

# Fungal Pathogen Genomics 2024

Wellcome Connecting Science

**Nishadi De Silva, Manuel Carbajo and Louise Paola Mirabueno**  
**Ensembl, EMBL-EBI**

Sunday, 2<sup>nd</sup> to Friday, 7<sup>th</sup> June 2024



# Ensembl Fungi: Introduction

**Database introductions,  
Fungal Pathogen Genomics 2024**

**Louisse Paola Mirabueno**

Ensembl Outreach Officer

Sunday, 2<sup>nd</sup> June 2024



# e!Ensembl

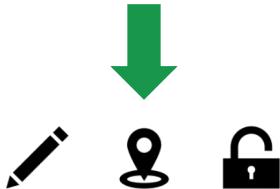
# Why do we need genome browsers?

CGGCCTTGGGCTCCGCCTCAGCTCAAGACTTAACCTCCCTCCCAGCTGTCCCAGATGACGCCATCTGAAATTCTTGAAACACGATCA  
CTTTAACGGAATATTGCTGTTGGGAAGTGTTCAGAGCTGGCACGCTGTATTGCCTACTTAAGCCCTGGTAATTGCTGTAT  
TCCGAAGACATGCTGATGGATTACCAAGCGCGTTGGTCTCTAAGCTGGAGCCCTGTCCCCACTAGCCACGCGTCAGTGGTAGCGTG  
ATTGAAACTAAATCGTATGAAAATCCTCTAGTCGCACTAGCCACGTTGAGTGCTTAATGTGGTAGTGGCACCGGTTGGACAG  
CACAGCTGTAAAATGTTCCCACAGTAAGCTGTTACCGTTCCAGGAGATGGACTGAATTAGAACAAATTCCAGCGCT  
TCTGAGTTTACCTCAGTCACATAATAAGGAATGCATCCCTGTGTAAGTGCATTGGCTTCTGTTGCAGACTTATTACCAAGCATT  
GGAGGAATATCGTAGGTAAAAATGCCTATTGGATCCAAAGAGAGGCCAACATTGGAAATTAAAGACACGCTGCAACAAAGCAGGTA  
TTGACAAATTATATAACTTATAAAATTACACCGAGAAAGTGTTCAAAAAATGCTGCTAAAACCAGTACGTACAGTGGCTT  
AGAACCATAAACTGTTCTTATGTGTATAAATCCAGTTAACACATAATCATCGTTGCAGGTTAACCATGATAAATAGAACGTC  
TAGTGGATAAAGAGGAAACTGGCCCTTGACTAGCAGTAGGAACAATTACTAACAAATCAGAACATTAAATGTTACTTTATGGCAGAAGTT  
GTCCAACCTTTGGTTTCAGTACTCCTTAACTCTTAAAGATCTAGGACCCCCGGAGTGCTTTGTTATGCTTACCATATTAGA  
AATTAAAACATAAGAATTAAAGGCTGGCGTGGCTCACGCCGTAAATCCCAGCACTTGGAGGCCAGGTGGCGGATCACTTGAGG  
CCAGAAGTTGAGACCAGCCTGGCAACATGGTAAACCCATTCTACTAAAAATACAAAAAATGTGCTGCGTGTGGTGCCTGCGCTG  
TAATCCCAGCTACACGGGAGGTGGAGGCAGGGAGAATCGCTGAACCTGGAGGCAGAGGTTGCAGTGAGCCAAGATCATGCCACTGC  
TAGCCTGGGCCACATAGCATGACTCTGTCTAAAACAAACAAACAAAAACTAAGAATTAAAGTTAAACTTACTTAAATAATGAA  
AGCTAACCCATTGCATATTATCACAAACATTCTAGGAAAAATAACTTTGAAAACAAGTGAGTGGAAATAGTTTACATTGCTTAC  
TCTTTAATGCTGGCTAAATAGAGATAGCTGGATTCACTTATCTGTGCTAATCTGTTATTGGTAGAAGTATGTGAAAAAAATTAACC  
TCACGTTGAAAAAAGGAATATTAAATAGTTTCAGTTACTTTGGTATTTCCTGTACTTGCTAGATTTCAAAGATCTAATAG  
ATATACCATAGGTCTTCCCAGTCGCAACATCATGCAGTGATTATTGGAAGATAGTGGTCTGAATTGCAATTGCTGCCGTTAAAGC

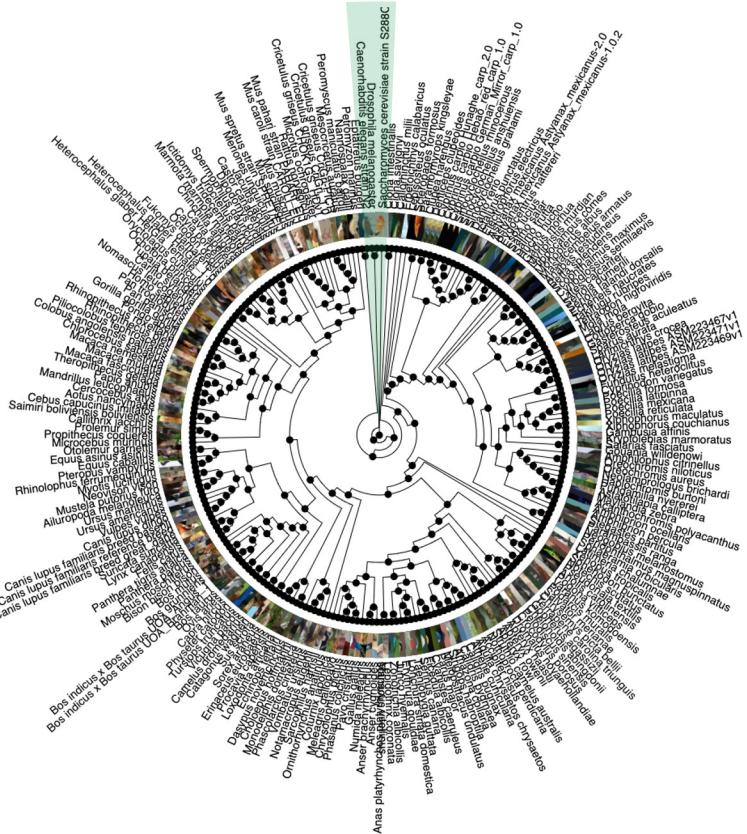
# What is Ensembl?

Ensembl annotates and maps genomic features from genome sequences, and brings together information from a wide range of other databases in a single site.

```
AGTGCTTAATGTGGCTAGTGGCACCGGTTGGACAGCACAGCTGTAAAATGTTCC  
CATCCTCACAGTAAGCTTACCGTTCCAGGAGATGGGACTGAATTAGAATTCAA  
ACAAATTTCAGCGCTCTGAGTTACCTCAGTCACATAATAAGGAATGCATC  
CCTGTGTAAGTGCATTTGGTCTCTGTTGCAGACTTATTACCAAGCATTGG  
AGGAATATCGTAGGTAAAATGCCTATTGGATCCAAGAGAGGCCAACATTTT  
GAAATTTTAAGACACGCTGCAACAAAGCAGGTATTGACAAATTATATAACCT  
TATAAATTACACCGAGAAAGTGTCTAAAAATGCTTGCTAAAACCCAGTAC  
GTCACAGTGTGCTTAGAACCATAAACTGTCCTATGTGTGTATAAATCCAGTT  
ACAACATAATCATCGTTGCAGGTTAACACATGATAAATATAGAACGTCTAGT  
GGATAAAGAGGAAACTGGCCCCTGACTAGCAGTAGGAACAATTACTAACAAATC
```



# Ensembl: vertebrate species



# Ensembl Genomes: non-vertebrate species

## EnsemblPlants

The screenshot shows the EnsemblPlants homepage. At the top, there's a search bar with "All species" selected and a placeholder "e.g. *Carboxy*\* or *chlor*". Below it, a "Wheat assemblies" section lists "Triticum aestivum" (IWGSC v1.0) and "A. thaliana" (TAIR10). A "Favourite genomes" dropdown is open, showing "Arabidopsis thaliana". Under "All genomes", a dropdown menu is open with "Select a species" and "Arabidopsis thaliana" highlighted. Other options include "Oryza sativa Japonica Group" (KASP v1.0) and "Triticum aestivum" (IWGSC). A link "View full list of all species" is at the bottom.

## EnsemblMetazoa

The screenshot shows the EnsemblMetazoa homepage. At the top, there's a search bar with "All species" selected and a placeholder "e.g. *CPRB* or *chitin*". Below it, a "What's New in Release 53" section lists "Updated data" for "Anophelis gambiae" (AgAPv4), "Drosophila melanogaster" (BDGP-32), and "Apia mellifera" (ApisMellifera1). A "Favourite genomes" dropdown is open, showing "Anophelis gambiae" (AgAPv4). Under "All genomes", a dropdown menu is open with "Select a species" and "Anophelis gambiae" highlighted. Other options include "Drosophila melanogaster" (BDGP-32) and "Apia mellifera" (ApisMellifera1). A link "View full list of all species" is at the bottom.

## EnsemblBacteria

The screenshot shows the EnsemblBacteria homepage. At the top, there's a search bar with "All species" selected and a placeholder "e.g. *Fe2+* or *uridine*\*". Below it, a "What's New in Release 53" section lists "New data" for "Culex pipienspallidus" (Southern house mosquito, GCA\_014932785.1) and "Ixodes scapularis" (Black-legged tick, GCA\_018007185.2). A "Favourite genomes" dropdown is open, showing "Culex pipienspallidus". Under "All genomes", a dropdown menu is open with "Select a species" and "Culex pipienspallidus" highlighted. Other options include "Ixodes scapularis" (Black-legged tick, GCA\_018007185.2) and "Desmoderma silvatica" (Marine stalk-eyed fly, GCA\_002371196.2). A link "View full list of all species" is at the bottom.

## EnsemblProtists

The screenshot shows the EnsemblProtists homepage. At the top, there's a search bar with "All species" selected and a placeholder "e.g. PF3D7\_0330500 or *cyst*". Below it, a "What's New in Release 53" section lists "No updated genomes from last release" and "Updated data" for "Plasmodium falciparum" (3D7) and "Dicyostelium discoideum" (dso-2.7). A "Favourite genomes" dropdown is open, showing "Plasmodium falciparum" (3D7). Under "All genomes", a dropdown menu is open with "Select a species" and "Plasmodium falciparum" (3D7) highlighted. Other options include "Dicyostelium discoideum" (dso-2.7) and "Phytophthora infestans" (ASMI-0294r). A link "View full list of all species" is at the bottom.

## EnsemblFungi

The screenshot shows the EnsemblFungi homepage. At the top, there's a search bar with "All species" selected and a placeholder "e.g. NAT2 or alcohol". Below it, a "What's New in Release 53" section lists "New assemblies with gene and protein annotation every two weeks" for "Saccharomyces cerevisiae" (R64-1) and "Aspergillus nidulans" (ASMI-101). A "Favourite genomes" dropdown is open, showing "Saccharomyces cerevisiae" (R64-1). Under "All genomes", a dropdown menu is open with "Select a species" and "Saccharomyces cerevisiae" (R64-1) highlighted. Other options include "Aspergillus nidulans" (ASMI-101) and "Puccinia graminis" (Asterisco1b). A link "View full list of all species" is at the bottom.

## EnsemblCOVID-19

The screenshot shows the EnsemblCOVID-19 homepage. At the top, there's a search bar with "SARS-CoV-2" selected. Below it, a "Reference Genome" section provides details about the genome assembly ASMB85889v3 (GCA\_000888893.4) and sequence MN909947. A "View in Genome browser" button is available. In the "COVID-19 resources" section, there are three buttons: "Example gene", "Example protein summary", and "Whole genome". A "Sequence and annotation data" section includes a "Download data" button. A note at the bottom states that the EnsemblGENCODE human protein-coding gene set used in EnsemblCOVID-19 can be found in Human.Gencode33.Genes.Annotation.gff.

# Assemblies and annotation



Genome assemblies

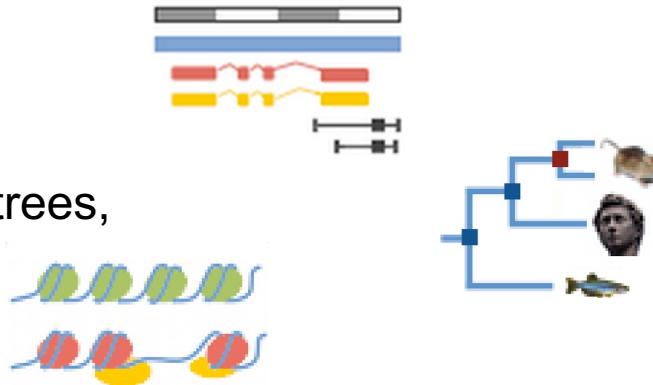


Data generators and  
collaborators, in-house annotation



# Ensembl Fungi features

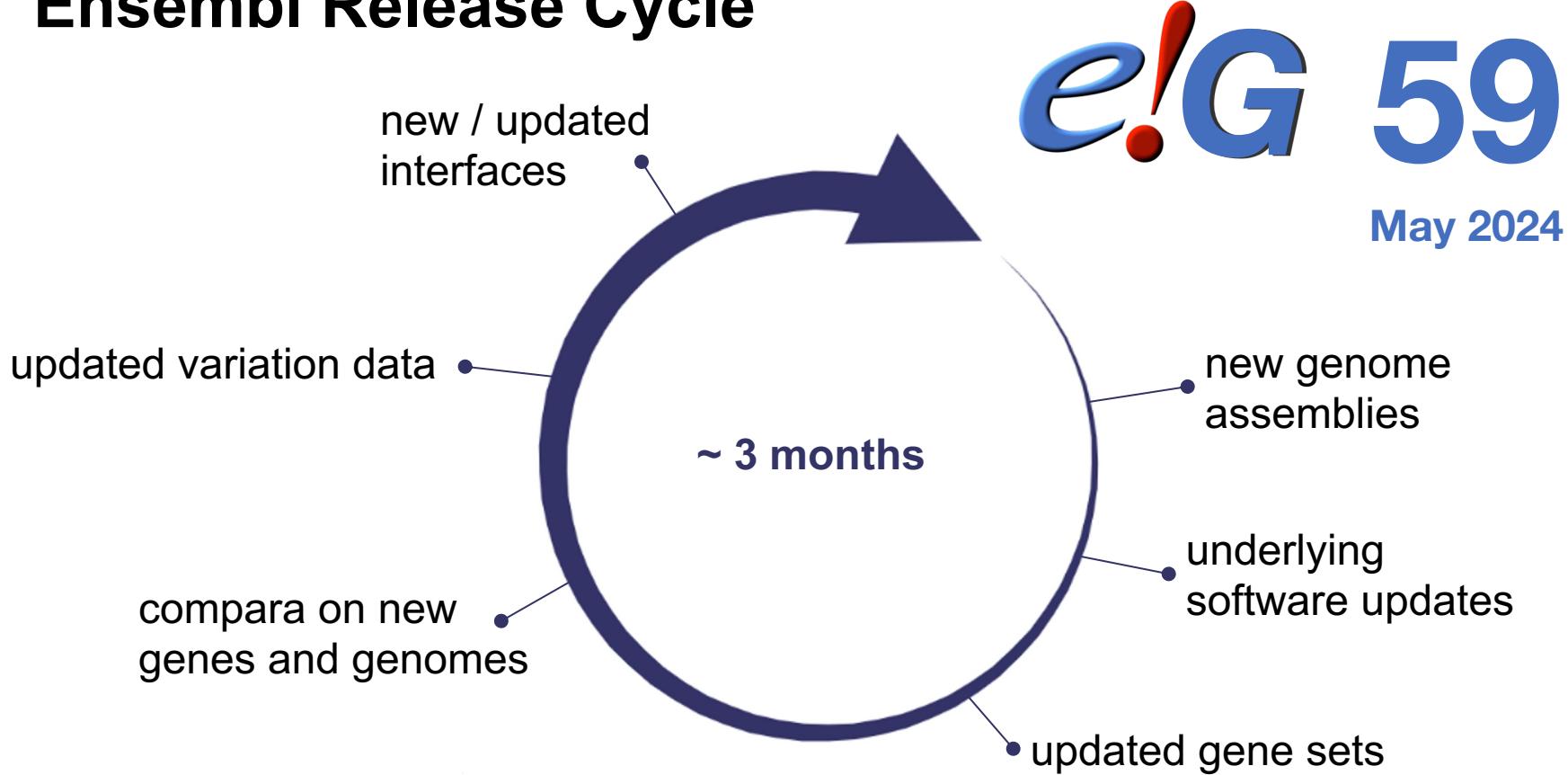
- Genome assemblies and gene builds
- Variation data and Ensembl VEP
- Comparative genomics (alignments, gene trees, homologues)
- Cross-species molecular interactions
- Microarray data
- TrackHubs
- BioMart (database query)
- Display your own data
- Programmatic access via APIs
- Completely open-source (FTP, GitHub)



**Ve!P**

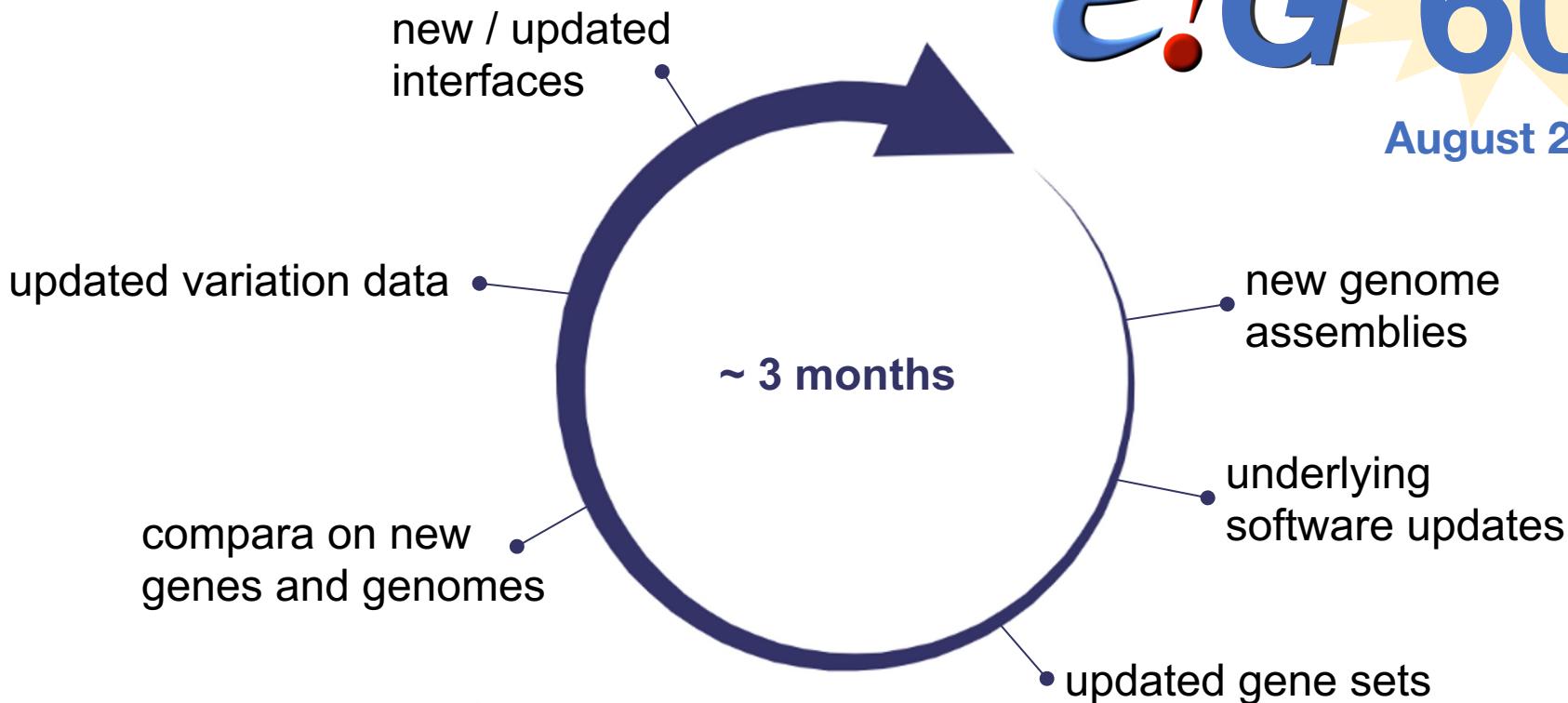


# Ensembl Release Cycle



e!G 59  
May 2024

# Ensembl Release Cycle



# Ensembl Rapid Release

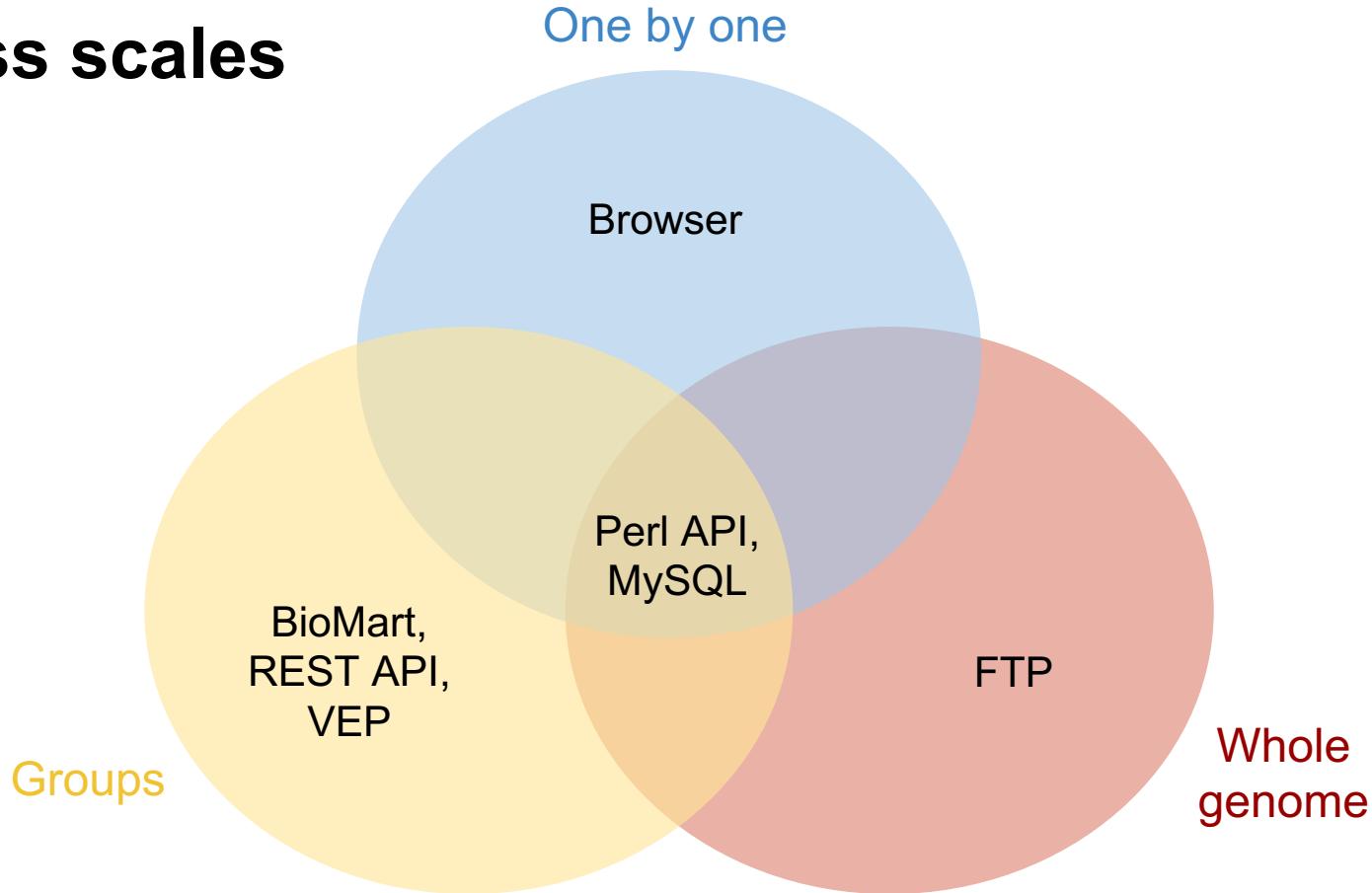
- Released every 2 weeks ✓
- Genome assemblies and gene builds ✓
- BLAST ✓
- Limited genetic variation data
- Comparative analyses are gradually rolled out across all species
- No BioMart ✗

The screenshot shows the Ensembl Rapid Release website. At the top, there's a navigation bar with links to BLAST, Tools, Downloads, Help & Docs, Known Bugs, and a Blog. Below the header, there's a search bar for BLAST with the placeholder "Search our genomes for your DNA or protein sequence". A dropdown menu is open, showing "Camarhynchus parvulus (Small)" selected. To the right of the search bar is a "Go" button. Below the search bar, there's a note: "e.g. Camarhynchus parvulus 2:361680-384534 or Clytia hemisphaerica IPR001650". On the right side of the page, under "Latest Genomes", there's a list of seven new genomes released:

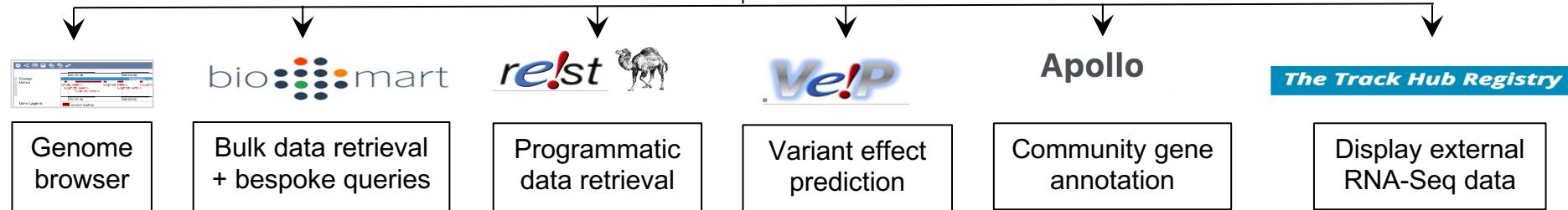
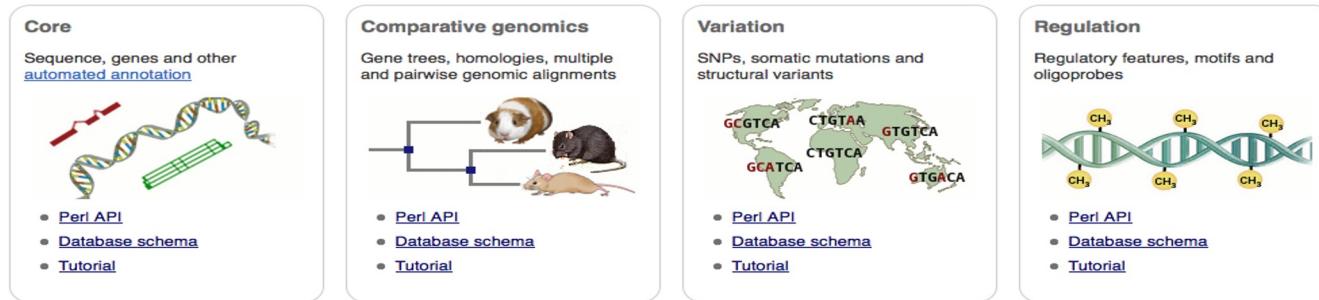
- *Acanthoagonus latus* (Yellowfin seabream) - GCA\_904848185.1
- *Chelmon rostratus* (Copperband butterflyfish) - GCA\_017976325.1
- *Hippoglossus hippoglossus* (Atlantic halibut) - GCA\_009819705.1
- *Meleagris gallopavo* (Turkey) - GCA\_905368555.1
- *Patella pellucida* (Gastropods) - GCA\_017208275.1
- *Synemphus aculeatus* (Greater pipefish) - GCA\_901708675.2
- *Toxotes jaculator* (Banded archerfish) - GCA\_017976425.1

Below this, there are links to "View all species and download data" and "Comparative analyses".

# Access scales

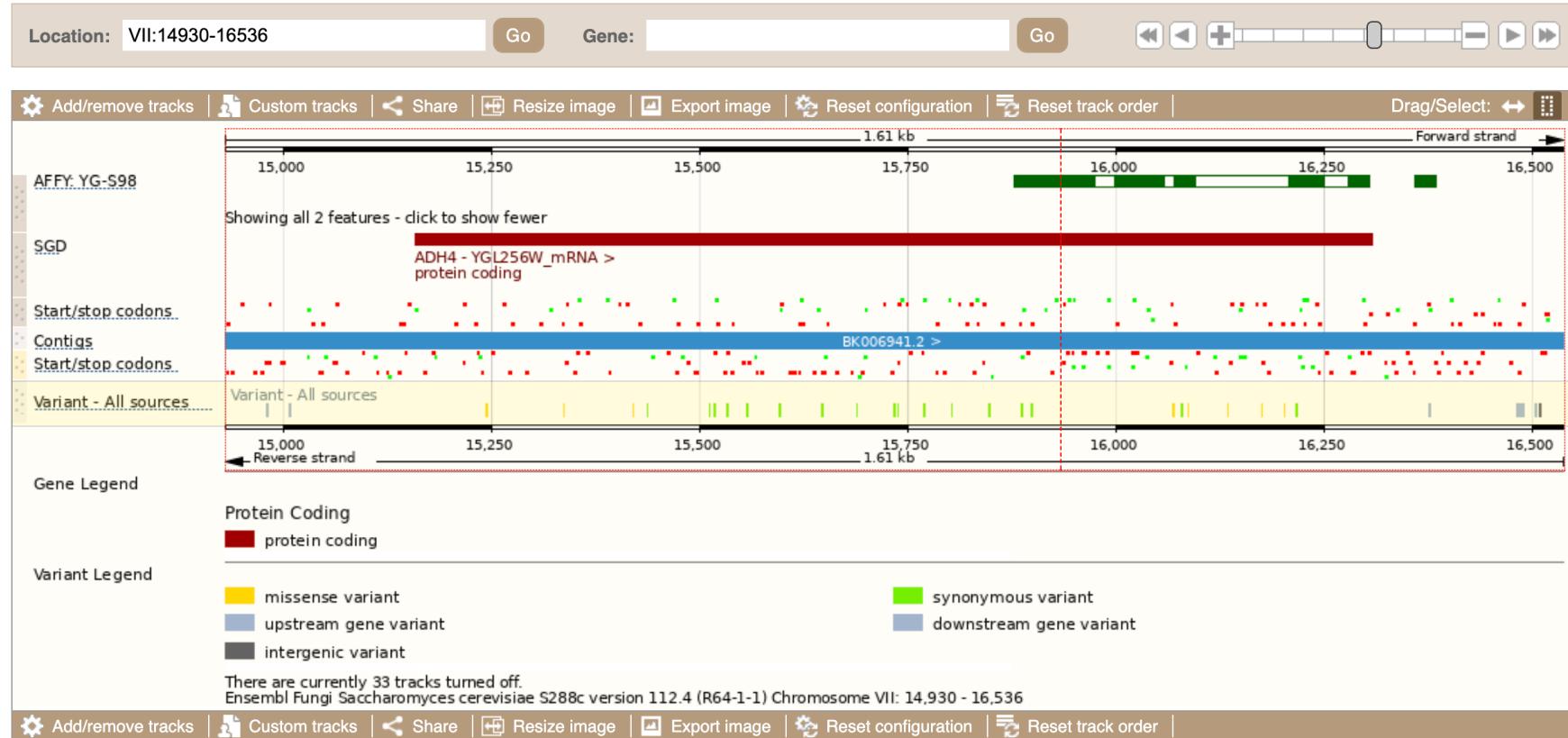


# Use the Ensembl infrastructure

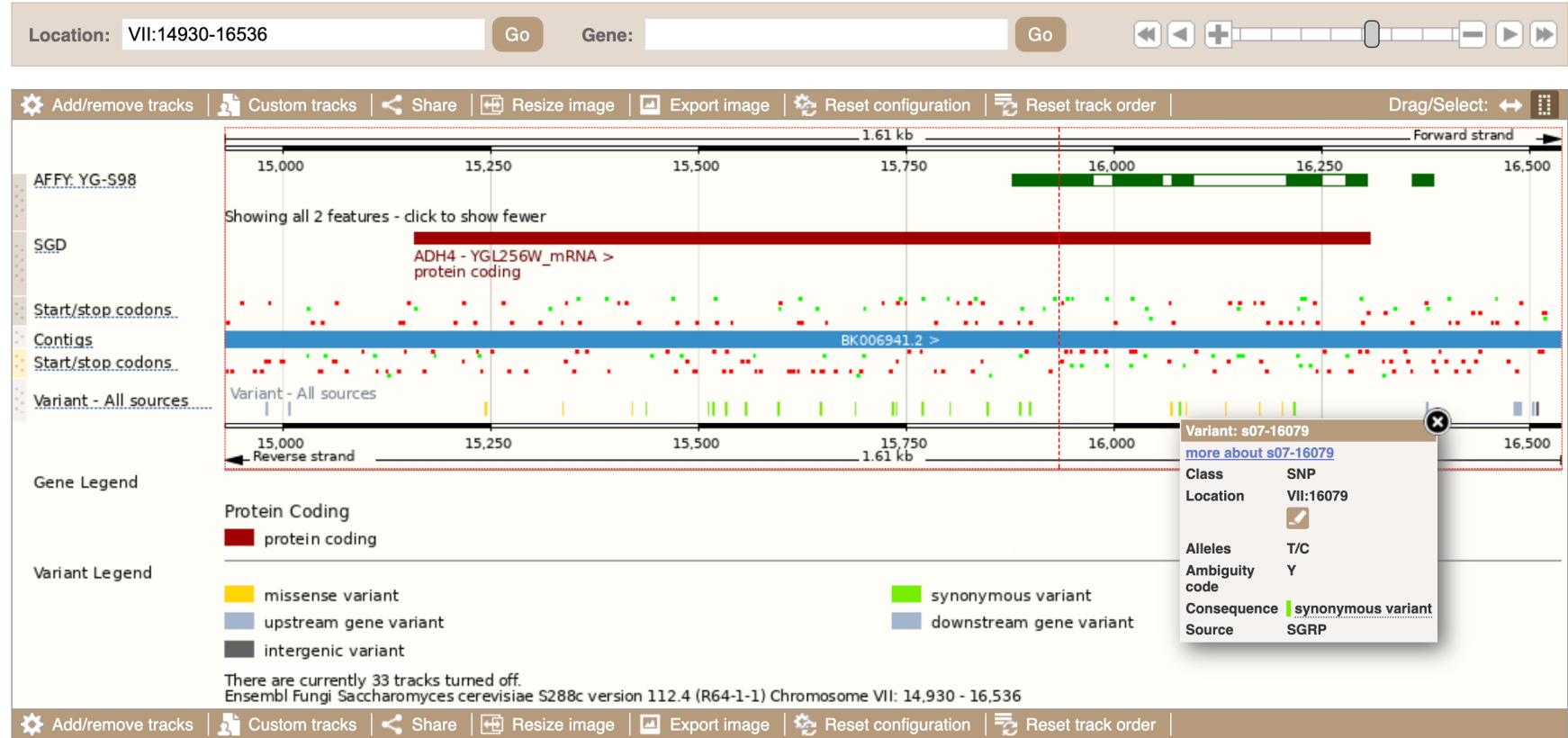


## Training

# Visualising genomic regions



# Visualising genomic regions





# Saccharomyces cerevisiae (R64-1-1) ▾

Location: VII:15,159-16,307

Gene: ADH4

Transcript: ADH4

Variant: s07-14981

## Gene-based displays

## Summary

- Splice variants
- Transcript comparison
- Gene alleles

## Sequence

- Secondary Structure
- Gene families

## Literature

- Fungal Compara
- Genomic alignments
- Gene tree
- Gene gain/loss tree
- Orthologues
- Paralogues

## Pan-taxonomic Compara

- Gene Tree
- Orthologues

## Ontologies

- GO: Cellular component
- GO: Biological process
- GO: Molecular function
- PHL: Phibase identifier

## Phenotypes

- Genetic Variation
- Variant table
- Variant image
- Structural variants

## Gene expression

- Pathway

## Molecular interactions

## Regulation

## External references

## Supporting evidence

## ID History

- Gene history

**Configure this page****Custom tracks****Export data****Share this page****Bookmark this page****Gene: ADH4 YGL256W**

## Description

Alcohol dehydrogenase isoenzyme type IV; dimeric enzyme demonstrated to be zinc-dependent despite sequence similarity to iron-activated alcohol dehydrogenases; transcription is induced in response to zinc deficiency [Source:SGD;Acc:S000003225]

## Gene Synonyms

NRC465, ZRG5

## Location

Chromosome VII: 15,159-16,307 forward strand.

## About this gene

R64-1-1:BK006941.2

## Transcripts

This gene has 1 transcript ([splice variant](#)), 306 orthologues and is a member of [1 Ensembl protein family](#).**Show transcript table****Summary** ⓘ

## Name

[ADH4](#) (SGD gene name)

## UniProtKB

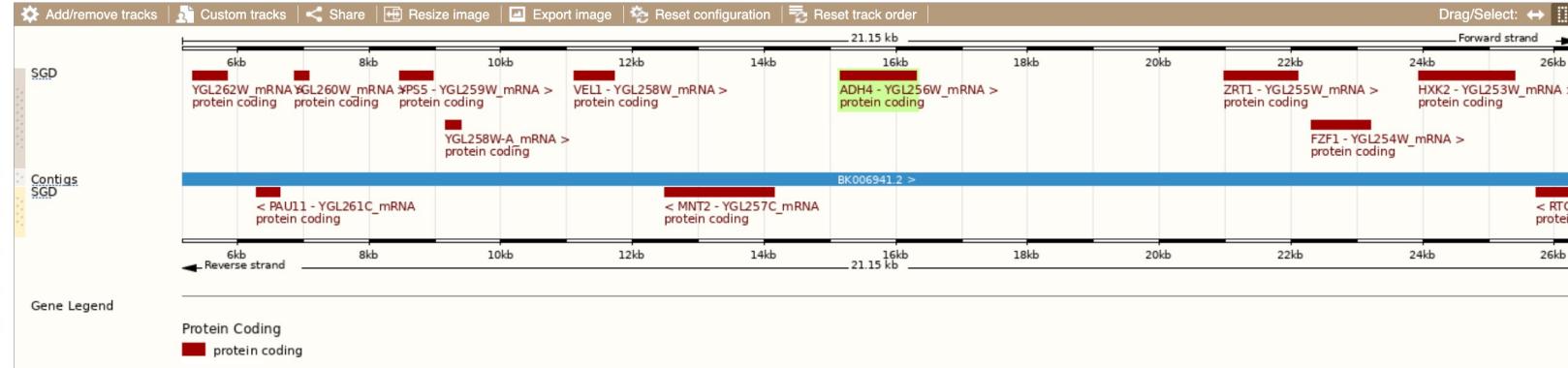
This gene has proteins that correspond to the following UniProtKB identifiers: [P10127](#)

## Gene type

Protein coding

## Annotation method

Annotation imported from SGD

[Go to Region in Detail](#) for more tracks and navigation options (e.g. zooming)



**Saccharomyces cerevisiae (R64-1-1) ▾**

Location: VII:15,159-16,307

Gene: ADH4

Transcript: ADH4

Variant: s07-14981

## Gene-based displays

## ↳ Summary

- Splice variants
- Transcript comparison
- Gene alleles

## Sequence

- Secondary Structure
- Gene families
- Literature

## Fungal Compara

- Genomic alignments
- Gene tree
- Gene gain/loss tree
- Orthologues
- Paralogues

## Pan-taxonomic Compara

- Gene Tree
- Orthologues

## Ontologies

- GO: Cellular component
- GO: Biological process
- GO: Molecular function
- PHI: Phibase identifier

## Phenotypes

## Genetic Variation

- Variant table
- Variant image
- Structural variants

## Gene expression

## Pathway

## Molecular interactions

## Regulation

## External references

## Supporting evidence

## ID History

- Gene history

Configure this page

Custom tracks

Export data

Share this page

Bookmark this page

**Gene: ADH4 YGL256W**

## Description

Alcohol dehydrogenase isoenzyme type IV; dimeric enzyme demonstrated to be zinc-dependent despite sequence similarity to iron-activated alcohol dehydrogenases; transcription is induced in response to zinc deficiency [Source:SGD;Acc:S000003225]

## Gene Synonyms

NRC465, ZRG5

## Location

Chromosome VII: 15,159-16,307 forward strand.

R64-1-1:BK006941.2

## About this gene

This gene has 1 transcript ([splice variant](#)), 306 orthologues and is a member of [1 Ensembl protein family](#).

## Transcripts

[Show transcript table](#)**Summary** ⓘ

## Name

## UniProt

## Gene

## Annotations

[HMMER](#) | [BLAST](#) | [BioMart](#) | [Tools](#) | [Downloads](#) | [Help & Docs](#) | [Blog](#)

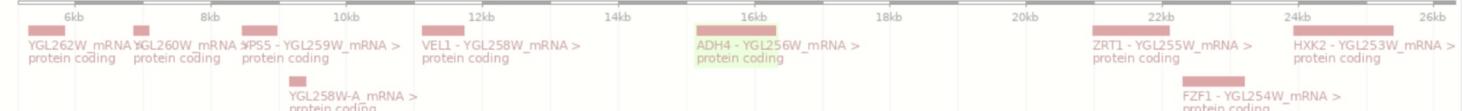
Go to Region In Detail for more tracks and navigation options (e.g. zooming)

[Add/remove tracks](#)[Custom tracks](#)[Share](#)[Resize image](#)[Export image](#)[Reset configuration](#)[Reset track order](#)

Drag&gt;Select:

Forward strand 

SGD



Contigs SGD

SGD

Gene Legend

Protein Coding

 protein coding


 Saccharomyces cerevisiae (R64-1-1) ▾

Location: VII:15,159-16,307

Gene: ADH4

Transcript: ADH4

Variant: s07-14981

## Gene-based displays

## Gene: ADH4 YGL256W

## Description

Alcohol dehydrogenase isoenzyme type IV; dimeric enzyme demonstrated to be zinc-dependent despite sequence similarity to iron-activated alcohol dehydrogenases; transcription is induced in response to zinc deficiency [Source:SGD;Acc:S000003225]

## Gene Synonyms

NRC465, ZRG5

## Location

Chromosome VII: 15,159-16,307 forward strand.

## About this gene

R64-1-1-BK006941.2

## Transcripts

This gene has 1 transcript ([splice variant](#)), 306 orthologues and is a member of [1 Ensembl protein family](#).[Show transcript table](#)

# Saccharomyces cerevisiae (R64-1-1) ▾

Location: VII:15,159-16,307

Gene: ADH4

Transcript: ADH4

Variant: s07-14981





## Saccharomyces cerevisiae (R64-1-1) ▾

Location: VII:15,159-16,307

Gene: ADH4

Transcript: ADH4

Variant: s07-14981

## Gene-based displays

## ⊖ Summary

- └ Splice variants
- └ Transcript comparison
- └ Gene alleles

## ⊖ Sequence

- └ Secondary Structure

## ⊖ Gene families

## ⊖ Literature

## ⊖ Fungal Compara

- └ Genomic alignments
- └ Gene tree
- └ Gene gain/loss tree
- └ Orthologues
- └ Paralogues

## ⊖ Pan-taxonomic Compara

- └ Gene Tree
- └ Orthologues

## ⊖ Ontologies

- └ GO: Cellular component
- └ GO: Biological process
- └ GO: Molecular function

## ⊖ Phenotypes

## ⊖ Genetic Variation

- └ Variant table
- └ Variant image
- └ Structural variants

## ⊖ Gene expression

## ⊖ Pathway

## ⊖ Molecular interactions

## ⊖ Regulation

## ⊖ External references

## ⊖ Supporting evidence

## ⊖ ID History

## ⊖ Gene history

[Configure this page](#)[Custom tracks](#)[Export data](#)[Share this page](#)[Bookmark this page](#)

Gene: ADH4 YGL256W

Description

Alcohol dehydrogenase I  
zinc deficiency [Source: SGD]

Gene synonyms

NRC465, ZRG5

Location

Chromosome VII: 15,159  
R64-1-1-BK006941.2

About this gene

This gene has 1 transcript

[Show transcript table](#)

## Summary ⓘ

Name

ADH4 ⓘ (SGD gene name)

UniProtKB

This gene has proteins in UniProtKB

Gene type

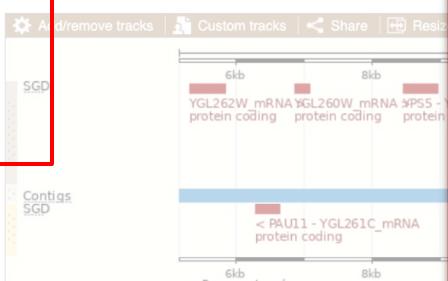
Protein coding

Annotation method

Annotation imported from SGD

[Go to Region in Detail for more tracks and navigation](#)

[Add/remove tracks](#) | [Custom tracks](#) | [Share](#) | [Results](#)



Gene Legend

Protein Coding

protein coding

Location: VII:15,159-16,307

Gene: ADH4

## Gene-based displays

## ⊖ Summary

- └ Splice variants
- └ Transcript comparison
- └ Gene alleles

## ⊖ Sequence

- └ Secondary Structure
- └ Gene families
- └ Literature

## ⊖ Fungal Compara

- └ Genomic alignments
- └ Gene tree
- └ Gene gain/loss tree
- └ Orthologues
- └ Paralogues

## ⊖ Pan-taxonomic Compara

- └ Gene Tree
- └ Orthologues

## ⊖ Ontologies

- └ GO: Cellular component
- └ GO: Biological process
- └ GO: Molecular function

## ⊖ Phenotypes

## ⊖ Genetic Variation

- └ Variant table
- └ Variant image
- └ Structural variants

## ⊖ Gene expression

## ⊖ Pathway

## ⊖ Molecular interactions

## ⊖ Regulation

## ⊖ External references

## ⊖ Supporting evidence

## ⊖ ID History

- └ Gene history

Gene

Gene

Description

despite sequence similarity to iron-activated alcohol dehydrogenases; transcription is induced in response to

Gene S

Location

About family

Transc

## Sum

Name

UniProt

Gene t

Annotation

Drag&gt;Select: ↪ ↮

Forward strand



Add

SGD

Gene

Legend



## Saccharomyces cerevisiae (R64-1-1) ▾

Location: VII:15,159-16,307

Gene: ADH4

Transcript: ADH4

Variant: s07-14981

## Gene-based displays

## ⊖ Summary

- Splice variants
- Transcript comparison
- Gene alleles

## ⊖ Sequence

- Secondary Structure
- Gene families
- Literature

## ⊖ Fungal Compara

- Genomic alignments
- Gene tree
- Gene gain/loss tree
- Orthologues
- Paralogues

## ⊖ Pan-taxonomic Compara

- Gene Tree
- Orthologues

## ⊖ Ontologies

- GO: Cellular component
- GO: Biological process
- GO: Molecular function
- PHL: Phibase identifier

## ⊖ Phenotypes

- Genetic Variation
- Variant table
- Variant image
- Structural variants

## - Gene expression

## - Pathway

## - Molecular interactions

## - Regulation

## - External references

## - Supporting evidence

## ⊖ ID History

## └ Gene history

## Configure this page

## Custom tracks

## Export data

## Share this page

## Bookmark this page

**Gene: ADH4 YGL256W**

## Description

Alcohol dehydrogenase isoenzyme type IV; dimeric enzyme demonstrated to be zinc-dependent despite sequence similarity to iron-activated alcohol dehydrogenases; transcription is induced in response to zinc deficiency [Source:SGD;Acc:S000003225]

## Gene Synonyms

NRC465, ZRC6

## Location

Chromosome

R64-1-1-BK

## About this gene

This gene has

## Transcripts

Show tra

**Summary** ⓘ

## Name

ADH4 ⓘ (SGD)

## UniProtKB

This gene has

## Gene type

Protein coding

## Annotation method

Annotation in

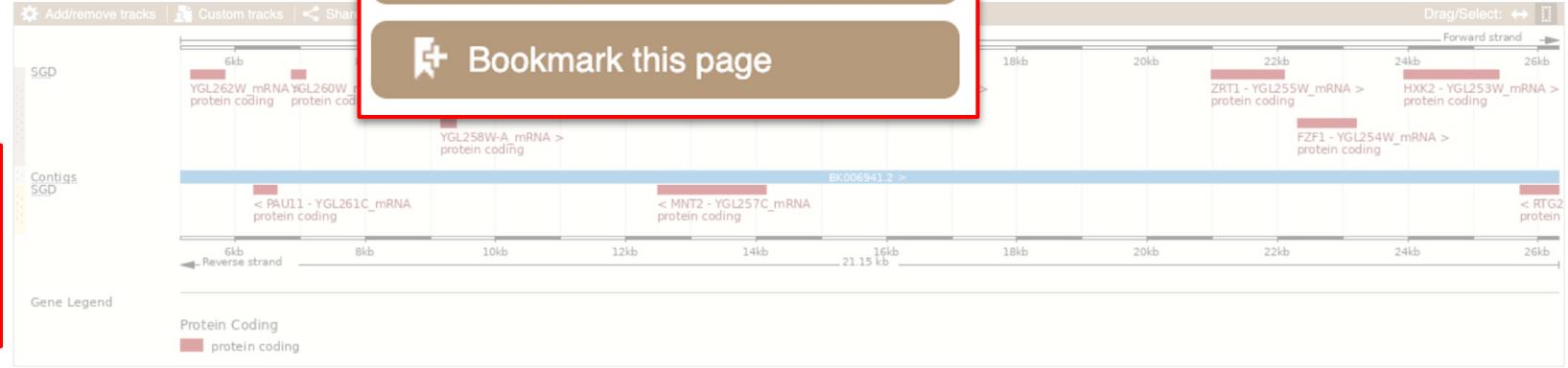


## Add/remove tracks

## Custom tracks

## Share this page

Configure this page
  
  
 Custom tracks
  
  
 Export data
  
  
 Share this page
  
  
 Bookmark this page



# What can you use Ensembl Fungi for?

- Visualise genomic regions and map custom annotations using tracks
- Learn more about genes of interest
- Learn more about variants of interest and annotate your own variants
- Investigate orthologues to explore behaviour in other pathogens
- Search for information in external databases

# Fungal Pathogen Genomics 2024: Ensembl Fungi

Sun, 2 June	Sun, 2 June	Mon, 3 June	Wed, 5 June
Introduction to Database Queries	Transcriptomics & Proteomics	Comparative Genomics & Orthology	SNPs & Variants
<ul style="list-style-type: none"><li>• Bulk Ensembl data download</li><li>• Molecular interactions</li></ul>	<ul style="list-style-type: none"><li>• Public RNAseq data</li><li>• Visualise your own data</li></ul>	<ul style="list-style-type: none"><li>• Gene trees</li><li>• Whole-genome alignments</li></ul>	<ul style="list-style-type: none"><li>• Variation data</li><li>• Ensembl Variant Effect Predictor (VEP)</li></ul>

# Ensembl Fungi: BioMart & molecular interactions

Introduction to database queries,  
Fungal Pathogen Genomics 2024

Louisse Paola Mirabueno

Ensembl Outreach Officer

Sunday, 2<sup>nd</sup> June 2023



# Ensembl BioMart

The 4-step principle

## Step 1: Dataset

Choose the database and species.

## Step 2: Filters

Narrow down the dataset.

## Step 3: Attributes

Specify your output and what to print on your table.

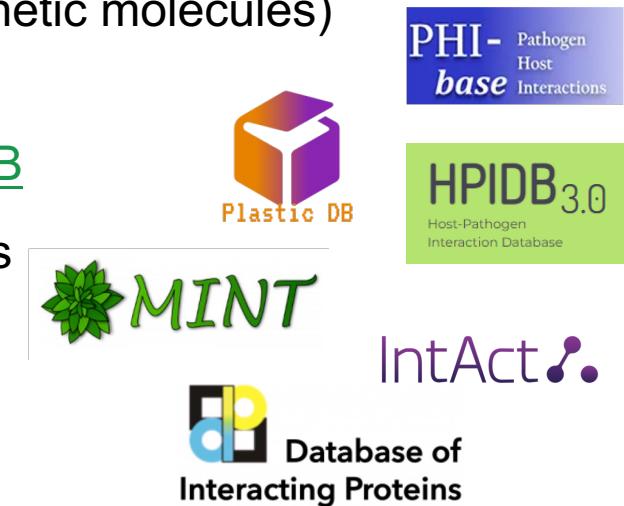
## Step 4: Results

Choose the format of your results and export.

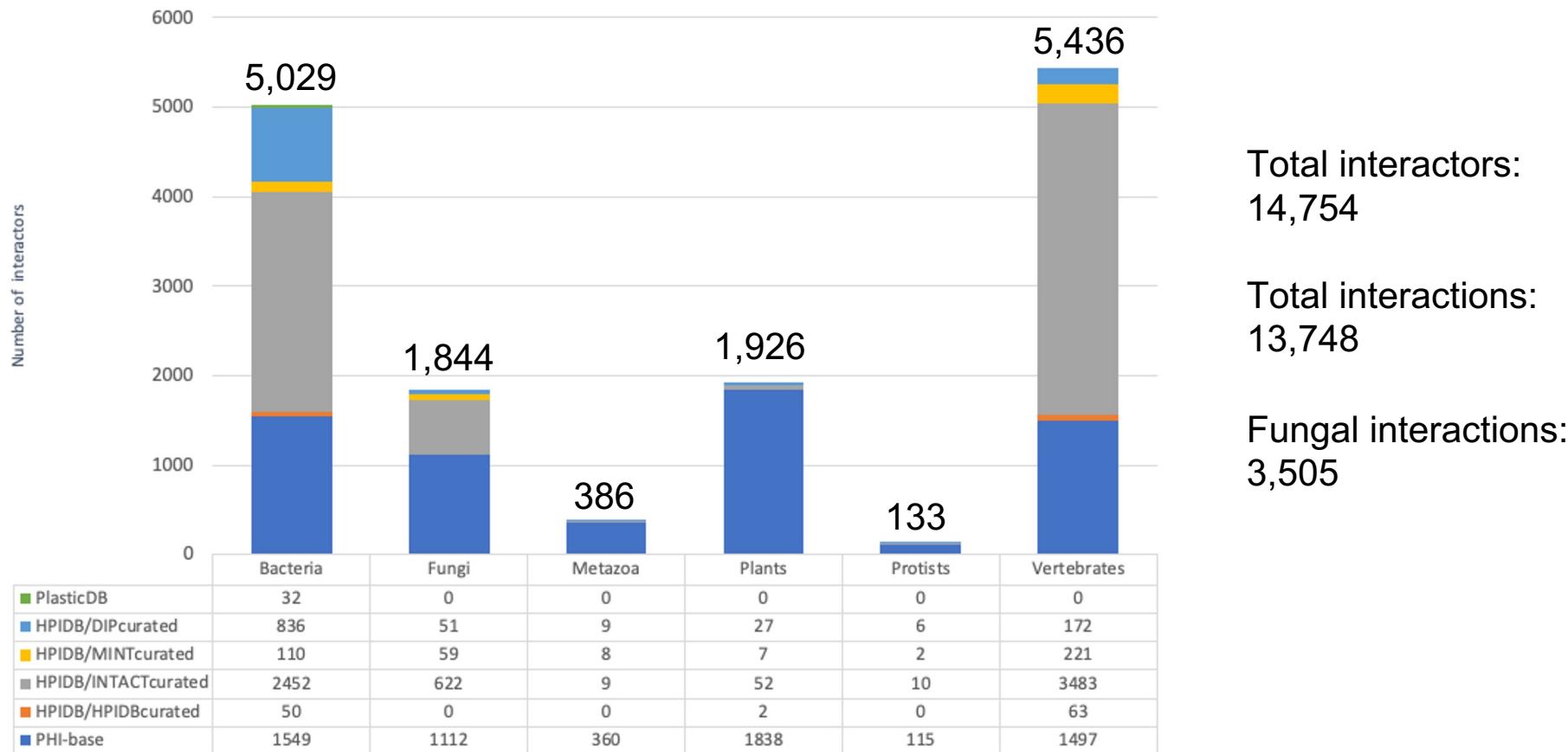
# Cross-species molecular interactions

A new Ensembl resource:

- 2 molecules + experimentally verified + published + curated
- Cross-species (species-species OR species-synthetic molecules)
- Imported from external interaction databases:  
[DIP](#), [HPIDB](#), [IntAct](#), [MINT](#), [PHI-base](#), [PlasticDB](#)
- Conservative import of exact protein/strain matches
- Data interlaced with other Ensembl resources



# Unique interactors per division and source database



# Accessing molecular interactions

# Accessing molecular interactions (browser)

EnsemblFungi ▾ HMMER | BLAST | BioMart | Tools | Downloads | Help & Docs | Blog

Phaeosphaeria nodorum (ASM14691v1) ▾

Location: supercont1.10:1,101,626-1,102,347 Gene: SNOG\_20078 Transcript: SNOT\_20078

Gene-based displays

- Summary
- Splice variants
- Transcript comparison
- Gene alleles
- Sequence
  - Secondary Structure
- Gene families
- Literature
- Fungal Compara
  - Genomic alignments
  - Gene tree
  - Gene gain/loss tree
  - Orthologues
  - Paralogues
- Pan-taxonomic Compara
  - Gene Tree
  - Orthologues
- Ontologies
  - GO: Biological process
  - GO: Molecular function
  - GO: Cellular component
  - PHI: Phibase identifier
- Phenotypes
- Genetic Variation
  - Variant table
  - Variant image
  - Structural variants
  - Gene expression
  - Pathway
- Molecular interactions
  - Regulation
  - External references
  - Supporting evidence
- ID History
  - Gene history

Gene: SNOG\_20078

Description Putative uncharacterized protein [Source:UniProtKB/TrEMBL;Acc:[A9JX75|e](#)]

Location SuperContig supercont1.10:1,101,626-1,102,347 reverse strand.  
ASM14691v1:CH445334.1

About this gene This gene has 1 transcript ([splice variant](#)).

Transcripts [Show transcript table](#)

**Molecular interactions** Cross-species interactions imported from PHI-base, HPIDB and PlasticDB with exact matches to proteins in Ensembl.

This species		Interacts with		<a href="#">Show metadata</a>
Species	Gene ID	Interactor	Identifier	Source DB
Phaeosphaeria nodorum	<a href="#">SNOG_20078</a>	protein	<a href="#">uniprot:A9JX75</a>	
SN15				

Species	Gene ID	Interactor	Identifier	Source DB
Triticum aestivum	UNDETERMINED	protein	UNDETERMINED	<a href="#">PHI-base</a>
Triticum aestivum	<a href="#">TraesCS1B02G004100</a>	protein	<a href="#">uniprot:W5AB81</a>	<a href="#">PHI-base</a>
Triticum aestivum	<a href="#">TraesCS1B02G004100</a>	protein	<a href="#">uniprot:W5AB81</a>	<a href="#">PHI-base</a>

Ensembl Fungi release 56 - Feb 2023 © [EMBL-EBI](#)

---

**About Us**

- [About us](#)
- [Contact us](#)
- [Citing Ensembl Genomes](#)
- [Privacy policy](#)
- [Disclaimer](#)

**Get help**

- [Using this website](#)
- [Documentation](#)
- [Adding custom tracks](#)
- [Downloading data](#)

**Our sister sites**

- [Ensembl](#)
- [Ensembl Bacteria](#)
- [Ensembl Plants](#)
- [Ensembl Protists](#)
- [Ensembl Metazoa](#)

**Follow us**

- [Blog](#)
- [Twitter](#)

# Accessing molecular interactions (REST API)

Swagger  
Supported by SMARTBEAR

<https://interactions.rest.ensembl.org/?format=openapi>

**API interactions doc v1**

[ Base URL: [interactions.rest.ensembl.org/](https://interactions.rest.ensembl.org/) ]  
<https://interactions.rest.ensembl.org/?format=openapi>

Ensembl Interactions REST API  
A list of endpoints to provide access to the Ensembl molecular interactions database.  
We welcome your suggestions to add endpoints to fit new use case scenarios. Please write to [helpdesk@ensemblgenomes.org](mailto:helpdesk@ensemblgenomes.org)

Schemes  
HTTPS ▾

Authorize 

Filter by tag

**display\_by\_gene** ^

**GET** /display\_by\_gene/{ens\_stbl\_id} **display\_by\_gene\_read** ▾ 

**ensembl\_gene** ^

**GET** /ensembl\_gene **ensembl\_gene\_list** ▾ 

**GET** /ensembl\_gene/ensembl\_name **ensembl\_gene\_ensembl\_name\_list** ▾ 

# Accessing molecular interactions

- API REST endpoints: <https://interactions.rest.ensembl.org/interaction>

?species\_A=zymoseptoria tritici

&species\_B=Triticum aestivum

&meta\_key=disease

Filters

Field filters

Interaction id:

Interactor name:

ensembl stable id:

species A:  
zymoseptoria tritici

species B:  
triticum aestivum

Source db:

meta value:

meta key:  
disease

Submit



```
[  
  {  
    "interaction_id": 18174,  
    "interactor_1": "uniprot:F9WR1",  
    "interactor_2": "uniprot:UNDETERMINED_PHI:4966_Triticum_aestivum",  
    "ensembl_gene_1": "Mycgr3G53658",  
    "ensembl_gene_2": "UNDETERMINED_PHI:4966_Triticum_aestivum",  
    "species_1": "zymoseptoria tritici",  
    "species_2": "Triticum aestivum",  
    "doi": "26092798",  
    "source_db": "PHI-base"  
  },  
  {  
    "interaction_id": 18254,  
    "interactor_1": "uniprot:F9X056",  
    "interactor_2": "uniprot:UNDETERMINED_PHI:2442_Triticum_aestivum",  
    "ensembl_gene_1": "Mycgr3G88451",  
    "ensembl_gene_2": "UNDETERMINED_PHI:2442_Triticum_aestivum",  
    "species_1": "zymoseptoria tritici",  
    "species_2": "Triticum aestivum",  
    "doi": "19522561",  
    "source_db": "PHI-base"  
  },  
  {  
    "interaction_id": 18277,  
    "interactor_1": "uniprot:F9X7N2",  
    "interactor_2": "uniprot:UNDETERMINED_PHI:4943_Triticum_aestivum",  
    "ensembl_gene_1": "Mycgr3G85940",  
    "ensembl_gene_2": "UNDETERMINED_PHI:4943_Triticum_aestivum",  
    "species_1": "zymoseptoria tritici",  
    "species_2": "Triticum aestivum",  
    "doi": "26092796",  
    "source_db": "PHI-base"  
  }  
]
```

# Ensembl Fungi: Custom tracks

Transcriptomics & proteomics  
Fungal Pathogen Genomics 2023

Louisse Paola Mirabueno

Ensembl Outreach Officer

Sunday, 2<sup>nd</sup> June 2023



# Track Hubs

- The [Track Hub registry](#) is a way of publicly sharing large amounts of data
- Many large collaborative projects make their data available in sets of tracks organised as a single hub
- You can create a Track Hub containing your own data to share with the public and collaborators
- You can search for data in the Track Hub registry and view this data directly on Ensembl

# Visualise your own data

Upload data:

- Data saved by Ensembl
- 20 MB limit (not possible for large files)

Attach remote file:

- URL-based (HTTP or FTP)
- No size-limit

# Data visualisation formats

## Genes and features

- GBrowse
- GFF / GTF
- BigBed

## Continuous values

- BedGraph
- WIG
- BigWig

## Sequence alignments

- PSL
- BAM

## Variants

- VCF

# Where to host URLs for larger files?

If your file is >20 MB in size, you will need to attach it by URL. You can create a URL using:

- Institute FTP site
- [CyVerse](#)
- [FigShare](#)

Note that Google Drive will not work and GitHub has proven unreliable.

# Exercise: Attaching Track Hubs to Ensembl Fungi

The **RNA-Seq alignment hub SRP062592** track in the TrackHub Registry is currently unavailable. Instead, please use the URL below (this is the same file):

<https://ftp.sra.ebi.ac.uk/vol1/analysis/ERZ492/ERZ492288/SRR2174493.cram>

# Ensembl Fungi: Compara

Comparative genomics & orthology  
Fungal Pathogen Genomics 2024

Louisse Paola Mirabueno

Ensembl Outreach Officer

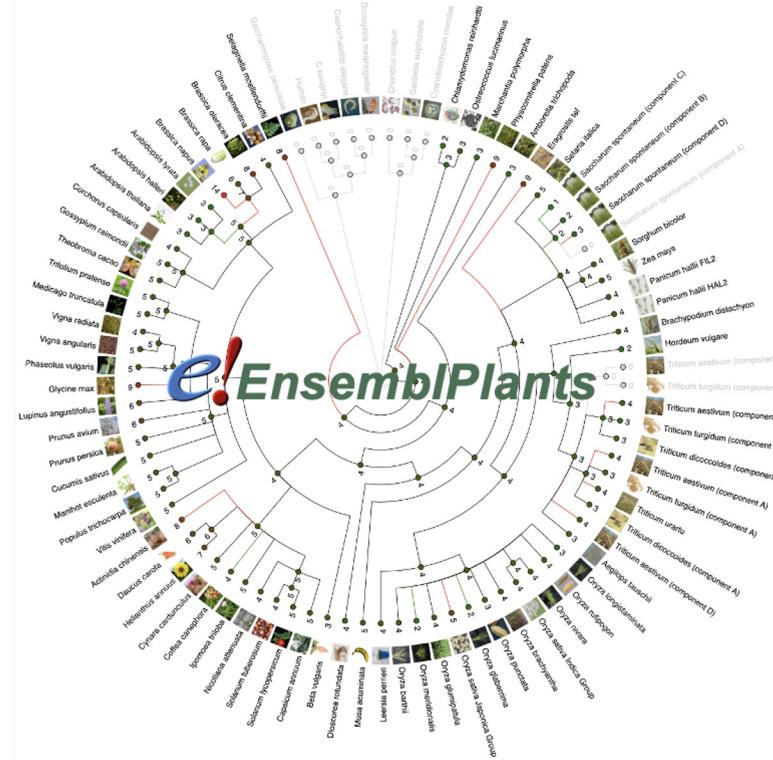
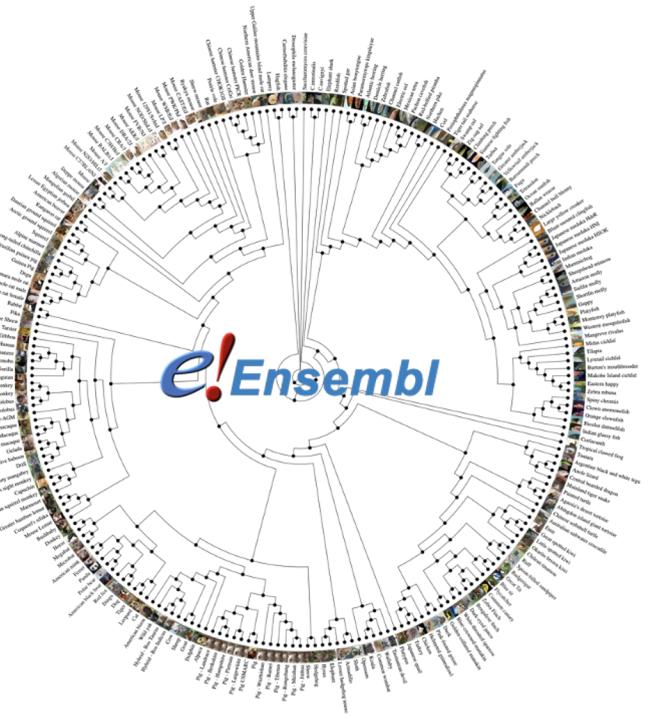
Monday, 3<sup>rd</sup> June 2024



# Overview

- Comparative genomics:
  - Gene trees
  - Homology predictions
- Whole-genome alignments (WGA):
  - Pairwise
  - Multiple
- Shared synteny

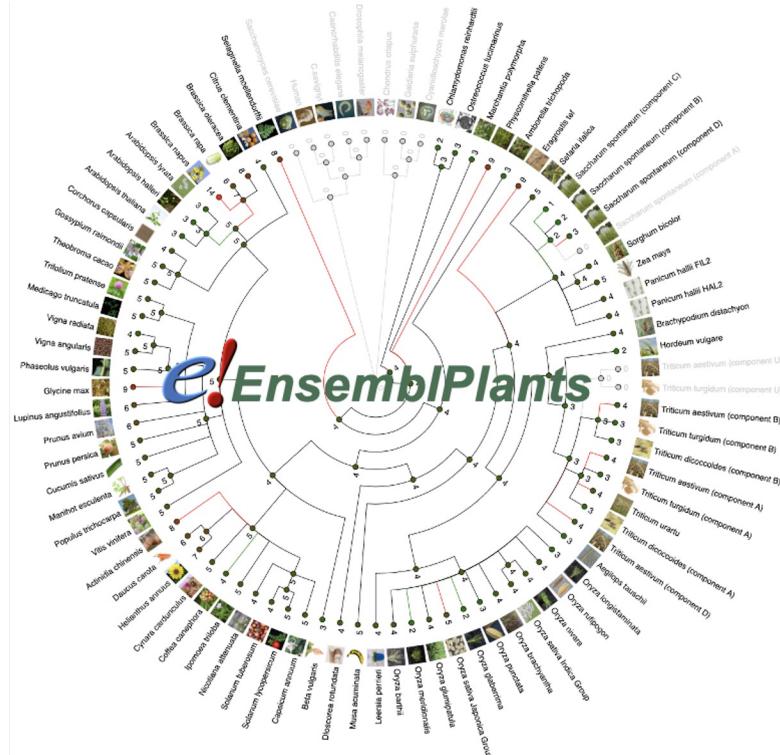
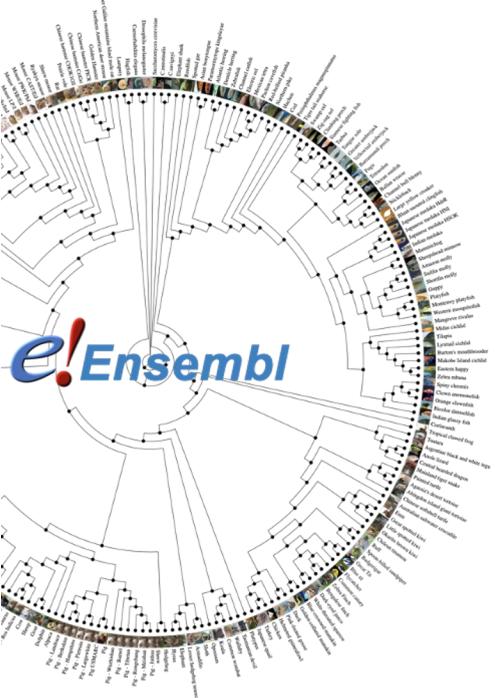
# Comparative analysis by taxa



# Comparative analysis by taxa

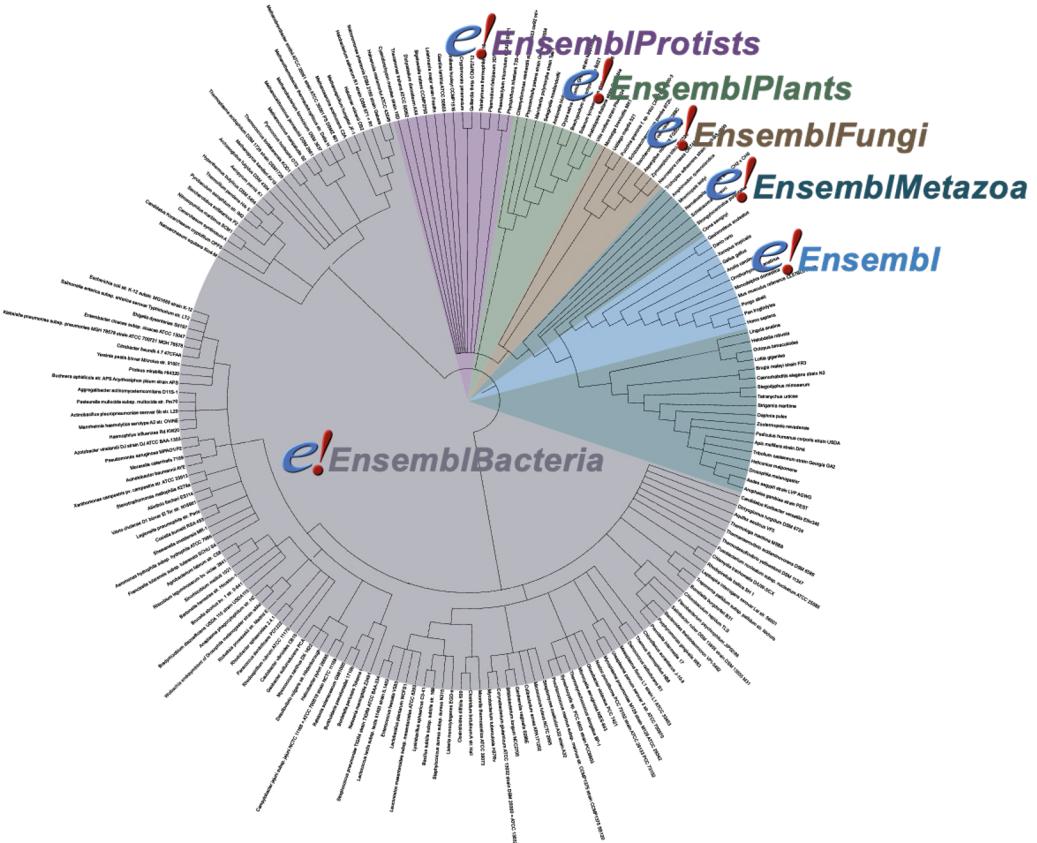
Gene-based displays

- Summary
  - Splice variants
  - Transcript comparison
  - Gene alleles
- Sequence
  - Secondary Structure
  - Gene families
  - Literature
- Fungal Compara
  - Genomic alignments
  - Gene tree
  - Gene gain/loss tree
  - Orthologues
  - Paralogues
- Pan-taxonomic Compara
  - Gene Tree
  - Orthologues
- Ontologies
  - GO: Biological process
  - GO: Molecular function
  - GO: Cellular component
  - PHI: Phibase identifier
- Phenotypes
- Genetic Variation
  - Variant table
  - Variant image
  - Structural variants
- Gene expression
- Pathway
- Molecular interactions
- Regulation
- External references
- Supporting evidence
- ID History
  - Gene history



<https://github.com/WCSCourses/Fungal2024>

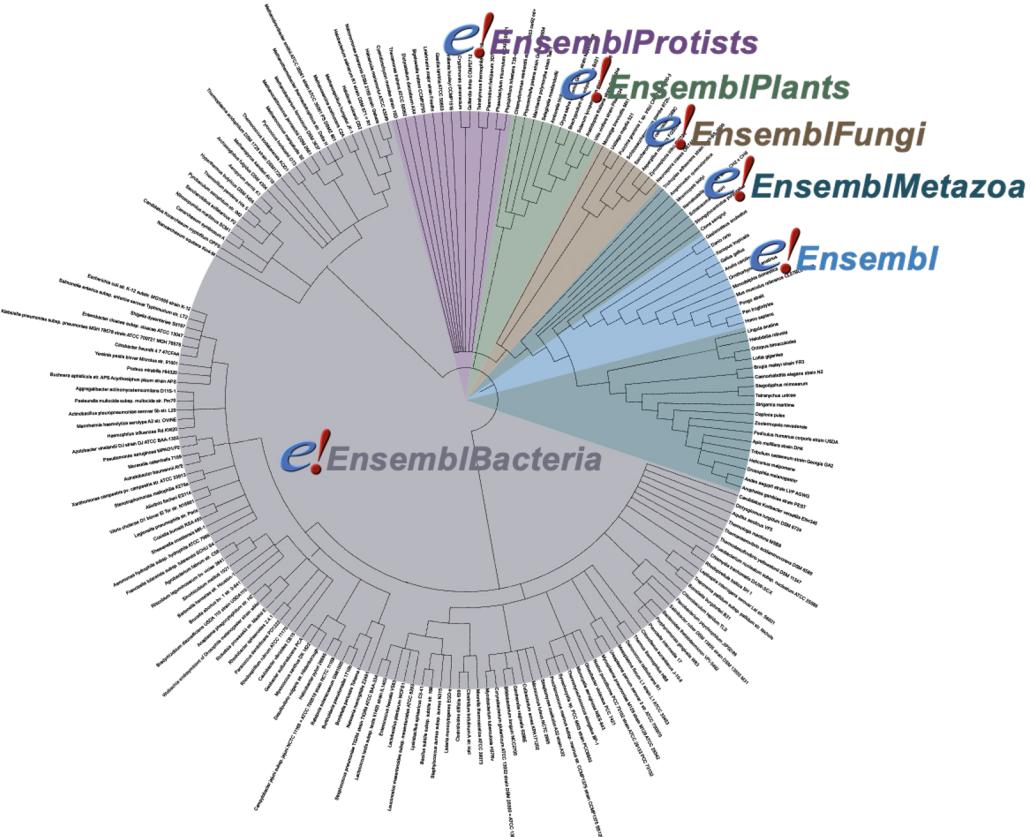
# Pan-taxonomic compara



# Pan-taxonomic compara

## Gene-based displays

- Summary**
    - Splice variants
    - Transcript comparison
    - Gene alleles
  - Sequence**
    - └ Secondary Structure
  - Gene families
  - Literature
  - Fungal Compara**
    - Genomic alignments
    - Gene tree
    - Gene gain/loss tree
    - Orthologues
    - Paralogous
  - Pan-taxonomic Compara**
    - Gene Tree
    - Orthologues
  - Ontologies**
    - GO: Biological process
    - GO: Molecular function
    - GO: Cellular component
    - PHI: Phibase identifier
  - Phenotypes**
  - Genetic Variation**
    - Variant table
    - Variant image
    - Structural variants
  - Gene expression
  - Pathway
  - Molecular interactions
  - Regulation
  - External references
  - Supporting evidence
  - ID History**
    - └ Gene history



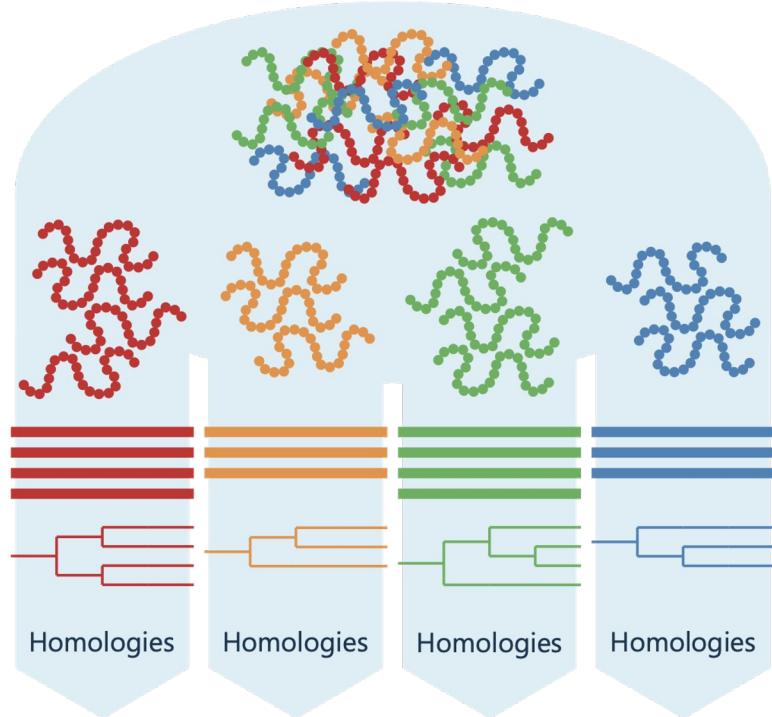
<https://github.com/WCSCourses/Fungal2024>

# Pan-taxonomic compara

Name	Classification	Taxon ID	Assembly
 <a href="#">Aspergillus nidulans</a>	Eurotiales	<a href="#">227321</a>	ASM1142v1
 <a href="#">Neurospora crassa</a>	Sordariales	<a href="#">367110</a>	NC12
 <a href="#">Puccinia graminis</a>	Pucciniales	<a href="#">418459</a>	ASM14992v1
 <a href="#">Saccharomyces cerevisiae</a>	Saccharomycetales	<a href="#">559292</a>	R64-1-1
 <a href="#">Schizosaccharomyces pombe</a>	Schizosaccharomycetales	<a href="#">284812</a>	ASM294v2
 <a href="#">Ustilago maydis</a>	Ustilaginales	<a href="#">237631</a>	Umaydis521_2.0
 <a href="#">Zymoseptoria tritici</a>	Capnodiales	<a href="#">336722</a>	MG2

# Gene / protein trees

1. Representative translation of each gene from all species
2. All-vs-all HMM search to classify into families or clustering
3. Multiple protein alignment
4. Phylogenetic tree for each aligned cluster and reconciliation against NCBI taxonomy
5. Ortho-/parologue inference



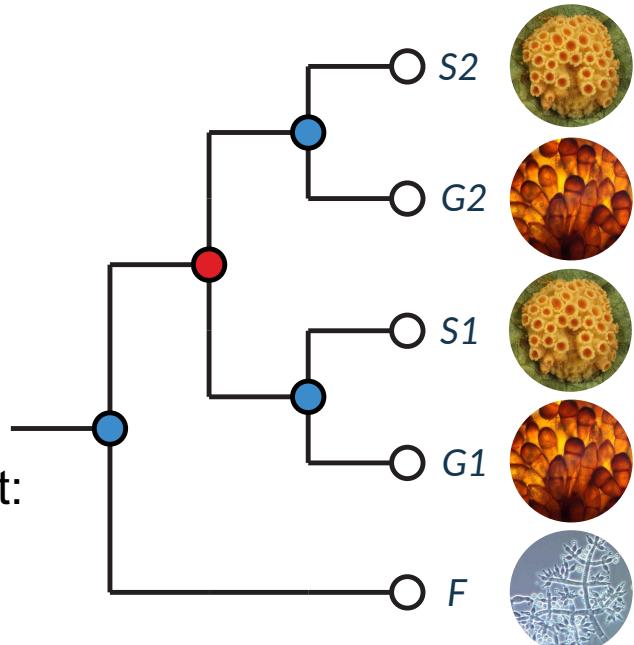
# Homology

Orthologues:

- Genes emerged through a **speciation** event:
  - e.g. G1 and S1; G2 and F; S2 and F
- 1-to-1: G1 and S1
- 1-to-many: F and S1, S2, G1, G2

Paralogues:

- Genes emerged through a **duplication** event:
  - e.g. G1 and G2, S1 and S2
  - Within species: G1 and G2
  - Between species: G1 and S2



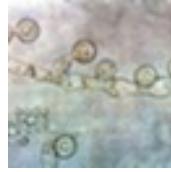
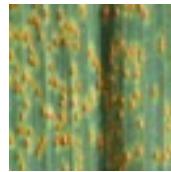
# Whole-genome alignments (WGA)

- To identify highly conserved regions
  - Sequences that evolve slowly
  - Regions likely to be functional
  - Both coding and non-coding
- To support problematic gene predictions
- To define syntenic regions
- Two types:
  - Pairwise
  - Multiple

# Pairwise alignments

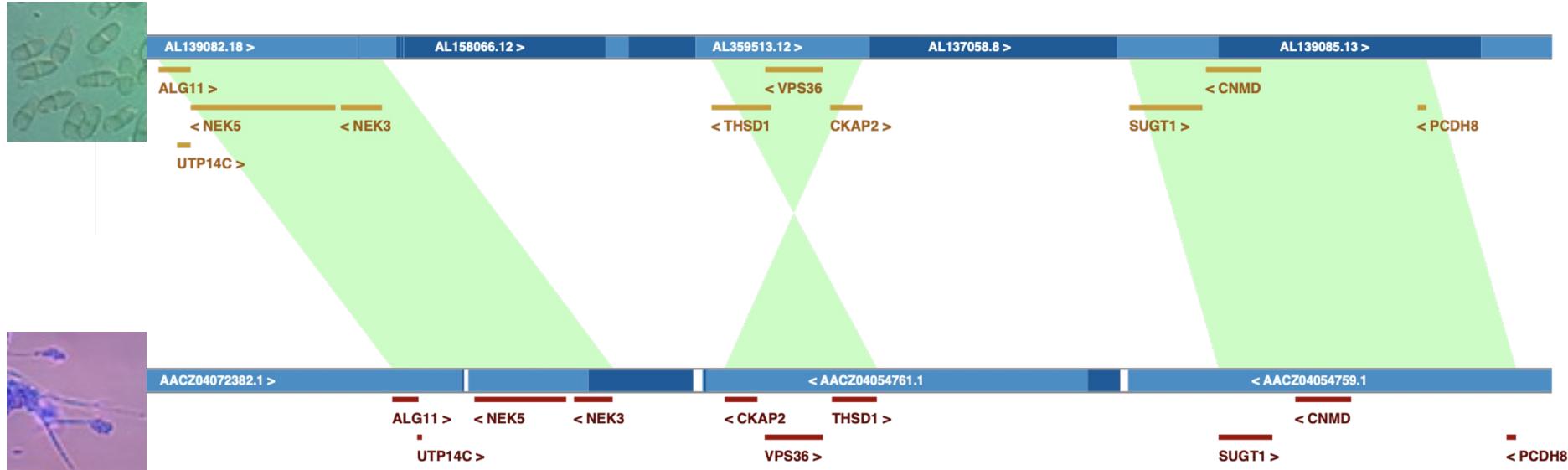
Pairwise alignments with BLASTZ (older) or LASTZ (newer) and tBLAT:

- Related species vs each other:  
Hypocreales, Magnaporthe, Puccinia, Schizosaccharomyces,  
Saccharomycetales



# Shared synteny

Conserved order of aligned homologous genomic blocks between species (irrespective of orientation):



# Types of data

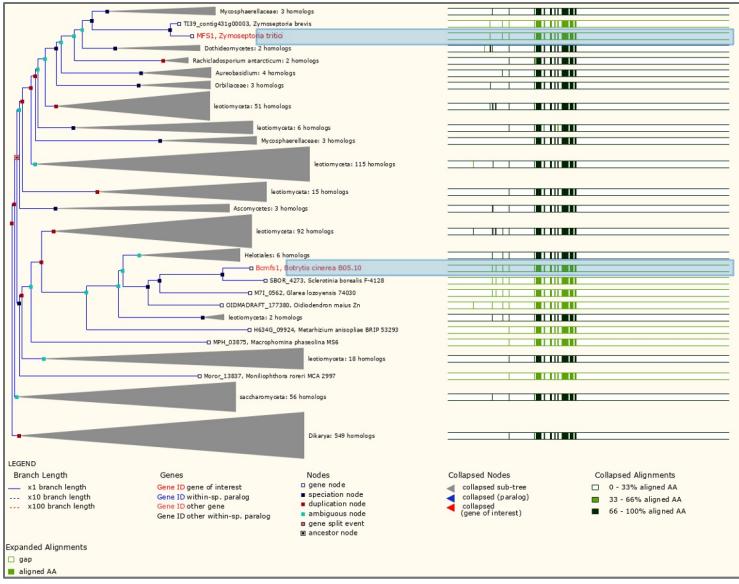
Gene-based resources  
(under the **Gene** tab):

- Phylogenetic trees and tree-inferred homology
  - Gene / protein trees
  - ncRNA trees
- Stable ID mapping
- Protein families

Sequence-based resources  
(under the **Location** tab):

- Whole-genome alignments
  - Ancestral sequences
  - Age of base
  - Conservation scores and constrained elements
  - Syntenies

# Explore behaviour in other pathogens



*Botrytis cinerea* infections

Pest Manag Sci. 2013 Jun;69(6):667-78. doi: 10.1002/ps.3506. Epub 2013 Apr 10.

French vineyards provide information that opens ways for effective resistance management of *Botrytis cinerea* (grey mould).

Walker AS<sup>1</sup>

Appl Environ Microbiol. 2013 Jan;79(1):159-67. doi: 10.1128/AEM.02655-12. Epub 2012 Oct 19.

Gray mold populations in german strawberry fields are resistant to multiple fungicides and dominated by a novel clade closely related to *Botrytis cinerea*.

Appl Environ Microbiol. 2002 Oct;68(10):4996-5004.

**Bcmfs1**, a novel major facilitator superfamily transporter from *Botrytis cinerea*, provides tolerance towards the natural toxic compounds camptothecin and cercosporin and towards fungicides.

Hayashi K<sup>1</sup>, Schoonbeek HJ, De Waard MA.

# Ensembl Fungi: Ensembl VEP

**SNPs & variants  
Fungal Pathogen Genomics**

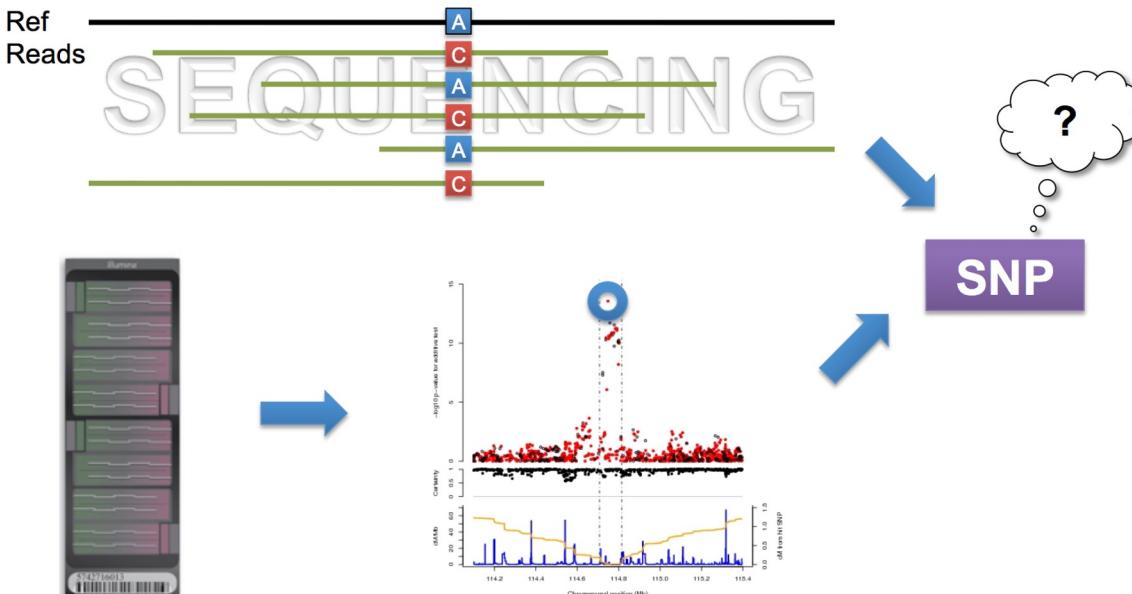
**Louisse Paola Mirabueno**  
Ensembl Outreach Officer

Wednesday, 5<sup>th</sup> June 2024



# What is the Variant Effect Predictor (VEP) for?

A tool to predict and annotate the **functional consequences** of variants (SNPs, insertions, deletions, CNVs, or structural variants).



# What can you do with the VEP?

Data input:

Variant  
coordinates

VCF

HGVS

Variant IDs

SPDI



VEP output:

Genes,  
transcripts  
affected

Pathogenicity  
 

Frequency  
data

Regulatory  
and splicing  
consequences

PubMed  
citations

# Where does the data come from?

## The Ensembl variation process

Variant import

Quality control

Linked data

Ensembl analysis

# Input formats

Ensembl default

#chr	start	end	allele	strand	id
------	-------	-----	--------	--------	----

VEP input:

1	65568	65568	A/C	1
2	265023	265023	C/T	1
3	319781	319781	A/-	1

VCF:

#chr	pos	id	ref	alt	qual	filt	info
1	65568	.	A	C	.	.	.
2	265023	.	C	T	.	.	.
3	319780	.	GA	G	.	.	.

# Input formats

**HGVS notation:** ENST00000641515.2:c.4A>C

ENST00000272065.10:c.43+16C>T

ENST00000256509.7:c.5del

9:g.22125504G>C

**Variant IDs:** rs1156485833

COSM327779

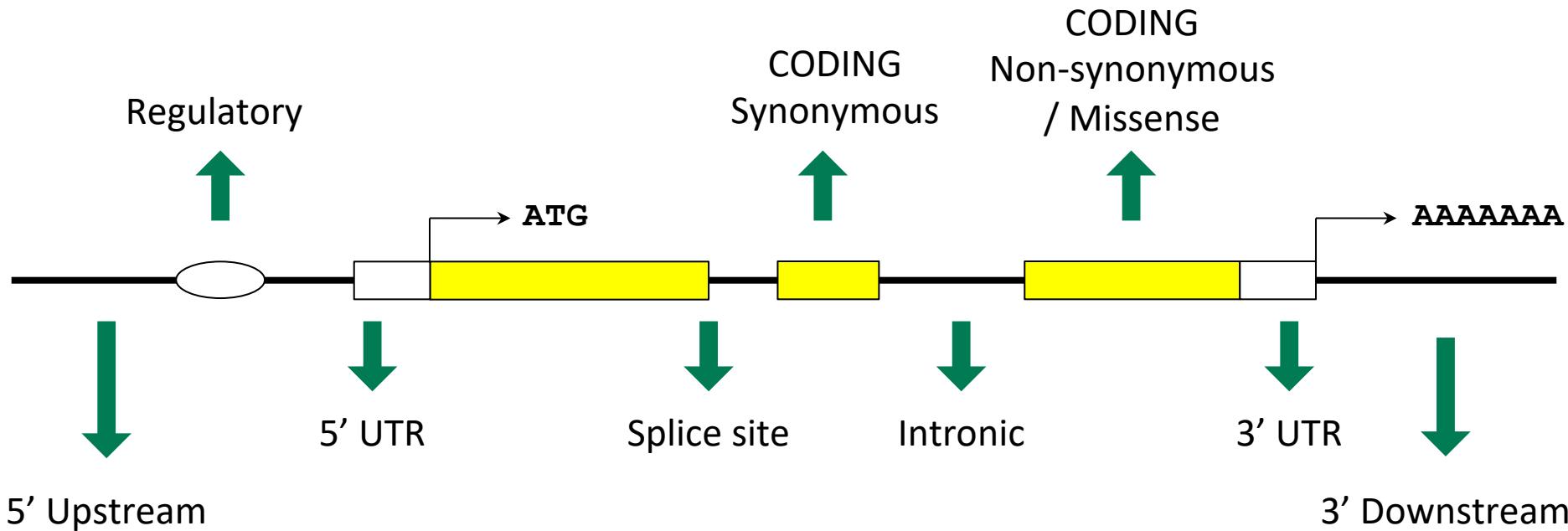
rs867704559

**SPDI:** NC\_00001.11:65567:A:C

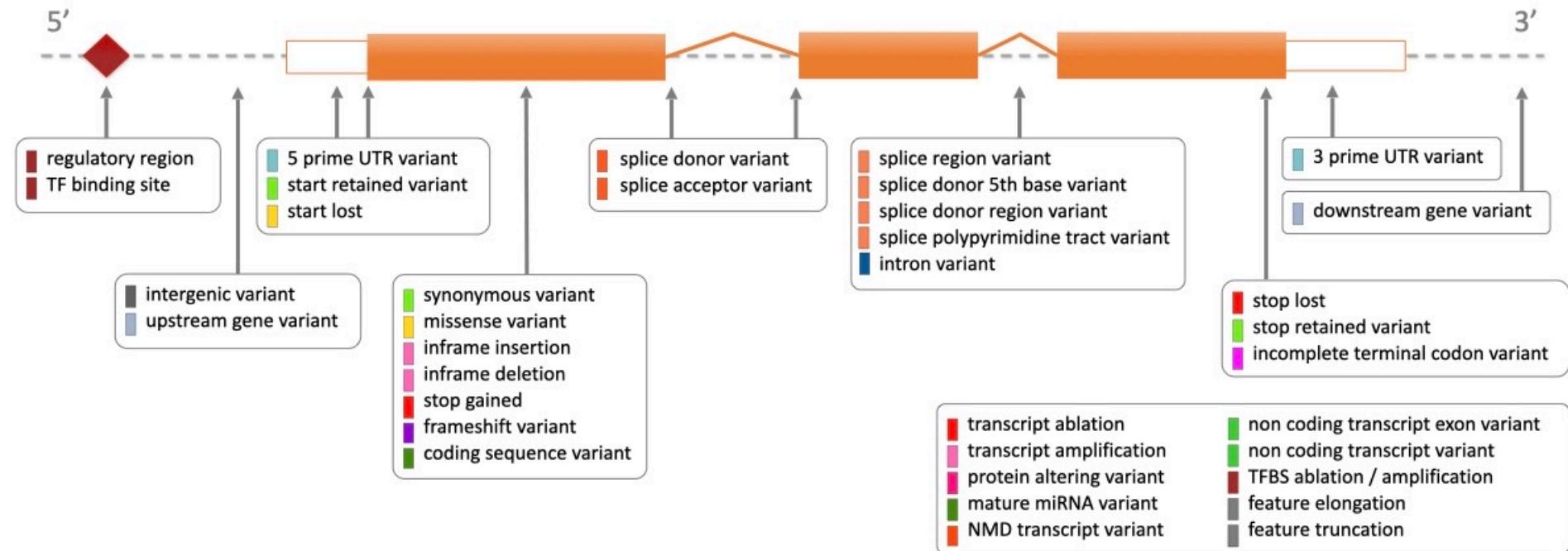
NC\_00002.12:265022:C:T

NC\_00003.12:319780:A:

# Variation consequences



# Sequence ontology (SO)



# Sequence ontology (SO)

SO term	SO description	SO accession	Display term	IMPACT
transcript_ablation	A feature ablation whereby the deleted region includes a transcript feature	<a href="#">SO_0001893</a>	Transcript ablation	HIGH
splice_acceptor_variant	A splice variant that changes the 2 base region at the 3' end of an intron	<a href="#">SO_0001574</a>	Splice acceptor variant	HIGH
splice_donor_variant	A splice variant that changes the 2 base region at the 5' end of an intron	<a href="#">SO_0001575</a>	Splice donor variant	HIGH
stop_gained	A sequence variant whereby at least one base of a codon is changed, resulting in a premature stop codon, leading to a shortened transcript	<a href="#">SO_0001587</a>	Stop gained	HIGH
frameshift_variant	A sequence variant which causes a disruption of the translational reading frame, because the number of nucleotides inserted or deleted is not a multiple of three	<a href="#">SO_0001588</a>	Frameshift variant	HIGH
stop_lost	A sequence variant where at least one base of the terminator codon (stop) is changed, resulting in an elongated transcript	<a href="#">SO_0001578</a>	Stop lost	HIGH
start_lost	A codon variant that changes at least one base of the canonical start codon	<a href="#">SO_0002012</a>	Start lost	HIGH
transcript_amplification	A feature amplification of a region containing a transcript	<a href="#">SO_0001889</a>	Transcript amplification	HIGH
inframe_insertion	An inframe non synonymous variant that inserts bases into in the coding sequence	<a href="#">SO_0001821</a>	Inframe insertion	MODERATE
inframe_deletion	An inframe non synonymous variant that deletes bases from the coding sequence	<a href="#">SO_0001822</a>	Inframe deletion	MODERATE
missense_variant	A sequence variant, that changes one or more bases, resulting in a different amino acid sequence but where the length is preserved	<a href="#">SO_0001583</a>	Missense variant	MODERATE
protein_altering_variant	A sequence variant which is predicted to change the protein encoded in the coding sequence	<a href="#">SO_0001818</a>	Protein altering variant	MODERATE
splice_region_variant	A sequence variant in which a change has occurred within the region of the splice site, either within 1-3 bases of the exon or 3-8 bases of the intron	<a href="#">SO_0001630</a>	Splice region variant	LOW
splice_donor_5th_base_variant	A sequence variant that causes a change at the 5th base pair after the start of the intron in the orientation of the transcript	<a href="#">SO_0001787</a>	Splice donor 5th base variant	LOW
splice_donor_region_variant	A sequence variant that falls in the region between the 3rd and 6th base after splice junction (5' end of intron)	<a href="#">SO_0002170</a>	Splice donor region variant	LOW
splice_polyprymidine_tract_variant	A sequence variant that falls in the polyprymidine tract at 3' end of intron between 17 and 3 bases from the end (acceptor -3 to acceptor -17)	<a href="#">SO_0002169</a>	Splice polyprymidine tract variant	LOW
incomplete_terminal_codon_variant	A sequence variant where at least one base of the final codon of an incompletely annotated transcript is changed	<a href="#">SO_0001626</a>	Incomplete terminal codon variant	LOW
start_retained_variant	A sequence variant where at least one base in the start codon is changed, but the start remains	<a href="#">SO_0002019</a>	Start retained variant	LOW
stop_retained_variant	A sequence variant where at least one base in the terminator codon is changed, but the terminator remains	<a href="#">SO_0001567</a>	Stop retained variant	LOW
synonymous_variant	A sequence variant where there is no resulting change to the encoded amino acid	<a href="#">SO_0001819</a>	Synonymous variant	LOW
coding_sequence_variant	A sequence variant that changes the coding sequence	<a href="#">SO_0001580</a>	Coding sequence variant	MODIFIER
mature_miRNA_variant	A transcript variant located with the sequence of the mature miRNA	<a href="#">SO_0001620</a>	Mature miRNA variant	MODIFIER
5_prime_UTR_variant	A UTR variant of the 5' UTR	<a href="#">SO_0001623</a>	5 prime UTR variant	MODIFIER
3_prime_UTR_variant	A UTR variant of the 3' UTR	<a href="#">SO_0001624</a>	3 prime UTR variant	MODIFIER
non_coding_transcript_exon_variant	A sequence variant that changes non-coding exon sequence in a non-coding transcript	<a href="#">SO_0001792</a>	Non coding transcript exon variant	MODIFIER
intron_variant	A transcript variant occurring within an intron	<a href="#">SO_0001627</a>	Intron variant	MODIFIER
NMD_transcript_variant	A variant in a transcript that is the target of NMD	<a href="#">SO_0001621</a>	NMD transcript variant	MODIFIER
non_coding_transcript_variant	A transcript variant of a non coding RNA gene	<a href="#">SO_0001619</a>	Non coding transcript variant	MODIFIER
upstream_gene_variant	A sequence variant located 5' of a gene	<a href="#">SO_0001631</a>	Upstream gene variant	MODIFIER
downstream_gene_variant	A sequence variant located 3' of a gene	<a href="#">SO_0001632</a>	Downstream gene variant	MODIFIER
TFBS_ablation	A feature ablation whereby the deleted region includes a transcription factor binding site	<a href="#">SO_0001895</a>	TFBS ablation	MODIFIER
TFBS_amplification	A feature amplification of a region containing a transcription factor binding site	<a href="#">SO_0001892</a>	TFBS amplification	MODIFIER
TF_binding_site_variant	A sequence variant located within a transcription factor binding site	<a href="#">SO_0001782</a>	TF binding site variant	MODIFIER
regulatory_region_ablation	A feature ablation whereby the deleted region includes a regulatory region	<a href="#">SO_0001894</a>	Regulatory region ablation	MODIFIER
regulatory_region_amplification	A feature amplification of a region containing a regulatory region	<a href="#">SO_0001891</a>	Regulatory region amplification	MODIFIER
feature_elongation	A sequence variant that causes the extension of a genomic feature, with regard to the reference sequence	<a href="#">SO_0001907</a>	Feature elongation	MODIFIER
regulatory_region_variant	A sequence variant located within a regulatory region	<a href="#">SO_0001566</a>	Regulatory region variant	MODIFIER
feature_truncation	A sequence variant that causes the reduction of a genomic feature, with regard to the reference sequence	<a href="#">SO_0001906</a>	Feature truncation	MODIFIER
intergenic_variant	A sequence variant located in the intergenic region, between genes	<a href="#">SO_0001628</a>	Intergenic variant	MODIFIER

# Species with variation data

Name	Classification	Taxon ID	Assembly
 <a href="#">Fusarium graminearum str. PH-1</a>	Fungi	<a href="#">5518</a>	RR1
 <a href="#">Fusarium oxysporum</a>	Hypocreales	<a href="#">426428</a>	FO2
 <a href="#">Puccinia graminis</a>	Pucciniales	<a href="#">418459</a>	ASM14992v1
 <a href="#">Puccinia graminis Ug99</a>	Pucciniales	<a href="#">1040990</a>	v1
 <a href="#">Saccharomyces cerevisiae</a>	Saccharomycetales	<a href="#">559292</a>	R64-1-1
 <a href="#">Schizosaccharomyces pombe</a>	Schizosaccharomycetales	<a href="#">284812</a>	ASM294v2
 <a href="#">Verticillium dahliae JR2</a>	Glomerellales	<a href="#">1202531</a>	VDAG_JR2v.4.0
 <a href="#">Zymoseptoria tritici</a>	Capnodiales	<a href="#">336722</a>	MG2

# Use the Ensembl VEP



## Web interface

- Point-and-click interface
- Suits smaller volumes of data

[Documentation](#)

[Launch the web interface](#)



## Standalone perl script

- More options, more flexibility
- For large volumes of data

[Documentation](#)

[Download latest version](#)



## REST API

- Language-independent API
- Simple URL-based queries
- GET single variants, POST many

[Documentation](#)

# Ensembl VEP command-line

- Can handle genome-wide data
- Can work 100% offline
- A lot of options and plugins
- Use your own genome (custom annotations)

# Ensembl VEP plugins

- Plugins add extra functionality to VEP
- They may extend, filter or manipulate the output of VEP
- Plugins may make use of external data or code
- Available on the web interface and command line



# Recommend us

## Browser course

One day course on the Ensembl browser, aimed at wet-lab scientists.

## REST API course

Half day course on the Ensembl REST API, aimed at bioinformaticians.

## Train the trainer course

One day course on delivering the Ensembl browser course.

We can teach an Ensembl course at any institute for free (plus trainers' expenses in high income countries).

[helpdesk@ensembl.org](mailto:helpdesk@ensembl.org)

<https://training.ensembl.org/hosting>



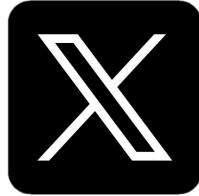
# Reach out to us



[/Ensembl.org](https://www.facebook.com/Ensembl.org)



[/EnsemblGenomeBrowser](https://www.linkedin.com/company/ensemblgenomrowser/)



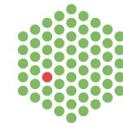
[@Ensembl](https://twitter.com/Ensembl)



[www.ensembl.info](https://www.ensembl.info)

# Help and documentation

Courses online <https://www.ebi.ac.uk/training/>



Tutorials <https://www.ensembl.org/info/website/tutorials/>

Flash animations <https://www.youtube.com/user/EnsemblHelpdesk>

<https://youku.com/u/UMzM1NjkzMTI0>



Job opportunities <https://www.ebi.ac.uk/careers>

Email [helpdesk@ensembl.org](mailto:helpdesk@ensembl.org)



Mailing lists [dev@ensembl.org](mailto:dev@ensembl.org)

[announce@ensembl.org](mailto:announce@ensembl.org)

# Publications

Harrison PW, Amode MR, Harrison PW, Austine-Orimoloye O et al. [Ensembl 2024](#). Nucleic Acids Research. 2024 Jan;52(D1):D891-D899. DOI: 10.1093/nar/gkad1049. PMID: 37953337; PMCID: PMC10767893.

Yates AD, Allen J, Amode RM, et al. [Ensembl Genomes 2022: an expanding genome resource for non-vertebrates](#). Nucleic Acids Research. 2022 Jan;50(D1):D996-D1003. DOI: 10.1093/nar/gkab1007. PMID: 34791415; PMCID: PMC8728113.

# Ensembl staff at FPG24



**Nishadi De Silva**  
Ensembl Microbes



**Manuel Carbajo**  
Ensembl Microbes



**Louisse Paola Mirabueno**  
Ensembl Outreach

# Acknowledgements and funding



National Human  
Genome  
Research Institute  
(NHGRI)

National Institute  
of Allergy and  
Infectious  
Diseases (NIAID)



Funded by the  
European Union

# Training materials

- Ensembl training materials are protected by a CC BY license:

<https://creativecommons.org/licenses/by/4.0/>



- If you wish to re-use these materials, please credit Ensembl for their creation.
- If you use Ensembl for your work, please cite our papers:

<https://www.ensembl.org/info/about/publications.html>