

Demo: Ensembl Fungi gene trees and homologues

Let's look at the homologues of *Saccharomyces cerevisiae* *YPR140W* aka Taz1. This gene is involved in stress response and conserved across different taxonomic domains. Search for the gene and go to the Gene tab.

Click on [Fungal compara: Gene tree](#), which will display the current gene in the context of a phylogenetic tree used to determine orthologues and paralogues.

GeneTree [ENSGT00930000005979](#) Unique gene tree ID

Summary statistics

- Number of genes: 186
- Number of speciation nodes: 158
- Number of duplication nodes: 18
- Number of ambiguous nodes: 9
- Number of gene split events: 0
- Highlight annotations: [Hide annotations table](#)

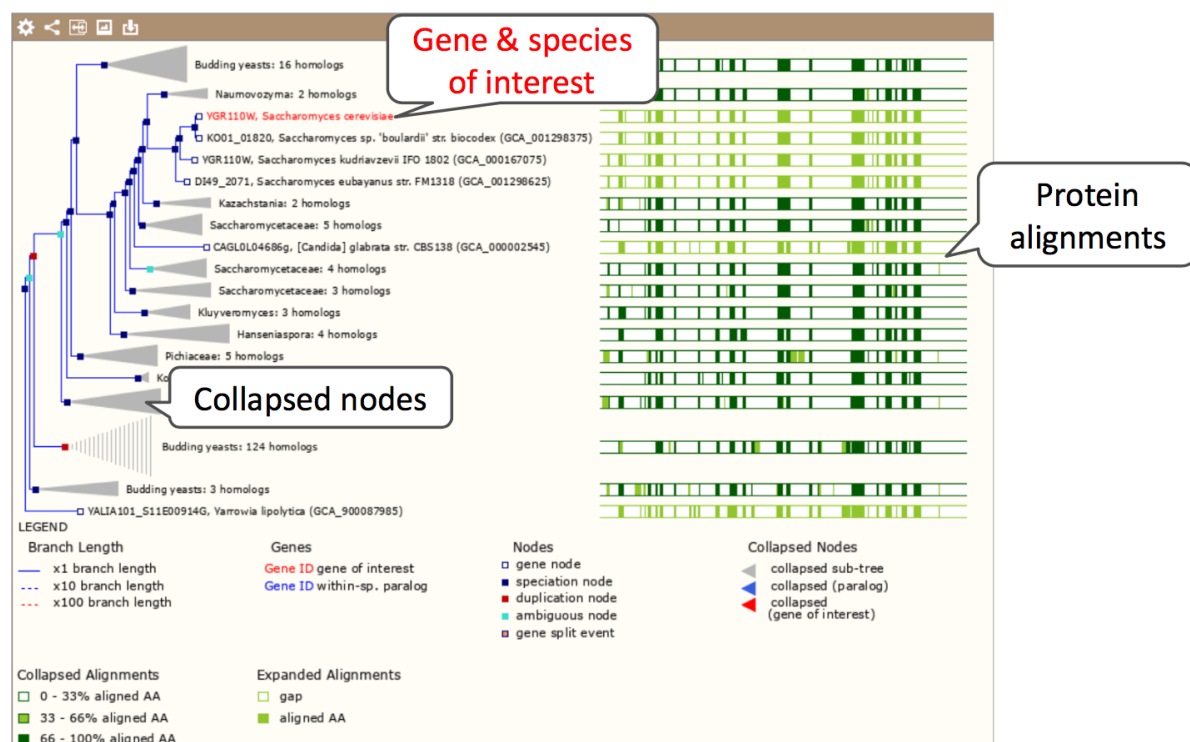
Show ☒ INTERPRO ☒ GO

Show 10 entries

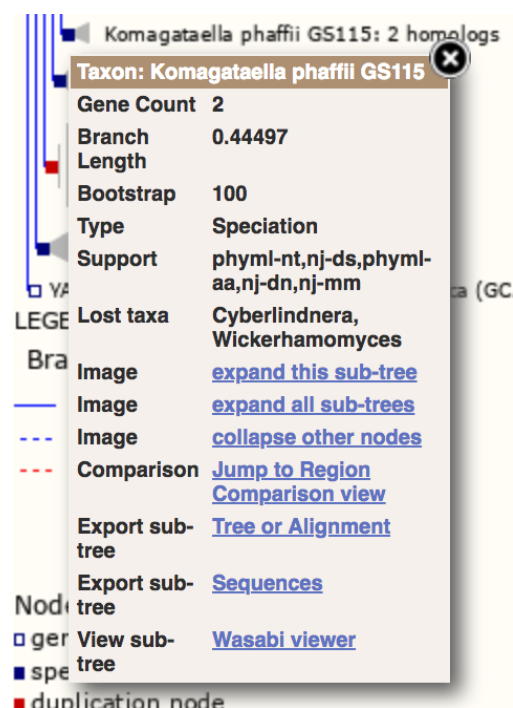
	Accession	Description
GO:0003674	GO:0003674	molecular_function
GO:0003824	GO:0003824	catalytic activity
GO:0008150	GO:0008150	biological_process
GO:0016787	GO:0016787	hydrolase activity
GO:0006629	GO:0006629	lipid metabolic process
GO:0006644	GO:0006644	phospholipid metabolic process
GO:0006650	GO:0006650	glycerophospholipid metabolic process
GO:0006793	GO:0006793	phosphorus metabolic process
GO:0006796	GO:0006796	phosphate-containing compound metabolic process
GO:0008152	GO:0008152	metabolic process

Showing 1 to 10 of 80 entries

Filter tree by **GO** terms or **Interpro** protein domains



Funnels indicate collapsed nodes. Click on a node (coloured square) to get a pop-up. We can then see what type of node this is, some statistics and options to expand or export the sub-tree.



There are some quick filtering options below the image, where you can add paralogues, and quickly expand or collapse nodes.


View options:

- [View current gene only](#) (Default)
- [View paralogs of current gene](#)
- [View all duplication nodes](#)
- [View fully expanded tree](#)
- Collapse all the nodes at the taxonomic rank

Use the 'configure page' link in the left panel to s

Ensembl release 92 - April 2018 © [EMBL-EBI](#)

✓ -- Select a rank--
Species
Genus
Family
Order
Class
Phylum
Kingdom

You can download the tree in a variety of formats. From the pop-up above you can click to export the sub-tree. Alternatively, click on the download icon  in the bar at the top of the image to get a pop-up where you can choose your format. You can preview this file before you download.

File name: TAZ1_gene_tree

File format: -- Choose Format --

[Preview](#) [Download](#) [Download Compressed](#)

Guide to file formats

CLUSTALW	FASTA	Mega
<pre>CLUSTAL W(1.81) Multiple sequence homo_sapiens/1-465588 CCTCAGGAC pan_troglodytes/1-465588 CCCAGGAC ** ***** homo_sapiens/1-465588 CCCAGTGCC pan_troglodytes/1-465588 CCCAGTGCC *****</pre>	<pre>>homo_sapiens/1-464388 CCTCAGGACCCAGCGCAACACCCAGAT CCAGTGCTCTGACCTCCCTCTGGGCG TGGGACAGAGAGAACACAGCTGGCTC AGGGCCCTGGTGGGGGTTAGATCAAA CCAGCTGGATCTGATATTGGCACCTC CCAGCTCTGTGCAAAAGGTGCTGTG AGGAAGACCGTGGCTCTGCTGTGTT AAAGATGGGGGGGTGCTGCTGCTCT GGGAGGGGGAAGAGAAAGGCCCTGGG PAACTCTGGGGCTGCTGCTGCTGCTG</pre>	<pre>#mega !Title: ProjectedMultiAlign !Format datatype=dna identical= #homo_sap CCTCAGGAC GACGCAAC #pan_trog ..C..... #homo_sap CCCAGTGCTCT GCGCTGCT #pan_trog #homo_sap TGGGACAGAG AGAGAACAC</pre>
MSF	Newick	Nexus
<pre>ProjectedMultiAlign MSF: 2 Type: Name: homo_sapiens/1-465588 Les Name: pan_troglodytes/1-465588 Les // homo_sapiens/1-465588 CCTCAGGAC G pan_troglodytes/1-465588 CCCAGGAC homo_sapiens/1-465588 GGGTCAACAC C</pre>	<pre>((((((((ENSRNOP00000015030_Trub_0.07 ENSRNOP00000002430_Trig_0.203149):0. ENSGMCP00000015199_Gacu_0.161942):0. ((ENSGPFP000000001573_Pfor_0.042925, ENSGNAP000000006983_Xmac_0.244447):0. ENSONIP000000006940_Osil_0.296133):0.0 ENSONIP000000004773_Olat_0.2502139):0. ENSGMCP00000010385_Omor_0.360666):0.1 ((ENSGMCP00000013440_Amek_0.474772, ENSGNAP00000009874_Drex_0.288918):0. ENSGMCP00000009962_Locu_0.219888):0. (((((((ENSGALP000000027524_Opal_0.03 ENSGNAP00000015940_Mual_0.2463749):0. ENSHAPL000000007411_Apia_0.122877):0.0</pre>	<pre>#NEXUS [TTITLE: ProjectedMultiAlign] begin data; dimensions ntax=2 nchar=465588; format interleave datatype=dna gap= matrix homo_sapiens CCTCAGGAC pan_troglodytes CCCAGGAC homo_sapiens GGGTCAACAC pan_troglodytes GGGCCACAC</pre>
NHX	OrthoXML	Pfam
<pre>{(((((((0.046083[64NHX:D=N:T=48 :0.065551[64NHX:D=N:T=8083])Poec :0.359035[64NHX:D=N:T=8128])Oval ((:0.077336[64NHX:D=N:T=31033], :0.099898[64NHX:D=N:T=99803])Tet :0.160116[64NHX:D=N:T=69293])Per :0.37365[64NHX:D=N:T=8049])Acant :0.780276[64NHX:D=N:T=8090])Acac :0.44137[64NHX:D=N:T=7994], :0.582768[64NHX:D=N:T=7955])Otop :n 225188[64NHX:T=N:T=75181]Moon</pre>	<pre><?xml version="1.0" encod <orthoXML xsi:schemaLocat xmlns:xsi="http://www.w3. <species NCBI:TaxId="925 <database name="Unkno <genes> <gene id="6053741 <gene id="5945247</pre>	<pre>homo_sapiens/1-465588 CCTCAGGA pan_troglodytes/1-465588 CCCAGGAC</pre>

We can look at homologues in the [Orthologues](#) and [Paralogues](#) pages, which can be accessed from the left-hand menu. If there are no orthologues or paralogues, then the name will be greyed out. Click on [Orthologues](#) to see the orthologues available.

Orthologues

[Download orthologues](#)

Summary of orthologues of this gene [Hide](#)

Click on 'Show details' to display the orthologues for one or more groups of species. Alternatively, click on 'Cont' species.

Orthologue types (mouse over for definitions)

Summary of orthologues by taxonomic group

Species set	Show details	With 1:1 orthologues	With 1:many orthologues	With many:many orthologues	Without orthologues
All (487 species)	<input checked="" type="checkbox"/>	53	6	0	428
Acidomyces (1 species)	<input type="checkbox"/>	0	0	0	1
Agaricales (24 species)	<input type="checkbox"/>	0	0	0	24
Atheliales (2 species)	<input type="checkbox"/>	0	0	0	2
Blastocladales (2 species)	<input type="checkbox"/>	0	0	0	2
Boletales (10 species)	<input type="checkbox"/>	0	0	0	10
Botryosphaerales (4 species)	<input type="checkbox"/>	0	0	0	4
Cantharellales (3 species)	<input type="checkbox"/>	0	0	0	3
Cantharellales (14 species)	<input type="checkbox"/>	0	0	0	14

Choose a taxon of interest

Similarity metrics

Orthologue details per species

Filter table

Species	Type	Orthologue	dN/dS	Target %id	Query %id	GOC Score	WGA
[Candida] auris str. 6684 (GCA_001189475)	1-to-1	QG37_03321 View Gene Tree Compare Regions (Scaffold21:284,573-285,871:1) View Sequence Alignments	n/a	37.96 %	36.85 %	n/a	n/a
[Candida] glabrata str. CBS138 (GCA_000002545)	1-to-1	CAGL0L04686g View Gene Tree Compare Regions (L:539,203-540,558:1) View Sequence Alignments	n/a	46.34 %	46.97 %	n/a	n/a
Ascoidea rubescens DSM 1968 (GCA_001661345)	1-to-1	ASCRUDRAFT_74793 View Gene Tree Compare Regions (KV454477:653,396-655,165:-1) View Sequence Alignments	n/a	29.16 %	33.48 %	n/a	n/a

Scroll to the bottom of the page to see a list of the species that do not have any orthologues with taz1 in *Saccharomyces cerevisiae*... there's a lot!

Species without orthologues


460 species are not shown in the table above because they don't have any orthologue with YPR140W.

- Absidia glauca
- Acidomyces richmondensis BFW
- Acremonium chrysogenum ATCC 11550
- Agaricus bisporus var. burnettii JB137-S8
- Allomyces macrogynus ATCC 38327
- Alternaria alternata
- Amanita muscaria Koide BX008

Saccharomyces cerevisiae is part of Pan-compara, which compares a subset of fungal species with species from other taxa, such as plants, bacteria and vertebrates.

Click on [Pan-taxonomic compara > Orthologues](#). Let's see if there are any orthologues of this gene in plants. Click the 'Show details' box for plants.

Species set	Show details	With 1:1 orthologues	With 1:many orthologues	With many:many orthologues	Without orthologues
All (187 species)	<input type="checkbox"/>	8	1	1	177
Vertebrates (0 species)	<input type="checkbox"/>	0	0	0	0
Metazoa (22 species)	<input type="checkbox"/>	0	0	0	22
Plants (9 species)	<input checked="" type="checkbox"/>	2	0	0	7
Fungi (6 species)	<input type="checkbox"/>	5	0	0	1
Protists (14 species)	<input type="checkbox"/>	1	1	1	11
Bacteria (98 species)	<input type="checkbox"/>	0	0	0	98
Archaea (25 species)	<input type="checkbox"/>	0	0	0	25

Selected orthologues [Hide](#) 

Show All entries		Show/hide columns		Filter					
Species	Type	Orthologue	dN/dS	Target %id	Query %id	GOC Score	WGA Coverage	High Confidence	
Chlamydomonas reinhardtii	1-to-1	CHLRE_01g000300v5	n/a	13.07 %	16.63 %	n/a	n/a	No	
	View Gene Tree	1:54,136-57,919:1							
		View Sequence Alignments							
Cyanidioschyzon merolae	1-to-1	CML191C	n/a	20.09 %	19.55 %	n/a	n/a	No	
	View Gene Tree	12:479,112-480,413:1							
		View Sequence Alignments							

There are two orthologues in plants, both are species of algae.

Exercises: Ensembl Fungi gene trees, homologues and alignments

Compara Exercise 1 – Orthologues for the *Schizosaccharomyces pombe mcm6* gene

(a) How many orthologues are predicted for this gene in Fungal Compara? What about in Pan-compara?

(b) Filter the second table to view the human orthologue. How much sequence identity does the human protein have to the pombe one? Click on the [Alignment](#) link next to the [Ensembl identifier](#) column to view a protein alignment in Clustal format.

Compara Exercise 2 - *Zymoseptoria* orthologues

Exploring an orthologue that we identified using BioMart, exercise 2. We identified 6 genes with a pathogenic phenotype of 'Loss of pathogenicity' in *Zymoseptoria tritici*. We then found a single gene low confidence gene orthologue in *Cryptococcus neoformans* which we will now explore further.

Search for CNC06590 in *Cryptococcus neoformans* var. *Neoformans* JEC21 to go to the gene page. Click on the gene ID [CNC06590](#) to go to the gene page.

- Does this gene in *C. neoformans* have a PHI-Base annotation?
- Is it the same as that in *Z. tritici* (loss of pathogenicity)? What host species was this linked to?
- Find the *Z. tritici* orthologue in the [Orthologues](#) page and view a protein alignment.
- At which end of the protein (3' or 5') does the alignment between these two genes become worse?

Compara Exercise 3 - Mushroom genes

We're going to take a look at the gene CC1G_05700 in *Coprinopsis cinerea* okayama7#130.

(a) From the gene tab, click to view the [Gene tree](#). At the bottom of the image click to collapse all the nodes at the taxonomic rank of [Class](#).

- What do you notice about the types of fungi shown in the gene tree?
- Does this match with what you would expect from the gene description?
- Based on the protein alignment shown at the right, can you predict which end of the gene/protein is most conserved?

(b) Click to view the [Orthologues](#) page. In the Selected orthologues table, find the entry for the species *Amanita thiersii* and click to view a protein alignment.

- Does this support your conclusion about the conserved region of the gene/protein?