
VariantGrid

CCB ACRF Cancer Genomics Facility

Oct 10, 2019

OVERVIEW:

1	Intro	3
1.1	Cloud servers	3
1.2	Private server	3
2	Analysis Intro	5
2.1	Analysis Nodes	5
2.2	Analysis screen	8
2.3	Analysis Grid	8
3	Analysis Nodes	9
3.1	Source Nodes	9
3.2	Filter Nodes	10
4	Analysis - advanced	15
4.1	Analysis settings	15
4.2	Node Counts	16
4.3	Column Summary	17
5	Variant Tagging	19
5.1	Create tags	19
5.2	Tagging variants	20
5.3	Using tags	21
6	Analysis Classification	23
7	Karyomapping	25
7.1	Background	25
7.2	Gene analysis	25
7.3	Genome-wide analysis	26
8	Annotation Details	27
8.1	Variant Level Annotation	27
8.2	Annotation Versions	27
8.3	IVAT	27
9	Variant Details	31
9.1	Classifications	31
9.2	Transcripts	31
9.3	Samples	31
10	Representative Transcript	33

11 Uploading Data	35
12 Managing data	37
12.1 Sharing data	38
12.2 Search	38
13 Somatic data	39
13.1 Allele Frequency	39
14 Mutational signatures	41
15 VCF / Samples	43
15.1 VCF import	43
15.2 Multi-sample VCFs	43
16 Search	45
17 Gene Page	47
18 Gene Lists	49
18.1 Creating Gene Lists	49
18.2 Using gene lists in analyses	50
19 Gene Grid	53
19.1 GeneGrid screen	53
19.2 Gene Info	54
20 Gene Coverage	55
20.1 Canonical Transcripts	55
20.2 Sample QC metrics	55
20.3 GeneGrid EnrichmentKit coverage	55
21 Pathology Tests	57
21.1 Active tests	57
21.2 Requesting gene changes	57
21.3 Accepting gene changes	58
22 Test Ordering	61
23 Patients	63
23.1 Searching	63
23.2 Patient records	64
23.3 Other sources of patients	64
23.4 Other	64
24 Phenotypes	65
24.1 Assigning Terms to Patients	65
24.2 How phenotype term matching works	65
25 Cohorts	67
25.1 Create a new cohort	67
25.2 Create from a larger cohort	67
25.3 Cohort Analyses	68
26 Trios	71
26.1 Creating a trio	71

26.2	Digital karyomapping	71
26.3	Trio inheritance analysis	72
27	Sequencing Runs	75
27.1	Gold Standard Runs	76
27.2	Finding sequencing data	77
28	User Settings	79
28.1	Customise columns	79
29	Customise columns	81
30	IGV integration	83
30.1	IGV Configuration	83
30.2	VariantGrid Configuration	84
30.3	Network drives and File Servers	85
31	Variant Classifications	87
31.1	Creating Classifications	87
31.2	Autopopulation	87
31.3	Editing	87
31.4	Configuring Fields	87
32	Variant Classification Form	89
32.1	View	89
32.2	Identify Errors	90
32.3	Change History / Diff	90
32.4	ACMG Guidelines	90
32.5	Actions	90
32.6	Delete / Withdraw	91
32.7	Export	91
32.8	Literature Citations	91
33	Classification Flags	93
33.1	Flag Dialog	94
33.2	Flag Types	95
34	Variant Classification Report	97
34.1	Running the report	97
34.2	Configuring the report	97
34.3	Values available for the report	97
35	Variant Classification REDCap	99
35.1	REDCap Definition	99
35.2	REDCap Rows	100
35.3	Technical Specifics	100
36	Variant Normalization	101
37	Indices and tables	103

VariantGrid is an open source variant database and web application for analyzing genetic data.

VariantGrid has a number of installations. Please visit the individual sites for login/registration details.

1.1 Cloud servers

- variantgrid.com - Research cloud server
- [runx1db](https://runx1db.org) - Rare disease exome sharing
- [Shariant](https://shariant.org) - [Australian Genomics](https://australian-genomics.org) variant classification sharing platform

1.2 Private server

There is a VariantGrid private server inside [SA Pathology](https://sa-pathology.org), the public pathology provider to the South Australian Health.

The advantages of a private server are being restricted to a private intranet, and being able to analyse private patient data without worrying about it being on the cloud.

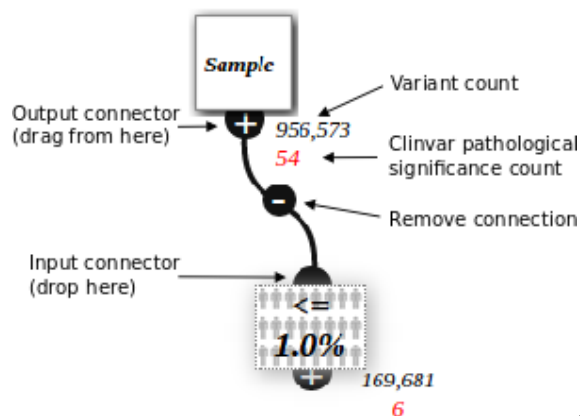
To install a local copy of VariantGrid, please see the [GitHub](https://github.com/variantgrid/variantgrid) page.

ANALYSIS INTRO

Create custom variant filters by connecting together nodes representing sources or filters of variants. See [analysis nodes](#)

Other variant databases allow similar creation of filters, but VariantGrid can construct nodes in real-time, enabling rapid exploration of large and difficult genomic data sets.

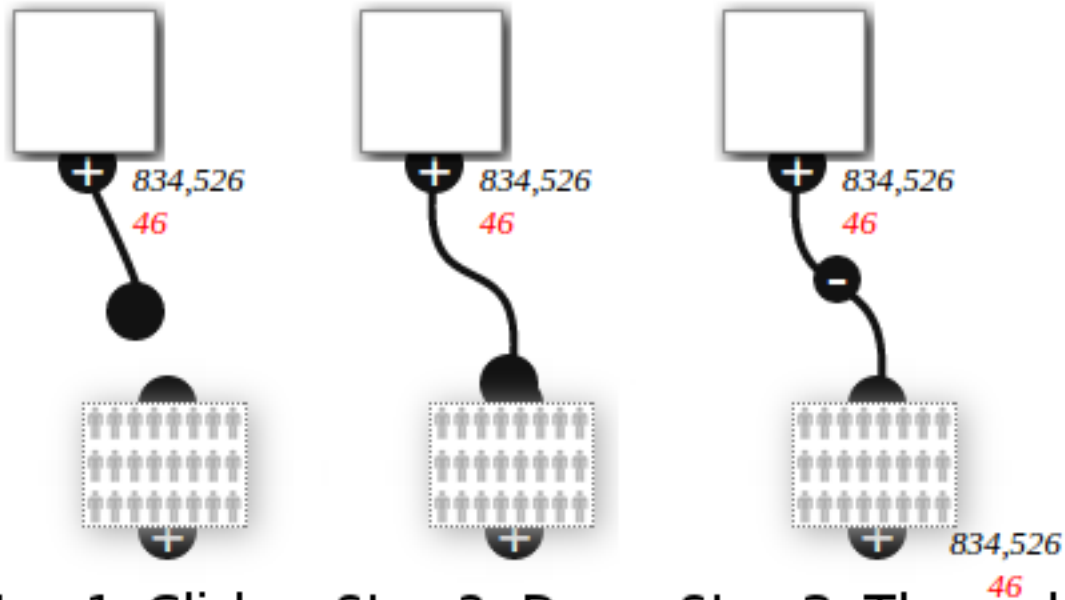
2.1 Analysis Nodes



Sample Node connected to a Population Filter Node

The top node is configured to show a particular patient exome (from an uploaded VCF).

These variants are then filtered to those that are less than 1% of the population.




Step 1: Click and drag on the on the "+" symbol.

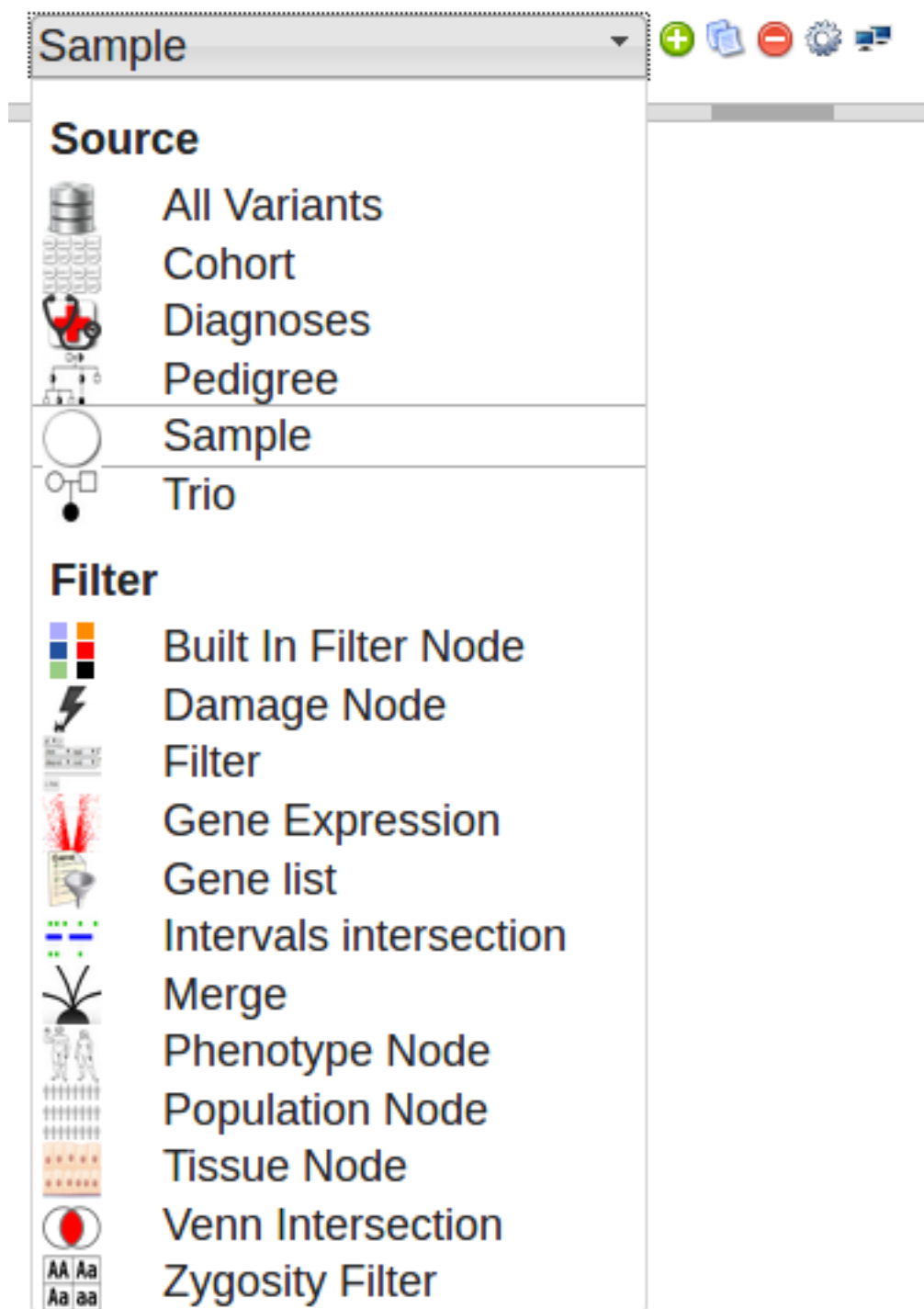
Step 2: Drop it on the top of another node.


Step 3: The nodes are connected and counts calculated

Connecting

Nodes

To add a node, select the node type from the drop down menu in the top left of the screen and click the  add button



Click and drag a node to move it around. You can select multiple nodes by drag-selecting a box around them. This allows you to copy, delete or move them as a group. Delete selected nodes by pressing DELETE, or click the  delete button.

2.2 Analysis screen

The screenshot shows the VariantGrid analysis screen. On the left is a node graph with nodes like HSS24, Case (Lof-6), and snpeff_impact. On the right is the node editor for 'Case (1 of 6)' with a Venn diagram and a comparison column set to 'variant'. Below the editor is a table of variants.

ID	chr	position	ref	alt	dbSNP rs id	gene symbol	snpeff transcript id	snpeff am	snpeff coc	snpeff effect	snpeff impact
102	1	1196863	T	C	rs6659787	UBE2J2	ENST00000400930	c.220+186		intron_variant	MODIFIER
112	1	235976	C	A	rs201583565	AP006222.2	ENST00000442116		1492	upstream_gene_variant	MODIFIER
149	1	1649842	G	T	rs113724695	CDK11A	ENST00000378633	c.325+955		intron_variant	MODIFIER
153	1	758324	T	C	rs3131955	RP11-206L10.1	ENST00000445118		4664	upstream_gene_variant	MODIFIER
256	1	1651071	T	C	rs372567872	CDK11A	ENST00000378633	c.228-177		intron_variant	MODIFIER
386	1	1310924	T	C	rs2765033	AURKAIP1	ENST00000338370		387	upstream_gene_variant	MODIFIER
405	1	1648946	T	C	rs909824	CDK11A	ENST00000378633	c.326-102		intron_variant	MODIFIER
453	1	1649866	C	T	rs74045994	CDK11A	ENST00000378633	c.325+931		intron_variant	MODIFIER
568	1	943907	C	G	rs2488992	ISG15	ENST00000379389		4896	upstream_gene_variant	MODIFIER
806	1	1654013	C	G	rs74045997	CDK11A	ENST00000378633	c.111+134		intron_variant	MODIFIER

The screenshot above shows the VariantGrid analysis screen. The node graph is on the left part of the screen, showing the user built filters.

Click a node to select it. This loads the node editor (top right) and a grid of the variants (see section below) in the node (bottom right).

Clicking on the node loads this editor window. The node editor is different depending on the [type of node](#).

2.3 Analysis Grid

The 1st column (ID) is special and contains a check box, a numbered link and an IGV logo. The check box is used to select rows manually. The link loads detailed information about that variant above the grid. The IGV link will view the locus in IGV (loading bam files associated with samples). See IGV Integration page. Clicking on a row highlights it. Select the “tagging” tab, then click on a label to tag/colour the row.

ANALYSIS NODES

3.1 Source Nodes

Provide a source of variants

3.1.1 All Variants



All variants in the database.

3.1.2 Cohort

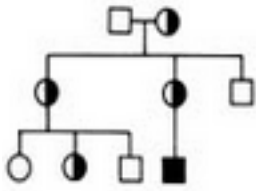


A collection of related samples, eg “control group” or “poor responders”

3.1.3 Classifications



3.1.4 Pedigree



Variants from family samples filtered by genotype according to inheritance models

3.1.5 Sample



A sample, usually one genotype (patient, cell or organism) with a set of variants.

3.1.6 Trio



Mother/Father/Proband - filter for recessive/dominant/denovo inheritance

3.2 Filter Nodes

These nodes filter variants connected to the top of them

3.2.1 Built In Filter



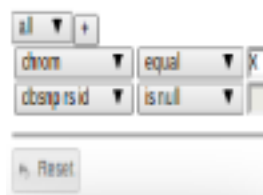
Built in filters used in node counts eg High or Moderate Impact / OMIM / ClinVar Pathological

3.2.2 Damage



Filter to damage predictions

3.2.3 Filter



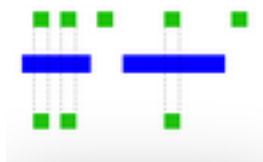
Filter based on column values

3.2.4 Gene List



Filter to a list of gene symbols

3.2.5 Intervals Intersection



Filter based on intersection with genomic ranges (eg .bed files)

3.2.6 Merge



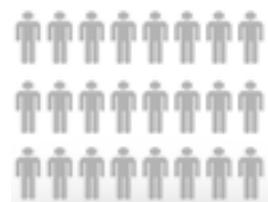
Merge variants from multiple sources

3.2.7 Phenotype



Filter to gene lists based on ontology keywords

3.2.8 Population



Filter on population frequencies in public databases (gnomAD/Exac/1KG/UK10K) or number of samples in this database.

Max population frequency of % in ticked database(s) below.

☒ [gnomAD](#) All genomes + exomes. 138,632 individuals.

- ☒ **African/African American**
- ☐ **Ashkenazi Jewish**
- ☐ **East Asian**
- ☐ **Finnish**
- ☐ **Latino / Mixed Amerindian**
- ☒ **Non-Finnish European**
- ☐ **South Asian**
- ☐ **Other**

☒ [1000 genomes 1kg Phase3_v5](#). Global pop. ~2,500 individuals

☒ [UK10K project](#) WGS for controls. 3,781 individuals

☐ [Exome Sequencing Project](#) Contains disease cohorts. All, ie EA+AA - 6,503 individuals

☒ [ExAC - Exome Aggregation Consortium](#) Unrelated, from disease and population studies. ~60,706 individuals

Restrict to samples in this database ☒

Keep internally classified (likely) pathogenic: ☒

Max percent: (Note: results vary over time with # of samples in database)

Max count: () (of the 1118 samples in the database)

save

3.2.9 Tags



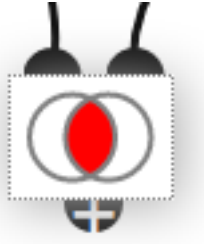
Filter variants to those that have been tagged

3.2.10 Tissue Expression



Filter based on tissue specific expression (from Human Protein Atlas)

3.2.11 Venn



A filter based on set intersections between parent nodes


3.2.12 Zygoty

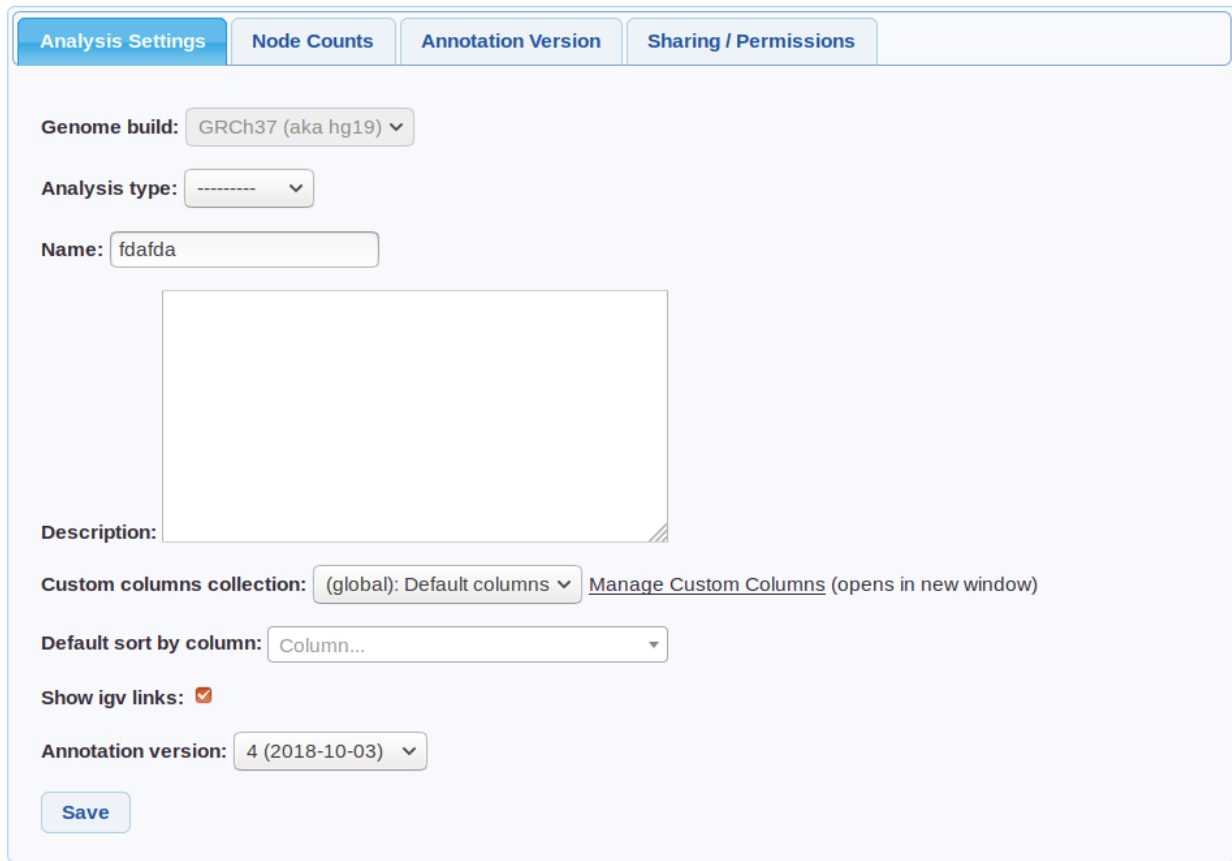


Compound HET and other Zygoty filters

ANALYSIS - ADVANCED

4.1 Analysis settings

In an analysis click the  Settings icon to open the analysis settings page.



[Force Reload Nodes](#) [Close](#)

Analysis

settings screenshot

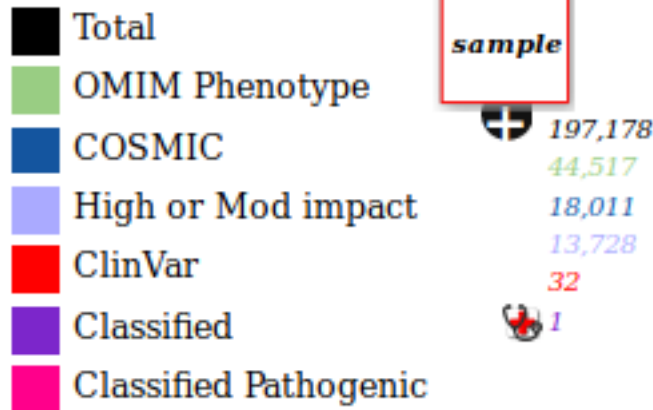
- **Genome build** - Cannot be changed. Only data (eg VCF samples) from this build can be used in the analysis.
- **Analysis type** - One of (Singleton/Cohort/Trio/Pedigree) set at creation if using an auto-analysis.
- **Custom columns** - Columns to use - from *customise columns*. Default set in *user settings*

- **Default sort by column** - Can be used for example to make the grid always sort by gene.
- **Annotation Version** - The *Annotation Version* used.

4.2 Node Counts

The numbers below a node are counts of variants that meet a certain criteria. The colours correspond to names in bottom left hand legend, eg in the image below, there are 32 ClinVar (Likely) Pathogenic variants in that node.

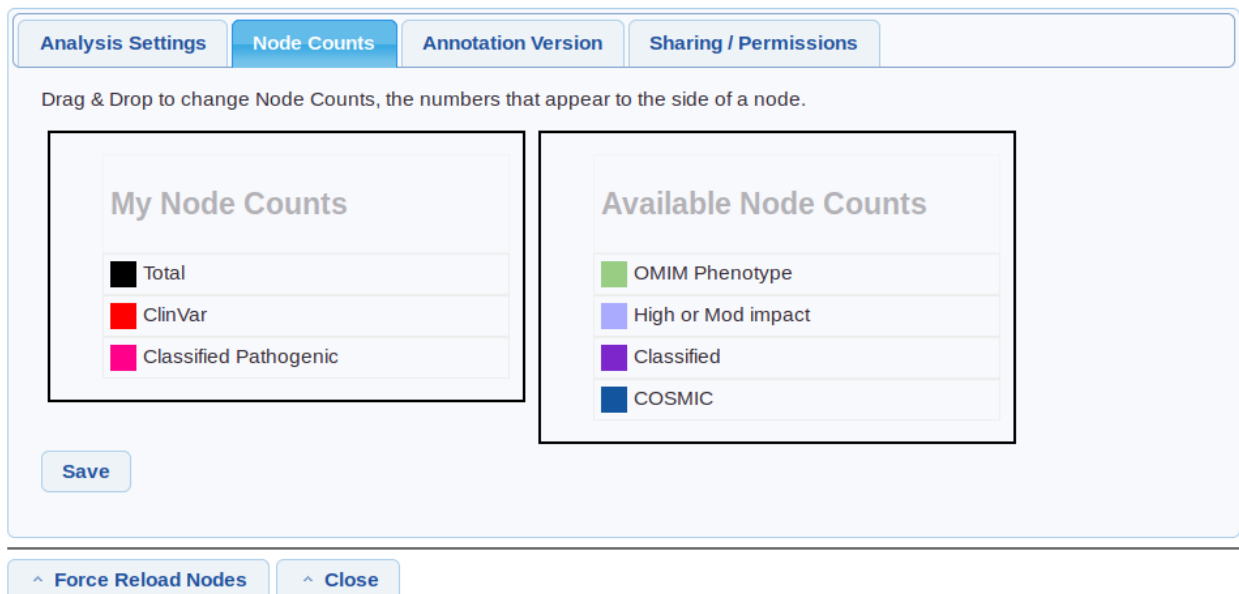
Node Counts:



Node with counts

Click on a count to load the variants in the node that meet that criteria, eg clicking on the red 32 would just load the ClinVar variants.

To edit which node counts are shown, open analysis settings, then select the “node counts” tab.

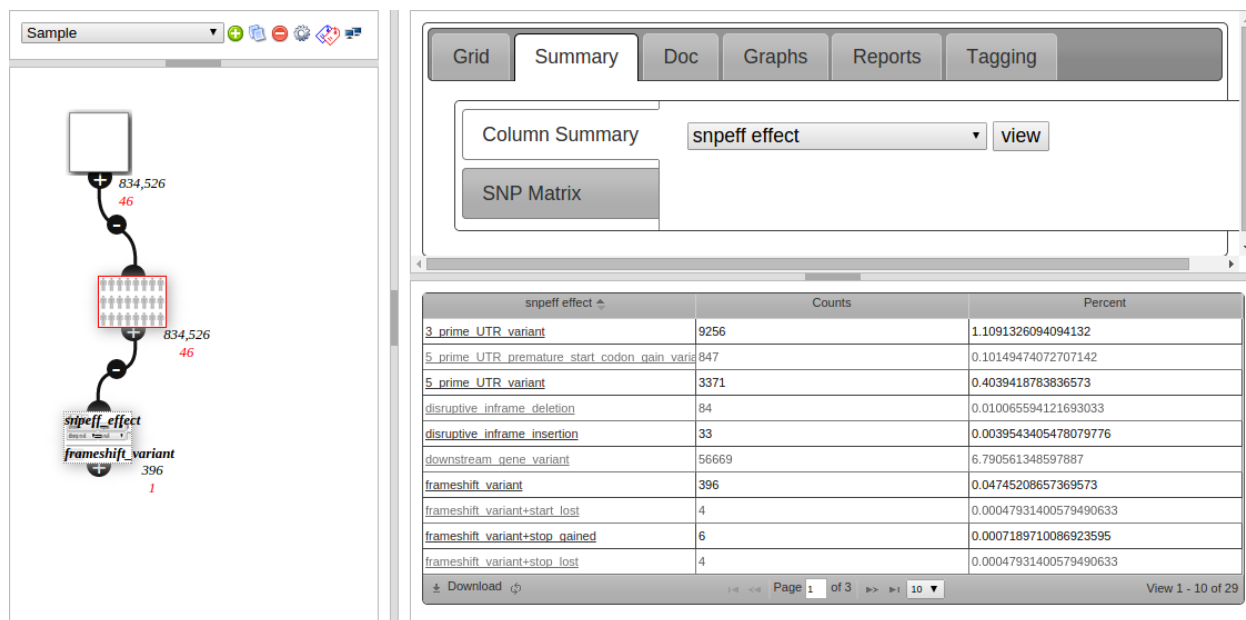


Settings/Node

counts

Drag and drop the node counts to show/hide them and change the order.

4.3 Column Summary



The screenshot displays the VariantGrid interface. On the left, a genomic track shows a red box highlighting a region. Below it, a filter node for 'snpeff_effect' is set to 'frameshift_variant' with 396 entries. On the right, the 'Summary' tab is active, showing a table of counts and percentages for various snpeff effects.

snpeff effect	Counts	Percent
3_prime_UTR_variant	9256	1.1091326094094132
5_prime_UTR_premature_start_codon_gain_variant	847	0.10149474072707142
5_prime_UTR_variant	3371	0.4039418783836573
disruptive_inframe_deletion	84	0.010065594121693033
disruptive_inframe_insertion	33	0.0039543405478079776
downstream_gene_variant	56669	6.790561348597887
frameshift_variant	396	0.04745208657369573
frameshift_variant+start_lost	4	0.00047931400579490633
frameshift_variant+stop_gained	6	0.0007189710086923595
frameshift_variant+stop_lost	4	0.00047931400579490633

At the bottom of the table, there is a 'Download' button and a pagination bar showing 'Page 1 of 3' and 'View 1 - 10 of 29'.

Node

Summary

The second tab (Summary) is used to view what values are in a column. Qualitative data is counted and shown in a grid, such as snpEFF Effect in the screenshot below:

Clicking on the link in the 1st column creates a child node filtering to that value. This is useful for getting an overview then drilling down into your data.

The screenshot shows 396 entries under “frameshift variant”, and the filter node created underneath the current (red bordered) node, which is configured to filter to snpeff_effect = frameshift variant, and also has 396 variants after filtering.

Quantative data (numbers, such as for the af_1kg column (1000 Genomes Alt Frequency)) is shown as a box-plot.

VARIANT TAGGING

A tag is a label (such as “Cancer” or “Investigate”) which you can use to label and track variants in an analysis.

5.1 Create tags

Menu: [settings] -> [tags]

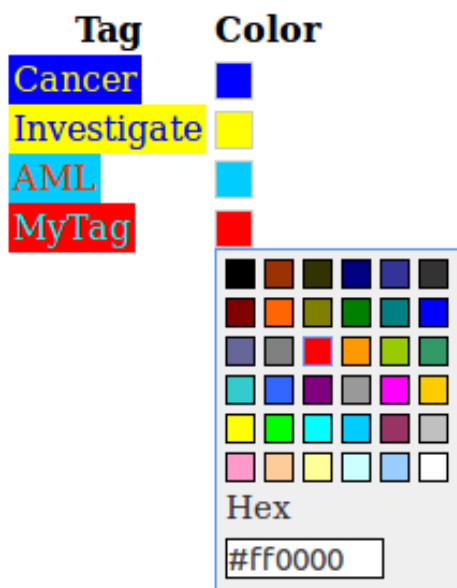
Tags

Tags names must be alphanumeric (no spaces or special characters)


Tag:

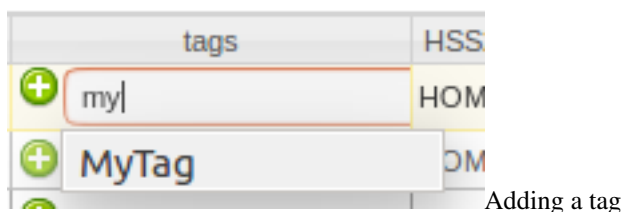
✓ Tag created successfully.


Click the colored box on the right to change background color



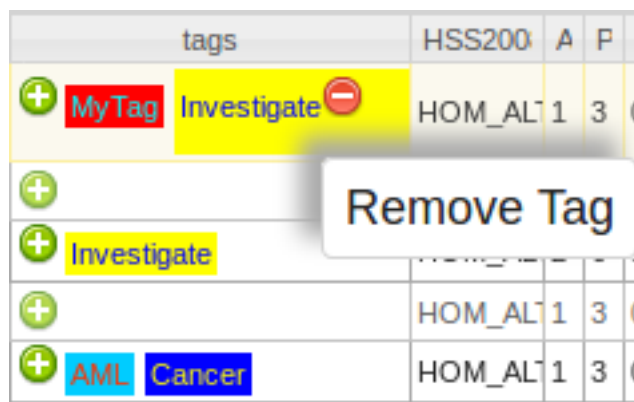
5.2 Tagging variants

In an analysis, click the  Add icon in the “tags” column then auto-complete your tag.



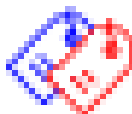
To remove a tag - clicking on the tag. The tag will grow in size, and a  delete symbol will appear. Click it to remove

the variant tag.



Removing a tag

5.3 Using tags

Click the  tag icon on the toolbar to view all Tags in an analysis

The 'Tags' panel shows a list of tags for the sample 'HSS2008 (HET, HOM_ALT)'. The 'All tags for analysis' section shows a dropdown menu for selecting a tag and a 'save' button.

ID	ch	position	ref	alt	dbsnp rs id	gene symbol	tags
<input type="checkbox"/>	16	34989694	G	A	rs34151874	5S_rRNA	MyTag
<input type="checkbox"/>	5	68281684	C	CA	rs563058488	7SK	Cancer
<input type="checkbox"/>	5	68309961	T	C	rs202010758	7SK	Investigate
<input type="checkbox"/>	5	68323790	A	G	rs10214127	7SK	AML Cancer
<input type="checkbox"/>	20	10030188	T	A	rs652633	ANKEF1	AML Cancer
<input type="checkbox"/>	19	55530035	C	T	rs1654416	GP6	Investigate
<input type="checkbox"/>	5	66459878	G	C	rs1705399	MAST4	Investigate

Page 1 of 1

To filter to specific tags - add a tag node, and use it like any other node to filter variants to just those that have been tagged.

Tag Filter

Grid | Summary | Doc | Graphs | SQL

Analysis wide: ☐
Tag: Cancer










ID	ch	position	ref	alt	dbsnp rs id	gene symbol	tags
<input type="checkbox"/>	5	68304573	T	C	rs4624745	ZSK	<input type="button" value="+"/> Cancer
<input type="checkbox"/>	5	68309961	T	C	rs202010758	ZSK	<input type="button" value="+"/> Cancer








You can view all tagged variants on a page, via menu: [analysis] -> [Tagged Variants]

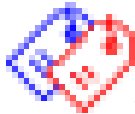
ANALYSIS CLASSIFICATION

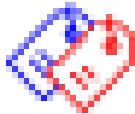
Recommended workflow to create a *classification* from a variant in an analysis:

1. *Tag* the variant with the “RequiresClassification” tag.

ID	ch	position	ref	alt	dbSNP rs id	gene symbol	tags	HSS2008	AD	AF	DP	GQ	PL	HSS2009	AD	AF	D
<input type="checkbox"/>  	12	49433599	T	G	rs147706410	KMT2D		HET	31	47.6	null	null	0	HET	24	44.4	ni
<input type="checkbox"/>  	12	49428694	T	C	rs146044282	KMT2D		HET	56	43.7	null	null	0	HET	56	49.1	ni
<input type="checkbox"/>  	6	10410466	T	G	rs776792762	TFAP2A	 RequiresClassification	HET	3	33.3	null	null	0	.	2	12.5	ni

 CSV
  VCF
 
 Page 1 of 1
 



 15



1. Click the  tags button, then then “Classification” tab.
2. Select the sample, then click the **[classify]** button.

KARYOMAPPING

7.1 Background

We handle the simpler case of a *Trio* with an affected child (ie proband/mother/father).

Variants are assigned to the following bins

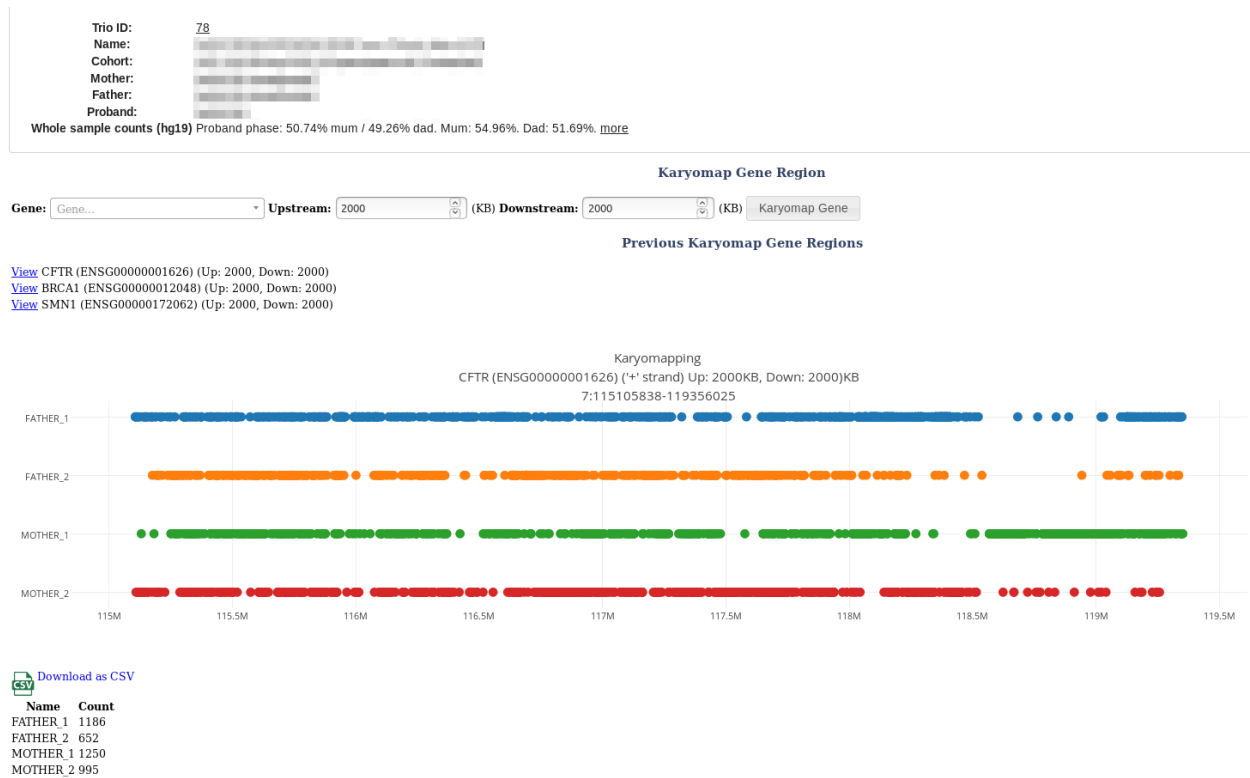
F1ALT: Paternally inherited, in phase with affected child, ALT variant. F1REF: Paternally inherited, in phase with affected child, REF variant. F2ALT: Paternally inherited, out of phase with affected child, ALT variant. F2REF: Paternally inherited, out of phase with affected child, REF variant.

And vice versa for the mother. The only variants that fall into each of these situations are:

7.2 Gene analysis

Menu: [analysis] -> [karyomapping]

Enter a gene name and click [**Karyomap Gene**] button.



7.3 Genome-wide analysis

A genome wide karyomap count is performed when you create a trio. This is useful for finding sample mixups. This is summarised as *Proband phase: 50.74% mum / 49.26% dad. Mum: 54.96%. Dad: 51.69%.* and is visible on the gene analysis screenshot above and the [Trio](#) page.

Proband phase shows the child’s marker percentage from each parent. Mum%/Dad% = Percent of parent markers that are in phase in proband.

Here are some examples for various Trios:

As a rough rule, you’d expect a minimum of 40% for an actual child.

ANNOTATION DETAILS

Annotation refers to all of the information about a variant, it is made from different components, including:

Variant-level annotation: Information specific to a base change. Examples include computational predictions and effects, and existing database entries (such as population frequency for the variant)

Gene-level annotation: Information about the gene (from RefSeq/Ensembl + other sources), matched from the variant's assigned transcript_id.

ClinVar: Clinical variant classifications from [ClinVar](#)

To see a description of each field, use menu: **[annotation]** -> **[descriptions]**

Annotation is shown on the [variant details](#) page, and in an [analysis](#), where it is used in filters and shown on the grid (see [customise columns](#))

8.1 Variant Level Annotation

The first time we see a variant, it is annotated by the variant annotation pipeline.

8.2 Annotation Versions

Each annotation component above is versioned and can be upgraded separately by the site administrator. To see the versions via menu: **[annotation]** -> **[versions]**

VariantGrid can store multiple annotation versions, which allows us to load historical analyses which return the same results as when they were first analysed, as well as updating from new sources regularly.

8.3 IVAT

VariantGrid uses IVAT developed by Jinghua (Frank) Feng from the CCB ACRF Cancer genomics facility.

8.3.1 SACGF Tiers

Tier 1

Novel variants, with evidence of being strongly damaging, and without any evidence of being artificial:

- Not in dbSNP, 1KG, UK10K, ExAC or ESP

- HIGH or MODERATE snpEFF impact, or mutating at branch point, at miRNA binding site, or at transcription factor binding site
- For a SNV: GERP > 4 or CADD > 30
- For an INDEL: not in LowComplexRegion
- Not in SegmentDup region
- No multi-ALT alleles were called

Tier 2

Extremely rare variants, with evidence of being strongly damaging, and without any evidence of being artificial:

- Not Tier 1
- Minor allele frequency (MAF) < 0.05% in 1KG, UK10K and ExAC.
- HIGH or MODERATE snpEFF impact, or mutating at branch point, at miRNA binding site, or at transcription factor binding site
- For a SNV: GERP > 3 or CADD > 20
- For an INDEL: not in LowComplexRegion
- Not in SegmentDup region
- No multi-ALT alleles were called

Tier 3

Very rare variants, with evidence of being potentially functional, and without any evidence of being artificial:

- Not Tier 1 or 2
- MAF < 0.2% in 1KG, UK10K and ExAC.
- HIGH, MODERATE or LOW snpEFF impact, or mutating at branch point, at miRNA binding site, or at transcription factor binding site
- For a SNV: GERP > 2 or CADD > 20
- For an INDEL: not in LowComplexRegion
- Not in SegmentDup region
- No multi-ALT alleles were called

Tier 4

Rare variants, with evidence of being potentially damaging. They can locate within the SegmentDup regions, and hence are with increased chance of being artificial:

- Not Tier 1, 2 or 3
- MAF < 0.5% in 1KG, UK10K and ExAC.
- HIGH, MODERATE or LOW snpEFF impact, or mutating at branch point, at miRNA binding site, or at transcription factor binding site
- For a SNV: GERP > 2 or CADD > 20

- For an INDEL: not in LowComplexRegion

Tier 5

Uncommon variants with potential damage effect, and can located in SegmentDup and LowComplexRegion and hence with significantly increased chance of being artificial:

- Not Tier 1, 2, 3 or 4
- MAF < 1% in 1KG, UK10K and ExAC
- Satisfying ***any one** of the three criteria below:
 - Annotated with HIGH, MODERATE or LOW snpEFF impact (aka. altering the exon or splice region)
 - Altering splicing branchpoint, miRNA binding site, or transcription factor binding site
 - GERP > 2 or CADD > 20

Tier 6

- Not Tier 1, 2, 3, 4 or 5


Notes:

A variant is classified as Tier 6, when all your samples are HOM-ALT at the variant and that ALT allele is common in 1KG, UK10K and ExAC (i.e. The frequency of the ALT allele is > 0.5 in anyone of 1KG, UK10K and ExAC). This applies before all the tiering above. From a trio sequenced with the Medical Exome Capture on our NextSeq machine in September 2016, below are the numbers of variants (called by GATK, mostly germline) for each tier:

VARIANT DETAILS

This page shows the [annotation](#) and other information about a variant.

The top of the page has an IGV link, and a link to the allele for this variant:

 **10:43615633 C>G (GRCh37 (aka hg19))**
Allele 350 (CA9034) (GRCh37, GRCh38)

An allele is genome build independent - ie hg19 and hg38 variants for same change point to same allele. The ID (CA9034) is from the [ClinGen Allele Registry](#)

9.1 Classifications

ID	HGVS	Clinical Significance	Condition	Curated Date	Flags
 My lab / vc0042	NM_000130.4(F5):c.1601G>A	Benign (1)		2019-08-06	 

Variant

Details - Classification section

This shows internal [classifications](#) for an allele (may have been classified against a different genome build)

The far right column contains [Classification Flags](#)

9.2 Transcripts

Variant annotation is calculated for each transcripts overlapping a variant. You can select each of the different transcripts to change which is being displayed.

9.3 Samples

At the bottom of the page is a grid of samples that contain the variant (and the zygosity and read information). Only samples you have permissions to view are shown, but a warning will be shown informing you that samples you don't have permission to see exist.

REPRESENTATIVE TRANSCRIPT

SnPEff calculates the damage effects for each transcript. The representative transcript is chosen as:

1. The most damaging transcript
2. If equally damaging, the canonical transcript defined by Ensembl is selected
3. If no canonical transcript exists, the longest transcript is selected. If more than one canonical transcript exists, the longest canonical transcript is selected.



UPLOADING DATA

Menu: **[data]** -> **[upload]**

Drag and drop VCF, bed, GeneList (.txt), CuffDiff and .PED (pedigree files) to upload.

Show last records

☐

✓  AS-145_WES_HiSeq_Variants.vcf	5.47 MB	<input type="button" value="Delete"/> <input type="checkbox"/>
✓  test.vcf	1.43 KB	<input type="button" value="Delete"/> <input type="checkbox"/>

You can either drag & drop files onto the page, or by selecting the **[Add Files]** button.

After the file has been transferred to the server, a spinning icon (⌂) will appear as the file is processed. The large link (eg “AS-145_WES_HiSeq_Variants.vcf”) takes you to the import processing page if you’d like to monitor the progress.

Once it has been successfully imported, a link will appear beneath the file (eg the “VCF” links above) allowing you to jump to the data page for this file.

MANAGING DATA

Menu: **[data]**

The data page displays all of your uploaded data such as (VCFs, Bed files, Pedigree Files etc)

Data is displayed in grids, with each data type in a separate tab.

You can enter parts of the name into an autocomplete search box to quickly find your files:

samples

VCF

bed file

gene lists

Pedigree .ped files

HSS232

HSS2326 (all_HM_samples.2017Jan.gatk.vcf.gz)

HSS2327 (all_HM_samples.2017Jan.gatk.vcf.gz)

HSS2328 (all_HM_samples.2017Jan.gatk.vcf.gz)

HSS2329 (all_HM_samples.2017Jan.gatk.vcf.gz)

HSS2320 (all_HM_samples.2017Jan.gatk.vcf.gz)

HSS2321 (all_HM_samples.2017Jan.gatk.vcf.gz)

HSS2322 (all_HM_samples.2017Jan.gatk.vcf.gz)

HSS2323 (all_HM_samples.2017Jan.gatk.vcf.gz)

HSS2324 (all_HM_samples.2017Jan.gatk.vcf.gz)

HSS2325 (all_HM_samples.2017Jan.gatk.vcf.gz)

	import status	
HSS2326	success	all
HSS2327	success	all
HSS2328	success	all
HSS2329	success	all
HSS2320	success	all
HSS2321	success	all
HSS2322	success	all
HSS2323	success	all
HSS2324	success	all
HSS2325	success	all
HSS2336	success	all
HSS2335	success	all
HSS2334	success	all
HSS2333	success	all

Click the link on the grid to view the file details page.

12.1 Sharing data

Users belong to groups (see [user settings](#)) that can share data. Ticking the **Show Group Data** checkbox will show this on a grid.

By default, you automatically share data (read-only) with your group.

To change data permissions, click the **[Data/Sharing]** tab:

151120_AHISEQTEST > VCF

SN1101

Details Variants Graphs QC Sharing / Permissions

Permissions

Group	Read	Write
my_group	<input type="checkbox"/>	<input type="checkbox"/>
public	<input checked="" type="checkbox"/>	<input type="checkbox"/>

save

Genelist Security

Set Genelist Security No Gene List Security set.

logged_in_users is a special group - and means everyone who has a VariantGrid account.

12.2 Search

Enter text into the search box in the top right hand corner and press enter or click Go.

search... Go help logout

Accepted inputs:

12.2.1 HGVS

We use [PyHGVS](#) library for parsing HGVS names, which supports 'c.', 'n.' and 'p'.

SOMATIC DATA

Somatic VCFs detected as somatic only (tumor minus normal) are analysed for [mutational signatures](#)

13.1 Allele Frequency

We do not import the AF value from the VCF, but instead *normalize* the data then recalculate AF to be $AD / \text{sum}(AD \text{ for all variants at locus})$

In an analysis, Sample, Cohort and Trio nodes can filter by allele frequency. For the Cohort and Trio nodes, **all** or **any** refers to requiring all samples to have allele frequency within the ranges or just one or more sample.

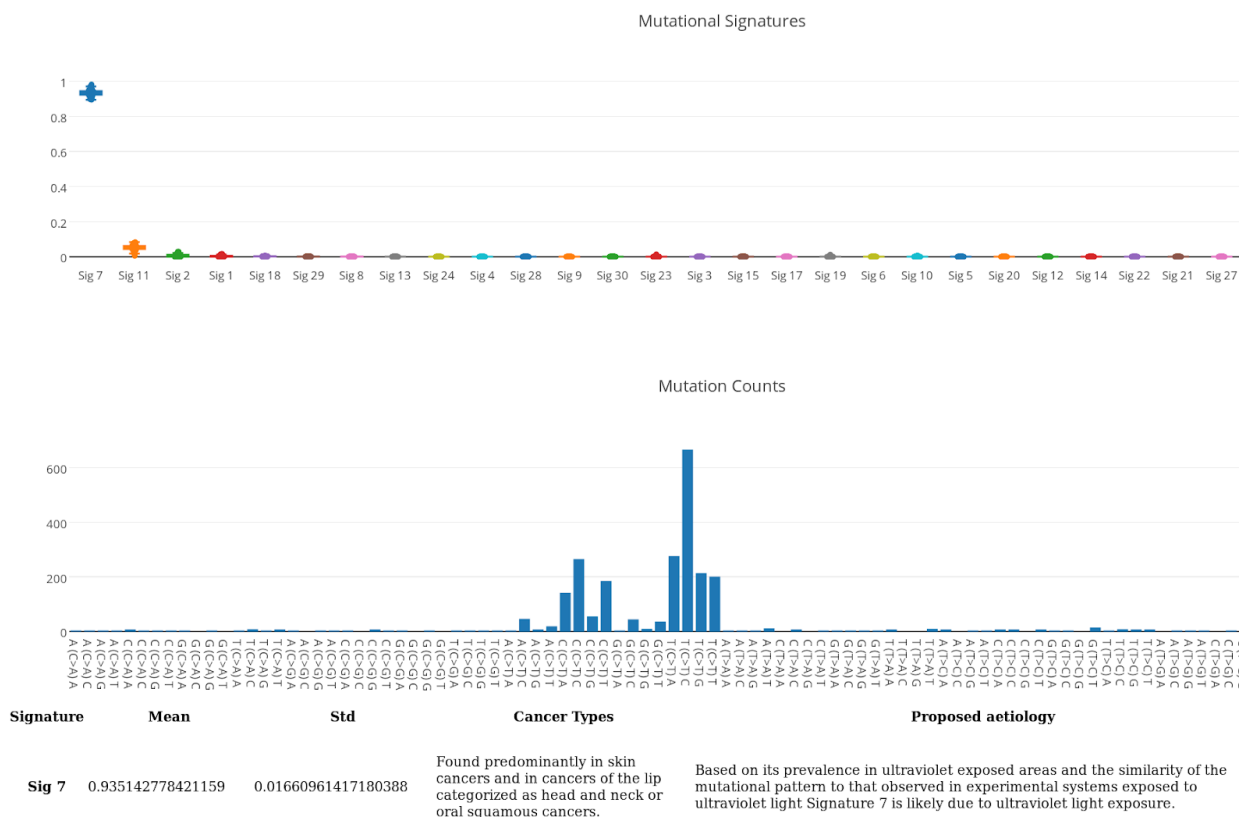
Allele Frequency all ▼ +

34	<input type="text"/>	<input type="text"/>	56	<input type="button" value="-"/>
0	<input type="text"/>	<input type="text"/>	16	<input type="button" value="-"/>
87	<input type="text"/>	<input type="text"/>	100	<input type="button" value="-"/>

MUTATIONAL SIGNATURES

Different types of cancer can have consistent somatic variants, see [Signatures of mutational processes in human cancer, Alexandrov et al 2013](#)

Sample [161130HamishScott_somvar_combined.snvs.HC.vcf \(161130HamishScott_somvar_combined.snvs.HC.vcf\)](#)
Summary Sig 7: 93.51%, Sig 11: 5.23%
snps 2375
Iterations 100
Sampling Fraction 0.8
Minimisation Strategy Least Squares



Mutational signatures are calculated during [VCF import](#) when the sample is detected as *somatic only*

Menu: **[data]** -> Sort samples grid by “Mutational Signature” column -> Click on entry.

Or click on the link in the “Mutational Signatures” at the bottom of the sample page.

Thanks to Paul Wang from the ACRF Cancer Genomics Facility for the code.

VCF / SAMPLES

15.1 VCF import

Variants are *normalized* upon import. We only import variants, filters and genotypes (we don't use INFO as we do our own annotations)

The VCF format can vary a lot, we have tested VCFs from the following variant callers:

- GATK
- FreeBayes

Each sample is assigned a “variants type” of *Unknown*, *Germline*, *Mixed* (single sample) or *Somatic only* (tumor minus normal).

This is determined by looking at the “source” entry in the VCF header, and matching it to an entry in **VCFSOURCE** object (setup by your administrator)

Samples with variants type of `_somatic only_` are checked for *mutational signatures*

15.2 Multi-sample VCFs

Multi-sample VCF files combined using bam files record the genotype for all samples at each variant position.

This allows you to differentiate between reference calls and no coverage - and is extremely important for Trios so that you can make correct calls about inheritance and denovo variants

You must use bam files, to re-call the genotypes for each position.

Consider 3 VCF files:

There's no way to tell if a variant not being present in a single sample VCF is due to having the reference allele or no coverage.

Merging just the VCFs (without supplying the bams) will give the genotypes of:

If you merge them using **GATK/Picard** using bam files - the caller will re-examine the reads over the locus, and make the genotype call.

Thus, if both parents had reference bases, the calls would be:

And you can be confident that it is a denovo variant, rather than just lacking coverage in one of the parent samples.

CHAPTER
SIXTEEN

SEARCH

TODO

GENE PAGE

Menu: **[genes]** -> **[genes]** then autocomplete a gene name.

You can also enter a gene name such as “GATA2” or “RUNX1” into the search box, or click on a link in [GeneGrid](#)

If you have [gene coverage](#) data, boxplots will be shown.

GENE LISTS

Menu: [genes]

18.1 Creating Gene Lists

Ways to create a gene list include:

- Upload a text file (see *upload*)
- Create via [GeneGrid](#)
- Creating manually (see screenshot below)

The screenshot shows the SA Pathology web interface. The top navigation bar includes links for sequencing, data, patients, tests, analysis, classifications, genes (highlighted), variants, and annotation. A search bar and a 'Go' button are also present. On the left sidebar, 'Gene Lists' is highlighted. The main content area shows a 'Jump to gene list' dropdown menu. Below this, a 'New GeneList' button is highlighted. A table titled 'Gene Lists' displays a list of gene lists with columns for ID, name, and Uploaded by. The table contains 10 rows of data. At the bottom of the table, there is a pagination bar showing 'Page 1 of 15' and a dropdown menu for '10' items per page.

ID	name	Uploaded by
COIs.txt		Im cintyre
COIs_per_line.txt		Im cintyre
Non-im m une fetal hydrops_20150826.txt		kbrion
HCM 061015.txt		DouglasEvelyn
Alport x3.txt		DouglasEvelyn
4_MED12_MC-45876.txt		Isanchez
MED12 gene list.txt		Isanchez
10_TICAM1_RM_46834.txt		Isanchez
130930_NB501009_0007_AHCNTTBCXX_10_FALCONI_ANAEMIA_MC_46600_S10_QC.txt		Irawlings
FALCONI_ANAEMIA_KABUKI_MC_46600.txt		Irawlings

on New GeneList

Click

SA
PATHOLOGY

sequencing | data | patients | tests | analysis | classifications | genes | variants | annotation

Gene Lists

Genes

GeneGrid

Jump to gene list:

Gene List...

Create New GeneList

Name:

Training SKS


ACTC1, MYL3, PLEX, CPA1D9, FRG2

Create GeneList






User



GeneInfo



Gene Lists

ID	name	
	GOIs.txt	Im cint
	GOIs_per_line.txt	Im cint
	Non-immune fetal hydrops_20150826.txt	kibion
	HCM 061015.txt	Dougle
	Alport x3.txt	Dougle
	4_MED12_MC-45876.txt	Isanoh

name, genes and click save

Enter

18.2 Using gene lists in analyses

To quickly filter to a gene list in an *analysis*

1. Add and connect a gene list node
2. Select “Custom Gene List” in the top right node editor
3. Enter the genes into the text box and click “Save”

Gene list

Grid Summary Doc Graphs SQL Tagging

Named Gene Lists

Custom Gene List

BRCA1 BRCA2

save

ID	chr	position	ref	alt	db/np rs id	gene symbol	snpeff transcript id	snpeff am
12361004	17	41219853	ATT	ATT		BRCA1	ENST00000471181	c.5050-14
12366274	17	41279968	T	G		BRCA1	ENST00000471181	
12376719	17	41197939	AT	ATT		BRCA1	ENST00000471181	c.5531-12

GENE GRID

Menu: [genes] -> [gene grid]

GeneGrid allows quick comparisons between gene lists and adding/removing genes from them. Genes are rows and gene lists are columns.

The screenshot shows the GeneGrid interface with the following configuration:

- SA Pathology current test: Pathology Test...
- SA Pathology historical test: alports_syndrome (v1)
- User: Gene List...
- Fulgent: Alport Syndrome NGS Panel (3 x gen...)
- GeneInfo: Gene List...
- Invitae: Gene List...
- Enrichment Kit: medical_exomes
- Panel App Panel: Panel App Panel...
- Human Phenotype Ontology: Phenotype...
- OMIM: OMIM:104200 ALPORT SYNDROME... x

Custom Gene List:

Name:

Gene names...

[Add Custom Gene List](#)

Evidence columns:

- CinGen
- PanelApp
- Color
- Coverage

Grid Instructions:

To add a gene to a column, hover over the empty space, then click the symbol that appears. To remove a gene, click the name, then the red button that appears.

Gene	roche_1k_disease (version 6)	medical_exomes	alports_syndrome (v1)	Training SKS	Alport Syndrome NGS Panel	OMIM: 104200 ALPORT SYNDROME, AUTOSOMAL DOMINANT
Gene...						
	% at 20x	% at 20x				
A2ML1	100.00	100.00		A2ML1		
ACTC1	100.00	100.00		ACTC1		
COL4A3	100.00	99.46	COL4A3		COL4A3	COL4A3
COL4A4	100.00	99.86	COL4A4		COL4A4	COL4A4
COL4A5	100.00	100.00	COL4A5		COL4A5	
CPAMD9				Matched A2ML1 UCSC Alias		
FRG2	-	9.29		FRG2		
MYL3	100.00	100.00		MYL3		
PLEX				Could not match gene symbol		

19.1 GeneGrid screen

You can copy/paste the URL at any time to re-create a particular comparison.

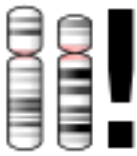
Choose lists from the top left select boxes, or manually paste in gene names into the **Custom Gene List** text entry box. Click the red delete button to remove a gene list column.

In the top right are optional evidence columns which provide information about genes.

See [Gene Coverage](#) for details on how the **% at 20x** values in the Enrichment Kit columns are calculated. Enrichment kits are automatically added when a *pathology test* that uses it is added to the grid.

19.2 Gene Info

Small icons next to gene names on the left of the grid indicate the gene has one of these attributes:



Alternative Haplotype



Pseudogenes



Triplet repeat disorders

GENE COVERAGE

Gene Coverage refers to how well a gene was covered by high throughput sequencing reads. This is useful to know how confident you can be about a lack of variant calls in a region.

Having gene coverage associated with a VCF sample allows you to be warned in an *analysis* when a gene in a gene list is below a threshold (default: 20x) and you may be missing some variants. The node will flash yellow, and the “genes” tab will be highlighted yellow so you can view which genes have low coverage.

Boxplots of sample coverage for genes are on the [gene page](#)

20.1 Canonical Transcripts

Many genes have multiple transcripts, but people want only one value for each gene.

This is achieved by choosing a single (representative or canonical) transcript, and use that transcripts value for the gene.

A CanonicalTranscriptCollection is a list of gene:transcript mappings imported into the system. The administrator can import different collections, linking them to EnrichmentKits and setting a system default.

20.2 Sample QC metrics

You can *upload* gene coverage files (.txt files) which use the system default canonical transcripts. You can then associate them with a *sample from a VCF*

Sample QC coverage loaded via *sequencing features* - and automatically choose transcripts based on EnrichmentKit

20.3 GeneGrid EnrichmentKit coverage

The per-gene QC metrics for an EnrichmentKit on the GeneGrid page are from *Gold Standard Runs*, using the canonical transcripts for that EnrichmentKit.

PATHOLOGY TESTS

Menu: [tests] -> [manage tests]

Pathology Tests are curated, versioned gene lists offered as a diagnostic test. There can be multiple versions of a test.

A Pathology Test Version is a specific versions of a pathology test.

21.1 Active tests

Each pathology test has at most one currently active test - the one available for test orders.

An active test is the most recent confirmed version of a pathology test.



Active test logo



All other versions of tests

The curator confirms & adds a time-stamp by clicking the **Confirm Test** button. Once a test has been confirmed it cannot be modified, and any further changes must create a new test version.

21.2 Requesting gene changes

Only the curator can modify a test, everyone else can make modification request but these must be approved by the curator. Contact an administrator to change curator for a test.

Make gene modification requests on the [GeneGrid](#) page.

Request gene addition

BRCA2

save

cancel

CDH1	CDH1
GATA2	+1
MLH1	MLH1, -1

The gene symbols in the pathology test column are always what is in the test. The +/- numbers (green background for add, blue for delete) in the image above are counts of requested additions/removals for that gene.

To request a gene addition: Add genes to the [GeneGrid](#), then click on an empty space where the gene should be. *To request a gene deletion:* Click on an existing gene, then the red delete symbol which appears.

In both cases a box will appear where you can enter a brief justification of the request. Only put a brief summary - please put in depth evidence such as linking a disease with a gene or adding literature on the gene page (click on the the gene name on the left column of the grid to open gene page in a new window).

21.3 Accepting gene changes

The curator can see any pending requests on the pathology test version page, where they can accept/reject them.

Gene Addition Requests

GATA2 ☒ Reject request ☐ Add Gene

Operation	User	Last modified	Comments
Add	dLawrence	Sept. 21, 2018, 10:42 a.m.	This gene should be part of the test

Gene Deletion Requests

MLH1 ☒ Reject request ☐ Remove Gene

Operation	User	Last modified	Comments
Remove	dLawrence	Sept. 21, 2018, 10:42 a.m.	This gene doesn't have enough evidence

Create new test version

Any genes added will have the user, date and brief justification comment from the addition request stored on the “Modification info column” which you can see on the grid of genes for a pathology test version.

The outcomes for any processed requests can be seen by all users at the bottom of the page:

Outcome	Operation	User	Last modified	Comments
Accepted	Add	dLawrence	Sept. 21, 2018, 10:44 a.m.	This gene should be part of the test
Outcome	Operation	User	Last modified	Comments
Rejected	Remove	dLawrence	Sept. 21, 2018, 10:44 a.m.	This gene doesn't have enough evidence

TEST ORDERING

Create patients to store phenotype information and link multiple samples (eg tumor/normal) together.

23.1 Searching

You can search by name, code or free text in the phenotype description.

Click the graph of phenotype terms to filter the grid to patients with that phenotype.



grid filtered to microcephaly

23.2 Patient records

Import a CSV to create patients in bulk. Click the **patient record imports** link at the top of the page, then can select to download an example CSV with your samples pre-filled, so it's easy to match your patients to your existing data.

You can also create patients one at time via a form, by clicking the **Create New Patient** link just above the grid.

23.3 Other sources of patients

Patients can be created via the pathology test ordering system.

On a private server (eg diagnostic lab intranet), patient records can be automatically created via your LIMS/Patient records system (speak to your administrator)

23.4 Other

Family Code is useful for linking together patients

The system can be configured to show/hide names, or convert birthdates to years depending on your privacy needs.

PHENOTYPES

It is useful to store phenotypes, diseases and genes for a patient. Having this information well structured and using controlled terms is very useful as it allows us to:

- Filter variants to genes associated with a disorder
- Know phenotypes for patients that share variants
- Perform analyses across disease cohorts (is the same variant or gene responsible for the disease or are they different?)
- Track per-disease solve rates

24.1 Assigning Terms to Patients

You can auto-complete terms in the boxes, which will be added to the bottom of the patient description.

Or, you can type plain text and we'll automatically match your words to Human Phenotype Ontology, OMIM and Gene Names.

Matched terms will be highlighted to the right of the description box.

The screenshot shows a patient record form with the following fields and content:

- First name:** [Redacted]
- Last name:** [Redacted]
- Date of birth:** [Redacted]
- Date of death:** [Redacted]
- Sex:** [Redacted]
- State:** [Redacted]
- Description:** (See Patient Phenotypes Guide)
 - From NGS Database on 2017-04-20
 - From Phenotype: Unexpected interstitial lung disease in [Redacted] (DO NOT proceed until array results in)
 - From Comments: Not to proceed until SNP array results available. ABCA3, AP3B1, CSF2RA, CSF2RB, SFTPB, SFTPC, FOXF1, NKX2-1, SFTPA2, SLC7A7, TERT, TINF2, HPS1, HPS4, DKC1, FLNA (plus genes associated with Primary Ciliary Dyskinesia as second phase analysis if required).
 - From Comments2: Emailed [Redacted] re array result - neg, go ahead with NGS [Redacted]

Buttons: reset, save

Patients

grid filtered to microcephaly

24.2 How phenotype term matching works

Everything after “--” on a line is ignored and can be used for comments.

The text is broken up into sentences based on punctuation and new lines.

The sentence is separated into words, and then sub sets of the words in order are created, and sorted largest to smallest. For instance:

```
The cat sat on the mat
cat sat on the mat
The cat sat on the
sat on the mat
cat sat on the
The cat sat on
The cat sat
on the mat
sat on the
cat sat on
the mat
cat sat
The cat
on the
sat on
mat
the
sat
cat
The
on
```

This allows us to find the biggest matches first. If a match occurs, the unmatched parts of the sentence continue to be searched until there is nothing left. If no match occurs for a sentence, we try the next smaller one.

Some filtering is done to avoid matching to common words and terms. For instance “Trio” is a gene name, but we will not match it as a gene if the sentence also contains the name of an enrichment_kit or one of the words: “exome”, “WES”, “father” or “mother”.

Matching occurs first against [Human Phenotype Ontology](#) terms and synonyms, and [OMIM](#) terms and aliases.

If no exact match is found, we try again using mismatches - 1 mismatch (including insertions/deletions) is allowed for two or more words.

For single words, we only allow mismatches if the word is more than 5 letters long and made entirely of letters (ie no digits or symbols).

Single words are then matched (exact with no mismatches) to gene names.

Sometimes there will be multiple matches, eg “PKD1” will map to both the OMIM term PKD1 (POLYCYSTIC KIDNEY DISEASE 1) and the gene PKD1. This is usually what people want as the gene is associated with the disorder.

COHORTS

Menu: [patients] -> [cohorts]

A cohort is a collection of samples, which you can analyse as a group. A multi-sample VCF automatically becomes a cohort, but you can create your own to organise your own samples.

25.1 Create a new cohort

From the cohort page, enter the name of a cohort and click the **Create** button.

This opens the Add/Remove samples tab. Add samples to your cohort by auto-completing sample names in the Enter to add box, or filter the grid, select the checkbox to the left of a sample, and click the green arrow to add, or red button to delete.

Once you have finished adding/removing samples, click save. This processes the cohort so it can be used in analyses.

25.2 Create from a larger cohort

You can create a smaller cohort from a larger one. Select at least 2 samples then click the [Create cohort from selected samples] button. Selecting exactly 3 samples allows you to create a **Trio** which allows for simpler analyses.

Details

Sharing / Permissions

Name:

190208HamishScott_WGS_1

Date:

2019-03-20 11:57:29

User:

dlawrence

Project:

Import status:

success

Processing

[View upload processing](#)

Samples

Bulk Set Fields

Sample	Variants (passed)	VCF Sample Name	Name	Patient	Specimen	BAM path
<input type="checkbox"/> Sample 2745	5251965 (4929662)	FD02523372	FD02523372	Patient...	Specimen...	
<input checked="" type="checkbox"/> Sample 2746	5317840 (4990108)	FD02523383	FD02523383	Patient...	Specimen...	
<input checked="" type="checkbox"/> Sample 2747	5238972 (4910634)	FD02523385	FD02523385	Patient...	Specimen...	
<input checked="" type="checkbox"/> Sample 2748	5254771 (4930620)	FD02523386	FD02523386	Patient...	Specimen...	

Update VCF

☐ Perform trio analysis using template

☒ Create cohort from selected samples

More than 3 samples - select exactly 3 using checkboxes on the left. 3 samples selected.

a sub-cohort

Creating

25.3 Cohort Analyses

Use the Cohort Node to filter by counts within the cohort (eg in 7 out of 8 of the samples) or zygosity. (see screenshot below).

Family 24421 (7 of 8)

2,107
220
362

private snps

4
1

GridSummaryDocGraphsSQL

Cohort: Family 24421 (9 samples) [View Cohort](#)

Minimum: 7, Maximum: 8 of 9 samples.

Show reference alts (non-variants) ☐

Simple Zygosity

Per Sample Zygosity

	Show In Grid	Hom Ref	Het	Hom Alt	No Record	Toggle Row
HSS2095	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
HSS2096	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
HSS2097	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
HSS2098	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
HSS2099	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
HSS2100	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
HSS2101	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
HSS2102	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Toggle Column	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	

save

Cohort

Node filtering by zygosity

Quickly create an analysis using the cohort by clicking “Create new analysis for cohort” on the details tab of the cohort page.

There are some other analyses you can perform from the cohort/VCF page, eg:

Legend

Stopgain	Frameshift	Missense	Splice site	Other	Not tested
----------	------------	----------	-------------	-------	------------

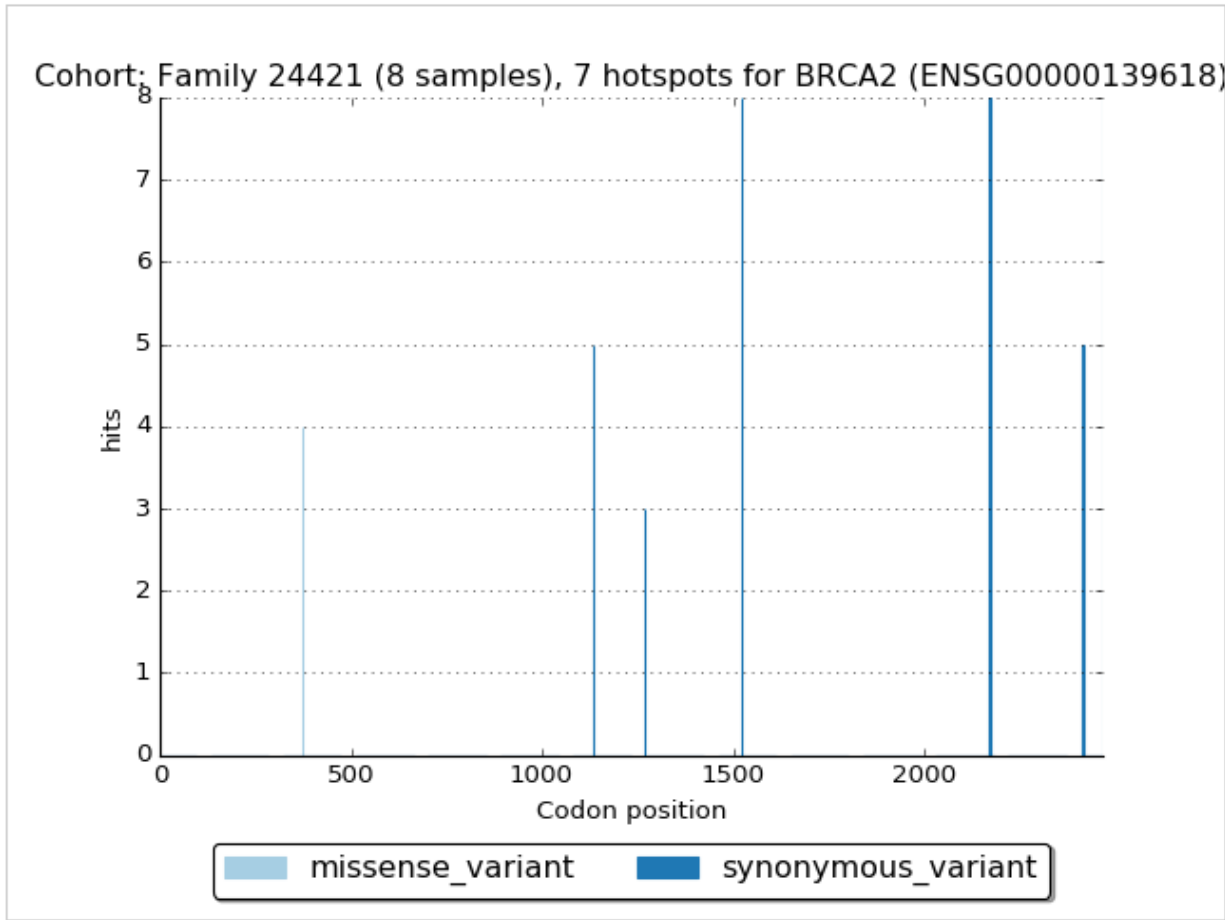
	S000	S001	S002	S004	S005	S006	S007	S008	S009	S010	S011	S012	S013	S014	S015	S016	
Age																	
HPO																	
MIM																	
RUNX1	1	1	1					1				1	1	1	1	3	
ASXL1			1											2	1	2	
BCOR	1	1						1				1	1		3	1	
BCORL1			1									1		1	5	6	
CBL																	
CDC25C			2	1	1	1	1							1			
CDC27																	
CEBPA																4	
CREBBP														2			
DNMT3A			1												1	1	
EZH2					1										3	4	
FLT3			1									3		1	1		
GATA2															1	5	
IDH1																	
IDH2																	
JAK1																	
JAK2														1			
JAK3	1	1	1					1									
KIT				1			1										
KMT2A														1	1	2	

Gene/Sample

Matrix

Cohort: Family 24421

Gene × BRCA2 (ENSG00000139618) View Gene



Hotspots graph

Cohort

TRIOS

Menu: [patient] -> [trios]

A trio is a collection of 3 samples (mother/father/proband) which are frequently analysed together in high throughput sequencing, as they have a number of standard analyses.

26.1 Creating a trio

It is far better to upload a trio within the same *multi-sample VCF*. If not, you must first create a cohort containing the 3 samples/

View the VCF or cohort, select exactly 3 samples then click the [Perform Trio Analysis using template] button.

Details
Sharing / Permissions

Name: 190208HamishScott_WGS...
Date: 2019-03-20 11:57:29
User: dlawrence
Project: -----
Import status: success
Processing [View upload processing](#)

Samples

[Bulk Set Fields](#)

Sample	Variants (passed)	VCF Sample Name	Name	Patient	Specimen	BAM path
<input type="checkbox"/> Sample 2745	5251965 (4929662)	FD02523372	FD02523372	Patient...	Specimen...	
<input checked="" type="checkbox"/> Sample 2746	5317840 (4990108)	FD02523383	FD02523383	Patient...	Specimen...	
<input checked="" type="checkbox"/> Sample 2747	5238972 (4910634)	FD02523385	FD02523385	Patient...	Specimen...	
<input checked="" type="checkbox"/> Sample 2748	5254771 (4930620)	FD02523386	FD02523386	Patient...	Specimen...	

Update VCF
☐ Perform trio analysis using template
☐ Create cohort from selected samples

More than 3 samples - select exactly 3 using checkboxes on the left. 3 samples selected.

Creating

a Trio

The Trio wizard will now open, showing the 3 samples and patient / phenotype info. Assign samples (1 each to mother/father/proband) and check mother or father affected if they also have the disorder.

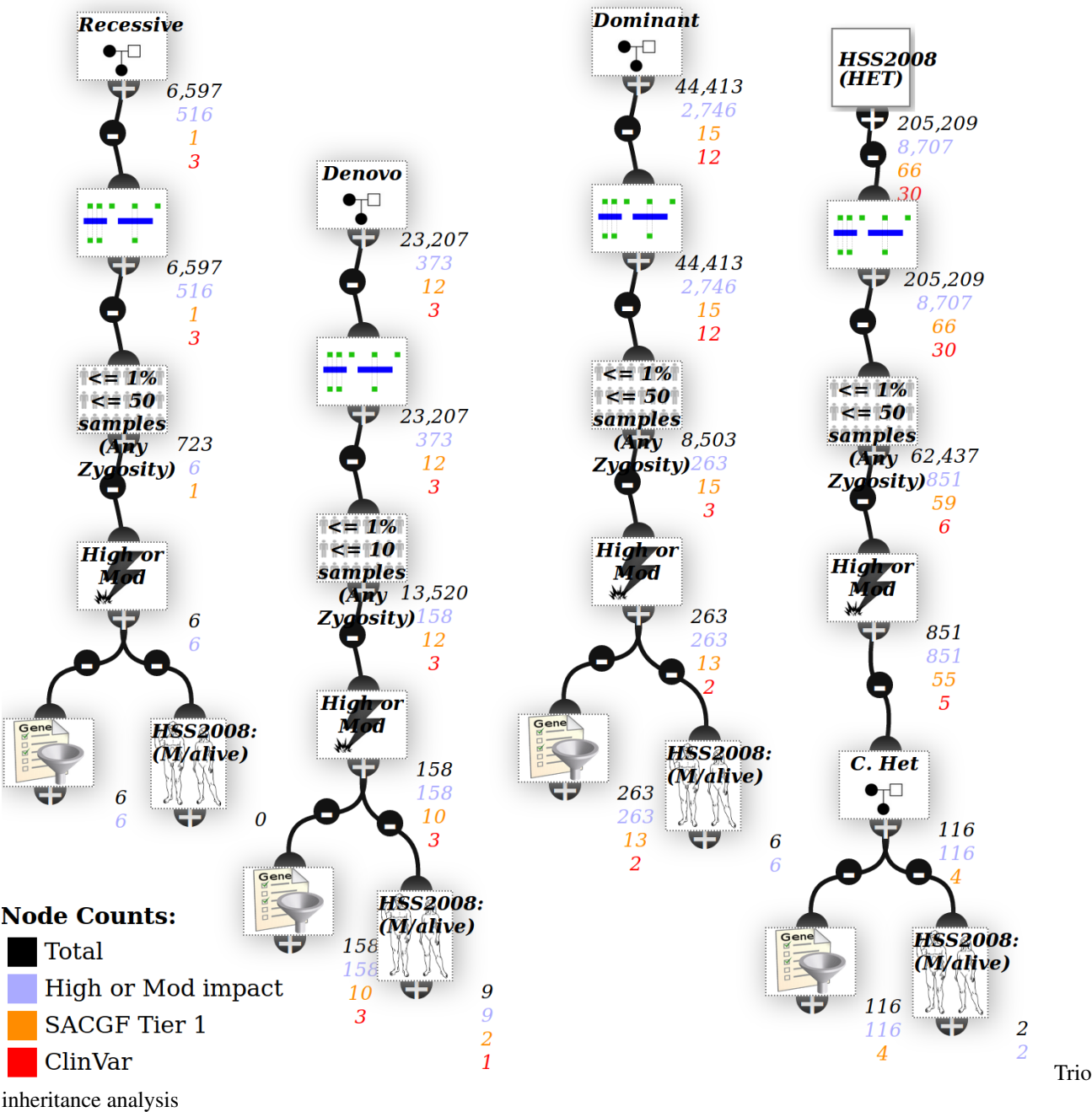
26.2 Digital karyomapping

By checking a trio's zygosity, it's possible to perform a number of relatedness calculations, see *karyomapping*.

A genome-wide count is automatically performed, and a summary provided on the trio page - this is useful for checking for sample mix-ups.

26.3 Trio inheritance analysis

An analysis is created using different inheritance models (see below). If either parent is affected it will also use an autosomal dominant inheritance model.



26.3.1 Require Zygosity Calls

By default, the filters are strict and require zygosity calls in all patients - for instance the recessive inheritance model requires a variant to be HOM in proband and HET in both parents.

However that may be overly strict - one parent may have low coverage, with no variants recorded at that locus.

Click on an Trio node to open the editor - unchecking the **require zygosity calls** box is less strict and allow for variants that are missing due to low coverage.

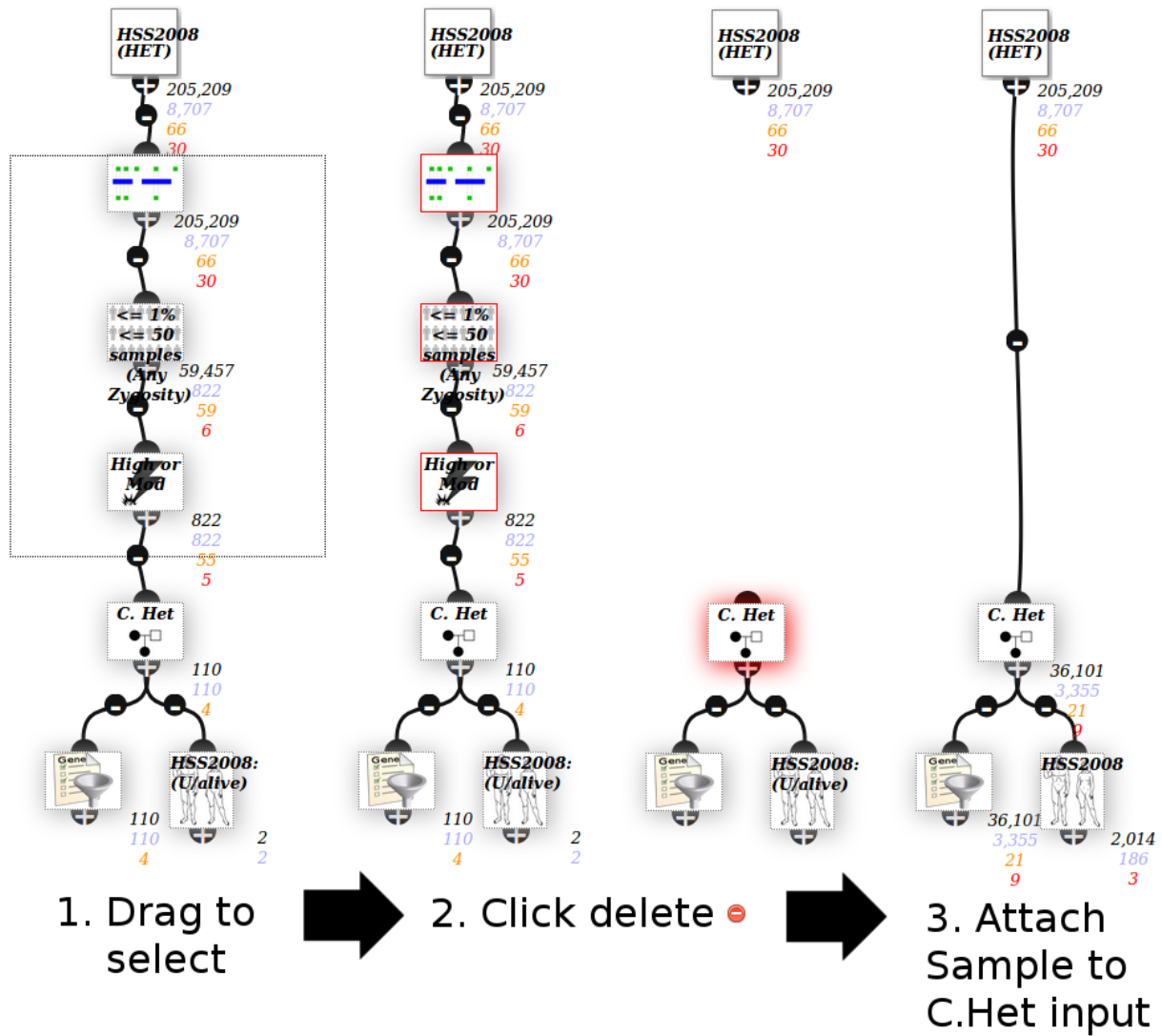
26.3.2 Compound Het filter

Compound heterozygous means 2 variants in the same gene from different parents.

The C. Het node in the bottom right of the screenshot above is a filter node - ie it has another node connected to the top, while the other inheritance models do not.

This is because you probably don't want every gene with ≥ 2 variants, but rather only ≥ 2 damaging/rare ones. Adjust the filters above the C.Het node to adjust this.

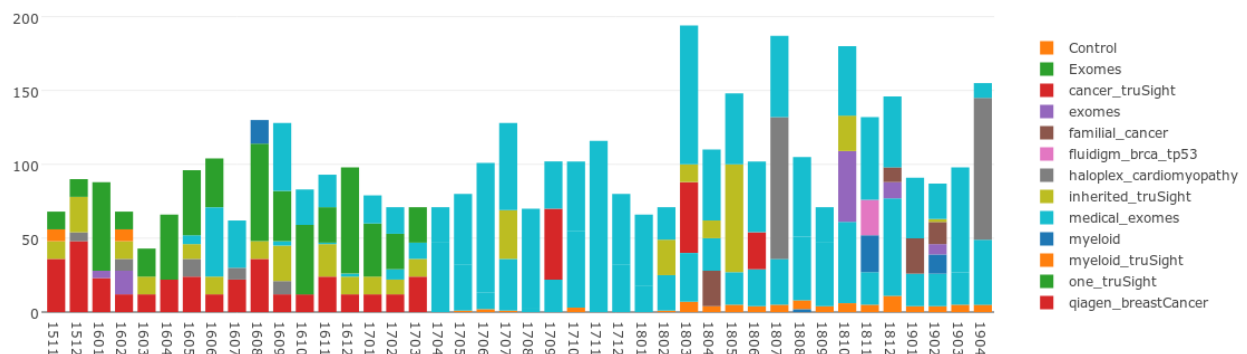
Modify the analysis as per instructions below to filter to all of them.



CHAPTER TWENTYSEVEN

SEQUENCING RUNS

When VariantGrid has access to a network drive (eg a diagnostic lab intranet) it can scan disks for sequencing runs to collect QC metrics, gene coverage and automatically load VCFs.



Sequencing

Samples over time

EnrichmentKit: roche_1k_disease (version 6)

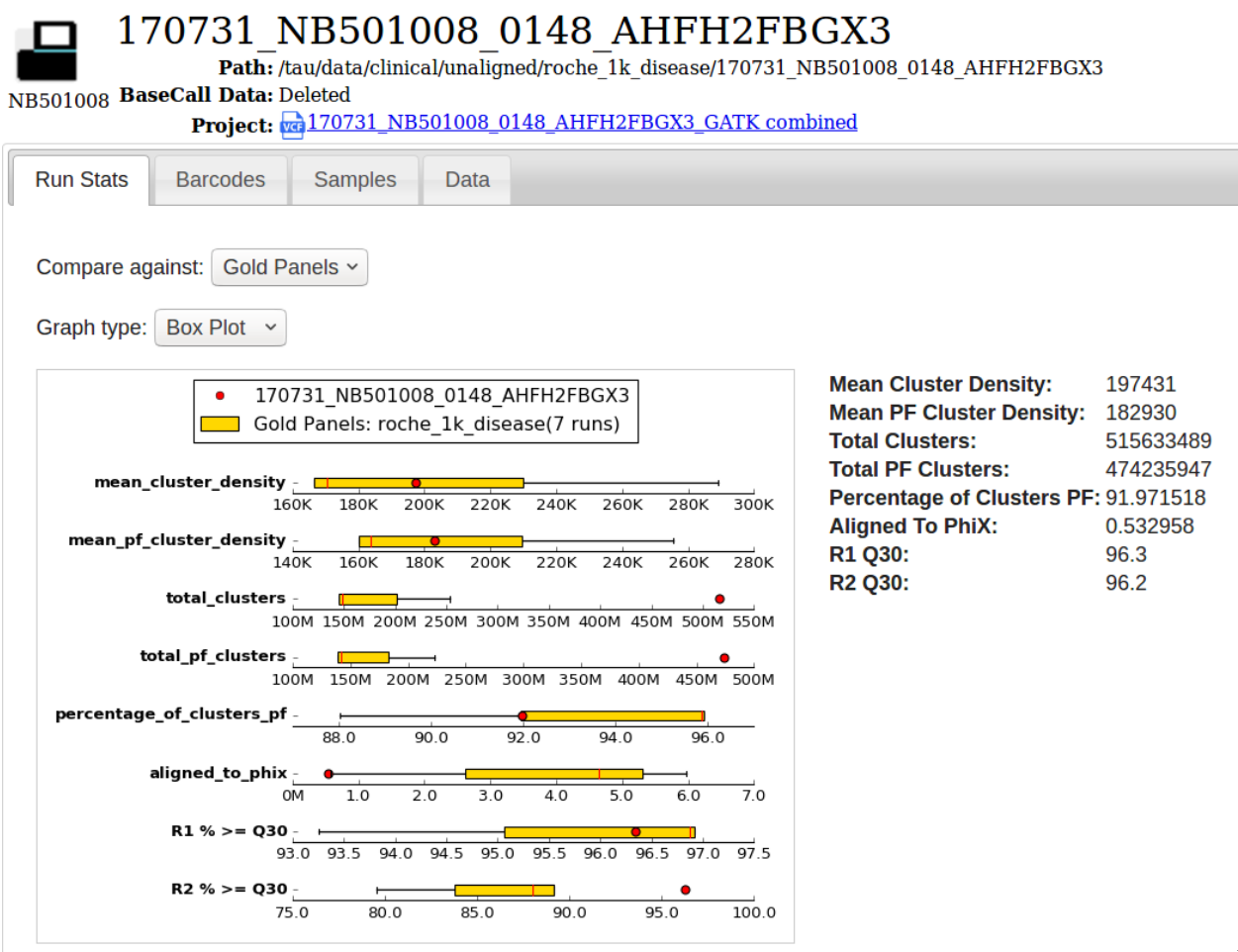
Filtering to Enrichment Kit. ☒ Show Incomplete Data: ☐ Show Hidden Data: ☐

SequencingRuns	name	Sample	Model	Sequencer	QC Lo	Experiment	EnrichmentKit	Kit ver	Gold	Hidden	Bad	VCF	path
190412	NB501008_0315_AH2HG5BGBXB	F11	NextSeq 500	NB501008	Complete	R1KD_19_009	roche_1k_disease	6					/tau/data/clinical/unaligned/roche_1k_disease/190412_NB501008_0315_AH2HG5BGBXB
190326	NB501009_0287_AHLFTKAFXY	24	NextSeq 500	NB501009	Complete	R1KD_19_008	roche_1k_disease	6					/tau/data/clinical/unaligned/roche_1k_disease/190326_NB501009_0287_AHLFTKAFXY
190324	NB501008_0308_AHFMM5AFXY	25	NextSeq 500	NB501008	Complete	R1KD_019_006	roche_1k_disease	6					/tau/data/clinical/unaligned/roche_1k_disease/190324_NB501008_0308_AHFMM5AFXY
190313	NB501009_0281_AHFVCKAFXY	25	NextSeq 500	NB501009	Complete	R1KD_019_004	roche_1k_disease	6					/tau/data/clinical/unaligned/roche_1k_disease/190313_NB501009_0281_AHFVCKAFXY
190215	NB501009_0274_AHHKYVAFXY	25	NextSeq 500	NB501009	Complete	R1KD_19_003_REPEAT	roche_1k_disease	6					/tau/data/clinical/unaligned/roche_1k_disease/190215_NB501009_0274_AHHKYVAFXY
190121	NB501008_0294_AHCNFGAFXY	21	NextSeq 500	NB501008	Complete	R1KD019_002	roche_1k_disease	6					/tau/data/clinical/unaligned/roche_1k_disease/190121_NB501008_0294_AHCNFGAFXY
190107	NB501009_0263_AHGLFYAFXY	22	NextSeq 500	NB501009	Complete	R1KD_19_001	roche_1k_disease	6					/tau/data/clinical/unaligned/roche_1k_disease/190107_NB501009_0263_AHGLFYAFXY
181217	NB501008_0283_AHHHWGAFXY	25	NextSeq 500	NB501008	Complete	R1KD18_028	roche_1k_disease	6					/tau/data/clinical/unaligned/roche_1k_disease/181217_NB501008_0283_AHHHWGAFXY
181203	NB501008_0276_AHGJUNAFXY	25	NextSeq 500	NB501008	Complete	R1KD_18_027_RECAPTU	roche_1k_disease	6					/tau/data/clinical/unaligned/roche_1k_disease/181203_NB501008_0276_AHGJUNAFXY
181119	NB501009_0244_AHFVC5AFXY	25	NextSeq 500	NB501009	Complete	R1KD_18_026	roche_1k_disease	6					/tau/data/clinical/unaligned/roche_1k_disease/181119_NB501009_0244_AHFVC5AFXY
181112	NB501008_0266_AHGJCNBGBX9	F19	NextSeq 500	NB501008	Complete	R1KD_18_025_FFPE	roche_1k_disease	6					/tau/data/clinical/unaligned/roche_1k_disease/181112_NB501008_0266_AHGJCNBGBX9
181105	NB501009_0239_AHFT2YAFXY	25	NextSeq 500	NB501009	Complete	R1KD_18_024	roche_1k_disease	6					/tau/data/clinical/unaligned/roche_1k_disease/181105_NB501009_0239_AHFT2YAFXY
181022	NB501009_0233_AHC7CLAFXY	25	NextSeq 500	NB501009	Complete	R1KD_18_023	roche_1k_disease	6					/tau/data/clinical/unaligned/roche_1k_disease/181022_NB501009_0233_AHC7CLAFXY
181008	NB501009_0222_AHC7F3AFXY	24	NextSeq 500	NB501009	Complete	R1KD_18_022	roche_1k_disease	6					/tau/data/clinical/unaligned/roche_1k_disease/181008_NB501009_0222_AHC7F3AFXY
180926	AHC7GVAFXY_AHC7KGAFXY_Me	25	NextSeq 500	NB501008	Error	R1KD18_021	roche_1k_disease	6					/tau/data/clinical/unaligned/roche_1k_disease/180926_AHC7GVAFXY_AHC7KGAFXY_Me
180830	NB551037_0234_AHCT3CAFXY	24	NextSeq 500	NB551037	Complete	R1KD_18_020	roche_1k_disease	6					/tau/data/clinical/unaligned/roche_1k_disease/180830_NB551037_0234_AHCT3CAFXY
180813	NB501008_0233_AHGG75BGBX7	F18	NextSeq 500	NB501008	Complete	R1KD_18_019_FFPE	roche_1k_disease	6					/tau/data/clinical/unaligned/roche_1k_disease/180813_NB501008_0233_AHGG75BGBX7
180806	NB501009_0204_AH7WHKAFXY	25	NextSeq 500	NB501009	Complete	R1KD_18_018	roche_1k_disease	6					/tau/data/clinical/unaligned/roche_1k_disease/180806_NB501009_0204_AH7WHKAFXY
180723	NB501009_0198_AH7GL3AFXY	25	NextSeq 500	NB501009	Complete	R1KD_18_017	roche_1k_disease	6					/tau/data/clinical/unaligned/roche_1k_disease/180723_NB501009_0198_AH7GL3AFXY
180709	NB501009_0195_AH7GH2AFXY	22	NextSeq 500	NB501009	Complete	R1KD18_016	roche_1k_disease	6					/tau/data/clinical/unaligned/roche_1k_disease/180709_NB501009_0195_AH7GH2AFXY
180702	NB501008_0221_AHK5G3BGBX7	11	NextSeq 500	NB501008	Complete	R1KD_18_015_FFPE	roche_1k_disease	6					/tau/data/clinical/unaligned/roche_1k_disease/180702_NB501008_0221_AHK5G3BGBX7
180625	NB501009_0189_AH7FVVFXY	25	NextSeq 500	NB501009	Complete	R1KD_18_014	roche_1k_disease	6					/tau/data/clinical/unaligned/roche_1k_disease/180625_NB501009_0189_AH7FVVFXY
180608	NB501009_0186_AH2JYWAFXY	25	NextSeq 500	NB501009	Complete	R1KD_18_013	roche_1k_disease	6					/tau/data/clinical/unaligned/roche_1k_disease/180608_NB501009_0186_AH2JYWAFXY
180531	NB501009_0184_AH27M2AFXY	25	NextSeq 500	NB501009	Complete	R1KD_18_012	roche_1k_disease	6					/tau/data/clinical/unaligned/roche_1k_disease/180531_NB501009_0184_AH27M2AFXY
180514	NB501009_0178_AH33KLAFXY	25	NextSeq 500	NB501009	Complete	R1KD_18_010	roche_1k_disease	6					/tau/data/clinical/unaligned/roche_1k_disease/180514_NB501009_0178_AH33KLAFXY
180430	NB501008_0209_AH33KWAFXY	25	NextSeq 500	NB501008	Complete	R1KD_18_009	roche_1k_disease	6					/tau/data/clinical/unaligned/roche_1k_disease/180430_NB501008_0209_AH33KWAFXY
180416	NB501009_0171_AH332YAFXY	25	NextSeq 500	NB501009	Complete	R1KD_18_008	roche_1k_disease	6					/tau/data/clinical/unaligned/roche_1k_disease/180416_NB501009_0171_AH332YAFXY
180329	NB501009_0169_AH2JWVAFXY	25	NextSeq 500	NB501009	Complete	R1KD_18_007	roche_1k_disease	6					/tau/data/clinical/unaligned/roche_1k_disease/180329_NB501009_0169_AH2JWVAFXY
180319	NB501009_0167_AHYGH3AFXX	20	NextSeq 500	NB501009	Complete	R1KD_18_006	roche_1k_disease	6					/tau/data/clinical/unaligned/roche_1k_disease/180319_NB501009_0167_AHYGH3AFXX
180309	NB501009_0165_AHMY7NBGX5	F13	NextSeq 500	NB501009	Complete	R1KD_005_FFPE	roche_1k_disease	6					/tau/data/clinical/unaligned/roche_1k_disease/180309_NB501009_0165_AHMY7NBGX5
180309	NB501009_0165_AHMY7NBGX5	F13	NextSeq 500	NB501009	Complete	R1KD_005_FFPE	roche_1k_disease	6					/tau/data/clinical/unaligned/roche_1k_disease/180309_NB501009_0165_AHMY7NBGX5

Page 1 of 1

Automatically

loaded sequencing runs + VCFs

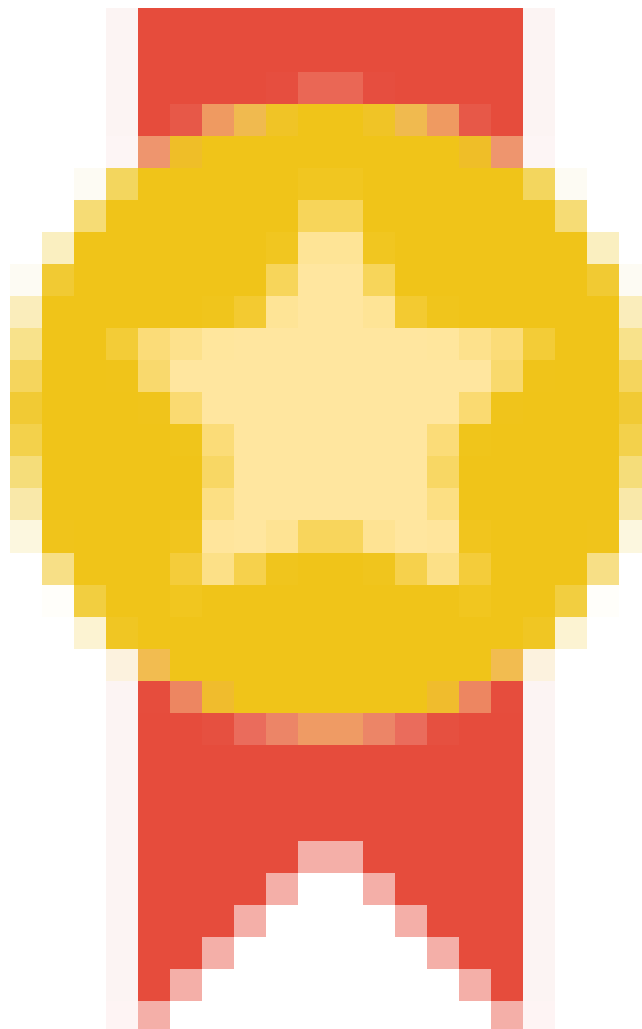


Sequencing Run

We collect Sequencing QC metrics and display them with interactive graphs. Collecting data over time allows us to see how this run compares to other runs over time (or vs *gold standard runs*).

27.1 Gold Standard Runs

The administrator can mark a [sequencing run](#) as “Gold Standard” - which means it has passed validation / is of sufficient quality to be used as a benchmark for other runs.



Gold standard runs have an icon (on the sequencing run grid.

Gold runs for an enrichment kit are used:

- In boxplots on QC metrics pages for a [sequencing run](#) or other sample QC graphs.
- To calculate average *gene coverage* on the *GeneGrid* page.

27.2 Finding sequencing data

Sequencing Runs are found by searching for the file 'RTAComplete.txt' on the server disks. You can ignore flow cells by putting a file ".variantgrid_skip_flowcell" in the directory.

USER SETTINGS

Lab Password

28.1 Customise columns

CUSTOMISE COLUMNS

You can customise grid columns on the **Customise Columns** ([user]->[customise columns]) page.

IGV INTEGRATION



Click the IGV link to automatically jump to your variants + BAM files in IGV.

ID	chr	position	ref	alt	dbsnp rs id	gene symbol
<input type="checkbox"/>	11	48166267	G	C	rs4752904	PTPRJ
<input type="checkbox"/>	11	48145166	G	A	rs2270993	PTPRJ
<input type="checkbox"/>	11	48145247	T	C	rs2270992	PTPRJ
<input type="checkbox"/>	12	25368462	C	T	rs4362222	KRAS
<input type="checkbox"/>	12	25362777	A	G	rs1137282	KRAS
<input checked="" type="checkbox"/>	14	75513883	T	C	rs175081	MLH3
<input type="checkbox"/>	14	75482813	T	C	rs123713	MLH3
<input type="checkbox"/>	15	40500986	C	T	rs11630664	BUB1B
<input type="checkbox"/>	15	40477831	G	A	rs1801376	BUB1B
<input type="checkbox"/>	17	17124815	C	T	rs3744124	FLCN
<input type="checkbox"/>	17	63554591	G	A	rs2240308	AXIN2
<input type="checkbox"/>	17	63533768	G	A	rs1133683	AXIN2
<input type="checkbox"/>	17	7579472	G	C	rs1042522	TP53
<input type="checkbox"/>	17	63533789	T	C	rs9915936	AXIN2

Open14:75513883 in IGV

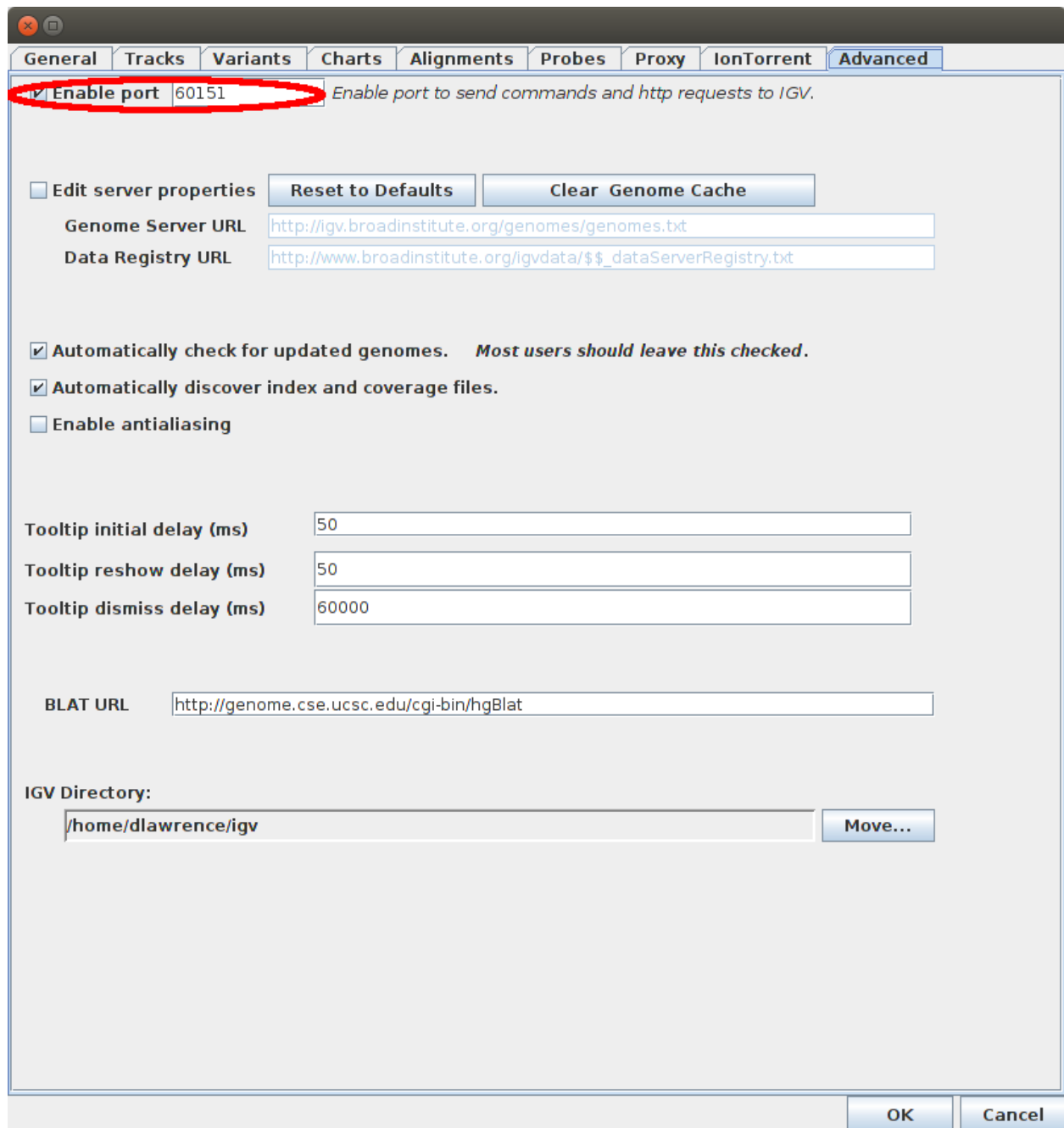
Sequence →
RefSeq Genes

GAATGGAAACTTCTCTGAGTTAAGGATGTGGCTTGCTGGTT
F P F K E S N L I H S A P Q
MLH3

30.1 IGV Configuration

IGV needs to be running, and have the Enable Port option ticked.


To check this open preferences in the IGV menu: [View] -> [Preferences] -> [Advanced] Tab.



30.2 VariantGrid Configuration

If the value of the IGV port is different from **60151** (default), you need to change the IGV Port option in your User Settings page.



Clicking the IGV link ( IGV link) will jump to the locus, and show BAM files associated with input samples (Sample or Cohort ancestors). These are the same samples that have their zygosity/allele depth shown on the grid.

Each sample has a bam file path entry. If your samples were automatically loaded from a server, this is probably already set. Otherwise you can change it on the Sample or VCF (VCF) page.

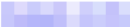


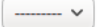
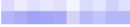
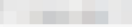

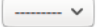



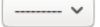

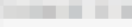

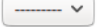

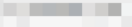
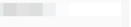
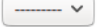


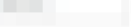
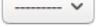
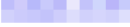
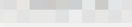

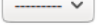
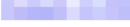

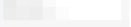
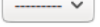
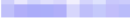


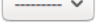
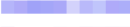
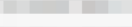
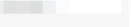
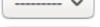

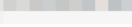
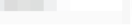
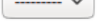
You can set all the samples in a VCF file at once in the vcf page, click Bulk Set Fields to set all samples according to a pattern based on the sample name.

Samples

Bulk Set Fields

BAM path

Public Data Toggle ☐

Sample	Variants (passed)	Name	Patient	Physical Sample	BAM path	Public Data
	12607 (12264)				<input type="text" value="/tau/data/clinical/align"/>	<input type="checkbox"/>
	12512 (12163)				<input type="text" value="/tau/data/clinical/align"/>	<input type="checkbox"/>
	12590 (12249)				<input type="text" value="/tau/data/clinical/align"/>	<input type="checkbox"/>
	12762 (12420)				<input type="text" value="/tau/data/clinical/align"/>	<input type="checkbox"/>
	12768 (12417)				<input type="text" value="/tau/data/clinical/align"/>	<input type="checkbox"/>
	12905 (12549)				<input type="text" value="/tau/data/clinical/align"/>	<input type="checkbox"/>
	12702 (12357)				<input type="text" value="/tau/data/clinical/align"/>	<input type="checkbox"/>
	12770 (12423)				<input type="text" value="/tau/data/clinical/align"/>	<input type="checkbox"/>
	12579 (12229)				<input type="text" value="/tau/data/clinical/align"/>	<input type="checkbox"/>
	12643 (12297)				<input type="text" value="/tau/data/clinical/align"/>	<input type="checkbox"/>
	12585 (12247)				<input type="text" value="/tau/data/clinical/align"/>	<input type="checkbox"/>

30.3 Network drives and File Servers

Many labs access data via servers, or network shares. These can be different on different computers.

It is recommended that you set bam file path to be the location on the server, so that it is consistent between users.

Different data access methods on different computers can then be managed by having users change their configuration on the IGV Integration page.

VARIANT CLASSIFICATIONS

31.1 Creating Classifications

- From an analysis (see *analysis classification workflow*)
- From the *variant details page*
- Via API (See *Shariant API docs*)

31.2 Autopopulation

When you create a classification from inside the system, a number of fields are auto-populated from annotation and sample information.

Variants created from the external API are not auto-populated with values from annotation.

31.3 Editing

See the *Classification Form*.

31.4 Configuring Fields

An administrator can add/remove EvidenceKeys which are used to create fields.

They can also hide visible fields on a per-lab basis.

VARIANT CLASSIFICATION FORM

The Classification Web Form can be used to create and edit classifications directly within VariantGrid.

32.1 View

Variant

ClinGen Canonical Allele Identifier

CA396457842

Ensembl Gene ID

ENSG00000039068

Gene symbol

CDH1

*Genome Build

GRCh37

Gene OMIM ID

192090

RefSeq Transcript ID

NM_004360.4

Ensembl Transcript ID

ENST00000261769

HGNC ID

HGNC:1748

UniProt ID

P12830

Variant coordinate

16:68842599 A>G

g.HGVS

NC_000016.9:g.68842599A>G

c.HGVS

NM_004360.4(CDH1):c.535A>G

p.HGVS

NP_004351.1:p.Lys179Glu

Molecular consequence

Missense variant

*Zygosity

Gene

*Condition under curation

Hereditary diffuse gastric cancer

Gene-disease validity

Definitive

Y-Path Zues Lab / vc768

NM_004360.4(CDH1):c.535A>G,
NP_004351.1:p.Lys179Glu
VUS (3)

Flags

Zygosity

blank

Zygosity in the tested individual.

Does the allele frequency agree with the zygosity? Be aware of mosaicism.

Status

Last Edited 05/Aug/2019 12:53

Last Shared Ver. 05/Aug/2019 11:33

Compare with

historical versions of this record

other classifications for this variant (Pathogenic x1, Unclassified x1)

Messages

Zygosity - Missing mandatory value

Links

ClinGen Allele Registry

ClinGen KB

Clinvar Variant

Genomizer

gnomAD

GTEX

Monarch Phenotypes

NCBI

OMIM (Gene)

PDB

UCSC

Uniprot ID

	BA	BS	BP	PP	PM	PS	PVS
P	/	//		1	1		
CP			///	/	//	/	/
F	/			/	/	/	
S	/			1			
D					/	/	
A			/		/		
DB			/	/			
O		/	/				

2xPP, 1xPM

Calculation: Uncertain significance

(1)

To quickly see all fields that have values for a classification, enter “*” into the filter box at the top of the classification. To see all possible fields, enter “**” in the filter box. To find an individual field, start typing the label of the field into the filter e.g. “gnomad”.

32.2 Identify Errors

A record might not be shared as there are outstanding validation errors. In the Messages box on the form it will list any errors. If possible fix those errors in your curation system and then they should be fixed on the next sync.

32.3 Change History / Diff

Each version of a record published in VariantGrid is recorded, by clicking on “Compare historical versions of this record”.

If there are other classifications for the same variant, there will be a link to compare them there too.

32.4 ACMG Guidelines

The classification form has fields for the ACMG Guidelines, e.g. PM4, BA1 - the meaning of each is given in the help. See [Guidelines](#)

VariantGrid displays a grid of ACMG fields with each row being a category of data, and each column representing the strength of evidence for benign or pathogenic.

- The number of met criteria for a given box will be shown as a number.
- Explicitly unmet criteria will show as “/”s.
- Criteria not yet marked as met or unmet will show as “?”s.

The various values will be plugged into the ACMG formulae and a recommended overall clinical significance will be displayed. This calculated value has no affect on any of the data, the user is still able to set the overall clinical significance to whatever (hopefully justifiable) value they like.

32.5 Actions

Actions		Share	
Export as	CSV	Clinvar	Report

Literature Citations

Sanguinarine, inhibitor of Na-K dependent ATP'ase.

Straub, K D, Carver, P

Biochem Biophys Res Commun. 1975 Feb 17;62(4):913-22. doi: 10.1016/0006-291x(75)90410-6. PubMed: 123445

At the bottom of the form there will be a list of action buttons.

The Tick icon re-submits the classification at its current change level. For any manual changes to be seen, this button will need to be ticked.

Next to it is a Share button that allows you to increase the scope of who can see the classification. Important, increasing the Share level is not un-doable. The share levels are

- Just your lab
- Anyone within your organisation (if your organisation has multiple labs)
- All Shariant Users
- 3rd Party Databases (this will allow us to upload the record to Clinvar at a later date)

32.6 Delete / Withdraw

If the classification has only been shared at the lab or organisation level, you are able to perform a hard delete on the record. If it has been shared, instead you have the option to “withdraw”. This will remove the record from most listings and search results, but will not remove it from any Discordance Reports that it had been involved in (it will no longer be a part of discordance calculations).

When a record has been withdrawn it can be unwithdrawn by clicking the same button (it should look like a rubbish bin with a raised lid now).

32.7 Export

You can also export the single record as CSV, a preview of the Clinvar format or as a report. (The report does require that your lab has a report template pre-configured.)

32.8 Literature Citations

Any PMID references in the form of PMID:123456 from anywhere within the classification will be summed together and listed at the bottom of the classification.

CLASSIFICATION FLAGS

Each classification flag indicates that there is an action that needs to be performed against the classification.


Many of the flags will be automatically raised by Shariant, though some of them you will be able to open yourself.

To look at the details of a specific open flag, simply click on it to be taken to the flag dialog.

33.1 Flag Dialog

Test X Lab One / vc850

In Progress Flags




Unshared Classification
10 days old


This classification is not yet shared outside of your lab or institution.

- 1 From the classification form, ensure there are no validation errors stopping this record from being published.
- 2 Review the content of the classification to make sure it's ready to be shared.
- 3 At the bottom of the form, click the Share to submit at a higher share level.


Resolved Flags



Internal Review
New



Suggestion
New



Suggestion
New

Raise New Flags



Internal Review

You can raise this flag to let people know the classification is currently in review, or raise it as "Completed" to record the fact that a review has recently taken place. Please record any internal reviews while a classification is marked as discordant.



Suggestion

If you have found some extra information that you think should be incorporated into this classification, you can raise a suggestion for the classification owner to accept or reject.

From the flag dialog you can view summaries about what flags are currently open, see a list of flags that have been resolved as well as raise new ones. Note that only important flags still show up when closed, e.g. suggestions and internal reviews and a few others.

In the provided screenshot we can see we have an open flag asking us to share the classification, a completed internal review, an accepted suggestion and a rejected suggestion, as well as the buttons to create new internal reviews and suggestions.

You can visit the details of an open flag, or a closed one by clicking on the icon.

From the details page of an open flag, depending on the type of flag, you can add a comment and potentially change the status of a flag.

You can raise a new flag by clicking on one of the icons near the bottom with a plus button.

(The kinds of actions you can take on flags will depend on if you're looking at a classification from your lab or another lab.)

See below for flags and how to solve them:

33.2 Flag Types

33.2.1 Discordance

This classification is in discordance with one or more classifications.

1. Ensure that you have completed an internal review of your lab's classification recently (within the last 12 months is recommended). If not, raise the internal review flag and complete an internal review of your lab's classification.
2. Review any outstanding suggestions against your lab's classification.
3. View the other classifications in the discordance report and view the evidence differing between multiple records via the diff page. If appropriate, raise suggestions against other lab classifications.
4. This Discordance flag will automatically be closed when concordance is reached.

This is discussed in the [Classification Discordance](#) page.

33.2.2 Internal Review

This classification is marked as currently being internally reviewed.

1. Once the internal review is complete, ensure you update the classification in your curation system.
2. Mark the internal review as Completed.

This is discussed in the [Classification Discordance](#) page.

33.2.3 Matching Variant

This variant has not been seen in this system previously. It should be linked to a variant given time.

33.2.4 Matching Variant Failed

We were unable to normalise the variant provided based on the c.hgvs and genome build values.

1. Please contact Shariant support for help in resolving this.

33.2.5 Outstand Edits

Edits have been made to this classification that are not included in a published version.

1. From the classification form, ensure there are no validation errors stopping this record from being published.
2. At the bottom of the form, click the tick to submit the outstanding changes.

33.2.6 Significance Changed

This classification has changed its clinical significance compared to a previously published version.

1. Set the status of this flag to reflect the primary reason behind the change in classification.
2. Please also add a comment providing some context.

This is discussed in more detail on the [Classification Discordance](#) page.

33.2.7 Suggestion

Someone has raised suggestion(s) against this classification.

1. Review the contents of each suggestion.
2. If appropriate, make changes in your curation system and mark the suggestion as Complete.
3. If you decline the suggestion, mark it as Rejected.

33.2.8 Unshared Classification

This classification is not yet shared outside of your lab or institution.

1. From the classification form, ensure there are no validation errors stopping this record from being published.
2. Review the content of the classification to make sure it's ready to be shared.
3. At the bottom of the form, click the Share to submit at a higher share level.

33.2.9 Withdrawn

This classification has been marked as withdrawn. It will be hidden from almost all searches and exports.

1. If the classification is not of high enough quality or in error, you may leave it as “withdrawn” indefinitely.
2. If you wish to un-withdraw the classification, click the open bin icon in actions from the variant classification form. (Note you can't open a Withdrawn flag, but you can Withdraw/Unwithdraw from the classification form)

VARIANT CLASSIFICATION REPORT

34.1 Running the report

To generate the report from a classification, open the classification and scroll to the bottom. You will see a button called “Report”. Click on it and you will then be able to copy & paste the report contents into a document.

34.2 Configuring the report

The report can only be configured by admin users. Each “organisation” within variantgrid uses its own report. To edit it go to the admin view, Organisations, (your organisation), and then edit the Classification report template.

The template is run using Django template and produces HTML

34.3 Values available for the report

34.3.1 Evidence Keys

All the fields in the classification are exposed here, see the Evidence Keys admin for a list of possible values, e.g. zygoty, mechanism_of_disease, mode_of_inheritance. In addition you can also suffix `_raw` or `_note` e.g.

```
The raw value for Mode of Inheritance is {{ mode_of_inheritance_raw }} and the note_
↳for it is {{ mode_of_inheritance_note }}
{% if mode_of_inheritance_raw == 'x_linked' %}
Special case for X Linked
{% endif %}
```

Typically you’ll only want to refer to the `_raw` value if you’re doing some logic for a specific drop down value. If you omit the `_raw` then you will get the human friendly label for the value which might subtly change in the future.

34.3.2 p.hgvs

You can reference the full `p_hgvs` or breakdown

```
full p.hgvs = {{ p_hgvs }}<br/>
p amino acid from = {{ p_hgvs_aa_from }}<br/>
p hgvs codon = {{ p_hgvs_codon }}<br/>
p hgvs amino acid to = {{ p_hgvs_aa_to }}
```

34.3.3 c.hgvs

You can reference the full c_hgvs or breakdown

```
full c.hgvs = {{ c_hgvs }}<br/>
c hgvs transcript = {{ c_hgvs_transcript }} or {{refseq_transcript_id}}<br/>
c hgvs gene symbole = {{ c_hgvs_gene_symbol }} or {{ gene_symbole }}<br/>
c hgvs short = c.{{ c_hgvs_short }} (this is the value in c_hgvs after "c.")
```

34.3.4 Evidence weights

A summary of the strength of ACMG critieria met can be accessed with

```
Evidence weights = {{ evidence_weights }}
```

34.3.5 Citations

PMIDs put anywhere in the classification can be accessed, and then specific attributes of those citations can be referenced. citations is an array that you must loop through, e.g.

```
{% for cit in citations %}
    <tr>
        <td>{{ cit.source }}</td>
        <td>{{ cit.citation_id }}</td>
        <td>{{ cit.citation_link }}</td>
        <td>{{ cit.journal }}</td>
        <td>{{ cit.journal_short }}</td>
        <td>{{ cit.title }}</td>
        <td>{{ cit.year }}</td>
        <td>{{ cit.authors }}</td>
        <td>{{ cit.authors_short }}</td>
        <td>{{ cit.abstract }}</td>
    </tr>
{% endfor %}
```

The example here is in a table but you can display it however you'd like, e.g.

```
{% for cit in citations %}
{{ cit.source }}:{{ cit.citation_id }}
{% endfor %}
```

Which would give you PMID:12334 PMID:4555 etc

VARIANT CLASSIFICATION REDCAP

Variantgrid supports the exporting of Variant Classification data into REDCap files. Note that this is currently the full extent of REDCap integration with Variantgrid, there is no support for importing REDCap records or exporting any other kinds of records in a REDCap format.

There are two parts to the REDCap export.

35.1 REDCap Definition

The data definition is available by opening the page help on the classification page.

sequencing | data | patients | tests | analysis | classifications | genes | variants | annotation
search...
Go
?
help
jandrews

?
Variant Classification Help

Click on the pie chart to filter to that classification.
Tick 'Mine' to only show ones you created.

Create a new Variant Classification by entering a HGVS sequence in the box below, on the variant details page or inside an analysis by tagging it as RequiresClassification.

Export the grid below by clicking the CSV or Redcap buttons on the bottom left of the grid.

You can import into REDCap using this [data definition](#).

HGVS / dbSNP / VCF coordinate
Classify Variant

Variant Classifications

- Benign
- Likely Benign
- VUS
- Likely Pathogenic
- Pathogenic
- Unclassified

Gene Classifications

☐ Mine
Gene:
Gene...

Simple Filter... or Advanced classification search

ID	Status	clinical significance	c.HGVS	Gene Symbol	Lab Name	Lab Record ID	User	Created
1	Pathogenic	Likely Pathogenic	NM_000130.4(F5):c.1601A>G	F5	Test X Lab One	vc21	admin_bot	2019-07-17 17:19

The definition is dynamically generated from the variant classification evidence key configuration. We do our best to ensure that changes to evidence keys are backwards compatible for REDCap definitions.

The definition is laid out in such a way that up to 10 records can be grouped together in one record e.g. `vc_zygosity_1`, `vc_zygosity_2`, `vc_zygosity_3` up to `vc_zygosity_10`. This is so that variants for the same patient can be consolidated.

Note that the REDCap definition is primarily used as a read only representation of the data, doing large edits of data in REDCap is not recommended.

35.2 REDCap Rows

Important: Variant Classifications will **ONLY** be exported if REDCap Record ID has a value. All rows that do not have a value for REDCap Record ID will be ignored in the export.

At the bottom of the classification table there will be a CSV and REDCap download button. Clicking the REDCap download will download records that are:

- Available in the current filter (if the results are split over multiple pages all will be downloaded). For example if you filter to show “Mine” the records in the download have to belong to you.
- Have a value for REDCap Record ID.

Records that have the same REDCap Record ID, regardless of any other factors, will be grouped together as described earlier, re `vc_zygosity_1`, `vc_zygosity_2` etc

35.3 Grid Technical Specifics

This means while single drop down fields work as you’d expect, multi-drop downs produce text that’s harder to report on.

The evidence key definitions for selects have an explicit index for each drop down option. If adding more options (regardless of insertion order) a new index should be assigned and existing options should retain their index. This is to help keep newer REDCap definitions compatible with older REDCap records.

VARIANT NORMALIZATION

INDICES AND TABLES

- `genindex`
- `modindex`
- `search`