Genomics England

Finding participants based on genotypes

Emily Perry and Eleni Christodoulou

11th June 2024

Presenters

Emily Perry
Research Engagement Manager

Eleni Christodoulou
Solutions Architect at Lifebit

Data security 

- This training session will include data from the GEL Research Environment
- As part of your IG training you have agreed to not distribute these data in any way
- You are not allowed to:
    - Invite colleagues to watch this training with you
    - Take any screenshots or videos of the training
    - Share your webinar link (we will remove anyone who is here twice)
- We will record this training and distribute the censored video afterwards

Questions

All your microphones are muted

Use the Zoom Q&A to ask questions

Upvote your favourite questions: if we are short on time we will prioritise those with the most votes

Questions

Matthieu Vizuete-Forster
Bioinformatician - Research Services

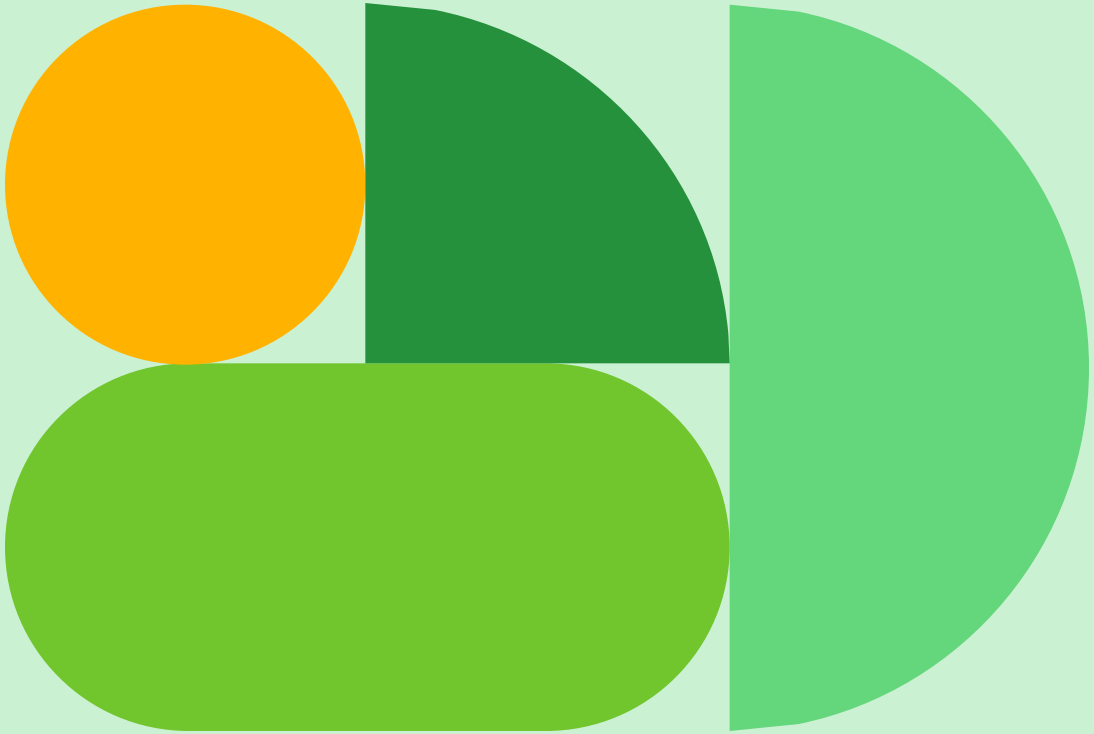
Asier Gonzalez-Uriarte
Bioinformatician - Research Services

Alex Ho
Bioinformatician - Research Services

Rosie Davies
Principal User Success Manager at Lifebit

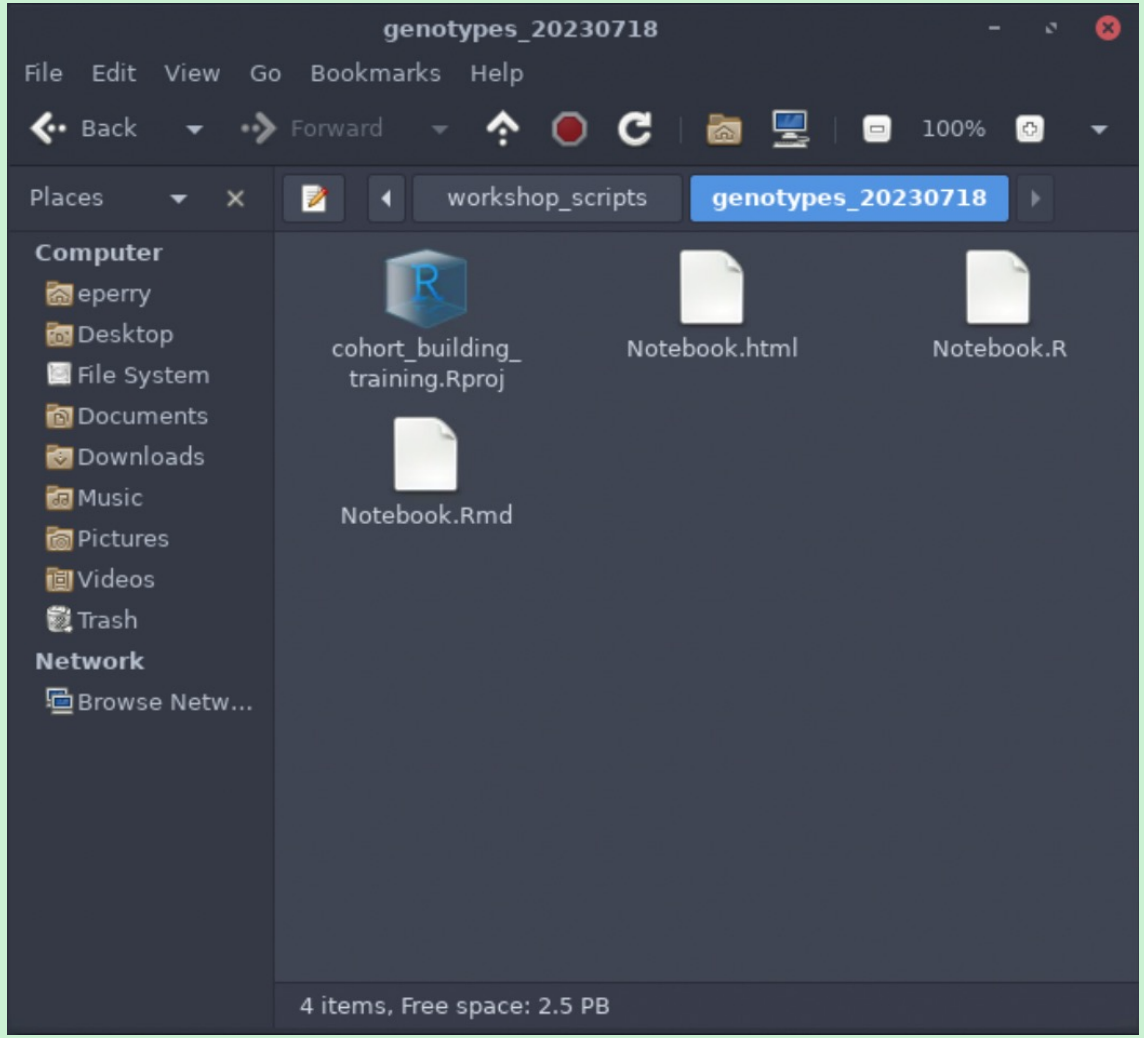
Agenda

1. Introduction and admin
2. Genome assembly
3. LabKey tables of variant genotypes
4. Finding genotypes with IVA and Cohort Browser
5. The Small Variant and Structural Variant workflows
6. Aggregated variant files
7. When/why you would use each method
8. Help and questions



Materials

- Slides and video will be sent out to you after the session
- Scripts available in `/gel_data_resources/example_scripts/workshop_scripts/genotypes_2024`



2. Genome Assembly

100,000 Genomes Project

Rare disease

GRCh37 (hg19)
GRCh38 (hg38)

Cancer

Germline GRCh37 (hg19)
Somatic GRCh37 (hg19)
Germline GRCh38 (hg38)
Somatic GRCh38 (hg38)

100,000 Genomes Project

Rare disease

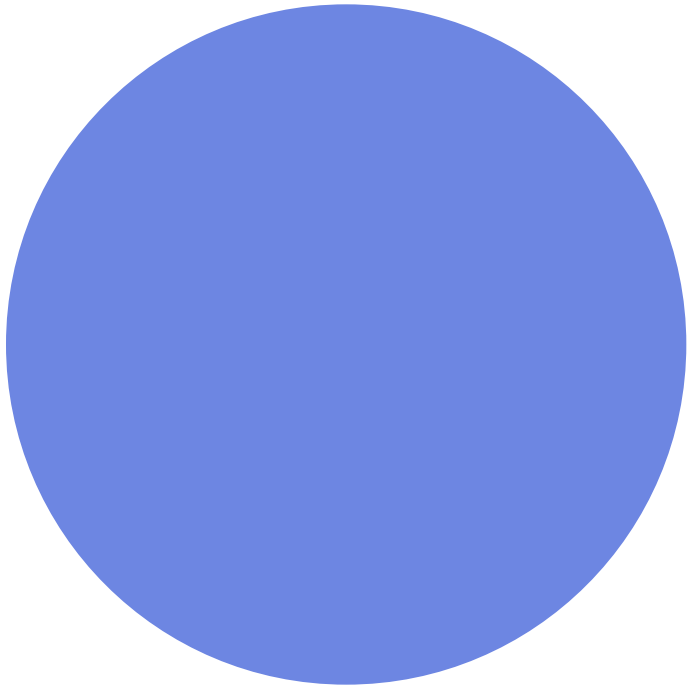
GRCh37 (hg19)
GRCh38 (hg38)

Cancer

~~Germline GRCh37 (hg19)~~
~~Somatic GRCh37 (hg19)~~
Germline GRCh38 (hg38)
Somatic GRCh38 (hg38)

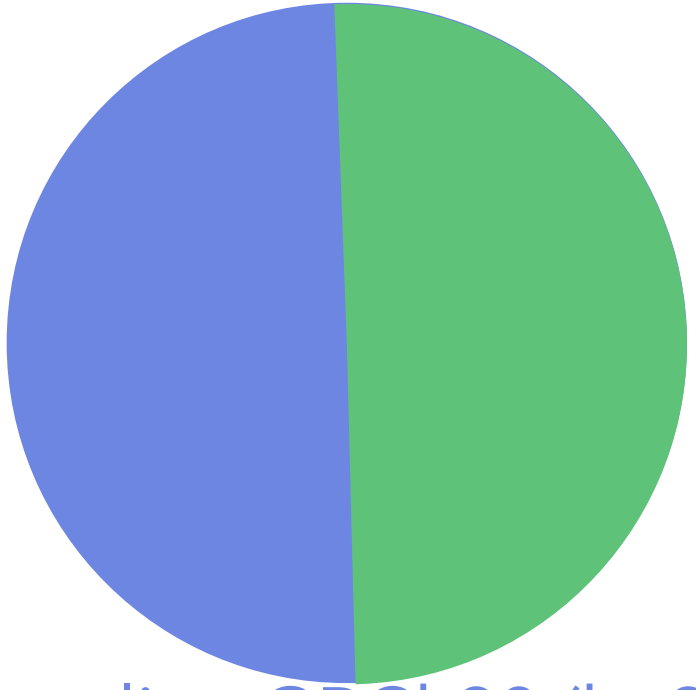
NHS GMS

Rare disease



GRCh38 (hg38)

Cancer



Germline GRCh38 (hg38)

Somatic GRCh38 (hg38)

Genome assembly coordinates

	GRCh37 (hg19)	GRCh38 (hg38)
ZAR1L	13:32,877,837-32,889,481	chr13:32,303,699- 32,315,363
ENST00000345108.6:c.931T>C	13:32,878,051	chr13:32,303,914

Converting between assemblies

Inside the RE

- Liftover tool on HPC
- Chain files in `public_data_resources`

Outside the RE

- Ensembl Assembly converter
- UCSC Liftover

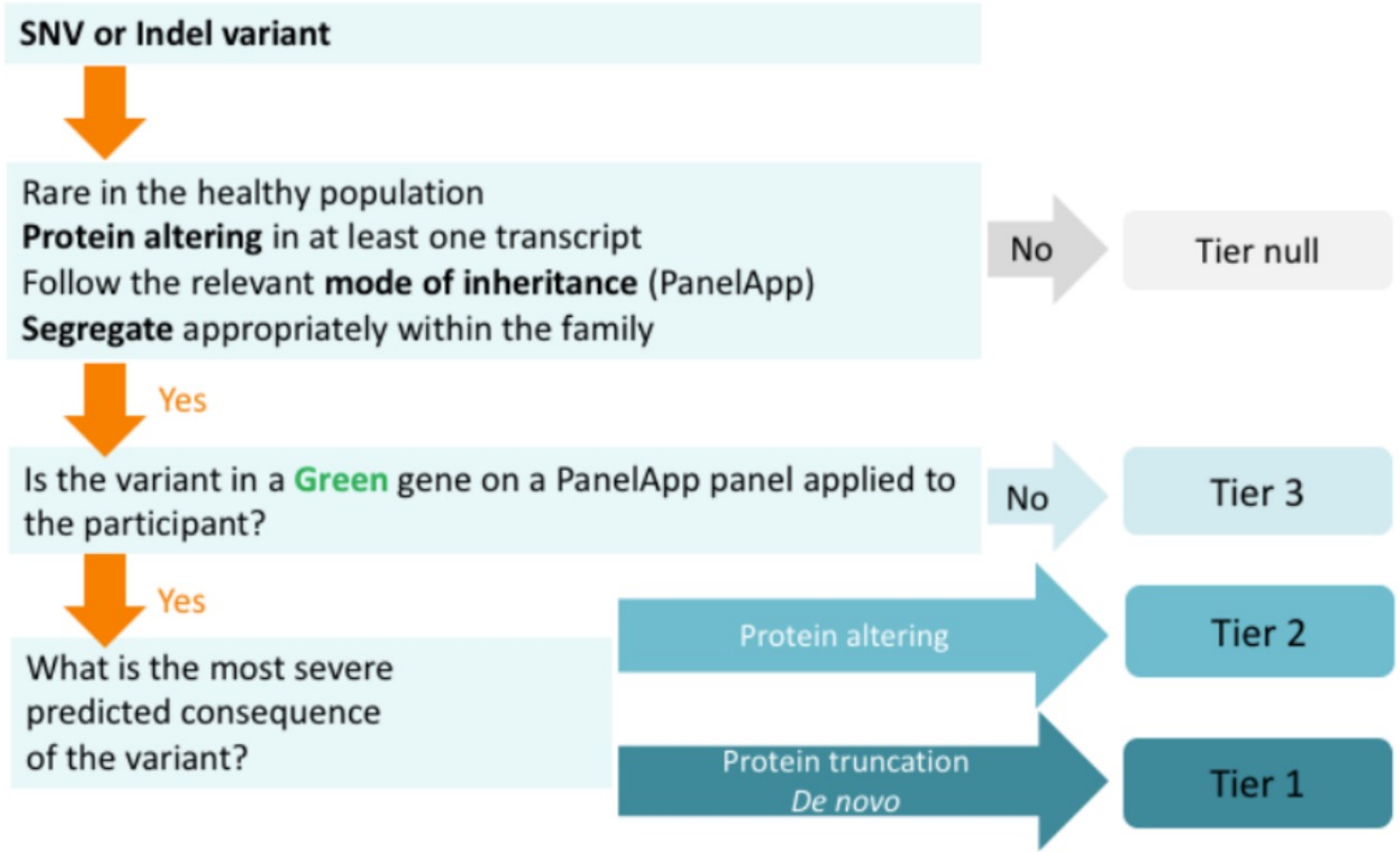
3. LabKey tables of variant genotypes

LabKey

- Participant details and family relationships
- Sample details
- Genomic file locations
- Clinical data
- Bioinformatics analysis results
  - Rare disease tiering
  - Cancer tiering
  - Exomiser

https://re-docs.genomicsengland.co.uk/labkey/

Rare disease tiering



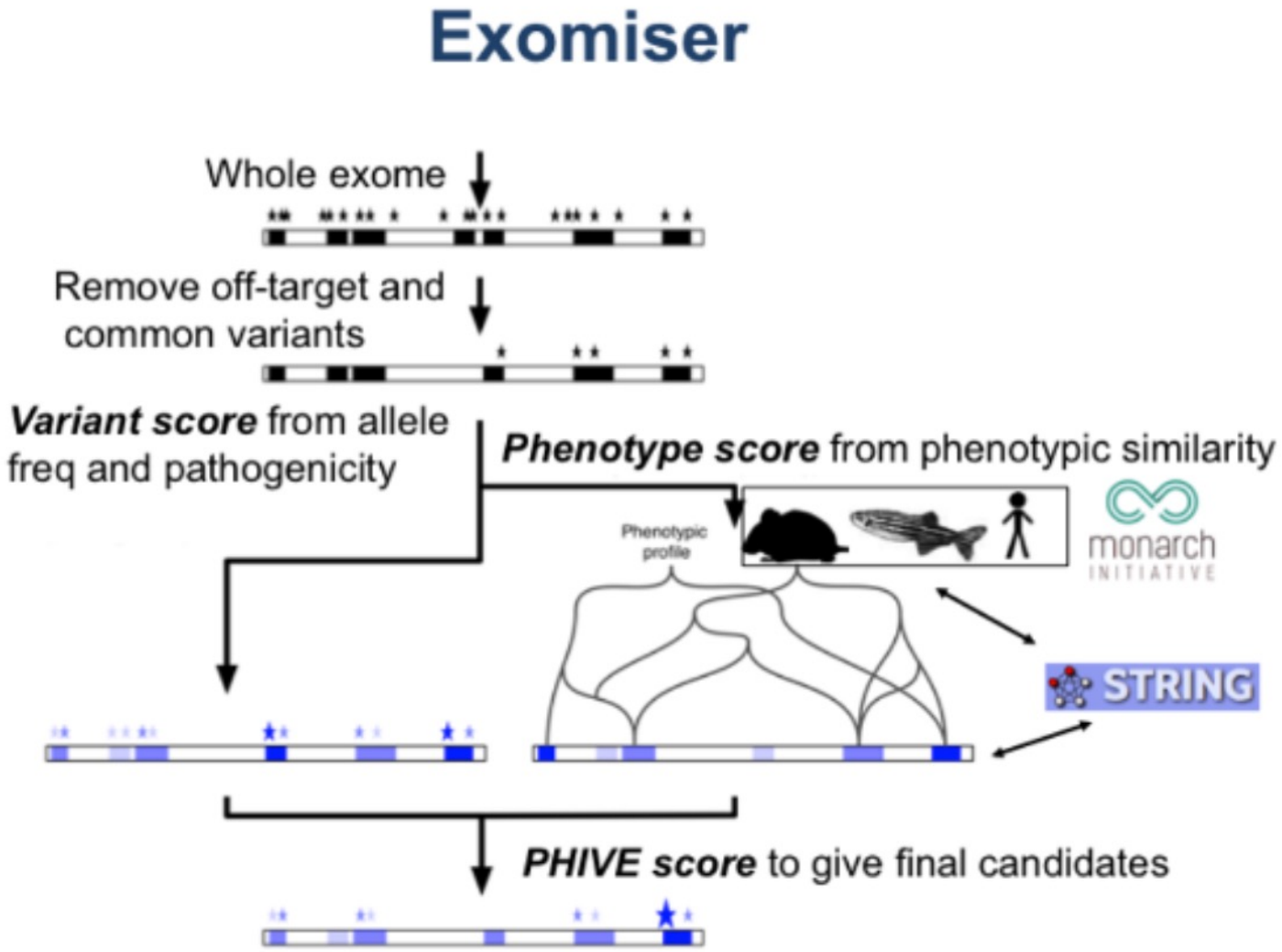
https://re-docs.genomicsengland.co.uk/tiering/

Rare disease tiering based on PanelApp genes

List ↑	Entity	Reviews	Mode of inheritance	Details
Filter Entities				8 Entities
Green	ATP1A3	1 review 1 green	MONOALLELIC, autosomal or pseudoautosomal, NOT imprinted	**Sources** • Expert Review • Expert Review Green **Phenotypes** • 601338 • 614820 **Tags**
Green	DFNB59	2 reviews 1 green	BIALLELIC, autosomal or pseudoautosomal	**Sources** • Expert Review • Expert Review Green **Phenotypes** • 610219 **Tags** new-gene-name
Green	OPA1	2 reviews 1 green	MONOALLELIC, autosomal or pseudoautosomal, NOT imprinted	**Sources** • Eligibility statement prior genetic testing • Expert Review Green **Phenotypes** • Optic atrophy 1, OMIM:165500 • Optic atrophy plus syndrome, OMIM:125250 **Tags**
Green	OTOF	1 review 1 green	BIALLELIC, autosomal or pseudoautosomal	**Sources** • Expert Review Green • Radboud University Medical Center, Nijmegen **Phenotypes** • 601071 **Tags**
Amber	DIAPH3	3 reviews 1 red	BOTH monoallelic and biallelic, autosomal or pseudoautosomal	**Sources** • Expert Review Amber • Radboud University Medical Center, Nijmegen **Phenotypes** • Auditory neuropathy, autosomal dominant, 1, 609129 **Tags**

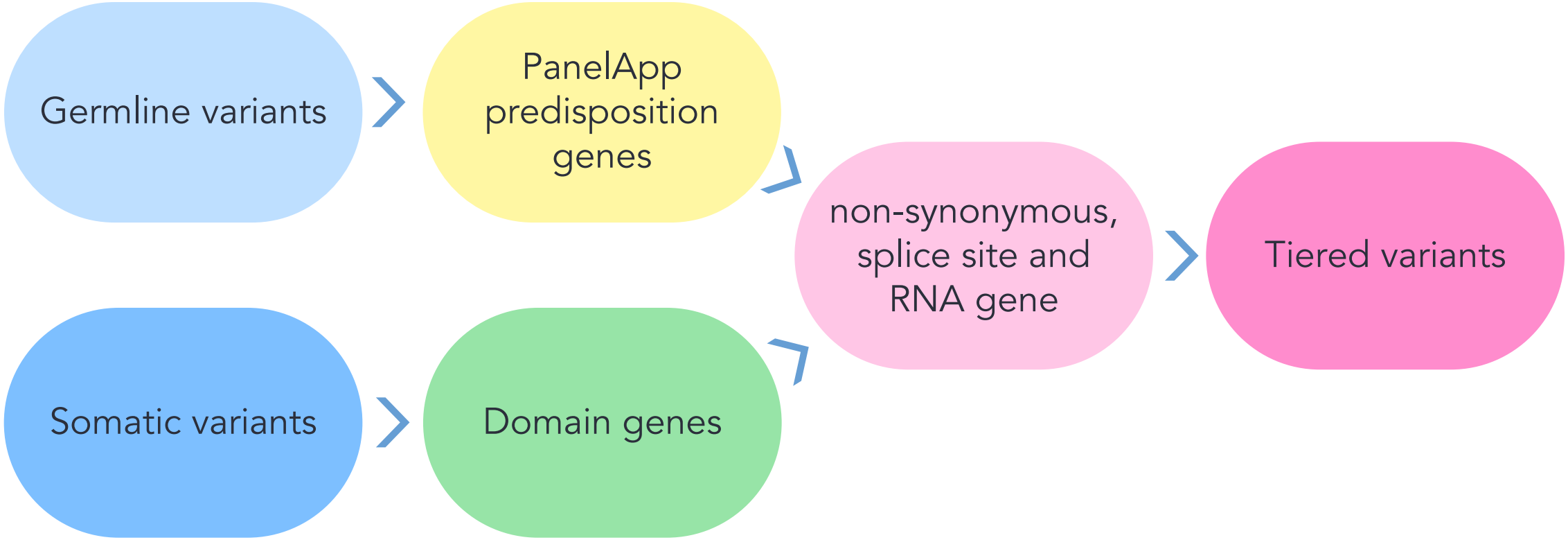
https://re-docs.genomicsengland.co.uk/panelapp/

Rare disease Exomiser



https://re-docs.genomicsengland.co.uk/exomiser/

Cancer tiering



https://re-docs.genomicsengland.co.uk/cancer_tiering/

100k tiering

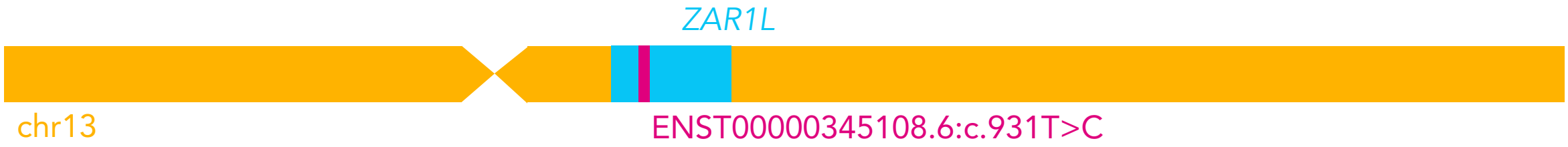
Rare disease

GRCh37 (hg19)
GRCh38 (hg38)

Cancer

Germline GRCh38 (hg38)
Somatic GRCh38 (hg38)

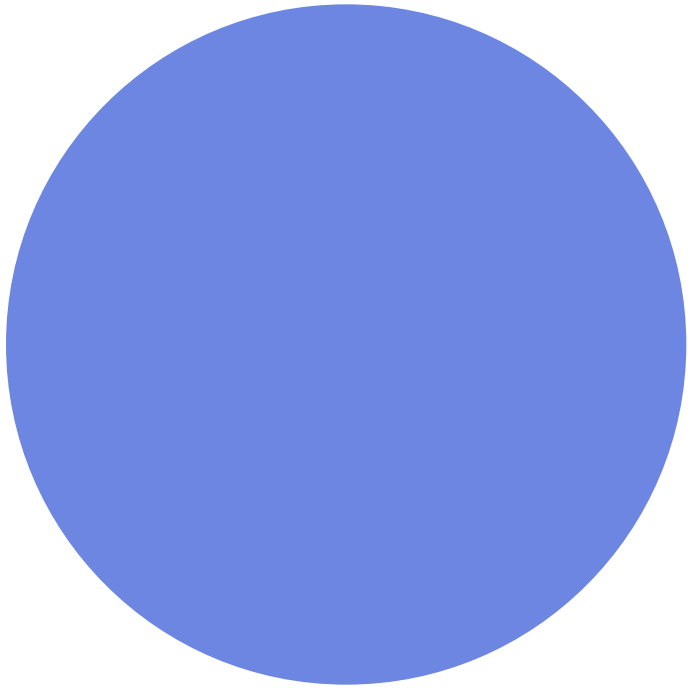
Tiering tables genome assembly – 100k



Search by	
gene	Should find all filter-passing variants in the gene on either assembly
coordinate	You must also specify the genome assembly
HGVS (exomiser only)	Should find all filter-passing variants that match your string on either assembly

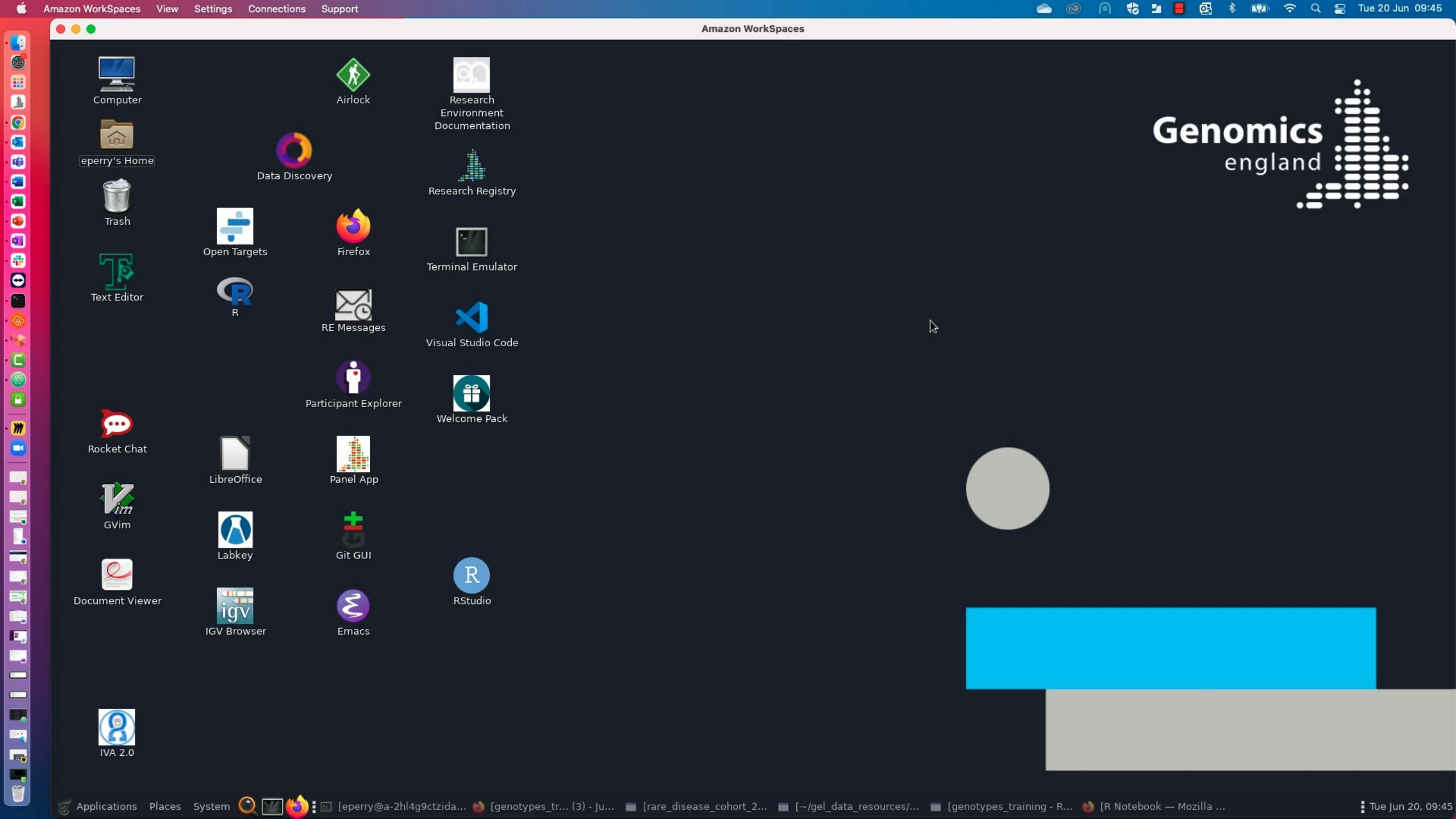
NHS GMS tiering

Rare disease



GRCh38 (hg38)

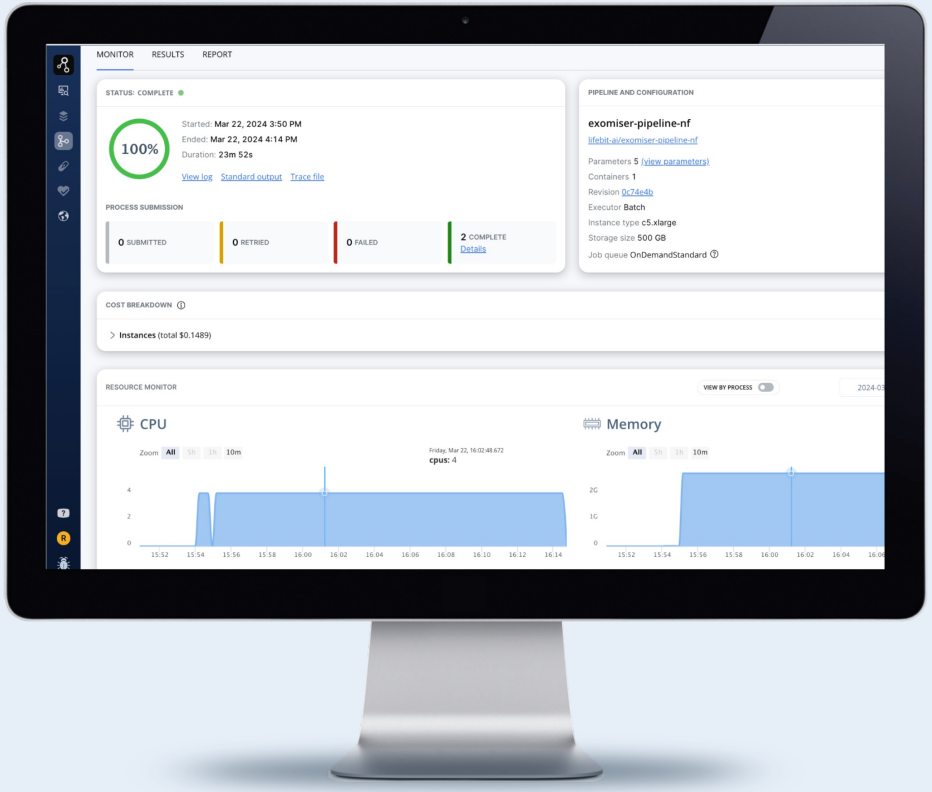
Variant data in LabKey demo


Amazon WorkSpaces View Settings Connections Support
Tue 20 Jun 09:45
Amazon WorkSpaces
Computer
Airlock
Research Environment Documentation
eperry's Home
Data Discovery
Research Registry
Trash
Open Targets
Firefox
Terminal Emulator
Text Editor
R
RE Messages
Visual Studio Code
Participant Explorer
Welcome Pack
Rocket Chat
LibreOffice
Panel App
GVim
Labkey
Git GUI
Document Viewer
RStudio
IGV Browser
Emacs
IVA 2.0
Genomics england
Applications Places System
[eperry@a-2hl4g9ctzida...
[genotypes_tr... (3) - Ju...
[rare_disease_cohort_2...
[~/gel_data_resources/...
[genotypes_training - R...
[R Notebook — Mozilla ...
Tue Jun 20, 09:45

Genomics England CloudOS

CloudOS is a Cloud (AWS) based tool within the RE which provides you with a secure way to access and analyse Genomics England data with scalable, on-demand compute.

Offers a complimentary solution alongside other RE tools.

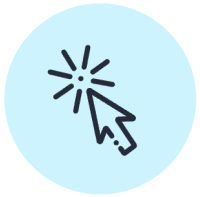


How CloudOS may help streamline your analysis

On demand scalable compute, so no need to queue for resources.

Collaborative workspaces for your project, making it easy to securely share tools and results.

Fully end-to-end, from data to analytics in one platform.



Simple and intuitive user interface.

Easy to bring in data, packages and tools.

CloudOS considerations

AWS S3 storage in place of Weka, so there are some limitations on data availability

- **Clinical:** 100k and Covid
- **Omic:** VCFs, aggVCFs & some BAMS/CRAMs.

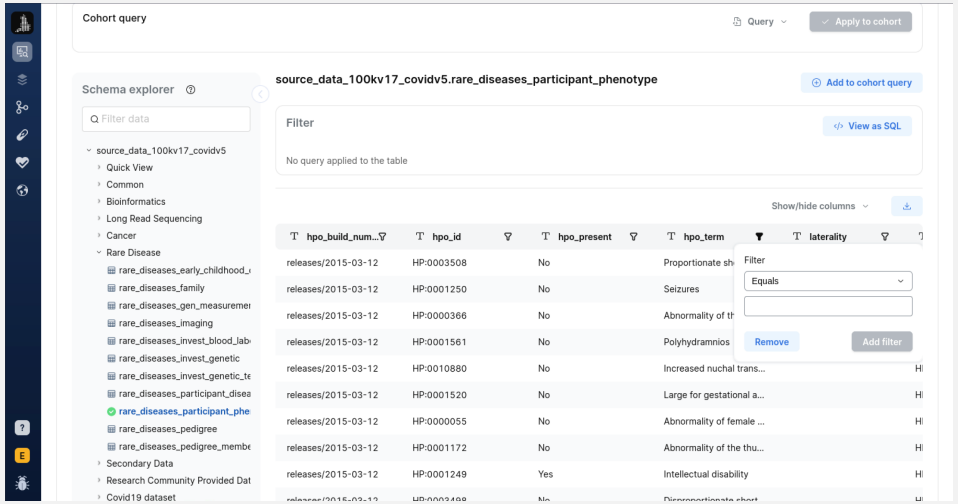
Genomics England provides free $1000 worth of compute per month to academic researchers

CloudOS Cohort Browser – Source tables

Filter participants using multiple phenotypic tables

Pull genomic file locations for filtered cohort

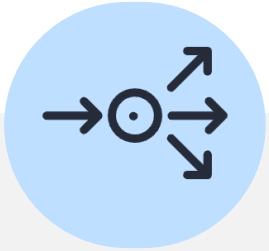
Build Phenotype files for downstream analysis



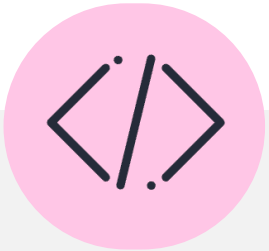
https://lifebit.atlassian.net/wiki/spaces/CD/pages/813106901/Build+a+cohort+using+source+data

Variant data in Cohort Browser demo

LabKey API



Combine queries between tables



Work in a variety of programming languages (support for Python and R) using SQL queries

Replicate queries between releases and analyses

Work locally and on the HPC

https://re-docs.genomicsengland.co.uk/labkey_api/

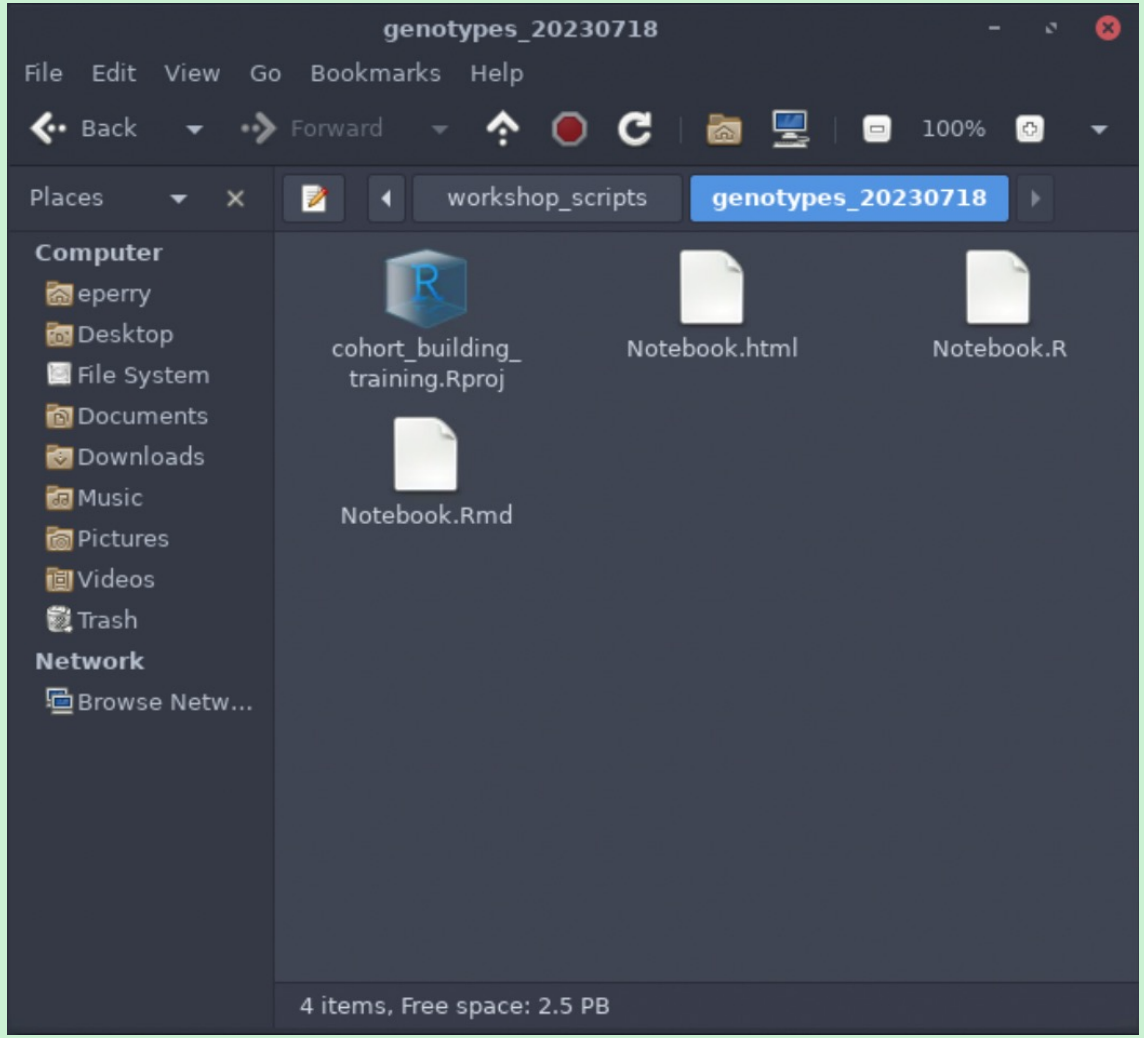
LabKey .netrc

- You can access the same data via the LabKey API as you can through other means
- You will need to configure access to the LabKey API with your username and password
    - In your home directory
    - On the HPC
- You do this by editing a file called .netrc

https://re-docs.genomicsengland.co.uk/labkey_api/

Materials

- Slides and video will be sent out to you after the session
- Scripts available in `/gel_data_resources/example_scripts/workshop_scripts/genotypes_2024`



To launch the notebooks

https://re-docs.genomicsengland.co.uk/enable_rstudio/

https://re-docs.genomicsengland.co.uk/hpc_jupyter/

Variant data in LabKey API demo

genotypes_tr... (3) - JupyterLab — Mozilla Firefox

Jupyter Lab on Helix - Ge | R Notebook | genotypes_tr... (3) - Jupy

127.0.0.1:9537/lab/tree/genotypes_training.ipynb

File Edit View Run Kernel Tabs Settings Help

GENOTYPES_TRAINING.IPYNB

rd_cohort_building_trainir | genotypes_training.ipynb | cohort_building_training.ip

Markdown | Python 3 (ipykernel)

Finding participants by genotypes in Python

This notebook will walk you through finding participants by genotypes. You are welcome to copy/paste any code from this notebook for your own scripts/notebooks.

Contents:

Use the LabKey API

Import Python modules you need

[32]:
```
import numpy as np
import functools
import labkey
import pandas as pd
```

Helper function to access the LabKey API with

Simple 0 3 Python 3 (ipykernel) | Idle Mode: Command Ln 1, Col 1 genotypes_training.ipynb

R Notebook — Mozilla Firefox

R Notebook

file:///nas/weka.gel.zone/pgen_int_work/BRS/emily/genotypes_2023/genotyp

R Notebook

Finding participants by genotypes in R

This notebook will walk you through finding participants by genotypes. You are welcome to copy/paste any code from this notebook for your own scripts/notebooks.

Contents:

Use the LabKey API

Import R libraries you need

```
library(tidyverse)
```

```
## ── Attaching core tidyverse packages ──────────────────────── tidyverse 2.0.0 ──
## ✔ dplyr     1.1.2     ✔ readr     2.1.4
## ✔ forcats   1.0.0     ✔ stringr   1.5.0
## ✔ ggplot2   3.4.2     ✔ tibble    3.2.1
## ✔ lubridate 1.9.2     ✔ tidyr     1.3.0
## ✔ purrr     1.0.1
## ── Conflicts ────────────────────────────────────────── tidyverse_conflicts() ──
## ✖ dplyr::filter() masks stats::filter()
## ✖ dplyr::lag()    masks stats::lag()
## ℹ Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```
library(Rlabkey)
```

4. Finding genotypes with IVA and Cohort Browser

Interactive variant analysis (IVA)

- Point-and-click interface to explore variants
- Filter by loci, consequences, population frequencies and inheritance
- Find participant genotypes

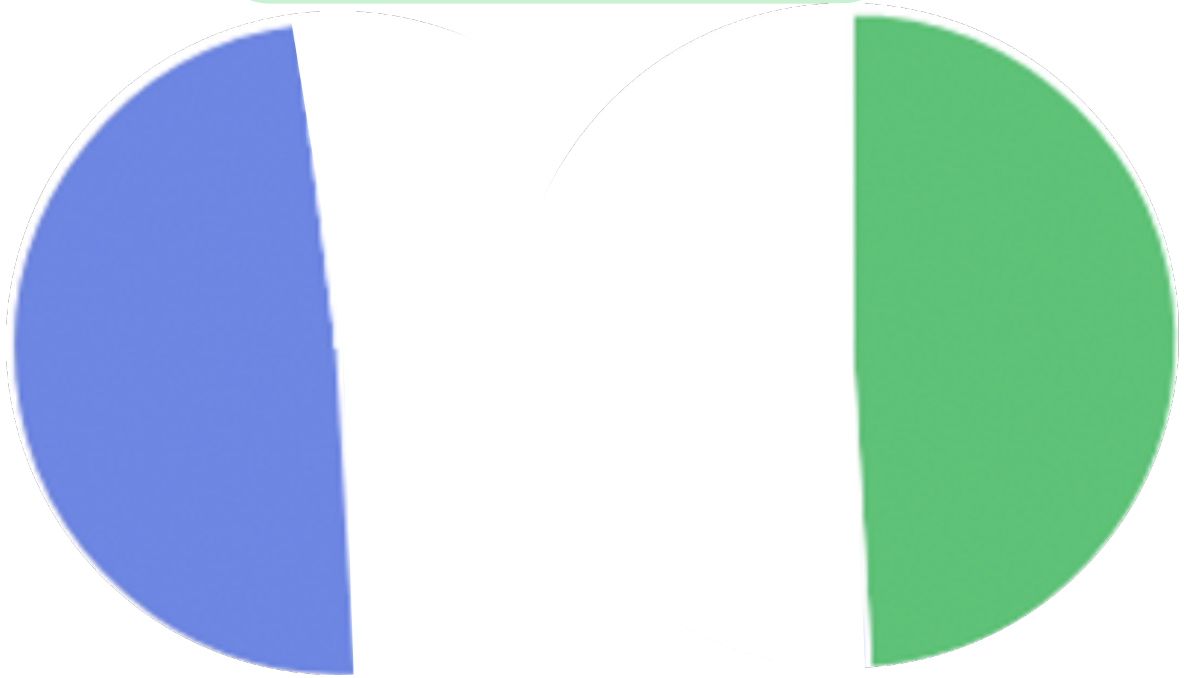
https://re-docs.genomicsengland.co.uk/iva/

100k in IVA

Rare disease

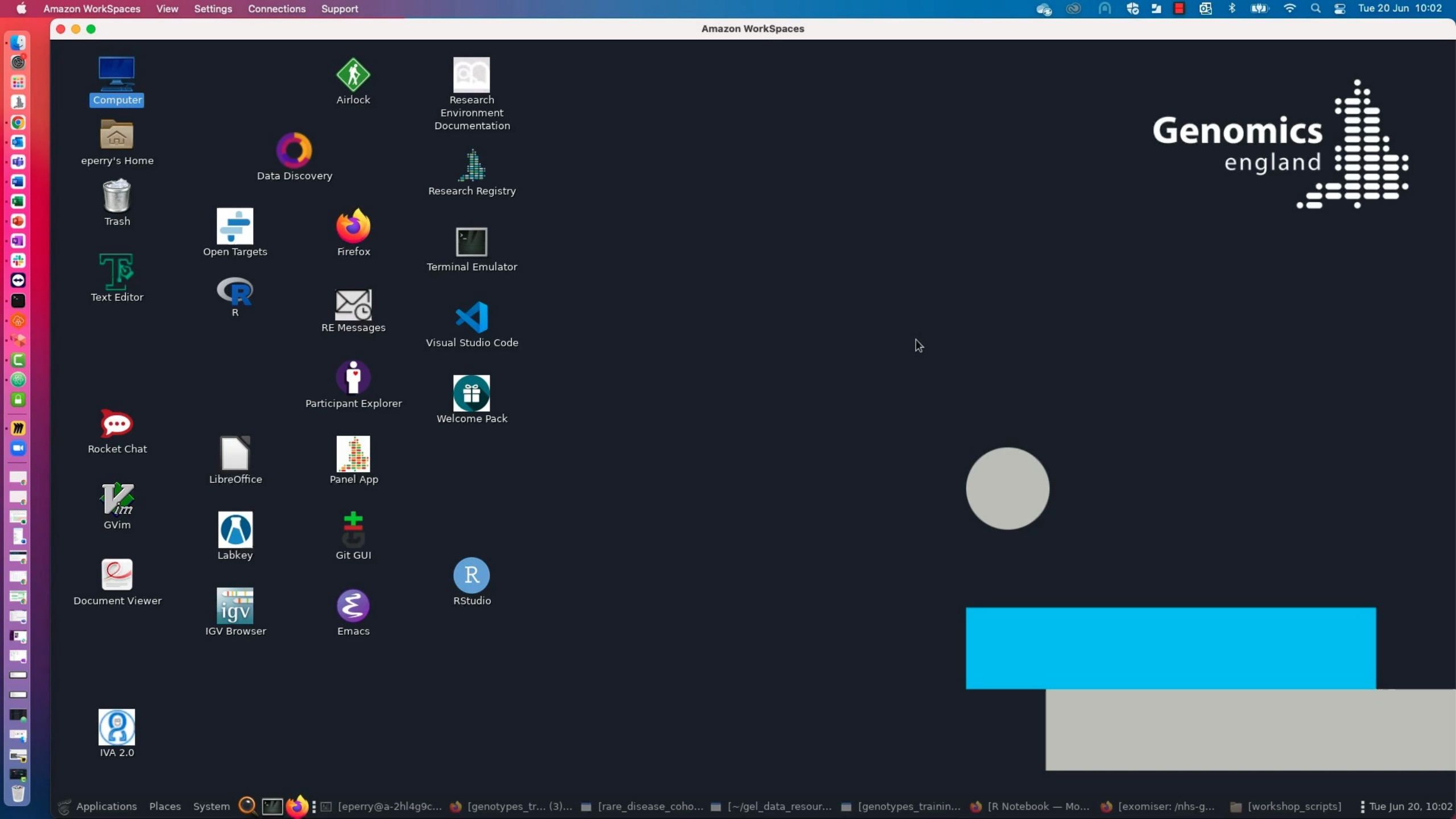
GRCh37 (hg19)
GRCh38 (hg38)

Cancer



Germline GRCh38 (hg38)
Somatic GRCh38 (hg38)

IVA demo


Amazon WorkSpaces View Settings Connections Support
Tue 20 Jun 10:02
Amazon WorkSpaces
Computer
Airlock
Research Environment Documentation
eperry's Home
Data Discovery
Research Registry
Trash
Open Targets
Firefox
Terminal Emulator
Text Editor
R
RE Messages
Visual Studio Code
Participant Explorer
Welcome Pack
Rocket Chat
LibreOffice
Panel App
GVim
Labkey
Git GUI
Document Viewer
RStudio
IGV Browser
Emacs
IVA 2.0
Genomics england
Applications Places System
[eperry@a-2hl4g9c...
[genotypes_tr... (3)...
[rare_disease_coho...
[~/gel_data_resour...
[genotypes_trainin...
[R Notebook — Mo...
[exomiser: /nhs-g...
[workshop_scripts]
Tue Jun 20, 10:02

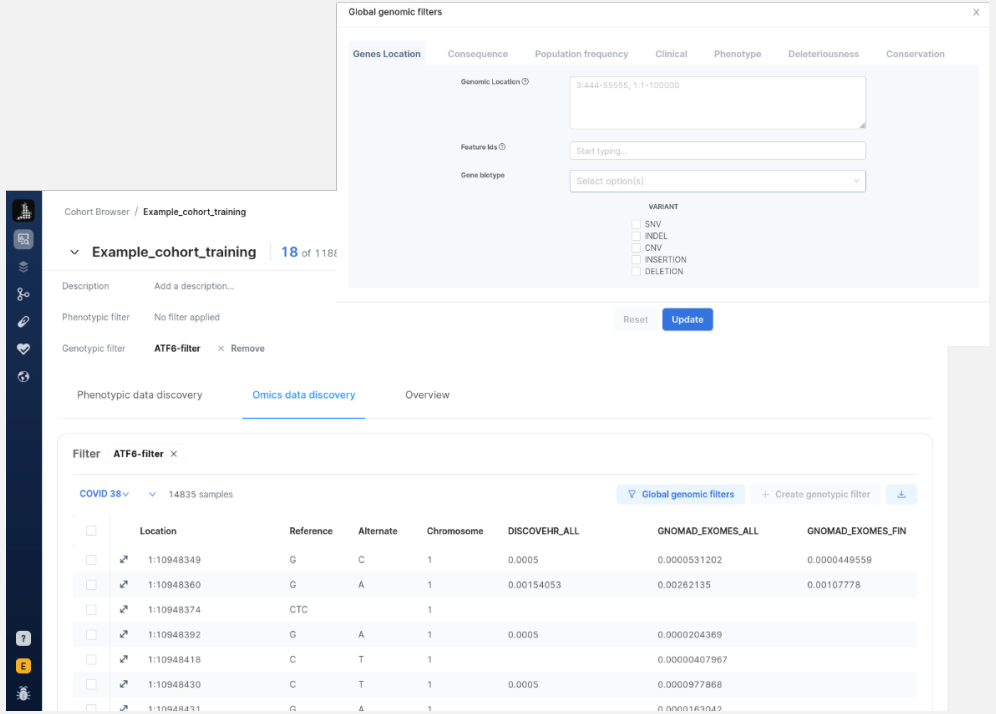
CloudOS Cohort Browser – genotypes

Filter genotypes and phenotypes in one interface

Point-and-click interface to explore variants

Filter by loci, consequences, population frequencies and inheritance.

Find participant genotypes.



https://lifebit.atlassian.net/wiki/spaces/CD/pages/813107122/Filter+cohort+by+genomic+data

CloudOS omics demo

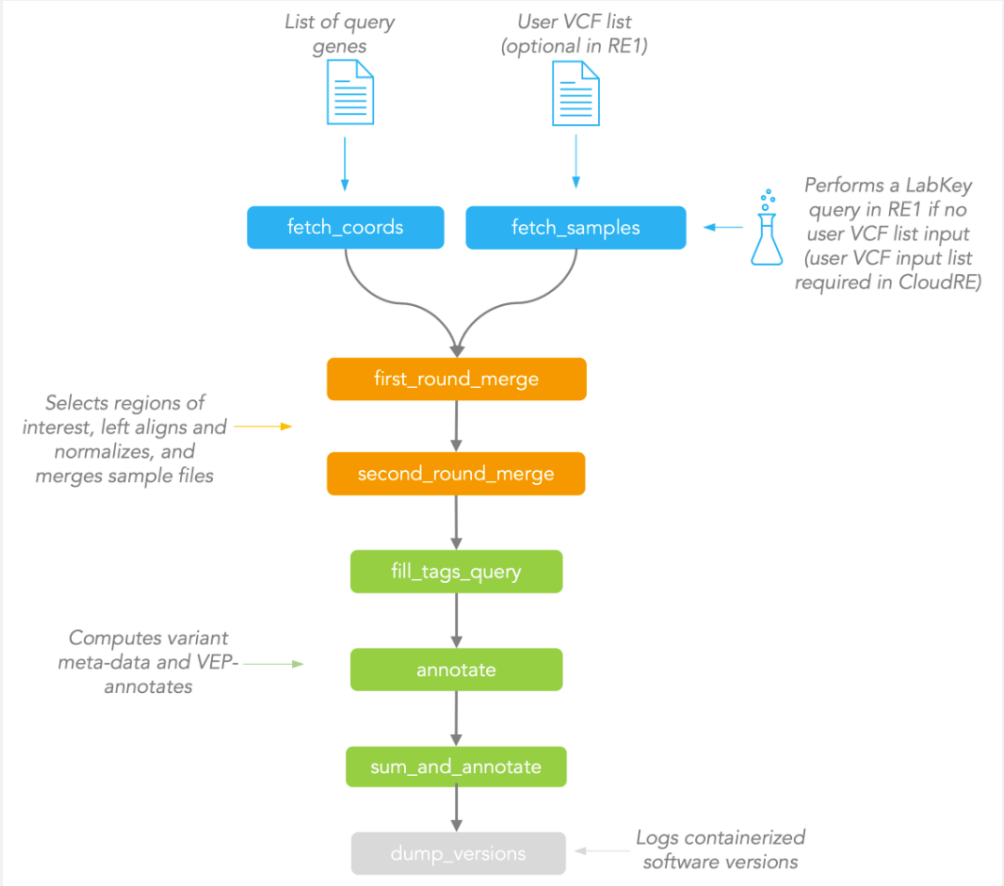
5. The Small Variant and Structural Variant workflows

Small variant workflow

Submit a list of genes

Find all short variants in these genes

Get 100k participants with these variants



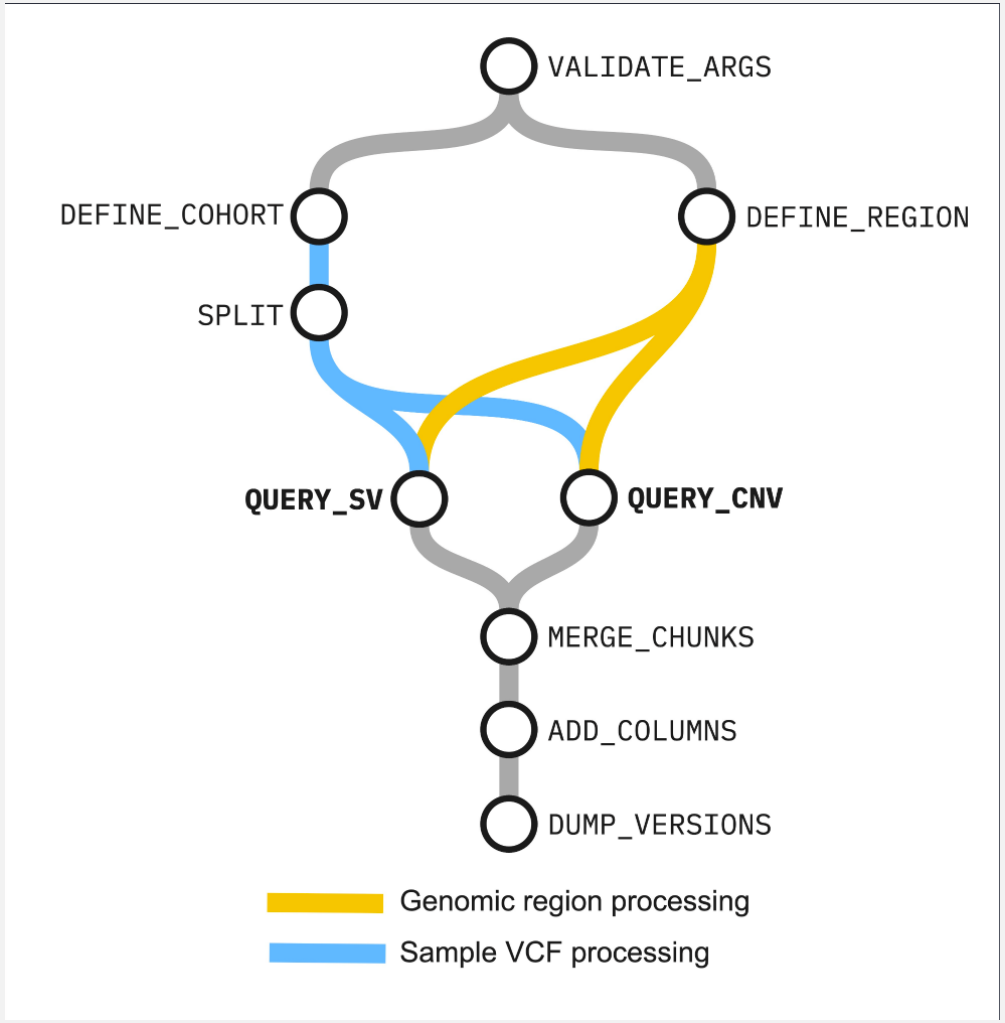
https://re-docs.genomicsengland.co.uk/small_variant/

Structural variant workflow

Submit a list of genes or regions

Find all structural variants overlapping these genes

Get 100k participants with these variants



https://re-docs.genomicsengland.co.uk/structural_variant/

Workflows genome assembly – 100k

Search by	
gene	Should find all variants in the gene(s) on either assembly
coordinates (structural only)	You must also specify the genome assembly

Running workflows on the HPC demo

Amazon WorkSpaces View Settings Connections Support

Wed 24 Apr 10:47

Amazon WorkSpaces

genotypes_tr... - JupyterLab — Mozilla Firefox

genotypes_tr... - Jupyter

127.0.0.1:8387/lab/tree/genotypes_training.ipynb

File Edit View Run Kernel Tabs Settings Help

Launcher | genotypes_training.ipynb

Markdown | Python 3 (ipykernel)

Filter files by name

/

Name	Last Modified
small_variant	2 days ago
structural_...	2 days ago
cohort_buil...	2 days ago
genotypes...	a day ago
Notebook....	20 hours ago
Notebook.R	2 years ago
Notebook....	20 hours ago

604 rows × 7 columns

Running workflows on the HPC

Small variant workflow

- `ssh` to the HPC with your usual login credentials
- `cd` into your working directory
- Make and cd into your working directory
  `mkdir small_variant_demo`
  `cd small_variant_demo`
- Copy the Small Variant workflow submission script into your folder
  `cp /pgen_int_data_resources/workflows/rdp_small_variant/main/submit.sh .`
- Make the file `gene_list.txt` and add your list of genes to it
  `vi gene_list.txt`
  Add `SMC3` to the file
  `Esc :wq`
- Edit the submission script:
  `vi submit.sh`
  Change the `project_code` to your code Change the `gene_input` to `gene_list.txt`
  `Esc :wq`
- Run the workflow
  `bsub < submit.sh`
- Find your results when the job is finished

[18]:
```
small_variant_results = pd.read_csv('small_variant/results/GRCh38_
small_variant_results
```

```
/resources/conda/miniconda3/envs/2022_base/lib/python3.7/site-pac
kages/IPython/core/interactiveshell.py:3186: DtypeWarning: Column
s (101) have mixed types.Specify dtype option on import or set lo
w_memory=False.
  interactivity=interactivity, compiler=compiler, result=result)
```

Simple 0 1 Python 3 (ipykernel) | Idle Mode: Command Ln 1

eperry@a-avtl60awn8z7:~

/nas/weka.gel.zone/pgen_int_work/BRS/emily/genotypes_2024/Notebook.html

Notebook.html | Open in Browser | Find | Publish

Running workflows on the HPC

Small variant workflow

- `ssh` to the HPC with your usual login credentials
- `cd` into your working directory
- Make and cd into your working directory `mkdir small_variant_demo` `cd small_variant_demo`
- Copy the Small Variant workflow submission script into your folder
  `cp /pgen_int_data_resources/workflows/rdp_small_variant/main/submit.sh .`
- Make the file `gene_list.txt` and add your list of genes to it `vi gene_list.txt` Add `SMC3` to the file `Esc :wq`
- Edit the submission script: `vi submit.sh` Change the `project_code` to your code Change the `gene_input` to `gene_list.txt` `Esc :wq`
- Run the workflow `bsub < submit.sh`
- Find your results when the job is finished

```
small_variant_results <- readr::read_tsv('small_variant/results/GRCh38_SMC3_ENSG00000108055_annotated_variants.tsv')
```

```
## Rows: 58169 Columns: 102
## ── Column specification ────────────────────────────────────────────────────
## Delimiter: "\t"
## chr (91): CHROM_variant, ID_variant, REF_variant, ALT_variant, Location_anno...
## dbl (11): POS_variant, AN_variant, AC_variant, AC_Hom_variant, AC_Het_varian...
##
## ℹ Use `spec()` to retrieve the full column specification for this data.
## ℹ Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
head(small_variant_results)
```

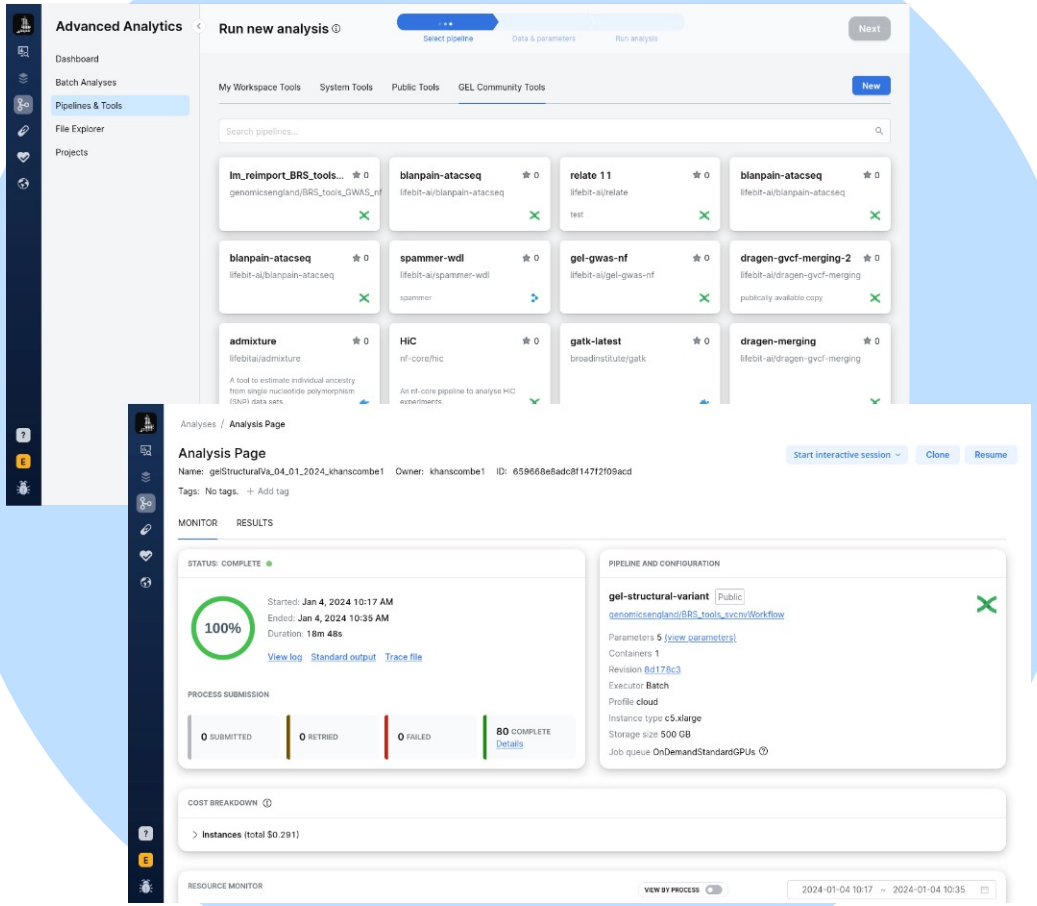
CHROM_variant <chr>	POS_variant <dbl>	ID_variant <chr>	REF_variant <chr>	ALT_variant <chr>	AN_variant <dbl>	AC_variant <dbl>
chr10	110567687	chr10_110567687_1	G	A	242	121
chr10	110567687	chr10_110567687_1	G	A	242	121
chr10	110567687	chr10_110567687_1	G	A	242	121
chr10	110567687	chr10_110567687_1	G	A	242	121
chr10	110567687	chr10_110567687_1	G	A	242	121
chr10	110567687	chr10_110567687_1	G	A	242	121

6 rows | 1-7 of 102 columns

Applications Places System [genotypes_2024] [workshop_scripts] [eperry@a-avtl60awn8... [genotypes_2024 - RSt... genotypes_tr... - Jupyte... /nas/weka.gel.zone/pge... [Homepage - Genomic... Wed Apr 24, 10:47

Running workflows on CloudOS

- Nextflow & WDL pipelines
- One-click System tool GWAS, PRS & VEP pipelines
- GEL Community Pipelines
- Bring your own Pipelines
- Monitor Pipeline in Real-time
- Clone or resume failed pipelines



https://lifebit.atlassian.net/wiki/spaces/CD/pages/813040723/Run+a+pipeline

Batch pipelines on CloudOS demo

6. Aggregate variant files

Aggregate VCFs

Combined gVCFs

Quality metrics

Functional annotation files

Genotype files – platekeys

1371 chunks

https://re-docs.genomicsengland.co.uk/aggv2/ https://re-docs.genomicsengland.co.uk/somAgg/

AggV2 – germline samples

Rare disease

GRCh37 (hg19)
GRCh38 (hg38)

Cancer

Germline GRCh38 (hg38)
Somatic GRCh38 (hg38)

https://re-docs.genomicsengland.co.uk/aggv2/

somAgg

Rare disease

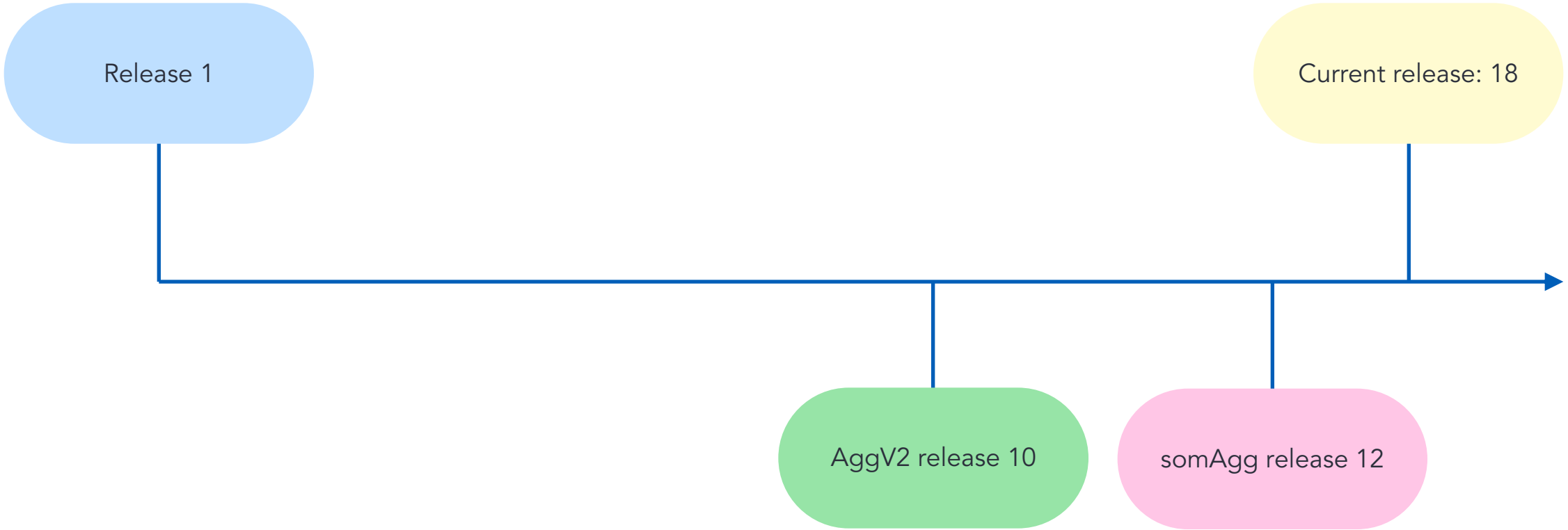
Cancer

GRCh37 (hg19)
GRCh38 (hg38)

Germline GRCh38 (hg38)
Somatic GRCh38 (hg38)

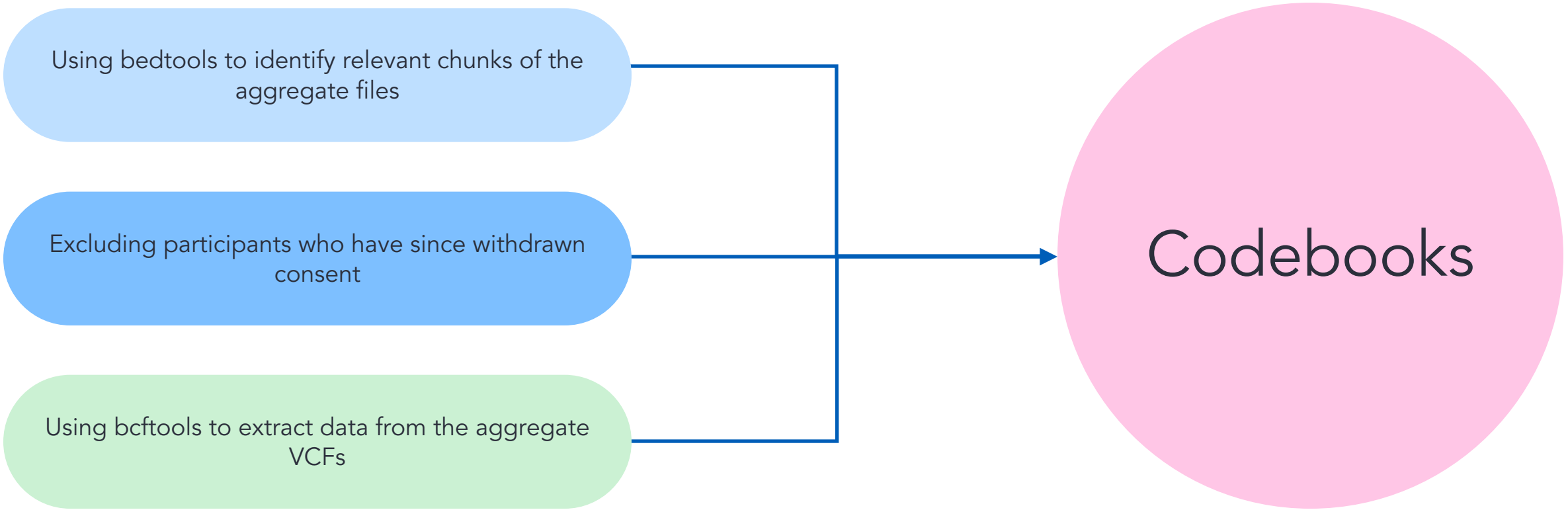
https://re-docs.genomicsengland.co.uk/somAgg/

Aggregate VCFs



https://re-docs.genomicsengland.co.uk/aggv2/ https://re-docs.genomicsengland.co.uk/somAgg/

Aggregate VCFs



https://re-docs.genomicsengland.co.uk/aggv2_code_book/ https://re-docs.genomicsengland.co.uk/somAgg_code_book/

Aggregate VCF chunks

- Locus-based queries must query the correct chunk file
- BED file of chunks available
- Create a sorted BED file of your own regions
- Intersect with BEDtools
- Code books with more information
- Also available in Plink2 format

https://re-docs.genomicsengland.co.uk/aggv2_code_book/ https://re-docs.genomicsengland.co.uk/somAgg_code_book/

Using bcftools on the HPC demo

Amazon WorkSpaces

AggV2 code book - Genomics England Research Environment User Guide — Mozilla Firefox

AggV2 code book - Geno × genotypes_tr... (3) - Jupy ×

https://re-docs.genomicsengland.co.uk/aggv2_code_book/

AggV2 code book

Overview

Back to top

We supply this Code Book to help you use aggV2 in your analysis. This is a live document and will be updated with new feedback and requests.

The code snippets assume that you are working in the **HPC environment and that you submit jobs to the cluster**. Please see About the HPC for more information.

Feedback and Requests

Please reach out via the Genomics England Service Desk for any issues related to the aggV2 aggregation or companion datasets, including "aggV2" in the title/description of your inquiry.

Applications

The majority of queries to aggV2 can be implemented using the applications below:

Application	Description
bcftools	A set of utilities that manipulate variant calls in the Variant Call Format (VCF). Use version 1.10.2 via `module load bio/BCFtools/1.10.2-GCC-8.3.0`
split-vep	A bcftools plug-in to parse VEP annotation (comes with bcftools version 1.10.2-GCC-8.3.0).
LabKey APIs	The LabKey client libraries (APIs) provide programmatic access to the clinical/phenotype data.
R / Python	For downstream processing.
bedtools	To intersect, merge, count, complement, and shuffle genomic intervals. Use version 2.27.1 via module load bio/BEDTools/2.27.1-foss-2018b

Genomics england

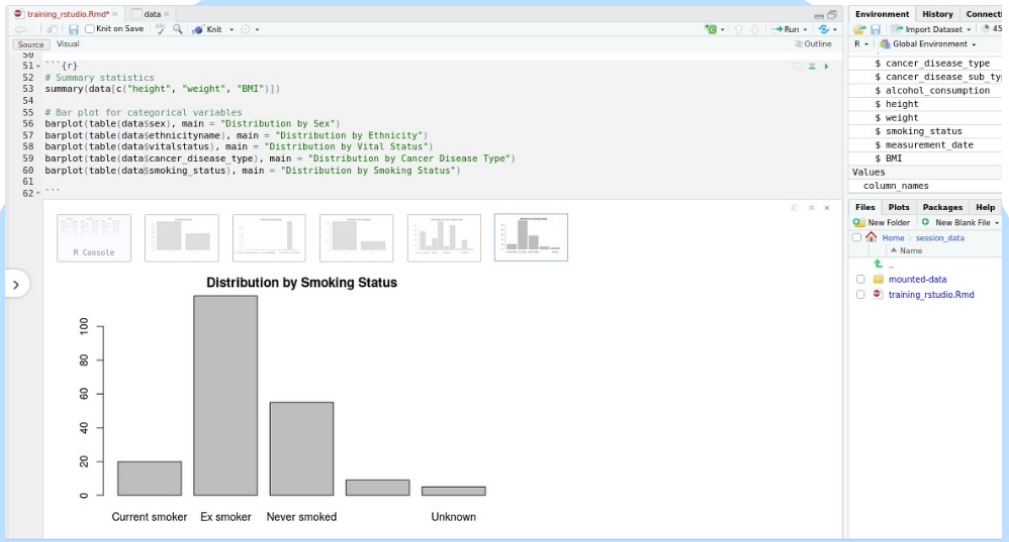
eperry@corp.gel.ac@phpgridzlogn003:/pgen_int_work/BRS/emily/svcnv_demo

File Edit View Search Terminal Tabs Help

```
.0.1/'
[eperry@corp.gel.ac@phpgridzlogn003 emily]$ cp -R /gel_data_resources/workflows/
BRS_tools_svcnvWorkflow/v2.0.1/ svcnv_demo
cp: cannot open '/gel_data_resources/workflows/BRS_tools_svcnvWorkflow/v2.0.1/.R
history' for reading: Permission denied
[eperry@corp.gel.ac@phpgridzlogn003 emily]$ cd svcnv_demo
[eperry@corp.gel.ac@phpgridzlogn003 svcnv_demo]$ ls
cromwell.conf           input         README.md          svcnvCatchCore.wdl
cromwell-executions     logs          src                svcnvCatch_output
cromwell-workflow-logs  makeClean.sh  submit_script.sh   svVW_report_out
[eperry@corp.gel.ac@phpgridzlogn003 svcnv_demo]$ vi input/gene_list.txt
[eperry@corp.gel.ac@phpgridzlogn003 svcnv_demo]$ vi submit_script.sh
[eperry@corp.gel.ac@phpgridzlogn003 svcnv_demo]$ bsub < submit_script.sh
Job <624884> is submitted to queue <inter>.
[eperry@corp.gel.ac@phpgridzlogn003 svcnv_demo]$ bjobs
JOBID   USER    STAT  QUEUE      FROM_HOST   EXEC_HOST   JOB_NAME   SUBMIT_TIME
473405  eperry@ RUN   inter      phpgridzlog phpgridzlsf bash       Jun 16 09:11
527673  eperry@ RUN   inter      phpgridzlog phpgridzlsf bash       Jun 19 09:52
359264  eperry@ RUN   inter      phpgridzlog phpgridzlsf /bin/bash  Jun 12 10:09
359443  eperry@ RUN   inter      phpgridzlog phpgridzlsf bash       Jun 12 10:31
524644  eperry@ RUN   inter      phpgridzlog phpgridzlsf bash       Jun 19 07:24
623719  eperry@ PEND  inter      phpgridzlog             *l_variant Jun 20 09:12
624884  eperry@ PEND  inter      phpgridzlog             svcnvCatch Jun 20 09:18
[eperry@corp.gel.ac@phpgridzlogn003 svcnv_demo]$
```

cloudOS Interactive Analysis Sessions

- Utilise Jupyter Notebooks, Spark Notebooks or RStudio
- Collaborate with multiple users in real-time
- Save snapshots of environments
- Use nextflow pipelines and docker containers
- Develop Dash and Shiny apps



https://lifebit.atlassian.net/wiki/spaces/CD/pages/813105209/Interactive+sessions

Using bcftools in a CloudOS interactive session demo

7. When/why you would use each method

Platform

Tiering and exomiser tables

IVA and Cohort browser omics

Small/ Structural variant workflows

Querying the aggregates

Search by

Tiering and exomiser tables

HGVS HUMAN GENOME VARIATION SOCIETY

IVA and Cohort browser omics

dbSNP Short Genetic Variations

Small/ Structural variant workflows

Querying the aggregates

Variants included

Tiering and exomiser tables

Variants that have passed filters

IVA and Cohort browser omics

All

Small/ Structural variant workflows

All

Querying the aggregates

Variants from GRCh38-aligned genomes from release 8 or 12

Datasets

Tiering and exomiser tables

100k and NHS GMS

IVA and Cohort browser omics

100k

Small/ Structural variant workflows

100k

Querying the aggregates

100k

Genome assembly

Tiering and exomiser tables

GRCh37 and GRCh38
Assembly as a separate column

IVA and Cohort browser omics

GRCh37 and GRCh38 in separate databases

Small/ Structural variant workflows

GRCh37 and GRCh38 queries simultaneously

Querying the aggregates

GRCh38 only

Underlying VCFs

Tiering and exomiser tables

Rare disease: Platypus
Cancer: Strelka

IVA and Cohort browser omics

Rare disease: Platypus
Cancer: Strelka

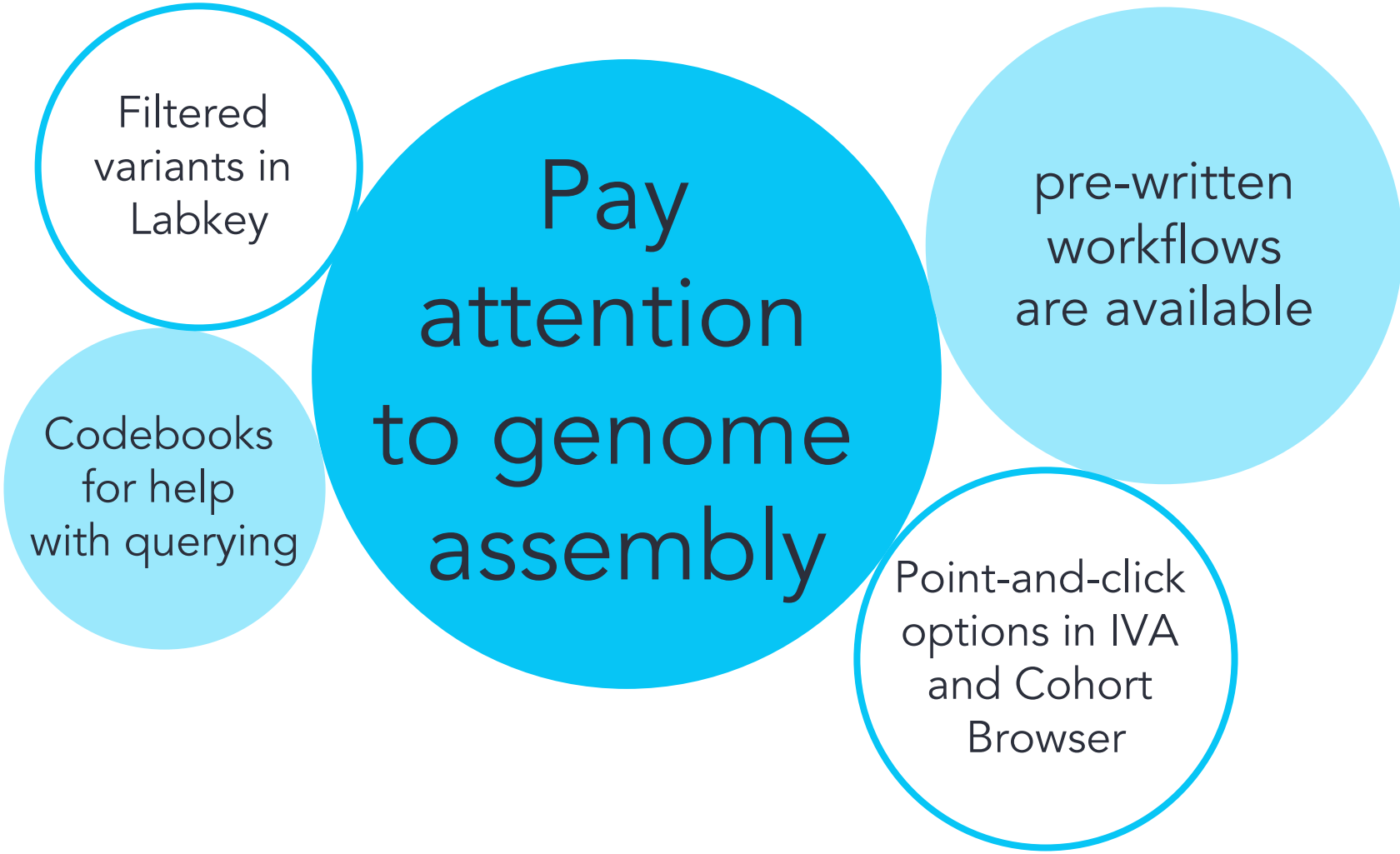
Small/ Structural variant workflows

Strelka

Querying the aggregates

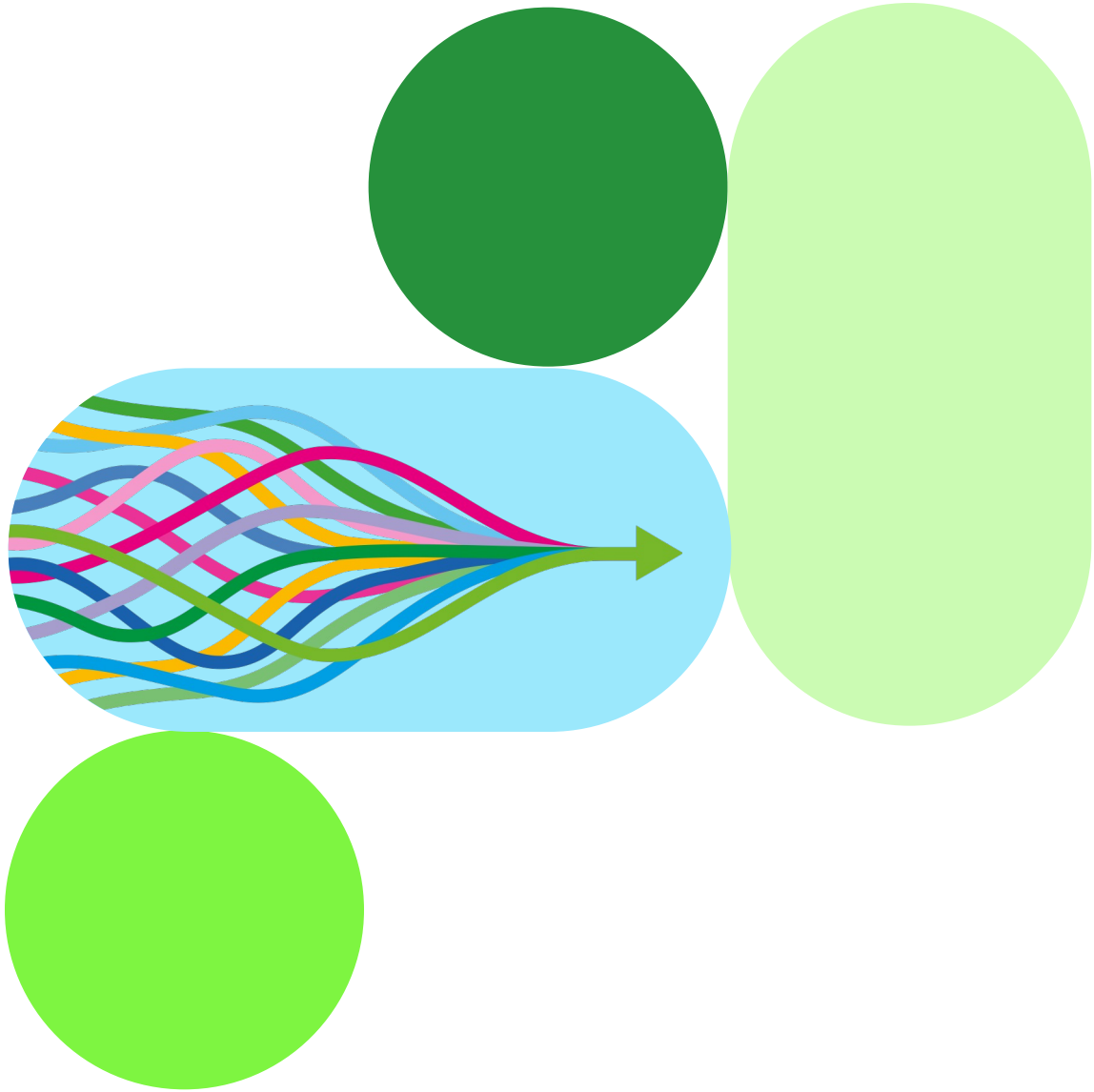
Strelka

Key takeways



Genomics England | Learning and Development

Optional exercise



Optional exercise - coding/command line

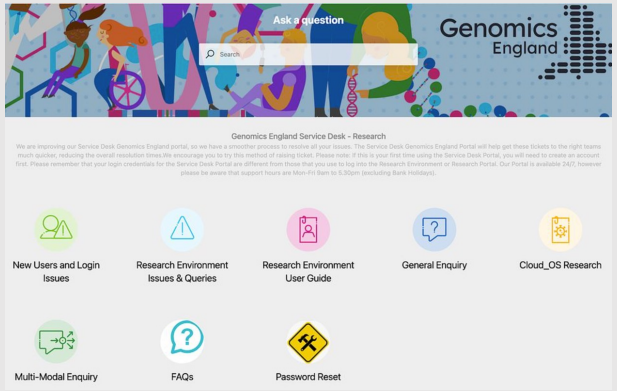
1. Use the LabKey API to look up participants with variants in the gene *JPH3* that have been selected by rare disease tiering, cancer tiering or exomiser. Repeat your rare disease tiering query with NHS GMS data.
2. Run the Small Variant and Structural variant workflows to find participants with all variants in *JPH3*.
3. Query the SomAgg aggregate VCF for all participants with an alternate allele at 16:87690170. Make sure you query the correct file chunk.

Optional exercise – point and click

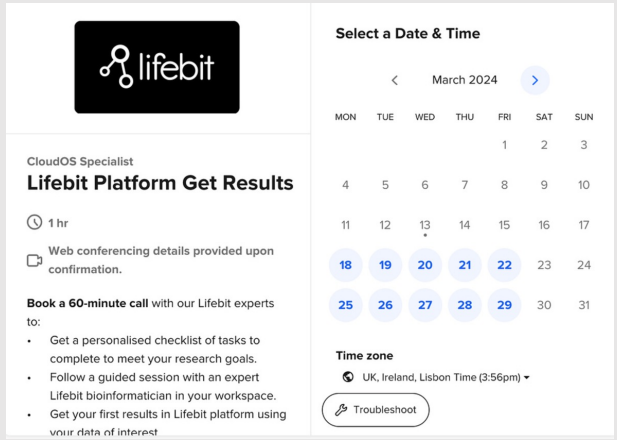
1. Use Labkey or Cohort Browser to look up participants with variants in the gene *JPH3* that have been selected by rare disease tiering, cancer tiering or exomiser. Repeat your rare disease tiering query with NHS GMS data.
2. Use IVA or Cohort Browser to find all participants with somatic variants in *JPH3*.
3. (CloudOS Cohort Browser only) Apply joint phenotypic-genotypic query.

8. Help and questions

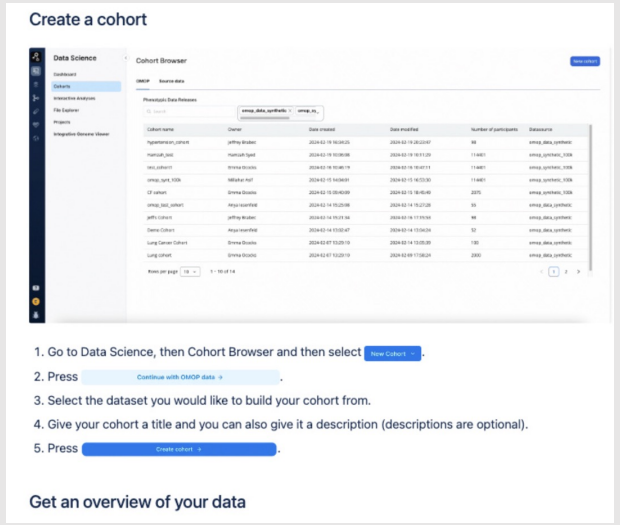
Getting started on CloudOS



Open a ticket via the Service Desk.



1:1 onboarding and introductory training session with a Lifebit bioinformatician.



Continuous support from the Lifebit team and documentation

Getting help

Check our documentation:
https://re-docs.genomicsengland.co.uk/
Click on the documentation icon in the environment

Contact our Service Desk:
https://jiraservicedesk.extge.co.uk/plugins/servlet/desk

Questions

All your microphones are muted

Use the Zoom Q&A to ask questions

Upvote your favourite questions: if we are short on time we will prioritise those with the most votes

Training sessions

16/7	Getting medical records for participants
10/9	Using GEL data for publications and reports
8/10	What tools and workflows should I use to fulfil an overall goal?
12/11	Running workflows on the HPC and Cloud
10/12	Introduction to the RE

https://re-docs.genomicsengland.co.uk/upcoming/

In-person training day

- For any registered RE users
- 20th November 2024
- Held in our offices at Canary Wharf
- Hybrid attendance
- Hands-on practicals and exercises

https://www.genomicsengland.co.uk/events/research-environment-training-event-november-2024

Feedback

Thank you

Visit: https://re-docs.genomicsengland.co.uk/