


Exercise: Exploring variants in Ensembl Fungi

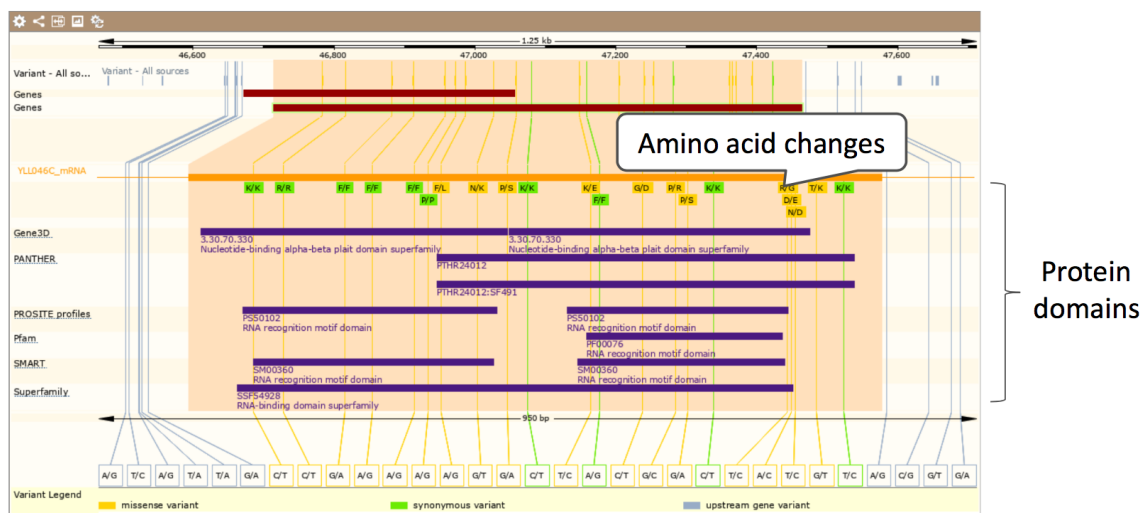
In any of the sequence views shown in the Gene and Transcript tabs, you can view variants on the sequence. You can do this by clicking on  [Configure this page](#) from any of these views.

Let's take a look at the Gene sequence view for *RNPI* (Gene Stable ID: YLL046C). This gene is a ribonuclease protein in *Saccharomyces cerevisiae*. Select *Saccharomyces cerevisiae* on the Ensembl Fungi homepage, search for **YLL046C** and go to the [Variant image](#) view.

Search:  for

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 *RNPI* is located on the reverse strand, so we are first shown variants downstream of the gene (starting at the 3' downstream 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 (5/27 on)

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

PTV = Protein Truncating Variant

transcript ablation (0) Off	inframe deletion (0) Off	mature miRNA variant (0) Off
splice donor variant (0) On	missense variant (11) On	5 prime UTR variant (0) Off
splice acceptor variant (0) On	protein altering variant (0) Off	3 prime UTR variant (0) Off
stop gained (0) On	splice region 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	stop retained variant (0) Off	NMD transcript variant (0) Off
start lost (0) Off	synonymous variant (10) Off	non coding transcript variant (0) Off
transcript amplification (0) Off	start retained variant (0) Off	upstream gene variant (11) Off
inframe insertion (0) Off	coding sequence variant (0) Off	downstream gene variant (8) Off

Apply » Cancel

gene variant

Filter SIFT: All Consequences: splice acceptor vari... (5/27) Filter Other Columns Table Filters

Show/hide columns Search...

Variant ID	Chr: bp	Alleles	Class	Source	Evidence	Clin. Sig.	Conseq. Type	AA	AA co-ord	SIFT	Transcript
s12-46985	XII:46986	A/G	SNP	SGRP	-	-	missense variant	F/L	160	0	YLL046C_mRNA
s12-47025	XII:47026	G/T	SNP	SGRP	-	-	missense variant	N/K	146	0.01	YLL046C_mRNA
s12-47057	XII:47058	G/A	SNP	SGRP	-	-	missense variant	P/S	136	0.19	YLL046C_mRNA
s12-47147	XII:47148	T/C	SNP	SGRP	-	-	missense variant	K/E	106	1	YLL046C_mRNA
s12-47203	XII:47204	C/T	SNP	SGRP	-	-	missense variant	G/D	87	0.72	YLL046C_mRNA
s12-47239	XII:47240	G/C	SNP	SGRP	-	-	missense variant	P/R	75	0.35	YLL046C_mRNA
s12-47252	XII:47253	G/A	SNP	SGRP	-	-	missense variant	P/S	71	0.72	YLL046C_mRNA
s12-47360	XII:47361	T/C	SNP	SGRP	-	-	missense variant	R/G	35	0.14	YLL046C_mRNA
s12-47364			SNP	SGRP	-	-	missense variant	D/E	33	0.51	YLL046C_mRNA
s12-47369			SNP	SGRP	-	-	missense variant	N/D	32	1	YLL046C_mRNA
s12-47392	XII:47393	G/T	SNP	SGRP	-	-	missense variant	T/K	24	0.85	YLL046C_mRNA

Variant IDs are links to variant tab

Protein pathogenicity predictions

Let's have a look at a specific variant. Click on the top result in the filtered table, or search for [s12-46985](#). This will open up the variation tab.

Saccharomyces cerevisiae (R64-1-1) ▾

Location: XII:32,153-49,473 Gene: YLL046C Transcript: YLL046C_mRNA Variant: s12-46985

Variant tab

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

s12-46985 SNP

Most severe consequence: missense variant | See all predicted consequences

Alleles: A/G | Highest population MAF: 0.08

Location: Chromosome XII:46986 (forward strand) | VCF: XII 46986 s12-46985 A G

HGVS names: This variant has 5 HGVS names - Show

Original source: Variation features from SGRP, with Ensembl identifiers | About SGRP

About this variant: This variant overlaps 2 transcripts and has 3 sample genotypes.

Explore this variant

Genomic context Genes and regulation Flanking sequence Population genetics Phenotype

Sample genotypes Linkage disequilibrium Phylogenetic context Citations

VCF format

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. 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										Filter		
Gene	Transcript (strand)	Allele (Tr. allele)	Consequence Type	Position in transcript	Position in CDS	Position in protein	AA	Codons	SIFT	Detail		
YLL046C	YLL046C_mRNA (-) biotype: protein_coding	G (C)	missense variant	478 (out of 750) <div><div></div></div>	478 (out of 750) <div><div></div></div>	160 (out of 249) <div><div></div></div>	F/L	TTT/CTT	<div><div>0</div></div>	Show		
YLL047W	YLL047W_mRNA (+) biotype: protein_coding	G (G)	synonymous variant	315 (out of 384) <div><div></div></div>	315 (out of 384) <div><div></div></div>	105 (out of 127) <div><div></div></div>	K	AAA/AAG	-	Show		

This variant overlaps two genes. It causes a change in the protein sequence (missense variant) in the YLL046C gene we were looking at, but it is a synonymous variant in the other gene, not causing a change in the amino acid. Only missense variants have SIFT scores, hence why this is missing for the other gene.

(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.

Show/hide columns		Filter	
Population	Allele: frequency (count)	Genotype: frequency (count)	
SGRP	A: 0.923 (39) G: 0.077 (39)	A: 0.923 (39) G: 0.077 (39)	

Additional Exercise – Variation data in *Fusarium oxysporum*

(a) Select the *Fusarium oxysporum* 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?

Show/hide columns		Filter			
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?

Show/hide columns		Filter		
Sample (Male/Female/Unknown)	Genotype (forward strand)	Population(s)	Father	Mother
886599 (U)	C/C	melonis	-	-
889404 (U)	C/C	melonis	-	-
889405 (U)	C/C	melonis	-	-
889406 (U)	C/C	melonis	-	-
889407 (U)	C/C	melonis	-	-
889408 (U)	C/C	melonis	-	-
889410 (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
```

Variant Effect Predictor ?

The screenshot shows the Ensembl VEP web interface. At the top is a 'New job' button. Below it, the 'Species:' field is set to 'Verticillium dahliae ...' with a dropdown arrow and a link to 'Add/remove species'. A callout bubble 'Change the species' points to this field. The 'Name for this job (optional):' field contains 'Exercise 1', with a callout bubble 'Name your job' pointing to it. The 'Input data:' section has a text area with the variant data. A callout bubble 'Paste or type in data...' points to this text area. Below the text area are options to '...or upload a data file...' (with a 'Choose File' button and 'no file selected' text) and '...or provide a link to a file hosted online' (with a text input field). A callout bubble 'See data format examples' points to the 'Examples:' link. A 'Run instant VEP for current line >' button is next to the text area, with a callout bubble 'See preview of the results for a selected line' pointing to it.

Species: Verticillium dahliae ... [Add/remove species](#)

Name for this job (optional):

Input data:

Either paste data:

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

[Examples: Ensembl default, VCF, Variant identifiers, HGVS notations](#)

no file selected

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*2 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: [tmp_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:

☐

HGVS:

☐

Which identifiers do you want in the output?

Variants and frequency data Co-located variants and frequency data

Variants and frequency data

Find co-located known variants:

Yes

☒

Variant synonyms:

☐

Include flagged variants:

☐

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):

Protein annotation

Protein domains: ☐

Predictions ▾ *Variant predictions, e.g. SIFT, PolyPhen*

Pathogenicity predictions

SIFT:

Filtering options ▾ *Pre-filter results by frequency or consequence type*

Filters

Return results for variants in coding regions only: ☐

Restrict results:

NB: Restricting results may exclude biologically important data!

Advanced options ▴ *Additional enhancements*

Run >

Run VEP

Hover over the options to see definitions. When you've selected everything you need, scroll right to the bottom and click **Run**.

Recent jobs ▾

This will count down and refresh the page every 10 seconds

Analysis	Jobs	Submitted at
Variant Effect Predictor	VEP analysis of pasted data in Verticillium_dahliaejr2 Done View results	23/04/2021, 16:23 (BST)

Click here to view results

Buttons to save, edit, share or delete the job

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.

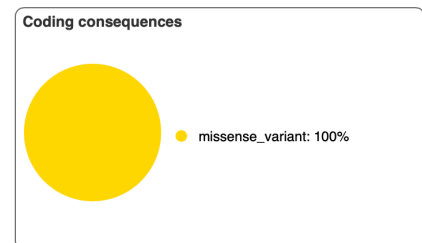
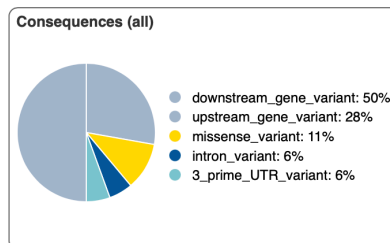
Click **View results** once your job is done. In your results you will see a graphical summary of your data, as well as a table of your results.

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?

Summary statistics

Category	Count
Variants processed	4
Variants filtered out	0
Novel / existing variants	3 (75.0) / 1 (25.0)
Overlapped genes	4
Overlapped transcripts	5
Overlapped regulatory features	-



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.

The screenshot shows the Ensembl Filters interface with the following components and annotations:

- Filters** (Search icon)
- Filter by column:** choose column name from the list (Annotation pointing to the dropdown menu)
- Logical operator** (Annotation pointing to the dropdown menu)
- Start typing to see all matching options** (Annotation pointing to the search input box)
- Apply your filter** (Annotation pointing to the Add button)
- Applied filter** (Annotation pointing to the filter bar)
- Edit filter** (Annotation pointing to the edit icon)
- Remove filter** (Annotation pointing to the remove icon)
- Add another filter** (Annotation pointing to the Add button)

The filter bar shows: **Consequence is missense_variant**

The dropdown menu shows: **Consequence** (selected), **is** (selected), **mi** (typed), **Add** (button)

The dropdown menu options are: **is** (checked), **is not**, **matches**, **<**, **>**, **<=**, **>=**, **in file**

The dropdown menu options are: **mature_miRNA_variant**, **missense_variant**, **incomplete_terminal_codon_variant**

The filter bar shows: **Uploaded variant** (selected), **is** (selected), **defined** (typed), **Add** (button)

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.

Navigation (per variant)

Filters

Download

New job

Show: 1 All variants

Uploaded variant is defined Add

All: VCF VEP TXT

Show/hide columns (10 hidden)

Uploaded variant	Location	Allele	Consequence	Gene	Feature type	Feature	Biotype	Exon	cDNA position	CDS position	Protein position	Amino acids	Codons	Existing variant	Feature strand
Variant 1	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
Variant 2	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
Variant 3	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/ACC	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/ACC	1
	5_700313_G/A	5:700313-700313	A	downstream_gene_variant	VDAG_JR2_Chr5g02171a	Transcript	VDAG_JR2_Chr5g02171a-00001.	protein_coding	-	-	-	-	-	-	-1
Variant 4	5_701484_C/A	5:701484-701484	A	downstream_gene_variant	VDAG_JR2_Chr5g02160a	Transcript	VDAG_JR2_Chr5g02160a-00001.	protein_coding	-	-	-	-	-	tmp_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	-	-	-	-	-	tmp_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	-	-	-	-	-	tmp_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	-	-	-	-	-	tmp_5_701484_C_A	-1

Filter this table

Download options

Show additional columns

Existing variants

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

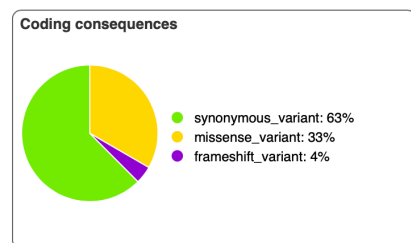
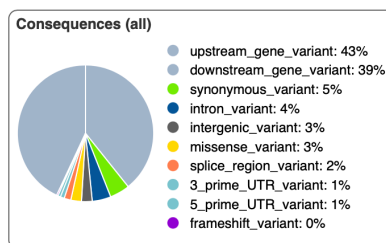
On the course file page, you will find a VCF file labelled VEP_exercise.vcf. This is a small subset of the outcome of *Puccinia graminis* Ug99 whole genome sequencing and variant calling experiment. This file can also be found on our FTP site under the following link:
http://ftp.ebi.ac.uk/pub/databases/ensembl/training/2021/FungalPathogens/VEP_exercise.vcf

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).

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

Summary statistics

Category	Count
Variants processed	100
Variants filtered out	0
Novel / existing variants	0 (0.0) / 100 (100.0)
Overlapped genes	258
Overlapped transcripts	303
Overlapped regulatory features	-



- (c) 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)													
Show: 1 5 10 50 All variants													
Consequence is missense_variant													
All: VCF VEP TXT													
Filtered: VCF 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_3.1594.801-801	T	missense_variant	GMQ_27112	3/3	GMQ_27112T0.c.265G>A	GMQ_27112T0.p.Glu89Lys	265	265	89	E/K	GAG/AAG	tmp_Supercontig_3.1594.801_C_T	Pfam:PF14303 PANTHER:PTHR45125 PANTHER:PTHR45125.SF3 MobiDB_lite.mobiDB-lite
Supercontig_3.156.127654-127654	C	missense_variant	GMQ_21813	2/6	GMQ_21813T0.c.235T>C	GMQ_21813T0.p.Ser79Pro	235	235	79	S/P	TCC/CCC	tmp_Supercontig_3.156.127654_T_C	Gene3D:2.102.10.10 PROSITE_profiles:PS51296 Superfamily:SSF50022
Supercontig_3.58.69935-69935	C	missense_variant	GMQ_20457	4/4	GMQ_20457T0.c.1003G>C	GMQ_20457T0.p.Asp335His	1003	1003	335	D/H	GAT/CAT	tmp_Supercontig_3.58.69935_G_C	Gene3D:3.40.720.10 PANTHER:PTHR23071 Superfamily:SSF3649 CDD:cd16023
Supercontig_3.48.118082-118082	T	missense_variant	GMQ_20311	1/4	GMQ_20311T0.c.73G>A	GMQ_20311T0.p.Glu25Lys	73	73	25	E/K	GAA/AAA	tmp_Supercontig_3.48.118082_C_T	-
Supercontig_3.41.7765-7765	G	missense_variant	GMQ_06767	6/15	GMQ_06767T0.c.1328A>C	GMQ_06767T0.p.Gln443Pro	1328	1328	443	Q/P	CAG/CCG	tmp_Supercontig_3.41.7765_T_G	PANTHER:PTHR46896 PANTHER:PTHR46896.SF3
Supercontig_3.16.171261-171261	T	missense_variant	GMQ_04080	1/2	GMQ_04080T0.c.287G>A	GMQ_04080T0.p.Gly96Glu	287	287	96	G/E	GGA/GAA	tmp_Supercontig_3.16.171261_C_T	-
Supercontig_3.73.160474-160474	G	missense_variant	GMQ_03045	2/3	GMQ_03045T0.c.407G>C	GMQ_03045T0.p.Arg136Thr	407	407	136	R/T	AGA/ACA	tmp_Supercontig_3.73.160474_C_G	Low_complexity_(Seg):seg PANTHER:PTHR31595.SF1 PANTHER:PTHR31595.SF1
Supercontig_3.427.55213-55213	A	missense_variant	GMQ_02814	2/2	GMQ_02814T0.c.99G>T	GMQ_02814T0.p.Gln33His	358	99	33	Q/H	CAG/CAT	tmp_Supercontig_3.427.55213_C_A	PANTHER:PTHR31361 PANTHER:PTHR31361.SF15 MobiDB_lite.mobiDB-lite

- (d) What are the HGVSp notations of missense variants falling in known protein domains?

Results preview

Navigation (per variant)

Show: 1 5 10 50 All variants

Filters

Domains is defined

Consequence is missense_variant

Clear filters

Match all of the above rules

Updated

Uploaded variant is defined

Add

Download

All: VCF VEP TXT

Filtered: VCF VEP TXT

New job

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_3.1594.801-801	T	missense_variant	GMQ_27112	3/3	GMQ_27112T0.c.265G>A	GMQ_27112T0.p.Glu89Lys	265	265	89	E/K	GAG/AAG	tmp_Supercontig_3.1594.801_C_T	Pfam:PF14303 PANTHER:PTHR45125:SF3 MobiDB_lite:mobidb-lite
Supercontig_3.156.127654-127654	C	missense_variant	GMQ_21813	2/6	GMQ_21813T0.c.235T>C	GMQ_21813T0.p.Ser79Pro	235	235	79	S/P	TCC/CCC	tmp_Supercontig_3.156.127654_T_C	Gene3D:2.102.10.10 PROSITE_profiles:PS1296 Superfamily:SSF50022
Supercontig_3.58.69935-69935	C	missense_variant	GMQ_20457	4/4	GMQ_20457T0.c.1003G>C	GMQ_20457T0.p.Asp335His	1003	1003	335	D/H	GAT/CAT	tmp_Supercontig_3.58.69935_G_C	Gene3D:3.40.720.10 PANTHER:PTHR23071 Superfamily:SSF53649 CDD:cd16023
Supercontig_3.41.7765-7765	G	missense_variant	GMQ_06767	6/15	GMQ_06767T0.c.1328A>C	GMQ_06767T0.p.Gln443Pro	1328	1328	443	Q/P	CAG/CGG	tmp_Supercontig_3.41.7765_T_G	PANTHER:PTHR46896 PANTHER:PTHR46896:SF3
Supercontig_3.73.160474-160474	G	missense_variant	GMQ_03045	2/3	GMQ_03045T0.c.407G>C	GMQ_03045T0.p.Arg136Thr	407	407	136	R/T	AGA/ACA	tmp_Supercontig_3.73.160474_C_G	Low_complexity_(Seg) PANTHER:PTHR31595 PANTHER:PTHR31595:SF1
Supercontig_3.427.55213-55213	A	missense_variant	GMQ_02814	2/2	GMQ_02814T0.c.99G>T	GMQ_02814T0.p.Gln33His	358	99	33	Q/H	CAG/CAT	tmp_Supercontig_3.427.55213_C_A	PANTHER:PTHR31361 PANTHER:PTHR31361:SF15 MobiDB_lite:mobidb-lite

(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)

Show: 1 5 10 50 All variants

Filters

Consequence is frameshift

Uploaded variant is defined

Add

Download

All: VCF VEP TXT

Filtered: VCF VEP TXT

New job

Show/hide columns (22 hidden)

Location	Allele	Consequence	Gene	Exon	HGVSc	HGVSp	cDNA position	CDS position	Protein position	Amino acids	Codons	Existing variant	UNIPARC
Supercontig_3.1482.1095-1097	G	frameshift_variant	GMQ_27001	1/3	GMQ_27001T0.c.221del	GMQ_27001T0.p.Pro74ArgfsTer27	220-221	220-221	74	P/X	CCG/CG	tmp_Supercontig_3.1482.1095_CGG_CG	UPI0004E9C5AE