

**Описание функциональных характеристик Информационной системы
“InUnio Flow: система анализа и аннотирования омиксных (геномных,
транскриптомных) данных”**

Оглавление

1. Функциональные характеристики	3
1.1. Краткое описание системы	3
1.2. Краткое описание текущего функционала и существующих ограничений	4
1.3. Ключевые функции	4
2. Информация, необходимая для установки и эксплуатации системы	5
2.1. Требования к рабочему месту пользователя	5
2.2. Требования к серверной инфраструктуре	5
3. Описание функционала	6
3.1. Общее расположение элементов	6
3.2. Экран “Список пациентов”	6
3.3. Экран “Информация о пациента”	7
3.3.1. Вкладка “Инфо”	7
3.3.2. Вкладка “образцы”	8
3.4. Экран “Профиль образца”	9
3.4.1. Вкладка “Инфо”	9
3.4.2. Вкладка “Файлы”	10
3.4.3. Вкладка “Workflow”	10
3.5. Экран “Исследования”	11
3.5.1. Вкладка “Список исследований”	11
3.5.2. Вкладка “Доступные виды исследований”	12
3.6. Экран “Информация об исследовании”	12
4. Основные операции пользователя	12
4.1. Вход на сайт и авторизация	12
4.2. Добавление нового пациента	13
4.3. Добавление образцов	14
4.4. Проведение исследования	16
4.5. Работа с результатами исследования:	17
4.6. Работа с ВАР-Фалами	18

1. Функциональные характеристики

1.1. Краткое описание системы

InUnio Flow - система анализа и аннотирования омиксных (геномных, транскриптомных) данных, представляет собой облачную среду исполнения биоинформатических пайплайнов, написанных с помощью технологии nextflow с графическим веб-интерфейсом пользователя.

На текущий момент в системе предустановлен только один биоинформатический протокол, исходный код которого находится в открытом доступе:

<https://github.com/zenomeplatform/nf-germline-mapping>

Данный протокол представляет собой биоинформатический рабочий процесс, предназначенный для первичного анализа данных экзомного секвенирования от сырых данных (fastq) до файлов с выровненными на референсный геном прочтениями (BAM). Обработка BAM файла происходит в соответствии лучшими практиками GATK.

Основные этапы рабочего процесса:

- * Контроль качества исходных данных и фильтрация технических последовательностей генома.
- * Выравнивание геномных карт на референс
- * Оценка глубины и полноты секвенирования
- * Оценка эффективности целевого обогащения ДНК-библиотеки экзомными последовательностями

1.2. Краткое описание текущего функционала и существующих ограничений

В настоящее время доступны следующие функции:

- Возможность добавлять и редактировать данные пациента
- Внутри карточки пациента можно добавлять информацию соответствующих пациенту образцов, указывать необходимую мета-информацию
- Для каждого образца возможно осуществлять загрузку файлов с данными высокопроизводительного секвенирования в формате *fastq (*fq.gz, *fastq.gz)
- Во вкладке исследования (на данном этапе разработки) можно выбрать необходимый тип исследования, указать образец/файлы к которым данное исследование будет применено и запустить анализ на облачном сервере
- После завершения вычислений, в карточке исследования доступны ссылки для скачивания полученных файлов и отчетов о проведении исследования, возможность открывать отчеты непосредственно в браузере
- Для файлов формата *bam (в доступной версии являются конечным результатом исследования) доступна возможность открывать их в веб версии геномного браузера IGV, не покидая интерфейса системы.

Существующие ограничения на данной стадии разработки:

- Возможен запуск анализа только для данных экзомного секвенирования человека
- Для выравнивания используется hg38 версия референсного генома
- Для оценки эффективности целевого обогащения используется bed-файл от панели Agilent SureSelectV7 (анализ данных экзотов, полученных с помощью других реагентов возможен, однако при вычислении эффективности целевого обогащения (fold enrichment) будут использованы геномные координаты указанной выше панели.
- В настоящее время для демо-версии используется сервер с ограниченным вычислительным ресурсом, поэтому анализ полных данных может занять продолжительное время (8 и более часов, в зависимости от размера загруженных файлов)

1.3. Ключевые функции

1. Авторизация пользователей по пред созданной учетной записи.
2. Возможность добавлять и редактировать данные пациента/ов.
3. Добавлять образцы с мета-информацией в существующие карточки пациентов.
4. Осуществлять загрузку файлов с данными высокопроизводительного секвенирования в формате *fastq (*fq.gz, *fastq.gz)

5. Проводить исследования (на текущий момент доступен только один один биоинформатический протокол.
6. Во вкладке исследования (на данном этапе разработки) можно выбрать необходимый тип исследования, указать образец/файлы к которым данное исследование будет применено и запустить анализ на облачном сервере
7. После завершения вычислений, в карточке исследования доступны ссылки для скачивания полученных файлов и отчетов о проведении исследования, возможность открывать отчеты непосредственно в браузере
8. Для файлов формата *bam (в доступной версии являются конечным результатом исследования) доступна возможность открывать их в веб версии геномного браузера IGV, не покидая интерфейса системы.

2. Информация, необходимая для установки и эксплуатации системы

2.1. Требования к рабочему месту пользователя

Рабочее место должно отвечать следующим требованиям:

1. Операционная система
 - a. Linux
 - b. Windows
 - c. MacOS
2. Веб-браузер
 - a. Chrome/Firefox/Safari актуальной версии

2.2. Требования к серверной инфраструктуре

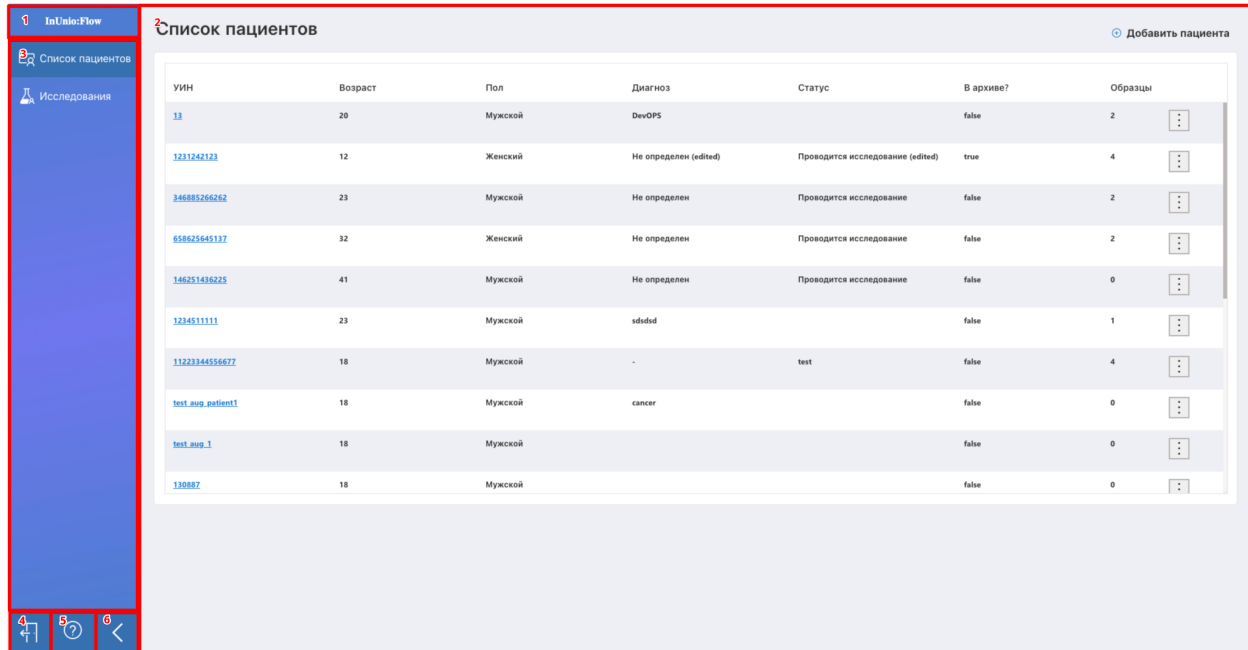
Для запуска серверной части системы необходимо подготовить сервер со следующими характеристиками:

1. CPU-4
2. Memory-8 GB
3. HDD-200 GB

3. Описание функционала

3.1. Общее расположение элементов

Пример компоновки элементов:



На экране можно выделить следующие элементы:

1. Название Информационной системы
2. Динамический элемент, зависит от страницы на которой находится пользователь.
3. Левое контекстное меню для быстрого перехода.
4. Кнопка для выхода из учетной записи.
5. Кнопка для вызова справки
6. Кнопка для минимизации левого меню.

3.2. Экран “Список пациентов”

Пример экрана:

Список пациентов 1 [Добавить пациента](#)

2

УИН	Возраст	Пол	Диагноз	Статус	В архиве?	Образцы
13	20	Мужской	DevOPS		false	2
1231242123	12	Женский	Не определен (edited)	Проводится исследование (edited)	true	3
346885266262	23	Мужской	Не определен	Проводится исследование	false	2
658625645137	32	Женский	Не определен	Проводится исследование	false	2
146251436225	41	Мужской	Не определен	Проводится исследование	false	0
1234511111	23	Мужской	sdsdsd		false	1
11223344556677	18	Мужской	-	test	false	4
text_aup_patient1	18	Мужской	cancer		false	0
text_aup_1	18	Мужской			false	0
130887	18	Мужской			false	0

3

Редактирование
Удаление

Экран содержит следующие элементы:

1. Добавление пациента. Для добавление нового пациента необходимо указать следующую информацию:
 - a. УИН - Уникальный идентификационный номер
 - b. Возраст - количество полных лет
 - c. Пол
 - d. Диагноз - в свободной форме
 - e. Комментарий - внутренний комментарий
 - f. Статус - в свободной форме
 - g. В архиве - Наличие записи в Архиве.
2. Таблица с данными существующих пациентов. Содержит информацию, указанную при добавлении пациента + количество загруженных образцов.
3. Модальное окно для редактирования/Удаления пользователей. Активируется при нажатии на кнопку, вызывающую контекстное меню "...".

3.3. Экран "Информация о пациента"

Экран содержит две вкладки "Инфо" - Содержащую в себе общую информацию о пациента и "Образцы" - Содержащую в себе информацию о загруженных образцах и возможность загрузить новые образцы.

3.3.1. Вкладка "Инфо"

Пример экрана:

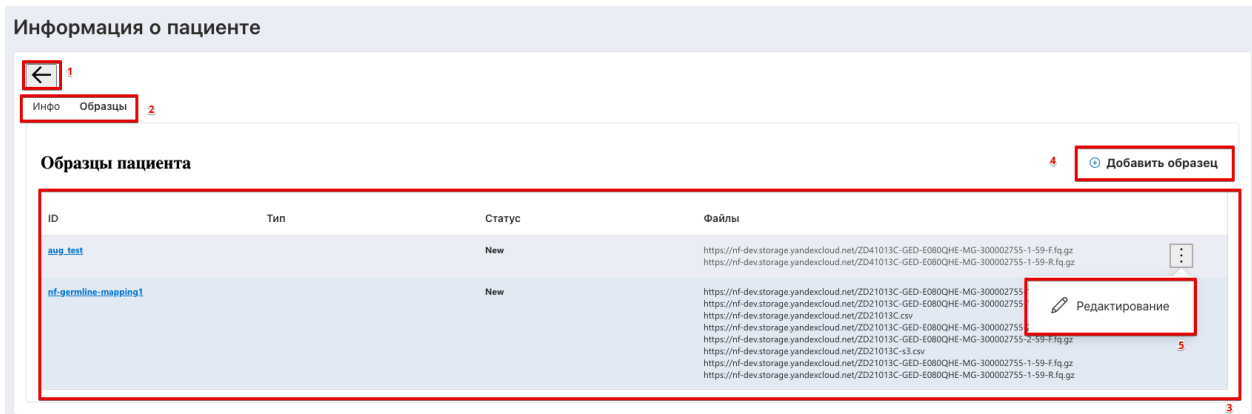


Экран содержит следующие элементы:

1. Кнопка возвращения на предыдущий экран.
2. Блок для переключения вкладок
3. Данные о пациенте.

3.3.2. Вкладка “образцы”

Пример экрана:



Экран содержит следующие элементы:

1. Кнопка возвращения на предыдущий экран.
2. Блок для переключения вкладок
3. Данные образца. Содержит следующие поля
 - a. ID - номер образца
 - b. Тип образца
 - c. Статус образца
 - d. Файлы - ссылка на загруженные файлы.
 - e. Контекстное меню
4. Кнопка добавления нового образца. При добавлении необходимо указать следующую информацию:

- a. ID образца - номер образца
 - b. Тип образца
 - c. Тип считывания
 - d. Устройство для забора
 - e. Тип биоматериала
 - f. Дата получения
 - g. Секвенатор
 - h. Комментарии
5. Контекстное меню, с помощью которого можно отредактировать информацию образца.

3.4. Экран “Профиль образца”

Экран содержит три вкладки:

- “Инфо” - Содержащую в себе общую информацию об образце
- “Файлы” - Содержащую в себе загруженные файлы образца + возможность загрузить новые
- “Workflow” - содержит в себе список созданных исследований по образцам + возможность создать новое исследование.

3.4.1. Вкладка “Инфо”

Пример экрана:

Профиль образца

← 1

Инфо Файлы Workflow 2

Информация об образце

ID образца	aug_test
Тип образца	
Тип считывания	
Устройство для забора	
Тип биоматериала	
Дата получения	
Секвенатор	
Комментарии	
Статус	New

Экран содержит следующие элементы:

1. Кнопка возвращения на предыдущий экран.

2. Блок для переключения вкладок
3. Информацию об образце.

3.4.2. Вкладка “Файлы”

Пример экрана:

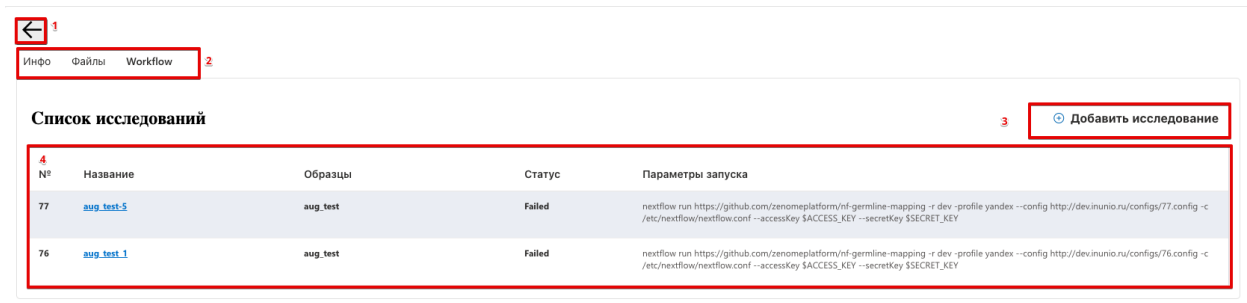


Экран содержит следующие элементы:

1. Кнопка возвращения на предыдущий экран.
2. Блок для переключения вкладок
3. Кнопка загрузки новой файловой пары. При загрузке новых файлов, ID образца должен совпадать с загружаемыми файлами, соответствующими этому образцу.
4. Состав уже загруженных файлов.

3.4.3. Вкладка “Workflow”

Пример экрана:



Экран содержит следующие элементы:

1. Кнопка возвращения на предыдущий экран.
2. Блок для переключения вкладок
3. Кнопка для добавления нового исследования. При добавлении нового исследования нужно указать Тип исследования. На текущий момент доступен только Germline mapping и название исследования.
4. Состав уже созданных исследований, по конкретному пациенту, содержащую следующую информацию
 - a. Номер исследования - порядковый номер исследования
 - b. Название исследования
 - c. Образцы - название образцов
 - d. Статус - Статус исследования
 - e. Параметры запуска.

3.5. Экран “Исследования”

Экран содержит две вкладки:

- “Список исследований” - список всех исследований созданных в системе.
- “Доступные виды исследований” - Содержащую в себе загруженные файлы образца + возможность загрузить новые

3.5.1. Вкладка “Список исследований”

Пример экрана:

№	Название	Образцы	Статус	Параметры запуска
83	ZD-210122-demo-test	ZD210122	Succeeded	nextflow run https://github.com/zenomeplatform/nf-germline-mapping -r dev -profile yandex --config http://dev.inunio.ru/configs/83.config -c /etc/nextflow/nextflow.conf --accessKey \$ACCESS_KEY --secretKey \$SECRET_KEY
82	ZD210122-GED-E080A56_full_test	ZD210122	Succeeded	nextflow run https://github.com/zenomeplatform/nf-germline-mapping -r dev -profile yandex --config http://dev.inunio.ru/configs/82.config -c /etc/nextflow/nextflow.conf --accessKey \$ACCESS_KEY --secretKey \$SECRET_KEY
81	test_aug_10-05-22	ZD21013C	Succeeded	nextflow run https://github.com/zenomeplatform/nf-germline-mapping -r dev -profile yandex --config http://dev.inunio.ru/configs/81.config -c /etc/nextflow/nextflow.conf --accessKey \$ACCESS_KEY --secretKey \$SECRET_KEY
80	testName-10	ZD210122	Succeeded	nextflow run https://github.com/zenomeplatform/nf-germline-mapping -r dev -profile yandex --config http://dev.inunio.ru/configs/80.config -c /etc/nextflow/nextflow.conf --accessKey \$ACCESS_KEY --secretKey \$SECRET_KEY
79	test_aug_7	ZD21013C	Succeeded	nextflow run https://github.com/zenomeplatform/nf-germline-mapping -r dev -profile yandex --config http://dev.inunio.ru/configs/79.config -c /etc/nextflow/nextflow.conf --accessKey \$ACCESS_KEY --secretKey \$SECRET_KEY
78	test_aug_6	test_aug_2	Failed	nextflow run https://github.com/zenomeplatform/nf-germline-mapping -r dev -profile yandex --config http://dev.inunio.ru/configs/78.config -c /etc/nextflow/nextflow.conf --accessKey \$ACCESS_KEY --secretKey \$SECRET_KEY
77	aug_test-5	aug_test	Failed	nextflow run https://github.com/zenomeplatform/nf-germline-mapping -r dev -profile yandex --config http://dev.inunio.ru/configs/77.config -c /etc/nextflow/nextflow.conf --accessKey \$ACCESS_KEY --secretKey \$SECRET_KEY
76	aug_test_1	aug_test	Failed	nextflow run https://github.com/zenomeplatform/nf-germline-mapping -r dev -profile yandex --config http://dev.inunio.ru/configs/76.config -c /etc/nextflow/nextflow.conf --accessKey \$ACCESS_KEY --secretKey \$SECRET_KEY
75	aug_test_3	nf-germline-mapping1	Failed	nextflow run https://github.com/zenomeplatform/nf-germline-mapping -r dev -profile yandex --config http://dev.inunio.ru/configs/75.config -c /etc/nextflow/nextflow.conf --accessKey \$ACCESS_KEY --secretKey \$SECRET_KEY
74	test_aug-2	Ex123Test	Failed	nextflow run https://github.com/zenomeplatform/nf-germline-mapping -r dev -profile yandex --config http://dev.inunio.ru/configs/74.config -c /etc/nextflow/nextflow.conf --accessKey \$ACCESS_KEY --secretKey \$SECRET_KEY

Экран содержит следующие элементы:

1. Блок для переключения вкладок
2. Кнопка для добавления нового исследования. При добавлении нового исследования нужно указать Тип исследования. На текущий момент доступен только Germline mapping и название исследования.
3. Состав уже созданных исследований, по конкретному пациенту, содержащую следующую информацию
 - a. Номер исследования - порядковый номер исследования
 - b. Название исследования
 - c. Образцы - название образцов
 - d. Статус - Статус исследования
 - e. Параметры запуска.

3.5.2. Вкладка “Доступные виды исследований”

Вкладка содержит доступные виды исследований. На текущий момент доступен только один биоинформатический протокол.

3.6. Экран “Информация об исследовании”

Содержит в себе результаты исследования (исполнения протокола) и содержит в себе общую информацию, результаты выполнения протокола и логи.

Инфо Результаты Лог

Список результатов Биоинформатический отчёт Анализ покрытия Workflow

fastqc
raw
fastqc_ZD210122-GED-E080QHE-MG-300002755-1-59_raw_logs
[ZD210122-GED-E080QHE-MG-300002755-1-59-F_fastqc.html](#)
[ZD210122-GED-E080QHE-MG-300002755-1-59-F_fastqc.zip](#)
[ZD210122-GED-E080QHE-MG-300002755-1-59-R_fastqc.html](#)
[ZD210122-GED-E080QHE-MG-300002755-1-59-R_fastqc.zip](#)

trimmed
fastqc_ZD210122-GED-E080QHE-MG-300002755-1-59_trimmed_logs
[ZD210122-GED-E080QHE-MG-300002755-1-59-F_trim_fastqc.html](#)
[ZD210122-GED-E080QHE-MG-300002755-1-59-F_trim_fastqc.zip](#)
[ZD210122-GED-E080QHE-MG-300002755-1-59-R_trim_fastqc.html](#)
[ZD210122-GED-E080QHE-MG-300002755-1-59-R_trim_fastqc.zip](#)

multiqc_postalignment_report
[multiqc_report.html](#)
ZD210122
[multiqc_report_ZD210122.html](#)

multiqc_prealignment_report
[multiqc_report.html](#)
ZD210122-GED-E080QHE-MG-300002755-1-59
[multiqc_report_ZD210122-GED-E080QHE-MG-300002755-1-59.html](#)

pipeline_info
[execution_report.html](#)
[execution_timeline.html](#)

4. Основные операции пользователя

4.1. Вход на сайт и авторизация

Для входа на сайт запустите браузер и наберите в адресной строке URL- адрес сайта <https://dev.inunio.ru>

На экране появится приглашение для входа и авторизации.

Вход

Логин *

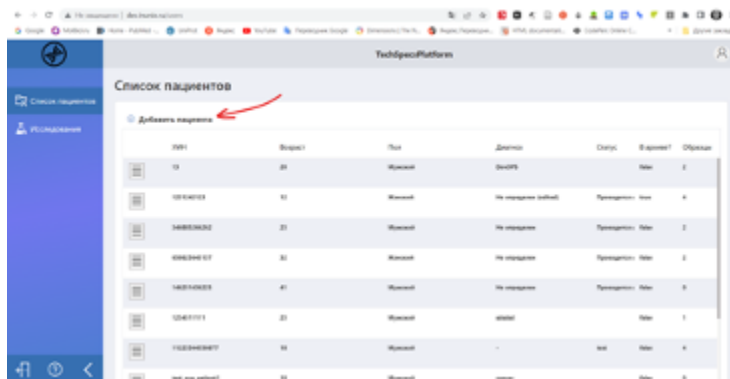
Пароль *

Войти

В открывшемся окне необходимо ввести данные учетной записи.

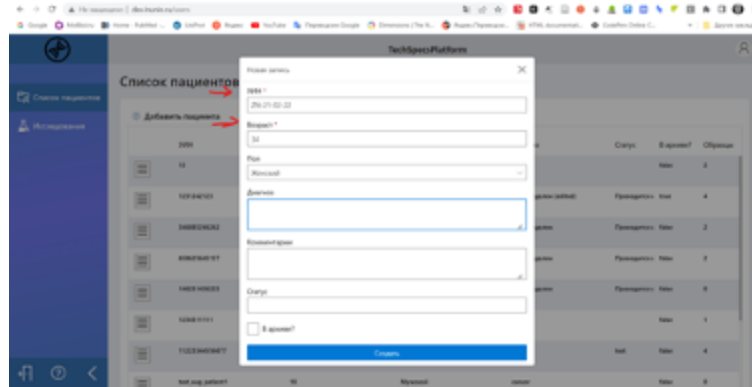
4.2. Добавление нового пациента

Шаг 1. Для добавления нового пациента необходимо перейти во вкладку «Список пациентов» и нажать на кнопку «Добавить пациента».



Шаг 2. В открывшемся окне необходимо заполнить следующую информацию:

- УИН (Уникальный идентификационный номер) – обязательно
- Возраст – обязательно
- Пол – опционально
- Диагноз – опционально
- Комментарий – опционально
- Статус – опционально
- Указать что карточка пациента находится в Архиве.



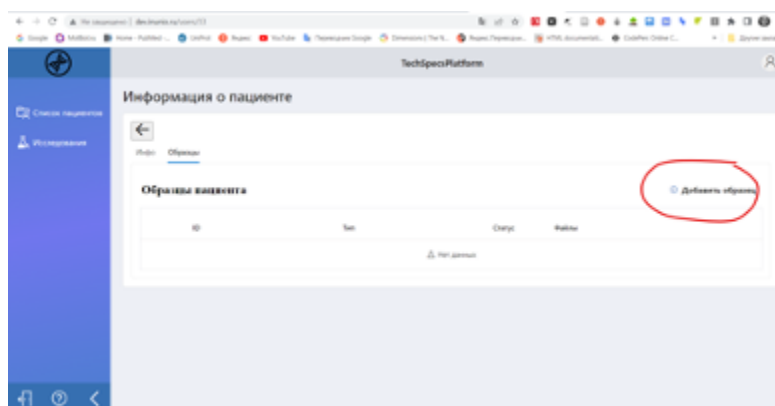
Шаг 3. Добавленная запись появится в общем списке. Данные можно скорректировать или удалить с помощью вызова контекстного меню и нажать «Редактировать» или «Удалить»



Шаг 4. Для перехода в карточку пациента можно нажать на его идентификационный номер. Карточка пациента содержит все информацию о пациент и загруженные образцы.

4.3. Добавление образцов

Шаг 1. Для добавление образцов необходимо перейти в карточку пациент и выбрать вкладку «Образцы» На данной вкладке доступно добавление образцов с последующей загрузкой fastq-файлов. Далее необходимо нажать «Добавить образец».



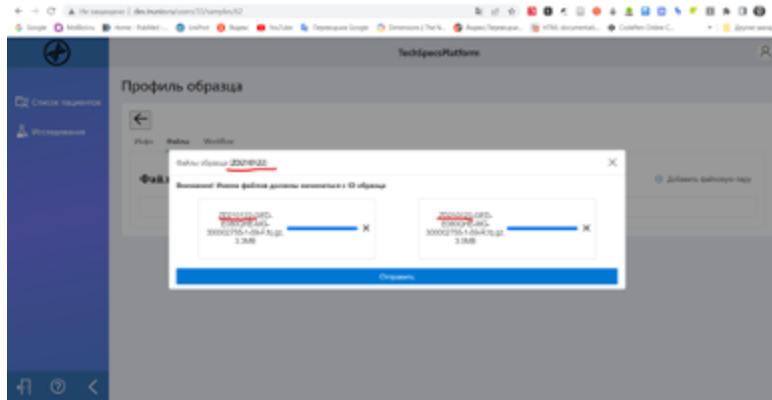
Шаг 2. По образцу который требуется добавить необходимо заполнить следующие данные:

- Номер образца – обязательно
- Тип образца – опционально
- Тип считывания – опционально
- Устройство для забора – опционально
- Тип биоматерала – опционально
- Дата получения – опционально
- Секвенатор -опционально
- Комментарий – опционально.

Это важно: При добавлении нового образца, ID должен совпадать с загружаемыми файлами, соответствующими этому образцу.

Шаг 4. Для загрузки файлов необходимо перейти в карточку образца и открыть вкладку «файлы»

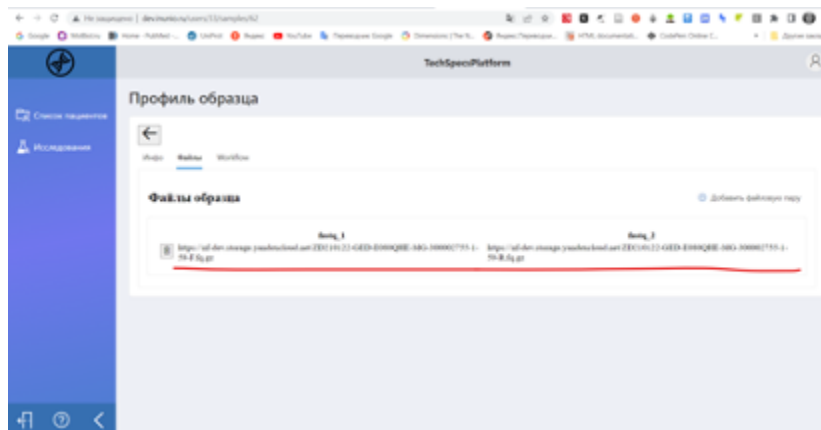
Шаг 5. Необходимо загрузить файлы с помощью кнопки «Добавить файловую пару».



Это важно: пожалуйста проверьте, что название файлов соответствует названию образца.

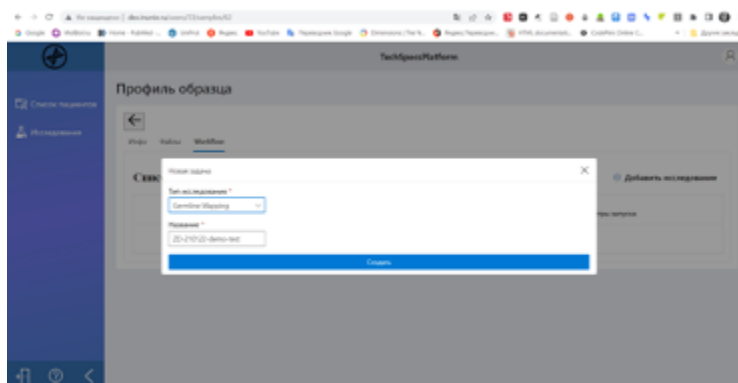
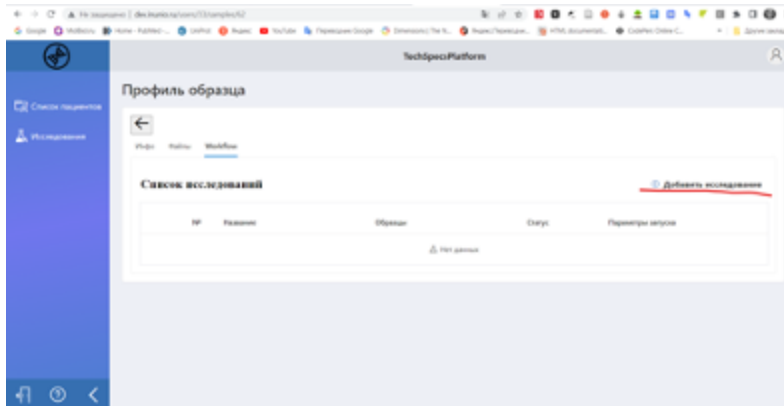
Правила наименования файлов в данном случае определяются разработанным протоколом анализа, подробнее о правилах названия можно прочитать по ссылке: <https://github.com/zenomeplatform/nf-germline-mapping> (см. Naming convention)

Шаг 6. После загрузки fastq файлов, они хранятся в облаке и могут быть скачаны по соответствующей ссылке:

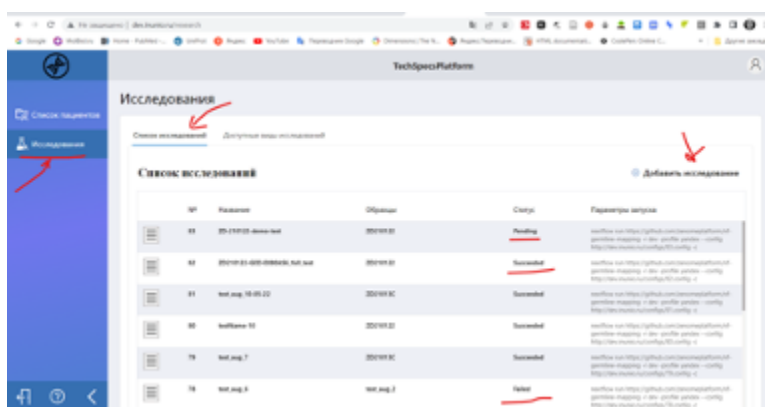


4.4. Проведение исследования

Шаг 1. После того, как пациент и соответствующие ему образцы/данные созданы и загружены в облачную платформу, можно выбрать доступный тип исследований и запустить анализ с помощью кнопки «Добавить исследование»:

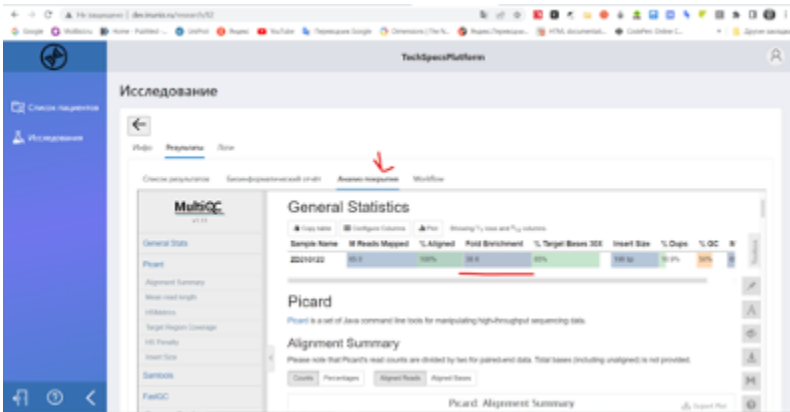
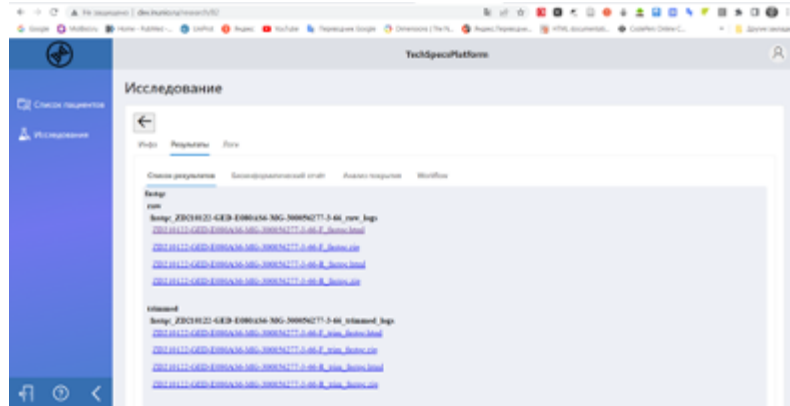


Шаг 2. В процессе проведения исследования можно отслеживать статус исследования, а также изучать информацию о ранее проведенных исследованиях во вкладке «Исследования»



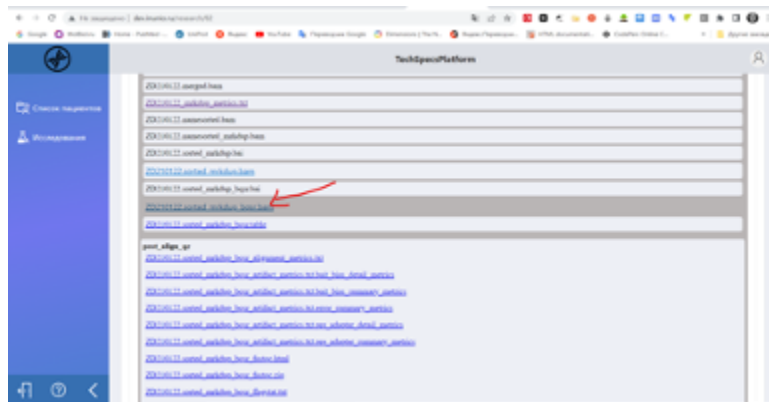
4.5. Работа с результатами исследования:

Шаг 1: После успешного завершения исследования, перейдите в карточку исследования, где можно посмотреть файлы и отчеты, содержащие в себе результаты исследования:

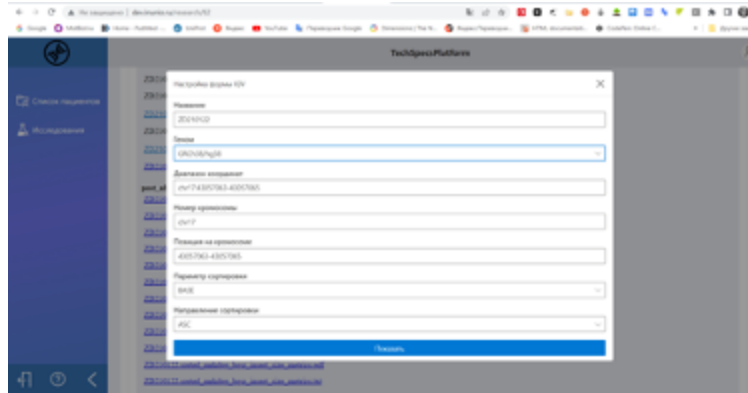


4.6. Работа с BAM-Фалами

Шаг 1: Для работы с BAM-файлов в IGV-геномном браузере непосредственно из интерфейса системы кликните на нужный BAM-файл.



Шаг 2. Укажите необходимый регион для просмотра



Можем проверить наличие интересующего геномного варианта (в данном примере инсерция в гене BRCA1)

