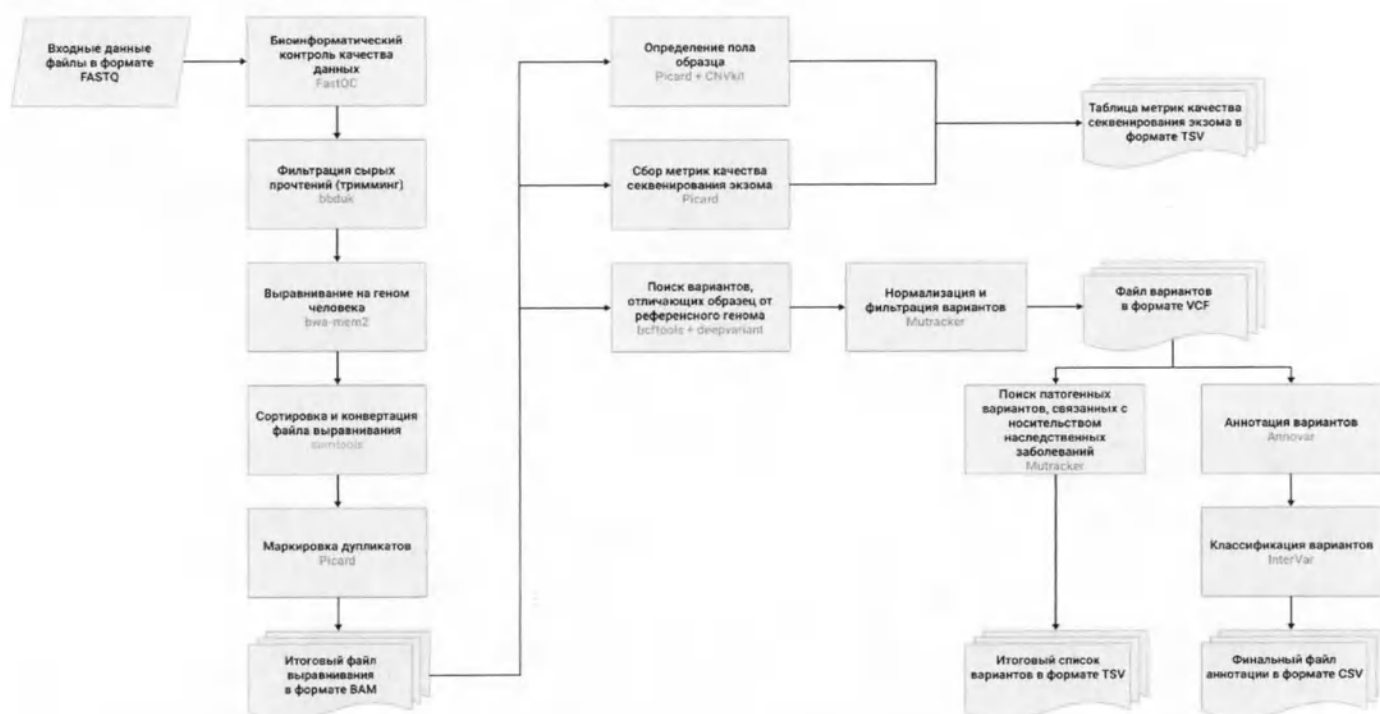


Рисунок 3. Схема алгоритма автоматизированных вычислений ПО Mutracker.

ПО Mutracker принимает на вход данные высокопроизводительного секвенирования в формате FASTQ-файлов.

Биоинформатический контроль качества данных проводится с использованием программы FastQC v0.11.9¹, которая оценивает длину и количество прочтений, среднее значение качества нуклеотидов, GC-состав, содержание адаптеров и дубликатов.

На основании результатов контроля качества проводится коррекция сырых прочтений с использованием ПО bbdduk v38.96².

Отфильтрованные и очищенные прочтения картируются на референсный геном человека GRCh38 с помощью ПО bwa-mem2 v2.2.1³, информация о картировании всех прочтений записывается в файл выравнивания (SAM-файл). Полученный SAM-файл конвертируется в бинарный формат (BAM-файл) и сортируется с помощью программы SAMtools v1.9⁴, после чего передается в ПО Picard v2.22.4⁵ для процедуры маркировки дубликатов.

Для оценки качества картирования прочтений на геном человека, а также для оценки различных параметров, отражающих качество исходного образца и подготовки библиотеки, используется ПО Picard v2.22.4. ПО Picard v2.22.4 также используется вместе с программой CNVkit⁶ для определения пола образца согласно результатам секвенирования.

ПО Mutracker использует две программы для поиска вариантов в полученном ранее файле выравнивания (BAM-файле) - ПО bcftools v1.9⁷ и Deepvariant v1.5.0⁸. Эти программы осуществляют поиск вариантов, отличающих анализируемый образец от референсного генома человека и записывают полученные варианты в файл формата VCF. Далее каждый VCF-файл проходит процедуры нормализации и фильтрации по bed-файлу, содержащему координаты регионов обогащения, для

¹ <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>

² <https://sourceforge.net/projects/bbmap/>

³ <https://github.com/bwa-mem2/bwa-mem2>

⁴ <https://github.com/samtools/samtools>

⁵ <https://broadinstitute.github.io/picard/>

⁶ <https://cnvkit.readthedocs.io/en/stable/index.html>

⁷ <https://github.com/samtools/bcftools>

⁸ <https://github.com/google/deepvariant>

фильтрации вариантов, вызванных ошибочным выравниванием прочтений на геном. После этого ПО Mutracker объединяет полученные варианты в один финальный VCF-файл, который будет использоваться в финальных шагах анализа.

Далее ПО Mutracker использует ПО AnnoVar⁹ и ПО Intervar v2.2.2¹⁰ для фильтрации, аннотации и классификации вариантов, после чего ПО Mutracker объединяет необходимые поля из vcf-файла и отдельные поля из полученной аннотации, формируя финальный файл аннотации.

Также ПО Mutracker определяет список однозначно патогенных вариантов, связанных с развитием наследственных заболеваний, в VCF-файле образца. Обнаруженные варианты записываются в файл формата tsv, содержащий в себе основные данные о варианте - его положение в геноме человека, транскрипт, кодировка HGVS, частота варианта согласно популяционной базе данных gnomAD (v3.1.1), глубина покрытия, а также информация о наследственном заболевании - название на русском и английском языках, коды по OMIM/Orphanet/МКБ-10, возраст дебюта и механизм наследования.

1.4 Роли пользователей

В ПО Mutracker реализовано разделение доступа к функциям в зависимости от роли пользователя:

- «Биоинформатик» – может проводить запуск всех функций ПО Mutracker, кроме формирования медицинского заключения по результатам фильтрации файла аннотации вариантов;
- «Врач-генетик» – может получать результаты оценки качества данных высокопроизводительного секвенирования и сбора статистических параметров качества секвенирования экзона человека и имеет доступ к

⁹ <https://annovar.openbioinformatics.org/en/latest/>

¹⁰ <https://github.com/WGLab/InterVar>

формированию медицинского заключения по результатам фильтрации файла аннотации вариантов;

- «Администратор» – имеет доступ ко всем функциям ПО Mutracker, может вносить изменения в процессы исполнения функций (обновление баз данных, версий основного и вспомогательного ПО и т.д.), может загружать файлы с данными высокопроизводительного секвенирования, менять настройки системы, обновлять базы данных, управлять доступом к системе и т.д.

Роль пользователя определяется Администратором со стороны компании-разработчика при регистрации нового пользователя.

1.5 Информационная безопасность

В ПО Mutracker не осуществляется хранение и обработка персональных данных, таким образом к этой программе не применимы нормы и требования по защите персональных данных.

В ПО Mutracker не предусмотрены встроенные средства защиты от несанкционированного доступа и от копирования. Защита от несанкционированного доступа осуществляется стандартными средствами операционной системы, а именно настройкой прав пользователей на доступ к ПО. Доступ к системе осуществляется по защищённому протоколу HTTPS. Аутентификация пользователей происходит по протоколу OAuth2.

2. УСЛОВИЯ ВЫПОЛНЕНИЯ ПРОГРАММЫ

ПО Mutracker выполнено по клиент-серверной архитектуре, состоящее из веб-служб обработки запросов к базе данных и вычислительным алгоритмам, установленным на сервере компании-разработчика, и веб-клиентов по формированию и отображению результатов их выполнения на персональном компьютере (ПК) пользователя. Для предоставления пользователям доступа и прав для работы с ПО требуется квалификация системного администратора. Процедура настройки системы выполняется сотрудниками компании разработчика либо квалифицированным специалистом, прошедшим обучение в компании разработчика. Для работы с ПО Mutracker на ПК пользователя необходимо выполнение требований к аппаратному и программному обеспечению, приведенных ниже. Для работы с ПО Mutracker необходимо проведение обучения пользователя, выполняемое силами сотрудников компании-разработчика.

Компания-разработчик несет ответственность за работу ПО Mutracker и ее характеристики только в том случае, если:

- ПО Mutracker используется в соответствии с настоящим руководством по эксплуатации (РЭ);
- Сервисное обслуживание и обновление ПО Mutracker выполняется лицами, уполномоченными компанией-разработчиком и имеющими соответствующую квалификацию и необходимое оборудование.

2.1 Аппаратное обеспечение

Минимальные требования к аппаратному обеспечению приведены в Таблица 1.

Таблица 1. Требования к аппаратному обеспечению рабочего ПК пользователя

Наименование	Требование
Процессор	процессор Intel Core i5 6-го поколения или более мощный
Объем оперативной памяти	Не менее 8 ГБ
Дисковая подсистема	Свободное пространство на жестком диске не менее 1000 МБ
Сетевой контроллер	Fast Ethernet 10/100/1000 Base-T или Wi-Fi (WLAN)
Видеомонитор	ЖК-монитор с рабочим разрешением не ниже 1920x1080 с глубиной цветности не ниже 16 бит
USB-разъем	Требуется наличие свободного порта в компьютере, соответствующего стандарту USB 2.0 High-speed
Клавиатура	Требуется наличие русифицированной клавиатуры, совместимой с ПК пользователя
Мышь	Требуется наличие двухкнопочной мыши

Со стороны серверной части ПО Mutracker минимальные требования к аппаратному обеспечению рабочего сервера компании-разработчика приведены в Таблица 2. За серверную часть ПО отвечает компания-разработчик, в ее задачи входит обеспечение технологических процессов для выполнения всех функций ПО Mutracker у пользователя.

Таблица 2. минимальные требования к аппаратному обеспечению рабочего сервера компании-разработчика.

Наименование	Требование
Процессор	4 процессора 4-го поколения или новее Количество ядер – не менее 10 на процессор Частота – не менее 2.40 Гц
Объем оперативной памяти	Не менее 512 Гб
Дисковая подсистема	Свободное пространство на жестком диске не менее 8 Тб
Сетевое подключение	От 100 Мбит/сек

2.2 Программное обеспечение

Для работы с ПО Mutracker на ПК пользователя достаточно любой операционной системы с поддержкой веб-браузера версии не ниже Microsoft Edge 40.x, Chrome 110, Firefox 68.x, а также наличия интернет-соединения либо локальной корпоративной сети с протоколом IP.

2.3 Уровень подготовки пользователя

Пользователь ПО Mutracker должен обладать следующими необходимыми знаниями и навыками для работы с программой:

- знание предметной области и положений нормативных и руководящих документов;
- практические навыки работы с ПК;
- практические навыки работы с графическим пользовательским интерфейсом ОС.

Перед началом работы с ПО Mutracker необходимо ознакомиться с данным Руководством пользователя.

2.4 Рекомендации по защите данных

Для защиты данных в ПО Mutracker пользователям рекомендуется:

- Сохранять конфиденциальность паролей;
- Избегать записи паролей на бумажных носителях;
- Запрашивать изменение паролей каждый раз, когда появляется любой признак их возможной компрометации;
- Завершать активные сеансы работы путем выхода из системы.

3. ИСПОЛЬЗОВАНИЕ ПРОГРАММЫ

3.1 Авторизация в ПО Mutracker

Откройте браузер, в его адресной строке введите <https://varsiever.genomecenter.ru> и нажмите клавишу Enter. Дождитесь загрузки страницы. Если Вы не авторизованы в системе, то на странице появится форма авторизации следующего вида Рисунок 4.

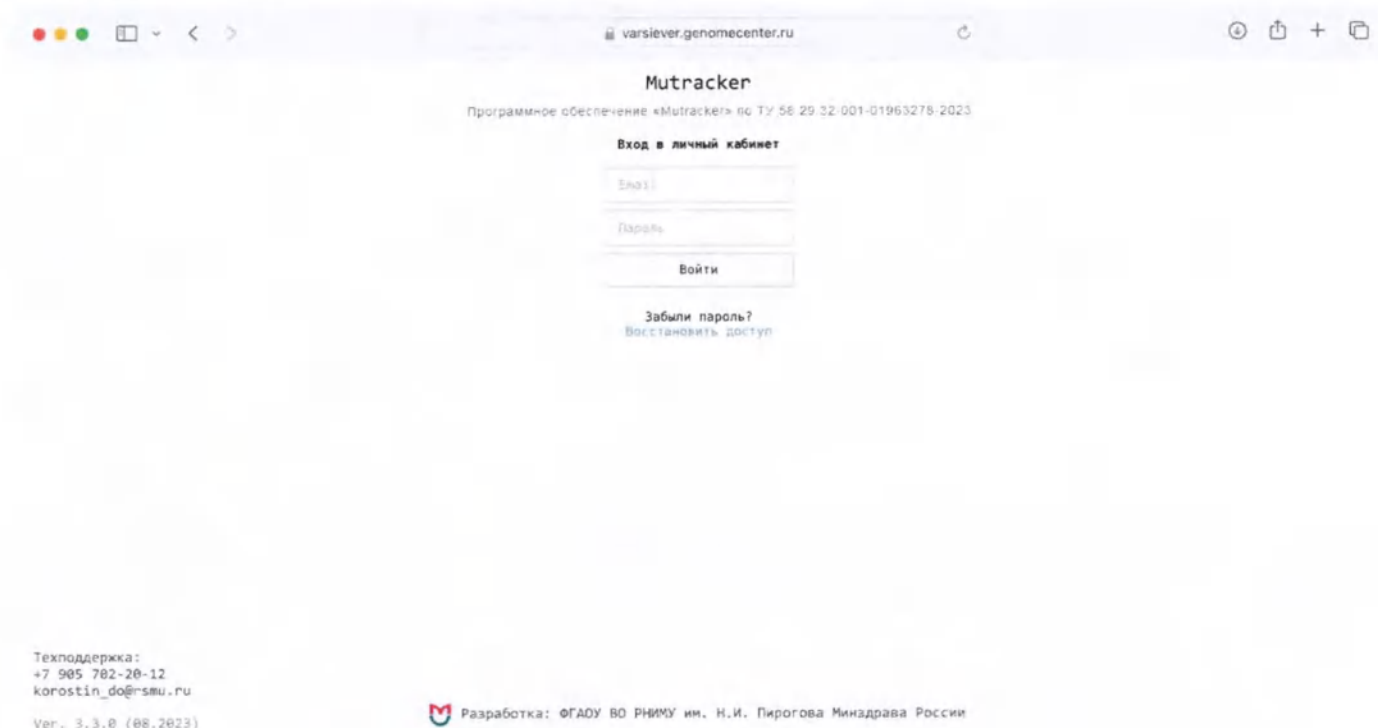


Рисунок 4. Форма авторизации в ПО Mutracker.

Укажите полученные от Администратора данные для авторизации, после чего нажмите кнопку "Войти". Система произведёт попытку авторизации. Дождитесь загрузки данных личного кабинета.

3.2 Личный кабинет пользователя

После успешной авторизации вы попадаете в личный кабинет. Слева расположены меню разделов, в основном окне представлен профиль пользователя. При первом входе нового пользователя необходимо прочитать правила политики обработки персональных данных и принять их, нажав на кнопку *я соглашаюсь с политикой обработки персональных данных и принимаю пользовательское соглашение*.

В меню слева доступны разделы *Образцы*, *Обработка данных*, *Панели*, *Настройки системы*, а также кнопки меню *Профиля* и *Выхода из системы* (деавторизации).

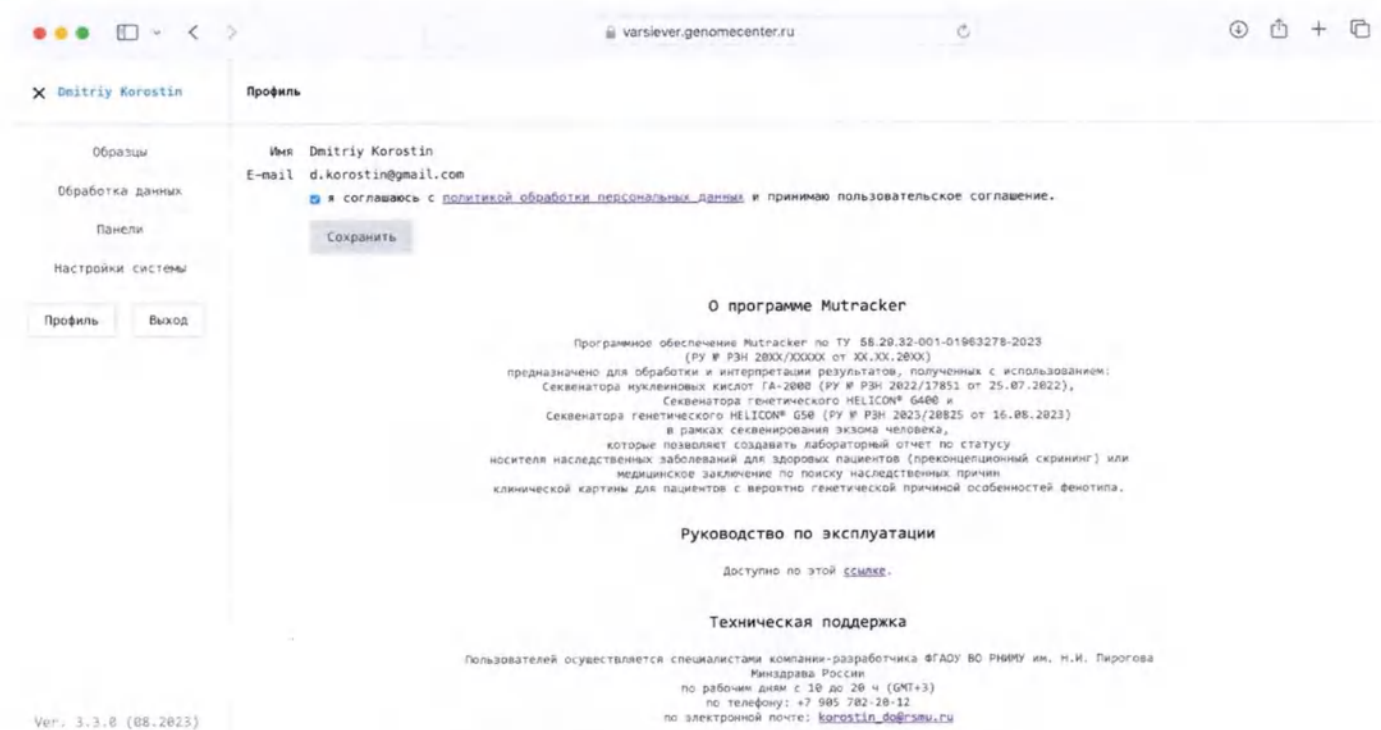


Рисунок 5. Личный кабинет пользователя.

3.3 Загрузка файлов секвенирования в ПО и запуск анализа

Для загрузки файлов результатов высокопроизводительного секвенирования в систему в меню выберите раздел *Обработка данных* (Рисунок 6).

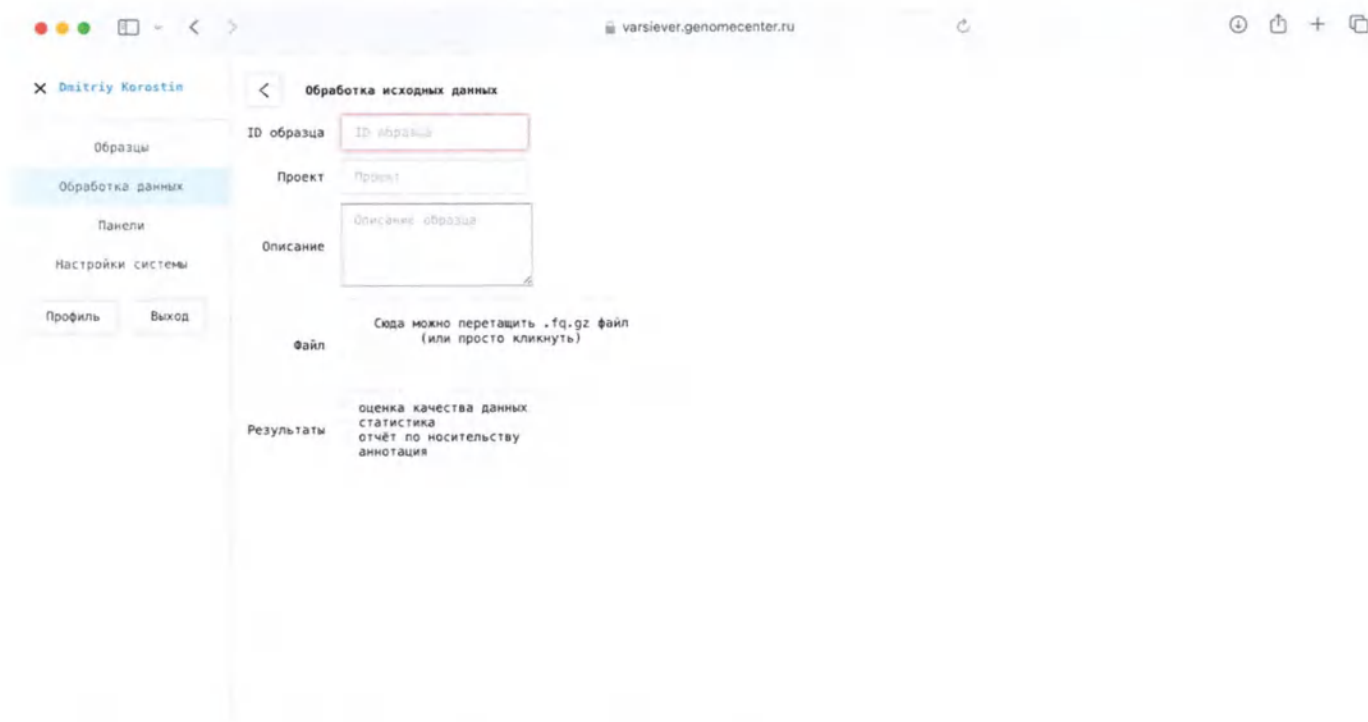


Рисунок 6. Меню выбора раздела *Обработка данных* для загрузки файлов секвенирования.

В открывшейся форме обязательным является поле идентификатор образца (ID образца). Это уникальный номер, который задается пользователем. Если в системе уже имеется образец с введенным ID, ПО выдаст предупреждение с рекомендацией сменить ID.

Опционально вы можете заполнить поля *Проект* и *Описание*. Поле *Проект* позволяет группировать образцы для удобства дальнейшей навигации: при нажатии на него пользователю выдается для выбора список существующих в ПО проектов. Если нужного проекта в списке нет, введите его с помощью клавиатуры. Новый проект будет создан автоматически. Поле *Описание* может содержать любые текстовые метаданные, необходимые для пользователя.

В поле *Файл* можно кликнуть и в открывшемся окне Проводника выбрать необходимый файл для анализа или перетащить его мышкой.

В поле *Результаты* необходимо выбрать, какого типа данные по образцу требуется получить на выходе. Это могут быть:

- **Оценка качества данных.** ПО Mutracker проведет первичную оценку качества сырых данных. Результатом является файл html-репорта FastQC или MultiQC.
- **Статистика.** ПО Mutracker проведет анализ данных до стадии коллинга вариантов и соберет таблицу метрик качества секвенирования образца. Результатом является csv файл «cov_stats».
- **Отчет по носительству.** ПО Mutracker проведет анализ данных до стадии коллинга вариантов и пересечет полученный из образца vcf файл с базой данных Carriage, встроенной в ПО Mutracker. Результатом является tsv файл «carriage», содержащий в себе данные о статусе носителя патогенных вариантов и соответствующих им наследственных заболеваний.
- **Аннотация.** ПО Mutracker проведет анализ исходных данных до финальной стадии – получения аннотации. В рамках этого процесса проводится классификация, фильтрация и аннотация списка вариантов, полученного в результате коллинга вариантов. Результатом является csv файл «annotation».

После выбора как минимум одного из типов результатов появится кнопка *Запустить обработку* (см Рисунок 7).

Выберите необходимые пункты для анализа, а затем нажмите кнопку *Запустить обработку*.

Файл: TMP001.fq.gz

Файл

Результаты

- ☒ оценка качества данных
- ☒ статистика
- ☒ отчёт по носительству
- ☒ аннотация

Запустить обработку

Рисунок 7. Меню выбора типа результатов.

Начнётся загрузка файла. Когда загрузка завершится, система автоматически запустит обработку. После завершения обработки система отправит уведомление на e-mail пользователя (совпадает с e-mail, используемым в качестве логина) о готовности результатов и пришлет файлы результатов в соответствии с выбранными пунктами.

3.4 Работа с Результатами

В зависимости от выбранных пунктов в меню *Результаты* пользователю будут присланы на e-mail от одного до четырех типов результатов в виде соответствующих файлов. Ниже представлено описание их содержимого и варианты дальнейших действий.

3.4.1 Результат Оценка качества данных



Рисунок 8). В меню репорта представлены разделы, содержащие данные о качестве:

- Basic Statistics. Основная статистика.
- Per base sequence quality. Распределения качества секвенирования по нуклеотидно.
- Per sequence quality scores. Распределение качества секвенирования по всем прочтениям (ридах) в исходном файле секвенирования.

- Per base sequence content. Распределение нуклеотидного состава в ридсах.
- Per sequence GC content. Распределение GC-состава в ридсах.
- Per base N content. Частота встречаемости нераспознанных (N) нуклеотидных оснований в ридсах.
- Sequence Length Distribution. Распределение длин ридов в исходном файле секвенирования.
- Sequence Duplication Levels. Измерение уровня дубликатов.
- Overrepresented sequences. Аномально представленные нуклеотидные последовательности и их частота в исходном файле.
- Adapter Content. Доля представленности технических (адаптерных) последовательностей в исходном файле.

Каждый из разделов может быть представлен в одном из трех состояний:

- Pass (зеленый значок). Указывает, что по этому типу проверки исходные данные корректны и пригодны для дальнейшего анализа.
- Warning (желтый значок). Указывает, что по этому типу проверки исходные данные имеют допустимые отклонения от нормальных. Следует внимательно отнестись к причинам их возникновения. Для этого необходимо связаться со специалистом компании-разработчика для получения консультации.
- Failed (красный значок). Указывает, что по этому типу проверки исходные данные имеют недопустимые отклонения от нормальных. **Компания-разработчик ПО Mutracker не гарантирует корректности получаемых на дальнейших этапах результатов!** Для этого необходимо связаться со специалистом компании-разработчика для получения консультации.

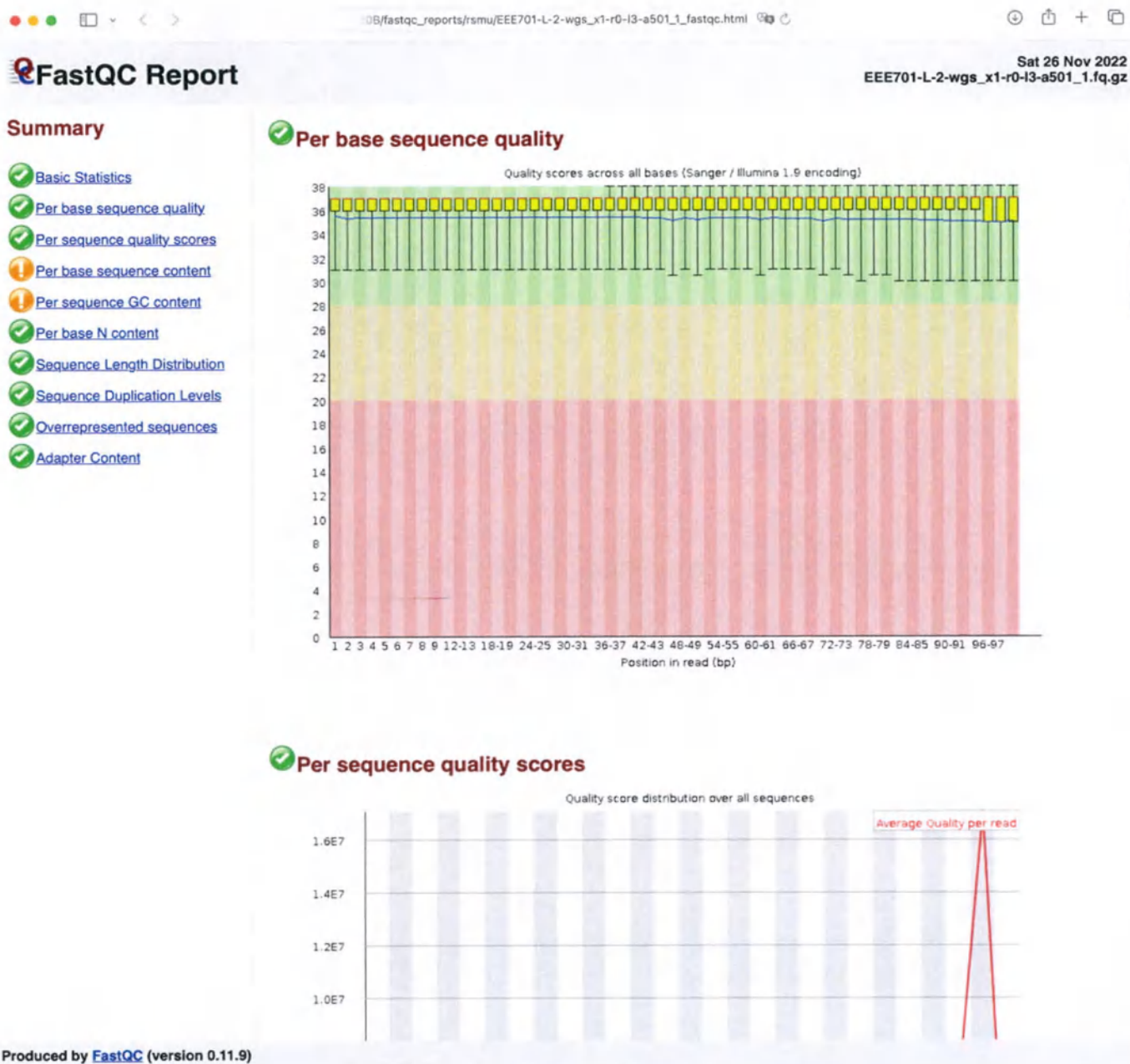


Рисунок 8. Пример html репорта раздела *Оценка качества данных*.

3.4.2 Результат Статистика

С помощью алгоритмов обработки данных в Модуле обработки данных секвенирования экзомной и геномной ДНК человека проводится контроль качества исходных файлов секвенирования; тримминг сырых прочтений по качеству; вырав-

нивание данных на референсный геном человека; сортировка результатов выравнивания; маркировка дубликатов; индексирование выровненных данных. Эти процессы автоматизированы и не требуют участия Пользователя.

На данном этапе анализа образца ПО Mutracker собирает информацию по основным метрикам качества секвенирования экзона (по регионам обогащения), определяет пол образца по покрытию гена *SRY* на Y хромосоме и при помощи программы алгоритма анализа числа копий хромосом, а также рассчитывает уровень гетерозиготности образца по каждой из соматических хромосом. Вся полученная на данном этапе информация совмещается с табличкой анализируемых образцов, содержащих основную, после чего ПО Mutracker составляет cov_stat – файл метрик качества в формате csv, содержащий в себе ключевые параметры качества секвенирования экзона человека. Он включает в себя следующие описанные выше поля – см Таблица 3.

Ключевыми для дальнейшего анализа являются 3 параметра:

- *gender_match*. Определенный биоинформатическим алгоритмом хромосомный пол образца.
- *MeanCovQC*. Средняя глубина прочтения по таргетной области.
- *x10CovQC*. Полнота 10х – доля таргетной области в процентах, имеющая не менее, чем десятикратное покрытие недублицированными ридями.

Эти параметры могут иметь три состояния:

- *Pass*. Прошел контроль качества.
- *Failed*. Не прошел контроль качества. **Необходимо связаться со специалистом компании-разработчика для получения консультации!**
- *No data*. Не удалось пройти контроль качества. **Необходимо связаться со специалистом компании-разработчика для получения консультации!**

Таблица 3. Пример результата раздела *Статистика* – файл метрик качества cov_stats в табличном виде.

Reference	hg37
sample	SVB211
gender_match	pass
MeanCovQC	pass
x10CovQC	pass
fq.gzGB	9,9
lis_lane	1
lis_adaptors	73
duplicates	9,6865
InsertSize	235
ELS	221 596 497
BAIT_SET	v8_Regions
BAIT_TERRITORY	35 241 032
BAIT_DESIGN_EFFICIENCY	1
ON_BAIT_BASES	5 509 072 860
NEAR_BAIT_BASES	4 364 493 411
OFF_BAIT_BASES	1 589 582 828
PCT_SELECTED_BASES	86,1%
PCT_OFF_BAIT	13,9%
ON_BAIT_VS_SELECTED	0,557962
MEAN_BAIT_COVERAGE	156,325526
PCT_USABLE_BASES_ON_BAIT	47,9%
PCT_USABLE_BASES_ON_TARGET	36,4%
FOLD_ENRICHMENT	42,216671
HS_LIBRARY_SIZE	184 763 022
HS_PENALTY_10X	3,625248
HS_PENALTY_20X	3,669062
HS_PENALTY_30X	3,714205
HS_PENALTY_40X	3,764659
HS_PENALTY_50X	3,81405
HS_PENALTY_100X	4,081719
TARGET_TERRITORY	35 241 032
GENOME_SIZE	3 095 693 983
TOTAL_READS	91 917 272
PF_READS	91 917 272
PF_BASES	11 489 659 000

PF_UNIQUE_READS	83 019 986
PF_UQ_READS_ALIGNED	82 955 242
PF_BASES_ALIGNED	11 463 149 099
PF_UQ_BASES_ALIGNED	10 352 877 126
ON_TARGET_BASES	4 178 088 928
PCT_PF_READS	100,0%
PCT_PF_UQ_READS	90,3%
PCT_PF_UQ_READS_ALIGNED	99,9%
MEAN_TARGET_COVERAGE	118,6
MEDIAN_TARGET_COVERAGE	114
MAX_TARGET_COVERAGE	3 164
MIN_TARGET_COVERAGE	0
ZERO_CVG_TARGETS_PCT	1,2%
PCT_EXC_DUPE	9,7%
PCT_EXC_ADAPTER	0,0%
PCT_EXC_MAPQ	3,6%
PCT_EXC_BASEQ	0,8%
PCT_EXC_OVERLAP	8,5%
PCT_EXC_OFF_TARGET	41,0%
FOLD_80_BASE_PENALTY	1,445823
PCT_TARGET_BASES_1X	97,8%
PCT_TARGET_BASES_2X	97,7%
PCT_TARGET_BASES_10X	97,3%
PCT_TARGET_BASES_20X	97,1%
PCT_TARGET_BASES_30X	96,9%
PCT_TARGET_BASES_40X	96,3%
PCT_TARGET_BASES_50X	95,0%
PCT_TARGET_BASES_100X	64,2%
AT_DROPOUT	0,315188
GC_DROPOUT	5,872065
HET_SNP_SENSITIVITY	0,976279
HET_SNP_Q	16
SRY mean bait coverage	308,243047
SRY hs penalty 10x	100596,018
SRY hs penalty 100x	113054,093
SRY on target bases	193375
SRY mean targer coverage	233,827086
SRY median target coverage	200
SRY max target coverage	345

SRY min target coverage	35
OurGender CNVkit	Male
total homo	15 028
total hetero	83 504
sum variants	98 532
het/homo ratio	5,6
homo/het ratio	18,0%
chr1 homo	1 595
chr1 hetero	8 407
chr2 homo	1 036
chr2 hetero	5 881
chr3 homo	785
chr3 hetero	4 710
chr4 homo	630
chr4 hetero	3 241
chr5 homo	560
chr5 hetero	3 523
chr6 homo	853
chr6 hetero	5 056
chr7 homo	616
chr7 hetero	4 452
chr8 homo	559
chr8 hetero	2 753
chr9 homo	647
chr9 hetero	3 367
chr10 homo	653
chr10 hetero	3 665
chr11 homo	1 120
chr11 hetero	5 249
chr12 homo	895
chr12 hetero	4 413
chr13 homo	323
chr13 hetero	1 447
chr14 homo	627
chr14 hetero	2 731
chr15 homo	591
chr15 hetero	2 587
chr16 homo	648
chr16 hetero	3 833

chr17 homo	751
chr17 hetero	4 895
chr18 homo	271
chr18 hetero	1 346
chr19 homo	777
chr19 hetero	6 611
chr20 homo	325
chr20 hetero	2 054
chr21 homo	184
chr21 hetero	1 232
chr22 homo	291
chr22 hetero	2 051

3.4.3 Получение файла вариантов, отличающих образце от референсного генома человека

ПО Mutracker автоматически осуществляет анализ файла выравнивания образца, сравнивая его с референсным геномом человека, и получает на выходе несколько файлов вариантов (vcf-файлы), отличающих образец от референсного генома, полученных разными алгоритмами. Далее каждый vcf-файл проходит процедуры нормализации и фильтрации по bed-файлу обогащения, для фильтрации вариантов, вызванных ошибочным выравниванием ридов на геном. После этого ПО Mutracker объединяет полученные vcf-файлы в один финальный vcf-файл, который будет использоваться в финальных шагах анализа.

3.4.4 Результат Отчет по носительству

ПО Mutracker анализирует объединенный vcf-файл, полученный на предыдущем шаге, на предмет наличия в нем однозначно патогенных вариантов, связанных с развитием наследственных заболеваний, путем сравнения с собственной базой данных. Обнаруженные варианты выписываются в файл формата tsv (см пример Таблица 4), в котором содержатся основные данные о варианте, его положении в

геноме человека, транскрипт, кодировки HGVS, частота варианта согласно популяционной базе данных gnomAD (v2.1.1), глубина покрытия, а также информация о наследственном заболевании - название на русском и английском языках, коды по OMIM/Orphanet/МКБ-10, возраст дебюта и механизм наследования.

Таблица 4. Пример результата раздела *Отчет по носительству* для образца с обнаруженным носительством семейной гиперхолестеринемии.

Chrom:Position	chr1:25870254
Ref/Alt	G/A
Gene	LDLRAP1
HGVSc	c.71del; c.71del c.65G>A; c.65G>A;
HGVSp	p.Gly24AlafsTer32; p.Gly24AlafsTer32; p.Trp22Ter;p.Trp22Ter;
Exon/Intron	EXON=1/9;EXON=1/9;EXON=1/9;EXON=1/9;
Transcript_c	ENST00000374338.4;NM_015627.3; ENST00000374338.4;NM_015627.3;
gnomAD_freq	0.0
Disease.1	HYPERCHOLESTEROLEMIA, FAMILIAL, 4; FHCL4.
Russian version	Гомозиготная семейная гиперхолестеринемия
OMIM.1/ORPHA.1/ICD-10	603813/391665/E78.0
Inheritance mode	AR
Age of Onset	Неонатальный
DP	32
Caller	bcftools

Файл *carriage* содержит следующие поля:

- *Chrom:Position*. Локализация патогенного варианта на хромосоме.

- *Ref/Alt*. Генотип референсного и альтернативного аллелей.
- *Gene*. Название гена, в котором обнаружен патогенный вариант.
- *HGVSc*. Обозначение типа замены в РНК по HGVS.
- *HGVSp*. Обозначение типа замены в белке по HGVS.
- *Exon/Intron*. Локализация варианта. Если это экзон, то порядковый номер экзона в гене.
- *Transcript_c*. Идентификатор транскрипта по UCSC и NCBI.
- *gnomAD_freq*. Частота альтернативного аллеля по БД gnomAD.
- *Disease.1*. Заболевание на английском языке.
- *Russian version*. Заболевание на русском языке.
- *OMIM.1/ORPHA.1/ICD-10*. Идентификатор заболевания в базах данных OMIM, Orphanet и МКБ-10.
- *Inheritance mode*. Тип наследования для заболевания.
- *Age of Onset*. Возраст манифестации заболевания.
- *DP*. Глубина покрытия варианта.
- *Caller*. Используемые для анализа коллеры.

Если по результатам анализа носительства у образца не обнаружится пересечений с БД Carriage, то файл на выходе будет содержать только хедер таблицы с прочерками в ячейках значений.

3.4.5 Результат Аннотация

При помощи ПО Mutracker происходит аннотации вариантов из объединенного vcf-файла образца. В рамках этого процесса проводится классификация, фильтрация и аннотация списка вариантов, отличающих образец от референсного генома. Затем ПО Mutracker объединяет необходимые поля из объединенного vcf-файла и отдельные поля из полученной аннотации, формируя финальный файл аннотации.

Файл аннотации представляет собой таблицу в формате csv, содержащую 60-70 колонок и 24 000 – 120 000 строк. Объем файла от 10 до 60 Мб.

3.5 Работа с аннотациями

ПО Mutracker в виде модуля Varsiever предоставляет широкий функционал для работы с файлами аннотаций, позволяя оптимизировать работу по интерпретации результатов полноэкзомного секвенирования специалистами за счет интегрированных в ПО генетических баз данных, а также удобной системы последовательных фильтров данных в файле аннотации, разработанной с учетом логики медико-генетического анализа результатов высокопроизводительного секвенирования геномных данных человека.

К модулю Varsiever относятся разделы Образцы и Панели в личном кабинете пользователя (см. Рисунок 5).

3.5.1 Загрузка новой аннотации

В личном кабинете в меню слева выберите раздел *Образцы*. Если у вас есть права на загрузку образцов, то в верхней части открывшейся страницы будет присутствовать раздел *Новые образцы* (см. Рисунок 9).

Нажмите кнопку *Добавить* (Рисунок 10), после этого на открывшейся странице Добавления образца в подразделе *Файл* кликните на поле для файла и в открывшемся окне выберите нужный файл. Допустимые типы файлов: .csv, .tsv и .txt. Можно также просто переместить файл в данное поле при помощи мыши (Рисунок 11).

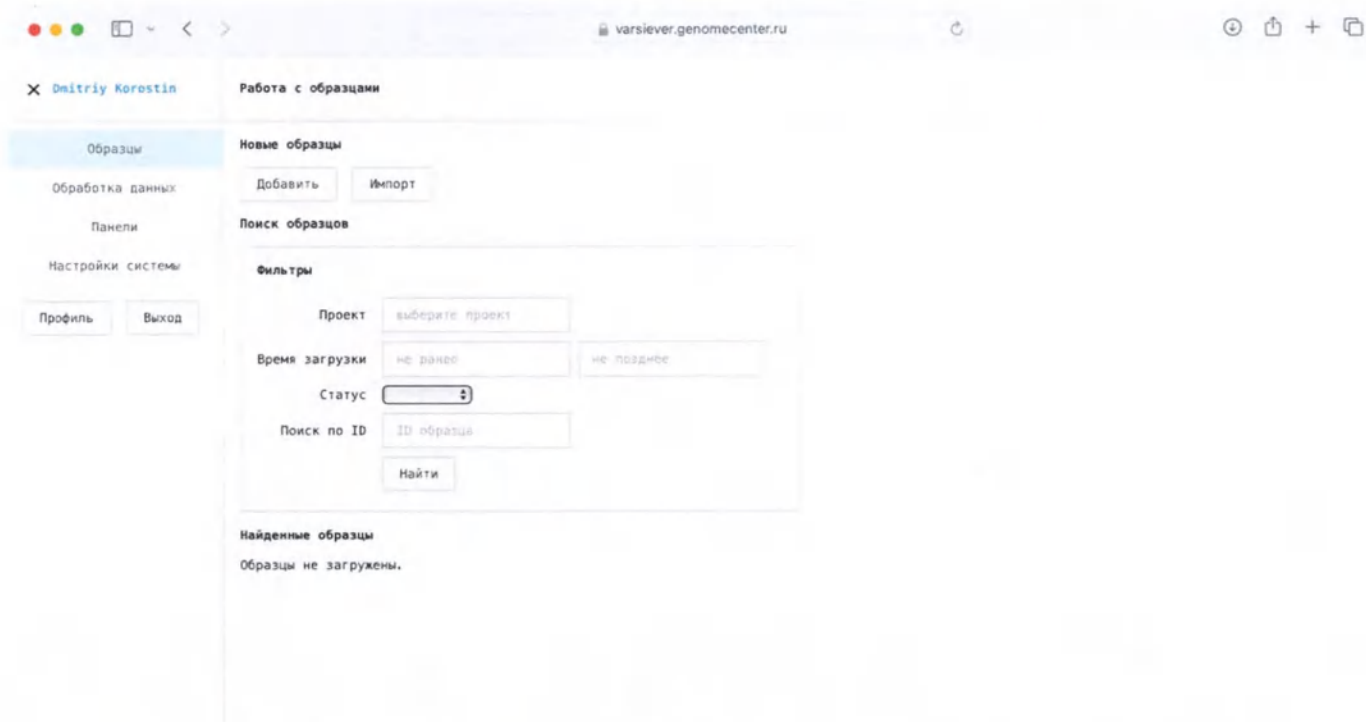


Рисунок 9. Меню добавления новой аннотации и поиска уже загруженных аннотаций в модуле Varsiever.

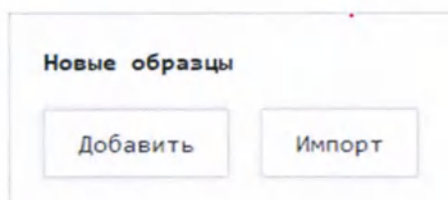


Рисунок 10. Форма ввода новых образцов в Varsiever.



Рисунок 11. Окно добавления новой аннотации.

На странице добавления образца аннотации для заполнения доступны поля:

- *Внутренний ID.* Идентификатор аннотации, присвоенный пользователем.
- *Внешний ID.* Идентификатор аннотации, который пользователь может связать с внешними источниками (например, номер ЭМК).
- *Версия сборки.* Версия сборки референсного генома (hg19 или hg38).
- *Зонды для обогащения.* Использованный в мокрой лаборатории набор для полноэкзомного обогащения.
- *Проект.* Поле *Проект* позволяет группировать образцы для удобства дальнейшей навигации: при нажатии на него пользователю выдается для выбора список существующих в ПО проектов. Если нужного проекта в списке нет, введите его с помощью клавиатуры. Новый проект будет создан автоматически.

- *Описание*. Поле *Описание* может содержать любые текстовые метаданные, необходимые для пользователя.

Если название файла аннотации соответствует формату наименования образцов, то поле *Внутренний ID* заполнится автоматически. Для загрузки аннотации необходимо, чтобы поля *Внутренний ID*, *Проект*, *Версия сборки*, *Зонды для обогащения* были заполнены.

После того, как файл выбран и необходимые поля заполнены, ниже поля для выбора файла появится кнопка *Загрузить*. Нажмите её и дождитесь ответа системы. Если все необходимые данные указаны верно, система выдаст сообщение *Образец добавлен в очередь обработки* и начнёт обработку файла. Так как процесс может занимать некоторое время, после завершения процесса на адрес электронной почты, указанный при авторизации в системе, придёт письмо с уведомлением о результатах обработки.

3.5.2 Поиск аннотации в списке уже загруженных

Для поиска нужного образца аннотации, в меню слева выберите раздел *Образцы*. После этого в окне справа в блоке *Фильтры* укажите необходимые параметры искомых образцов (проект; когда образцы были загружены в систему; текущий статус) и нажмите кнопку *Найти* (см Рисунок 9). Ниже, в разделе *Найденные образцы* отобразится таблица с образцами, удовлетворяющими условиям поиска (см пример поиска на Рисунок 12).

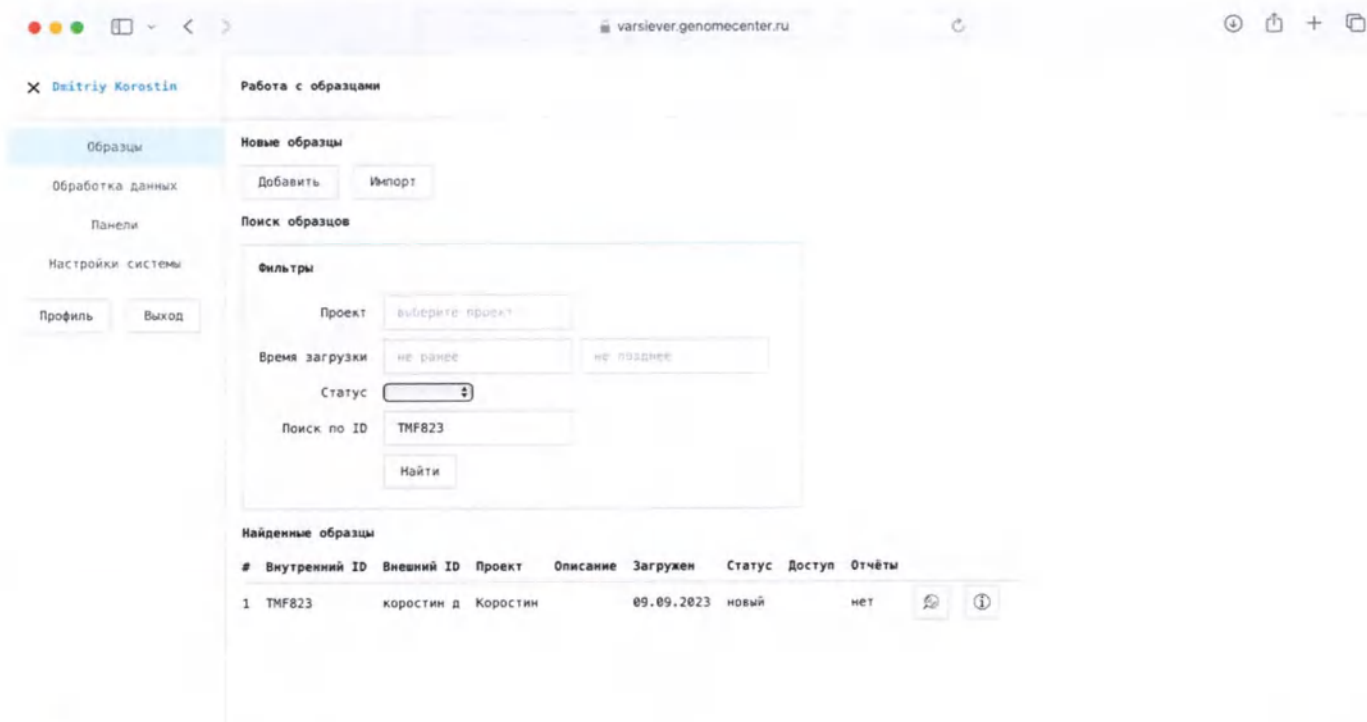


Рисунок 12. Меню поиска уже загруженных аннотаций по полю *Поиск по ID* на примере TMF823.

В самой правой колонке таблицы напротив каждого образца расположены кнопки (Рисунок 13) просмотра описания образца и просмотра аннотации.

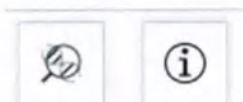


Рисунок 13. Кнопки просмотра описания образца и просмотра аннотации.

Для просмотра описания образца нажмите правую кнопку (с буквой *i*, вписанной в круг), для просмотра аннотации (с изображением ДНК под лупой) — левую кнопку. Если нажать на просмотр описания образца, откроется страница описания аннотации (см Рисунок 14). Нажав на кнопку *Карандаша* вверху, можно отредактировать содержимое полей описания образца.



Рисунок 14. Меню просмотра описания аннотации образца.

Количество строк, указанное в поле Аннотация, равно количеству обнаруженных в аннотации образца вариантов. Из меню поиска уже загруженных аннотаций или из меню *Описания образца* можно перейти в раздел работы с аннотациями (просмотр вариантов, фильтрация вариантов), нажав на кнопку с изображением ДНК под лупой. Для просмотра аннотации из раздела *Описание образца* необходимо нажать эту кнопку в строке *Аннотация*.

3.5.3 Просмотр вариантов

Список вариантов из аннотации представлен в виде таблицы в нижней части экрана раздела *Аннотация* (см Рисунок 15). Все варианты разбиты на страницы по 10 вариантов на каждой, таким образом в таблице может быть отображено не более 10 строк (Рисунок 16).

varsiever.genomecenter.ru

X Dmitry Korostin

Образцы

Обработка данных

Панели

Настройки системы

Профиль Выход

Панели генов

Поиск по названию

Новый фильтр

Сохранённые фильтры

строка	хромосома	начало	конец	ref	alt	зиготность	покрытие	ген	локализация варианта	тип замены в экзоне	пат
1	1	249141454	249141454	C	T	hetero	47	ZNF672	UTR5		
2	1	152192978	152192978	T	C	homo	48	HRNR	exonic	nonsynonymous SNV	
3	1	152193052	152193052	A	G	homo	70	HRNR	exonic	synonymous SNV	
4	1	47013364	47013364	T	C	homo	46	MKNK1-AS1	ncRNA intronic		
5	1	47016917	47016917	C	T	hetero	39	MKNK1-AS1	ncRNA intronic		
6	1	47025840	47025840	C	-	hetero	47	MKNK1-AS1	ncRNA intronic		
7	1	152193121	152193121	G	T	homo	80	HRNR	exonic	synonymous SNV	
8	1	152192825	152192825	C	T	homo	46	HRNR	exonic	nonsynonymous SNV	
9	1	152193286	152193286	G	T	homo	92	HRNR	exonic	nonsynonymous SNV	
10	1	152193544	152193544	A	G	homo	102	HRNR	exonic	synonymous SNV	

Сохранить в файл

Рисунок 15. Экран раздела *Аннотация*.

☐	строка	хромосома	начало	конец	ref	alt	зиготность	покрытие	ген	локализация варианта
☐	1	1	249141454	249141454	C	T	homo	56	ZNF672	UTR5
☐	2	1	152681680	152681680	-	AGCTCTGGGGCTGCTGT	hetero	112	LCE4A	exonic
☐	3	1	152681694	152681694	-	CTGTAGCTCTGGGGCCTC	hetero	155	LCE4A	exonic
☐	4	1	152691943	152691943	G	A	hetero	36	C10RF68	upstream
☐	5	1	152692074	152692074	G	C	hetero	182	C10RF68	exonic
☐	6	1	152692472	152692472	A	C	hetero	180	C10RF68	exonic
☐	7	1	152692588	152692588	A	G	hetero	131	C10RF68	exonic
☐	8	1	152681543	152681543	T	C	homo	123	LCE4A	UTR5
☐	9	1	152692781	152692781	G	T	hetero	104	C10RF68	UTR3
☐	10	1	152732436	152732436	G	C	hetero	141	KPRP	exonic

Сохранить в файл

Рисунок 16. Окно списка вариантов аннотации.

Список страниц и навигация по ним расположены сразу после таблицы. Так, для просмотра вариантов с 71 по 80 нужно выбрать страницу 8. Кнопки < и > позволяют перейти к предыдущему и следующему десятку страниц, соответственно. Кнопки << и >> перемещают указатель по 100 страниц.

В таблице аннотации вариантов представлены метаданные по вариантам:

- *Строка*. Техническая колонка.
- *Хромосома*. Хромосома, в которой расположен вариант.
- *Начало*. Начало варианта относительно референсного генома.
- *Конец*. Конец варианта относительно референсного генома.
- *Ref*. Генотип референсного генома в этом диапазоне координат.
- *Alt*. Генотип варианта в образце в этом диапазоне координат.
- *Зиготность*. Зиготность варианта (гомо- или гетерозигота).
- *Покрывтие*. Усредненное количество ридов в этом диапазоне координат.
 - **Раскрывающийся список**: количество ридов согласно результатам работы отдельного коллера (bcftools, deepvariant, strelka и т.д.).
- *Ген*. Название гена, в котором расположен вариант.
 - **Раскрывающийся список**: расшифровка названия гена, альтернативные названия гена.
- *Локализация варианта*. В какой части гена расположен вариант (экзон, интрон, промотор, межгенное пространство и т.д.).
- *Тип замены в экзоне*. Если вариант локализован в экзоне, то тип вызываемого им варианта (миссенс, нонсенс, синонимичная, сдвиг рамки считывания и т.д.).
- *Патогенность*. Классификация варианта по категории патогенности согласно ACMG.
 - **Раскрывающийся список**: ссылки на внешние ресурсы с подробной информацией о варианте: OMIM, varsome, Ensembl, NCBI, HGNC (Рисунок 17).

- *gnomAD*. Усредненная частота варианта по БД • *gnomAD*.
 - **Раскрывающийся список:** частоты варианта по различным популяциям (африканская, восточноазиатская, европейская без скандинавов, коренная американская, евреи-ашкеназы, скандинавская, другие популяции) – см Рисунок 18.
- *Аминокислотная замена*. Кодировка варианта по HGVS в геномных, транскриптомных и белковых координатах. Если записей более одной, это означает, что представлены кодировки для мажорных представителей альтернативного сплайсинга в этом гене.

Для просмотра значений в раскрывающихся списках необходимо кликнуть по зеленому кружку в соответствующей ячейке. При необходимости требуется горизонтальная прокрутка окна аннотации для просмотра всех колонок значений.

Образцы

Обработка данных

Панели

Настройки системы

Профиль

Выход

Применить фильтры

покрытие	ген	локализация варианта	тип замены в экзоне	патогенность	gnomAD	аминокисл
51	GPR88	exonic	frameshift deletion			GPR88:N
43	FCN3 название: ficolin 3	exonic	frameshift deletion	pathogenic OMIM: 604973 varsome: chr1:27699671:G:- Ensembl gene: ENSG00000142748 ClinVar: 20324 NCBI: 8547 HGNC: 3625	0.0163	FCN3:NM FCN3:NM
120	MST1L	exonic	frameshift deletion		0.4837	MST1L:N
244	PDE4DIP	exonic	frameshift deletion		0.3463	PDE4DIP 105 PDE4DIP PDE4DIP 43 PDE4DIP 105 PDE4DIP 105 PDE4DIP 43
246	PDE4DIP	exonic	frameshift deletion		0.3245	PDE4DIP PDE4DIP PDE4DIP PDE4DIP PDE4DIP PDE4DIP PDE4DIP PDE4DIP PDE4DIP
160	OR2T7	exonic	frameshift deletion		0.2833	OR2T7:N
93	OR2B11	exonic	frameshift deletion		0.4025	OR2B11:
153	OR2T12	exonic	frameshift deletion		0.1613	OR2T12:
91	ACTN2	exonic	frameshift deletion		0.6422	A

Рисунок 17. Раскрывающийся список по значению *патогенность*.

Применить фильтры

Образцы

Обработка данных

Панели

Настройки системы

Профиль

Выход

локализация варианта	тип замены в экзоне	патогенность	gnomAD	аминокислотная замена
exonic	frameshift deletion			GPR88:NM_022049:exon2:c.329_341del:p.S1
exonic	frameshift deletion	pathogenic OMIM: 604973 varsome: chr1:27699671:G:- Ensembl gene: ENSG00000142748 ClinVar: 20324 NCBI: 8547 HGNC: HGNC:3625	0.0163 gnomAD AFR: 0.0267 gnomAD EAS: 0 gnomAD NFE: 0.015 gnomAD AMR: 0.0107 gnomAD ASJ: 0.0232 gnomAD FIN: 0.0063 gnomAD OTH: 0.0092	FCN3:NM_173452:exon4:c.316delC:p.L106Sf FCN3:NM_003665:exon5:c.349delC:p.L117Sf
exonic	frameshift deletion		0.4837	MST1L:NM_001271733:exon2:c.119_123del:p
exonic	frameshift deletion		0.3463	PDE4DIP:NM_001350520:exon27:c.5480delA:105 PDE4DIP:NM_001198834:exon31:c.4994delA:105 PDE4DIP:NM_014644:exon31:c.4994delA:p.Q PDE4DIP:NM_001198832:exon34:c.4862delA:43 PDE4DIP:NM_001350521:exon34:c.5405delA:105 PDE4DIP:NM_001350522:exon34:c.5192delA:105 PDE4DIP:NM_001350523:exon34:c.4859delA:43
exonic	frameshift deletion		0.3245	PDE4DIP:NM_001002811:exon7:c.1947delT:p PDE4DIP:NM_001350520:exon7:c.1947delT:p PDE4DIP:NM_001002812:exon11:c.1458delT:p PDE4DIP:NM_001198834:exon11:c.1458delT:p PDE4DIP:NM_014644:exon11:c.1458delT:p.A PDE4DIP:NM_001198832:exon14:c.1656delT:p PDE4DIP:NM_001350521:exon14:c.1869delT:p PDE4DIP:NM_001350522:exon14:c.1656delT:p PDE4DIP:NM_001350523:exon14:c.1656delT:p
exonic	frameshift deletion		0.2833	OR2T7:NM_001355223:exon1:c.38_39del:p.I
exonic	frameshift deletion		0.4025	OR2B11:NM_001004492:exon1:c.23delT:p.F8
exonic	frameshift deletion		0.1613	OR2T12:NM_001004692:exon1:c.5delA:G
exonic	frameshift deletion		0.6422	ACTN2:NM_001278344:exon10:c.91delC:p.P3

Рисунок 18. Раскрывающийся список по значению *gnomAD*.

Модуль Varsiever позволяет осуществлять фильтрацию и сортировку вариантов. В таблице отображаются только те варианты, которые удовлетворяют применённым фильтрам, и в порядке, соответствующем применённой сортировке.

3.5.4 Сортировка вариантов

Для сортировки вариантов необходимо кликнуть на название столбца, по которому требуется отсортировать варианты. Кликая несколько раз на один и тот же столбец, можно сортировать варианты по возрастанию или убыванию значения. Последовательная сортировка по нескольким столбцам осуществляется последовательными кликами по этим столбцам. Так, например, если необходимо отсортировать варианты по хромосомам и положению в них, достаточно сначала кликнуть на столбец *начало*, а потом – на столбец *хромосома*.

3.5.5 Панели генов

В ПО Mutracker встроен раздел *Панели*, доступ к которому есть из личного кабинета пользователя (см. Рисунок 5). Этот раздел содержит два типа панелей генов, используемых для фильтрации результатов аннотации: *Пользовательские панели* и панели *PanelApp* (Рисунок 19).

При переходе в раздел *Панели* пользователю отображается список введенных им пользовательских панелей. До введения пользователем панелей этот список пуст, присутствует только кнопка *Добавить*. Если пользователем уже введены панели, но они отобразятся в виде списка (см. Рисунок 19). В конце списка также будет находиться кнопка *Добавить*.

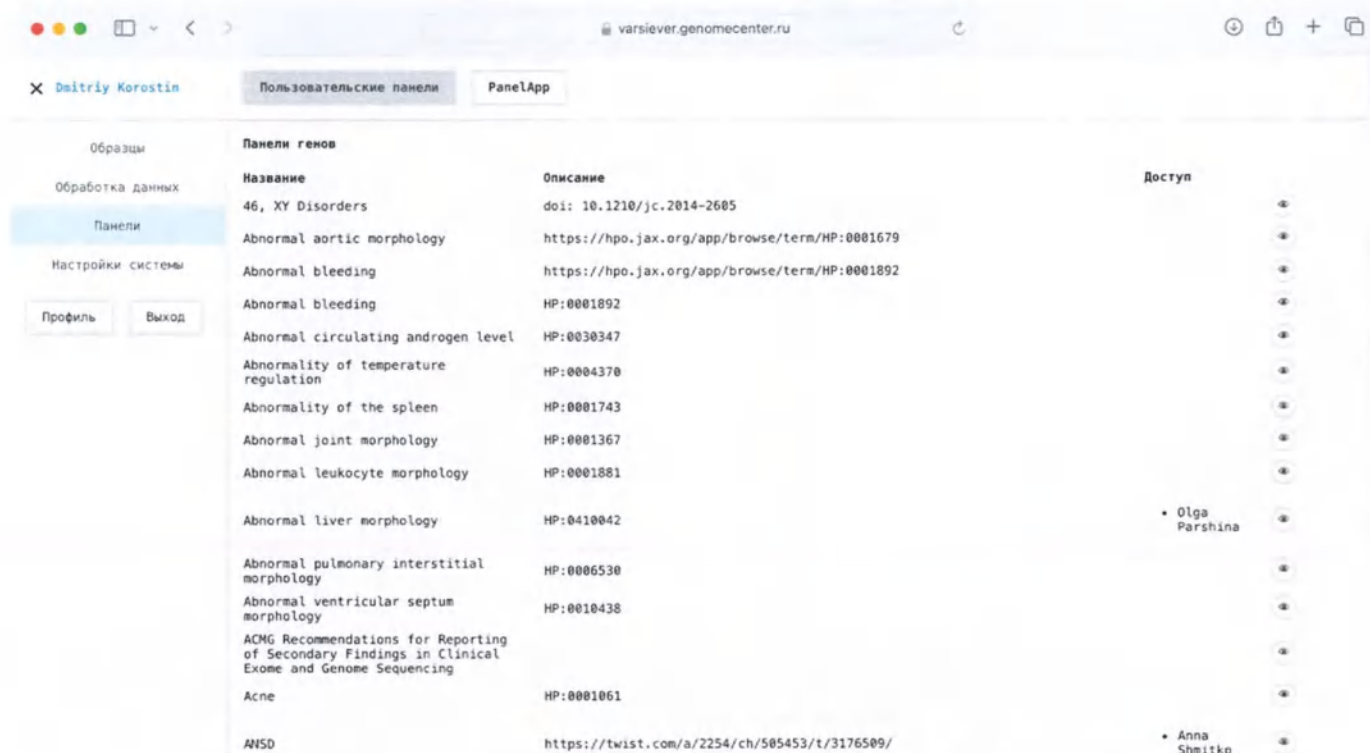


Рисунок 19. Меню *пользовательских панелей*.

3.5.5.1 Добавление новой пользовательской панели

Для добавления новой панели необходимо нажать на кнопку *Добавить* в конце списка пользовательских панелей. При этом отобразится меню добавления панели (Рисунок 20).

В меню доступны для заполнения поля:

- *Название*. Для введения названия панели.
- *Сохранить*. Кнопка сохранения новой панели.
- *Раздел*. Сходные панели можно упаковывать в аналог Проектов для аннотаций – разделы.
- *Описание*. В текстовой форме пользователь вводит любую информацию о панели генов.
- *Добавить*. Поле ввода генов по одной шт, при этом при введении части названия гена, система будет предлагать подходящие гены из листа всех генов человека.

- *Добавить список.* Если необходимо ввести много генов одновременно, поддерживаются списки генов, разделенные запятой.
- *Доступ.* Поле, доступное только Администратору – для управления контролем доступа к пользовательским панелям.

После сохранения новой панели, она отображается в списке пользовательских панелей.

Доступ к пользовательским панелям разным пользователям помимо создателя конкретной панели может раздавать Администратор.

Рисунок 20. Меню добавления новой пользовательской панели.

3.5.5.2 Просмотр пользовательских панелей

Для просмотра пользовательской панели нужно нажать на кнопку глаза в соответствующей строке панели в меню пользовательских панелей. Пользователь окажется на странице конкретной панели, где можно посмотреть сведения из полей:

- *Название.*
- *Раздел.*
- *Описание.*
- *Список генов.* Лист введенных в панель генов.
- *Доступ.* Поле, доступное только Администратору – для управления контролем доступа к пользовательским панелям.



Рисунок 21. Меню просмотра пользовательской панели.

3.5.5.3 Просмотр панелей PanelApp

Раздел PanelApp представляет собой иерархический список из более чем 300 панелей генов, обновляемый несколько раз в год. Источником списка панелей служит специальный сайт консорциума Genomics England. Лист панелей формирует

сообщество экспертов со всего мира, специализирующихся на генетике и геномике человека. На сегодня это наиболее актуальная система для фильтрации аннотаций, которая используется в лучших клиниках и научных заведениях планеты.

Для просмотра панелей PanelApp необходимо из личного кабинета пользователя перейти в раздел Панели (Рисунок 19), а затем нажать на кнопку *PanelApp* вверху.

Откроется меню просмотра панелей PanelApp (см Рисунок 22), содержащее таблицу со следующими колонками:

Разделы. Верхний уровень иерархии панелей – группировка панелей по типам нозологий.

Название. Название панели.

ID. Идентификатор панели.

Версия. Версия панели.

Обновлена. Дата последнего обновления листа генов в этой панели.

Список генов. Список генов, входящих в панель, через запятую.

Кол-во. Общее количество генов, входящих в панель.

Разделы	Название	ID	Версия	Обновлена	Список генов	Кол-во
Neurology and neurodevelopmental disorders	Acute rhabdomyolysis	1141	1.4	14.04.2023	ACAD9, ACADM, ACADVL, AGL, ALDOA, AMPD1, ANOS, CACNA1S, CAV3, CHKB, COO4, COO8A, CPT2, DGUOK, DMD, DYSF, ENO3, ETFA, ETFB, ETFDH, FDX2, FXRP, FLAD1, GAA, GBE1, GNPPB, GYG1, GYS1, HADHA, HADHB, ISCU, LAMP2, LDHA, LPIN1, COX1, COX2, OBSCN, PFKM, PGAM2, PGK1, PGM1, PHKA1, PHKB, POLG, POLG2, PYGM, RYR1, SCN4A, SGCA, SIL1, SLC22A5, TANGO2, TK2, MLIP, PRKAG2, RBCK1, RRM2B, SUCLA2, CYP2C8, PHKG1, SLC22A12, SLC2A9, TSEN54, TSFM, TYMP	65
	Additional findings health related	399	0.114	14.04.2023	APC, APOB, BRCA1, BRCA2, LDLR, MEN1, MLH1, MSH2, MSH6, MUTYH, PCSK9, RET, VHL, CFTR	14
	Additional findings health related - adults	933	2.2	14.04.2023	APC, APOB, BRCA1, BRCA2, LDLR, MEN1, MLH1, MSH2, MSH6, MUTYH, PCSK9, RET, VHL	13
	Additional findings health related - adult specific	929	1.2	14.04.2023	BRCA1, BRCA2, MLH1, MSH2, MSH6, MUTYH	6
	Additional findings health related - children	930	1.0	14.04.2023	APC, APOB, LDLR, MEN1, PCSK9, RET, VHL	7
	Additional findings health related - CNV analysis adults	934	2.2	14.04.2023	APC, BRCA1, BRCA2, LDLR, MEN1, MLH1, MSH2, MSH6, MUTYH, VHL	10
	Additional findings health related - CNV analysis adult specific	931	1.2	14.04.2023	BRCA1, BRCA2, MLH1, MSH2, MSH6, MUTYH	6
	Additional findings health related - CNV analysis children	932	1.0	14.04.2023	APC, LDLR, MEN1, VHL	4
	Additional findings reproductive carrier status	774	1.0	14.04.2023	CFTR	1

Рисунок 22. Меню просмотра панелей по PanelApp.

3.5.6 Фильтрация вариантов

Модуль Varsiever позволяет фильтровать варианты в аннотации образца по значениям всех колонок в аннотации, а также использовать Панели генов (Пользовательские и/или PanelApp) в соответствии с направительным диагнозом или симптомокомплексом. Выбрать панель генов или набор панелей можно по названию или по разделу.

3.5.6.1 Фильтрация по панели

Чтобы отфильтровать аннотацию по панели генов, в окне аннотации (см. Рисунок 15) в разделе *Панели генов* нажмите кнопку *Выбрать по разделу* или в поле *Ввести по названию* начните вводить название панели (неважно, пользовательской или PanelApp) – система предложит подходящие варианты (см. Рисунок 23). После чего выберите подходящий вариант из выпадающего списка

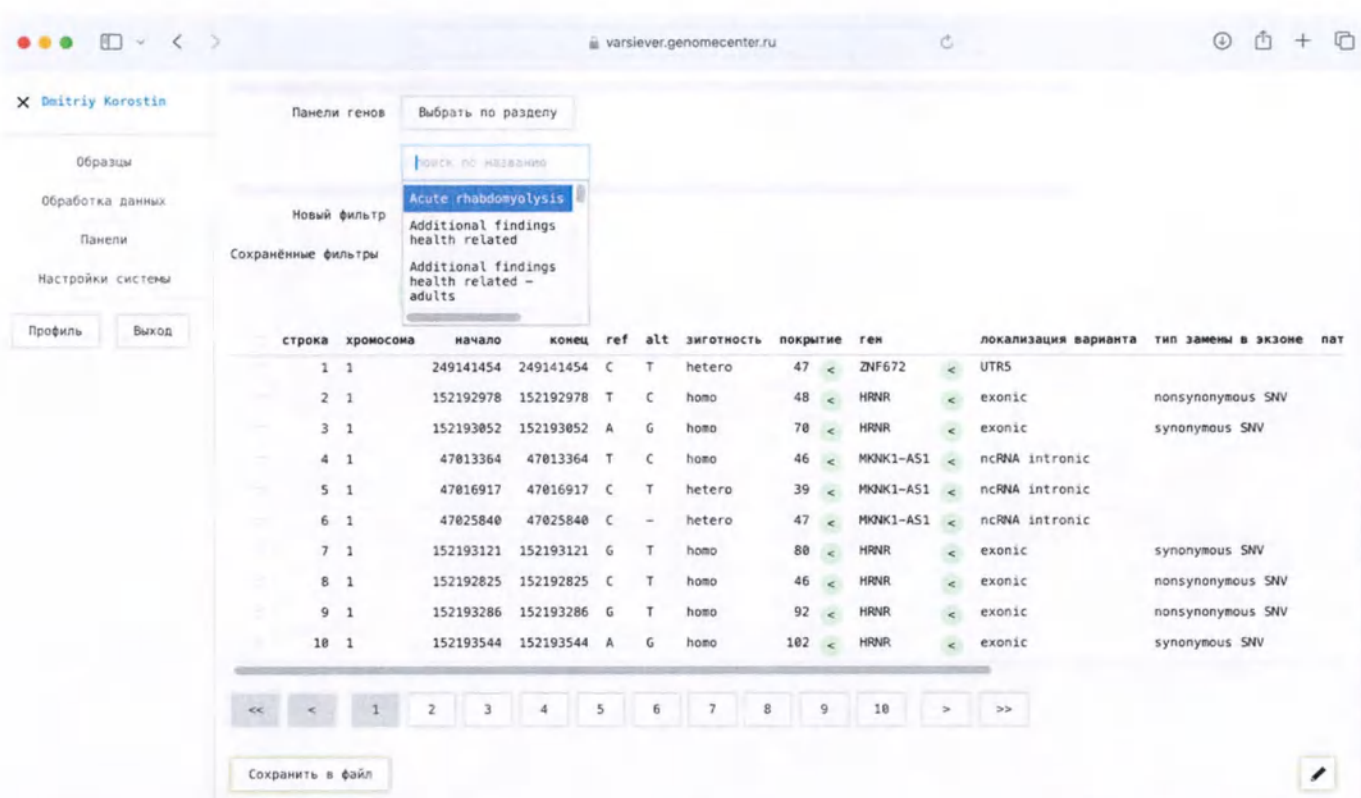


Рисунок 23. Поле *Ввести по названию* фильтрации аннотации по панели генов.

Если вы не знаете точное название панели или хотите посмотреть имеющийся набор панелей, нажмите кнопку *Выбрать по разделу*. При этом отобразится список имеющихся разделов (Рисунок 24), каждый из которых отдельно раскрывается в список подразделов и панелей. В списке панелей можно посмотреть описание каждой панели (источник и список генов) – см. Рисунок 25.

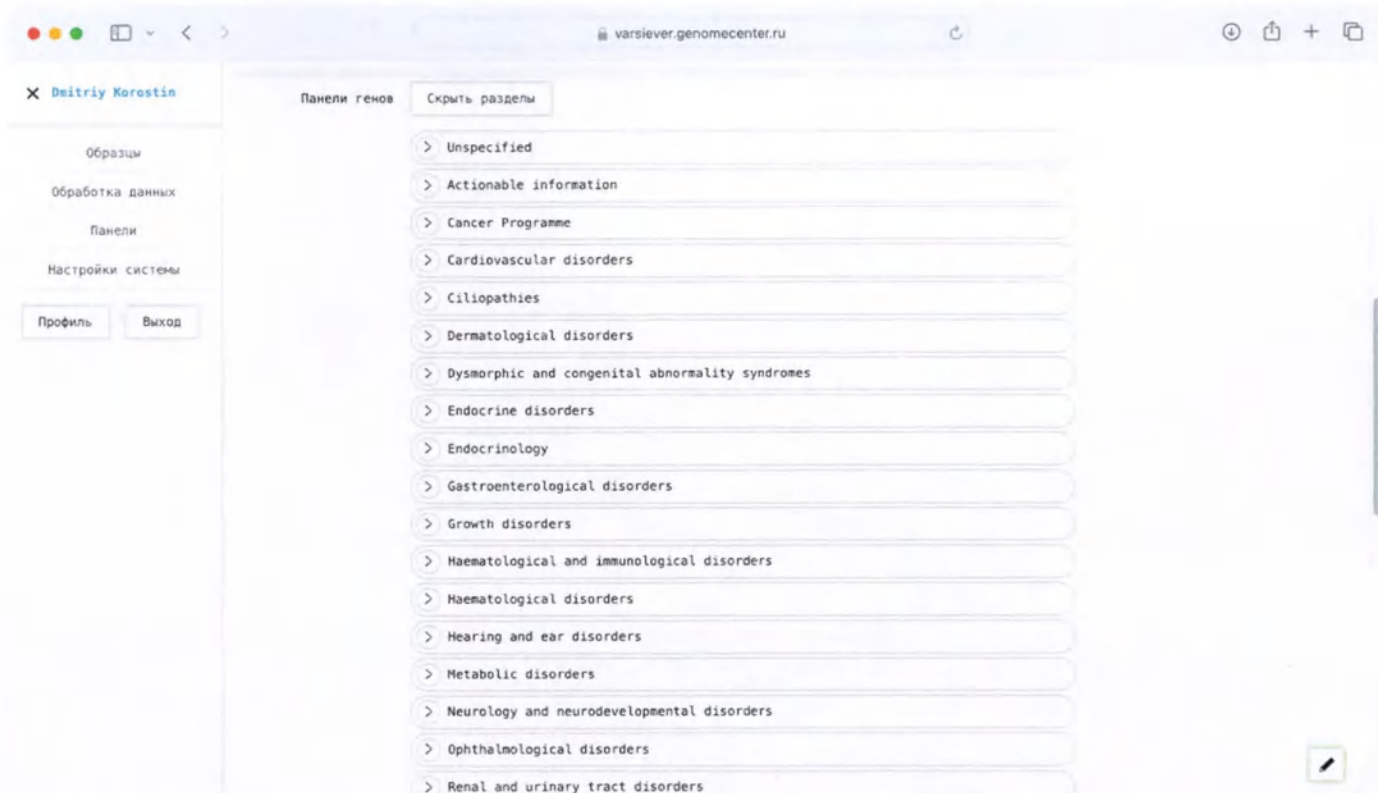


Рисунок 24. Просмотр иерархического списка разделов с панелями генов.

Для выбора панели из списка, нажмите кнопку *Добавить* в описании панели. Обратите внимание, что, при фильтрации по нескольким панелям генов одновременно, в выдаче будут отображаться все варианты, соответствующие хотя бы одному гену из хотя бы одной выбранной панели. Для применения фильтров по панелям необходимо нажать кнопку *Применить фильтры*.

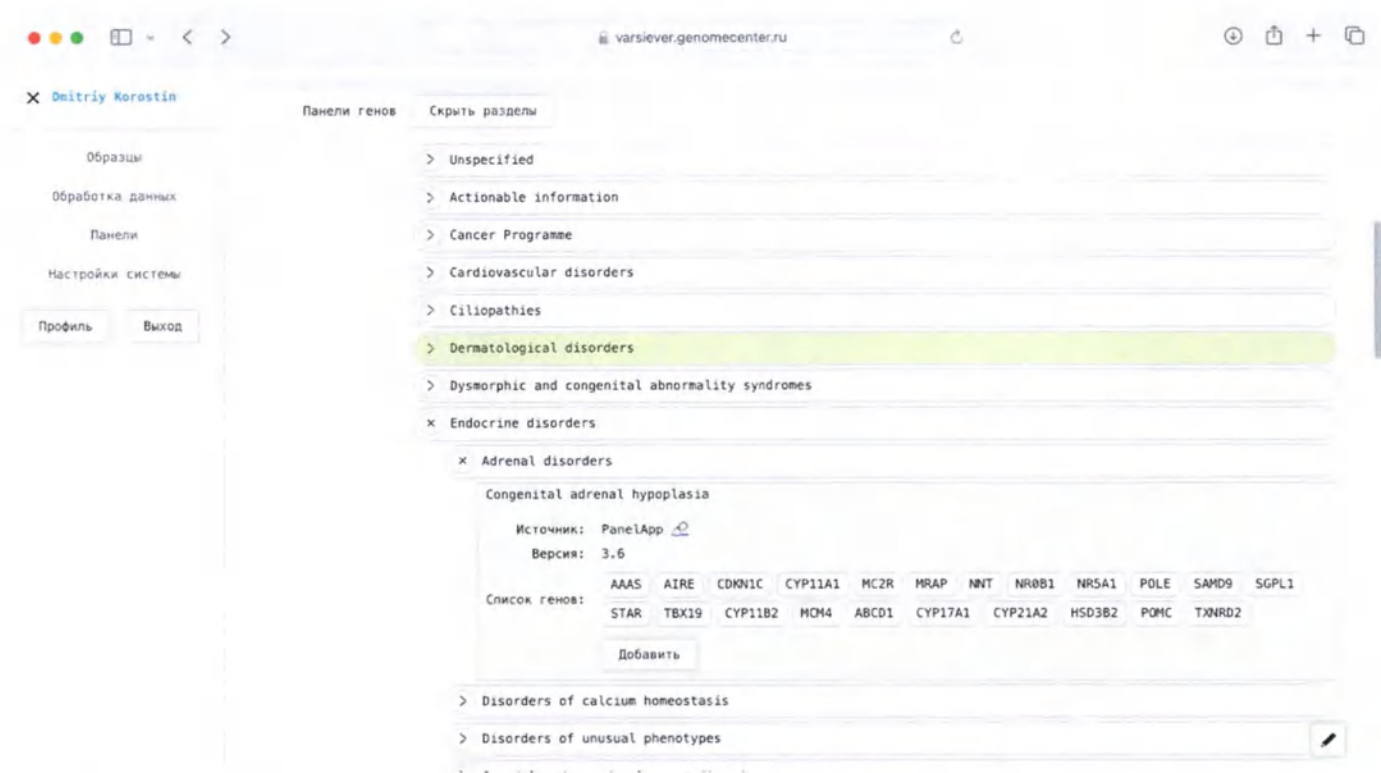


Рисунок 25. Просмотр иерархического списка разделов с панелями генов на уровне описания панели генов.

3.5.6.2 Фильтрация по содержимому аннотации

Чтобы отфильтровать аннотацию по одной или нескольким колонкам аннотации, в окне аннотации (см. Рисунок 15) в разделе *Новый фильтр* введите название параметра, который необходимо отфильтровать первым (Рисунок 26).

Далее появится меню для ввода диапазона или выбора возможных значений соответствующего параметра. Например, при выборе на первом этапе параметра *патогенность*, в появившемся меню нужно выбрать один или более значений патогенности, которые необходимо отобразить в списке вариантов ниже (см. Рисунок 27).

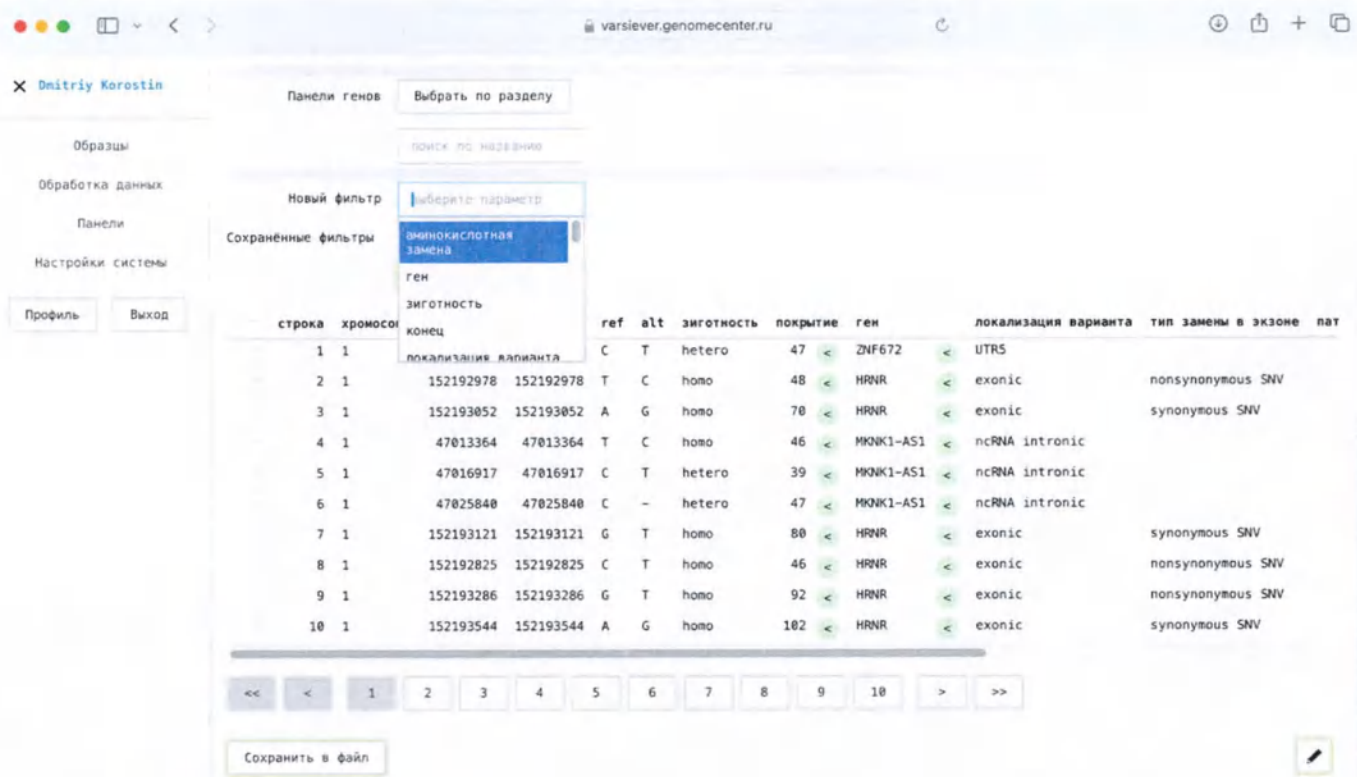


Рисунок 26. Поле *Новый фильтр* для фильтрации аннотации по содержимому.

После того, как допустимые значения для нового фильтра указаны, необходимо нажать кнопку *Добавить*. При этом данный фильтр со списком значений отобразится над строкой *Новый фильтр*.

При необходимости, добавленный фильтр можно удалить, нажав красную кнопку слева от списка выбранных значений. Кроме того, фильтр можно сохранить для быстрого доступа при последующем использовании. Для этого надо нажать зелёную кнопку со знаком + справа от списка значений, в появившемся поле ввода указать желаемое название такого фильтра и ещё раз нажать зелёную кнопку для сохранения. Сохранённые фильтры можно посмотреть в строке *Сохранённые фильтры*.

После выбора значений параметра или параметров фильтрации необходимо нажать на кнопку *Применить фильтры*. Система выдаст лист вариантов в соответствии с выбранными параметрами. Параметры можно комбинировать – см пример

на Рисунок 28: по комбинации параметров (критериев) локализация варианта – *exonic* и тип замены в экзоне – *frameshift deletion*.

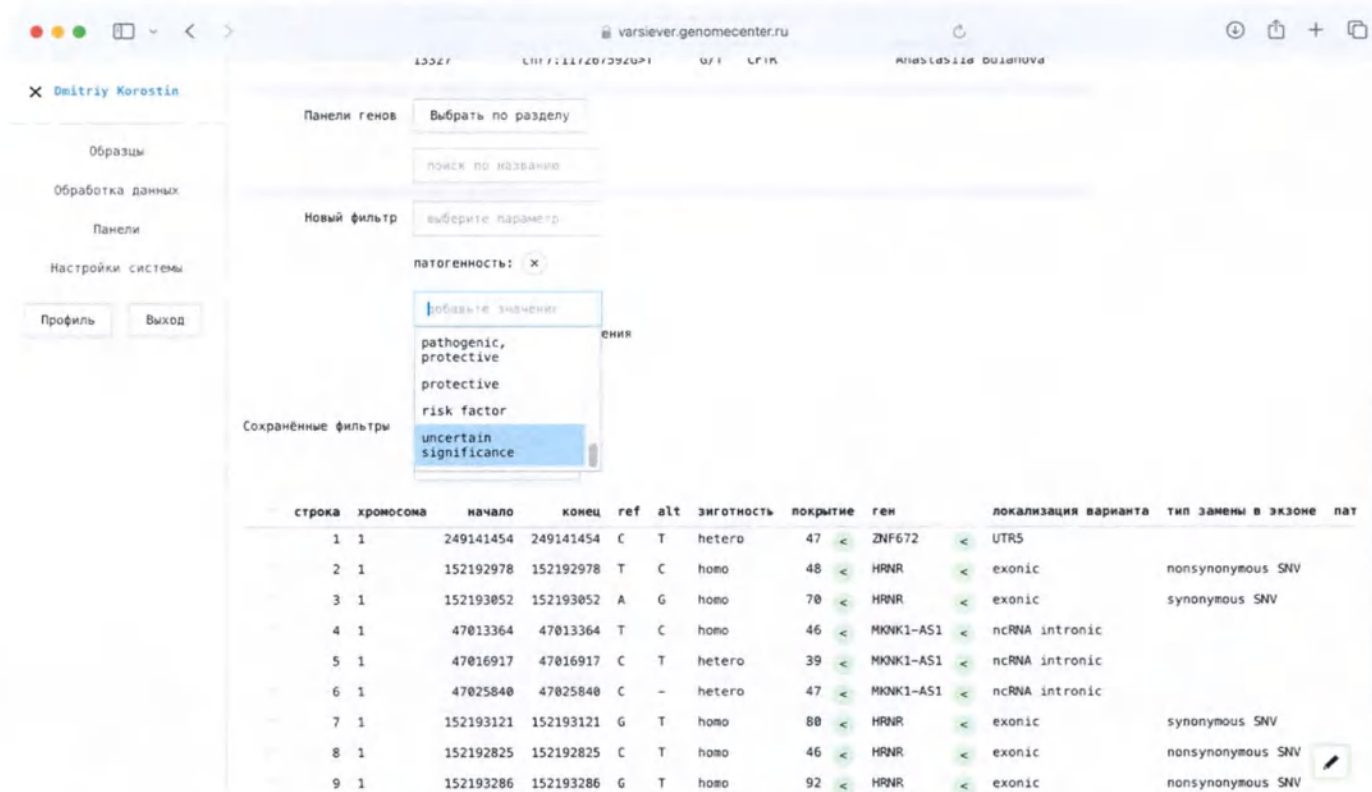


Рисунок 27. Меню выбора значений для фильтрации аннотации по параметру *патогенность*.

Следует обратить внимание на чекбокс *показывать пустые значения*: если он отмечен, то все варианты, для которых выбранное значение не указано, будут отображены, а в противном случае скрыты.

Надо отметить, что фильтрация по панелям в сочетании с фильтрацией по параметру *ген* будет давать пересечение результатов. Таким образом, если список генов из панелей не пересекается со списком значений, указанных для параметра *ген*, результат фильтрации всегда будет пустым.

Varsiever.genomecenter.ru

Панели генов:

Образцы

Обработка данных

Панели

Настройки системы

Профиль Выход

Локализация варианта:

Тип замены в экзоне:

Новый фильтр:

Сохранённые фильтры:

строка	хромосома	начало	конец	ref	alt	зиготность	покрытие	ген	локализация варианта	тип замены в экз
1235	1	101004851	101004863	GCAGCTACCGCCT	-	hetero	51	GPR88	exonic	frameshift delet
1731	1	27699671	27699671	G	-	hetero	43	FCN3	exonic	frameshift delet
2024	1	17087542	17087546	GTGCT	-	hetero	120	MST1L	exonic	frameshift delet
2270	1	144873963	144873963	T	-	hetero	244	PDE4DIP	exonic	frameshift delet
2310	1	144917828	144917828	A	-	hetero	246	PDE4DIP	exonic	frameshift delet

Рисунок 28. Пример фильтрации аннотации по комбинации параметров (критериев) локализация варианта – *exonic* и тип замены в экзоне – *frameshift deletion*.

3.5.7 Выбор вариантов для ввода в заключение

В окне аннотации (см. Рисунок 15) модуль Varsiever позволяет отмечать варианты для дальнейшего рассмотрения без указания их клинической значимости (т. н. *заметки*), либо с указанием категории для дальнейшего включения в заключение (т. н. *значимые варианты*).

3.5.7.1 Внесение вариантов в заметки

Для создания заметки, нужно отметить галочками в чекбоксах (крайний левый столбец таблицы вариантов) интересные варианты (строки с выбранными вариантами при этом будут отмечены зелёным цветом), после чего нажать на кнопку с изображением карандаша в правом нижнем углу экрана (см Рисунок 29).

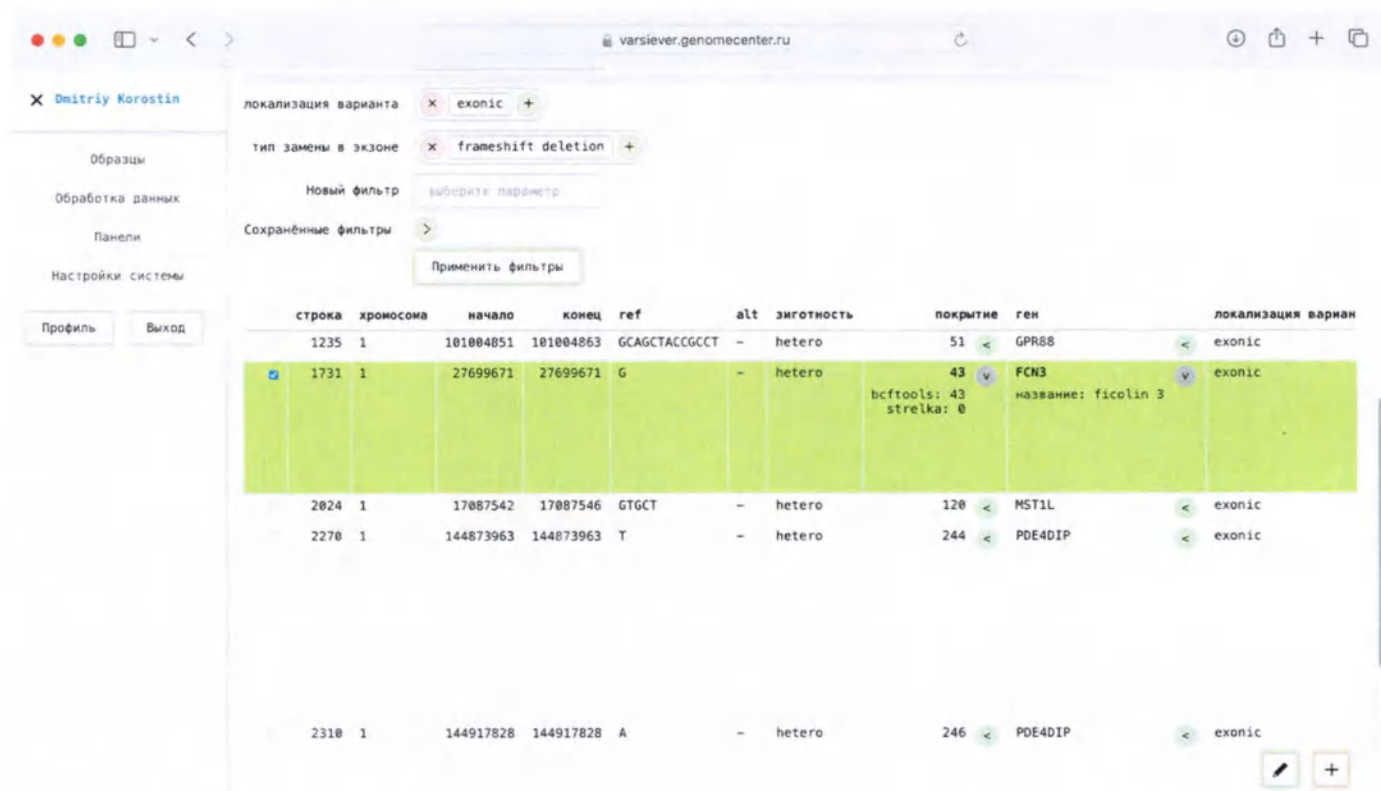


Рисунок 29. Окно аннотации с одним выбранным вариантом (строка выделена зеленым).

При этом в нижней части экрана появится форма для ввода текста заметки и список номеров строк выбранных вариантов (Рисунок 30). Необходимо ввести комментарий и нажать кнопку *Добавить*, чтобы сохранить заметку. Заметка появится над таблицей вариантов.

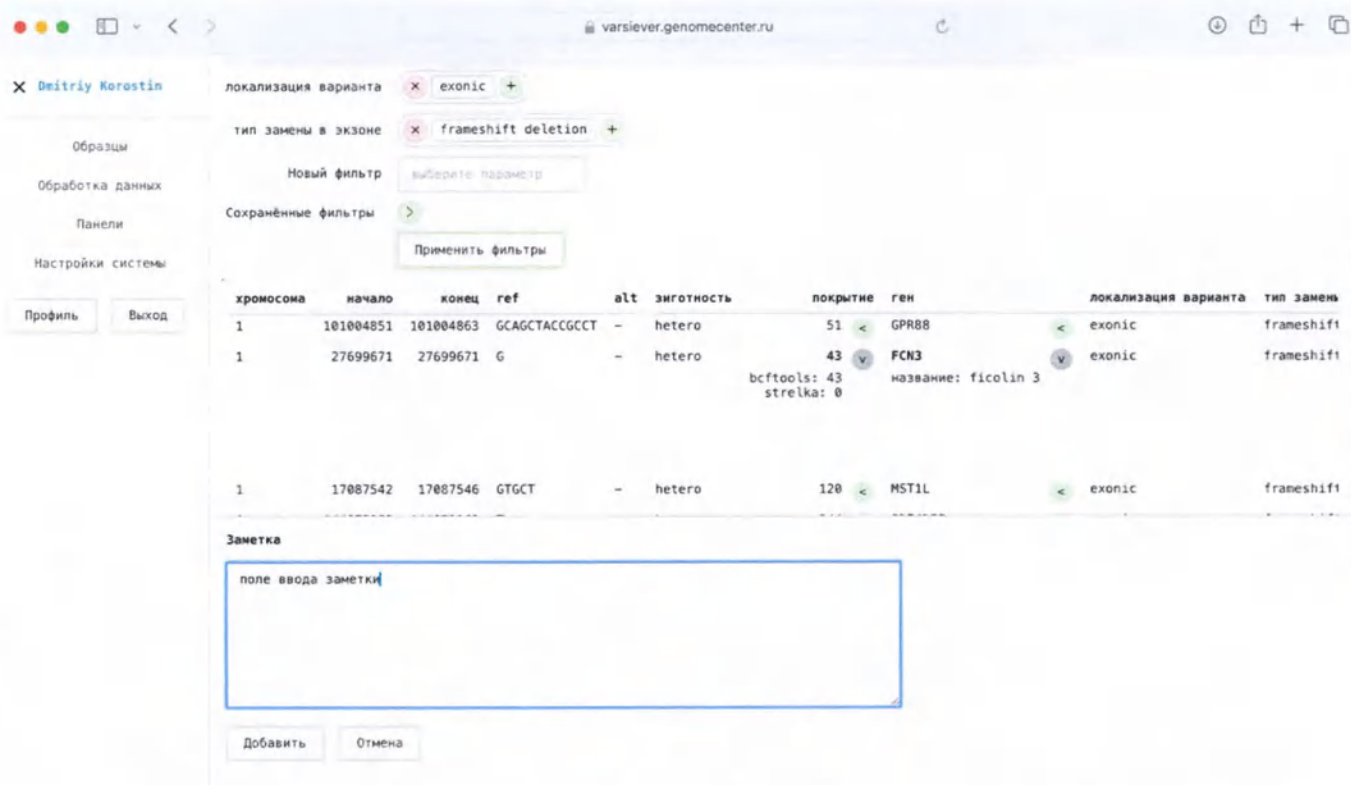


Рисунок 30. Поле ввода текста к добавляемым в заметки вариантам.

3.5.7.2 Внесение вариантов в заключение

Чтобы отметить вариант с указанием клинической значимости, необходимо выбрать соответствующую строку в таблице (будет выделена зеленым), после чего нажать кнопку со знаком "+" в правом нижнем углу экрана. В нижней части экрана появится форма для добавления варианта. Необходимо указать категорию варианта, комментарий (при необходимости). Если всё верно, и выбранный вариант нужно **сохранить**, требуется нажать кнопку с галочкой в нижней части формы (Рисунок 31).

Варианты в зависимости от своей клинической значимости могут быть отнесены к одной из трех категорий:

- Патогенные.
- Вероятно патогенные.
- С неизвестной клинической значимостью.

Выбор категории остается за пользователем несмотря на то, что система предлагает автоматическую классификацию.

Значимые мутации

Транскрипт NM_144702

категория Вероятно патогенные

комментарий комментарий

Рисунок 31. Окно ввода информации для значимого варианта.

Сохранённый вариант появится в соответствующей категории в строке *Значимые мутации* в верхней части страницы (Рисунок 32).

The screenshot shows a web application interface for viewing genetic variants. The user is logged in as 'Dmitriy Korostin'. The main content area displays a list of variants under the heading 'Патогенные варианты' (Pathogenic variants). The variants are organized into three sections: 'Патогенные варианты' (Pathogenic variants), 'Вероятно патогенные' (Probably pathogenic), and 'Неизвестное значение' (Unknown significance). Each section contains a table with columns: 'Строка' (Line), 'Положение в геноме hg19' (Position in hg19 genome), 'Генотип' (Genotype), 'Ген' (Gene), 'Комментарий' (Comment), and 'Добавил' (Added by).

Строка	Положение в геноме hg19	Генотип	Ген	Комментарий	Добавил
18212	chr11:108183151G>T	G/T	ATM		Anastasiia Buianova
Вероятно патогенные					
28721	chr12:112891184G>T	G/T	PTPN11		Anastasiia Buianova
32574	chr21:45753871C>G	C/G	CFAP410		Anastasiia Buianova
Неизвестное значение					
5185	chr2:29543695T>G	T/G	ALK		Anastasiia Buianova
14868	chr8:98994978T>A	T/A	NBN		Anastasiia Buianova
26865	chr17:59876546G>A	G/A	BRIP1		Anastasiia Buianova
13327	chr7:117267592G>T	G/T	CFTR		Anastasiia Buianova

At the bottom of the variant list, there is a section 'Панели генов' (Gene panels) with a button 'Выбрать по разделу' (Select by section).

Рисунок 32. Отображение выбранных для заключения вариантов из аннотации, относящихся к трем категориям патогенности.

3.5.8 Получение файла заключения по результатам фильтрации и интерпретации вариантов

Когда пользователь считает, что все значимые варианты добавлены через механизм внесения вариантов в заключение, и необходимо сформировать заключение, нужно перейти в раздел Описание образца (см. Рисунок 14). Для этого из раздела Аннотация нужно нажать кнопку со стрелкой влево вверху страницы (Рисунок 33).



Рисунок 33. Переход из окна работы с аннотацией к режиму формирования заключения.

В разделе Описание образца в строке Отчёты необходимо выбрать отчёт для редактирования либо сформировать новый нажав кнопку со знаком +. Для нового отчёта сначала необходимо указать название и выбрать шаблон из списка, после чего сохранить его, нажав кнопку с галочкой. Далее система отобразит список полей из выбранного шаблона, которые нужно заполнить. Для редактирования поля нужно нажимать на кнопки с изображением карандаша напротив названия соответствующего поля (см. Рисунок 34 и Рисунок 35).

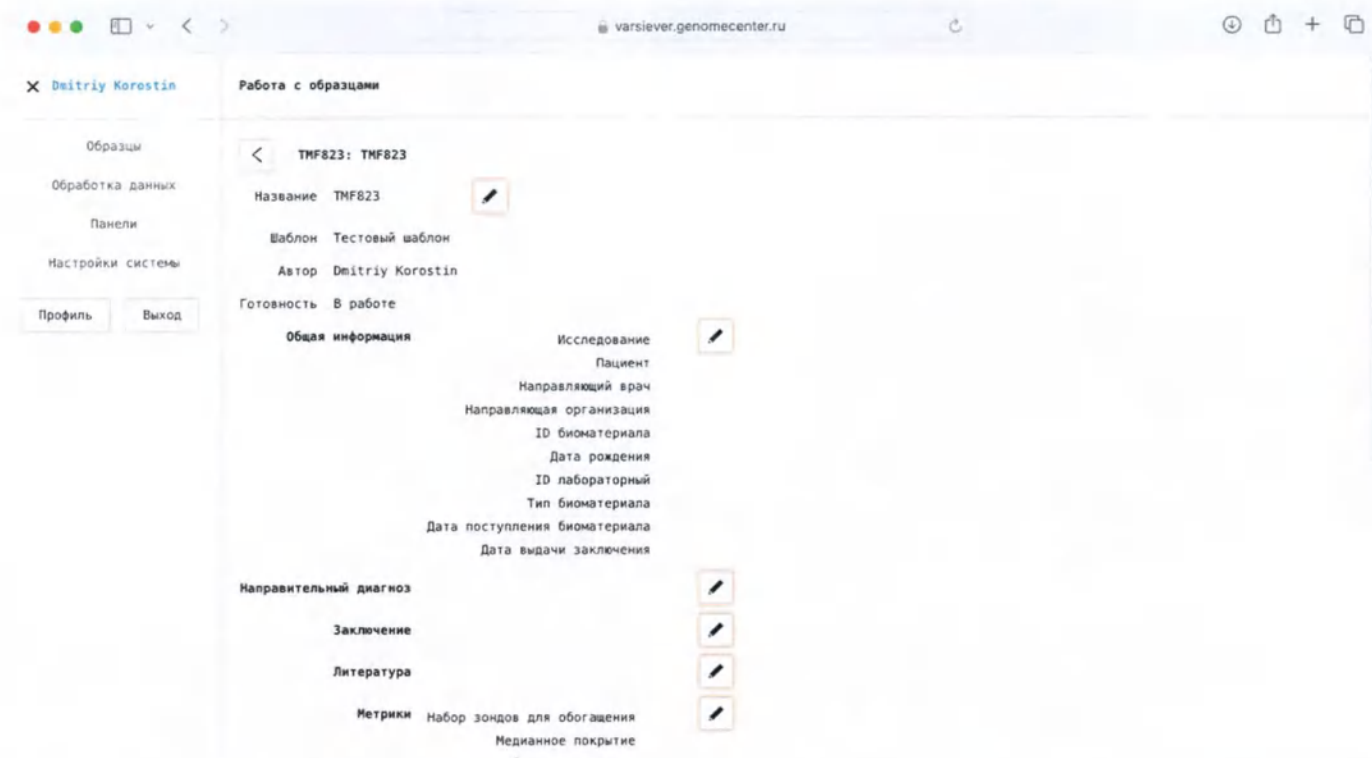


Рисунок 34. Окно редактирования содержимого разделов файла заключения по результатам фильтрации и интерпретации вариантов.

Если требуется сформировать уникальный, специфический для пользователя шаблон файла заключения, необходимо связаться с сотрудником компании-разработчика!

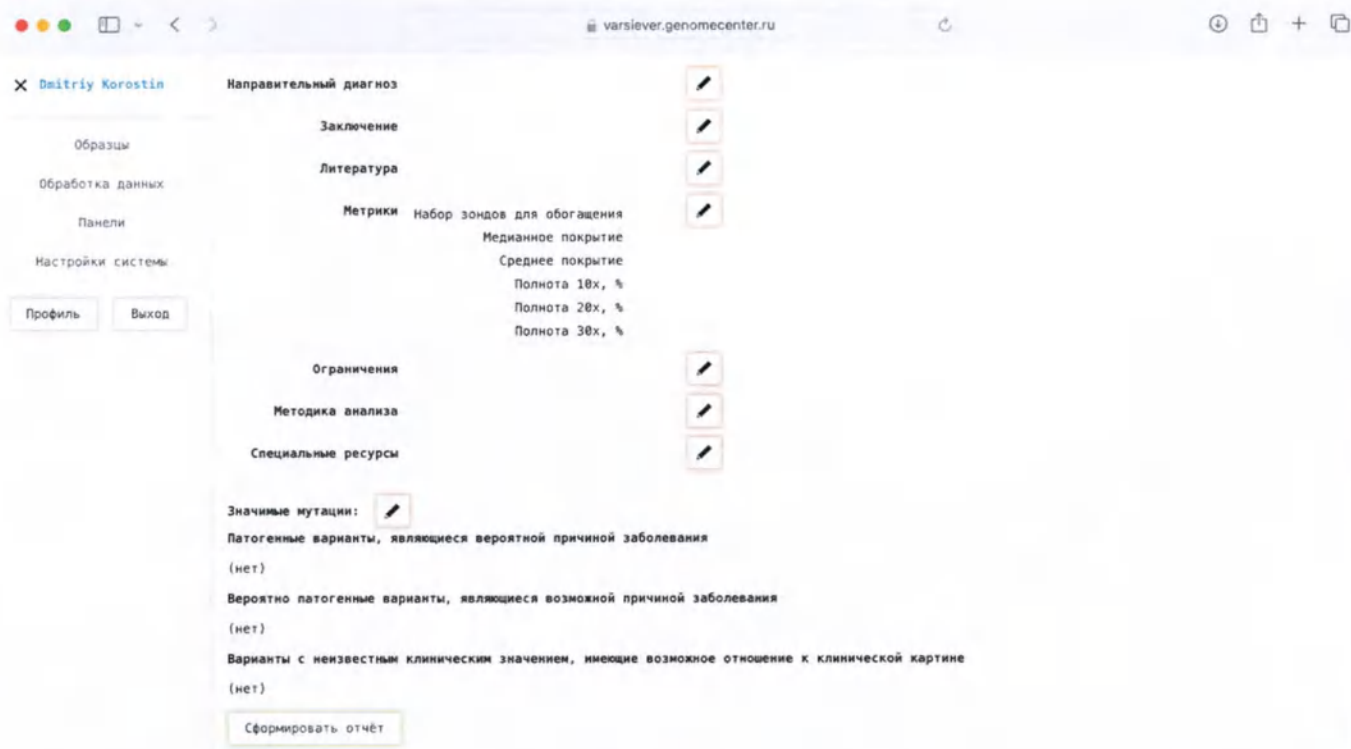


Рисунок 35. Окно редактирования содержимого разделов файла заключения по результатам фильтрации и интерпретации вариантов (*продолжение*).

Если выбранное поле не заполнено и для него существует значение по умолчанию, то при редактировании данного поля будет отображена кнопка со стрелкой влево (Рисунок 36).



Рисунок 36. Поле ввода данных Ограничения для заключения.

При нажатии на эту кнопку поле будет заполнено стандартным значением. Для сохранения значения необходимо нажать кнопку с галочкой (с зелёным фоном). Для отмены редактирования без сохранения — кнопку с крестиком.

Если необходимо, чтобы в отчёте выбранное поле отображалось с новой страницы, под полем ввода необходимо отметить галочкой пункт *с новой страницы*.

Кроме обычных текстовых полей, в отчёте присутствует поле *Значимые мутации*. В нём отображаются лишь те варианты, которые были ранее отмечены пользователем при анализе аннотации. Из этих вариантов необходимо отметить галочкой те, которые пользователь хочет включить в отчёт. Неотмеченные варианты в отчёт не попадут.

Когда все необходимые поля заполнены, нужно нажать кнопку *Сформировать отчёт*. Система откроет в новой вкладке браузера HTML-версию отчёта (см. Рисунок 37). Для формирования PDF-версии нужно открыть в браузере диалог для печати страницы (сочетание клавиш Ctrl+P), после чего выбрать вариант *Сохранить в PDF* и нажать кнопку *Сохранить* (см. Рисунок 38-41).

varsiever.genomecenter.ru

Report

TMF823

Исследование

test

Пациент

test

Направляющий врач

test

Направляющая организация

test

ID биоматериала

test

Дата рождения

test

ID лабораторный

test

Тип биоматериала

test

Дата поступления биоматериала

Дата выдачи заключения

Направительный диагноз

test

Результаты анализа

Патогенные варианты

Положение в геноме hg19	Генотип	Ген	кДНК замена	А/х замена	Экзон	Транскрипт	Частота аллеля по gnomAD	Глубина прочтения
chr11:50818351C>T	C/T	ATM	c.C5932T	p.E1978X	экзон40	NM_000051		146

Вероятно патогенные

Положение в геноме hg19	Генотип	Ген	кДНК замена	А/х замена	Экзон	Транскрипт	Частота аллеля по gnomAD	Глубина прочтения
chr12:112897184C>T	C/T	PTPN21	c.C538T	p.R1773L	экзон4	NM_001330437		84
chr21:45753071C>G	C/G	CFAP410	c.C278C	p.R73P	экзон4	NM_001271440	0.0005	62

Неизвестное значение

Положение в геноме hg19	Генотип	Ген	кДНК замена	А/х замена	Экзон	Транскрипт	Частота аллеля по gnomAD	Глубина прочтения
chr2:29543695T>G	T/G	ALK	c.A1468C	p.T490P	экзон7	NM_004304		143
chr8:90994970T>A	T/A	NBN	c.A151T	p.N51Y	экзон2	NM_002485		146
chr17:59876546G>A	G/A	BRIAP1	c.C1255T	p.R419W	экзон9	NM_032043	0.0002	176
chr7:73726799C>T	C/T	CFTR	c.C3485T	p.R1162L	экзон22	NM_000492	0.0008	90

Заключение

test

Литературные источники заключения

test

Метрики качества анализа

test

Набор зондов для обогащения	Медианное покрытие	Среднее покрытие	Полнота 10x, %	Полнота 20x, %	Полнота 30x, %
v8					

Анализ не позволяет выявить

1. Делеции и инсерции длиной более 40 п.о.;
2. Варианты, связанные с экспансией коротких повторов;
3. Варианты изменения числа нуклеотидов в гомополимерных регионах длиной более 10 п.о.;
4. Варианты в интронах (за исключением канонических сайтов сплайсинга);
5. Варианты в генах, имеющих псевдогены;
6. Мозаичные варианты;
7. Полиплоидии, анеуплоидии;
8. Эпигенетические модификации

программное обеспечение Mutracker Ver. 3.3.0 (08.2023) по ТУ 58.29.32-001-01963278-2023

Рисунок 37. HTML-версия отчёта.

62

ООО ЦИРМИ INFO@CIRMI.RU WWW.GOSZDRAVNADZOR.RU



test

Исследование	test
Пациент	test
Направляющий врач	test
Направляющая организация	test
ID биоматериала	test
Дата рождения	
ID лабораторный	test
Тип биоматериала	test
Дата поступления биоматериала	
Дата выдачи заключения	

Направительный диагноз

test

Результаты анализа

Патогенные варианты

Положение в геноме hg19	Генотип	Ген	кДНК замена	А/к замена	Экзон	Транскрипт	Частота аллеля по gnomAD	Глубина прочтения
chr11:105183151G>T	C/T	ATM	c.C5932T	p.E1978X	exon40	NM_000051		146

Вероятно патогенные

Положение в геноме hg19	Генотип	Ген	кДНК замена	А/к замена	Экзон	Транскрипт	Частота аллеля по gnomAD	Глубина прочтения
chr12:112591184G>T	G/T	PTPN11	c.C518T	p.R173L	exon4	NM_001330437		84
chr21:45753071C>G	C/G	CFAP410	c.C218C	p.R73P	exon4	NM_001271440	0.0005	62

Неизвестное значение

Положение в геноме hg19	Генотип	Ген	кДНК замена	А/к замена	Экзон	Транскрипт	Частота аллеля по gnomAD	Глубина прочтения
-------------------------	---------	-----	-------------	------------	-------	------------	--------------------------	-------------------

<https://varsiever.genomecenter.ru/report.html>

Страница 1 из 4

Рисунок 38. Первая страница выдачи по тестовому шаблону файла заключения по результатам фильтрации и интерпретации вариантов.

геном hg19			замена	замена			аллеля по gnomAD	прочтения
chr2:29543695T>G	T/G	ALK	c.A1468C	p.T490P	exon7	NM_004304		143
chr8:90994970T>A	T/A	NBN	c.A151T	p.N51Y	exon2	NM_002485		146
chr17:59876546G>A	G/A	BRIP1	c.C1255T	p.R419W	exon9	NM_032043	0.0002	176
chr7:117267592G>T	G/T	CFTR	c.G3485T	p.R1162L	exon22	NM_000492	0.0008	90

Заключение

test

Литературные источники заключения

test

Метрики качества анализа

test

Набор зондов для обогащения	Медианное покрытие	Среднее покрытие	Полнота 10x, %	Полнота 20x, %	Полнота 30x, %
Agilent all-exon v8	114	118.7	97.8	97.5	97.1

Анализ не позволяет выявить

1. Делеции и инсерции длиной более 40 п.о.;
2. Варианты, связанные с экспансией коротких повторов;
3. Варианты изменения числа нуклеотидов в гомополимерных регионах длиной более 10 п.о.;
4. Варианты в интронах (за исключением канонических сайтов сплайсинга);
5. Варианты в генах, имеющих псевдогены;
6. Мозаичные варианты;
7. Полиплоидии, анеуплоидии;
8. Эпигенетические модификации.

Методика анализа

Подготовка образцов к секвенированию: выделение геномной ДНК из образцов венозной крови производится с помощью наборов для

<https://varsiever.genomecenter.ru/report.html>

Страница 2 из 4

Рисунок 39. Вторая страница выдачи по тестовому шаблону файла заключения по результатам фильтрации и интерпретации вариантов.

выделения ДНК (Qiagen) по протоколу производителя. Дальнейшая пробоподготовка включает ультразвуковую фрагментацию геномной ДНК и стандартные этапы подготовки библиотек фрагментов ДНК. Обогащение библиотек фрагментов ДНК последовательностями экзонов проводится с помощью зондов по лабораторному протоколу.

Секвенирование белок-кодирующих последовательностей осуществляется методом парно-концевых прочтений по протоколу производителя.

Анализ данных. Использован автоматизированный пайплайн, написанный на языке python, включающий в себя оценку качества секвенирования (FastQC), удаление разбалансированных оснований и ридов с низким качеством (Trimmomatic), выравнивание прочтений на референсную сборку генома человека GRCh37/hg19 при помощи Bwa MEM, коллинг вариантов с помощью BCftools. Метрики покрытия были получены программой Picard. Аннотация полученных вариантов была реализована при помощи OMIM и ANNOVAR.

Интерпретация. Оценка клинической значимости выявленных вариантов производится в соответствии с критериями ACMG 2,3, с использованием баз данных вариантов и литературных источников.

Список анализируемых генов: ATM, BARD1, BRCA1, BRCA2, BRIP1, CDH1, CHEK2, NBN, NF1, PALB2, PMS2, PTEN, RAD50, RAD51C, RAD51D, STK11, TP53, MLH1, MRE11A, MSH2, MSH6, XRCC2.

Использованные специализированные ресурсы

1. Belova, V. et al. System analysis of the sequencing quality of human whole exome samples on BGI NGS platform. Scientific Reports 12, (2022).
2. Richards S. et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology //Genetics in medicine. – 2015. – Т. 17. – №. 5. – С. 405-423.
3. Рыжкова О. П. и др. Руководство по интерпретации данных последовательности ДНК человека, полученных методами массового параллельного секвенирования (MPS) (редакция 2018, версия 2) //Медицинская генетика. – 2019. – Т. 18. – №. 2. – С. 3-23.
4. HUGO Gene Nomenclature Committee (HGNC) – Committee of the Human Genome Organisation that sets the standards for human gene nomenclature. <https://www.genenames.org/>
5. GENCODE – High quality reference gene annotation and experimental validation. <https://www.gencodegenes.org/>
6. Ensemble – Genome browsers for the retrieval of genomic information.

Рисунок 40. Третья страница выдачи по тестовому шаблону файла заключения по результатам фильтрации и интерпретации вариантов.

- <https://www.ensembl.org/index.html>
7. dbSNP – Database of short genetic variants.
<https://www.ncbi.nlm.nih.gov/snp/>
 8. gnomAD – Germline variant frequencies exome-wide.
<https://gnomad.broadinstitute.org/>
 9. 1000Genomes – Germline variant frequencies genome-wide.
<https://www.internationalgenome.org/>
 10. RUSeq – База данных генетических вариантов, встречающихся у жителей РФ. <http://ruseq.ru/>
 11. Pfam – Collection of protein families/domains. <https://pfam.xfam.org/>
 12. UniProtKB – Comprehensive resource of protein sequence and functional information. <https://www.uniprot.org/>
 13. CORUM – The comprehensive resource of mammalian protein complexes. <http://mips.helmholtz-muenchen.de/corum/>
 14. KEGG – Knowledge base on the molecular interaction, reaction and relation networks. <https://www.genome.jp/kegg/network.html>
 15. ClinVar – Database of genomic variants of clinical significance.
<https://www.ncbi.nlm.nih.gov/clinvar/>
 16. VarSome – The human genomic variant search engine.
<https://varsome.com/>
 17. Orphanet – Database with the aspiration of gathering, providing and improving knowledge on rare diseases. <http://www.orpha.net/>
 18. HPO – Vocabulary that contains phenotypic information about genes or product genes. <https://hpo.jax.org/>
 19. OMIM – Catalog of all known human genes and genetic phenotypes.
<https://www.omim.org/>
 20. CancerMine – Literature-mined database of tumor suppressor genes/proto-oncogenes. <http://bionlp.bcgsc.ca/cancermine/>
 21. CIViC – Clinical interpretations of variants in cancer. <https://civicdb.org/>
 22. NCCN – The National Comprehensive Cancer Network, an alliance of leading cancer centers that creates clinical practice guidelines.
<https://www.nccn.org>
 23. OncoBRCA – База данных частот вариантов в генах системы репарации ДНК, полученных в рамках проекта «Наследственные синдромы в РФ». <https://oncobrca.ru/>
 24. OncoTree – Open-source ontology developed at MSK-CC for standardization of cancer type diagnosis. <http://oncotree.mskcc.org/>
 25. CPIC – ресурс по фармакогеномике <https://cpicpgx.org/genes-drugs/>
 26. PharmGKB – база данных по фармакогеномике
<https://www.pharmgkb.org>

Mutracker Ver. 3.3.0 (08.2023)
ТУ SB 29-32-001-01952278-2023
(ПУ № РЭН 28XX/XXXXX от XX.XX.28XX)

Разработка: ФГАУ ВО РНИМУ им. Н.И. Пирогова Минздрава России
Техподдержка: +7 905 782-20-12, korostin_d@rsmu.ru

<https://varsiever.genomecenter.ru/report.html>

Страница 4 из 4

Рисунок 41. Четвертая страница выдачи по тестовому шаблону файла заключения по результатам фильтрации и интерпретации вариантов.

Это Руководство пользователя предоставляется в форме электронного доку-
мента и размещено в сети интернет по адресу
<https://varsiever.genomecenter.ru/instructions.pdf>

4. АДРЕС ДЛЯ ОБРАЩЕНИЯ

Производитель: ФГАОУ ВО "Российский Национальный Исследовательский Медицинский Университет им. Н.И. Пирогова" Министерства здравоохранения Российской Федерации, Россия.

Адрес производителя: ФГАОУ ВО РНИМУ им. Н.И. Пирогова Минздрава России, Российская Федерация, 117513, г. Москва, ул. Островитянова, дом 1, стр. 6.

Место производства: ФГАОУ ВО РНИМУ им. Н.И. Пирогова Минздрава России: Российская Федерация, город Москва, вн.тер.г. муниципальный округ Обручевский, улица Островитянова, Дом 1, Строение 1.

По вопросам работы ПО Mutracker следует обращаться по адресу: ФГАОУ ВО РНИМУ им. Н.И. Пирогова Минздрава России, Российская Федерация, г. Москва 117513, ул. Островитянова д. 1, стр. 1, пом. 527, тел./факс +7 (495) 434 14 22, <https://genomecenter.ru>

Служба клиентской поддержки:

Телефон: +7 (905) 702 20 12

Email: korostin_do@rsmu.ru, gencentrsmu@gmail.com, lab@genomecenter-rsmu.ru

По телефону можно обращаться по будням с 10:00 до 20:00 по московскому времени.

Время ожидания ответа от специалиста поддержки по электронной почте: по будням – в течение рабочего дня, в выходной – до трех дней.

ПРИЛОЖЕНИЕ А

ПЕРЕЧЕНЬ ПРИМЕНЯЕМЫХ НАЦИОНАЛЬНЫХ СТАНДАРТОВ

Перечень национальных стандартов, применяемых к ПО Mutracker, включает:

1. ГОСТ Р ИСО/МЭК 12119-2000 Информационная технология. Пакеты программ. Требования к качеству и тестирование;
2. ГОСТ Р ИСО/МЭК 9126-93 Информационная технология. Оценка программной продукции. Характеристики качества и руководства по их применению;
3. ГОСТ Р ИСО 9127-94 Системы обработки информации. Документация пользователя и информация на упаковке для потребительских программных пакетов;
4. ГОСТ 28195-89 Оценка качества программных средств. Общие положения;
5. ГОСТ Р ИСО/МЭК ТО 9294-93 Информационная технология. Руководство по управлению документированием программного обеспечения;
6. ГОСТ Р 51188-98 Защита информации. Испытания программных средств на наличие компьютерных вирусов. Типовое руководство;
7. ГОСТ Р МЭК 62304-2013 Изделия медицинские. Программное обеспечение. Процессы жизненного цикла.

Примечание: указанные выше стандарты действительны на момент утверждения Руководства пользователя.

ПЕРЕЧЕНЬ СОКРАЩЕНИЙ

NGS	next generation sequencing, высокопроизводительное секвенирование
БД	база данных
ГА-2000	секвенатор нуклеиновых кислот ГА-2000 (ПУ № РЗН 2022/17851 от 25.07.2022)
ГОСТ	межгосударственный стандарт
ГОСТ Р	государственный стандарт России
ИСО (ISO)	Международная организация по стандартизации
ЛИС	лабораторная информационная система
МЭК (IEC)	Международная электротехническая комиссия
НД	несанкционированный доступ
ООО	общество с ограниченной ответственностью
ОС	операционная система
ПК	персональный компьютер
ПО	программное обеспечение
vcf	variant call format, формат файла вариантов при анализе NGS

Пронумеровано, прошито
и скреплено печатью на 70
_____ листах

