

# step 1

## select data set in Mosaic

Welcome to Mosaic!

An intuitive data hub for visually organizing and exploring genomic data.

[Tutorials](#) [Get Started](#) [Support](#)

1000G WGS	Distributed Data	Ewing's Sarcoma
Sample Count 851 Created On 29 August 2018 Last Updated 29 August 2018 1000+ samples from different ethnic groups around the world	Sample Count 4 Created On 29 August 2018 Last Updated 29 August 2018 Data on S3 cloud and Simons hardware	Sample Count 1,049 Created On 29 August 2018 Last Updated 29 August 2018 A cancer that most often occurs in and around the bones

H1K	Platinum Data Set	SPARK_Pilot
Sample Count 753 Created On 29 August 2018 Last Updated 29 August 2018 investigating 1k priority patients	Sample Count 4 Created On 29 August 2018 Last Updated 29 August 2018	Sample Count 1,394 Created On 29 August 2018 Last Updated 29 August 2018 pilot samples

SSC GRCh37 WES	SSC GRCh37 WGS	SSC GRCh38 WGS
Sample Count 9,046 Created On 29 August 2018 Last Updated 29 August 2018 The SSC Pilot, Phase1, Phase 2, Phase3-1, Phase3-2, and Phase 4 exome data aligned to the GRCh37 reference. The Simons Simplex Collection (SSC) is a core resource of the Simons Foundation Autism Research Initiative (SFARI) and contains genetic samples and associated phenotypic...	Sample Count 2,135 Created On 29 August 2018 Last Updated 29 August 2018 The SSC Pilot and Phase 1 wgs data aligned to the GRCh37 reference. The Simons Simplex Collection (SSC) is a core resource of the Simons Foundation Autism Research Initiative (SFARI) and contains genetic samples and associated phenotypic data from 2,600 simplex families	Sample Count 6,951 Created On 29 August 2018 Last Updated 29 August 2018 The SSC Phase 2, Phase3-1, Phase3-2, and Phase 4 wgs data aligned to the GRCh38 reference. The Simons Simplex Collection (SSC) is a core resource of the Simons Foundation Autism Research Initiative (SFARI) and contains genetic samples and associated phenotypic data from 2,600...

**Select desired project**

© 2018 Frameshift Labs, Inc.

Blog Contact Help

# step 2

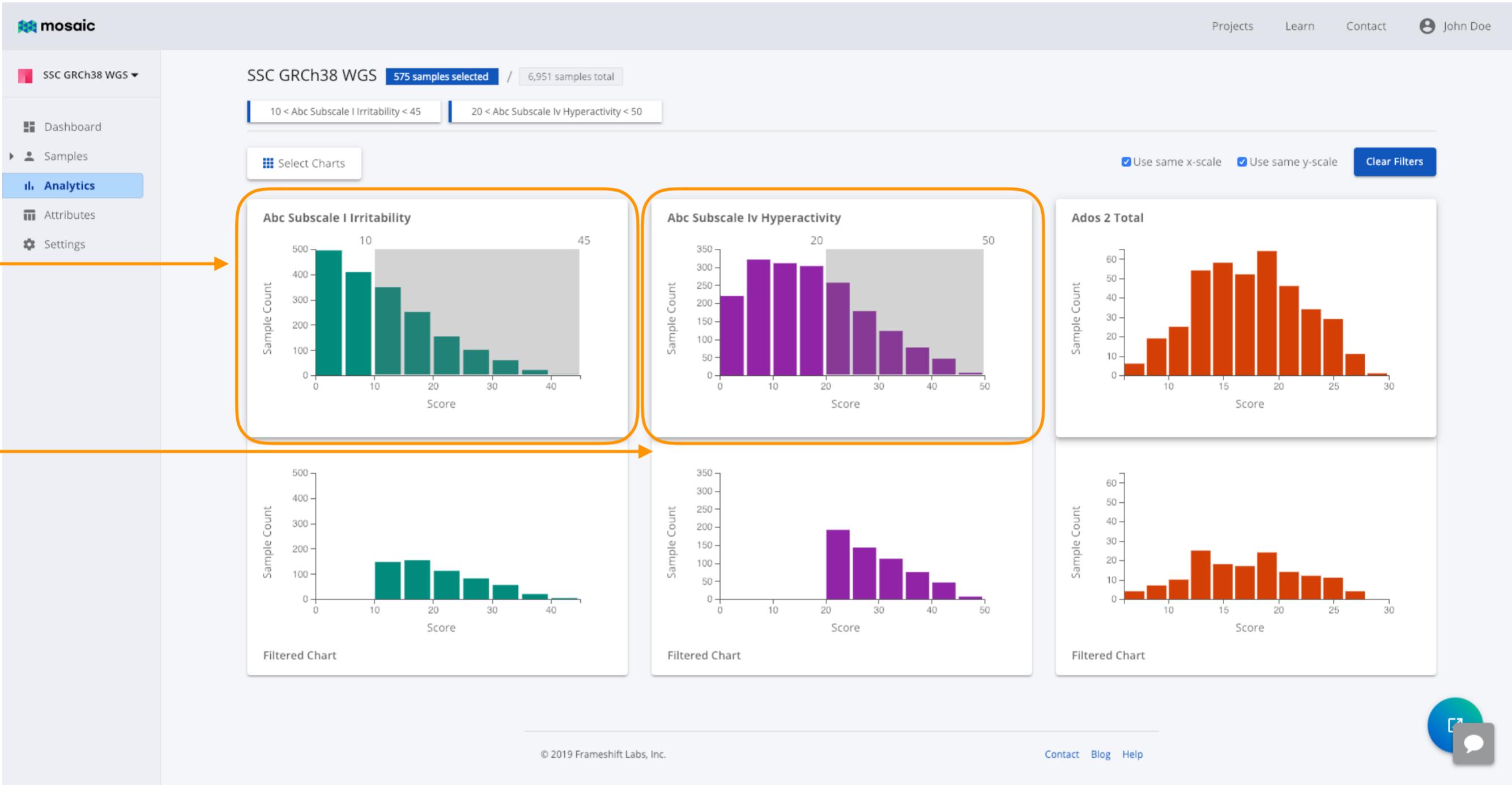
## choose filtering analytics

The screenshot shows the mosaic web interface. On the left, a sidebar menu includes 'Dashboard', 'Samples', and 'Analytics' (which is highlighted with a blue border). A secondary sidebar shows a project named 'SSC GRCh38 WGS' with 6,951 samples total. In the center, there are three charts: 'Abc Subscale IV Hyperactivity' (teal bars), 'Abc Total Score' (purple bars), and 'Ados 2 Total' (orange bars). An orange arrow points from the 'Analytics' button in the sidebar to the 'Select Charts' button at the top of the main content area.

This screenshot shows the 'Chart Selection' dialog box overlaid on the main mosaic interface. The dialog box has a title 'Chart Selection' and a sub-header '6,951 samples selected'. It contains a search bar and a dropdown menu set to 'Alphabetical Order'. Below these are several chart entries, each with a checkbox and a histogram icon. The first entry, 'abc subscale i irritability', has its checkbox checked and is highlighted with a yellow arrow. Other entries include 'abc subscale ii lethargy', 'abc subscale iv hyperactivity', 'abc subscale v inappropriate speech', and 'abc total score'. To the right of the dialog box, a portion of the main dashboard is visible, showing the 'Ados 2 Total' chart.

# step 3

## select filtered subset



# step 4

## select gene & launch cohort

SSC GRCh38 WGS 575 samples selected / 6,951 samples total

10 < Abc Subscale I Irritability < 45    20 < Abc Subscale Iv Hyperactivity < 50

Select Charts    Use same x-scale    Use same y-scale    Clear Filters

**Abc Subscale I Irritability**

Score Range	Sample Count
0-10	500
10-20	400
20-30	300
30-40	200
40-50	100
50+ (Grey)	450

**Abc Subscale Iv Hyperactivity**

Score Range	Sample Count
0-10	200
10-20	300
20-30	300
30-40	200
40-50 (Grey)	50

**Ados 2 Total**

Score Range	Sample Count
0-10	50
10-20	50
20-30	50
30-40	50
40-50	60

**Cohort-gene.iobio** An application that identifies and visualizes variants that are enriched in a selected subset relative to all probands. Analysis

**Multibam.iobio** Visualizes alignment quality control metrics for multiple samples (from several samples up to thousands) at the same time. Quality Control

**Steps**

- 575 samples selected
- Select a subset from the project analytics charts ⓘ
- Enter gene of interest

Please type in the exact name of the gene.

POGZ

Launch    Cancel

© 2019 Frameshift Labs, Inc.

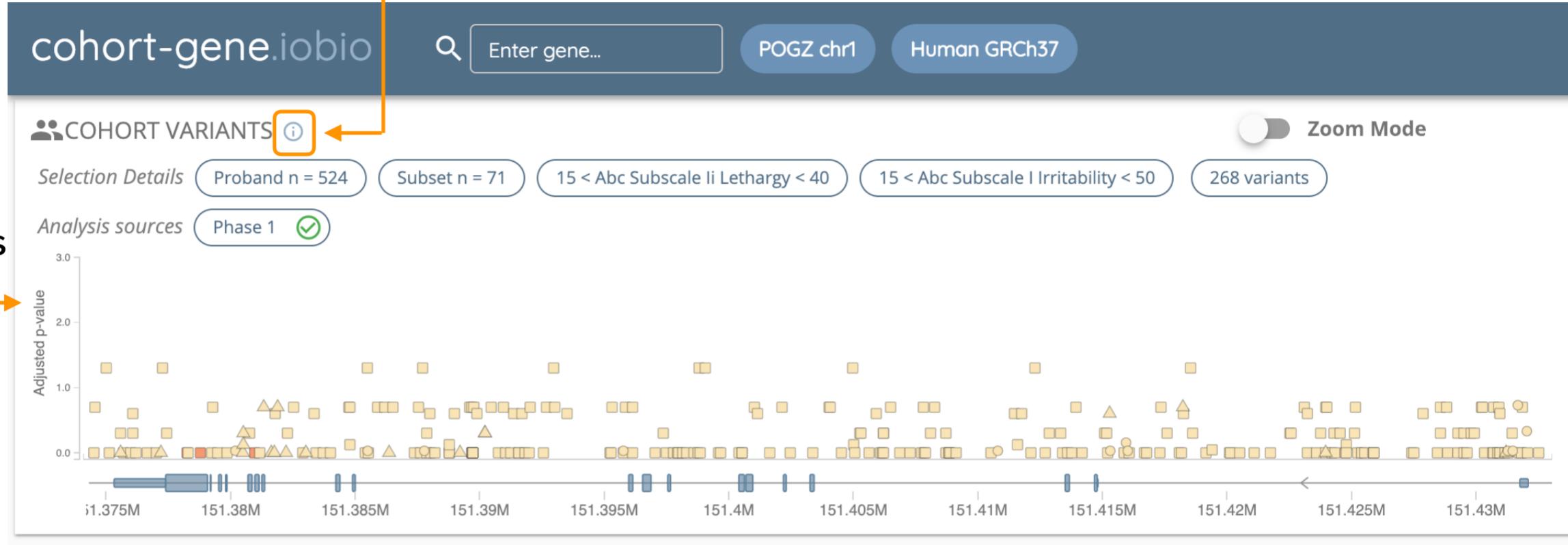
# step 5

## review enriched variants

Variants are arranged by enrichment within the subset cohort along the y-axis

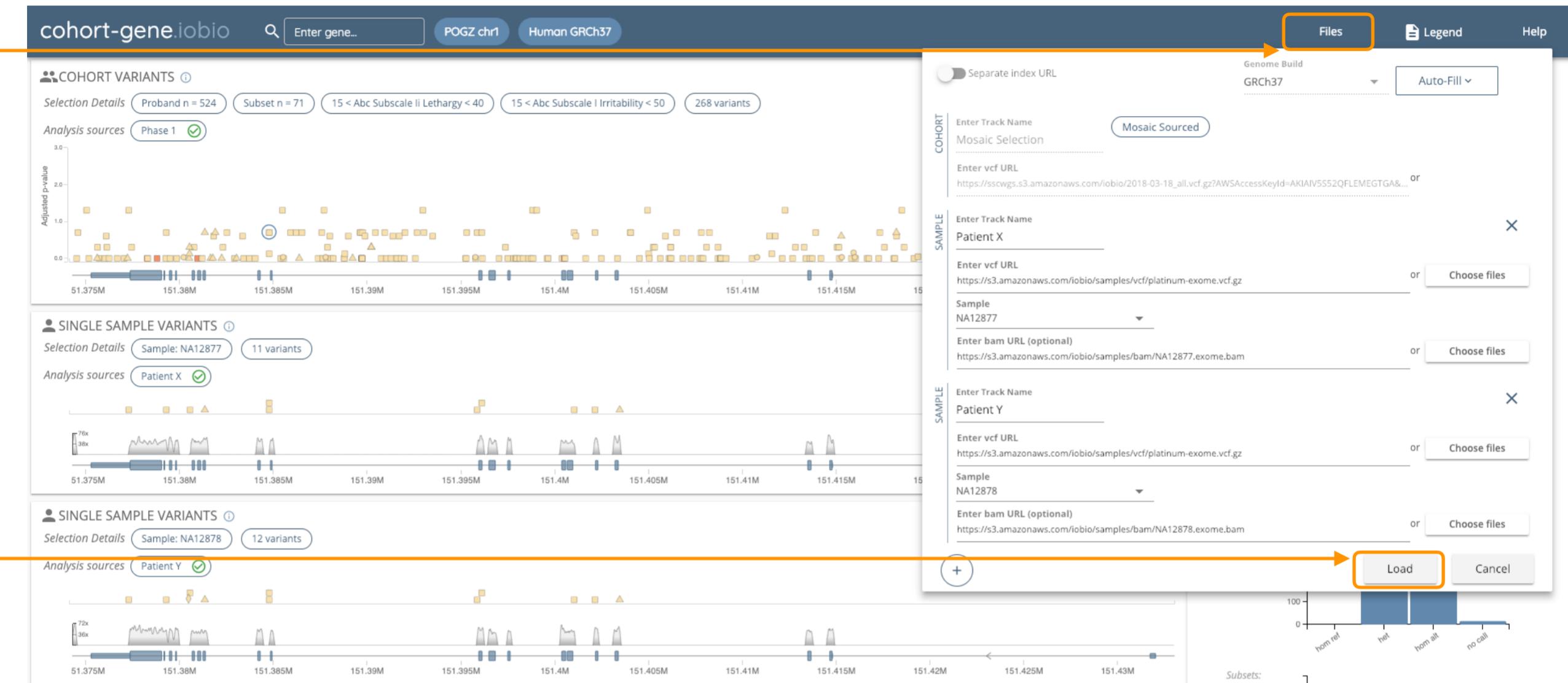
The higher the variant, the more enriched

More details about enrichment statistics can be viewed in the info panel



# step 6

## upload your own data



Entered data can be easily uploaded on subsequent analyses (see step 11)

When launched from Mosaic, cohort data and build cannot be altered

Enter vcf url or select 'Choose files' to upload local data

Select sample name from dropdown

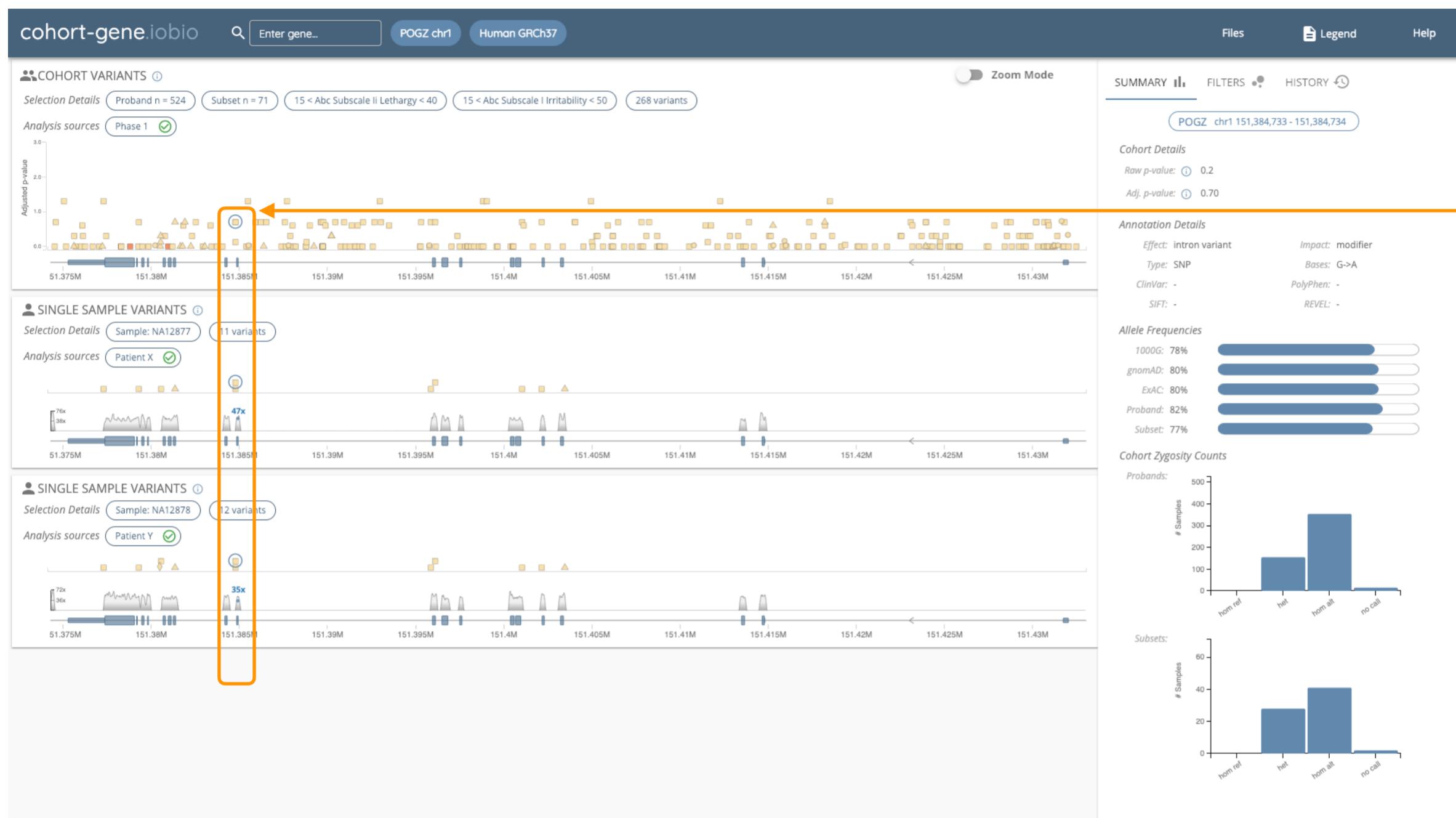
Optionally enter bam url or choose file to display coverage information

Add as many tracks as desired

This screenshot shows the 'Auto-Fill' modal window, which is a simplified version of the upload interface. It includes sections for COHORT (Mosaic Sourced) and SAMPLE (Patient X, Patient Y). Each sample section has a 'vcf URL' field with a sample VCF link and an optional 'bam URL' field with a sample BAM link. A '+' button is available to add more samples. The 'Load' button at the bottom right is highlighted with an orange arrow.

# step 7

## select individual variant for details



Click on a variant in any track to view it across all samples

Review enrichment and annotation details in summary tab

**Cohort Details**

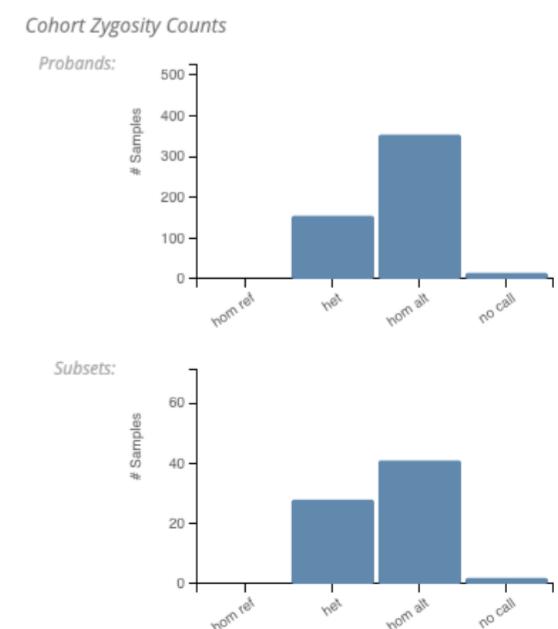
- Raw p-value: 0.2
- Adj. p-value: 0.70

**Annotation Details**

Effect: missense variant	Impact: moderate
Type: SNP	Bases: T->A
ClinVar: -	PolyPhen: benign
SIFT: tolerated	REVEL: 0.157

**Allele Frequencies**

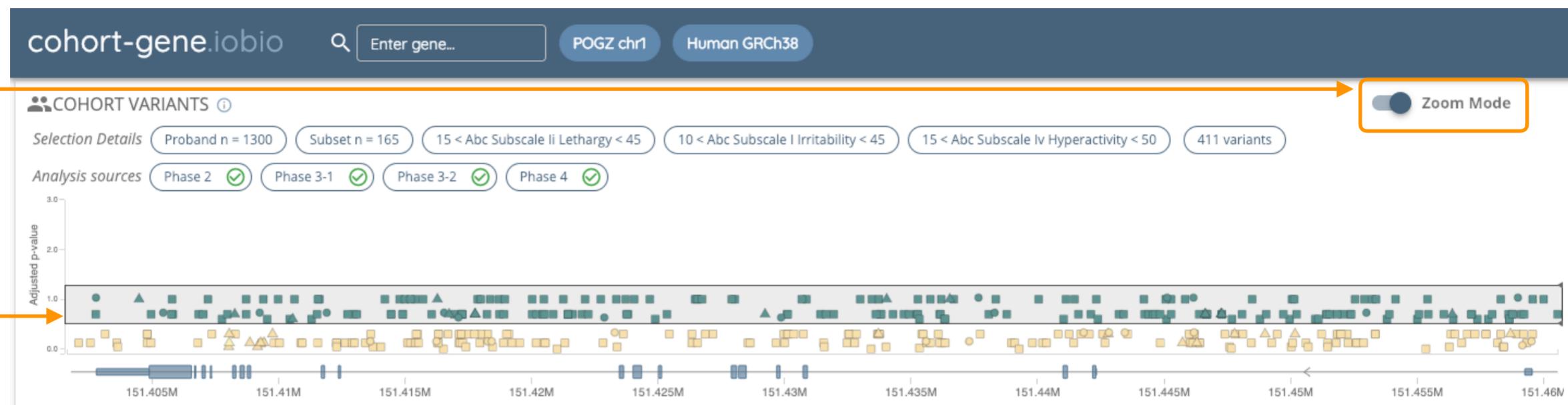
Dataset	Frequency
1000G	78%
gnomAD	80%
ExAC	80%
Proband	82%
Subset	77%



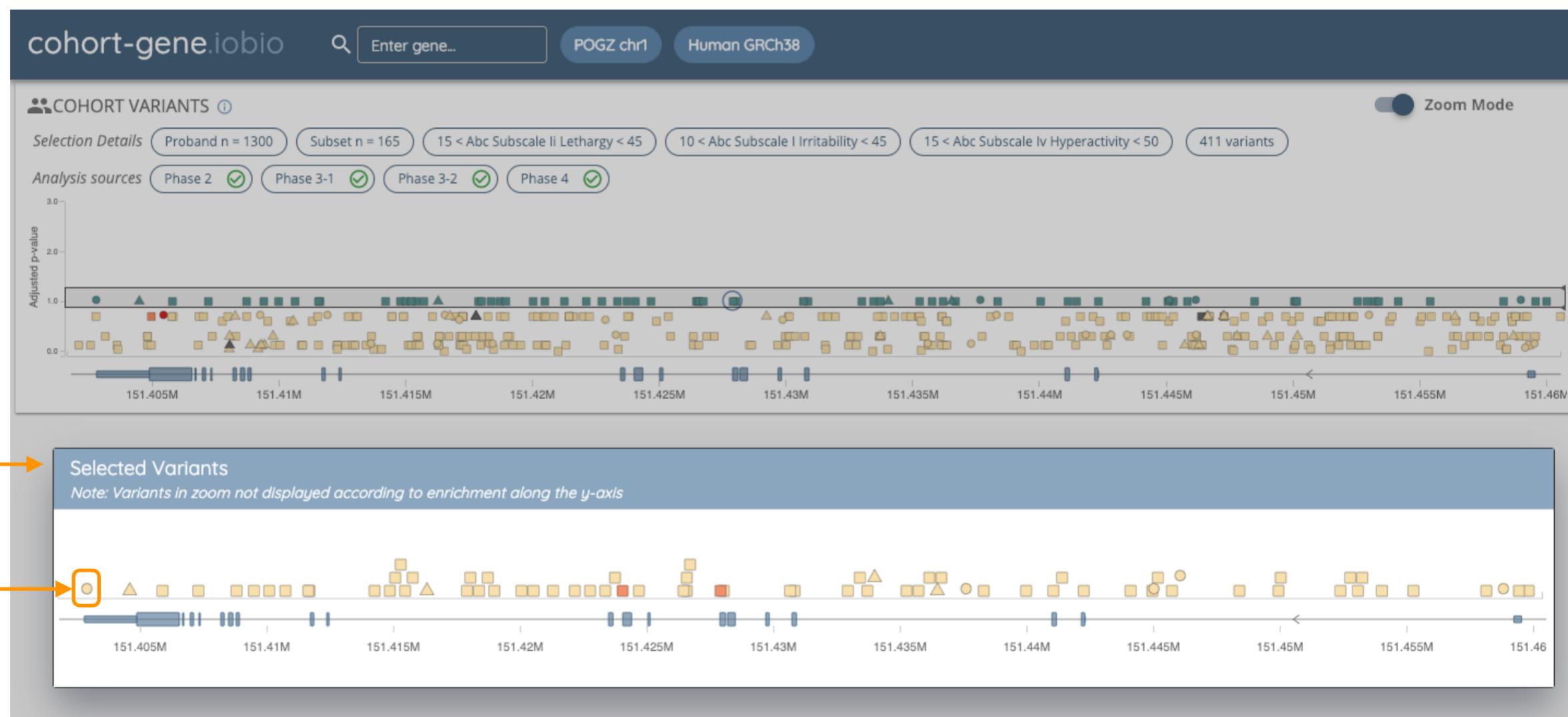
# step 8

zoom in on variants for better separation

Switch  
toggle



Brush  
vertically

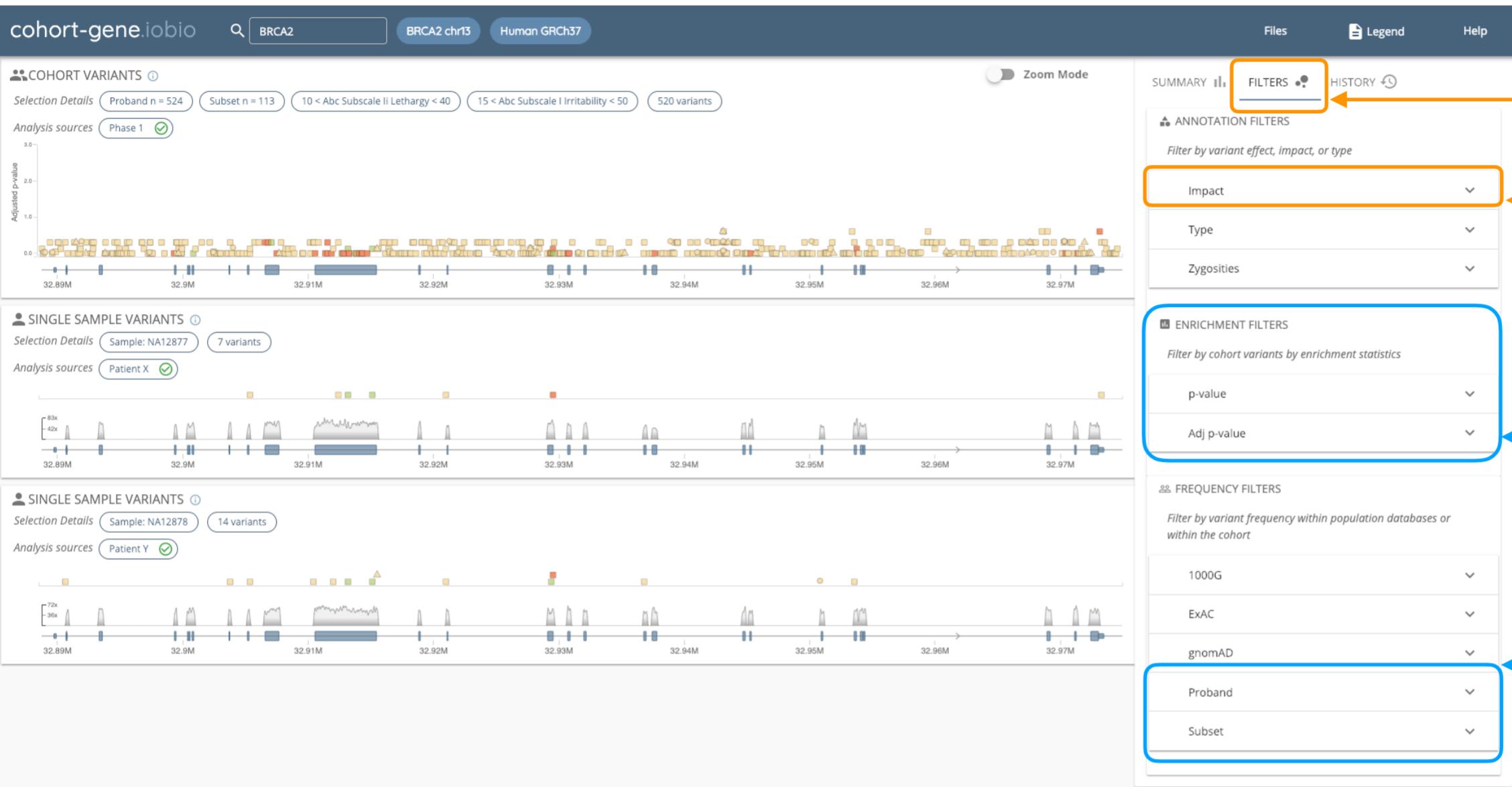


Modal is  
movable to  
view other  
tracks  
while  
zooming

Variants  
within  
zoom  
modal are  
selectable

# step 9

## filter variants



Select filter tab

Click on filter title to expand or contract options

Note:  
Certain filters only apply to cohort variants

Check and uncheck category filters

Enter criteria for range filters and click on green check to apply

Once applied, the check will turn gold and an active indicator dot will appear

Annotation Filters (Impact):

- HIGH
- MODERATE
- MODIFIER
- LOW

Frequency Filters (1000G):

Filter by variant frequency within population databases or within the cohort

1000G: <= 20 %

ExAC, gnomAD, Proband, Subset

Frequency Filters (1000G):

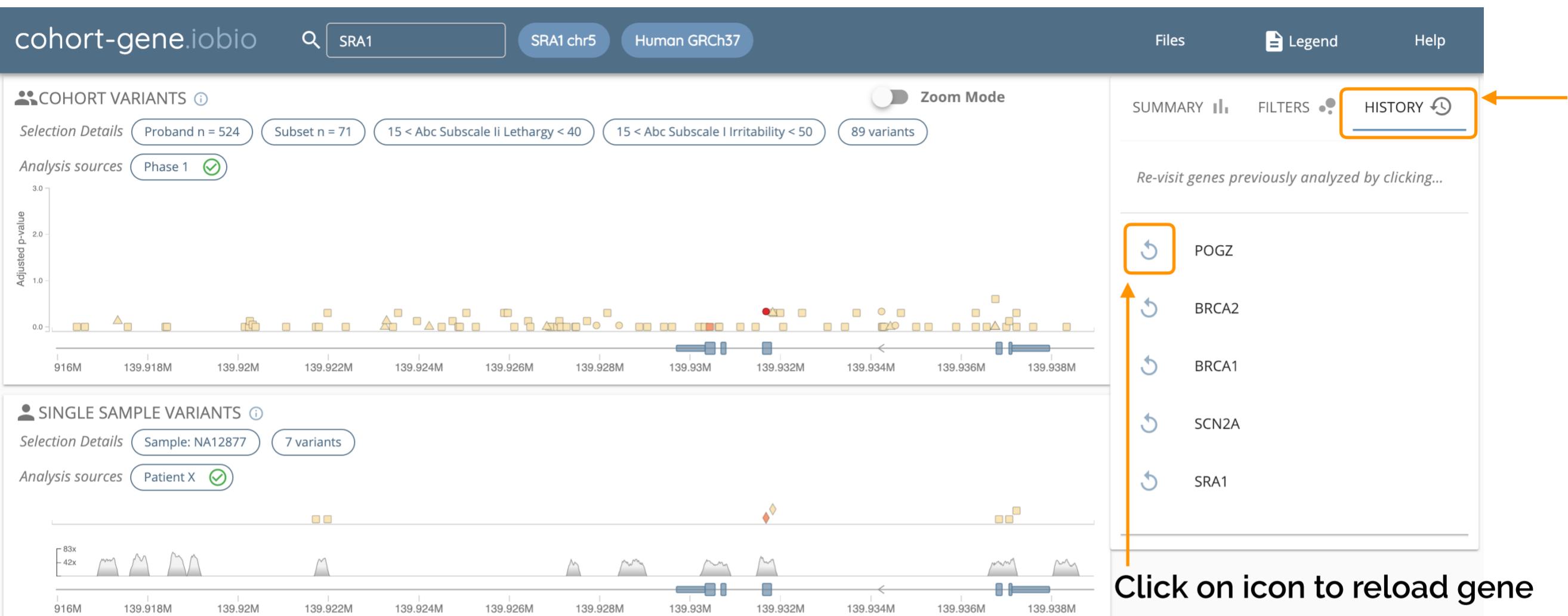
Filter by variant frequency within population databases or within the cohort

1000G: <= 20 %

ExAC, gnomAD, Proband, Subset

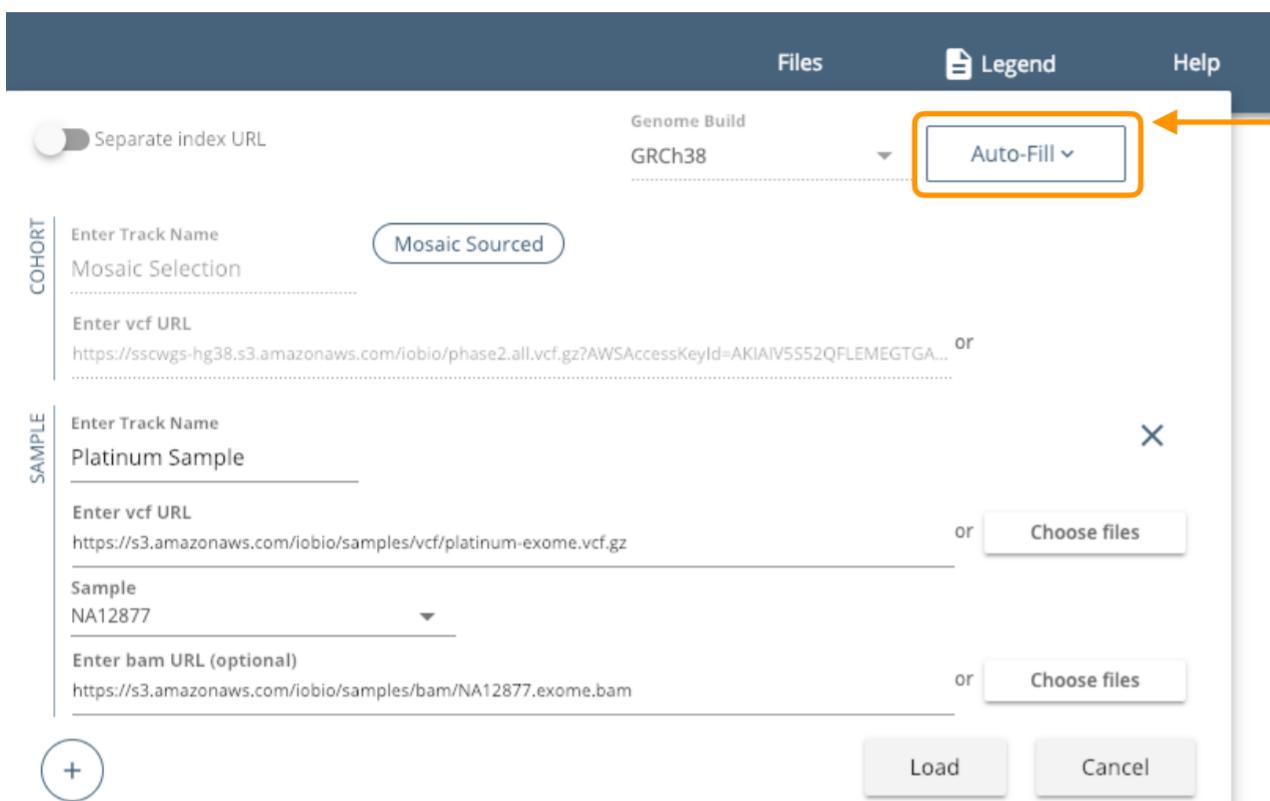
# step 10

## review analyzed loci

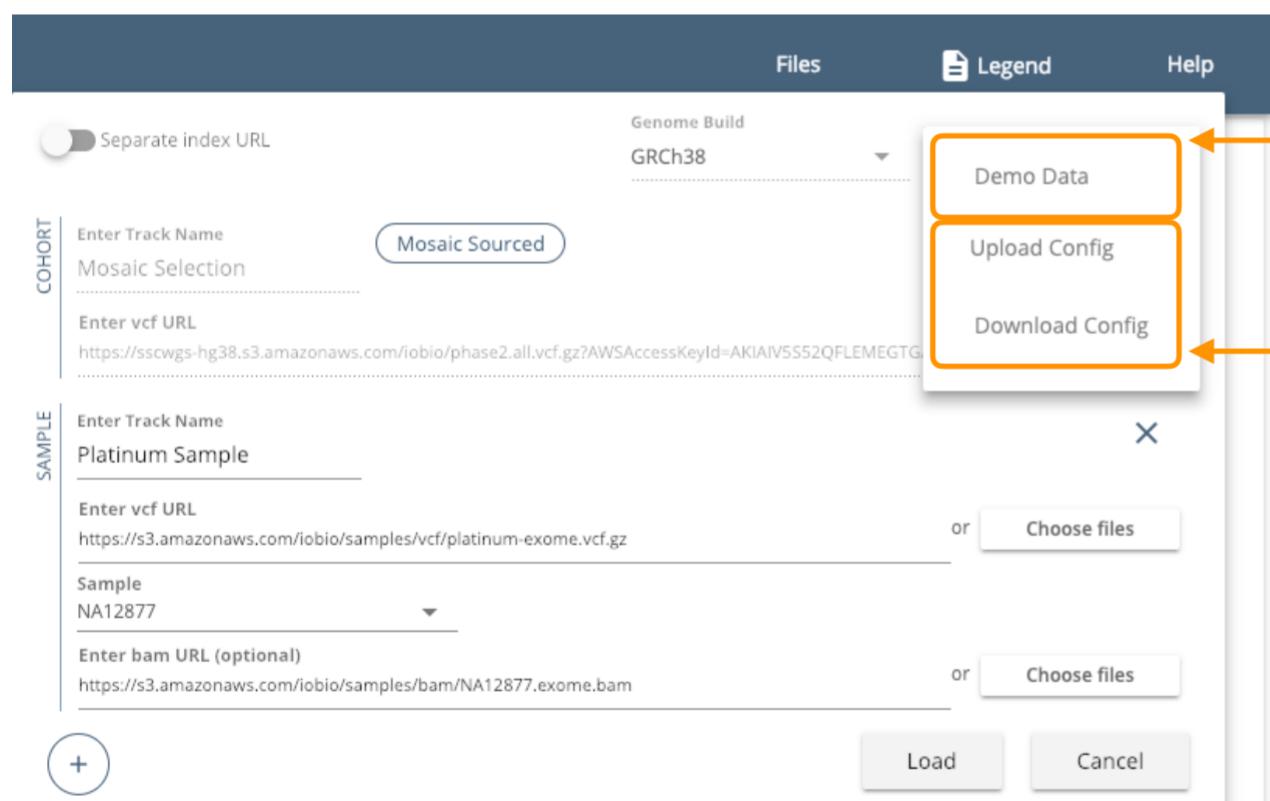


# step 11

## alternative launch options



Select auto-fill dropdown



Demo data from the GRCh37 Platinum dataset may be analyzed

When valid urls are entered, easy-entry configuration files may be downloaded. Upon subsequent analysis with the same samples, simply upload the configuration. Note: Mosaic data and local files may *not* be launched this way.

To upload cohort data hosted outside of Mosaic, navigate directly to [cohortgene.iobio.io](http://cohortgene.iobio.io)

