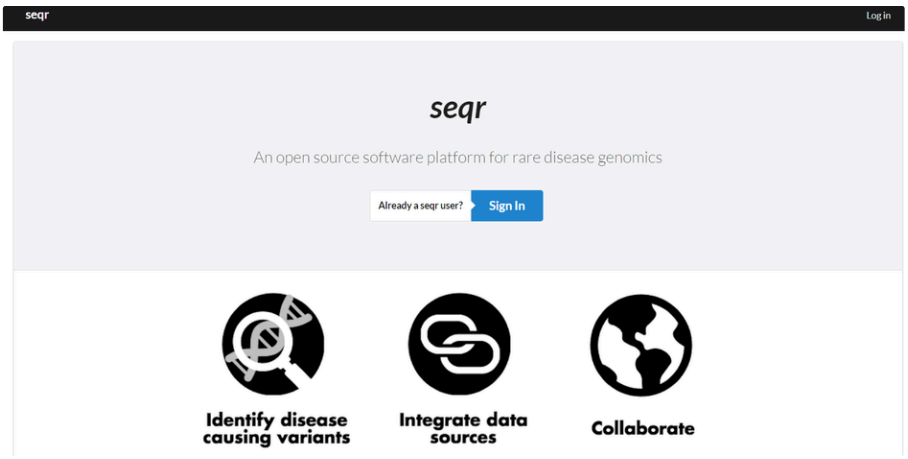
seqr documentation

Seqr is an open-source genomic data interpretation platform. It provides a user-friendly web interface for rare disease diagnosis and gene discovery. It allows uploading VCF (variant call format) files and annotates them with relevant information for Mendelian analysis. The platform supports variant filtering and searching in gene lists to prioritise variants meeting certain criteria.



Uploading VCF file

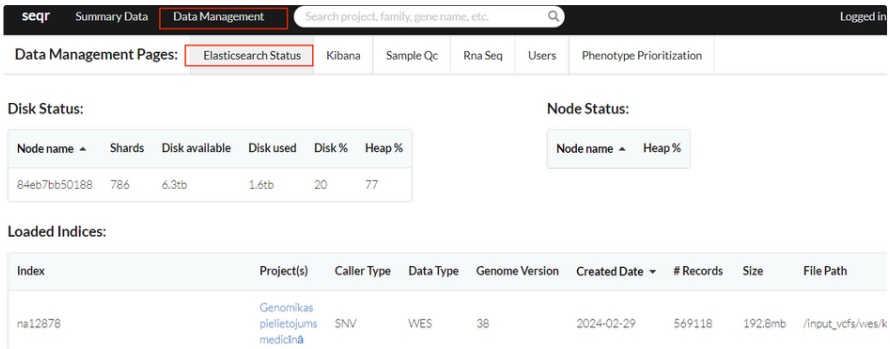
```
#Before executing the commands, ensure that the pipeline-runner container is active
#For more information on setting up an on-premises seqr instance, refer to
#https://github.com/broadinstitute/seqr/blob/master/deploy/LOCAL_INSTALL.md

BUILD_VERSION=38                # can be 37 or 38
SAMPLE_TYPE=WES                 # can be WES or WGS
INDEX_NAME=your-dataset-name    # the desired index name to output.

INPUT_FILE_PATH=${FILE_PATH}/${FILENAME}

docker-compose exec pipeline-runner load_data.sh $BUILD_VERSION $SAMPLE_TYPE $INDEX_NAME $INPUT_FILE_PATH
```

⚠ Make sure your VCF is annotated with fields AC, AN and AF, if not, run bcftools +fill-tags:

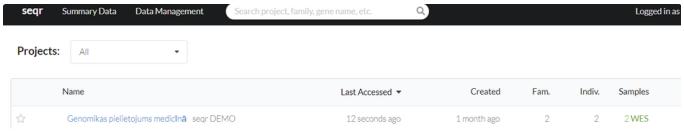
```
bcftools +fill-tags file.vcf -- -t AC,AN,AF
```

Adding loaded data to a project

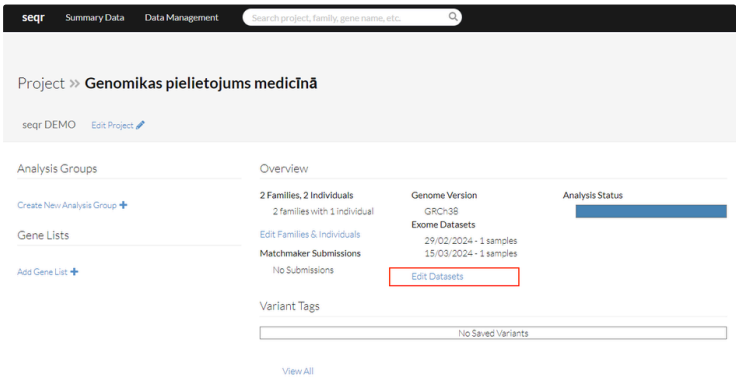
You can find loaded data indices under the 'Elasticsearch Status' tab in the 'Data Management' section. Admin rights are required to load data into projects.



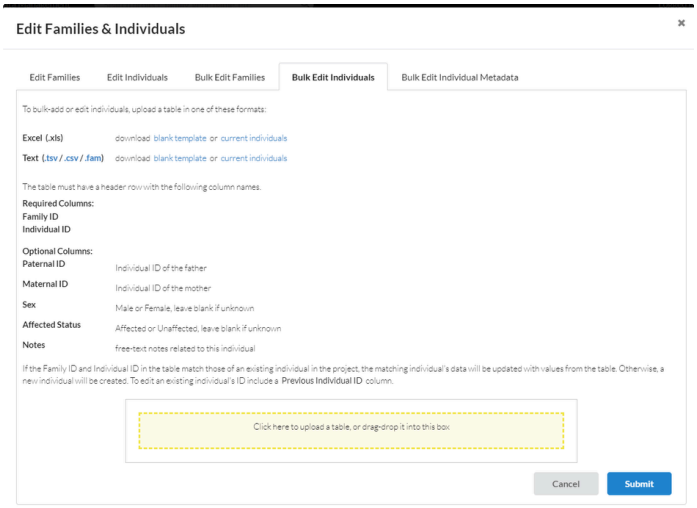
To add data to specific project, on the start page click on the project you want to load data into:



Click on “Edit Families & Individuals“:



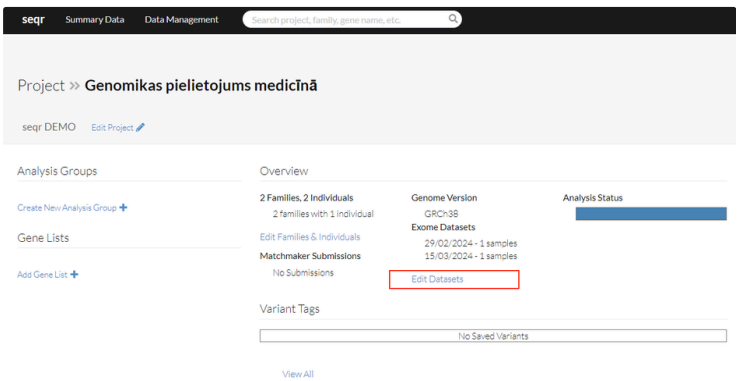
Then, under “Bulk edit Individuals“ upload a .csv file that is formatted as in the instructions and press “Submit“:

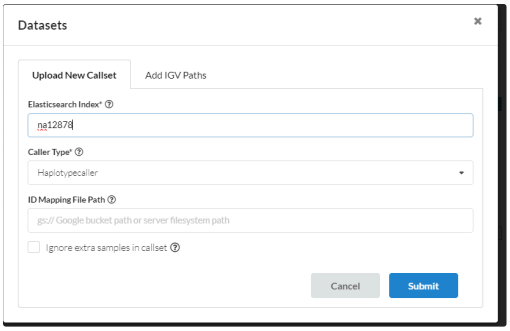


⚠ Make sure that **“Family ID”** is the same as the index specified when loading the data and can be seen under Data Management tab.

“Individual ID” should be the same as the sample ID specified in the VCF file.

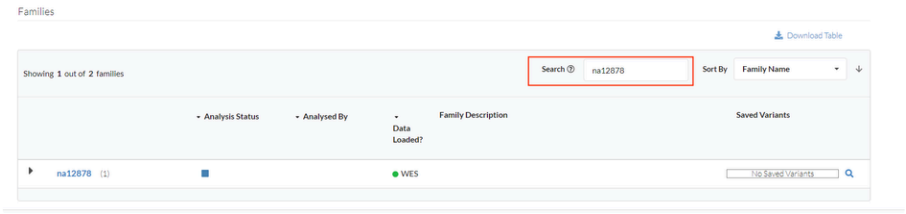
Then, click on “Edit Datasets“, enter the index to add data to the project and press “Submit“:



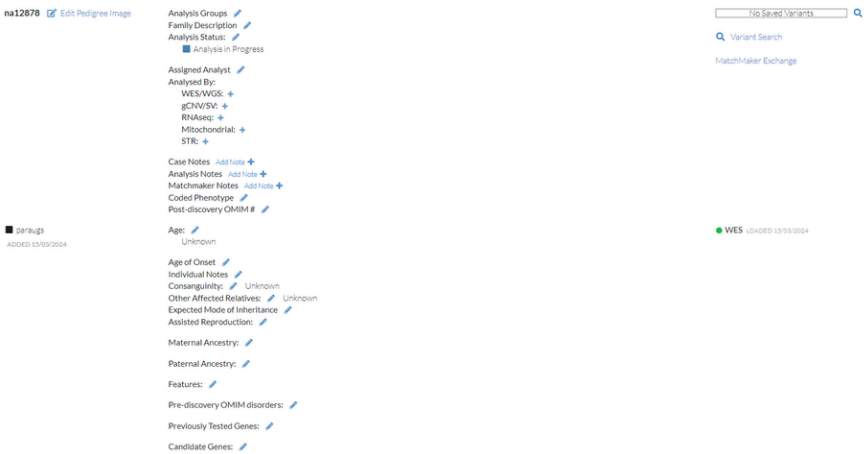


Adding patient information

On the project page, scroll down to “Families“ and find sample by typing index into the search bar:



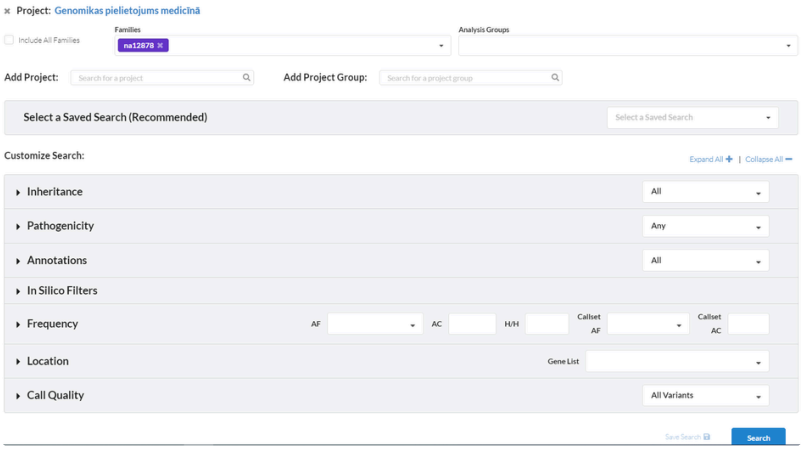
Click on the index name and open the sample page. Here, you can add information about the patient's phenotype and associate data analysts.



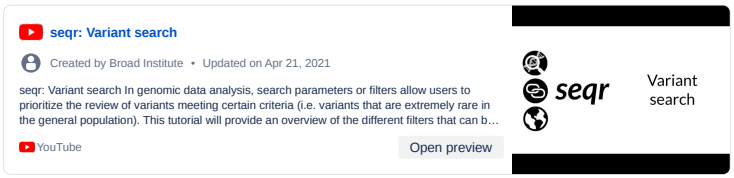
Variant search & filtering

To start filtering variants click on “Variant Search“ on patient’s page:

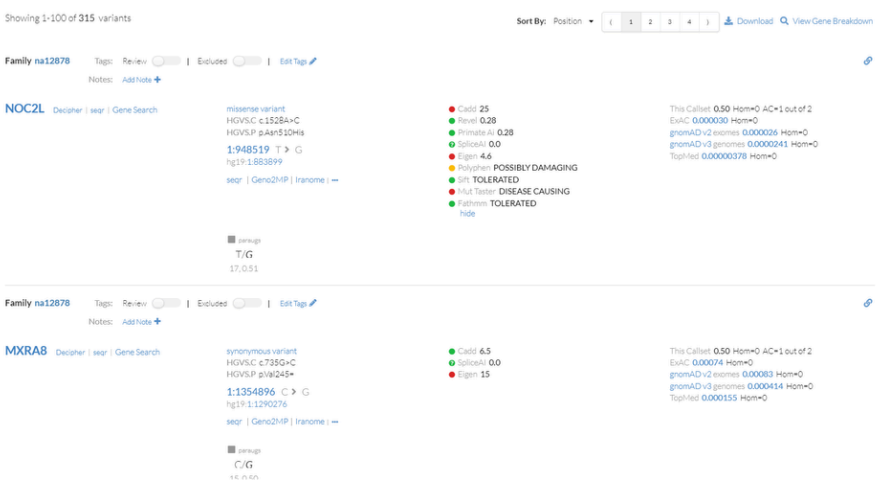
Then, adjust the filters and gene lists before pressing 'Search'.



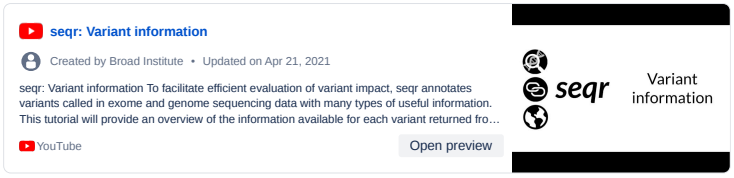
More on how to apply filters can be found in this video:



After applying filters you will see the filtered variants:

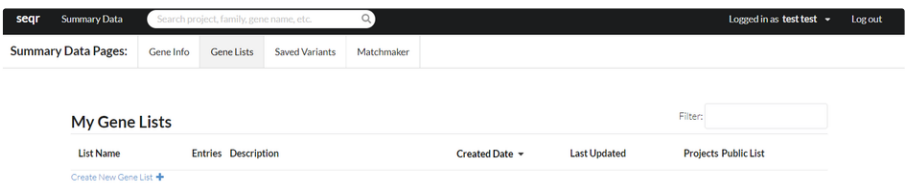


More on information that can be seen about each variant can be found in this video:

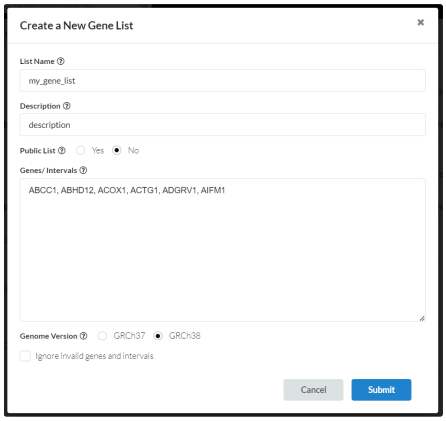


Adding gene lists

To add a gene list go to Summary Data → Gene Lists → Create New Gene List:



Add genes or intervals, name, description and the genome version:



More information about seqr:

Broad Institute's official seqr YouTube playlist: seqr: An open source software platform for rare disease genomics

seqr GitHub page: GitHub - broadinstitute/seqr: web-based analysis tool for rare disease genomics