

---

# **VariantGrid**

**CCB ACRF Cancer Genomics Facility**

**Oct 10, 2019**



## OVERVIEW:

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

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

26.2	Digital karyomapping . . . . .	71
26.3	Trio inheritance analysis . . . . .	72
<b>27</b>	<b>Sequencing Runs</b>	<b>75</b>
27.1	Gold Standard Runs . . . . .	76
27.2	Finding sequencing data . . . . .	77
<b>28</b>	<b>User Settings</b>	<b>79</b>
28.1	Customise columns . . . . .	79
<b>29</b>	<b>Customise columns</b>	<b>81</b>
<b>30</b>	<b>IGV integration</b>	<b>83</b>
30.1	IGV Configuration . . . . .	83
30.2	VariantGrid Configuration . . . . .	84
30.3	Network drives and File Servers . . . . .	85
<b>31</b>	<b>Variant Classifications</b>	<b>87</b>
31.1	Creating Classifications . . . . .	87
31.2	Autopopulation . . . . .	87
31.3	Editing . . . . .	87
31.4	Configuring Fields . . . . .	87
<b>32</b>	<b>Variant Classification Form</b>	<b>89</b>
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
<b>33</b>	<b>Classification Flags</b>	<b>93</b>
33.1	Flag Dialog . . . . .	94
33.2	Flag Types . . . . .	95
<b>34</b>	<b>Variant Classification Report</b>	<b>97</b>
34.1	Running the report . . . . .	97
34.2	Configuring the report . . . . .	97
34.3	Values available for the report . . . . .	97
<b>35</b>	<b>Variant Classification REDCap</b>	<b>99</b>
35.1	REDCap Definition . . . . .	99
35.2	REDCap Rows . . . . .	100
35.3	Technical Specifics . . . . .	100
<b>36</b>	<b>Variant Normalization</b>	<b>101</b>
<b>37</b>	<b>Indices and tables</b>	<b>103</b>



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

## 1.2 Private server

There is a VariantGrid private server inside [SA Pathology](#), 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 page](#).

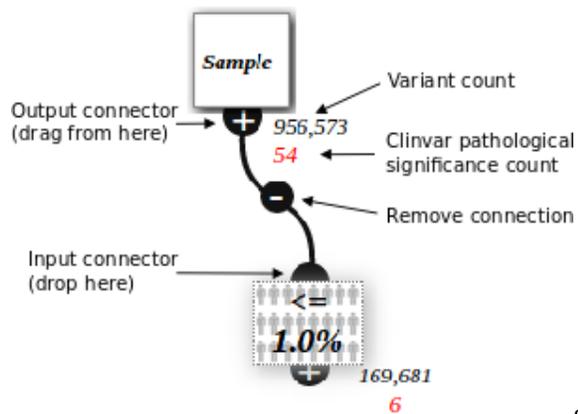


## ANALYSIS INTRO

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.

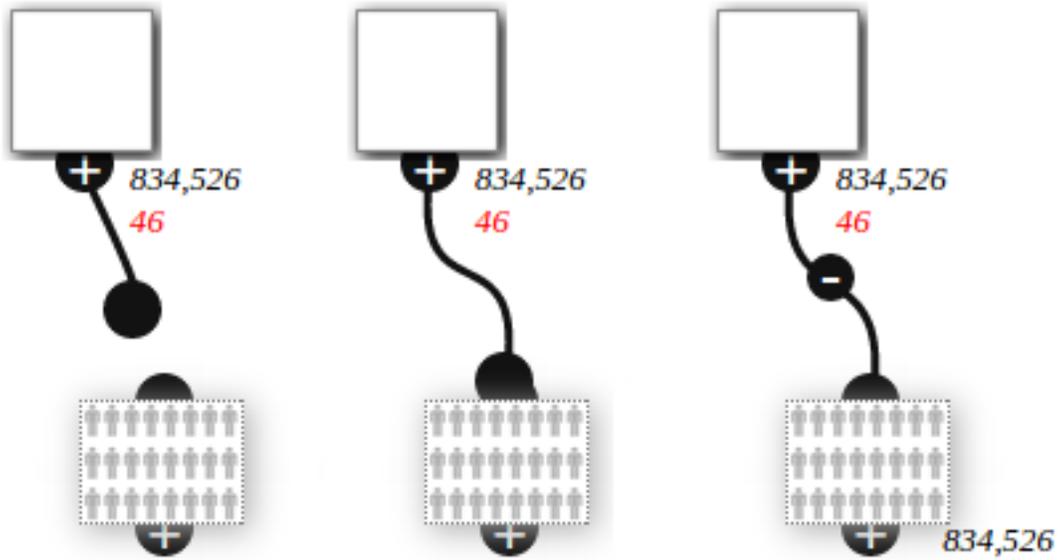
### 2.1 Analysis Nodes



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.**

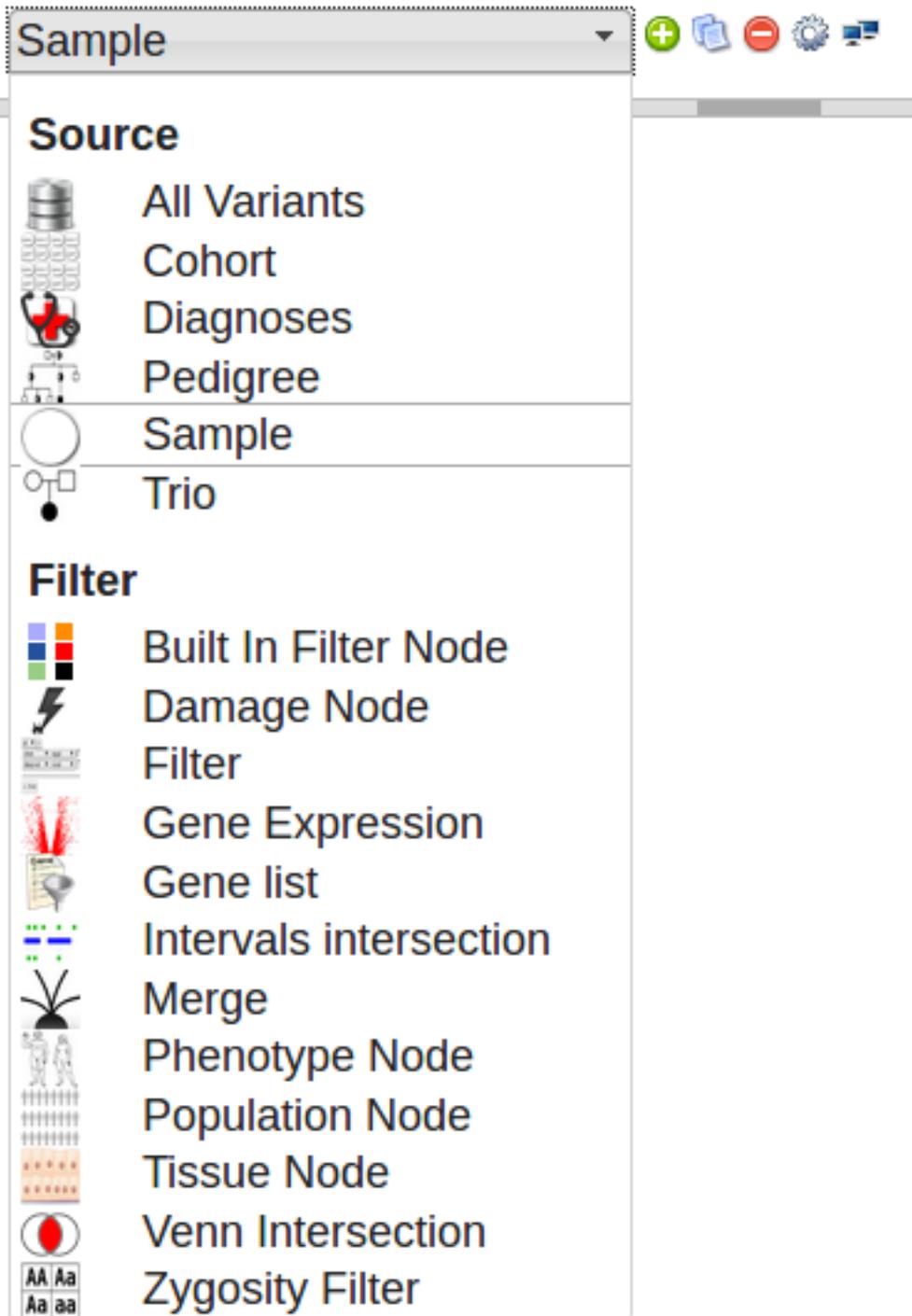
**Step 2: Drop it on the top of another node.**

**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 displays the VariantGrid analysis interface. On the left, a node graph visualizes filters: 'HSS24' (561,733 variants), 'Case (Lof 6)' (1,390,038 variants), '<= 1.0 %' (103,960 variants), 'Lymphedema (39 genes)' (126 variants), and 'snpeff\_impact = HIGH' (1 variant). The right panel shows the 'Grid' tab with a Venn diagram for comparing 'A: HSS24' and 'B: Case (1 of 6)'. The 'Comparison column' is set to 'variant'. Below is a table of variants:

ID	chr	position	ref	alt	dbsnp rs id	gene symbol	snpeff transcript id	snpeff am	snpeff coc	snpeff effect	snp	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	rs113724699	CDK11A	ENST00000378633	c.325+955		intron_variant		MODIFIER
153	1	758324	T	C	rs3131955	RP11-206L10.11	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

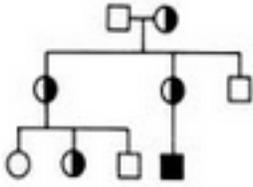


A collection of related samples, eg “control group” or “poor responders”

#### 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



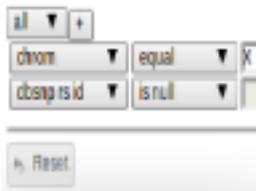
Built in filters used in node counts eg High or Moderate Impact / OMIM / ClinVar Pathological

### 3.2.2 Damage



Filter to damage predictions

### 3.2.3 Filter



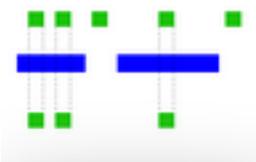
Filter based on column values

### 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



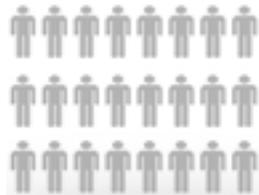
Merge variants from multiple sources

### 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

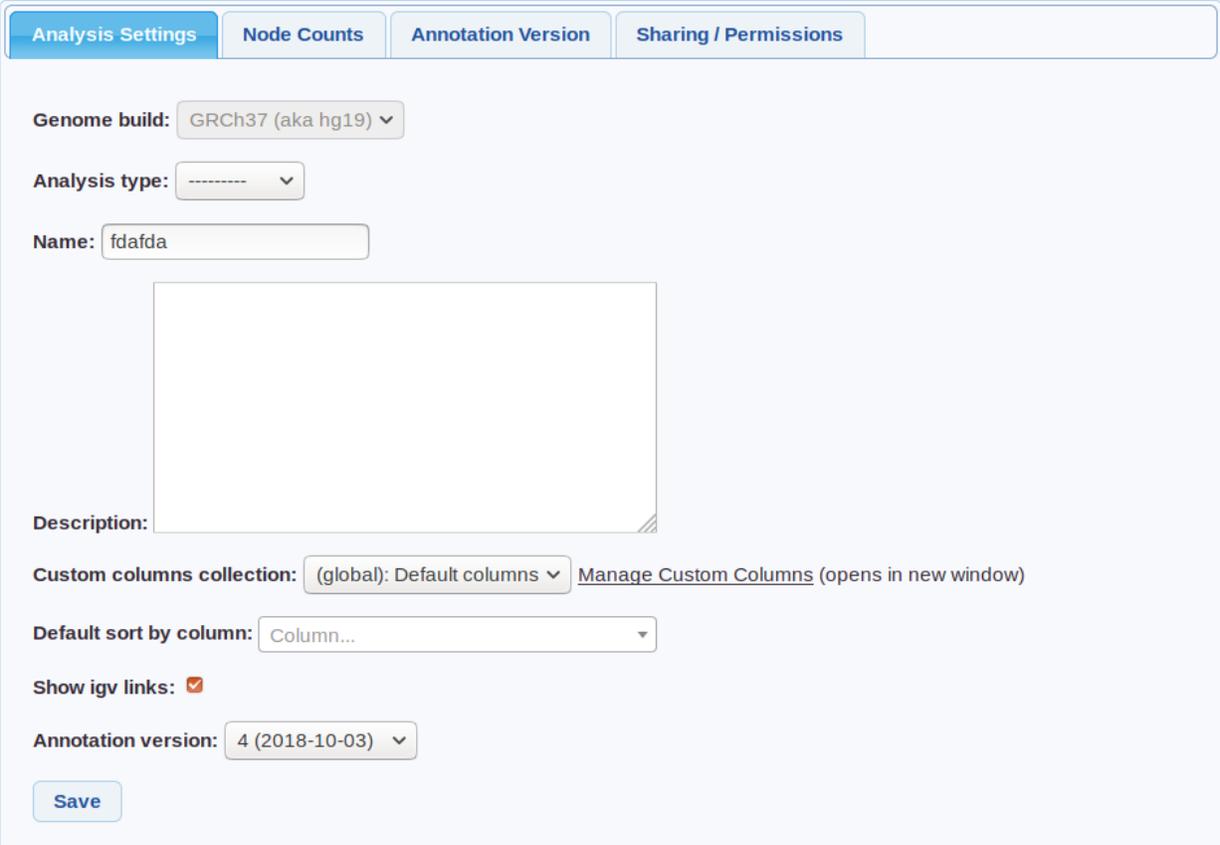


Compound HET and other Zygoty filters

## ANALYSIS - ADVANCED

### 4.1 Analysis settings

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



Force Reload Nodes

Close

settings screenshot

Analysis

- **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:

	Total		197,178
	OMIM Phenotype		44,517
	COSMIC		18,011
	High or Mod impact		13,728
	ClinVar		32
	Classified		1
	Classified Pathogenic		

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.

Analysis Settings
Node Counts
Annotation Version
Sharing / Permissions

Drag & Drop to change Node Counts, the numbers that appear to the side of a node.

My Node Counts

- Total
- ClinVar
- Classified Pathogenic

Available Node Counts

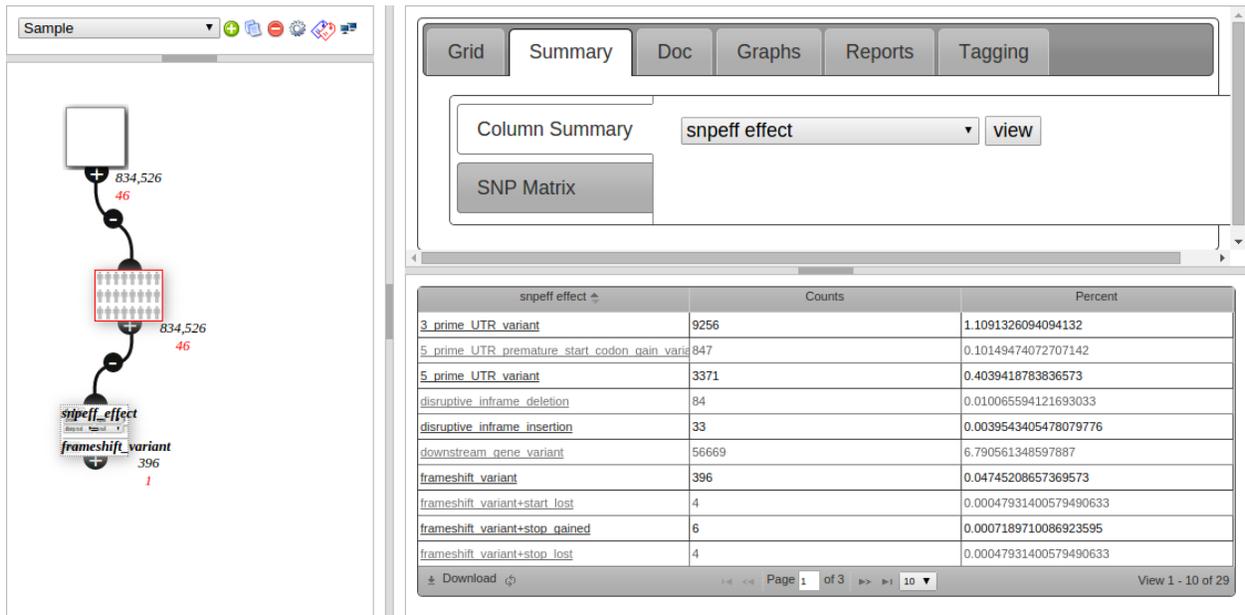
- OMIM Phenotype
- High or Mod impact
- Classified
- COSMIC

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 hierarchical tree shows nodes for 'snpeff\_effect' (396) and 'frameshift\_variant' (396). The right panel shows the 'Summary' tab with a dropdown menu set to 'snpeff effect' and a 'view' button. Below this is a table with the following data:

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

The table footer indicates '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.

Quantitative 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]

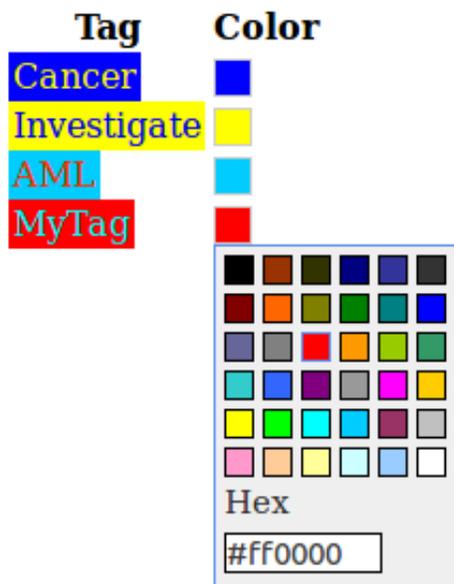
## Tags

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

Tag:

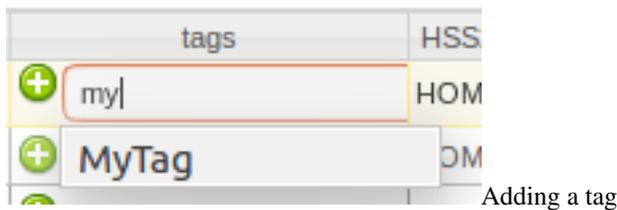
✓ Tag created successfully.

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.



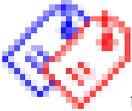
To remove a tag - clicking on the tag. The tag will grow in size, and a  delete symbol will appear. Click it to remove

the variant tag.

tags	HSS200	A	P
MyTag Investigate	HOM_ALT	1	3
Investigate			
	HOM_ALT	1	3
AML Cancer	HOM_ALT	1	3

Removing a tag

### 5.3 Using tags

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

Sample:

**Tags**

**HSS2008**  
**(HET,HOM\_ALT)**

 937,956  
13,959  
73  
20

Grid

**All tags for analysis**

Tag:

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	ZSK	Cancer
<input type="checkbox"/>	5	68309961	T	C	rs202010758	ZSK	Investigate
<input type="checkbox"/>	5	68323790	A	G	rs10214127	ZSK	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 15

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

**HSS2008 (HET, HOM\_ALT)**  
937,956  
13,959  
73  
20

1 Tagged Investigate

2 Tagged Cancer

Grid Summary Doc Graphs SQL

Analysis wide:

Tag: Cancer

save

ID	ch	position	ref	alt	dbsnp rs id	gene symbol	tags
<input type="checkbox"/>	5	68304573	T	C	rs4624745	7SK	<input type="checkbox"/> Cancer
<input type="checkbox"/>	5	68309961	T	C	rs202010758	7SK	<input type="checkbox"/> Cancer

CSV VCF

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

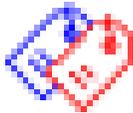
## ANALYSIS CLASSIFICATION

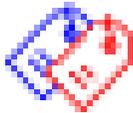
Recommended workflow to create a *classification* from a variant in an analysis:

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 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.

Trio ID: 78  
 Name: [redacted]  
 Cohort: [redacted]  
 Mother: [redacted]  
 Father: [redacted]  
 Proband: [redacted]

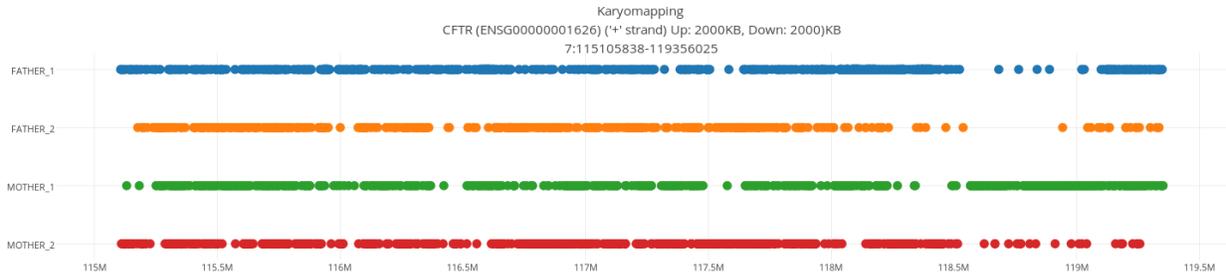
Whole sample counts (hg19) Proband phase: 50.74% mum / 49.26% dad. Mum: 54.96%. Dad: 51.69%. [more](#)

**Karyomap Gene Region**

Gene:  Upstream:  (KB) Downstream:  (KB)

**Previous Karyomap Gene Regions**

- [View](#) CFTR (ENSG0000001626) (Up: 2000, Down: 2000)
- [View](#) BRCA1 (ENSG0000012048) (Up: 2000, Down: 2000)
- [View](#) SMN1 (ENSG00000172062) (Up: 2000, Down: 2000)



[Download as CSV](#)

Name	Count
FATHER_1	1186
FATHER_2	652
MOTHER_1	1250
MOTHER_2	995

### 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
 <a href="#">My lab / vc0042</a>	NM_000130.4(F5):c.1601G>A	<a href="#">Benign (1)</a>		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

<input checked="" type="checkbox"/>	 <a href="#">AS-145_WES_HiSeq_Variants.vcf</a>	5.47 MB	<input type="button" value="Delete"/>	<input type="checkbox"/>
	<a href="#">VCF</a>			
<input checked="" type="checkbox"/>	 <a href="#">test.vcf</a>	1.43 KB	<input type="button" value="Delete"/>	<input type="checkbox"/>
	<a href="#">VCF</a>			

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:

The screenshot shows a web interface for managing data. At the top, there are five tabs: 'samples', 'VCF', 'bed file', 'gene lists', and 'Pedigree .ped files'. Below the tabs is a search box containing the text 'HSS232'. An autocomplete dropdown menu is open, showing a list of file names with their full paths, such as 'HSS2326 (all\_HM\_samples.2017Jan.gatk.vcf.gz)'. The file 'HSS2320 (all\_HM\_samples.2017Jan.gatk.vcf.gz)' is highlighted in blue. Below the search box is a table with columns for file name, status, and import status. The table contains several rows of data, including 'HSS2336', 'HSS2335', 'HSS2334', and 'HSS2333', all with a status of 'success' and an import status of 'al'.

import status	
al	HSS2326 (all_HM_samples.2017Jan.gatk.vcf.gz)
al	HSS2327 (all_HM_samples.2017Jan.gatk.vcf.gz)
al	HSS2328 (all_HM_samples.2017Jan.gatk.vcf.gz)
al	HSS2329 (all_HM_samples.2017Jan.gatk.vcf.gz)
al	HSS2320 (all_HM_samples.2017Jan.gatk.vcf.gz)
al	HSS2321 (all_HM_samples.2017Jan.gatk.vcf.gz)
al	HSS2322 (all_HM_samples.2017Jan.gatk.vcf.gz)
al	HSS2323 (all_HM_samples.2017Jan.gatk.vcf.gz)
al	HSS2324 (all_HM_samples.2017Jan.gatk.vcf.gz)
al	HSS2325 (all_HM_samples.2017Jan.gatk.vcf.gz)
al	HSS2336 success
al	HSS2335 success
al	HSS2334 success
al	HSS2333 success

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:

The screenshot shows the VariantGrid interface for a VCF file named '151120\_AHISEQTEST'. The interface includes a navigation bar with tabs for 'Details', 'Variants', 'Graphs', 'QC', and 'Sharing / Permissions'. The 'Sharing / Permissions' tab is active, displaying the 'Permissions' section. This section has a table with columns for 'Group', 'Read', and 'Write'. The 'public' group has its 'Read' checkbox checked, while 'my\_group' has both 'Read' and 'Write' checkboxes unchecked. A 'save' button is located below the table. Below the permissions section is the 'Genelist Security' section, which features a padlock icon and the text '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.

The screenshot shows the top right corner of the VariantGrid interface. It features a search bar with the placeholder text 'search...', a 'Go' button, and navigation links for 'help', an envelope icon, and '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 ▾ +

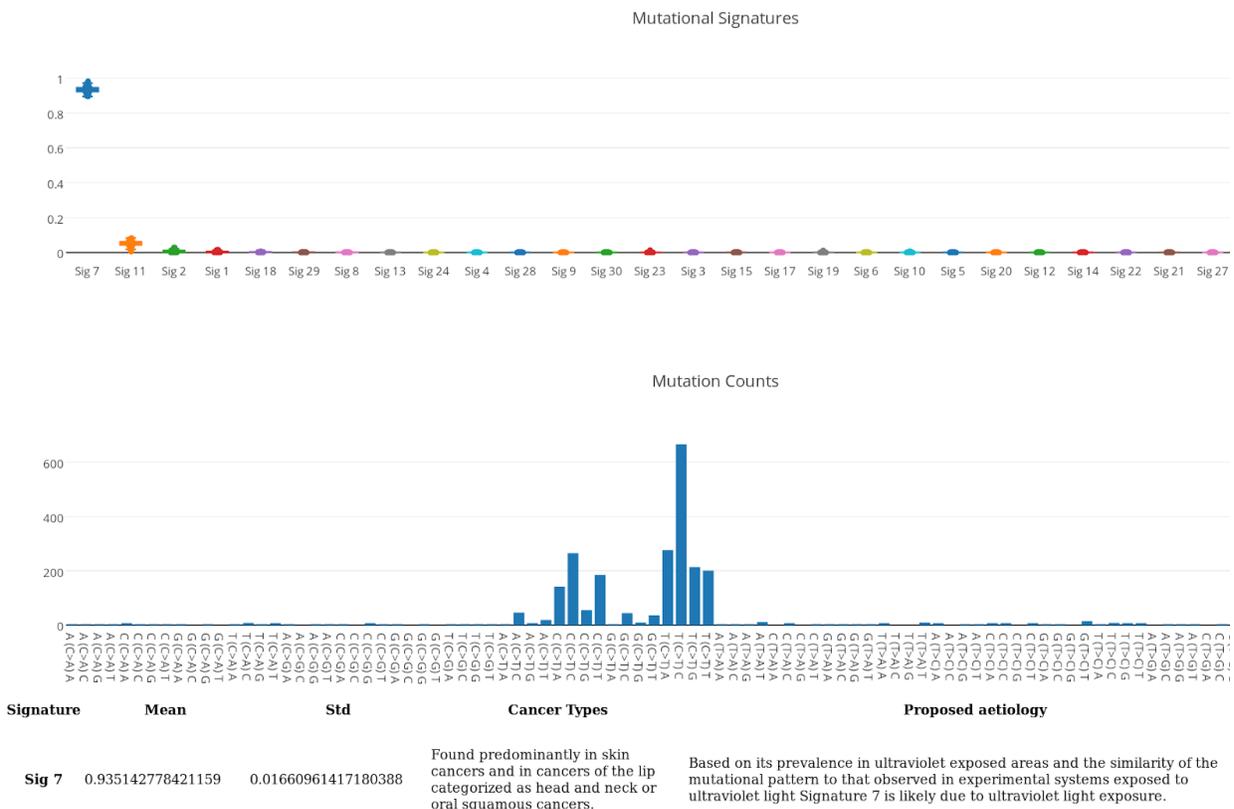
34	<input type="range"/>	56	<input type="button" value="-"/>
0	<input type="range"/>	16	<input type="button" value="-"/>
87	<input type="range"/>	100	<input type="button" value="-"/>



## 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.



## 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.



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 'sequencing | data | patients | tests | analysis | classifications | genes | variants | annotation' with a search box and a 'Go' button. The user 'sking-smith' is logged in. On the left sidebar, 'Gene Lists' is selected. The main content area has a 'Jump to gene list:' dropdown menu. Below it, a 'New GeneList' button is highlighted. A toolbar contains 'User', 'fulgent', 'GeneInfo', and 'iVITE' icons. A table titled 'Gene Lists' displays the following data:

ID	name	Uploaded by
COIs.txt		Im cintyre
COIs_per_line.txt		Im cintyre
Non-im in 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_46534.txt		Isanchez
150950_NB501009_0007_AHCNTTBQXX_10_FALCONI_ANAEMIA_MC_46600_S10_QC.txt		Irawlings
FALCONI_ANAEMIA_KABUKI_MC_46600.txt		Irawlings

At the bottom of the table, it shows 'Page 1 of 15' and a dropdown menu set to '10'.

on New GeneList

Click

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

Gene Lists | Genes | GeneGrid

Jump to gene list:

Create New GeneList

Name:

User: GenInfo:

ID	name	
OOIs.txt		Im cint
OOIs_per_line.txt		Im cint
Non-immune fetal hydrops_20150826.txt		kibron
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	dbnsfp rs id	gene symbol	snpeff transcript id	snpeff am
<input type="checkbox"/> 12361004	17	41219853	ATT	ATT		BRCA1	ENST00000471181	c.5050-14
<input type="checkbox"/> 12366274	17	41279868	T	G		BRCA1	ENST00000471181	
<input type="checkbox"/> 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. On the left, there are several dropdown menus for 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...), and OMIM (OMIM: 104200 ALPORT SYNDROME...). In the center, there is a 'Custom Gene List' section with a 'Name' field and a 'Gene names...' text area, and an 'Add Custom Gene List' button. On the right, there are 'Evidence columns' with checkboxes for CinGen, PanelApp, Color, and Coverage. Below this is a table of gene comparisons.

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						
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.

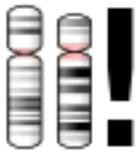
Choose lists from the top left select boxes, or manually paste in gene names into the **Custom Gene List** text entry box. Click the  red delete button to remove a gene list column.

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

<b>BRCA2</b>	<input type="text"/>
	<input type="button" value="save"/> <input type="button" value="cancel"/>
<b>CDH1</b>	CDH1
<b>GATA2</b>	+1
<b>MLH1</b>	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

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:

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



---

CHAPTER  
**TWENTYTWO**

---

**TEST ORDERING**



Menu: [patients]

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.

Click the graphs to filter patients, or enter terms:

**Human Phenotype Ontology**  
Phenotype...

**OMIM**  
MIM description...

**Genes**  
Gene...

**Jump to Patient:**  
Patient Name...

[Create New Patient...](#)

Filtering grid to **Microcephaly...** [show all](#)

Human Phenotype Ontology Patient Count

OMIM Patient Count

Patients							Phenotype	OMIM	Genes
ID	desci	last name	first name	sex	date of birth	date of dea			
							Hearing impairment Intellectual disability Microcephaly Precocious puberty Short stature		
							Edema Microcephaly		
							Abnormality of neuronal migration Global developmental delay Lissencephaly Microcephaly Seizures		TUBG1
							Ataxia Central hypotonia Developmental regression Global developmental delay Microcephaly Seizures Tremor Truncal ataxia		TPP1
							Microcephaly Seizures Short stature	ALPHA-THALASSEMIA/MENTAL RETARDATION SYNDROME, LINKED-ATRFX	ATRX CUL4B
							Delayed speech and language development		

grid filtered to microcephaly

Patients

## 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 interface with the following elements:

- Navigation tabs:** Patient, Physical Samples (1), Modifications.
- Form fields:** First name, Last name, Date of birth, Date of death, Sex, State.
- Right-hand summary:** Human Phenotype Ontology (with dropdown), OMIM (with dropdown), Genes (with dropdown).
- Description box:** Contains text from NGS Database and Phenotype, with auto-suggested terms highlighted in orange and green. The highlighted terms include: *Unexpected interstitial lung disease*, *ABCA3*, *AP3B1*, *CSF2RA*, *CSF2RB*, *SFTPB*, *SFTPC*, *FOXF1*, *NKX2-1*, *SFTPA2*, *SLC7A7*, *TERT*, *TINF2*, *HPS1*, *HPS4*, *DKC1*, *FLNA*, and *Ciliary Dyskinesia*.
- 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.

The screenshot shows the cohort management interface. At the top, there are tabs for 'Details' and 'Sharing / Permissions'. Below the tabs, there are several input fields: 'Name' (190208HamishScott\_WGS\_...), 'Date' (2019-03-20 11:57:29), 'User' (dlawrence), 'Project' (-----), and 'Import status' (success). Below these fields, there is a 'Processing' section with a 'View upload processing' link.

The main section is titled 'Samples' and contains a table with the following columns: 'Sample', 'Variants (passed)', 'VCF Sample Name', 'Name', 'Patient', 'Specimen', and 'BAM path'. The table has four rows of data, with the last three rows selected (checkboxes checked).

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...	

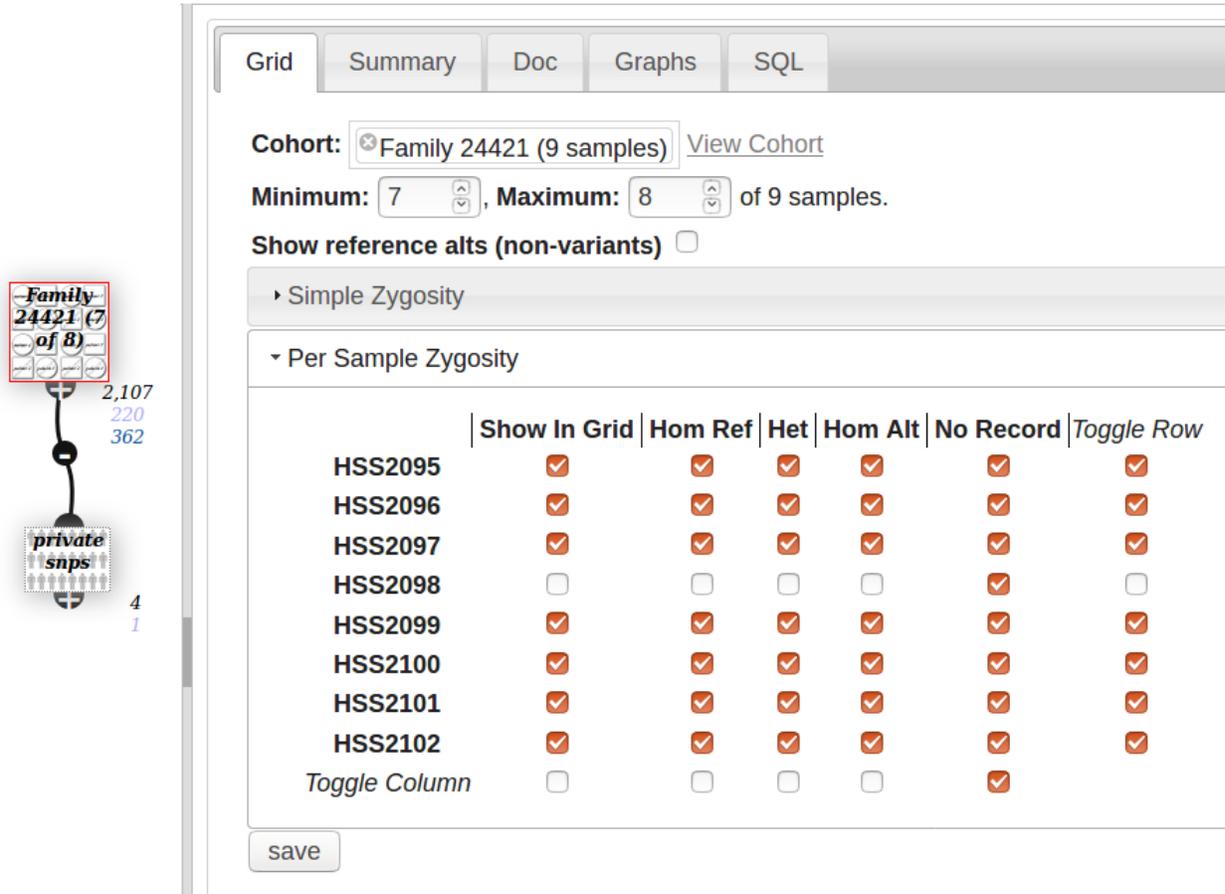
Below the table, there are several buttons: 'Update VCF', 'Perform trio analysis using template', and 'Create cohort from selected samples'. A message below the buttons states: 'More than 3 samples - select exactly 3 using checkboxes on the left. 3 samples selected.'

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).



The screenshot shows the VariantGrid interface for a cohort. On the left, a tree view shows a node for 'Family 24421 (of 8)' with 2,107 variants, and a sub-node for 'private snps' with 4 variants. The main panel has tabs for 'Grid', 'Summary', 'Doc', 'Graphs', and 'SQL'. The 'Grid' tab is active, showing a 'Cohort: Family 24421 (9 samples)' with a 'View Cohort' link. Below this, there are filters for 'Minimum: 7' and 'Maximum: 8' of 9 samples, and a checkbox for 'Show reference alts (non-variants)'. The main table is titled 'Simple Zygosity' and 'Per Sample Zygosity'. It has columns for 'Show In Grid', 'Hom Ref', 'Het', 'Hom Alt', 'No Record', and 'Toggle Row'. The table lists variants HSS2095 through HSS2102, with checkboxes for each column. A 'Toggle Column' row is at the bottom. A 'save' button is at the bottom left of the table area.

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

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:

Cohort

Legend



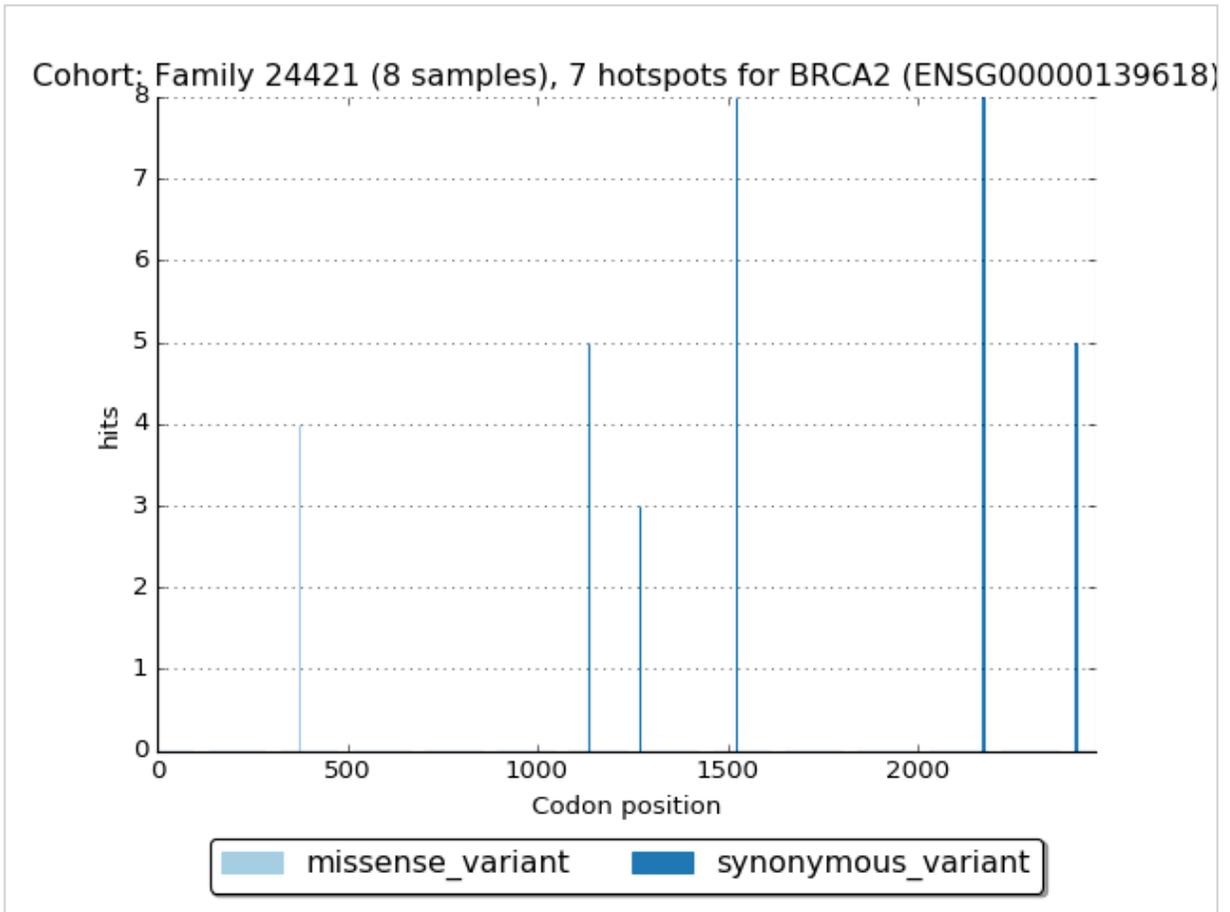
	S000	S001	S002	S004	S005	S006	S007	S008	S009	S010	S011	S012	S013	S014	S015	S016
Age																
HPO																
MIM																
<b>RUNX1</b>	1	1	1					1				1	1	1	1	3
ASXL1			1											2	1	2
BCOR	1	1						1				1	1		3	1
BCORL1			1									1		1	5	6
CBL																
CDC25C			2	1	1	1	1							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

Matrix

Gene/Sample

## Cohort: Family 24421

Gene × BRCA2 (ENSG00000139618) View Gene



Hotspots graph

Cohort

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\_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...	

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.

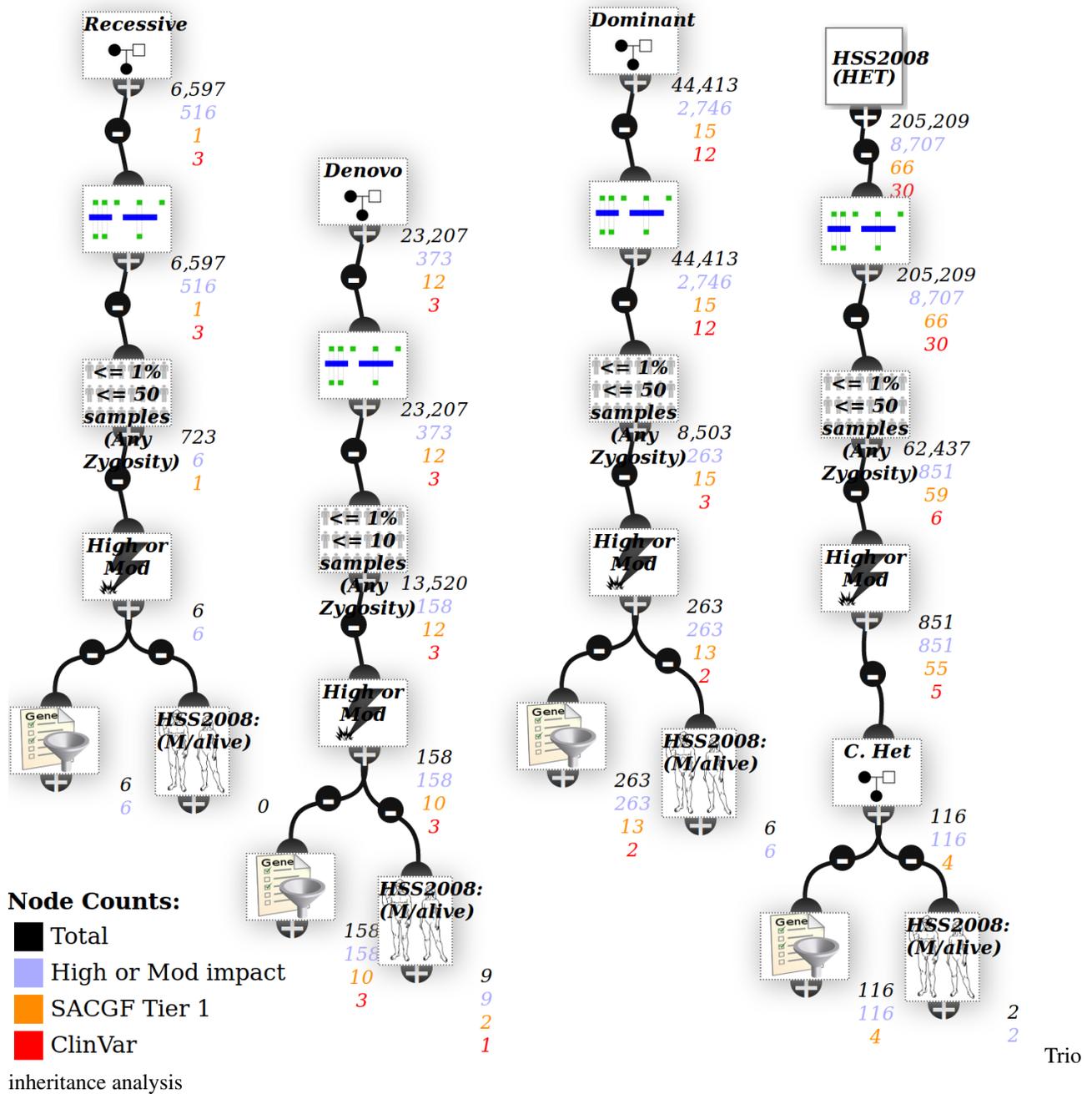
## 26.2 Digital karyomapping

By checking a trio's zygoty, 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.

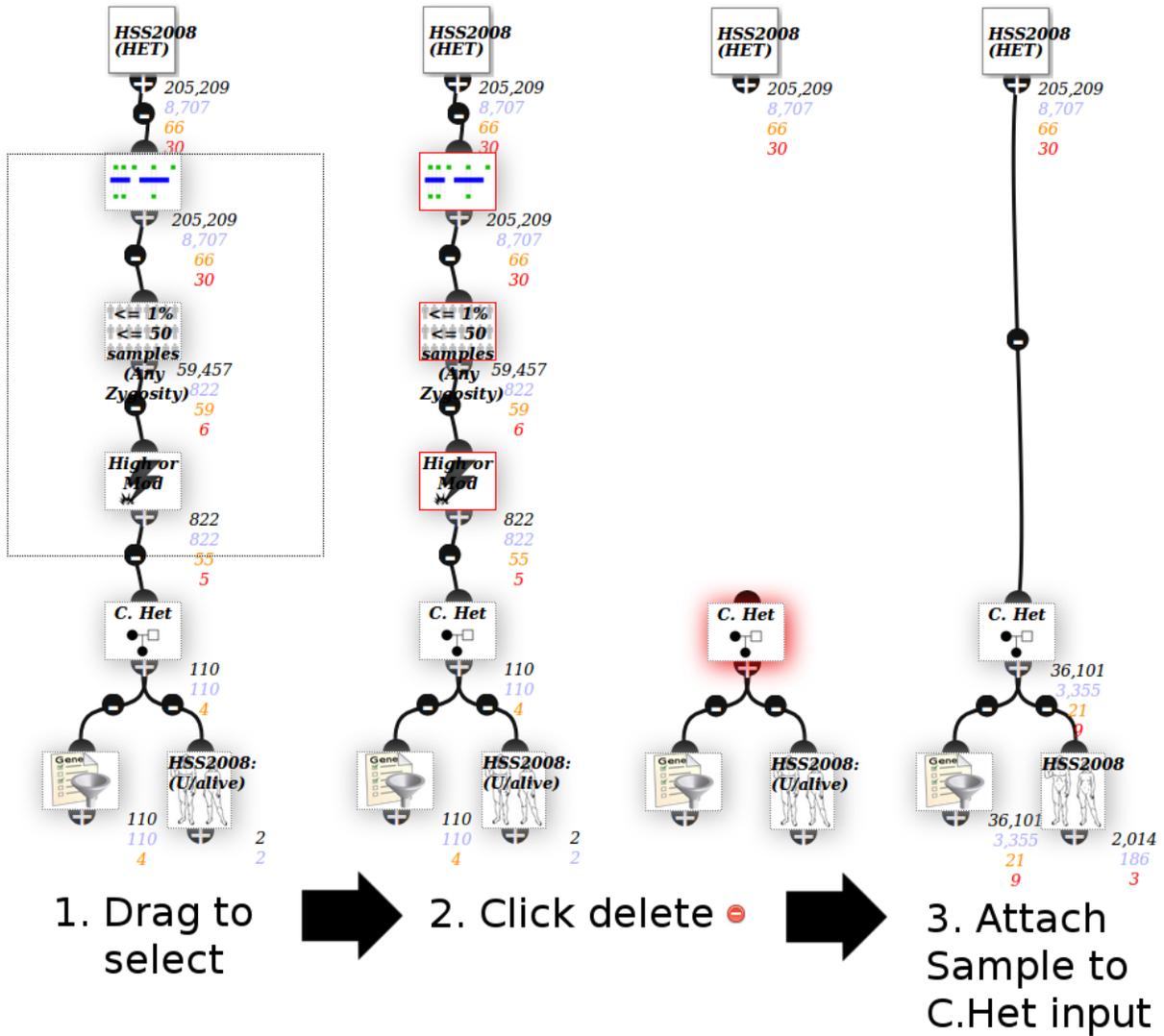
### 26.3.2 Compound Het filter

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.

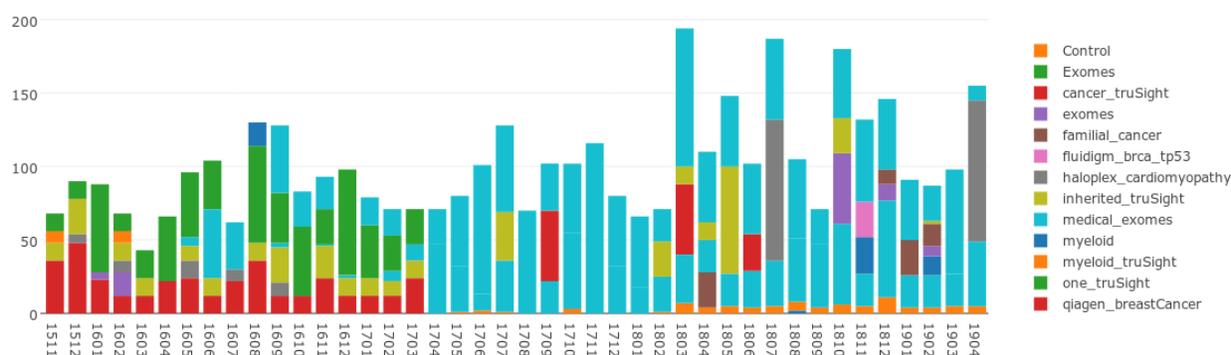
This is because you probably don't want every gene with  $\geq 2$  variants, but rather only  $\geq 2$  damaging/rare ones. Adjust the filters above the C.Het node to adjust this.

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



## 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 All](#)

Show Incomplete Data:  Show Hidden Data:

name	Sampl	Model	Sequencer	QC Lo	Experiment	EnrichmentKit	Kit ver	Gold	Hidden	Bad	VCF	path
190412 NB501008 0315 AH2HG5BGBX	F 11	NextSeq 500	NB501008	Complete	R1KD_19_009	roche_1k_disease	6					/tau/data/clinical/unaligned/roche_1k_disease/190412_NB501008_0
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_0
190324 NB501008 0308 AHFMM5AFXY	25	NextSeq 500	NB501008	Complete	R1KD_019_004	roche_1k_disease	6					/tau/data/clinical/unaligned/roche_1k_disease/190324_NB501008_0
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_0
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_0
190121 NB501008 0294 AHCNFGAFXY	21	NextSeq 500	NB501008	Complete	R1KD019_002	roche_1k_disease	6					/tau/data/clinical/unaligned/roche_1k_disease/190121_NB501008_0
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_0
181217 NB501008 0283 AHHWGAFXY	25	NextSeq 500	NB501008	Complete	R1KD18_028	roche_1k_disease	6					/tau/data/clinical/unaligned/roche_1k_disease/181217_NB501008_0
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_0
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_0
181112 NB501008 0266 AHGJCNBGBX	F 19	NextSeq 500	NB501008	Complete	R1KD_18_025_FFPE	roche_1k_disease	6					/tau/data/clinical/unaligned/roche_1k_disease/181112_NB501008_0
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_0
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_0
181008 NB501009 0227 AHC7F3AFXY	24	NextSeq 500	NB501009	Complete	R1KD_18_022	roche_1k_disease	6					/tau/data/clinical/unaligned/roche_1k_disease/181008_NB501009_0
180926 AHC7GAFXY AHC7KAFXY Me	25	NextSeq 500	NB501008	Error	R1KD18_021	roche_1k_disease	6					/tau/data/clinical/unaligned/roche_1k_disease/180926_AHC7GAFXY
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_0
180813 NB501008 0233 AHGG75BGBX	F 8	NextSeq 500	NB501008	Complete	R1KD_18_019_FFPE	roche_1k_disease	6					/tau/data/clinical/unaligned/roche_1k_disease/180813_NB501008_0
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_0
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_0
180709 NB501009 0195 AH7GH2AFXY	22	NextSeq 500	NB501009	Complete	R1KD18_016	roche_1k_disease	6					/tau/data/clinical/unaligned/roche_1k_disease/180709_NB501009_0
180702 NB501008 0221 AHK5G3BGBX	11	NextSeq 500	NB501008	Complete	R1KD_18_015_FFPE	roche_1k_disease	6					/tau/data/clinical/unaligned/roche_1k_disease/180702_NB501008_0
180625 NB501009 0189 AH7FVAFXY	25	NextSeq 500	NB501009	Complete	R1KD_18_014	roche_1k_disease	6					/tau/data/clinical/unaligned/roche_1k_disease/180625_NB501009_0
180608 NB501009 0186 AH2JWAFXY	25	NextSeq 500	NB501009	Complete	R1KD_18_013	roche_1k_disease	6					/tau/data/clinical/unaligned/roche_1k_disease/180608_NB501009_0
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_0
180514 NB501009 0178 AH33KAFXY	25	NextSeq 500	NB501009	Complete	R1KD_18_010	roche_1k_disease	6					/tau/data/clinical/unaligned/roche_1k_disease/180514_NB501009_0
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_0
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_0
180329 NB501009 0169 AH2JWAFXY	25	NextSeq 500	NB501009	Complete	R1KD_18_007	roche_1k_disease	6					/tau/data/clinical/unaligned/roche_1k_disease/180329_NB501009_0
180319 NB501009 0167 AHYGH3AFXY	20	NextSeq 500	NB501009	Complete	R1KD_18_006	roche_1k_disease	6					/tau/data/clinical/unaligned/roche_1k_disease/180319_NB501009_0
180309 NB501009 0165 AHMY7NBGX5	F 13	NextSeq 500	NB501009	Complete	R1KD_005_FFPE	roche_1k_disease	6					/tau/data/clinical/unaligned/roche_1k_disease/180309_NB501009_0
180309 NB501009 0165 AHMY7NBGX5	F 13	NextSeq 500	NB501009	Complete	R1KD_005_FFPE	roche_1k_disease	6					/tau/data/clinical/unaligned/roche_1k_disease/180309_NB501009_0

Automatically

loaded sequencing runs + VCFs

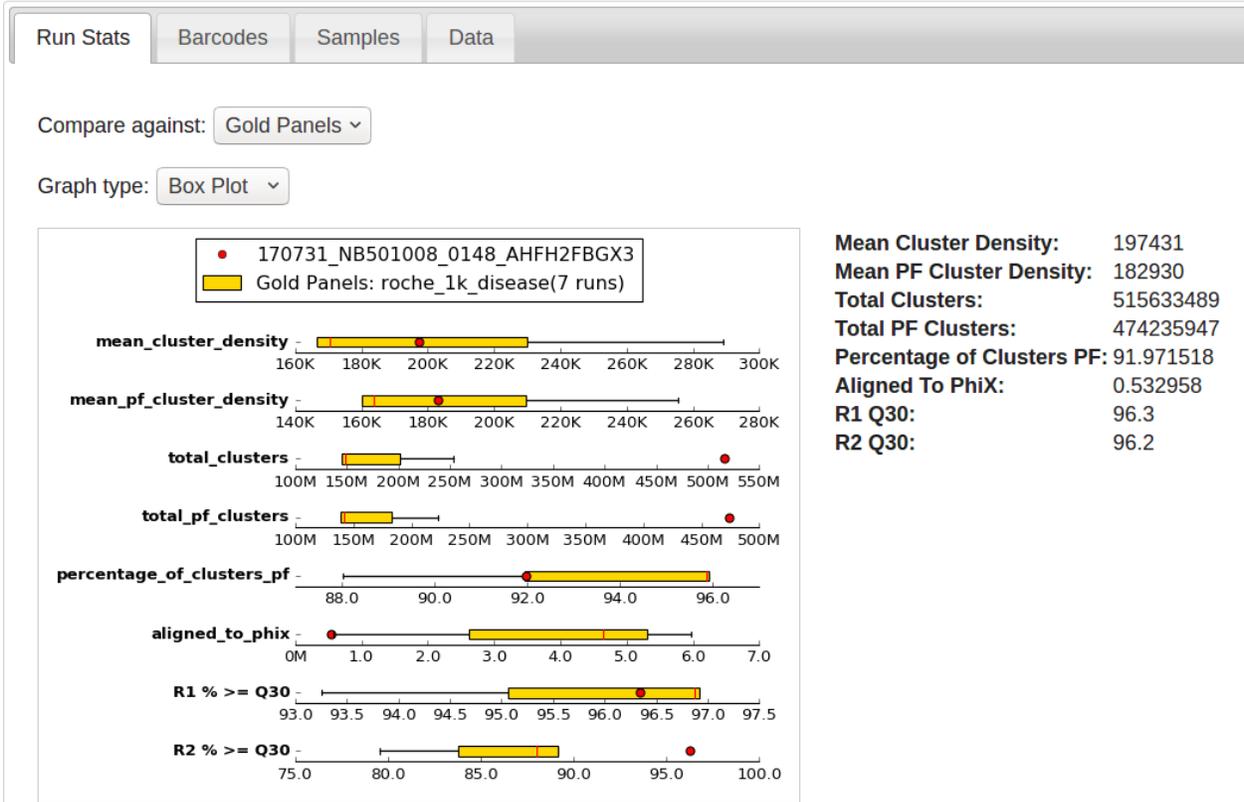


# 170731\_NB501008\_0148\_AHFH2FBGX3

Path: /tau/data/clinical/unaligned/roche\_1k\_disease/170731\_NB501008\_0148\_AHFH2FBGX3

NB501008 BaseCall Data: Deleted

Project: [170731\\_NB501008\\_0148\\_AHFH2FBGX3\\_GATK\\_combined](#)



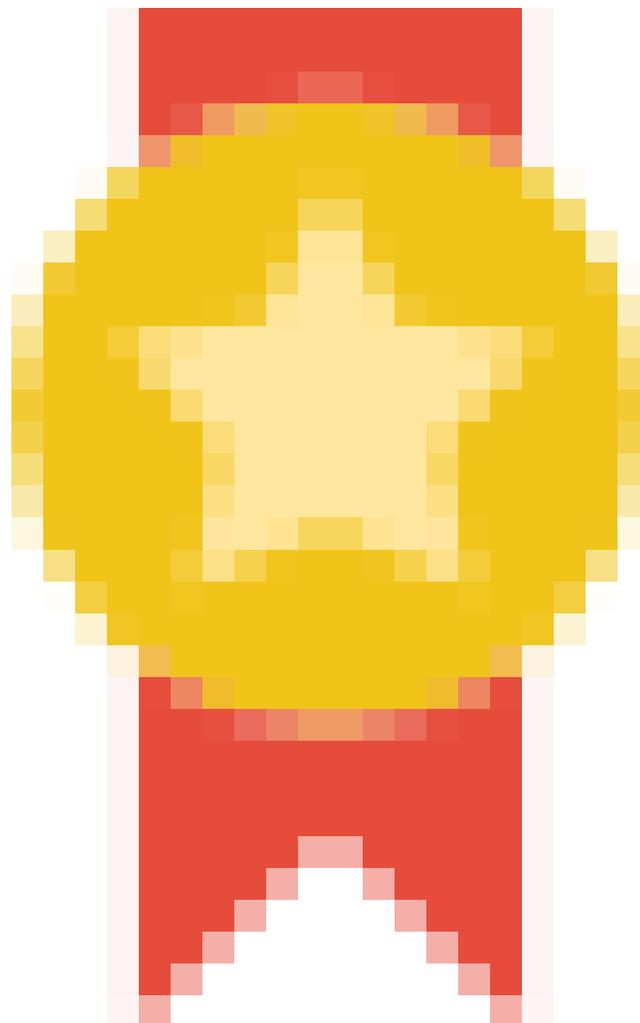
A

## 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.



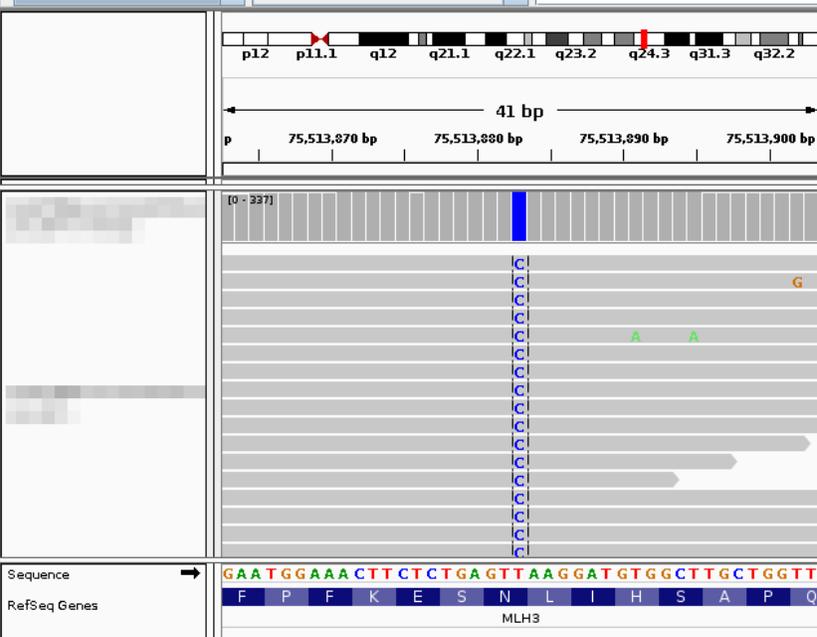
## IGV INTEGRATION



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	75483813	T	C	rs12713	MLH3
<input type="checkbox"/>	14	75483813	T	C	rs12713	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

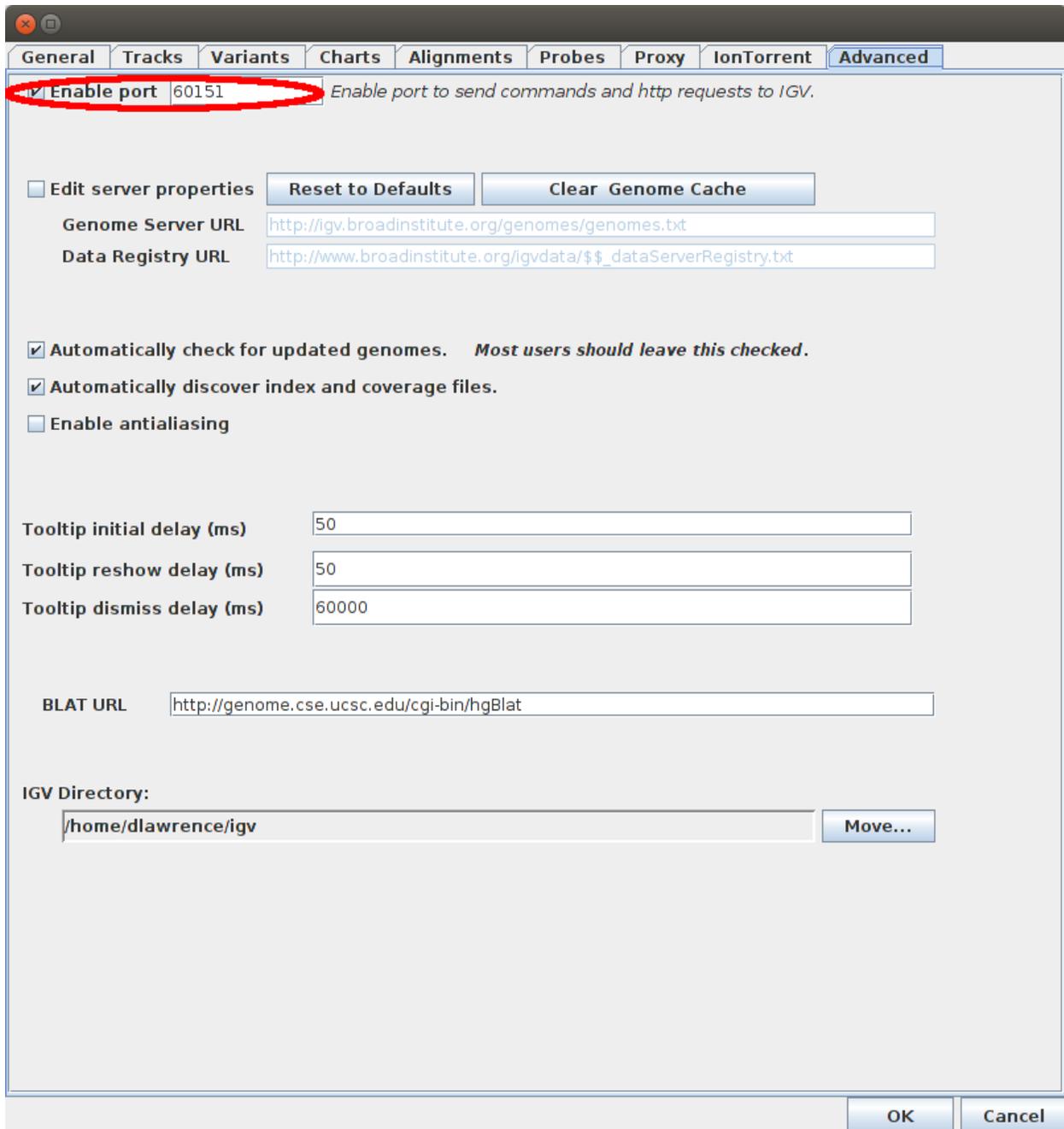


Sequence → GAATGGAAACTTCTCTGAGTTAAGGATGTGGCTTGCTGGTT  
 RefSeq Genes → F P F K E S N L I H S A P Q  
 MLH3

### 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.

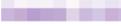
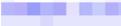
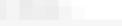
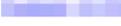
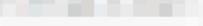
You can set all the samples in a VCF file at once in the vcf page, click Bulk Set Fields to set all samples according to a pattern based on the sample name.

### 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

Clear Filter

**Variant**

ClinGen Canonical Allele Identifier

Ensembl Gene ID

Gene symbol

\*Genome Build

Gene OMIM ID

RefSeq Transcript ID

Ensembl Transcript ID

HGNC ID

UniProt ID

Variant coordinate

g.HGVS

c.HGVS

p.HGVS

Molecular consequence

\*Zygoty

**Y-Path Zues Lab / vc768**

NM\_004360.4(CDH1):c.535A>G,  
NP\_004351.1:p.Lys179Glu  
VUS (3)

**Links**

ClinGen Allele Registry	ClinGen KB
Clinvar Variant	Genomizer
gnomAD	GTEX
Monarch Phenotypes	NCBI
OMIM (Gene)	PDB
UCSC	Uniprot ID

**Flags**

**Zygoty**

blank

Zygoty in the tested individual.

Does the allele frequency agree with the zygoty? 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**

**Zygoty** - Missing mandatory value

**Gene**

\*Condition under curation

Gene-disease validity

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

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



### 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 criteria 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   ? 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

Variant Classifications

Gene Classifications

Mine Gene:

[Simple Filter...](#) or [Advanced classification search](#)

VariantClassification									
ID	Status	clinical significance	c.HGVS	Gene Symbol	Lab Name	Lab Record ID	User	Created	
		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 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](#)