



# The philosophy behind PlasmoDB

first public release in June 2000

- Rapid access to unfinished data.** Unfinished data is critical for understanding large eukaryotic genomes.
- Curated annotation and analyses.** Predicted gene functions, pathways, and interactions.
- Integration of diverse data sources.** Genome & EST sequences, other organisms and related taxa, gene expression analysis, publications, clinical data.

- Enable laboratory researchers to ask their own questions.** An Encyclopedia of Malaria is not the goal: correct answers are less important than defining hypotheses that can be tested in the laboratory.
- Worldwide distribution, universal access.** 3,000+ hits/day, around the clock, from more than 50 countries. CD-ROM distribution for investigators without reliable high-speed Internet access.

VOLUME 31 NUMBER 1 DATABASE ISSUE JANUARY 1, 2003

NAR  
ONLINE  
<http://www.nar.oxjournals.org>

**Nucleic Acids Research**

Welcome to PlasmoDB 4.0!

PlasmoDB: The Plasmodium Genome Resource

PlasmoDB: The Plasmodium Genome Resource

Community Data

Genomic View

Gene Summary View

Query Interface

Expression Analysis

VEuPathDB

Eukaryotic Pathogen, Vector & Host Informatics Resources



## Search for...

[expand all](#) | [collapse all](#)

Filter the searches below... ?

- Genes
- Organisms
- Popset Isolate Sequences
- Genomic Sequences
- Genomic Segments
- ESTs

**BILL & MELINDA GATES foundation****wellcome trust****BURROUGHS WELLCOME FUND****USDA**  
**CIENCIA SEM FRONTEIRAS**  
**TDR** For research on diseases of poverty  
UNICEF • UNDP • World Bank • WHO

## The philosophy behind VEuPathDB databases ... 25 years on!

- ❖ **Curated annotation, integrated with automated analyses, in a semantically-defined context to facilitate cross-study analysis.** Predicted genes, proteins, motifs, structures, functions, locations, pathways, etc ... for diverse microbes and their host/vector species.
- ❖ **Incorporation & integration of diverse data types:** genome sequences, variation data, field & clinical isolates (with relevant metadata), chromatin marks, transcriptomic & proteomic profiling (multiple platforms), localization data, interactomes, pathways, phenotypes, structural genomics, metabolomics, comparative genomics & orthology-based inference, etc.
- ❖ **Rapid, FAIR access to both finished & draft datasets.** Unfinished (and unpublished) data is increasingly important, as genomic-scale datasets are typically incomplete or incompletely analyzed; most VEuPathDB datasets are now deposited and/or released *prior* to publication ... *but depositors always control release schedules.*
- ❖ **Robust, sustainable production resources, free universal access.** ~35K citations to date; ~10M hits/month from >50K unique users in >80 countries (the average user returns approx weekly); ~12 Tb/yr downloads; extensive scientific outreach & education activities
- ❖ **Provide researchers with the data and tools to ask their own questions!**



COMMUNITY CHAT



# Production Resources, Professional Staff

My Strategies Searches Tools My Workspace Data About Help Contact Us



## Alternative Career Paths in Biology & Bioinformatics:

- Independent faculty
- Collaborative faculty
- Research faculty
- Bioinformatics/Data wrangling
- Software development
- Data visualization
- Data management
- Biocuration & annotation
- Ontology development
- Scientific outreach/education
- Scientific project management

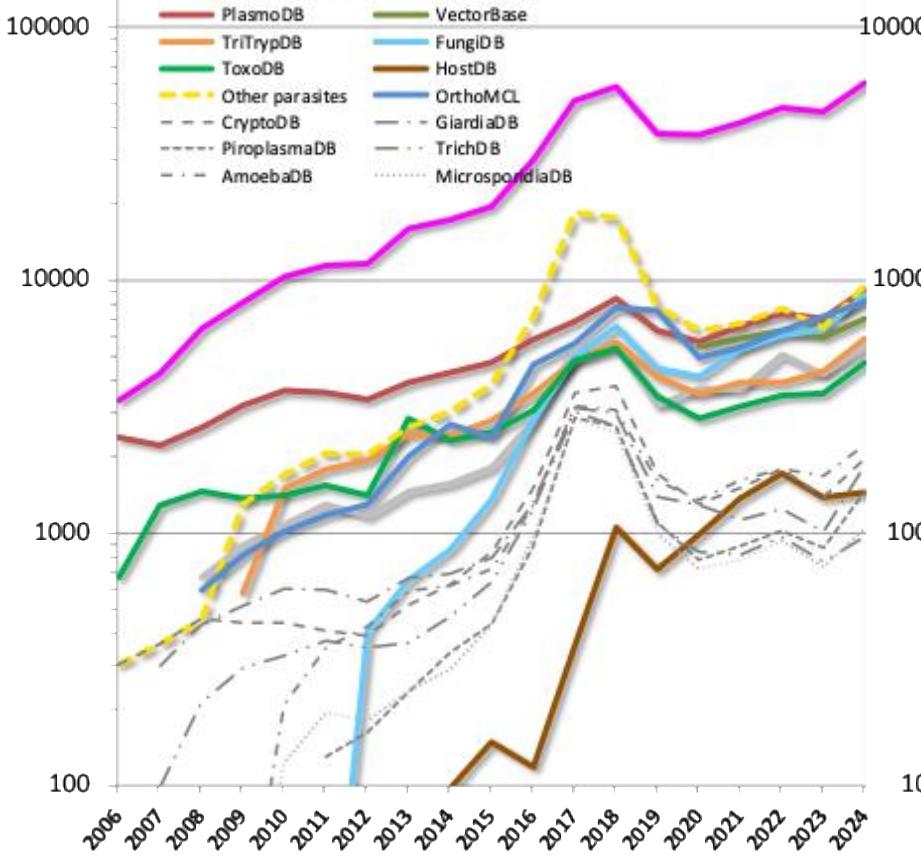


# Usage Statistics

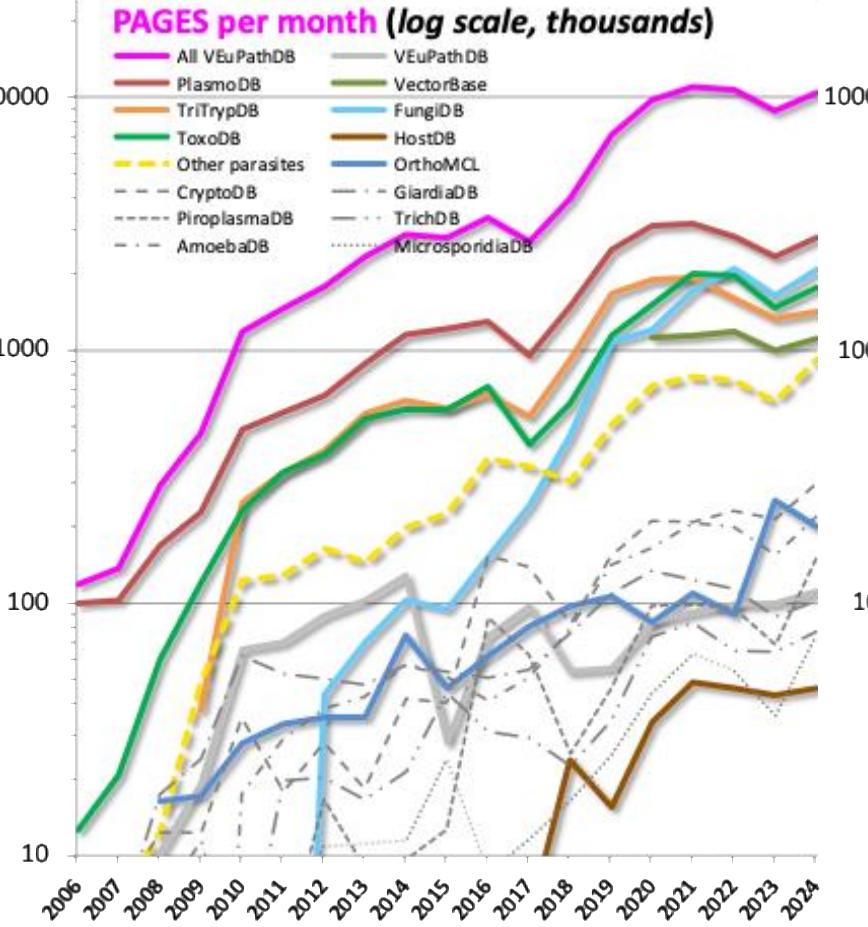
My Strategies Searches Tools My Workspace Data About Help Subscriptions Contact Us

Guest

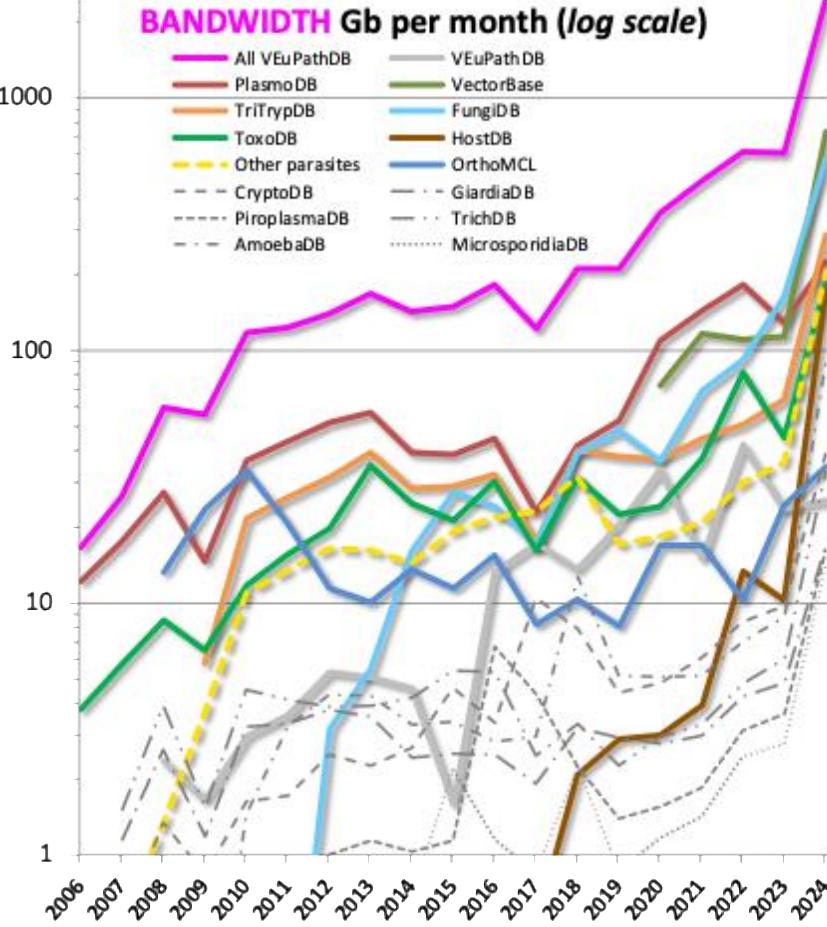
Unique USERS per month (*log scale*)



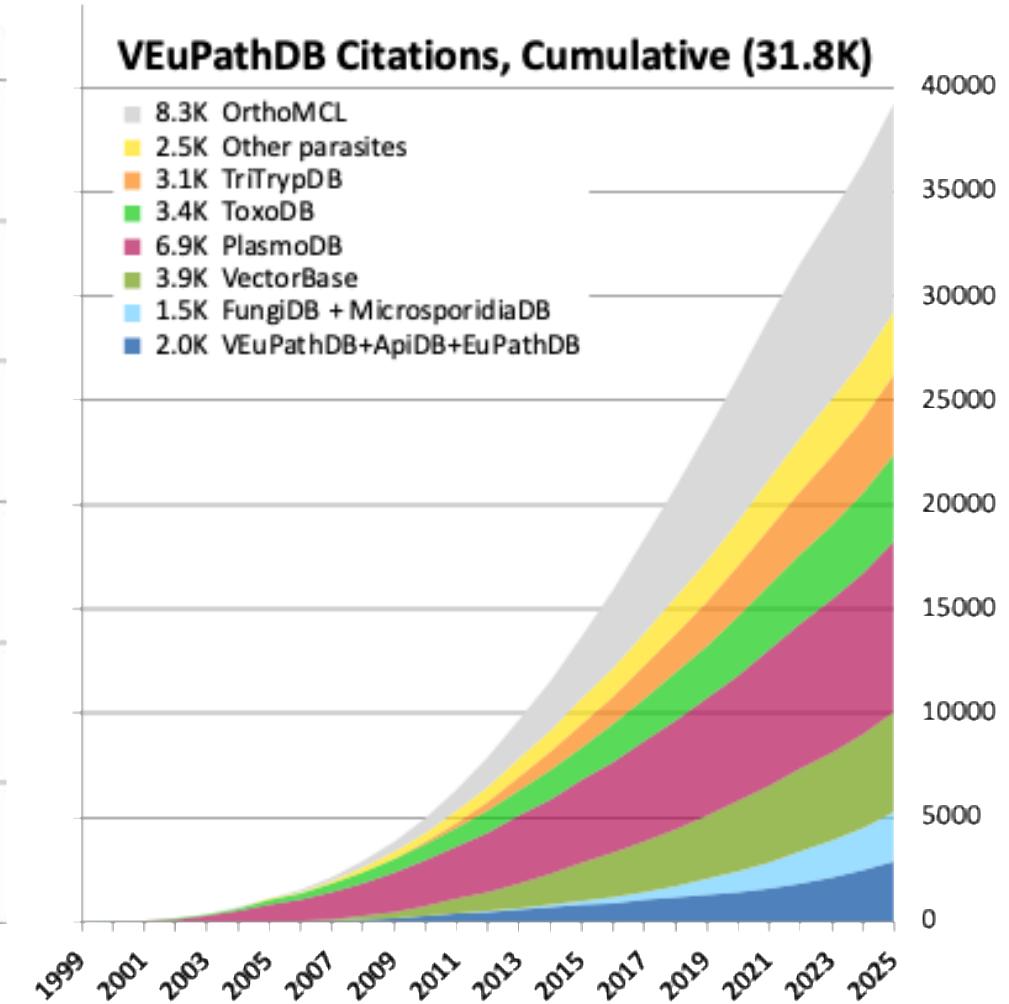
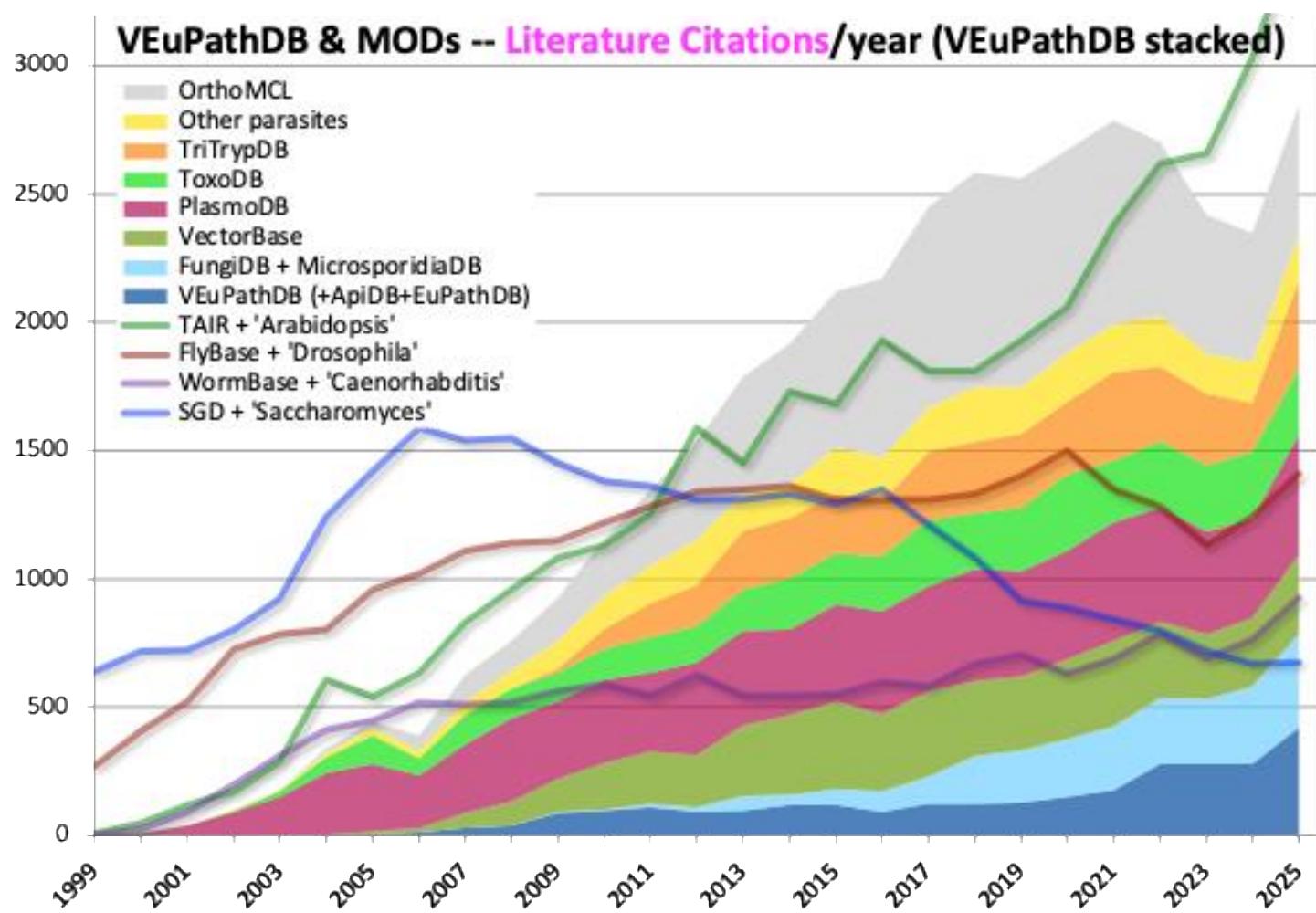
PAGES per month (*log scale, thousands*)



BANDWIDTH Gb per month (*log scale*)



## **Literature Citations**





VEuPathDB  
Release 68  
7 May 2024

Eukaryotic Pathogen, Vector & Host Informatics Resources

# Global Engagement

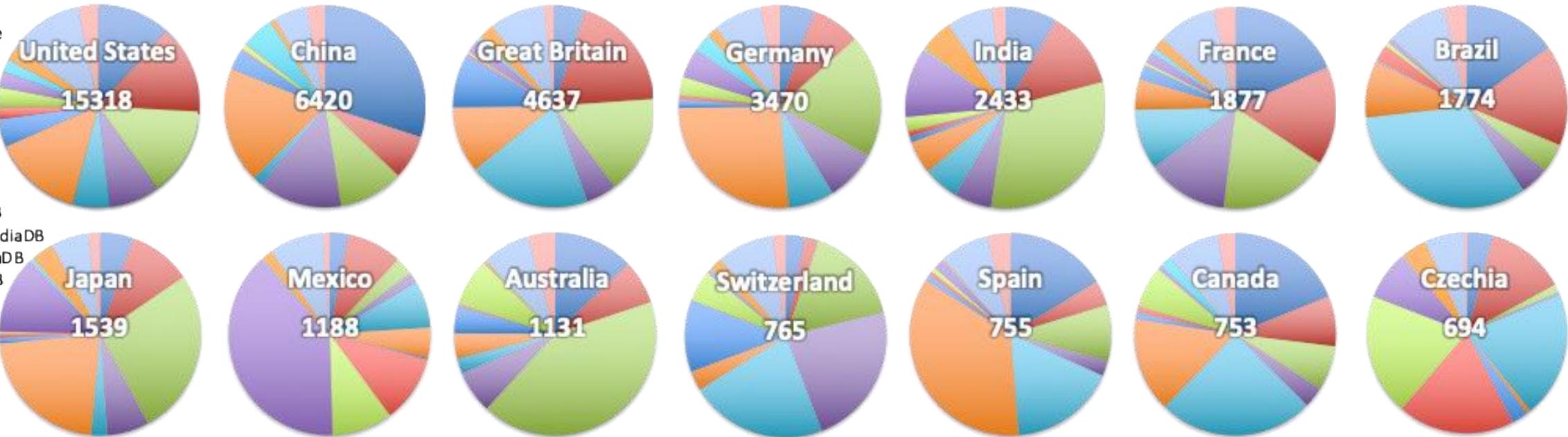


My Strategies Searches Tools My Workspace Data About Help Subscriptions Contact Us



Guest

- OrthoMCL
- VectorBase
- PlasmoDB
- ToxoDB
- TriTrypDB
- FungiDB
- CryptoDB
- TrichDB
- GiardiaDB
- AmoebaDB
- MicrosporidiaDB
- PiroplasmaDB
- VEuPathDB
- HostDB



GLOBAL  
CORE  
BIODATA  
RESOURCE





 **FungiDB**  
Fungal & Oomycete Informatics Resources

## Search for...

expand all | collapse all

Filter the searches below...

## ▼ Genes

- ▶ Annotation, curation and identifiers
- ▶ Epigenomics
- ▶ Function prediction
- ▶ Gene models
- ▶ Genetic variation
- ▶ Genomic Location
- ▶ Immunology
- ▶ Orthology and synteny
- ▶ Pathways and interactions
- ▶ Phenotype
- ▶ Protein features and properties
- ▶ Protein targeting and localization
- ▶ Proteomics
- ▶ Sequence analysis
- ▶ Structure analysis
- ▶ Taxonomy
- ▶ Text
- ▶ Transcriptomics

## ▶ Organisms

- ▶ Popset Isolate Sequences

## ▶ Genomic Sequences

## ▶ Genomic Segments

## ▶ SNPs

## ▶ ESTs



- ❖ **FungiDB is your resource:** tutorials & other on-line materials, webinars & workshops, 'Contact Us' link, virtual meetings w/ your lab, **Visit the help desk during poster sessions**
- ❖ **Help us to identify & prioritize key datasets (and features):** genomes, population diversity data, transcriptomes (incl long-read & single cell RNAseq data), proteomes, metabolomics datasets, interactomes, functional genomics data, phenotypic data, epidemiology
- ❖ **Contribute to the database:** User Comments, Apollo annotation interface
- ❖ **Did you know that you can:**
  - Define your own questions *in silico*, to ID genes of interest based on genome loc'n, taxonomic distrib'n, synteny, orthology, signatures of selection, RNA & protein expression, protein domains / motifs / structures, subcellular location, experimental metadata, etc ... and share with others
  - Visualize & interrogate expression profiles ... based on multiple transcriptomic and proteomic datasets (now including support for long-read & single-cell RNAseq data)
  - Assess gene annotation accuracy, and revise (if warranted) using the Apollo annotation platform, and/or add expert knowledge (references, functions, etc) as User Comments
  - Leverage orthology ... to infer structure & function based on data in related species
  - Determine GO-term & metabolic pathway enrichment ... for your searches or favorite gene lists
  - Analyze your own – or any publicly-available – data (privately) ... map RNAseq reads, or identify SNPs in mutant strains or field isolates, via a user-friendly interface to a cloud-based Galaxy platform



### Databases Are Offline

#### Parasitologists up in arms as NIH ends funding for key database

Trove of data-mining resources on malaria and other killers set to go dark this weekend

11 SEP 2024 • 5:55 PM ET • BY MEREDITH WADMAN



A lab technician works on blood samples at a hospital in Iquitos, Peru. Researchers fear the halt in funding for a database widely used by parasitologists will hinder research on malaria and other diseases with a heavy burden in the Global South. CESAR VON BANCLOS/AFP VIA GETTY IMAGES



**Philippos A Papathanos** • 2nd Senior Lecturer at Hebrew Universit...

Nothing in mosquito biology makes sense except in the light of #VectorBase... and now the lights are going out. #SaveVectorBase

## THE LANCET Microbe

### Critical loss: the effects of VEuPathDB defunding on global health

The scientific community worldwide is shocked by the recent announcement on the EuPathDB website, which states that the National Institute of Allergy and Infectious Diseases (NIH) contract supporting VEuPathDB will expire on Sept 14, 2024. This surprising decision could be catastrophic and seriously hinder scientific advancement, especially in the study of eukaryotic pathogens, as the decision indicates the impending termination of support for one of the most important database resources available to researchers. The gravity of the situation is reflected in the urgent notice imploring researchers to download any essential data, including query strategy results, saved or uploaded DataGalaxy results, and other valuable resources. The ramifications of the funding reduction are extensive and profound, posing a risk to ongoing research projects and delaying the rate of advancement in areas important for global health. The database receives inputs from several basic research initiatives supporting translational science projects that could aid in pioneering discoveries. The scientific community has found EuPathDB, which includes VEuPathDB, to be an invaluable resource (testimonials provided in the panel). This platform incorporates abundant functional proteomic and genomic data from a broad spectrum of eukaryotic pathogens, including those causing some of the most severe diseases in human history.<sup>1,2</sup> Trypanosomiasis, leishmaniasis, malaria, and various fungal infections are caused by eukaryotic pathogens, including fungi, helminths, and protozoa. These pathogens present intrinsic and unique biological features and are often difficult to study because of their complex lifecycles, genetic diversity, and complex host interactions. Research on these pathogens would be severely hampered by the loss of VEuPathDB, slowing the development of novel and critical medications, vaccines, and diagnostic solutions.

Malaria (caused by *Plasmodium* species) is one of the deadliest diseases in the world, especially in sub-Saharan Africa. Despite decades of research and intervention efforts, this disease claims hundreds of thousands of lives annually, with pregnant women and children being particularly susceptible.<sup>3</sup> The understanding of *Plasmodium* biology, drug-resistance mechanisms, and identification of potential vaccine targets have improved owing to the

genomic  
VEuPathDB  
without a  
halt progr...

Another  
affects mil  
is most co  
the Middle  
group of  
These par  
tations, fr  
diseases.<sup>4</sup>  
intricate g  
system. V  
resources  
pinpoint p  
the molec  
this resou  
developm  
severely se

In addi  
leishmania  
strated to  
trypanoso  
fungal inf  
associated  
platform c  
sharing ar  
also holds  
The tool h  
and ther  
important  
in VEuP  
diseases,  
drug-resis  
been usef  
ination of  
design tre  
VEuPathD  
uncharact  
protein st  
In the co  
the search



**Jason Stajich**  
@hyphaltip

Morning fueled by FungiDB mug. Hoping that the efforts for alternative funding avenues will allow these resource to keep providing their unique and key services and grow.



Take the Survey

**FungiDB.org**

## Search for...

expand all | collapse all




## ▼ Genes

- ▶ Annotation, curation and identifiers
- ▶ Epigenomics
- ▶ Function prediction
- ▶ Gene models
- ▶ Genetic variation
- ▶ Genomic Location
- ▶ Immunology
- ▶ Orthology and synteny
- ▶ Pathways and interactions
- ▶ Phenotype
- ▶ Protein features and properties
- ▶ Protein targeting and localization
- ▶ Proteomics
- ▶ Sequence analysis
- ▶ Structure analysis
- ▶ Taxonomy
- ▶ Text
- ▶ Transcriptomics

## ▶ Organisms

- ▶ Popset Isolate Sequences

- ▶ Genomic Sequences

- ▶ Genomic Segments

- ▶ SNPs

- ▶ ESTs

## What will become of VEuPathDB?

*... including FungiDB, VectorBase, TriTrypDB, PlasmoDB, ToxoDB, CryptoDB, etc*

- ❖ **NIH support for VEuPathDB genomics resources (including FungiDB) was terminated on 15 Sept 2024** ... contrary to study section recommendations, despite certification as a Global Core Biodata Resource, without community consultation, and notwithstanding adverse press, social media, and many *thousands* of outraged emails from affected users worldwide.
- ❖ Concerned about the impact of this decision on global public health, and the tens of thousands of mycology / parasitology / vector biology researchers who depend on VEuPathDB for their work, **several groups have provided funds to bring VEuPathDB back on line, at least temporarily:**



- ❖ **FungiDB is once again available, as a static resource**, i.e. without new data, datatypes, or analytical tools, and with minimal outreach activity. Our supporters have also provided partial funding to implement infrastructure that we expect will allow the **resumption of regular database releases (including new genomes, functional datasets & analytical tools) in 2025**.
- ❖ Long-term financial sustainability will depend upon additional support, likely involving a combination of grants, philanthropic contributions and community buy-in. **This will require careful assessment of costs, value, and what the community needs and is willing to support.**



# VEuPathDB Subscription Pricing

## Why are we implementing a subscription service?

makes it difficult to sustain essential infrastructure through consultation with our user communities, and are based

## for Academic Research Groups:

To determine your fair share of VEuPathDB annual budget

Use the table at right to define appropriate annual budget details associated with:

- individual grant(s) ... *OK to separate invoices*
- your entire lab's research program, or
- a larger group (e.g. your department) ... *sum organization, less a 5% group discount*

Calculate applicable subscription rates based on (1) *Base usage and (2) the scale of your operation, and the production of data to be shared via VEuPathDB*. **include PI salaries and all participating personnel budget(s), including indirect costs (overhead)**

**OR**

Use [fee calculator](#) (also downloadable as an [Excel](#) file) to calculate fees based on lab budget & resource usage. Usage by various individuals or projects.

**Note:** Financial and personnel details are not included in the fee calculator.

**Additional Questions?** Please see [FAQ](#).

### Use this form to generate an invoice for "VEuPathDB open access and data dissemination"

We accept payment via credit card, PayPal, check, or wire transfer; see below for payment details (also provided on the invoice). Your order will not be processed until payment is received.

1. Fill out required fields below, then click the button to generate an invoice

Your name full name:

Email address (please double check your entry):

Street Address:

Department:

Institution:

City:

State/Province:

Postal Code:

Country:

Annual payment amount in USD (from table or fee calculator):

Please explain rationale if this amount differs from table/fee calculator determination of your fair share  
(Note: this information is required if you are proposing to pay \$0 for access to VEuPathDB resources)

NA

Number of subscription years:

Check if paying by credit card:

*Note: 3% surcharge to cover credit card fees*

2. Click here to Generate Invoice **Required, even if you are proposing to pay \$0 for access to VEuPathDB**

3. Payment details (unless proposing that your fair share of VEuPathDB resources = \$0):

• by electronic transfer wire to: Wells Fargo Bank NA ABA number: 121-000-248  
(preferred) 420 Montgomery Street acct number: 2000018692644  
San Francisco CA 94104 USA SWIFT code: WFBIUS6S  
credit to: Trustees of the University of Pennsylvania CHIPS: 0407  
attn: VEuPathDB subscriptions

• by check, through any bank with branches in the US  
make checks payable to: Trustees of the University of Pennsylvania  
mail checks to: VEuPathDB Subscriptions – attn: Mary Ellen Sarno  
3600 Market Street / Suite 570  
University of Pennsylvania, Philadelphia PA 19104 USA

• by credit card (American Express, Discover, Visa or MasterCard)  
use [online payment](#) system to provide credit card ... *expected to be available in April 2025*

## SAMPLE INVOICE

### The University of Pennsylvania

VEuPathDB.org

c/o Mary Ellen Sarno  
3600 Market Street, Suite 570  
Philadelphia PA 19104  
USA  
[subscriptions@veupathdb.org](mailto:subscriptions@veupathdb.org)

INVOICE #: VEuPathDB-1194-2747  
DATE: 13 Feb 2025



### TO

Aspergillus gambiae  
123 Laboratory Ave  
Department of Microbial Genomics  
University of Cool Science  
Anywhere OK 12345-6789  
United States

Description (services provided)	Number of years	Price per year	Amount (in USD)
Data dissemination and open access to VEuPathDB websites			
Annual subscription (calendar year)	3	\$ 5,000	\$ 15,000
Credit card processing fee (3%)			
<b>Total amount due (in US dollars): \$ 15,450</b>			

### Additional notes (optional)

Covers research by the Aspergillus gambiae lab funded by grant #12345

### Payment methods and details

Electronic wire transfer (preferred)	Wells Fargo Bank NA 420 Montgomery St San Francisco CA 94104 USA	ABA number: 121-000-248 Acct number: 2000018692644 SWIFT code: WFBIUS6S CHIPS: 0407 Credit to: Trustees of the University of Pennsylvania attn: VEuPathDB subscriptions
--------------------------------------	--	--

Check (bank must have branches in the US)	Make payable to: Trustees of the University of Pennsylvania	Mail to: VEuPathDB subscriptions attn: Mary Ellen Sarno 3600 Market St / Suite 570 Univ Pennsylvania Philadelphia PA 19104 USA
---	--	---

Credit cards (Visa, Discover, MasterCard, American Express)	Online payment system ... available in April 2025
---	--

Thank you for supporting VEuPathDB!



***Sent on behalf of the Fungal Policy Committee***

Dear colleague,

As you know, [FungiDB \(part of VEuPathDB\)](#), which supports many of our research programs, is in the process of establishing a new funding model in order to continue providing open data access to all mycology researchers worldwide. This effort has been guided by input from [community meetings](#), [user surveys](#), and individual communications. The VEuPathDB subscription service is now available [here](#).

**All members of the Fungal Genetics Policy Committee have subscribed to this service. We are unanimous in viewing this as an important community effort, which we encourage each of you to join.** Generating invoices is easy and straightforward using the above website, and the folks at FungiDB are available to help if you need any additional assistance. Purchasing departments may require PIs to provide some additional information to establish a new vendor, which can take 1-2 weeks, depending on your institution. If you prefer to pay by credit card, that option will be available soon.

**We hope that all PIs will use the above link to determine their appropriate subscription level, based on the size of their operations & frequency of database usage, and complete the associated invoice form.** We realize that subscription fees for access to VEuPathDB have not been part of any of our budgets to date, making it difficult to mobilize funds, but please

consider the importance of these resources for your productivity, in comparison to annual expenditures for data analysis software, cell culture, microscopy, enzymes, animals, conference travel, publication charges, etc. Most funding agencies are willing to accept a line item covering subscription services for “website access, data-mining, and open access data sharing.”

Various philanthropic groups have helped keep VEuPathDB online for the past six months, and may be willing to provide limited additional support, but these database resources will not continue to exist without substantial contributions from all of us as end-users. **Without financial commitments from users before June 2025, VEuPathDB will lose the staff necessary to keep this resource alive, with little possibility to resurrect it in future. Your subscription to FungiDB is needed now.**

Please feel free to forward this email to other PIs. Further information is available on the VEuPathDB [FAQ page](#), and you are also welcome to contact [subscriptions@veupathdb.org](mailto:subscriptions@veupathdb.org) with any additional questions.

Sincerely,

[Xiaorong Lin](#), University of Georgia, US (2019–2025; current chair)

[Luis Larrondo](#), Catholic University, Chile (2019–2025)

[Vera Meyer](#), Technical University, Berlin, Germany (2019–2025)

[Reinhard Fischer](#), Karlsruhe Institute of Technology, Germany (2022–2028)

[Li-Jun Ma](#), University of Massachusetts Amherst, US (2022–2028)

[Oded Yarden](#), Hebrew University, Jerusalem, Israel (2022–2028)

[Luis Corrochano](#), University of Seville, Spain (2025–2031)

[Alex Idnurm](#), University of Melbourne, Australia (2025–2031)

[Rebecca Shapiro](#), University of Guelph, Canada (2025–2031)



**Fungal  
Genetics Conference**

MARCH 17-22, 2026  
PACIFIC GROVE, CA  
#Fungal26



# Browser Views: Assessing gene models

My Organism Preferences (317 of 317)  enabled

Genome Track View Help

0 500,000 1,000,000 1,500,000 2,000,000 2,500,000 3,000,000 3,500,000 4,000,000 4,500,000 5,000,000 5,500,000 6,000,000 6,500,000 7,000,000 7,500,000 8,000,000 8,500,000 9,000,000 9,500,000

00 2,190,000 2,195,000 2,200,000 2,205,000 2,210,000 2,215,000

Select tracks

CM002236 CM002236:2184997..2218330 (33.33) Go X

00 2,190,000 2,195,000 2,200,000 2,205,000 2,210,000 2,215,000

Scaffolds and Gaps

GC Content

Tandem Repeats

Low Complexity Regions

Annotated Transcripts (UTRs in White when available)

NCU01875 ein hypothetical protein

NCU01874 hypothetical protein

NCU01873 hypothetical protein

NCU01872 hypothetical protein

NCU01870 enolase-1

NCU01868 sugar transporter-25

NCU01869 hypothetical protein

NCU01867 endoglucanase II

NCU01866 pentose phosphate metabolism-9

NCU01865 hypothetical protein

NCU01864 unspecified product

NCU15018 tRNA-Gly

SNP by coding potential

Population genomics transcriptomes Density - Unique Only

Syntenic Sequences and Genes (Shaded by Orthology)

Neurospora crassa OR74A gene

Neurospora crassa OR74A span

Arabidopsis thaliana FGSC8579 gene

Arabidopsis thaliana FGSC8579 span

Arabidopsis thaliana FGSC2508 gene

Arabidopsis thaliana FGSC2508 span

Arabidopsis thaliana Smat gene

# Browser Views: Assessing gene models

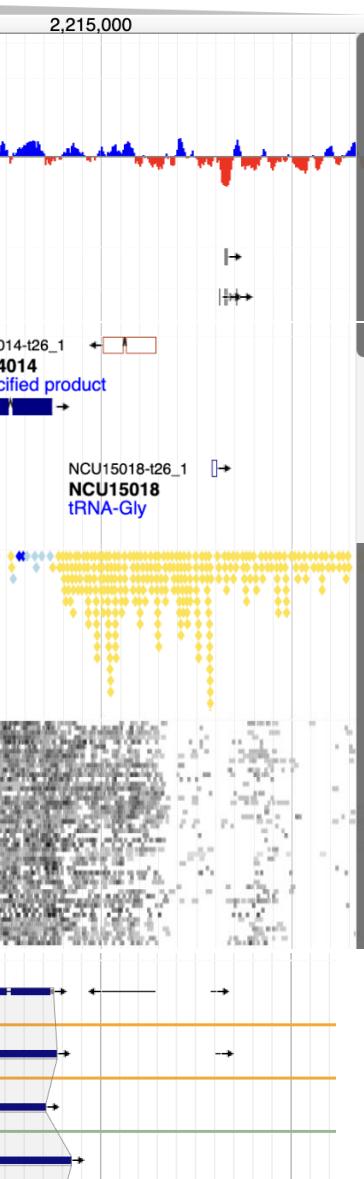
My Organism Preferences (317 of 317)  enabled

Genome	Track	View	Help	Neurospora crassa OR74A	Share
0	500,000	1,000,000	1,500,000	2,000,000	2,500,000 3,000,000 3,500,000 4,000,000 4,500,000 5,000,000 5,500,000 6,000,000 6,500,000 7,000,000 7,500,000 8,000,000 8,500,000 9,000,000 9,500,000

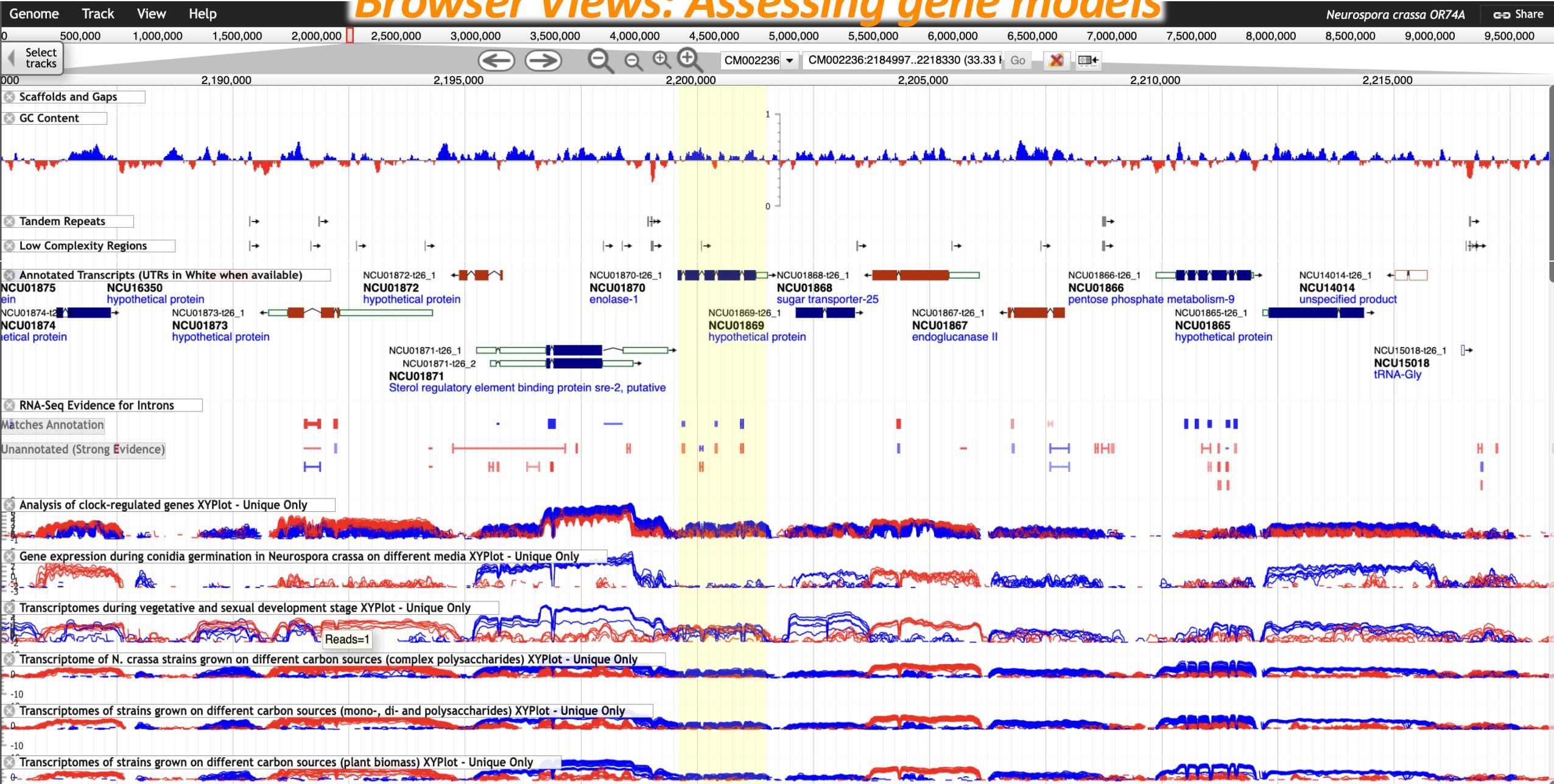
## Select Tracks

My Tracks	
Currently Active	
Recently Used	
Category	
1 Comparative Genomics	
36 Epigenomics	
3 Gene Models	
186 Genetic Variation	
8 Sequence Analysis	
1,218 Transcriptomics	
Subcategory	
2 (no data)	
3 BLAT and Blast Alignments	
36 ChIP-Seq	
1 DNA polymorphism	
185 DNA-Seq	
1 Orthology and Synteny	
1,214 RNA-Seq	
1 Sequence assembly	
3 Sequence composition, complexity and repeats	
1 Sequence sites, features and motifs	
2 Transcripts	
3 small non-coding RNA	
Dataset	
15 (no data)	
9 ADV-1 gene targets ChIP-seq	
272 Analysis of clock-regulated genes	
14 Characterization of light regulated genes	
1 Combined all RNA-Seq data for NcrassaOR74A	
12 Comparative transcriptome of Neurospora crassa WT and delta mms52 strains	
4 Dap-Seq of Neurospora crassa	
16 Gene expression analysis of ADV-1 and PP-1 mutants	
36 Gene expression during conidia germination in Neurospora crassa on different media	
36 Genome-wide transcriptome during sexual development	
23 Identification of direct target genes of the Neurospora crassa essential plant biomass deconstruction transcription factors CLR-1, CLR-2 and XLR-1 (ChIP-Seq)	
100 Population genomics transcriptomes	
36 Profiling of transcriptional response of clr-1 and clr-2 mutants to cellulose	
100 RNA-sequencing of WT and delta mms52 strains	

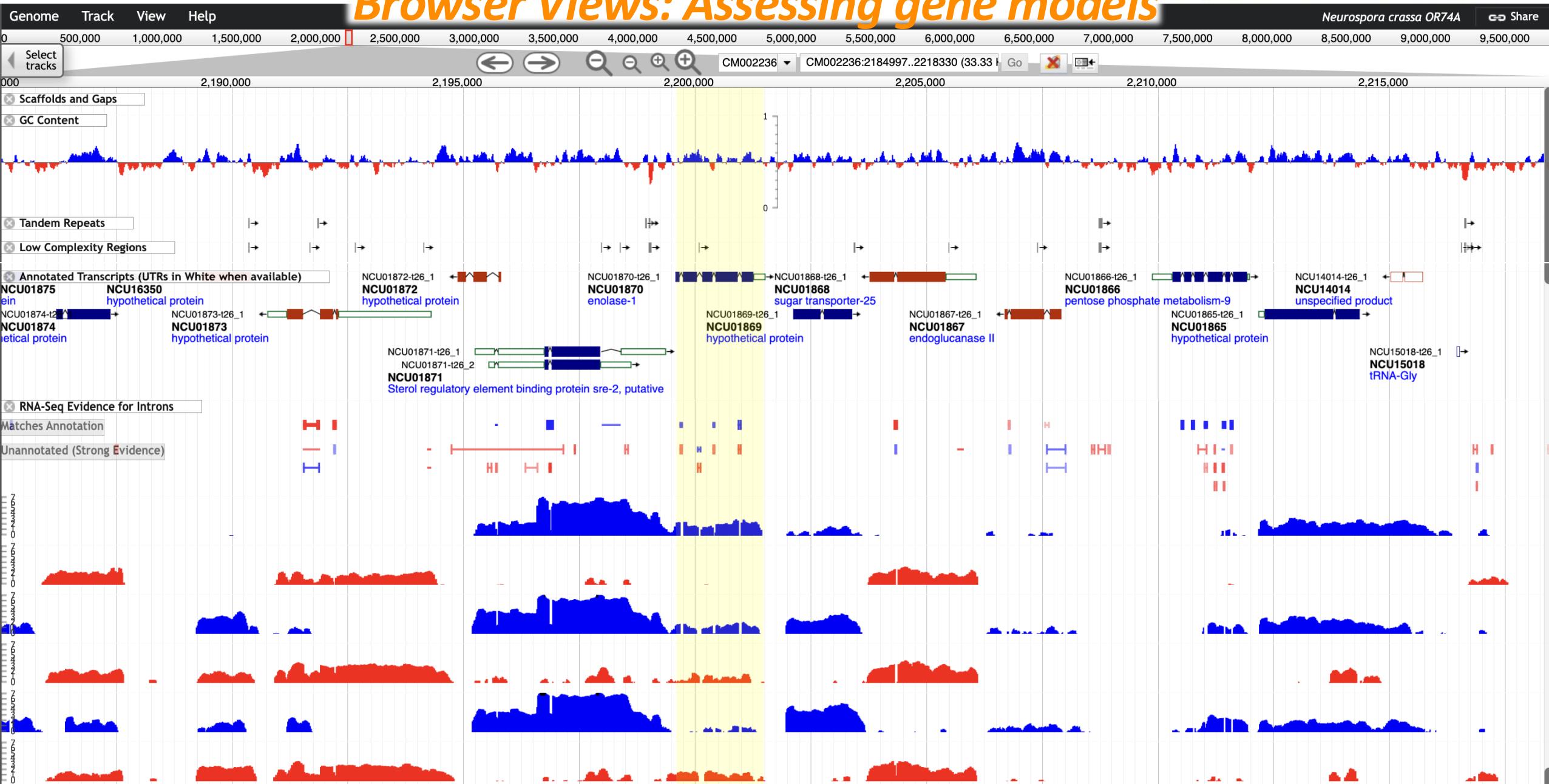
Back to browser	Clear All Filters	Contains text <input type="text"/>	1,452 tracks
<input type="checkbox"/>	Name	Category	Subcategory
<input type="checkbox"/>	2PF Small Non-coding RNAs	Transcriptomics	small non-coding RNA
<input type="checkbox"/>	4PF Small Non-coding RNAs	Transcriptomics	small non-coding RNA
<input type="checkbox"/>	ADV-1 gene targets ChIP-seq - ADV-1_dark_1 Coverage	Epigenomics	ChIP-Seq
<input type="checkbox"/>	ADV-1 gene targets ChIP-seq - ADV-1_dark_2 Coverage	Epigenomics	ChIP-Seq
<input type="checkbox"/>	ADV-1 gene targets ChIP-seq - ADV-1_light_15min_1 Coverage	Epigenomics	ChIP-Seq
<input type="checkbox"/>	ADV-1 gene targets ChIP-seq - ADV-1_light_15min_2 Coverage	Epigenomics	ChIP-Seq
<input type="checkbox"/>	ADV-1 gene targets ChIP-seq - ADV-1_light_30min_1 Coverage	Epigenomics	ChIP-Seq
<input type="checkbox"/>	ADV-1 gene targets ChIP-seq - ADV-1_light_30min_2 Coverage	Epigenomics	ChIP-Seq
<input type="checkbox"/>	ADV-1 gene targets ChIP-seq - ADV-1_light_60min_1 Coverage	Epigenomics	ChIP-Seq
<input type="checkbox"/>	ADV-1 gene targets ChIP-seq - ADV-1_light_60min_2 Coverage	Epigenomics	ChIP-Seq
<input type="checkbox"/>	ADV-1 gene targets ChIP-seq Density - Unique Only	Epigenomics	ChIP-Seq
<input type="checkbox"/>	Analysis of clock-regulated genes - TM1_DD10_CT23 - time course 1 (non-unique forward) Coverage	Transcriptomics	RNA-Seq
<input type="checkbox"/>	Analysis of clock-regulated genes - TM1_DD10_CT23 - time course 1 (non-unique reverse) Coverage	Transcriptomics	RNA-Seq
<input type="checkbox"/>	Analysis of clock-regulated genes - TM1_DD10_CT23 - time course 1 (unique forward) Coverage	Transcriptomics	RNA-Seq
<input type="checkbox"/>	Analysis of clock-regulated genes - TM1_DD10_CT23 - time course 1 (unique reverse) Coverage	Transcriptomics	RNA-Seq
<input type="checkbox"/>	Analysis of clock-regulated genes - TM1_DD12_CT1.2 - time course 1 (non-unique forward) Coverage	Transcriptomics	RNA-Seq
<input type="checkbox"/>	Analysis of clock-regulated genes - TM1_DD12_CT1.2 - time course 1 (non-unique reverse) Coverage	Transcriptomics	RNA-Seq
<input type="checkbox"/>	Analysis of clock-regulated genes - TM1_DD12_CT1.2 - time course 1 (unique forward) Coverage	Transcriptomics	RNA-Seq
<input type="checkbox"/>	Analysis of clock-regulated genes - TM1_DD12_CT1.2 - time course 1 (unique reverse) Coverage	Transcriptomics	RNA-Seq



# Browser Views: Assessing gene models

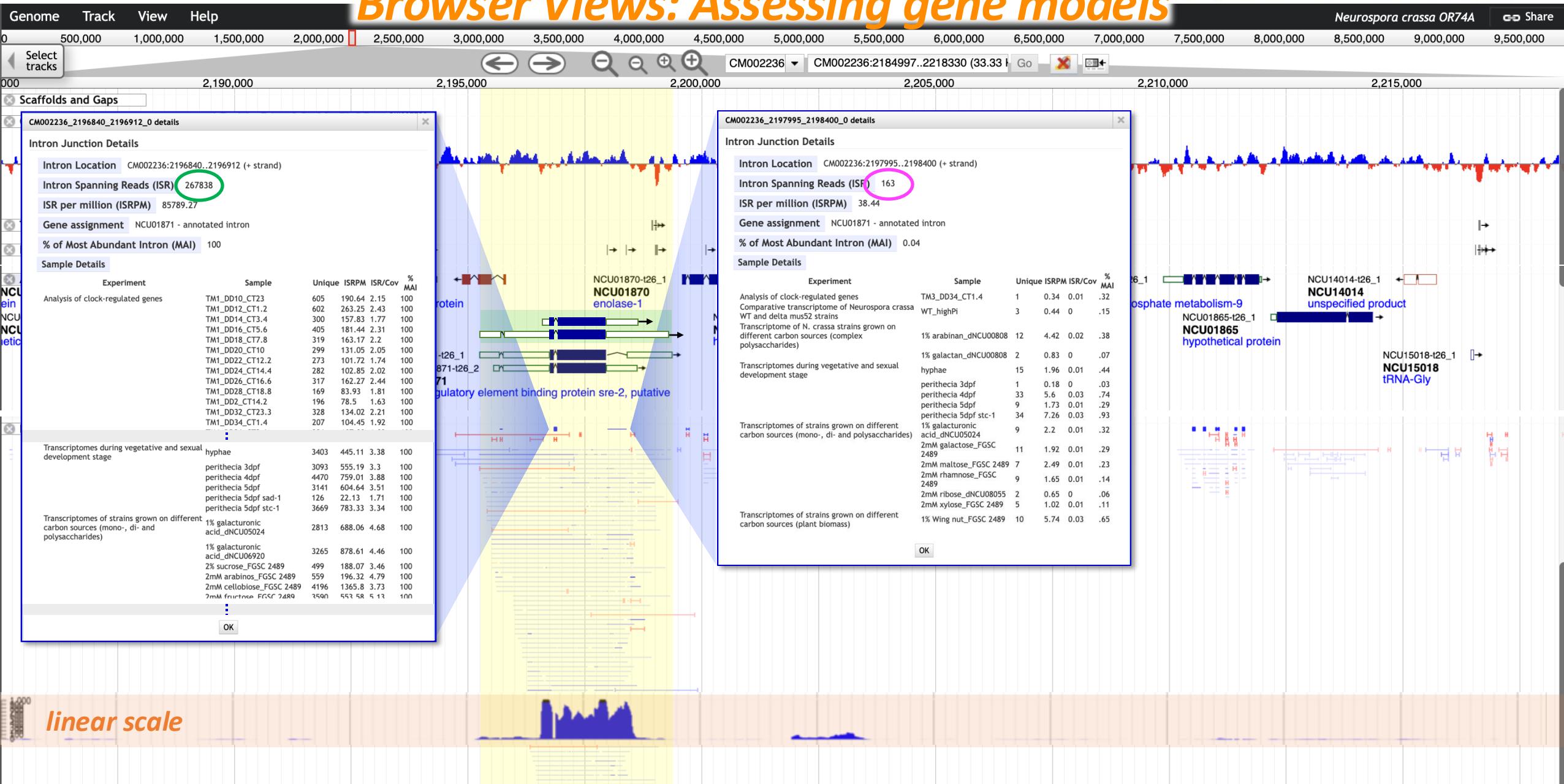


# Browser Views: Assessing gene models





# Browser Views: Assessing gene models



Add to basket Add to favorites Download Gene

## FGRAMPH1\_01G13253 Small secreted cysteine-rich protein, putative

Gene Type protein coding gene

Biotype Classification protein\_coding

Chromosome 2

Location HG970333:5,564,661..5,565,001(-)

Species *Fusarium graminearum*

Strain PH-1

Status Reference Strain

User Comments [Add the first](#)

Community Annotations [View / Update in Apollo editor](#)

### Shortcuts



Also see FGRAMPH1\_01G13253 in the Genome Browser or Protein Browser

Collapse all sections for better performance

FGRAMPH1\_01G13253

«

### Gene models

[expand all](#) | [collapse all](#)

▶ Gene models

▶ [Annotation, curation and identifiers](#)

▶ [Link outs](#)

# Exons in Gene 2

# Transcripts 1

GFF format of gene and transcript features [Click to open GFF in a new tab](#)

▼ Gene Models

This gene is available in **Apollo** for community annotation. To find out more about Apollo, please visit [this help page](#).

FGRAMPH1\_01G13253

**Literature**
[expand all](#) | [collapse all](#)


- ▶ Gene models
- ▶ Annotation, curation and identifiers
- ▶ Link outs
- ▶ Genomic Location
- ▶ Literature**
- ▶ Taxonomy
- ▶ Orthology and synteny
- ▶ Phenotype
- ▶ Genetic variation
- ▶ Transcriptomics
- ▶ Sequences
- ▶ Sequence analysis
- ▶ Structure analysis
- ▶ Protein features and properties
- ▶ Function prediction
- ▶ Pathways and interactions
- ▶ Proteomics

**PubMed**
 [Download](#)
 [Data sets](#)

Shown are references from the NCBI gene2pubmed database and from manual assignment by curators.

? 3 rows

PubMed ID	DOI	Title	Authors
27507961	<a href="#">10.3389/fmicb.2016.01113</a>	Comparison of Fusarium graminearum Transcriptomes on Living or Dead Wheat Differentiates Substrate-Responsive and Defense-Responsive Genes.	Boedi S, Berger H, Sieber C, Münsterkötter M, Maloku...
26524547	<a href="#">10.1094/PHYTO-09-15-0215-R</a>	Genome-Wide Analysis of Small Secreted Cysteine-Rich Proteins Identifies Candidate Effector Proteins Potentially Involved in Fusarium graminearum-Wheat Interactions.	Lu S, Edwards MC
17823352	<a href="#">10.1126/science.1143708</a>	The Fusarium graminearum genome reveals a link between localized polymorphism and pathogen specialization.	Cuomo CA, Güldener U, Xu JR, Trail F, Turgeon BG, Di...

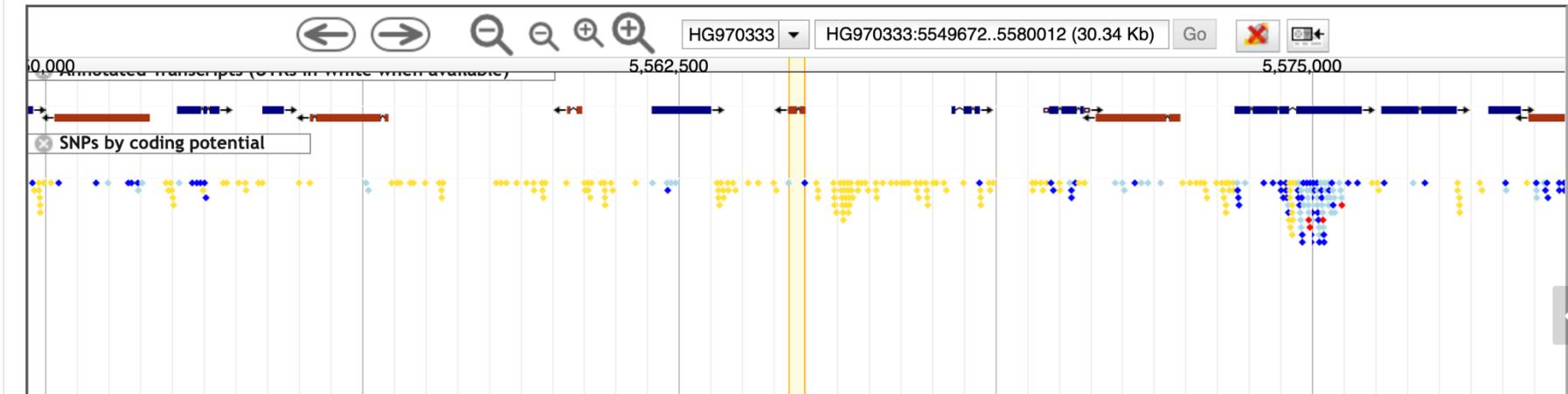
**Taxonomy**
> [Taxonomy](#)
**Orthology and synteny**
 Ortholog Group ? OG6\_135803

> [Orthologs and Paralogs within FungiDB](#)

FGRAMPH1\_01G13253

**Genetic variation**[expand all](#) | [collapse all](#)
 Search section names...

- ▶ Gene models
- ▶ Annotation, curation and identifiers
- ▶ Link outs
- ▶ Genomic Location
- ▶ Literature
- ▶ Taxonomy
- ▶ Orthology and synteny
- ▶ Phenotype
- ▶ **Genetic variation**
- ▶ Transcriptomics
- ▶ Sequences
- ▶ Sequence analysis
- ▶ Structure analysis
- ▶ Protein features and properties
- ▶ Function prediction
- ▶ Pathways and interactions
- ▶ Proteomics

**▼ DNA polymorphism**
Non-Coding SNPs All Strains [?](#) 0NonSyn/Syn SNP Ratio All Strains [?](#) 1NonSynonymous SNPs All Strains [?](#) 1SNPs with Stop Codons All Strains [?](#) 0Synonymous SNPs All Strains [?](#) 1Total SNPs All Strains [?](#) 2
**▼ SNPs**
[View in JBrowse genome browser](#)[Reset view](#)

## FGRAMPH1\_01G13253

[expand all](#) | [collapse all](#)
▼ [Search section names...](#)

- ▶ Gene models
- ▶ Annotation, curation and identifiers
- ▶ Link outs
- ▶ Genomic Location
- ▶ Literature
- ▶ Taxonomy
- ▶ Orthology and synteny
- ▶ Phenotype
- ▶ Genetic variation
- ▶ Transcriptomics
- ▶ Sequences
- ▶ Sequence analysis
- ▶ **Structure analysis**

[Similarities to Protein Data Bank \(PDB\)](#)  
[Chains](#)
[SSGCID pipeline](#)
[AlphaFold](#)
[AlphaFold Structure Prediction Visualization](#)
▼ AlphaFold Structure Prediction Visualization ?

VEuPathDB uses UniProt IDs to associate AlphaFold predictions with genes.

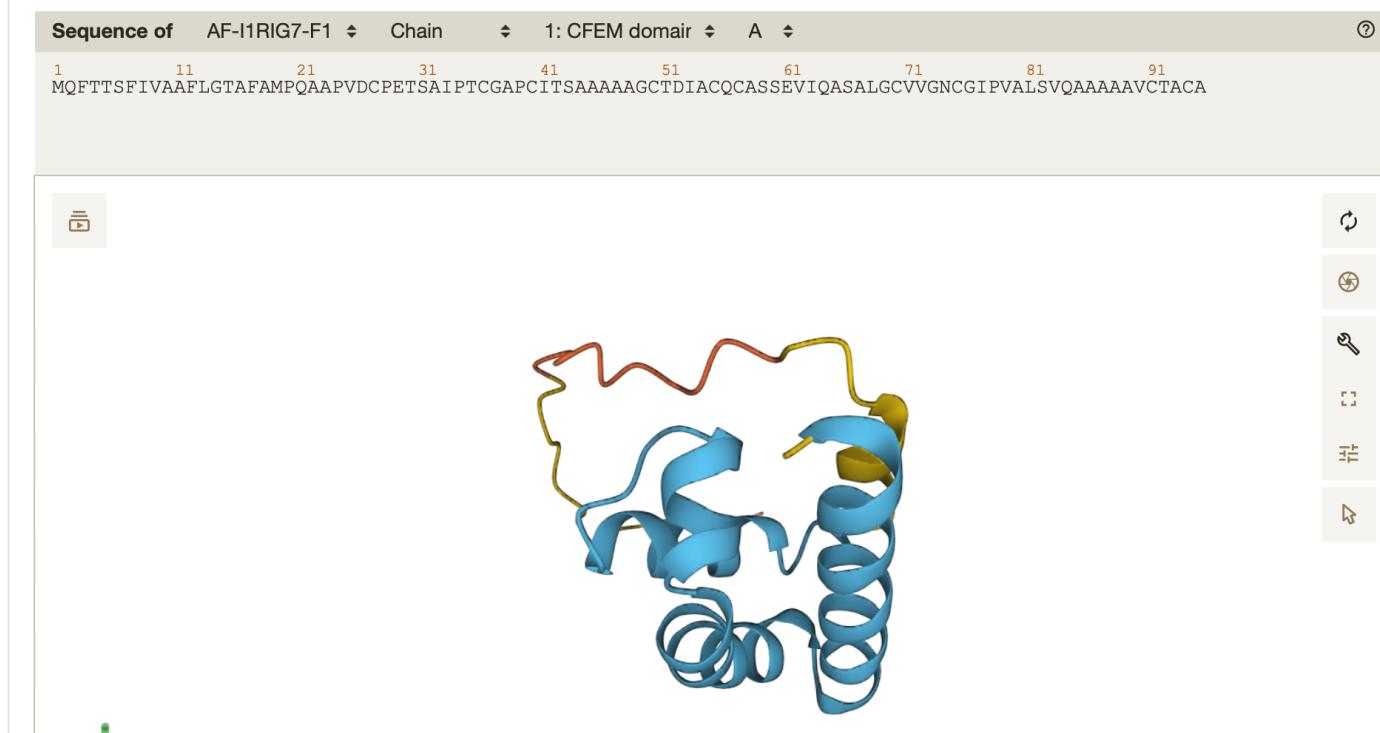
In some cases where specific organisms or genes are not in UniProt, associations between genes and UniProt entries have been made by sequence similarity. This may result in multiple UniProt IDs mapping to a gene.

In these cases, we have selected a representative structure to show on this page. Where possible, this is a direct UniProt entry. Where only associations using sequence similarity are available, we have shown the longest prediction. A full list of AlphaFold predictions putatively associated with this gene is available in the AlphaFold table above.

**Key:**

AlphaFold structures are colored using a per-residue confidence metric called pLDDT, which is scaled from 1 - 100.

Very high (pLDDT > 90) | Confident (90 > pLDDT > 70) | Low (70 > pLDDT > 50) | Very low (pLDDT < 50)



## FGRAMPH1\_01G13253

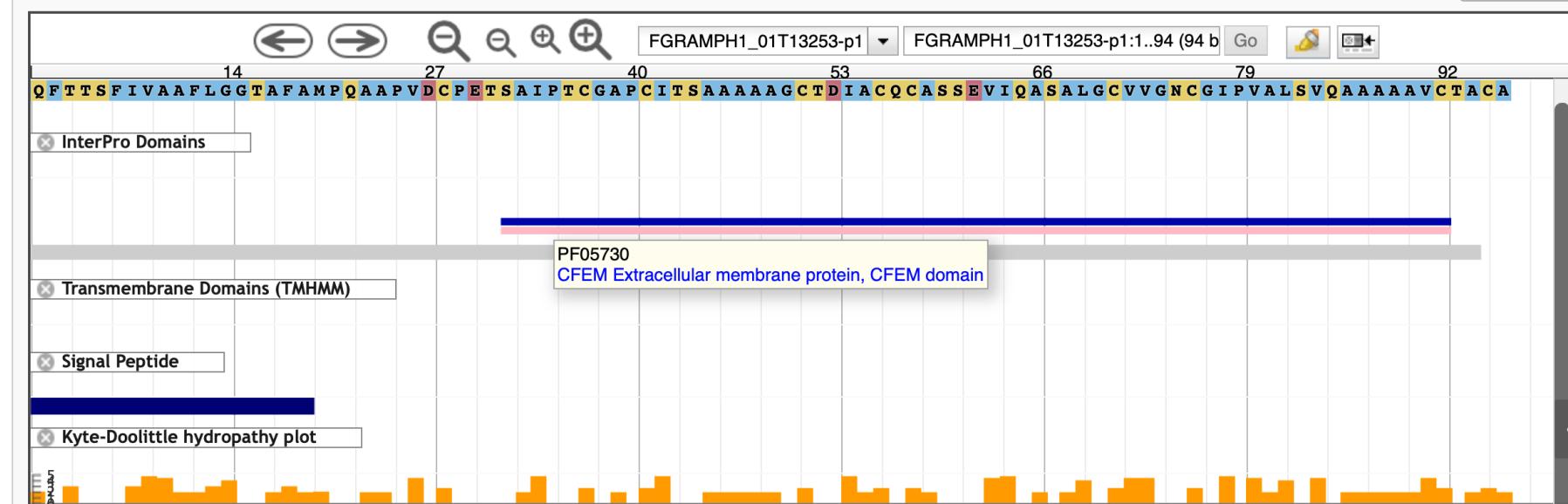

[expand all](#) | [collapse all](#)
▼ Search section names...

- ▶ Gene models
- ▶ Annotation, curation and identifiers
- ▶ Link outs
- ▶ Genomic Location
- ▶ Literature
- ▶ Taxonomy
- ▶ Orthology and synteny
- ▶ Phenotype
- ▶ Genetic variation
- ▶ Transcriptomics
- ▶ Sequences
- ▶ Sequence analysis
- ▶ Structure analysis
- ▶ Protein features and properties

[Attributes and Protein Browser](#)
[BLASTP \(protein-protein BLAST\)](#)
[GPI anchor prediction: big-PI Predictor](#)
[GPI anchor prediction: big-PI Predictor in](#)
▼ Attributes and Protein Browser ?
[Download](#) [Data sets](#)

Protein Browser (JBrowse) view of protein structure prediction, domain prediction and protein features including isoelectric point, molecular weight, Signal peptide (SignalP) domain, length, sequence, transmembrane domains (TMHMM), InterPro domains, predicted export domains, hydropathy plot, low complexity regions, PSIPRED Helix, Strand and Coil predictions.

	Transcript ID	Isoelectric Point	Molecular Weight	Has SignalP	Has TMHMM	Protein Length	Protein Browser
▼	FGRAMPH1_01T13253	3.29	9076	yes	no	95	Interactive

[View in protein browser](#)
[Reset view](#)


FGRAMPH1\_01G13253



## Transcriptomics

[expand all](#) | [collapse all](#)
 Search section names...

- ▶ Gene models
- ▶ Annotation, curation and identifiers

- ▶ Link outs

- ▶ Genomic Location

- ▶ Literature

- ▶ Taxonomy

- ▶ Orthology and synteny

- ▶ Phenotype

- ▶ Genetic variation

- ▼ Transcriptomics

- [RNA-Seq Transcription Summary](#)

- [Transcript Expression](#)

- [User Dataset Transcriptomics Graphs](#)

- ▶ Sequences

- ▶ Sequence analysis

- ▶ Structure analysis

- ▶ Protein features and properties

- ▶ Function prediction

### ▼ RNA-Seq Transcription Summary

#### Data sets

► Summary of expression values. Each row represents a distinct RNA-Seq experiment. [Click to read more...](#)

#### Gene: FGRAMPH1\_01G13253

0 2 4 6 8 10 12  
Fusarium transcriptome in low and high nitrogen conditions, Connolly et al. 2013

Genome-wide gene expression analysis of spores and mycelium, Zhao et al. 2014

RNA-seq Analysis of FlbD-, AbaA-, and WetA-dependent Genes in Fusarium graminearum Z-3639, Son et al. 2013

RNA-seq analysis of FSS1-dependent genes in Fusarium graminearum Z-3639, Son et al. 2015

RNA-Seq analysis of transcriptomes at 1 day after sexual induction, Son et al. 2016

Stage-specific and genome-wide gene expression during the different developmental stages of Fusarium graminearum, Liu et al. 2016

Transcriptome in response to an active or passive host, Boedi et al. 2016

Transcriptome in heat-stressed *Fusarium graminearum* Z-3639, Bui et al. 2016

Transcriptomic analysis during vegetative and infectious growth of *Fusarium graminearum* PH-1, Wang et al. 2017

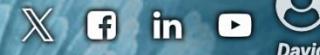
Control 0.31  
mycelia

host\_alive 2.01

host\_dead 313.12

infecte

# Gene pages: AI summarization

[My Strategies](#) [Searches](#) [Tools](#) [My Workspace](#) [Data](#) [About](#) [Help](#) [Contact Us](#)


FGRAMPH1\_01G13253


[expand all](#) | [collapse all](#)
 Search section names...

- ▶ Gene models
- ▶ Annotation, curation and identifiers
- ▶ Link outs
- ▶ Genomic Location
- ▶ Literature
- ▶ Taxonomy
- ▶ Orthology and synteny
- ▶ Phenotype
- ▶ Genetic variation
- ▼ Transcriptomics
  - # Microarray Datasets
  - # RNASeq Datasets
  - ▶ AI Expression Summary 
  - ▶ RNA-Seq Transcription Summary
  - Transcript Expression
  - User Dataset Transcriptomics Graphs
- ▶ Sequences
- ▶ Sequence analysis

## Transcriptomics

# Microarray Datasets 3

# RNASeq Datasets 9

▼ **AI Expression Summary** 

### Differential Expression of a Gene in *Fusarium graminearum* Under Pathogenic and Developmental Contexts

The analyzed gene demonstrates noteworthy patterns of expression across diverse experimental conditions involving *Fusarium graminearum*.

- During pathogenesis, the gene exhibits upregulation during wheat and maize infection, with peak expression observed at specific time points in both pathologies.
- Higher expression is also noted in developmental stages like conidia and perithecia, suggesting a potential role in developmental transitions.
- Conversely, the gene shows differential expression under specific mutant conditions and no significant activity in nitrogen limitation and heat stress experiments.

These observations imply potential roles in pathogenesis, lifecycle transitions, or response to environmental cues.

*The results from 12 experiments have been organized into the 3 topics below. The AI was instructed to present the most biologically relevant information first. As this method is still evolving, results may vary.*

▼	Topic	Summary	# Datasets
▼	Elevated Expression During Pathogen-Host Interactions in <i>Fusarium graminearum</i>	The gene is expressed significantly during wheat and maize infections, implicating its potential role in plant-pathogen interactions.	5

- Transcriptome of *Fusarium graminearum* during the infection of wheat (array)  
The gene exhibits a substantial increase in expression reaching peak levels about 48-96 hours post-inoculation during the wheat infection process.
- Transcriptomic analysis during vegetative and infectious growth of *Fusarium graminearum* PH-1 (RNA-Seq)  
This gene exhibits significantly higher expression in infected conditions compared to vegetative conditions.
- Gene expression of *F. graminearum* during maize stalk infection (array)  
The gene exhibits an increase in expression over time, peaking significantly at 144 hours after inoculation in the dataset.
- *Fusarium graminearum* transcriptome during symptomless and symptomatic wheat infection (array)  
The gene is highly expressed in symptomatic samples and progressively less expressed in intermediate and symptomless conditions, indicating

# Gene pages: AI summarization

[My Strategies](#) [Searches](#) [Tools](#) [My Workspace](#) [Data](#) [About](#) [Help](#) [Contact Us](#)
    David

FGRAMPH1\_01G13253

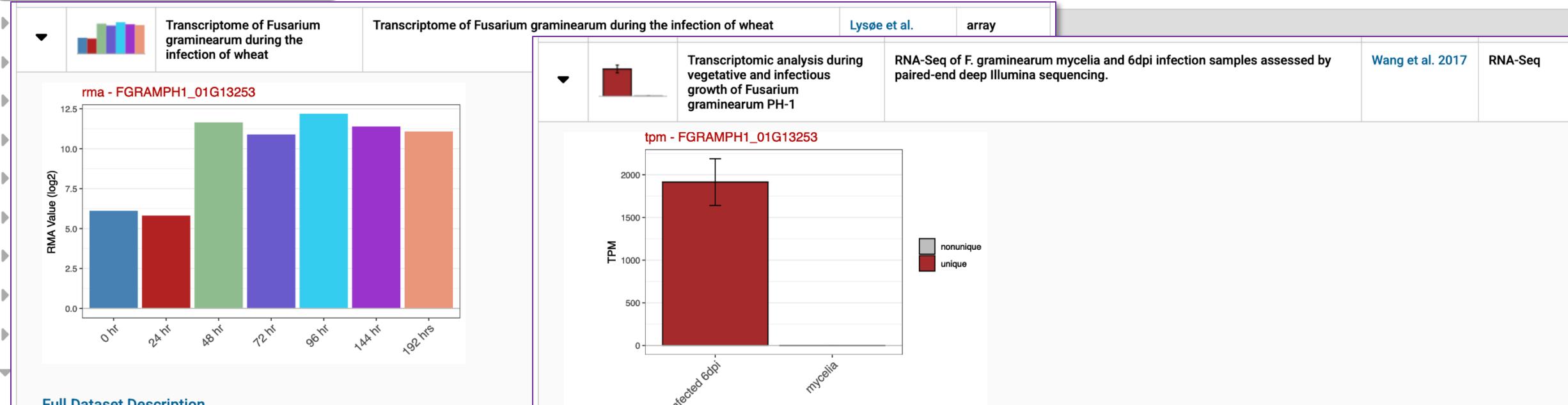


## Transcriptomics

[expand all](#) | [collapse all](#)

# Microarray Datasets 3

# RNASeq Datasets 9

 Search section names...

[Full Dataset Description](#)

# RNASeq Datasets

AI Expression Summary

RNA-Seq Transcription Summary

Transcript Expression

User Dataset Transcriptomics Graphs

Sequences

Sequence analysis

Topic	Full Dataset Description	
Elevated Expression During Pathogen-Host Interactions in Fusarium graminearum	The gene is expressed significantly during wheat and maize infections, implicating its potential role in plant-pathogen interactions.	5

- Transcriptome of Fusarium graminearum during the infection of wheat (array)  
The gene exhibits a substantial increase in expression reaching peak levels about 48-96 hours post-inoculation during the wheat infection process.
- Transcriptomic analysis during vegetative and infectious growth of Fusarium graminearum PH-1 (RNA-Seq)  
This gene exhibits significantly higher expression in infected conditions compared to vegetative conditions.
- Gene expression of F. graminearum during maize stalk infection (array)  
The gene exhibits an increase in expression over time, peaking significantly at 144 hours after inoculation in the dataset.
- Fusarium graminearum transcriptome during symptomless and symptomatic wheat infection (array)  
The gene is highly expressed in symptomatic samples and progressively less expressed in intermediate and symptomless conditions, indicating

# Gene pages: AI summarization

[My Strategies](#) [Searches](#) [Tools](#) [My Workspace](#) [Data](#) [About](#) [Help](#) [Contact Us](#)
X f in David

FGRAMPH1\_01G13253


[expand all](#) | [collapse all](#)
Search section names...

- ▶ Gene models
- ▶ Annotation, curation and identifiers
- ▶ Link outs
- ▶ Genomic Location
- ▶ Literature
- ▶ Taxonomy
- ▶ Orthology and synteny
- ▶ Phenotype
- ▶ Genetic variation
- ▶ **Transcriptomics**
  - # Microarray Datasets
  - # RNASeq Datasets
  - ▶ AI Expression Summary 
  - RNA-Seq Transcription Summary
  - Transcript Expression
  - User Dataset Transcriptomics Graphs
- ▶ Sequences
- ▶ Sequence analysis

<p>▼ Elevated Expression During Pathogen-Host Interactions in <i>Fusarium graminearum</i></p> <ul style="list-style-type: none"> <li>• <a href="#">Transcriptome of <i>Fusarium graminearum</i> during the infection of wheat (array)</a> The gene exhibits a substantial increase in expression reaching peak levels about 48-96 hours post-inoculation during the wheat infection process.</li> <li>• <a href="#">Transcriptomic analysis during vegetative and infectious growth of <i>Fusarium graminearum</i> PH-1 (RNA-Seq)</a> This gene exhibits significantly higher expression in infected conditions compared to vegetative conditions.</li> <li>• <a href="#">Gene expression of <i>F. graminearum</i> during maize stalk infection (array)</a> The gene exhibits an increase in expression over time, peaking significantly at 144 hours after inoculation in the dataset.</li> <li>• <a href="#">Fusarium graminearum transcriptome during symptomless and symptomatic wheat infection (array)</a> The gene is highly expressed in symptomatic samples and progressively less expressed in intermediate and symptomless conditions, indicating downregulation correlating with reduced symptom severity.</li> <li>• <a href="#">Transcriptome in response to an active or passive host (RNA-Seq)</a> The gene is substantially upregulated in <i>Fusarium graminearum</i> when interacting with dead plant tissues compared to alive plant tissues or control conditions.</li> </ul>	The gene is expressed significantly during wheat and maize infections, implicating its potential role in plant-pathogen interactions.	5
<p>▼ Developmental Stage-Specific Expression in <i>Fusarium graminearum</i></p> <ul style="list-style-type: none"> <li>• <a href="#">RNA-Seq analysis of transcriptomes at 1 day after sexual induction (RNA-Seq)</a> In this experiment, the gene exhibits moderate upregulation in the mutant strain 'delta abaA' relative to the wild-type strain.</li> <li>• <a href="#">Stage-specific and genome-wide gene expression during the different developmental stages of <i>Fusarium graminearum</i> (RNA-Seq)</a> The gene exhibits higher expression levels in the conidia and perithecia stages when compared to the vegetative hyphae stage under the examined conditions in <i>Fusarium graminearum</i>.</li> <li>• <a href="#">RNA-seq Analysis of FlbD-, AbaA-, and WetA-dependent Genes in <i>Fusarium graminearum</i> Z-3639 (RNA-Seq)</a> This gene shows notable differential expression under specific conditions related to the gene deletions studied in <i>Fusarium graminearum</i>.</li> <li>• <a href="#">RNA-seq analysis of FSS1-dependent genes in <i>Fusarium graminearum</i> Z-3639 (RNA-Seq)</a> The gene shows increased expression in the <math>\Delta fss1</math> mutant strain compared to the wild type and the <i>fss1</i> overexpression strain, while all other conditions showed negligible levels.</li> </ul>	This gene shows higher expression during reproductive stages (conidia and perithecia), suggesting developmental importance.	4
<p>▼ Other</p> <ul style="list-style-type: none"> <li>• <a href="#">Genome-wide gene expression analysis of spores and mycelium (RNA-Seq)</a> This gene exhibits comparable expression levels across spores and mycelium stages in <i>Fusarium graminearum</i>.</li> <li>• <a href="#">Transcriptome of heat-stressed <i>Fusarium graminearum</i> Z-3639 (RNA-Seq)</a> This gene shows no detectable expression across all experimental conditions tested in <i>Fusarium graminearum</i>.</li> <li>• <a href="#">Fusarium transcriptome in low and high nitrogen conditions (RNA-Seq)</a> This gene shows no significant expression above background levels across all conditions in this experiment.</li> </ul>	The AI ordered these experiments by biological importance but did not group them into topics.	3

NCU09033



[expand all](#) | [collapse all](#)

 [Search section names...](#)

-  [Gene models](#)
-  [Annotation, curation and identifiers](#)
-  [Link outs](#)
-  [Genomic Location](#)
-  [Literature](#)
-  [Taxonomy](#)
-  [Orthology and synteny](#)
-  [Phenotype](#)
-  [Genetic variation](#)
-  [Transcriptomics](#)

# Microarray Datasets

# RNASeq Datasets

AI Expression Summary

RNA-Seq Transcription Summary

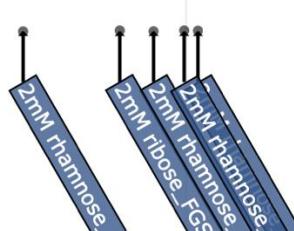
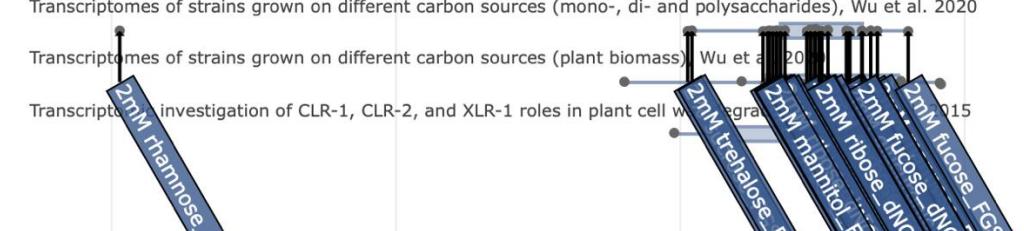
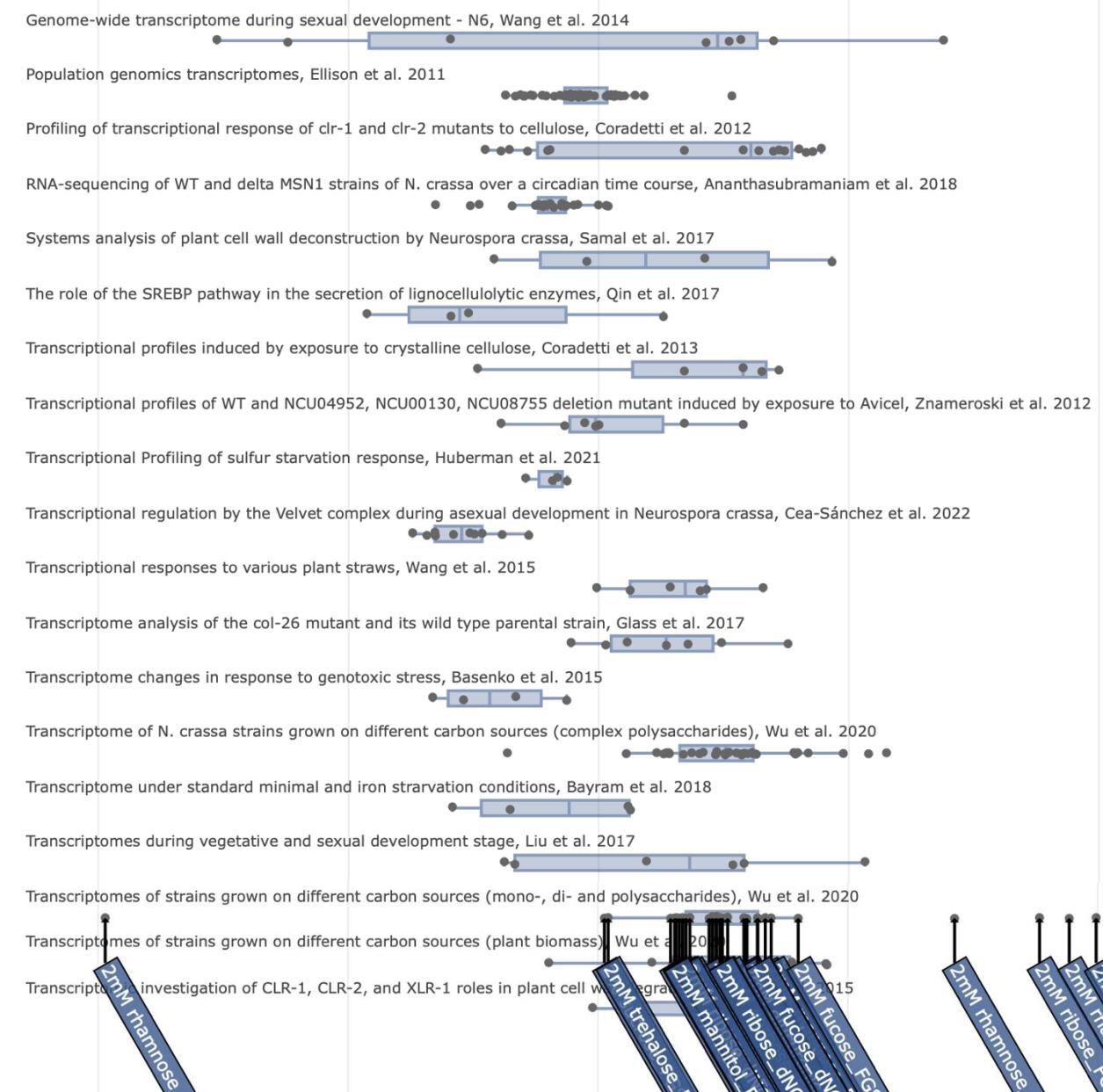
Transcript Expression

User Dataset Transcriptomics Graphs

-  [Sequences](#)

-  [Sequence analysis](#)

-  [Structure analysis](#)



**FungiDB** Ref ID: NCU09033

Gene expression during conidia germination in *Neurospora crassa* on different media

Help Subscriptions Contact Us

AI summary

This gene exhibits significant upregulation during the conidia germination of *Neurospora crassa* on many media, particularly at the later stages.

Experimental data

tpm\_sense - NCU09033

Condition	TPM
Bird_15min	~5
Bird_120min	~5
Bird_240min	~6
Bird_360min	~7
Maple sap_15min	~5
Maple sap_120min	~10
Maple sap_240min	~15
Maple sap_360min	~30

Gene expression

AI summary

The gene exhibits increased expression during conidia germination, with stronger expression on Bird (B) stages.

Experimental data

tpm\_sense

Condition	TPM
conidia_15m_B	~5.0
polar_growth_120m_B	~8.5
long_axis_240	~8.0

The results from the AI analysis

AI summary

The examined gene in *Neurospora crassa* shows significant upregulation during conidia germination.

Experimental data

tpm\_sense - NCU09033

Condition	TPM
no carbon_FGSC 2489	~35
no carbon_dhNCU02307	~28
1% inulin_FGSC 2489	~40
1% avicel_dhNCU06173	~22
1% avicel_dhNCU03699	~22
1% avicel_FGSC 2489	~30
1% avicel_d07705	~32
1% avicel_d08042	~30
1% xylan_dhNCU05909	~28
1% xylan_FGSC 2489	~38
1% xylan_d01074	~35
1% xyloglucan_FGSC 2489	~30
1% galactomannan_FGSC 2489	~32
Pectin	~38

VEuPathDB Eukaryotic Pathogen, Vector & Host Informatics Resources

How well do you know this particular gene?

Not familiar

Very familiar

How accurate do you think the overall AI Expression Summary for this gene is?

- Accurate
- Partially accurate
- Inaccurate
- Not sure
- Other

AI organizes experiments into topics (see the table below the overall summary). Is the first organizational topic the most important conclusion that can be drawn about this gene, based on the available expression data?

- Yes
- No
- Not sure
- Other

How about the rest of the organizational topics in that table? Are they useful for this gene?

- Useful
- Not useful
- Not sure
- Other

After exploring the results of the AI summary, did you learn something new about this gene or was the summary redundant?

- I learned something new

Was this AI expression summary for NCU09033 helpful?

Thumbs Up Thumbs Down

AI content warning

Was this AI expression summary for NCU09033 helpful?

Comments Count: 0

# Datasets

4A 2

3

nonunique unique

489 FGSC 2489 dNcu06173 dNcu07728 dNcu01209

no carbon Link Glucan\_dhNCU01209

1% inulin Link Glucan\_dhNCU01209

1% avicel Link Glucan\_dhNCU01209

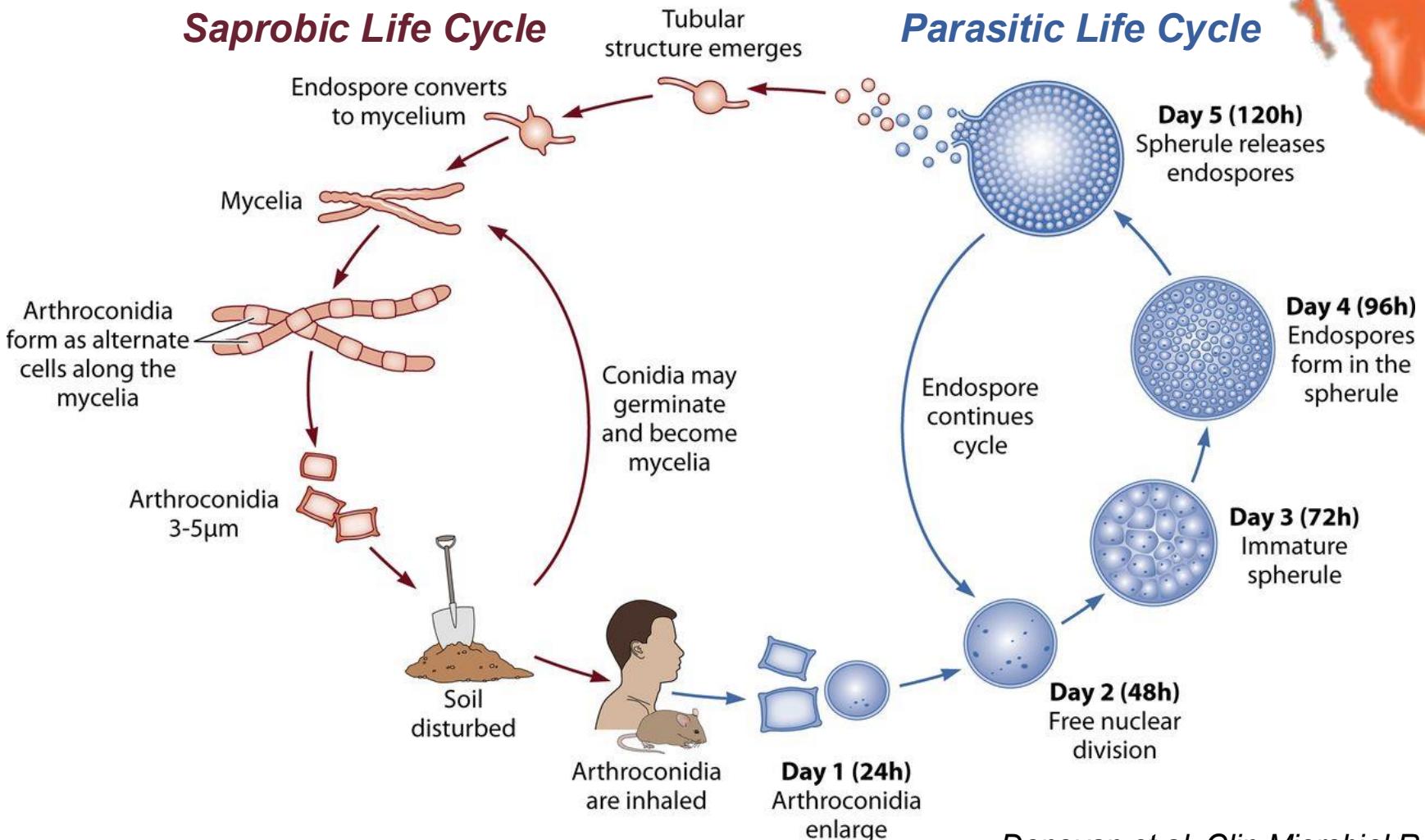
1% xylan Link Glucan\_dhNCU01209

1% xyloglucan Link Glucan\_dhNCU01209

1% galactomannan Link Glucan\_dhNCU01209

Pectin

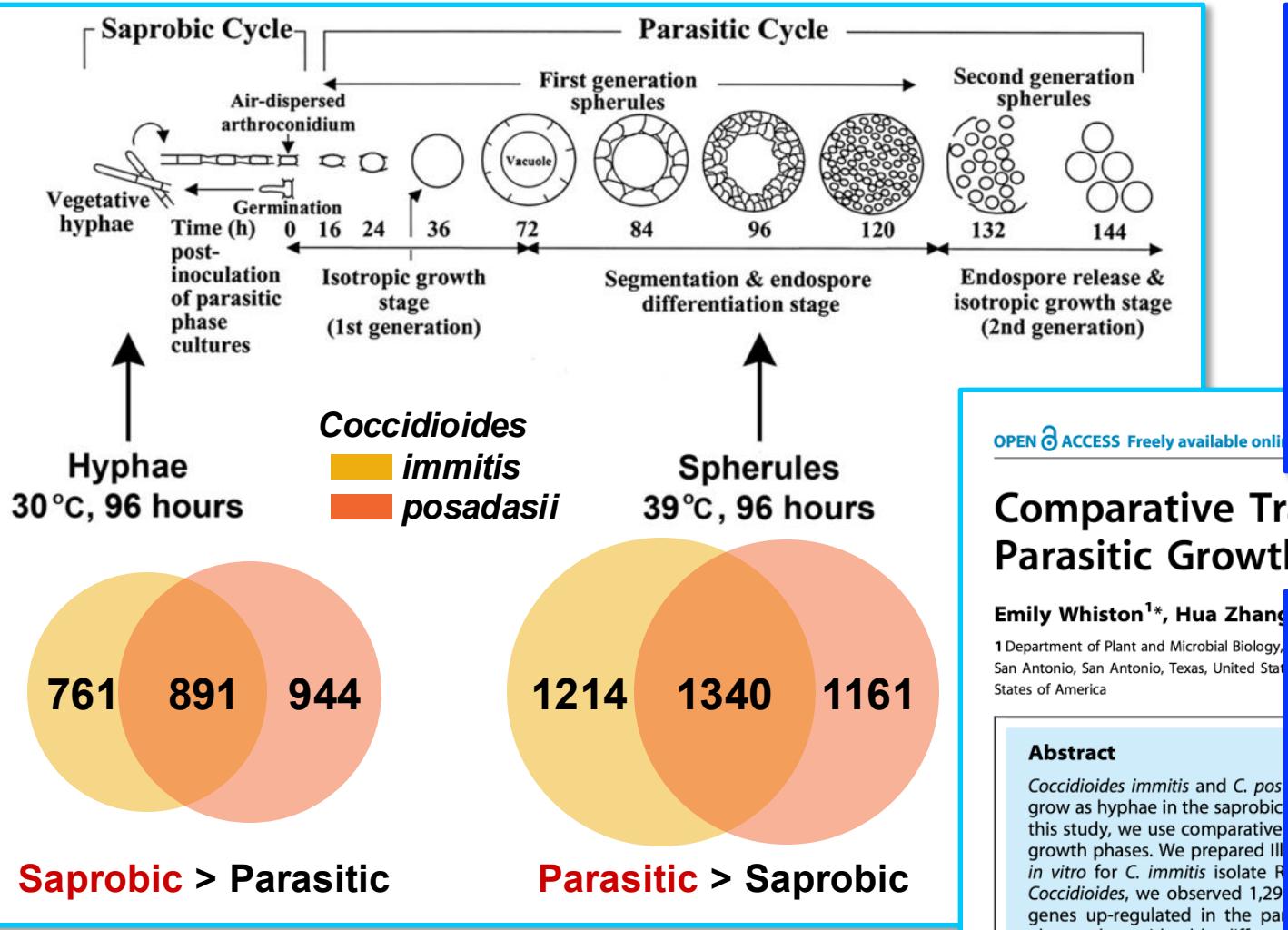
# What molecular changes are associated with pathogenic coccidiomycosis?



Donovan et al, Clin Microbiol Rev 2019

Ashraf et al, Mycopathologica 2020

# *RNA-seq analysis of stage-specific transcription in parasitic vs saprobic stage *C. immitis* & *posadasii**



**Table 1.** Top 15 genes with significantly higher expression (up-regulated) in the saprobic phase.

**Up in SAPROBIC stage *Coccidioides immitis* & *C.posadasii***

Fold difference*	Annotation	Gene ID
185	Conserved protein (insect antifreeze protein repeat domain, predicted secreted)	CIMG_00925
166	Acetyltransferase	CIMG_07556
106	Acetamidase	CIMG_02374
101	Conserved hypothetical protein (predicted secreted)	CIMG_03870
94	Fungal hydrophobin (predicted secreted)	CIMG_06615
75	Conserved protein (PAN domain, predicted secreted)	CIMG_09824
53	Conserved protein (zinc-finger domain)	CIMG_00099
53	Conserved hypothetical protein	CIMG_06344
43	Putative serine proteinase	CIMG_09304
41	Cell wall synthesis protein (beta-glucosidase domain, SUN family, predicted secreted)	CIMG_05254
34	Hypothetical protein (predicted secreted)	CIMG_07839
31	Hypothetical protein	CIMG_13374
31	Helix-loop-helix transcription factor	CIMG_02390
29	Conserved hypothetical protein (pyridine nucleotide-disulphide oxidoreductase domain, predicted secreted)	CIMG_07557
24	Prp4 (CROW domain-containing protein, predicted secreted)	CIMG_07303

# Comparative Transcriptomics of the Saprobic and Parasitic Growth Phases in *Coccidioides* spp

Emily Whiston<sup>1\*</sup>, Hua Zhan

**1** Department of Plant and Microbial Biology,  
San Antonio, San Antonio, Texas, United States  
of America

### **Abstract**

*Coccidioides immitis* and *C. posadasii* grow as hyphae in the saprobe. In this study, we use comparative genomics to study growth phases. We prepared cultures *in vitro* for *C. immitis* isolate B-92. In *Coccidioides*, we observed 1,200 genes up-regulated in the parasitic stage. We observed considerable differential expression of genes between the two pathogens. These included algal-like genes. We observed differential expression of genes involved in comparative genomics studies and virulence in *Coccidioides*. Our results provide context for studies in

**Table 2.** Top 15 genes with significantly higher expression (up-regulated) in the parasitic phase

Up in PARASITIC stage *Coccidioides immitis* & *C.posadasii*

Fold difference*	Annotation	Gene ID
125	Conserved hypothetical protein (DUF 536)	CIMG_09539
68	Heat shock protein 30 (Hsp20/alpha-crystallin domain)	CIMG_01749
31	Conserved hypothetical protein	CIMG_12822
30	Conserved hypothetical protein (YCI1-related domain)	CIMG_07089
29	Conserved hypothetical protein	CIMG_13084
28	Hypothetical protein	CIMG_11522
28	Hypothetical protein	CIMG_05235
26	Polysaccharide deacetylase (Arp2/3 complex subunit Arc16)	CIMG_02628
24	Conserved hypothetical protein (predicted secreted)	CIMG_00509
23	Hypothetical protein	CIMG_11203
19	Spherule outer-wall glycoprotein (SOWgp, predicted secreted)	CIMG_04613
18	Conserved hypothetical protein	CIMG_10488
18	Hypothetical protein	CIMG_10670
17	Sphingosine hydroxylase	CIMG_01209
17	Conserved hypothetical protein	CIMG_04740

# Query Strategies: experiments in silico

My Strategies Searches Tools My Workspace Data About Help Contact Us



Search for...

expand all | collapse all

Filter the searches below...

▼ Genes

- ▶ Annotation, curation and identifiers
- ▶ Function prediction
- ▶ Gene models
- ▶ Genetic variation
- ▶ Genomic Location
- ▶ Immunology
- ▶ Orthology and synteny
- ▶ Pathways and interactions
- ▶ Phenotype
- ▶ Protein features and properties
- ▶ Protein targeting and localization
- ▶ Proteomics
- ▶ Sequence analysis
- ▶ Structure analysis
- ▶ Taxonomy
- ▶ Text
- ▶ Transcriptomics
- ▶ Organisms



## Getting Started

VEuPathDB is packed with data, tools and workflows, and present the results for you to work with.

**SITE SEARCH:** Explore the site; find what you want.

Enter a term or ID in the site search bar to find matches. It's easy to find genes, pathways,

[Read More](#)

## Tutorials and Exercises

### Apollo: Manual gene annotation

Structural and functional community curation with Apollo, a real time collaborative genome annotation and exploration.

### Gene Pages

Explore the data and images on gene pages

[Gene Pages](#)

expand all | collapse all

Filter the searches below...

▼ Genes

- ▶ Annotation, curation and identifiers
- ▶ Function prediction
- ▶ Gene models
- ▶ Genetic variation
- ▶ Genomic Location
- ▶ Immunology
- ▶ Orthology and synteny
- ▶ Pathways and interactions
- ▶ Phenotype
- ▶ Protein features and properties
- ▶ Protein targeting and localization
- ▶ Proteomics
- ▶ Sequence analysis
- ▶ Structure analysis
- ▶ Taxonomy
- ▶ Text
- ▶ Transcriptomics
- ▶ Organisms
- ▶ Popset Isolate Sequences
- ▶ Genomic Sequences
- ▶ Genomic Segments
- ▶ SNPs
- ▶ ESTs
- ▶ Metabolic Pathways
- ▶ Compounds



Differential Expression Fold Change Percentile

Identify Genes based on C. posadasii C735 delta SOWgp Saprobic vs Parasitic Growth RNA-Seq (fold change)

[Reset values](#)

For the Experiment

Saprobic vs Parasitic Growth unstranded

return protein coding Genes  
that are up-regulated  
with a Fold change >= 8  
between each gene's minimum expression value  
(or a Floor of 10 reads)  
in the following Reference Samples

Saprobic Hyphae  
 Parasitic Spherules

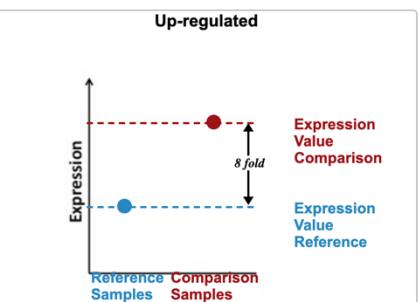
and its maximum expression value  
(or the Floor selected above)

in the following Comparison Samples

Saprobic Hyphae  
 Parasitic Spherules

Example showing one gene that would meet search criteria

(Dots represent this gene's expression values for selected samples)



For each gene, the search calculates:

$$\text{fold change} = \frac{\text{comparison expression value}}{\text{reference expression value}}$$

and returns genes when fold change >= 8.

You are searching for genes that are up-regulated between one reference sample and one comparison sample.

[Get Answer](#)

Build a Web Services URL from this Search &gt;

Give this search a name (optional)

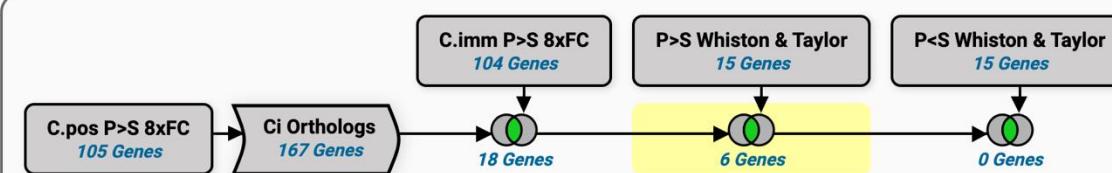
Give this search a weight (optional)



Opened (1) All (312) Public (70) Help

Coccidioides Parasitic > Saproic FC (Whiston & Taylor) (6) \*

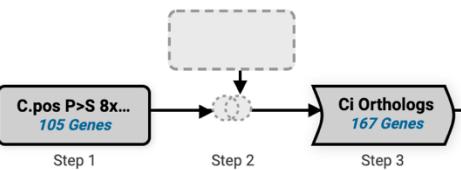
[Edit](#)



[Edit](#) [Save](#) [Share](#) [Delete](#) | [X](#)

<https://tinyurl.com/CoccidioidesPvStrcripts>

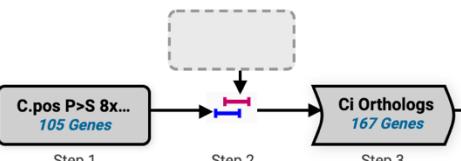
### Combine with other Genes



### Transform into related records

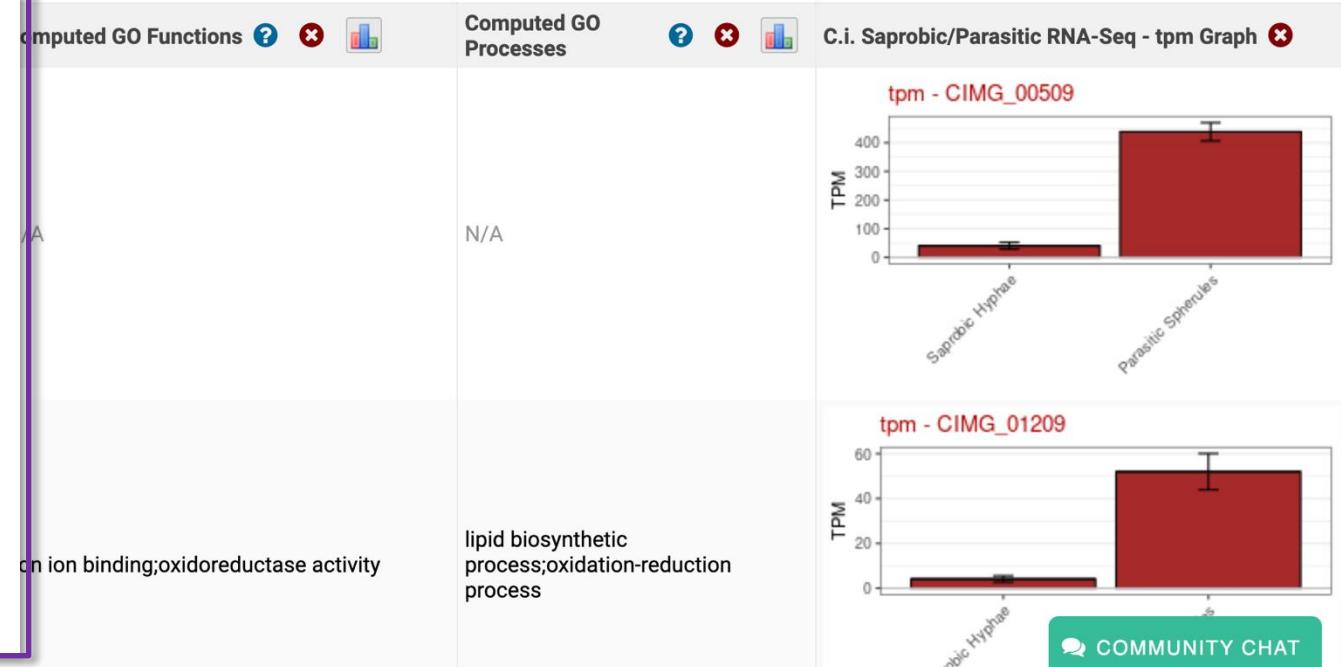


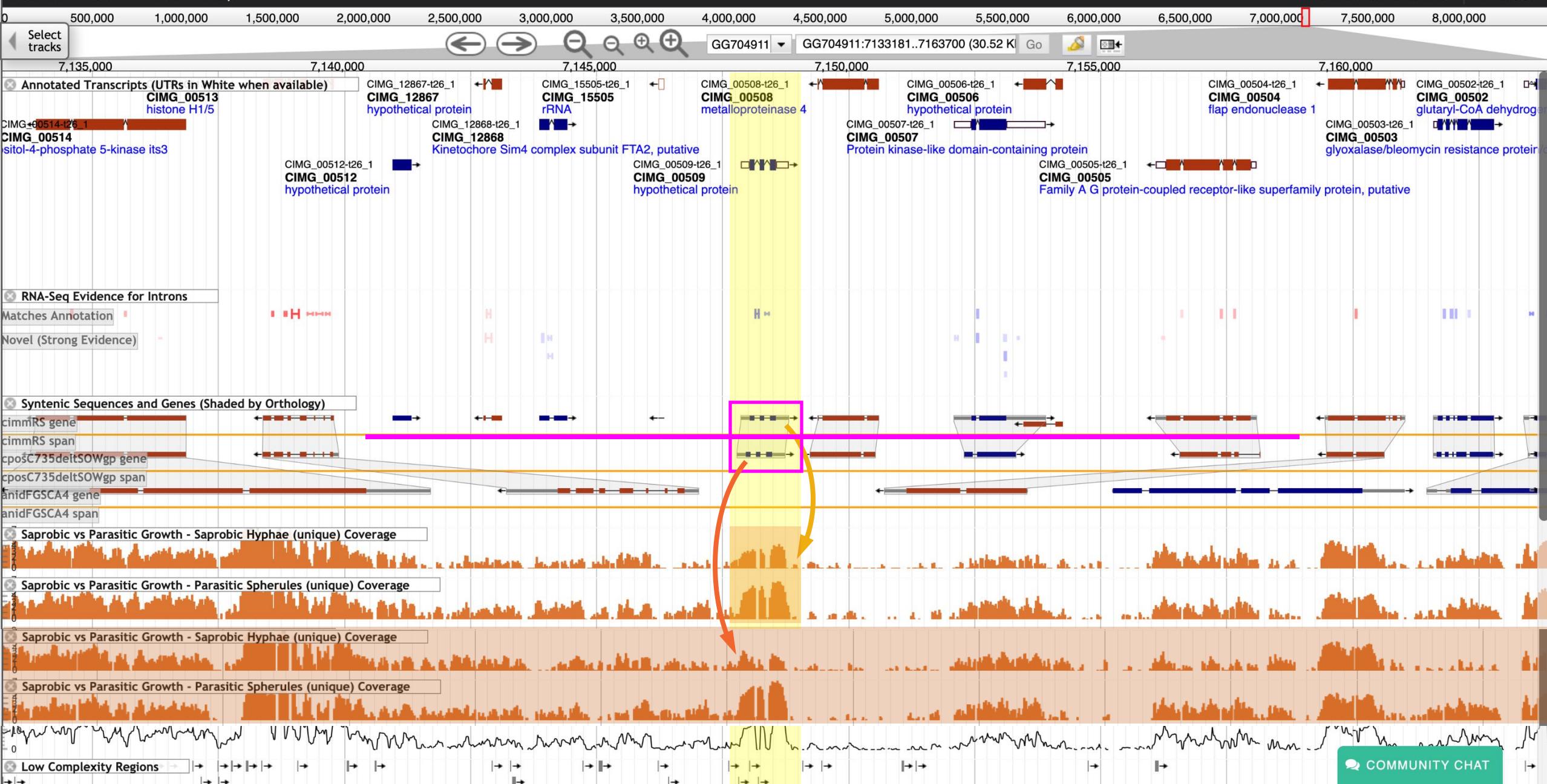
### Use Genomic Colocation to combine with other features



Transform 105 Genes into...

Orthologs





Opened (1) All (312) Public (70) Help

Coccidioides Parasitic > Saproic FC (Whiston & Taylor) (6) \*



9 Genes (9 ortholog groups)

! Some Genes in your combined result have Transcripts that were not returned by one or both of the two input searches. [Explore](#)

Gene Results Genome View Analyze Results

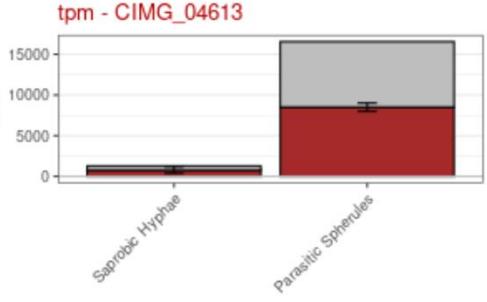
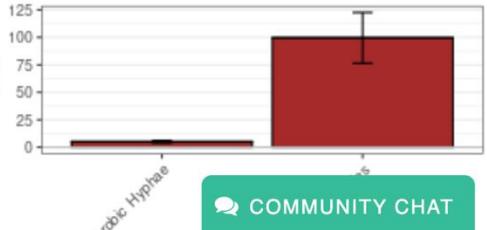
Rows per page: 50

 Download

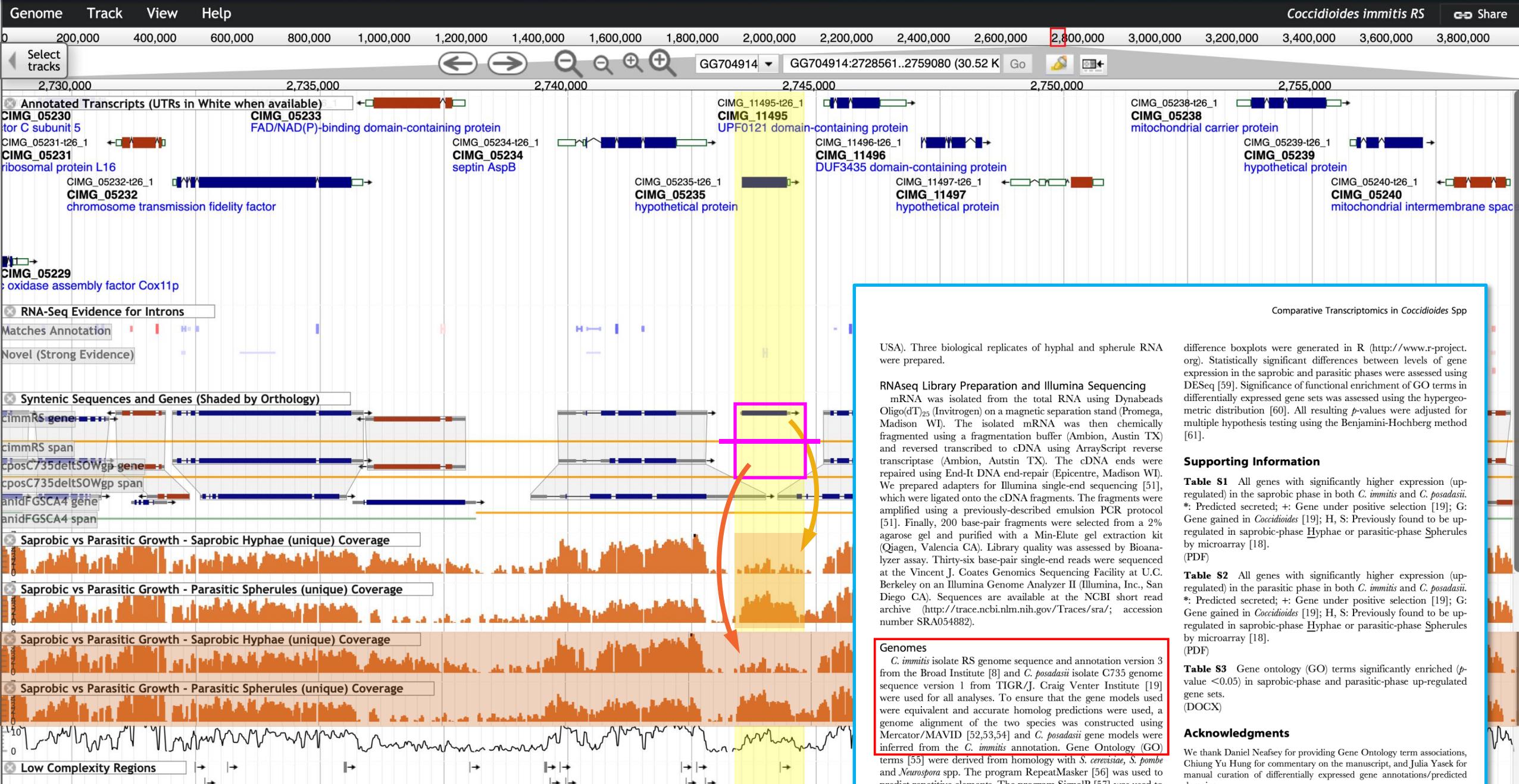
 Add to Basket

 Add Columns

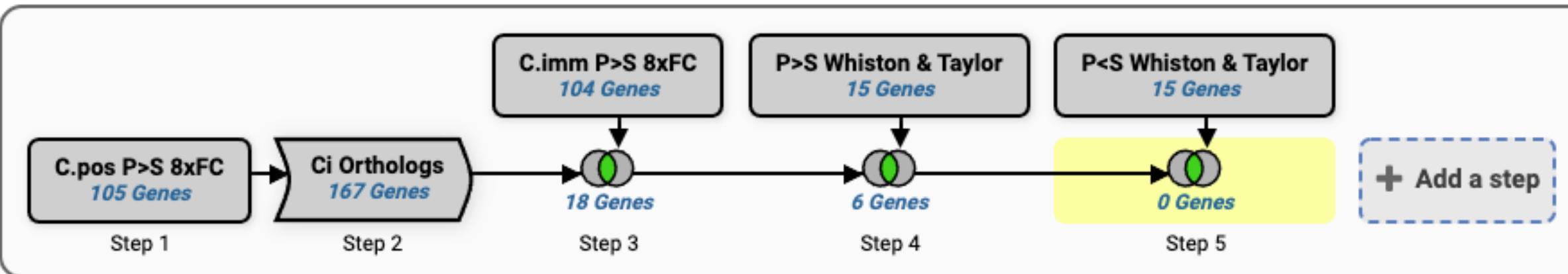
Show Organism Filter

	Gene ID	Transcript ID	Genomic Location (Gene)	Product Description	Computed GO Functions	Computed GO Processes	C.i. Saproic/Parasitic RNA-Seq - tpm Graph						
	CIMG_04613	CIMG_04613-t26_1	GG704914:970,094..971,690(+)	hypothetical protein	N/A	N/A	 <p>tpm - CIMG_04613</p> <table border="1"> <caption>Estimated TPM values for CIMG_04613</caption> <thead> <tr> <th>Condition</th> <th>TPM</th> </tr> </thead> <tbody> <tr> <td>Saproic Hyphae</td> <td>~100</td> </tr> <tr> <td>Parasitic Spherules</td> <td>~8000</td> </tr> </tbody> </table>	Condition	TPM	Saproic Hyphae	~100	Parasitic Spherules	~8000
Condition	TPM												
Saproic Hyphae	~100												
Parasitic Spherules	~8000												
	CIMG_05235	CIMG_05235-t26_1	GG704914:2,743,484..2,744,494(+)	hypothetical protein	N/A	N/A	 <p>tpm - CIMG_05235</p> <table border="1"> <caption>Estimated TPM values for CIMG_05235</caption> <thead> <tr> <th>Condition</th> <th>TPM</th> </tr> </thead> <tbody> <tr> <td>Saproic Hyphae</td> <td>~10</td> </tr> <tr> <td>Parasitic Spherules</td> <td>~100</td> </tr> </tbody> </table>	Condition	TPM	Saproic Hyphae	~10	Parasitic Spherules	~100
Condition	TPM												
Saproic Hyphae	~10												
Parasitic Spherules	~100												

COMMUNITY CHAT



## Coccidioides Parasitic > Saproic FC (Whiston & Taylor 2012) \*



<https://tinyurl.com/CoccidioidesPvTranscripts>

### RNA-Seq Evidence search

Dataset = *C. posadasii*

C735 delta SOWgp

Saprobic vs Parasitic

**8 fold up-regulated genes**

Reference = Saprobic

Comparison = Parasitic

### Transform into Related Records (Orthologs)

Organism = *C. immitis*

### RNA-Seq Evidence search

*Identical to Step 1, except:*

Dataset = *C. immitis*

Saprobic vs Parasitic

Boolean operator =

intersection

### GeneID(s) search

Dataset = list of 15 gene

IDs reported by Whiston *et al* as most up-regulated in the **Parasitic** stage

Boolean operator =

intersection

### GeneID(s) search

*Identical to Step 4, except:*

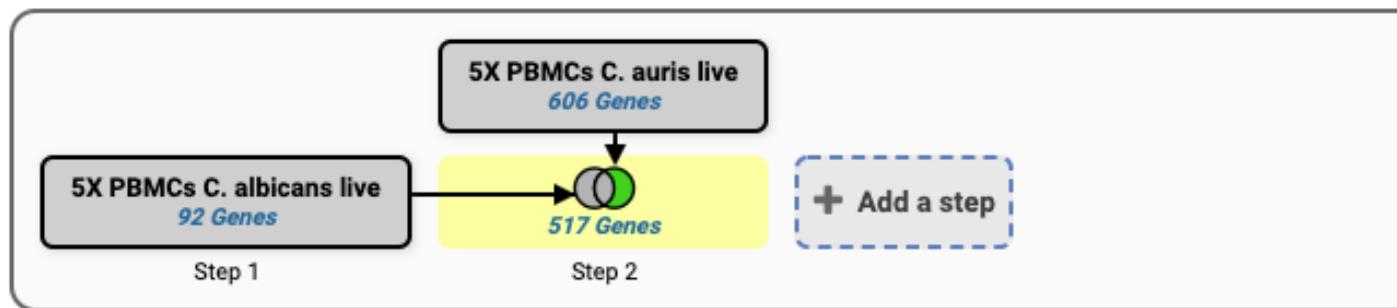
Dataset = list of genes

most up-regulated in the

**Saprobic** stage

**Exercise #1:** Modify Boolean operation in Step 4 to identify genes found by Whiston *et al*, but **not** identified by this Query Strategy.

**Exercise #2:** Reconfigure searches in Steps 1 & 3 to identify genes up-regulated in **Saprobic** stage. Also: modify boolean operator in Step 4 to ignore this step, so as to compare results with the list reported as up-regulated in Saprobic stages by Whiston *et al*.

MOMY2021 demo: PBMCs stimulated with *C. albicans* and *C. auris* live 

<https://hostdb.org/hostdb/app/workspace/strategies/import/0d7b69add3c0632c>

RNA-Seq Evidence search

**Dataset** = PBMCs stimulated for 4h and 24h with live *Candida* spp. and their cell wall components

**5 fold up-regulated genes**

**Reference** = RPMI\_24hr

**Comparison** = calb\_live\_24hr

RNA-Seq Evidence search

**Dataset** = PBMCs stimulated for 4h and 24h with live *Candida* spp. and their cell wall components

**5 fold up-regulated genes**

**Reference** = RPMI\_24hr

**Comparison** = caur\_KTClive\_24hr

nature > nature microbiology > articles > article

Article | Published: 24 August 2020

## Transcriptional and functional insights into the host immune response against the emerging fungal pathogen *Candida auris*

Mariolina Bruno , Simone Kersten, Judith M. Bain, Martin Jaeger, Diletta Rosati, Michael D. Krappa, Douglas W. Lowman, Peter J. Rice, Bridget Graves, Zuchao Ma, Yue Ning Jiao, Anuradha Chowdhary, George Renieris, Frank L. van de Veerdonk, Bart-Jan Kullberg, Evangelos J. Giamarellos-Bourboulis, Alexander Hoischen, Neil A. R. Gow, Alistair J. P. Brown, Jacques F. Meis, David L. Williams & Mihai G. Netea 

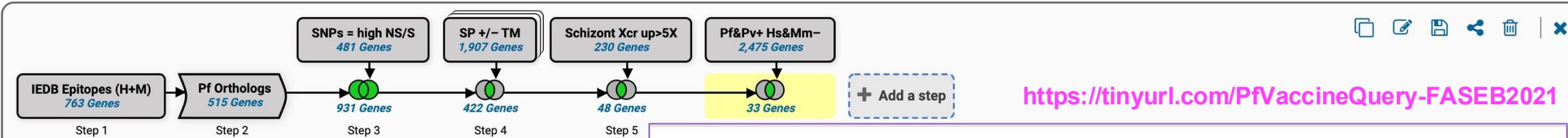
*Nature Microbiology* 5, 1516–1531 (2020) | Cite this article

3242 Accesses | 12 Citations | 83 Altmetric | Metrics

# My Search Strategies

[Opened \(2\)](#) All (378) Public (51) Help

Plasmodium vaccine antigens?



33 Genes (24 ortholog groups)

active targets (positive controls)  
 other targets of interest  
 novel candidates?

[Gene Results](#) [Genome View](#) [Analyze Results](#)
Genes: 33 Transcripts: 34 (hiding 1)  Show Only One Transcript Per Gene

Rows per page: 50

Gene ID	Transcript ID	Product Description	?	X	grid icon
PF3D7_1133400	PF3D7_1133400.1	apical membrane antigen 1			
PF3D7_1352900	PF3D7_1352900.1	Plasmodium exported protein, unknown function			
PF3D7_0202100	PF3D7_0202100.1	liver stage associated protein 2			
PF3D7_0930300	PF3D7_0930300.1	merozoite surface protein 1			
PF3D7_0508000	PF3D7_0508000.1	6-cysteine protein			
PF3D7_1335100	PF3D7_1335100.1	merozoite surface protein 7			
PF3D7_0207500	PF3D7_0207500.1	serine repeat antigen 6			
PF3D7_0831600	PF3D7_0831600.1	cytoadherence linked asexual protein 8			

Research a

Immunoreactive (manual curation, from Immune Epitope DB)

*P. falciparum* Orthologs (of any *Plasmodium* antigen)

Under diversifying selection (dN/dS)

Surface expression (SP ± TM)

Expressed at the desired stage (schizonts)

Orthologs in *P. falciparum* & *P. vivax*; not in humans/mice

2344 Mol. Biol. Evol. 27(10):2344–2351. 2010 doi:10.1093/molbev/msq119 Advance Access publication May 9, 2010

## Allele Frequency-Based and Polymorphism-Versus-Divergence Indices of Balancing Selection in a New Filtered Set of Polymorphic Genes in *Plasmodium falciparum*

Lynette Isabella Ochola,<sup>1</sup> Kevin K. A. Tetteh,<sup>2</sup> Lindsay B. Stewart,<sup>2</sup> Victor Riitho,<sup>1</sup> Kevin Marsh,<sup>1</sup> and David J. Conway<sup>\*2</sup>

<sup>1</sup>Kenya Medical Research Institute, Centre for Geographic Medicine Research Coast, Kilifi, Kenya

<sup>2</sup>Department of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, London, United Kingdom

\*Corresponding author: E-mail: dconway@mrcgm, david.conway@lshtm.ac.uk.

Associate editor: John H. McDonald

### Abstract

Signatures of balancing selection operating on specific gene loci in endemic pathogens can identify candidate targets of naturally acquired immunity. In malaria parasites, several leading vaccine candidates convincingly show such signatures when subjected to several tests of neutrality, but the discovery of new targets affected by selection to a similar extent has been slow. A small minority of all genes are under such selection, as indicated by a recent study of 26 *Plasmodium falciparum* merozoite-stage genes that were not previously prioritized as vaccine candidates, of which only one (locus PF10\_0348) showed a strong signature. Therefore, to focus discovery efforts on genes that are polymorphic, we scanned all available shotgun genome

# Did you know that you can ...

[My Strategies](#) [Searches](#) [Tools](#) [My Workspace](#) [Data](#) [About](#) [Help](#) [Contact Us](#)


## Search for...

[expand all](#) | [collapse all](#)
 Filter the searches below...

### Genes

- ▶ Annotation, curation and identifiers
- ▶ Epigenomics
- ▶ Function prediction
- ▶ Gene models
- ▶ Genetic variation
- ▶ Genomic Location
- ▶ Immunology
- ▶ Orthology and synteny
- ▶ Pathways and interactions
- ▶ Phenotype
- ▶ Protein features and properties
- ▶ Protein targeting and localization
- ▶ Proteomics
- ▶ Sequence analysis
- ▶ Structure analysis
- ▶ Taxonomy
- ▶ Text
- ▶ Transcriptomics

### Organisms

- ▶ Popset Isolate Sequences
- ▶ Genomic Sequences
- ▶ Genomic Segments
- ▶ SNPs
- ▶ ESTs

- ❖ **Explore diverse datatypes**, identifying genes of interest based on signatures of selection, species distribution, expression patterns, subcellular localization, phenotypic impact, etc
- ❖ **Leverage orthology** ... to infer structure & function based on data in related species
- ❖ **Visualize & interrogate expression profiles** ... based on multiple transcriptomic and proteomic datasets (now including single-cell RNAseq data!)
- ❖ **Assess gene annotation accuracy**, and revise (if warranted) using the Apollo annotation platform, and/or add expert knowledge (references, functions, etc) as User Comments
- ❖ **Define your own questions *in silico*** ... share with friends, colleagues, reviewers, readers
- ❖ **Determine GO-term & metabolic pathway enrichment** ... associated with your gene lists
- ❖ **Analyze your own data (privately)** ... map RNAseq reads, or identify SNPs in mutant strains or field isolates, via a user-friendly interface to a cloud-based Galaxy platform
- ❖ **Download query results and datasets for further analysis off-line**
- ❖ **Propose features or datasets** (of your own or others) for incorporation into VEuPathDB
- ❖ **Obtain further help:** on-line resources, email support, Virtual Lab Meeting with your group ... ask during breaks, poster sessions, or click the 'Contact Us' link