


Exercise: Exploring variants in Ensembl Fungi

In any of the sequence views shown in the Gene and Transcript tabs, you can view variants on the sequence. You can do this by clicking on  [Configure this page](#) from any of these views.

Let's take a look at the Gene sequence view for *ADH4* (Gene Stable ID: YGL256W). This gene is a ribonuclease protein in *Saccharomyces cerevisiae* R64-1-1. Select *Saccharomyces cerevisiae* R64-1-1 under [Favourite Genomes](#) on the Ensembl Fungi homepage, search for **YGL256W** and go to the [Variant image](#) view.

Search:

Saccharomyces cerevisiae

▼

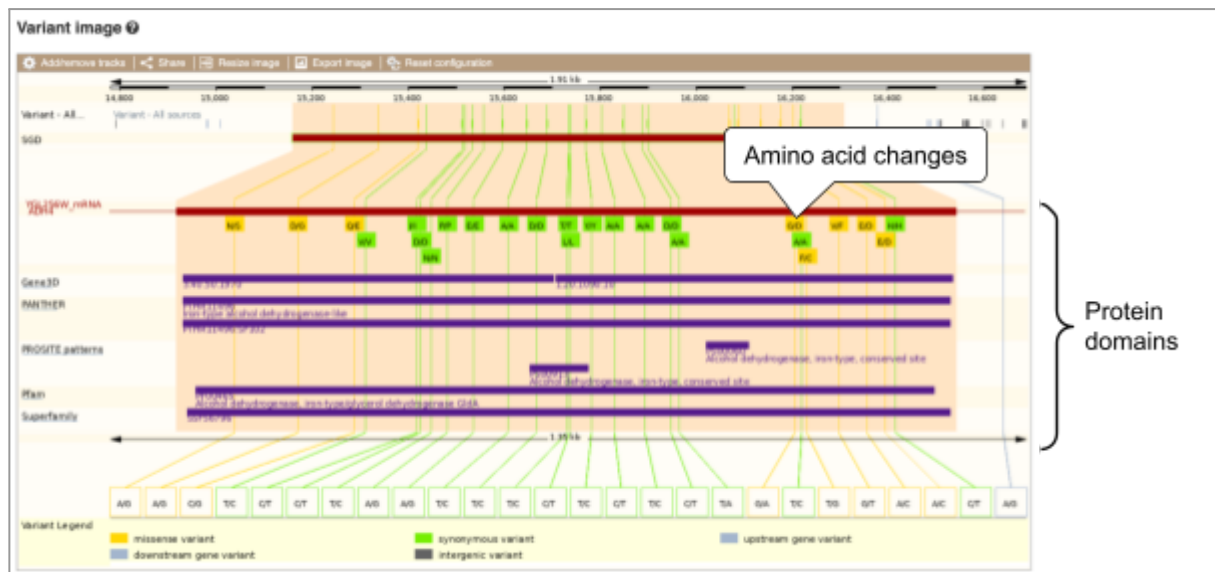
for

YGL256W

Go

e.g. **NAT2** or **alcohol***

This view shows variants mapped to the gene structure and protein domains.



We can examine all variants and filter to see ones we are interested in using the variant table. Click on the [Variant table](#) link.

This table shows the variants in order of their occurrence through the genome, and they are reported on the forward strand. The gene *ADH4* is located on the forward strand, so we are first shown variants upstream of the gene (starting at the 5' upstream region).

(a) How many variants in this gene are predicted to be missense?

You can filter the table to view variants that alter the protein sequence. Click on the [Consequences: All](#) button above the table. Click the option '[PTV and Missense](#)' in the pop

up, then [Apply](#). You can also filter by other columns such as variant [Class](#).

(b) Are there any known variants in this gene predicted to be deleterious?

The SIFT scores predict the consequence of the variant on the function of the protein taking into account chemical changes and conservation of amino acids. Scores <0.05 and coloured red are 'deleterious' while scores >0.05 and coloured green are tolerated.

Consequences: All ☐ Filter Other Columns

Consequences (5/30 on)

Turn All Off PTV PTV & Missense Only Exonic Turn All On

PTV = Protein Truncating Variant

transcript ablation (0) Off	inframe deletion (0) Off	coding sequence variant (0) Off
splice donor variant (0) On	protein altering variant (0) Off	mature miRNA variant (0) Off
splice acceptor variant (0) On	missense variant (8) On	5 prime UTR variant (0) Off
splice donor 5th base variant (0) Off	splice region variant (0) Off	3 prime UTR variant (0) Off
stop gained (0) On	splice polypyrimidine tract variant (0) Off	non coding transcript exon variant (0) Off
frameshift variant (0) On	incomplete terminal codon variant (0) Off	intron variant (0) Off
stop lost (0) Off	splice donor region variant (0) Off	NMD transcript variant (0) Off
start lost (0) Off	synonymous variant (17) Off	non coding transcript variant (0) Off
transcript amplification (0) Off	start retained variant (0) Off	upstream gene variant (2) Off
inframe insertion (0) Off	stop retained variant (0) Off	downstream gene variant (6) Off

Apply Cancel

VII:15518 C/T SNP SGRP - - synonymous D

Table filters

Filter SIFT: All ☒ Consequences: splice donor variant... (5/30) ☒ Filter Other Columns

Showhide columns

Protein pathogenicity predictions

Variant ID	Chr: bp	Alleles	Class	Source	Evidence	Ctn. Sig.	Conseq. Type	AA	AA co-ord	SIFT	Transcript
s07-15244	VII:15244	A/G	SNP	SGRP	-	-	missense variant	N/S	29	1	YGL256W_mRNA
s07-15337	VII:15337	A/G	SNP	SGRP	-	-	missense variant	D/G	60	1	YGL256W_mRNA
s07-15420	VII:15420	C/G	SNP	SGRP	-	-	missense variant	Q/E	88	0.72	YGL256W_mRNA
s07-16069	VII:16069	G/A	SNP	SGRP	-	-	missense variant	G/Q	304	0.1	YGL256W_mRNA
s07-16087	VII:16087	T/G	SNP	SGRP	-	-	missense variant	F/C	310	0	YGL256W_mRNA
s07-16134	VII:16134	G/T	SNP	SGRP	-	-	missense variant	V/F	326	0.03	YGL256W_mRNA
s07-16179	VII:16179	A/C	SNP	SGRP	-	-	missense variant	E/D	339	0.26	YGL256W_mRNA
s07-16202	VII:16202	A/C	SNP	SGRP	-	-	missense variant	E/D	348	0.67	YGL256W_mRNA

Variant IDs are links to variant tab

Let's have a look at a specific variant. Click on the top result in the filtered table, or search for [s07-15244](#). This will open up the variation tab.

EnsemblFungi | HMMER | BLAST | BioMart | Tools | D | **Variant tab** | Blog | Login/Register

Saccharomyces cerevisiae (R64-1-1) ▾

Location: VII:15,159-16,307 | Gene: ADH4 | Transcript: ADH4 | Variant: s07-15244

Variant displays

- Explore this variant
 - Genomic context
 - Genes and regulation
 - Flanking sequence
 - Genotype frequency
 - Phenotype data
 - Sample genotypes
 - Linkage disequilibrium
 - Phylogenetic context
 - Citations
 - 3D Protein model
- Configure this page
- Custom tracks
- Export data
- Share this page
- Bookmark this page

s07-15244 SNP

Most severe consequence: missense variant | [See all predicted consequences](#)

Alleles: AAG | Highest population MAF: 0.46

Location: [Chromosome VII:15244](#) (forward strand) | VCF: VII:15244:s07-15244:A:G

HGVs names: This variant has 3 HGVs names - [Show](#)

External Links

Original source

About this variant

Variation features from SGRP, with Ensembl identifiers | [About SGRP](#)

This variant overlaps [1 transcript](#) and has [18 sample genotypes](#).

Explore this variant

Genomic context | Genes and regulation | Flanking sequence | Population genetics | Phenotype data | Sample genotypes | Linkage disequilibrium | Phylogenetic context | Citations

VCF

Variation icons (these go to the same places as the links in the left-hand navigation panel)

The icons show you what information is available for this variant.

(c) What are the genomic coordinates of this variant?

(d) What is the reference allele? (Hint: Ensembl always reports alleles on the forward strand. The reference allele is given first.)

(e) How many genes are affected by this variant? Does it have the same consequence across different transcripts of different genes?

Click on [Genes and regulation](#), or follow the link at the left.

Genes and regulation

Gene and Transcript consequences

Show/hide columns

Gene	Transcript (strand)	Allele (Tx, allele)	Consequence Type	Position in transcript	Position in CDS	Position in protein	AA	Codons	SIFT	Detail
YGL256W	YGL256W.mRNA (+)	G (G)	missense variant	86 (out of 1149)	86 (out of 1149)	29 (out of 382)	N/S	AAC/AGC	1	Show

No overlap with Ensembl Regulatory features

No overlap with Ensembl Motif features

This variant overlaps one gene. It causes a change in the protein sequence (missense variant) in the YGL256W gene we were looking at (note that only missense variants have SIFT scores).

(f) Which allele is major in the SGRP study?

Click on [Genotype frequency](#) in the left-hand menu. Note that the reference allele is more frequent than alternative allele in this case.

Genotype frequency ⓘ

Frequency data (1) ⓘ

Show/hide columns		Filter			
Population		Allele: frequency (count)		Genotype: frequency (count)	
SGRP	<div><div></div><div></div></div>	A: 0.538 (39)	G: 0.462 (39)	A: 0.538 (39)	G: 0.462 (39)

Additional Exercise – Variation data in *Fusarium oxysporum*

(a) Select the *Fusarium oxysporum* FO2 genome and search for **FOXG_13574T0** gene. One of its upstream variants is SNP **tmp_10_6610**. What are the possible alleles for this polymorphic position? Which one is on the reference genome?

tmp_10_6610 SNP

Most severe consequence: [upstream gene variant](#) | [See all predicted consequences](#)

Alleles: **C/T** | Highest population MAF: 0.15

Location: [Chromosome 10:6610](#) (forward strand) | VCF: 10 6610 tmp_10_6610 C T

HGVS name: [10:g.6610C>T](#)

External Links
Original source
About this variant

This variant overlaps [4 transcripts](#) and has [10 sample genotypes](#).

(b) What is the most frequent allele at this position? How many heterozygous individuals were observed in the melonis population?

Genotype frequency

Frequency data (1)

Population	Allele: frequency (count)	Genotype: frequency (count)
melonis	C : 0.850 (17) T : 0.150 (3)	C/C : 0.800 (8) C/T : 0.100 (1) T/T : 0.100 (1)

(c) Which individuals have got genotypes C/T and T/T?

Sample genotypes

Search for a sample: Search (e.g. NA18507)

[\(back to top\)](#)

Genotypes for melonis

Sample (Male/Female/Unknown)	Genotype (forward strand)	Population(s)	Father	Mother
866399 (U)	C/C	melonis	-	-
869404 (U)	C/C	melonis	-	-
869405 (U)	C/C	melonis	-	-
869406 (U)	C/C	melonis	-	-
869407 (U)	C/C	melonis	-	-
869408 (U)	C/C	melonis	-	-
869410 (U)	C/C	melonis	-	-
909453 (U)	C/C	melonis	-	-
909454 (U)	C/T	melonis	-	-
909455 (U)	T/T	melonis	-	-

Exercise: The Ensembl Fungi Variant Effect Predictor (VEP)

We have identified four variants in *Verticillium dahliae* JR2: chromosome 5, C->G at 698711, G->T at 698935, G->A at 700313 and C->A at 701484.

Use the Ensembl VEP to determine:

- (a) Are your variants novel or have they already been annotated in Ensembl? (b) What genes are affected by your variants?
- (c) Do any of your variants affect gene regulation?

Click on [Tools](#) in the top brown bar from any Ensembl Fungi page, then [Variant Effect Predictor](#) to open the input form. You will need to change the species to *Verticillium dahliae* JR2 and paste your input data in the provided text box.

The VEP recognises a number of input formats including the Ensembl default format, VCF, Variant identifiers and HGVS notations.

The Ensembl default format is composed of four compulsory columns and additional 'strand' column: Chromosome, Start Position, End Position, Alleles (reference/alternate), Strand (1 for forward; -1 for reverse), with one line per variant. Your variants in this format would look like this:

```
5 698711 698711 C/G
5 698935 698935 G/T
5 700313 700313 G/A
5 701484 701484 C/A
```

The screenshot shows the Ensembl Variant Effect Predictor (VEP) web interface. It includes a 'New job' button and a 'Clear form' link. The 'Species' section shows 'Verticillium dahliae' selected, with a 'Change species' link. The 'Name for this job (optional):' field contains 'Fungal Pathogens VEP Exercise'. The 'Input data:' section has a text box with the following content: '5 698711 698711 C/G', '5 698935 698935 G/T', '5 700313 700313 G/A', and '5 701484 701484 C/A'. Below the text box is a 'Run instant VEP for current line' button. To the left of the text box is a 'Paste or type in data...' label. Below the text box are links for 'Examples: Ensembl default, VCF, Variant identifiers, HGVS notations'. To the right of the text box is a 'See preview of the results for a selected line' label. Below the text box are two options: '...or upload a file...' with a 'Choose file' button and 'No file chosen' text, and '...or provide a URL to a file hosted online' with a text input field.

Annotations on the screenshot:

- Change the species
- Name your job
- Paste or type in data...
- See preview of the results for a selected line
- See data format examples
- ...or upload a file...
- ...or provide a URL to a file hosted online

The VEP will automatically detect that the data is in Ensembl default format. Clicking on the ‘Run instant VEP for current line’ will generate a pop-up with summarised results for that individual variant.

Instant results for 5 701484 701484 C/A

Instant VEP

The below is a preview of results using the *Verticillium dahliae* Ensembl transcript database and does not include all data fields present in the full results set. To obtain these please close this preview window and submit the job using the **Run** button below.

Most severe consequence: [upstream_gene_variant](#)

Colocated variants: [imp_5_701484_C_A](#)

Gene/Feature/Type	Consequence	Details
VDAG_JR2_Chr5g02160a:VDAG_JR2_Chr5g02160a-00001 Type: protein_coding	downstream_gene_variant	Distance to transcript: 2165bp
VDAG_JR2_Chr5g02170a:VDAG_JR2_Chr5g02170a-00001 Type: protein_coding	downstream_gene_variant	Distance to transcript: 742bp
VDAG_JR2_Chr5g02170a:VDAG_JR2_Chr5g02170a-00002 Type: protein_coding	downstream_gene_variant	Distance to transcript: 778bp
VDAG_JR2_Chr5g02171a:VDAG_JR2_Chr5g02171a-00001 Type: protein_coding	upstream_gene_variant	Distance to transcript: 64bp

There are further options that you can choose for your output. These are categorised as [Identifiers](#), [Variants and frequency data](#), [Additional annotations](#), [Predictions](#), [Filtering options](#) and [Advanced options](#). Let's open all the menus and take a look.

Identifiers ☐ Additional identifiers for genes, transcripts and variants

Identifiers

Gene symbol: ☒

Transcript version: ☒

Protein: ☐

UniProt: ☐

NCBI: ☐

Variants and frequency data ☐ Co-located variants and frequency data

Variants and frequency data

Find co-located known variants:

Variant synonyms: ☐

Include flagged variants: ☐

Which identifiers do you want in the output?

Does this variant already exist?

Additional annotations Additional transcript, protein and regulatory annotations

Transcript annotation

Transcript biotype: ☒

Exon and intron numbers: ☐

Identify canonical transcripts: ☐

Upstream/Downstream distance (bp):

mRNA structure: ☐

NMD: ☐

UTRAnnotation: ☐

Protein annotation

Protein matches: ☐

IntAct: ☒ Disabled ☐ Enabled

Predictions Variant predictions, e.g. SIFT, PolyPhen

Splicing predictions

dbSNV: ☐

MaxEntScan: ☐

SpliceAI: ☒ Disabled ☐ Enabled

Conservation

BLOSUM62: ☐

Ancestral allele: ☐

Phenotype data and citations

Phenotypes: ☐

Gene Ontology: ☐

DisGeNET: ☐

Mastermind: ☐

Filtering options Pre-filter results by frequency or consequence type

Filters

Return results for variants in coding regions only: ☐ **Show only coding variants**

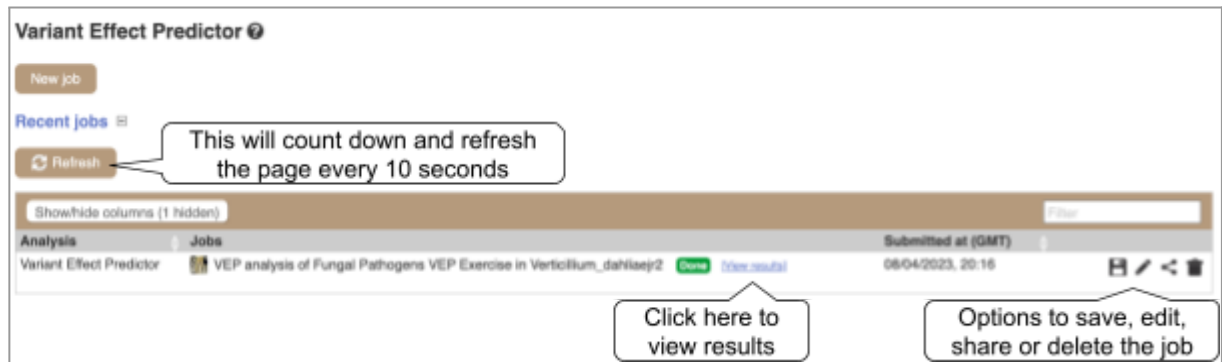
Restrict results: **More filter**

NB: Restricting results may exclude biologically important data

Advanced options Additional enhancements

Run VEP

Hover over the options to see definitions. When you've selected everything you need, scroll right to the bottom and click [Run](#).

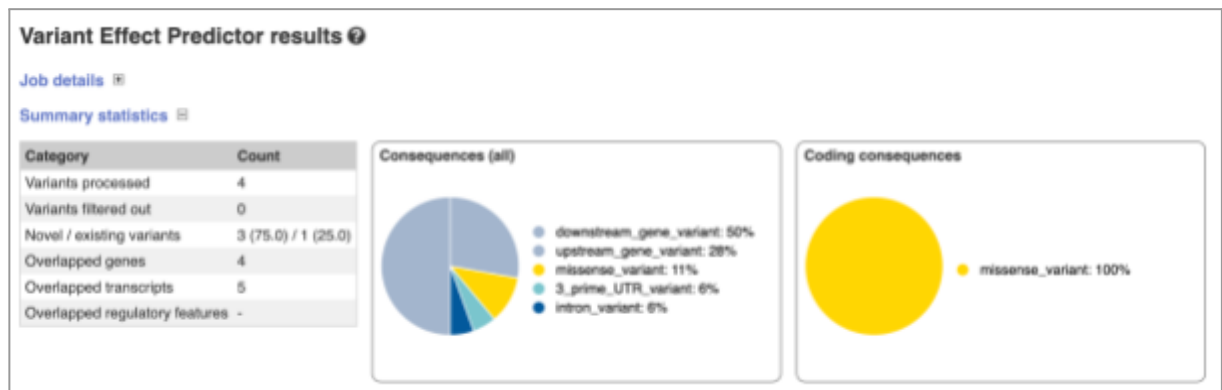


A table display will show you the status of your job. It will say **Queued**, then automatically switch to **Done** when the job is done, you do not need to refresh the page. You can edit or discard your job at this time. If you have submitted multiple jobs, they will all appear here.

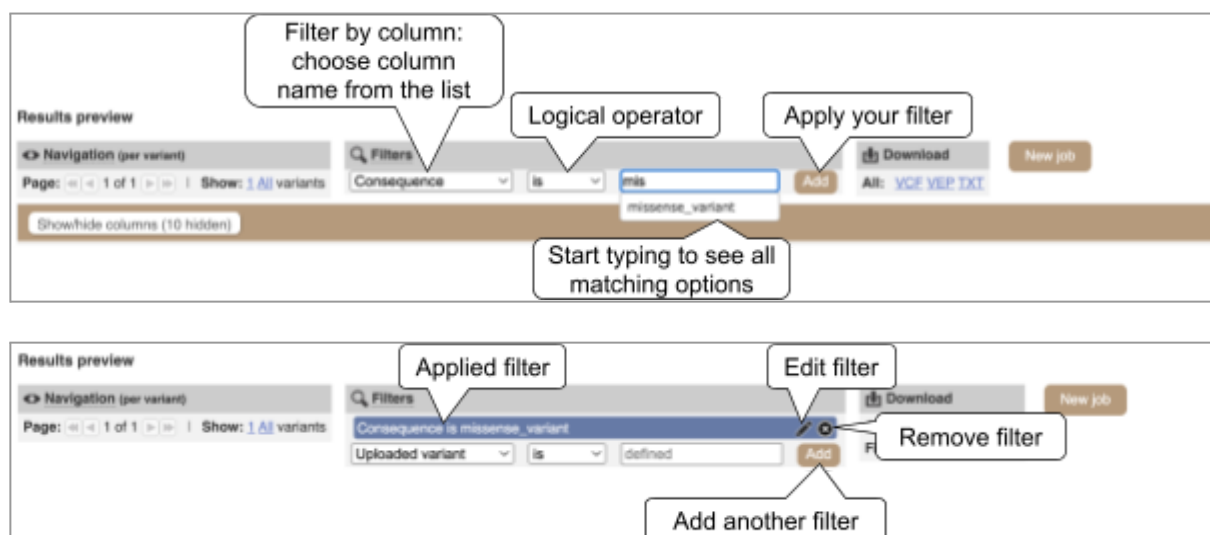
Click **View results** once your job is done. In your results you will see a graphical summary of your data, as well as a table of your results.

Let's come back to our questions:

- Are your variants novel or have they already been annotated in Ensembl?
- What genes are affected by your variants?
- Do any of your variants affect gene regulation?



The output table reports one variant consequence per row. If your variants have multiple alternate alleles, hit multiple genes or transcripts, you'll find few lines per variant. If the output table is large, you might want to use the filter option to narrow it down. Once you've added a filter, it will appear in the filter box, allowing you to add other filters.



Filter text box is by default set to 'defined', which can be used to filter out empty values, e.g. 'Existing variant' 'is' 'defined' will filter out variants with empty values in the 'Existing variant' column, leaving you with known variants only. Note that you should not type 'define' in the search box, just leave it as it is.

Filter this table

Download options

Results preview

Navigation (per variant)

Filters

Download

New job

Page: 1 of 1

Show: 1 variants

Uploaded variant

is

defined

Add

All: VCF VEP TXT

Show/hide columns (10 hidden)

Show additional columns

Uploaded variant	Location	Allele		Biotype	Exon	cDNA position	CDS position	Protein position	Amino acids	Codons	Existing variant	Feature strand			
5_698711_C/G	5_698711-698711	G	downstream_gene_variant	VDAG_JR2_Chr5g02150a	Transcript	VDAG_JR2_Chr5g02150a-00001.	protein_coding	-	-	-	-	1			
5_698711_C/G	5_698711-698711	G	intron_variant	VDAG_JR2_Chr5g02160a	Transcript	VDAG_JR2_Chr5g02160a-00001.	protein_coding	-	-	-	-	1			
5_698711_C/G	5_698711-698711	G	upstream_gene_variant	VDAG_JR2_Chr5g02170a	Transcript	VDAG_JR2_Chr5g02170a-00001.	protein_coding	-	-	-	-	1			
5_698711_C/G	5_698711-698711	G	upstream_gene_variant	VDAG_JR2_Chr5g02170a	Transcript	VDAG_JR2_Chr5g02170a-00002.	protein_coding	-	-	-	-	1			
5_698711_C/G	5_698711-698711	G	downstream_gene_variant	VDAG_JR2_Chr5g02171a	Transcript	VDAG_JR2_Chr5g02171a-00001.	protein_coding	-	-	-	-	-1			
5_698935_G/T	5_698935-698935	T	downstream_gene_variant	VDAG_JR2_Chr5g02150a	Transcript	VDAG_JR2_Chr5g02150a-00001.	protein_coding	-	-	-	-	1			
5_698935_G/T	5_698935-698935	T	3_prime_UTR_variant	VDAG_JR2_Chr5g02160a	Transcript	VDAG_JR2_Chr5g02160a-00001.	protein_coding	8/8	1679	-	-	1			
5_698935_G/T	5_698935-698935	T	upstream_gene_variant	VDAG_JR2_Chr5g02170a	Transcript	VDAG_JR2_Chr5g02170a-00001.	protein_coding	-	-	-	-	1			
5_698935_G/T	5_698935-698935	T	upstream_gene_variant	VDAG_JR2_Chr5g02170a	Transcript	VDAG_JR2_Chr5g02170a-00002.	protein_coding	-	-	-	-	1			
5_698935_G/T	5_698935-698935	T	downstream_gene_variant	VDAG_JR2_Chr5g02171a	Transcript	VDAG_JR2_Chr5g02171a-00001.	protein_coding	-	-	-	-	-1			
5_700313_G/A	5_700313-700313	A	downstream_gene_variant	VDAG_JR2_Chr5g02160a	Transcript	VDAG_JR2_Chr5g02160a-00001.	protein_coding	-	-	-	-	1			
5_700313_G/A	5_700313-700313	A	missense_variant	VDAG_JR2_Chr5g02170a	Transcript	VDAG_JR2_Chr5g02170a-00001.	protein_coding	2/2	155	52	18	A/T	GCC/AAC	-	1
5_700313_G/A	5_700313-700313	A	missense_variant	VDAG_JR2_Chr5g02170a	Transcript	VDAG_JR2_Chr5g02170a-00002.	protein_coding	2/2	161	52	18	A/T	GCC/AAC	-	1
5_700313_G/A	5_700313-700313	A	downstream_gene_variant	VDAG_JR2_Chr5g02171a	Transcript	VDAG_JR2_Chr5g02171a-00001.	protein_coding	-	-	-	-	-	-1		
5_701484_C/A	5_701484-701484	A	downstream_gene_variant	VDAG_JR2_Chr5g02160a	Transcript	VDAG_JR2_Chr5g02160a-00001.	protein_coding	-	-	-	-	tmo_5_701484_C_A	1		
5_701484_C/A	5_701484-701484	A	downstream_gene_variant	VDAG_JR2_Chr5g02170a	Transcript	VDAG_JR2_Chr5g02170a-00001.	protein_coding	-	-	-	-	tmo_5_701484_C_A	1		
5_701484_C/A	5_701484-701484	A	downstream_gene_variant	VDAG_JR2_Chr5g02170a	Transcript	VDAG_JR2_Chr5g02170a-00002.	protein_coding	-	-	-	-	tmo_5_701484_C_A	1		
5_701484_C/A	5_701484-701484	A	upstream_gene_variant	VDAG_JR2_Chr5g02171a	Transcript	VDAG_JR2_Chr5g02171a-00001.	protein_coding	-	-	-	-	tmo_5_701484_C_A	-1		

Existing variants

Variant 1

Variant 2

Variant 3

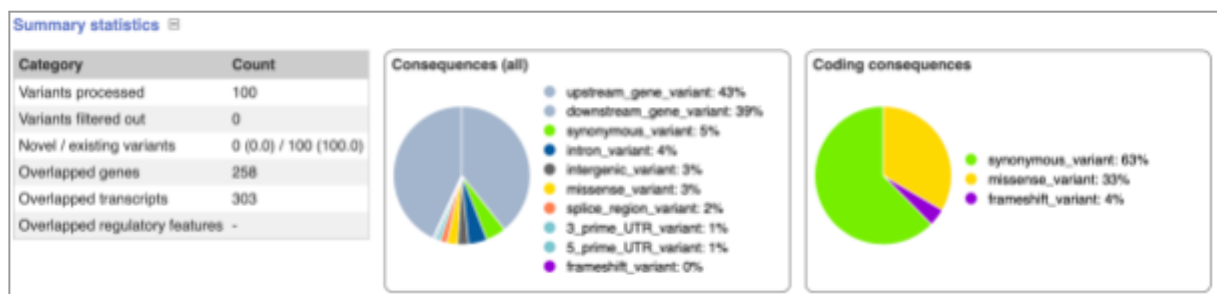
Variant 4

Additional Exercise: The Ensembl Fungi Variant Effect Predictor (VEP)

On the course file page, you will find a VCF file labelled VEP_exercise.vcf. This is a small subset of the outcome of *Puccinia graminis* Ug99 whole genome sequencing and variant calling experiment. This file can also be found on our FTP site under the following link:
http://ftp.ebi.ac.uk/pub/databases/ensembl/training/2021/FungalPathogens/VEP_exercise.vcf

Run the file through the VEP by downloading and uploading it from your computer, or alternatively by attaching it as a remote file hosted online (you will need to provide the FTP file URL).

- How many variants have been processed?
- How many genes and transcripts are overlapped by variants in this file?



- Do any of the variants change the amino acid sequences of any proteins? What genes? What is the amino acid change? (Hint: use the filters above the table to filter by consequences.)

Results preview

Navigation (per variant) | Filters | Download | New job

Show: 1 to 10 of 24 variants | Filter: Consequence is missense_variant | Filtered: VEP VEP TXT

Show/hide columns (22 hidden)

Location	Allele	Consequence	Gene	Exon	HGVSc	HGVSp	cDNA position	CDS position	Protein position	Amino acids	Codons	Existing variant	Domains
Supercontig_3.1381.801: 801	T	missense_variant	GMQ_27102	30	GMQ_27110T:c.284G>A	GMQ_27110T:p.G94G>S	289	289	89	E/K	GAG>AAG	ins_Supercontig_3.1381.801_C.T	Plex PF14303 PAWTHGRPT01845-125 PAWTHGRPT01845-125-SP3 ModDB_8a-mod45-Ba
Supercontig_3.1361.127684: 127684	C	missense_variant	GMQ_21813	26	GMQ_21813T:c.2387T>C	GMQ_21813T:p.Ser79Phe	233	236	79	S/P	TCC>CCC	ins_Supercontig_3.1361.127684_T.C	Geno3D.3.108.15.10 PROSITE_profiles:P5012904 Superfamily:SP575022
Supercontig_3.68.88930: 88930	C	missense_variant	GMQ_28452	404	GMQ_28452T:c.10893G>C	GMQ_28452T:p.Arg329His	1089	1089	329	D/N	GAT>GAT	ins_Supercontig_3.68.88930_G.C	Geno3D.3.40.735.10 PAWTHGRPT01845-125-SP1 Superfamily:SP575022 CDD:cd18023
Supercontig_3.48.118082: 118082	T	missense_variant	GMQ_28211	114	GMQ_28211T:c.739G>A	GMQ_28211T:p.Glu246Lys	73	73	25	E/K	GAA>AAA	ins_Supercontig_3.48.118082_G.T	-
Supercontig_3.41.7785: 7785	G	missense_variant	GMQ_08702	615	GMQ_08702T:c.1228A>C	GMQ_08702T:p.Gln409Pro	1228	1228	409	Q/P	CAG>CGG	ins_Supercontig_3.41.7785_T.G	PAWTHGRPT01845-125-SP3 PAWTHGRPT01845-125-SP3
Supercontig_3.18.171261: 171261	T	missense_variant	GMQ_04380	102	GMQ_04380T:c.287G>A	GMQ_04380T:p.Gly96Glu	287	287	96	G/E	GGA>GAA	ins_Supercontig_3.18.171261_G.T	-
Supercontig_3.73.169474: 169474	G	missense_variant	GMQ_03045	203	GMQ_03045T:c.407G>C	GMQ_03045T:p.Arg136Trp	407	407	136	R/T	AGA>ACA	ins_Supercontig_3.73.169474_G.G	Low_complexity_1Reg149 PAWTHGRPT01845-125-SP1 PAWTHGRPT01845-125-SP1
Supercontig_3.427.36213: 36213	A	missense_variant	GMQ_02814	202	GMQ_02814T:c.96G>T	GMQ_02814T:p.Gln322His	368	36	33	Q/H	CAG>CAT	ins_Supercontig_3.427.36213_G.A	PAWTHGRPT01845-125-SP1 PAWTHGRPT01845-125-SP15 ModDB_8a-mod45-Ba

- What are the HGVS notations of missense variants falling in known protein domains?

Results preview

Navigation (per variant) Filters

Show: 1 to 10 of 10 variants

Filters:

Match of the above rules

Uploaded variant: is defined

Show/hide columns (22 hidden)

Location	Allele	Consequence	Gene	Exon	HGVSc	HGVSp	cDNA position	CDS position	Protein position	Amino acids	Codon	Existing variant	Domains
Supercontig_3.1594.801.801	T	missense_variant	OMG_27112	30	OMG_27112T>C.2690G>A	OMG_27112T>C.Glu48>Lys	261	261	89	E/K	GAQAAQ	ins_Supercontig_3.1594.801_C_T	Plant PF14353 PANTHER:PTHR44128 PANTHER:PTHR44128_SF3 MotifDB_Neurofibin-like
Supercontig_3.1596.127664.127664	C	missense_variant	OMG_21813	26	OMG_21813T>C.235T>C	OMG_21813T>C.Ser79>Pro	235	235	79	S/P	TCCACC	ins_Supercontig_3.1596.127664_T_C	Geno3D:2.102.10.10 PROSITE_profile:PS012964 Superfamily:SGF5032
Supercontig_3.158.89935.89935	C	missense_variant	OMG_23457	414	OMG_23457T>C.1889G>C	OMG_23457T>C.Arg339>His	1809	1809	339	D/H	GATCAT	ins_Supercontig_3.158.89935_C_C	Geno3D:3.40.720.10 PANTHER:PTHR23371 Superfamily:SGF50348 CODon19083
Supercontig_3.41.7765.7765	G	missense_variant	OMG_28152	615	OMG_28152T>C.1325A>C	OMG_28152T>C.Glu443>Pro	1325	1325	443	Q/P	GAQCCG	ins_Supercontig_3.41.7765_T_G	PANTHER:PTHR48896 PANTHER:PTHR48896_SF3
Supercontig_3.73.160474.160474	G	missense_variant	OMG_33045	20	OMG_33045T>C.467G>C	OMG_33045T>C.Arg138>Thr	467	467	136	R/T	AGAADA	ins_Supercontig_3.73.160474_C_G	Low complexity (Seg-seq) PANTHER:PTHR21595 PANTHER:PTHR21595_SF1
Supercontig_3.427.55213.55213	A	missense_variant	OMG_33814	30	OMG_33814T>C.98G>T	OMG_33814T>C.Glu33>His	358	361	33	Q/H	CAGCAT	ins_Supercontig_3.427.55213_C_A	PANTHER:PTHR21361 PANTHER:PTHR21361_SF15 MotifDB_Neurofibin-like

(e) How many variants are frameshift? Which gene(s) do they fall in and which exons? Can you find a UniParc ID of protein(s) affected by this variant?

Results preview

Navigation (per variant) Filters

Show: 1 to 10 of 10 variants

Filters:

Match of the above rules

Uploaded variant: is defined

Show/hide columns (22 hidden)

Location	Allele	Consequence	Gene	Exon	HGVSc	HGVSp	cDNA position	CDS position	Protein position	Amino acids	Codon	Existing variant	UNIPARC
Supercontig_3.1482.1086.1087	G	frameshift_variant	OMG_27061	10	OMG_27061T>C.221del	OMG_27061T>C.Pro74>Kglu75del	220-221	220-221	74	R/K	CCGCG	ins_Supercontig_3.1482.1086_CCG_CG	UP000480546.P1087