

Demo: Ensembl Fungi species

Clickable links shown in [blue](#), text to be entered shown in [red](#).

Navigate to fungi.ensembl.org. You'll see a homepage similar to this:

1: Website header - Points to the Ensembl Fungi logo and navigation menu.

1a: Log in to Ensembl - Points to the Login/Register button in the top right.

2: Main search box - Points to the search input field with the placeholder text "All species".

1b: Quick search - Points to the search button and the search results area.

3: Genome and species directory - Points to the "All genomes" and "Favourite genomes" sections.

4: Release details - Points to the "What's New in Release 42" section.

Click on '[View full list of all Ensembl Fungi species](#)', which you can find in section 3: **Genome and species directory** shown above.

Download table - Points to the download icon in the top right of the table.

Type in your species to search the table - Points to the search input field above the table.

Name	Classification	Taxon ID	Assembly	Accession	Variation database	Regulation database	Whole genome alignments	Other alignments	In peptide compara	In pan-taxonomic compara
Absidia glauca	Mucoromycotina	4829	AG_v1	GCA_900079185.1	-	-	-	-	-	-
Absidia repens str. NRRL 1336	Mucoromycotina	90262	Absrep1	GCA_002106175.1	-	-	-	-	✓	-
Acaromyces ingoldii str. MCA 4198 (GCA_003144295)		215250	Acain1	GCA_003144295.1	-	-	-	-	✓	-
Acidomyces richmondensis (GCA_001572075)	Acidomyces	245562	ASM157207v1	GCA_001572075.1	-	-	-	-	-	-
Acidomyces richmondensis BFW (GCA_001592465)	Acidomyces	768039	Acidomyces_richmondensis_BFW_v1.0	GCA_001592465.1	-	-	-	-	✓	-
Acremonium chrysogenum ATCC 11550 (GCA_000769265)	Hypocreales	657340	ASM76926v1	GCA_000769265.1	-	-	-	-	✓	-
Agaricus bisporus var. burnettii JB137-S8 (GCA_000300555)	Agaricales	597362	Agabi_varbur_1	GCA_000300555.1	-	-	-	-	-	-

Data types available for this assembly - Points to the table headers.

Click on the latin name of your species of interest to go to the species homepage. We'll click on [Ashbya gossypii](#).

Search

Search Ashbya gossypii...

e.g. AGOS_AER342C or V:1263836-1265731 or synthetase

Search for a feature in this genome

About the Ashbya gossypii genome

More than 90% of Ashbya gossypii genes show both homology and a particular pattern of synteny with Saccharomyces cerevisiae. It was originally isolated from cotton as a pathogen causing stigmatomycosis by Ashby and Novell in 1926. The Ashbya gossypii genome project was initiated when conservation of gene order and orientation (synteny) to Saccharomyces cerevisiae was noted.

Ashbya gossypii became recognized as an attractive model to study the growth of long and multinucleate fungal cells (hyphae) because of its small genome, haploid nuclei, and efficient gene targeting methods. It is generally assumed that a better understanding of filamentous fungal growth will greatly stimulate the development of novel fungicides.

Taxonomy ID [284811](#)

Data source [Ashbya Genome Database](#)

[More information and statistics](#)

About this species

Links to example features in Ensembl

Genome assembly: [ASM9102v1](#)

[More information and statistics](#)

[Download DNA sequence \(FASTA\)](#)

[Display your data in Ensembl Fungi](#)



View karyotype



Example region

About this genome assembly

Gene annotation

What can I find? Protein-coding and non-coding genes, splice variants, cDNA and protein sequences, non-coding RNAs.

[More about this genebuild](#)

[Download genes, cDNAs, ncRNA, proteins - FASTA - GFF3](#)

[Update your old Ensembl IDs](#)



Example gene



Example transcript

To find out more about the genome assembly and gene annotation, click on [More information and statistics](#).

Ashbya gossypii Assembly and Gene Annotation

About the Ashbya gossypii genome

More than 90% of Ashbya gossypii genes show both homology and a particular pattern of synteny with Saccharomyces cerevisiae. It was originally isolated from cotton as a pathogen causing stigmatomycosis by Ashby and Novell in 1926. The Ashbya gossypii genome project was initiated when conservation of gene order and orientation (synteny) to Saccharomyces cerevisiae was noted.

Who did the gene annotation?

recognized as an attractive model to study the growth of long (hyphae) because of its small genome, haploid nuclei, methods. It is generally assumed that a better fungal growth will greatly stimulate the development of

Annotation

Annotation of Ashbya gossypii was imported from the [European Nucleotide Archive](#). Non coding RNA genes have been annotated using tRNAScan-SE (Lowe, T.M. and Eddy, S.R. 1997), Rfam (Griffiths-Jones et al 2005), and RNAmmer (Lagesen K., et al 2007); additional analysis tools have also been applied.

Other Data

Probe mapping data has been loaded for the experiment [A-AFFY-105](#).

References

1. [The Ashbya gossypii genome as a tool for mapping the ancient Saccharomyces cerevisiae genome](#). Dietrich FS, Voegelé S, Brachat S, Lerch A, Gates K, Steiner S, Morin R, Luedi P, Choi S et al. 2004. Science. 304:304-307.

More information

General information about this species can be found in [Wikipedia](#).

Statistics

Summary

Assembly	ASM9102v1 (Ashbya gossypii ATCC 10895 assembly from BioProject 13834), INSDC Assembly GCA_000091025.3, Oct 2010
Database version	95.1
Base Pairs	9,119,312
Golden Path Length	9,119,312
Genebuild by	AGD
Genebuild method	Generated from ENA and UniProtKB annotation
Data source	Ashbya Genome Database

Interesting statistics about the assembly and annotation

Gene counts

Coding genes	4,776
Non coding genes	725
Small non coding genes	725
Pseudogenes	19
Gene transcripts	5,520

A transcript is the operational unit of a gene. In a genomic context, transcripts consist of one or more exons, with adjoining exons being separated by introns. The exons/introns are transcribed and then the introns spliced out. Transcripts may or may not encode a protein

Terms underlined have mouse-over definitions

Exercises: Searching Ensembl Fungi species

Exercise – *Ustilago maydis*

(a) Navigate to the species homepage for *Ustilago maydis*. What is the name of the genome assembly for *Ustilago maydis*?

(b) Click on [More information and statistics](#). How long is the *Ustilago maydis* genome (in bp)? How many genes have been annotated?

Exercise - *Bipolaris* species

(a) How many genome assemblies are there for the genus *Bipolaris* in Ensembl Fungi?

(b) What is the INSDC accession number for *Bipolaris oryzae*? What institute submitted the data to INSDC?

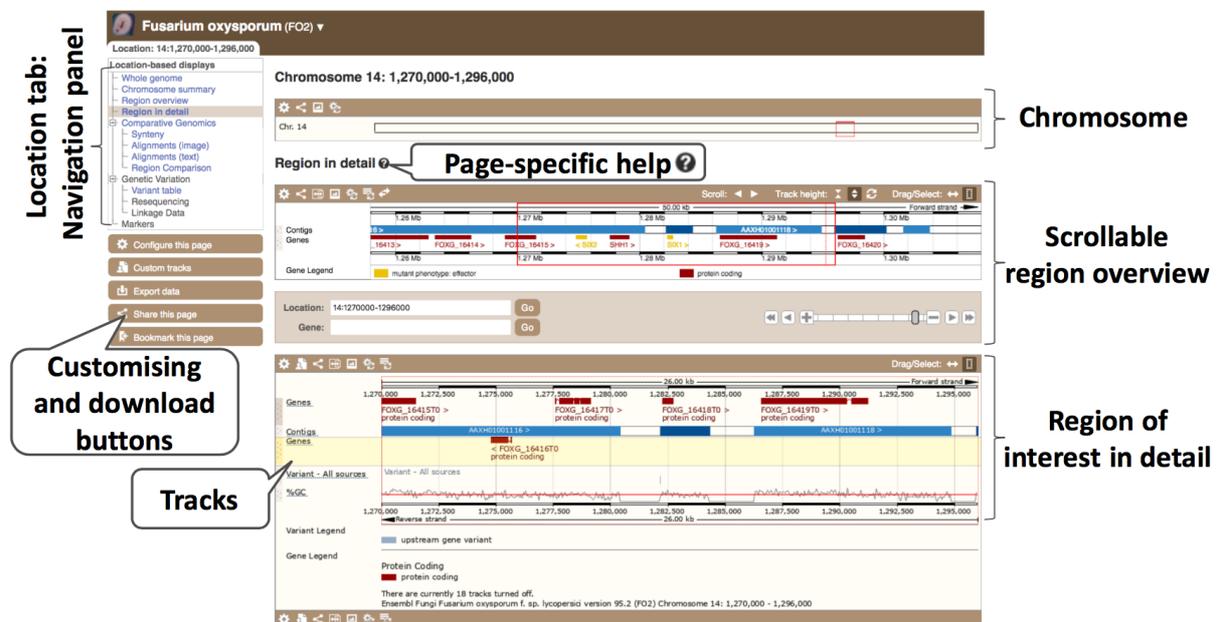
Demo: The Ensembl Fungi Region in detail view

Start at the Ensembl Fungi front page, fungi.ensembl.org. You can search for a region by typing it into a search box, but you have to specify the species.

Find *Fusarium oxysporum*, then type (or copy and paste) **14:1270000-1296000** into the search box. Press enter or click **Go** to jump directly to the **Region in detail** Page.



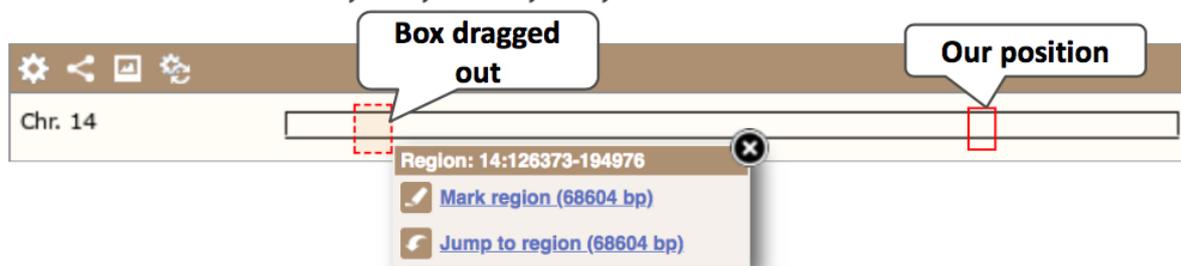
Click on the button  to view page-specific help. The help pages provide links to Frequently Asked Questions, a Glossary, Video Tutorials, and a form to Contact HelpDesk. There is a help video on this page at <http://youtu.be/tTKEvgPUq94>.



The Region in detail page is made up of three images, let's look at each one on detail.

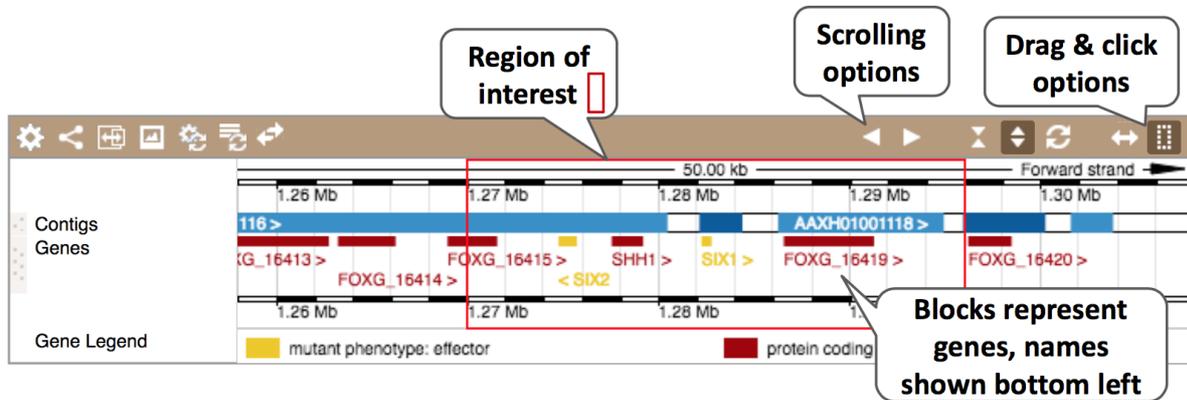
The first image shows the chromosome. You can jump to a different region by dragging out a box in this image. Drag out a box on the chromosome; a pop-up menu will appear.

Chromosome 14: 1,270,000-1,296,000

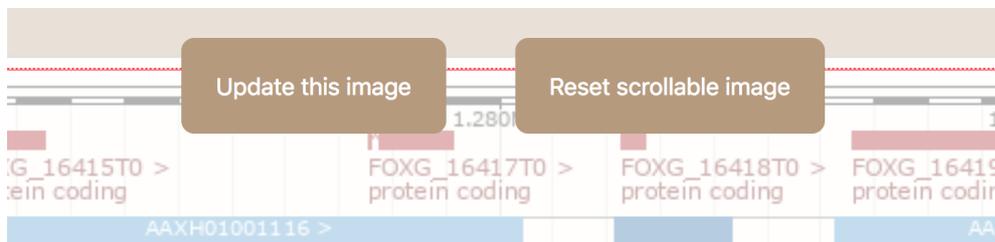


If you would like to move to the region, you could click on [Jump to region \(### bp\)](#). To highlight it, click on [Mark region \(### bp\)](#). For now, we'll close the pop-up by clicking on the [X](#) on the corner.

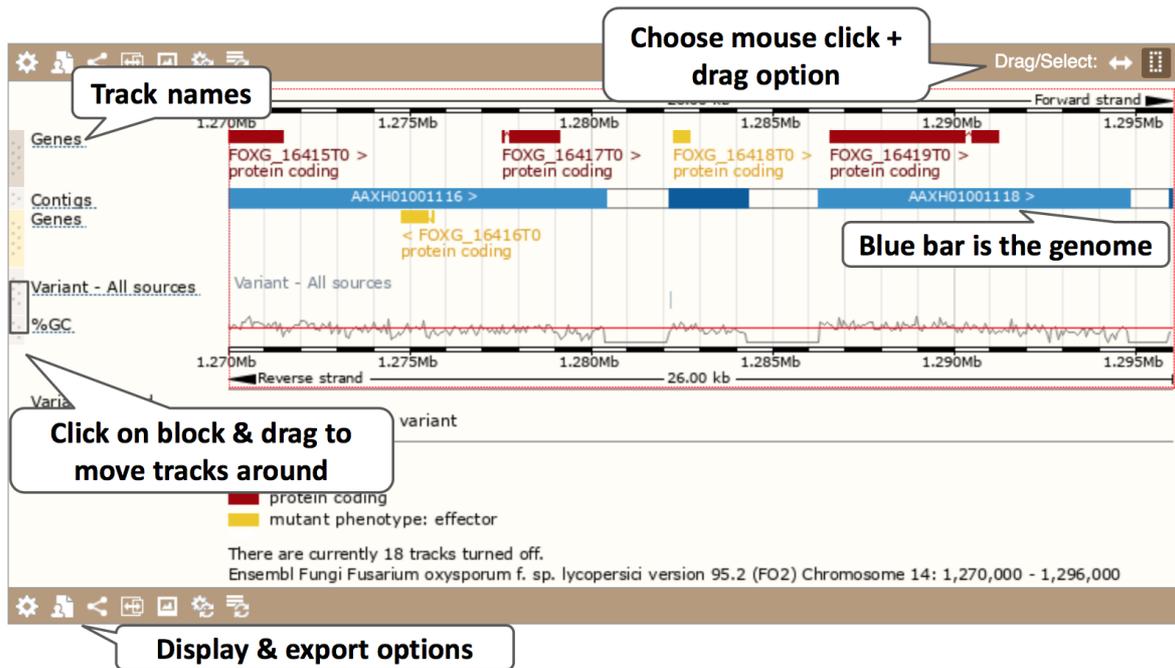
The second image shows a 50 kb region around our selected region. This view allows you to scroll back and forth along the chromosome.



Click on the [Drag/Select button](#) **Drag/Select:**  to change the action of your mouse click. Now you can scroll along the chromosome by clicking and dragging within the image. As you do this you'll see the image below grey out and two blue buttons appear. Clicking on [Update this image](#) would jump the lower image to the region central to the scrollable image. We want to go back to where we started, so we'll click on [Reset scrollable image](#).

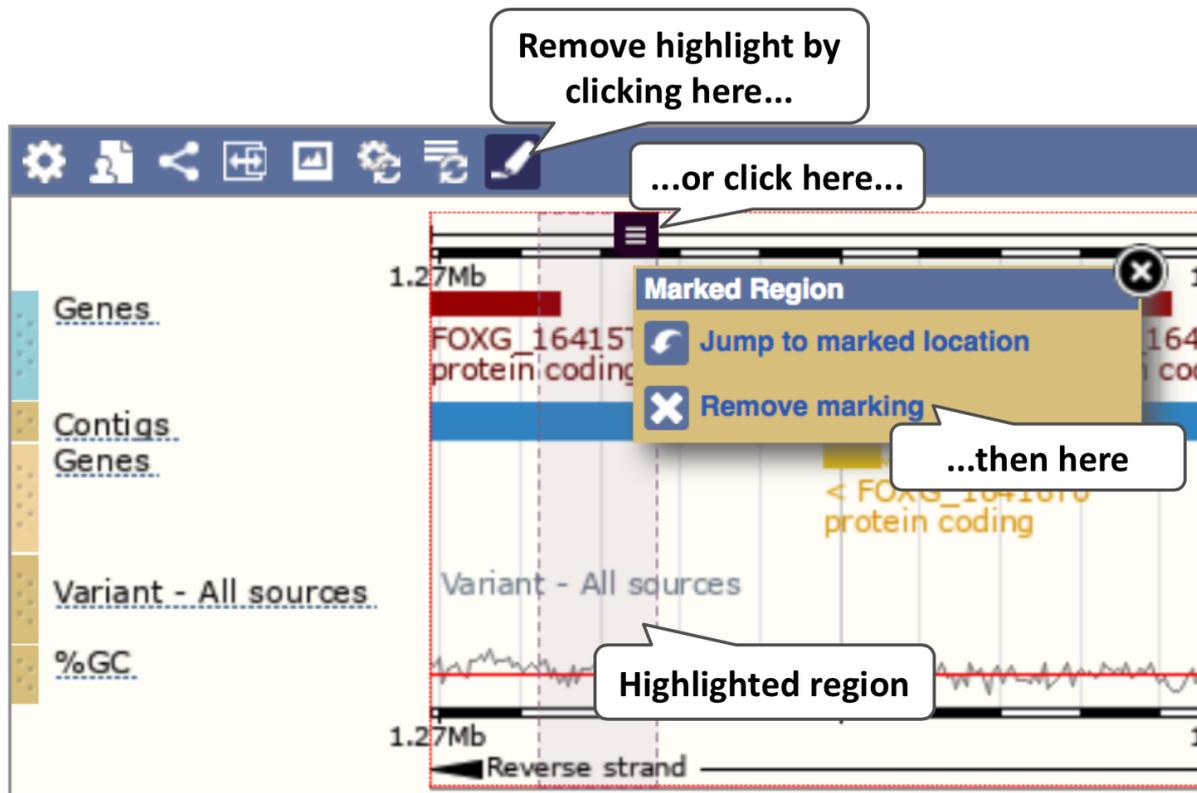


The third image is a detailed, configurable view of the region.



Click on the [Drag/Select](#) option at the top or bottom right to switch mouse action. On [Drag](#), you can click and drag left or right to move along the genome, the page will reload when you drop the mouse button. On [Select](#) you can drag out a box to highlight or zoom in on a region of interest.

With the tool set to [Select](#), drag out a box around an exon and choose [Mark region](#).



The highlight will remain in place if you zoom in and out or move around the region. This allows you to keep track of regions or features of interest.

We can edit what we see on this page by clicking on the [Configure this page](#) menu at the left.



This will open a menu that allows you to change the image. You can put some tracks on in different styles; more details are in this FAQ: <http://www.ensembl.org/Help/Faq?id=335>.

The screenshot shows the 'Configure Region Image' interface. On the left is a sidebar with track categories: Active tracks, Favourite tracks, Track order, Search results, Sequence and assembly (2/4), Genes and transcripts (1/1), mRNA and protein alignments (0/2), Comparative genomics (0/3), Variation (1/2), Repeats (0/10), and Information (0/11). Below these are buttons for Custom tracks, Manage configurations, Reset configuration, and Reset track order. The main area shows 'Active tracks' under 'Sequence and assembly' with tracks like Contigs, Scaffolds, Genes, and Variants. A 'Change track style' menu is open for the 'Genes' track, showing options like 'Off', 'No exon structure without labels', 'Expanded without labels', 'Collapsed without labels', and 'Coding transcripts only (in coding genes)'. Callouts point to: 'Search for a track' (pointing to the search bar), 'Click on / for track info' (pointing to the info icon), 'Click box to turn on/off & change style' (pointing to the track style menu), and 'Track categories' (pointing to the sidebar).

Let's add some tracks to this image. Add:

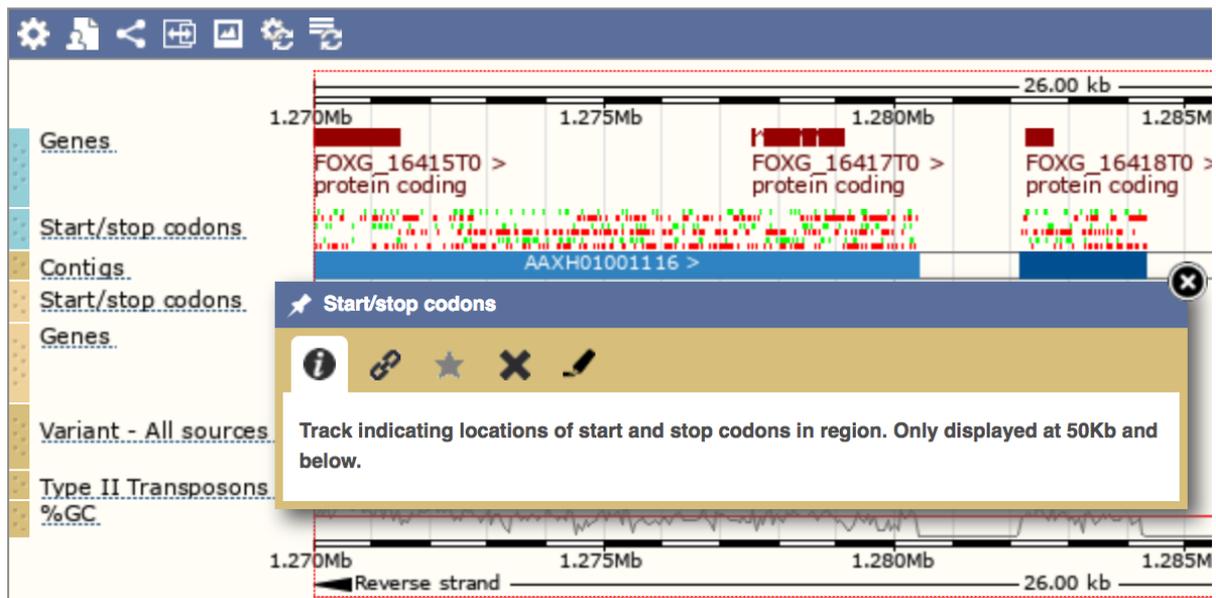
Start/stop codons

Type II Transposons

Now click on the tick in the top right hand tick to save and close the menu. Alternatively, click anywhere outside of the menu. We can now see the tracks in the image.

We can also change the way the tracks appear by clicking on the track name to open a menu. We can move tracks around by clicking and dragging on the coloured dotted block/bar to the left of the track name.

E.g.



Now that you've got the view how you want it, you might like to show something you've found to a colleague or collaborator. Click on the [Share this page](#) button to generate a link.



Email the link to someone else, so that they can see the same view as you, including all the tracks you've added. These links contain the Ensembl release number, so if a new release or even assembly comes out, your link will just take you to the archive site for the release it was made on.

To return this to the default view, go to [Configure this page](#) and select [Reset configuration](#) at the bottom of the menu.

Exercises: Ensembl Fungi Region in Detail

Region Exercise 1 – Exploring a *Coprinopsis cinerea okayama* region

- Go to the region 7:1400000-1425000 in *Coprinopsis cinerea okayama*.
- How many complete genes are found in this region? How many on the forward and how many on the reverse strand?
- Zoom in on the largest gene [EFI27358](#). How many exons does this gene have?
- Export the genomic sequence for this region.

Region Exercise 2 – Exploring a region in *Zymoseptoria tritici*

- (a) Go to the region [3:1310171-1318171](#) in *Zymoseptoria tritici* MG2.
- (b) How many genes are annotated in this region? Are they all annotated by the same institute?
- (c) Turn on the all repeat regions track. Are there any repeat regions identified in this region? Do they overlap any of the genes?

Region Exercise 3 - Exploring a region in *Schizosaccharomyces pombe*

- (a) We're exploring the region surrounding the gene Tor1. Search *Schizosaccharomyces pombe* for the region [II:3075647-3085541](#).
- (b) Turn on all the tracks for Polyadenylation sites. Which studies show data for the gene Tor1? Click on the track names to see descriptions.
- (c) Zoom into the 3' UTR for tor1 to see the peaks. What do you notice about the y-axis scales on the tracks? Can you change the Mata (2013) tracks to have the same Y-axis limits at the Schlackow (2013) tracks? (Hint: Click on the track name and explore the options in the pop-up).

Demo: The Ensembl Fungi gene tab

We're going to look at the gene *ATG8* in *Magnaporthe oryzae*. This gene is involved in autophagy, and targeted silencing of this gene inhibits infection (further info in Wilson and Talbot, *Nature Reviews Microbiology* volume 7, pages 185–195 (2009)).

From fungi.ensembl.org, type *ATG8* into the Search for a gene search bar, click the drop-down menu and select *Magnaporthe oryzae* and click the Go button.

Click on the gene name *ATG8* in the results. The **Gene tab** should open:

The screenshot displays the Ensembl Fungi Gene tab for *ATG8* in *Magnaporthe oryzae*. The page is divided into several sections:

- Gene-based displays:** A navigation panel on the left lists various tracks and data types, including Summary, Splice variants, Gene alleles, Sequence, Secondary Structure, Gene families, Literature, Fungal Compara, Genomic alignments, Gene tree, Gene gain/loss tree, Orthologues, Paralogs, Pan-taxonomic Compara, Gene Tree, Orthologues, Ontologies, GO: Molecular function, GO: Biological process, GO: Cellular component, Pfam: Pfam identifier, Phenotypes, Genetic Variation, Variant table, Variant image, Structural variants, Gene expression, Pathway, Regulation, External references, Supporting evidence, ID History, and Gene history.
- Gene: ATG8 MGG_01062:** Core information shown on all gene pages, including Description (Autophagy-related protein 8), Location (Chromosome 5: 641,522-643,203 reverse strand), and About this gene (This gene has 1 transcript (splice variant), 344 orthologues and is a member of 2 Ensembl protein families).
- Transcript table:** A table showing transcript details. The table has columns for Name, Transcript ID, bp, Protein, Biotype, UniProt, and Flags. The transcript MGG_01062T0 is highlighted in green.
- Summary:** A section providing additional details about the gene, such as Name (ATG8 (UniProtKB Gene Name)), UniProtKB (This gene has proteins that correspond to the following UniProtKB identifiers: Q51MW4), Gene type (Protein coding), and Annotation method (Protein coding genes annotation from the Broad Institute).
- Simple browser view, gene highlighted:** A genomic browser view showing the gene's location on the chromosome. The gene *ATG8* is highlighted in green, and the surrounding genomic context is shown with various tracks and annotations.

The *ATG8* gene is highlighted in green and in the centre of the display as it is the gene of interest.

Let's walk through some of the links in the left hand navigation column. How can we view the genomic sequence? Click [Sequence](#) at the left of the page.

Marked-up sequence




Exons ATG8 exons All exons in this region

Markup loaded

Download/BLAST whole gene sequence

```

>chromosome:MG8:5:640922:643803:-1
AATGATGTCCATCCCAGTGGCCTTGTCTCCCCTTGATGAATCAATGAATCATCTGGCTTT
GACCCTCTACTCCGTAGATGAGTCGCATACAGTACGGCGTAAATGCATAGGCTGGATGGA
ACGGAGAGAAATAACAAAGGCAGTAATTTGTAACAGCTAGACAGGTACCTTGAACAGTT
AGTTATGCGCAGCTGTTGATTCCCCACACCCCTTACTCTGATGCCTAGATCCGAATG
ACTGAATGTTCCACGTGAATGGATCTTTGTGAATCATTTCACTCGACGCCTTGTCCCTGC
CCCCACCAACAGTGAGAAATCCTGAGAGGAGAAAGGCCGCAAGCTTGATCCGATTGTA
ACCATCTGCCCAATCAACTCAAAGGTGTGCCACAAGGCGAACGGTCGGTCGGTCGCATG
CTGCAGTTAAAGTGAAGCTTGTAAAAGGTACTCACACCAGATAGTACTGATGGACCTAA
TGTGAAATGTGGAGCTCATCGTCTGGGGAAAAAGGCCAGGAATTCATGTCATTTGATC
CCGACGCACTGTAGCTGTGCTTTCGCCCGCTTACAGCATCGGGCCCTCTCCAACC
CGTTTTTGGTTGACCAGTCCAGTCCCTTGTTCCCGCTTCTCAAGGCGGCAACCACCT
GAGGAGCTACGGTGCCTGCAGTCCACAGCCGCCATTTCAAGGTCTACGCAATCCCTCT
TCCTTCCATTACCCCTCCACCCAGAACCTAATCTGTACTCCAACCTTTCATCATCAT
CATCCTGAACTCGACAAGCAACAACAAGTAGCCTGCAAAATAATTATCACTTTCCC
TTTGCTGTTTCCCGAACAGCAAGGACGCATCAAT
AAACAAGTCTCAATCAAGTACGCTCCAAGT
GACGCATCAAT
BLAST selected sequence
GACGCATCAAT
GCGCTCCAAGT
GATCCCCTTAG
GACGTTGATC
GTCATTTGCGA
GTTCCGGCCGA
TCTCCCGAGAA
CATGTCGAGCAT
CTACGAGTTACACAAGGACGAGGACGGGTGAGTCACTCATACCCTTTTACAAGCAGAAT
ACTTGTGTTACGGAGTAATAGACGGCTGACTCAAATACCTCTGTCCCAACAATAAGAT
TCCGTGATACACTACTCCGGCGAGAACACCTTCGGCGACCTGTTGAGGAAGTCGAGT
GACCTTCTCTACCCCGGGGGGCAAAATGCTTCTCTGACGGGACTGCAATGACGAAACCA
    
```

Upstream sequence (600bp)

1st exon of ATG8 gene

Highlight portion of sequence to BLAST

The sequence is shown in FASTA format. Take a look at the FASTA header:

Genome assembly Base pair start coordinate Reverse strand (1 is forward)

```

>chromosome:MG8:5:640922:643803:-1
AATGATGTCCATCCCAGTGGCCTTGTCTCCCCTTGATGA
GACCCTCTACTCCGTAGATGAGTCGCATACAGTACGGCGTAAATGCATAGGCTGGATGGA
ACGGAGAGAAATAACAAAGGCAGTAATTTGTAACAGCT
    
```

Chromosome Base pair end coordinate

Exons are highlighted within the genomic sequence. If you click on [Configure this page](#) you can change display options, and for species with variation databases you can highlight variants on this view.

You can download this sequence by clicking in the  button above the sequence.

This will open a dialogue box that allows you to pick between plain [FASTA sequence](#), or [sequence in RTF](#), which includes all the coloured annotations and can be opened in a word processor. This button is available for all sequence views.

File name:	Magnaporthe_oryzae_ATG8_sequence
File format:	-- Choose Format -- ✓ FASTA RTF (Word-compatible)
<hr/>	
  	

Settings

Sequences to export:

- Select/deselect all**
- cDNA (transcripts)
- Coding sequences (CDS)
- Amino acid sequences
- 5' UTRs
- 3' UTRs
- Exons
- Introns
- Genomic sequence

5' Flanking sequence (upstream):

600 * (Maximum of 1000000)

3' Flanking sequence (downstream):

600 * (Maximum of 1000000)

If we are interested in finding out about gene functions, the Gene Ontology (GO) annotations can tell us where the protein is located, the biological processes it is involved in and it's molecular function.

Click on [GO: Biological process](#). This page shows all linked GO annotations, some of these are linked as GO terms are hierarchical. For example if you click on 'Positive regulation of macroautophagy' you will be taken to the GO pages, which shows that this is a child term to 'Autophagy' which is also shown on the GO pages in Ensembl.

GO: Biological process

Accession	Term	Evidence	Annotation source	Mapped using	Transcript IDs
GO:0006914	autophagy	IEA	UniProtKB/Swiss-Prot:ATG8_MAGO7		MGG_01062T0 <ul style="list-style-type: none"> Search BioMart View on karyotype
GO:0009405	pathogenesis	IMP	UniProtKB/Swiss-Prot:ATG8_MAGO7		MGG_01062T0 <ul style="list-style-type: none"> Search BioMart View on karyotype
Inferred from Mutant Phenotype					
GO:0015031	protein transport	IEA	UniProtKB/Swiss-Prot:ATG8_MAGO7		MGG_01062T0 <ul style="list-style-type: none"> Search BioMart View on karyotype
GO:0016239	positive regulation of macroautophagy	IMP	UniProtKB/Swiss-Prot:ATG8_MAGO7		MGG_01062T0 <ul style="list-style-type: none"> Search BioMart View on karyotype
GO:0048102	autophagic cell death	IMP	UniProtKB/Swiss-Prot:ATG8_MAGO7		MGG_01062T0 <ul style="list-style-type: none"> Search BioMart View on karyotype

For some pathogenic species in Ensembl Fungi we have Pathogen-Host Interactions (PHI-base) annotations. Click on the [PHI: Phibase identifier](#) link in the left-hand menu.

There are four results here, listed by the PHI-base ID. Click on the link to go to the PHI-base website to view more information about this annotation.

PHI: Phibase identifier

Accession	Term	Evidence	Annotation source	Mapped using	Transcript IDs
PHI:2061	2061	ND	Sequence Publications:19115483		MGG_01062T0 <ul style="list-style-type: none"> Search BioMart View on karyotype
PHI:2076	2076	ND	Sequence Publications:19747456		MGG_01062T0 <ul style="list-style-type: none"> Search BioMart
PHI:2139	2139				
PHI:768	768				

Pathogen Gene	Mutant Phenotype	Pathogen Species	Disease	Host Species
Gph1	loss of pathogenicity	Magnaporthe oryzae	Rice Blast	Hordeum vulgare (related: barley)
ATG8	loss of pathogenicity	Magnaporthe oryzae	Rice Blast	Hordeum vulgare (related: barley)

Pathogen Gene	Allele	Pathogen	Host
Gene:ATG8 PHI-base entry:PHI:2061 Gene ID:XP_368182 Protein ID: Q51MW4 Sequence strain:70-15		Pathogen species:Magnaporthe oryzae Pathogen ID: 318829 Pathogen strain:B157	Host species:Hordeum vulgare (related: barley) Host classification:Monocots Host ID: 4513 Host strains:subsp. vulgare (related: domesticated barley) Tissue:leaf

Reference	Comments	PHI Phenotype	Pathogen Phenotype
Pmid:19115483 Ref source:Pubmed Year:2009 Author reference:Yi Zhen Deng		Phenotype:loss of pathogenicity Disease name:Rice Blast Tissue:leaf Host response:slight induction of hypersensitive reaction Experimental technique:Gene deletion: full	

Pathogen Gene	Mutant Phenotype	Pathogen Species	Disease	Host Species
ATG8	loss of pathogenicity	Magnaporthe oryzae	Rice Blast	Hordeum vulgare (related: barley)

Demo: The Ensembl Fungi transcript tab

Many genes have multiple transcripts which can be seen in the transcript table. Click on

[Show transcript table](#)

We can go to the transcript tab either by clicking on the transcript ID, [MGG_01062T0](#), in the table, or on the transcript tab at the top of the page. You are now in the Transcript tab on the summary page. Some summary information about the number of exons, length etc is shown at the bottom of the page under the diagram.

The left hand navigation column provides several options for the transcript. Click on the [Exons](#) link.

Exons/ Introns [Translated sequence](#) [Flanking sequence](#) [Intron sequence](#) [UTR](#)

Markup loaded

No.	Exon / Intron	Start	End	Start Phase	End Phase	Length	Sequence
1	MGG_01062-E1	643,203	642,789	-	0	415gtagctgctggtttgccccgcttcacagcatoggggccctctgcaaacg CGTTTTGGTTGACCCAGTCCAGTCCCTTGTCCCGCTTCTCACAGGGGGG GAGGAGCTACGGTGGTGCAGTCCACAGCCGCCATTTCAGGTCTACGCA TCCTTCCATTACCCCTCCACCCAGAACCTAATCTGTACTCCAACTTTC CATCTGAACTCGACAAGCAACACACAAGTAGCCTGCAATAAATTATC TTTCGTGTTCCCGAACGCACTCCGACGATCAATATCTCACGACAA AAACAAGTCTCAATCACCCGCCCTTCGGCTCAAGTCAAGGCGGACCCCTT GAAGCGCAAGGCTGAAGCCGAGCCATTCGCCAGAAGTATACCGACCCGATTCC
2	MGG_01062-E2	642,788	642,457	0	2	218	gttagtattccctctactccgggtt.....taatacctctgctcatctacaacag GTCATTTGCGAGAAGTAGAAAAGTGGACATTGCCACCATCGACAAGAAGTACCTG GTTCCGGCCGACTGACTGTGGCCAGTTCGTCTACGTGATCCGACGCGCATCAAGCTG TCTCCGAGAAAGCCATTTTCATCTTCGTCCAGGACACCCCTGCCCGACCGCGCACTC ATGTCGAGCATCTACGAGTTACACAAGGACGAGGAGCG
3	MGG_01062-E3	642,456	641,522	2	-	844	gtgagtcactcataccoccttttaca.....aatacctctgtccccaaacaatag ATTCTGTACATCACTACTCCGGGAGAACACTTCGGCGACTGTTGAGGAAGTCGA GTGAGCTTGCTCACCCGGCCGGGCAATCTGCTGTGACGCGGAGTGGATACCGAACC CAGTCCAGCCAGAACTGAAATGGACTTTGGGAGGACTCACACCCGCTGTTTCTGTTTT TGTTCCTCTGATGGACCGTTTTTCGACTTGTGATTGGCGTTAAAGAGGGGTTTTTC TATTATTGACAAGCTGGATTGGGACTTGGGACTTCTATGTGCTGTTTTTTACTTCTTT TTTTATCTGCTGTTGTACGGCACTCAGTCAACCCCTGGGGAGCGGTGAGAGCGAGAA ACCACGGGGTTTTGATGTGGCATCTCCAATATCTCGGCTCGGATTGCCACGGTCA ATCGAATATGACGAGGGCACCCACACTGTATTTCCTGAGT GAGGGAAGGAATGGCACTATGATGGAAATAGAATTCCTTTC AGCCTTCCCGACTCTCCCGAGGACTGAGATATTGGACCTT ACACATGAGTGGGAGGAGTGAAGCTTTGTAGGAGCGGAAGCCA AATCACTGCGGTGTGCTCCCTACATATGCTACATTCATTCCTT TCCAGCATGTTCCCATGTTAAAATTGGCTACATTCGATTTT CCAATCTAGTCTAGACTGTGACTTTTCTAACAGGGGTGAAAC

5' upstream sequence

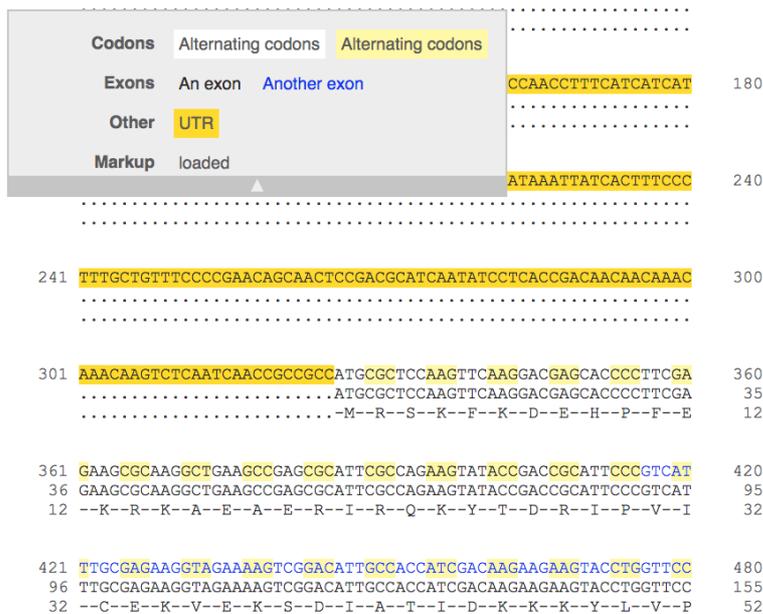
3' downstream sequence

Partially coding exon

Introns & up/down-stream sequence abbreviated & in lower case

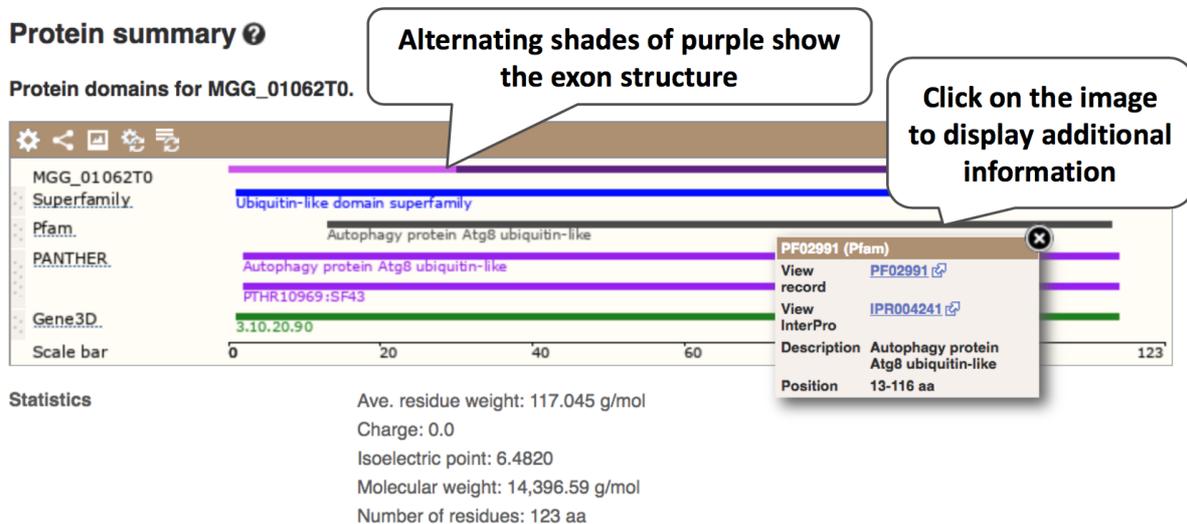
You may want to change the display (for example, to show more flanking sequence, or to show full introns). In order to do so click on [Configure this page](#) and change the display options accordingly.

Now click on the [cDNA](#) link to see the spliced transcript sequence.



UnTranslated Regions (UTRs) are highlighted in dark yellow, codons are highlighted in light yellow, and exon sequence is shown in black or blue letters to show exon divides.

We can look at the protein sequence in more detail, finding domains and structural information. Click on [Protein summary](#) to view domains from SignalP, Pfam, PROSITE, Superfamily, InterPro, and more.



Clicking on [Domains & features](#) shows a table of this information.

Next, follow the [General identifiers](#) link at the left.

This page shows information from other databases such as ENA, UniProtKB, INSDC and others, that match to the Ensembl transcript and protein.

General identifiers

This transcript corresponds to the following database identifiers:

External database	Database identifier
European Nucleotide Archive	CH476836  [align] [view all locations] CM001235  [align] [view all locations]
INSDC protein ID	EDK02260.1  [align] [view all locations] EHA48293.1  [align] [view all locations]
RefSeq mRNA predicted	XM_368182.1  [Target %id: 25; Query %id: 100] [align] 70-15 autophagy protein 8 (MGG_01062) partial mRNA [view all locations]
RefSeq peptide predicted	XP_368182.1  [Target %id: 100; Query %id: 100] [align] autophagy protein 8 [view all locations]
UniParc	UPI000021B9B9  [view all locations]
UniProtKB/Swiss-Prot	ATG8_MAGO7  [Target %id: 100; Query %id: 100] [align] Autophagy-related protein 8 [view all locations]

Exercises: Ensembl Fungi Genes and transcripts

Exercise – Exploring the *Zyloseptoria tritici* LEUC gene

- (a) Search Ensembl Fungi for the *LEUC* gene in *Zyloseptoria tritici* MG2. On which chromosome and which strand of the genome is this gene located?
- (b) What biological processes have been associated with LEUC?
- (c) View the gene sequence. Use the Configure this page option to show variants on the sequence and also the line numbering relative to the coordinate sequence, to this view.
- Are all exons shown in this display part of the LEUC gene? How can you tell?
 - Can you find the Stop Gained mutation? What letter is it represented by, looking at the letters and alleles of these surrounding variants what do you think this letter represents (these are [IUPAC ambiguity codes](#))
 - Which exon does the stop gained mutation fall in?
 - Export this sequence in RTF format
- (d) Click to go to the transcript tab by clicking on the transcript ID [Mycgr3T103221](#), and click to view the [Protein summary](#) page.
- Can you see the stop gained mutation we saw in (c) here?
 - Will this variant cause the deletion of an entire protein domain?
 - Which one(s)?

Exercise – Exploring a *Trichoderma reesei* gene

Find the *Trichoderma reesei* genome and search for the gene TRIREDRAFT_5868.

- (a) What are the molecular functions of this gene?
- (b) Go to the transcript tab for the transcript. How many exons does it have? Which one is the longest?
- (c) What domains can be found in the protein product of this transcript? How many different domain prediction methods agree with each of these domains?