

FungiDB: Navigating gene record pages

1. Components of gene record pages

Navigate to the gene records page for *Aspergillus fumigatus* **Afu2g13260**, which is known to be important in virulence of the filamentous fungus *Aspergillus fumigatus*.

- You can search for the gene using the **Gene ID** shortcut search at the top of the FungiDB main page:

A screenshot of the FungiDB gene record page for Afu2g13260. The page title is 'Afu2g13260 Putative regulator of adherence, host cell interactions and virulence'. The 'Name' is 'medA', 'Type' is 'protein coding', 'Chromosome' is '2', and 'Location' is 'Chr2_A_fumigatus_Af293:3,408,135..3,413,423(-)'. The 'Species' is 'Aspergillus fumigatus' and the 'Strain' is 'Af293'. The 'Status' is 'Curated Reference Strain'. A 'Contents' menu on the left lists various data types like 'Gene models', 'Transcriptomics', and 'Proteomics'. A 'Shortcuts' section at the top right provides quick access to 'Synteny', 'BLAT Alignments', 'SNPs', 'Transcriptomics', and 'Protein Features'. Below this, there is a 'Gene models' section showing '1 Gene models' and '5 Exons in Gene 5'. A 'View in genome browser' section shows a genomic track with 'Annotated Transcripts' and 'RNASeq evidence for introns'.

- Notice that **Shortcuts** at the top of the gene page will take you directly to the section containing SNPs, Transcriptomics, Protein Features, etc., while the **Contents** menu on the left brings up additional functionality and data.
- Note: Each thumbnail within the Shortcuts shows the number of datasets available within a given category. Clicking on the magnifying glass symbol within the thumbnail will open a preview screen of the evidence.*
- The top section of the gene page also provides information about gene name, type (protein coding, etc.), chromosome number and exact location. If a strain is actively under manual EuPathDB curation it will be indicated as shown below as well.
- User comments can be added via the *Add the user comment* link. User comments help improve genome annotations: provide evidence for alternative gene models,

phenotypes, subcellular localization, gene products, *etc.* User comments can be added in bulk (contact help@fungidb.org), and can be edited at any time.

- Gene record page can be saved by clicking on the **Add to basket** link. To bookmark a gene of interest, click on **Add to favorites**.
 - The *Add to basket* function saves the gene record to a basket associated with a user's account. Basket items can be found in the My Strategies section when a user is logged in. The basket serves as a shopping cart where genes in the basket can be downloaded or transferred to a search strategy.
 - Adding a gene to favorites creates a bookmark to that gene in the My Favorites section available within the grey menu bar. In the My Favorites section, users can also add private notes and project descriptions about saved items.
- Practice to export gene records. Click on the **Download Gene** will open up a selection page with available gene records for download, including sequences in FASTA (highlighted):

Download Gene: Afu2g13260

Choose a Report: Text: choose from columns and/or tables [?](#)
 FASTA (sequence retrieval, configurable) [?](#)

Choose Attributes select all | clear all | expand all | collapse all

Search Attributes... [?](#)

- Gene models
- Annotation, curation and identifiers
- Genomic Location
- Orthology and synteny
- Genetic variation
- Organism

select all | clear all | expand all | collapse all

Download Type

Text File
 Show in Browser

Additional Options

Include empty tables

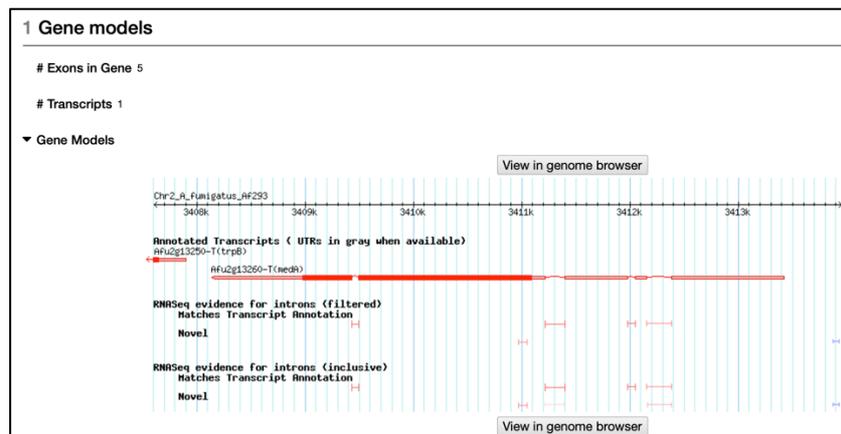
Choose Tables select all | clear all | expand all | collapse all

Search Tables... [?](#)

- Gene models
- Annotation, curation and identifiers
- Link outs
- Genomic Location
- Literature
- Orthology and synteny
- Phenotype
- Sequence analysis
- Structure analysis
- Protein features and properties
- Protein targeting and localization
- Function prediction
- Pathways and interactions
- Proteomics
- Immunology

- Return to the gene record page and look at the **Contents** section on the left. The individual sections comprising the Contents can be quickly identified by using the search box at the top of the Contents menu. In addition, sections in a gene page can be hidden by unchecking the box to the right of any section name in the Contents menu.

- The **Gene Models** section contains information about the structure of the gene such as exon count, transcript number, annotated UTRs and introns (when available), alternative transcripts, visual GBrowse representation of gene location, *etc.* The color of the gene represents which strand it is encoded on. Red is on the reverse strand (from right to left) and blue is on the forward strand (from left to right). More details are available via the **View in genome browser** button.



- How many introns are reported for this gene?
 - 2
 - 4
 - 9
- Is there any RNA-Seq evidence supporting the intron data? Is there evidence for a possible alternative gene model (introns with RNAseq evidence that are not part of the official gene model)?
- The **Annotation, curation and identifiers** section offers alternate product descriptions, previous identifiers, and aliases, and is populated using data from internal curation, other fungal specialized resources (*e.g.* AspGD, Ensembl, *etc.*), or user-submitted data (user comments, also see below).

2 Annotation, curation and identifiers

▼ **Alternate Product Descriptions** [Data sets](#)
No data available

External DB Version 2015-09-27

Gene Name or Symbol medA

▼ **Names, Previous Identifiers, and Aliases** [Download](#) [Data sets](#)
Search this table... Showing 4 rows

Name/ID/Alias	Type
MEDA	name
AFUA_2G13260	previous ID
AFUB_028890	previous ID
CADAFUAG00004866	previous ID

▼ **Notes from Annotator** [Download](#) [Data sets](#)
Search this table... Showing 3 rows

Transcript ID(s)	Date	Note
Afu2g13260-T	2010-07-28	NOTE TYPE=Standard Name NOTE=medA: 19889083 Gravelat FN, et al. (2010) Aspergillus fumigatus MedA governs adherence, host cell interactions and virulence. Cell Microbiol 12(4):473-88
Afu2g13260-T	2011-01-08	NOTE TYPE=Sequence Annotation Notes NOTE=Sequence information added to the database.
Afu2g13260-T	2012-03-06	NOTE TYPE=Alias Name NOTE=CADAFUAG00004866: AspGD (2012) Update of Aspergillus fumigatus Af293 reference annotation

- The **Link outs** section offers redirection to other resources (e.g. CGD, Ensembl, MycoCosm, etc.).
 - Find curated information for host-pathogen interactions

Hint: Click on the PHI-Base link out.

Pathogen Gene	Mutant Phenotype	Pathogen Species	Disease	Host Species
MedA	reduced virulence	Aspergillus fumigatus	Invasive Pulmonary Aspergillosis	Galleria mellonella (related: greater wax moth)

Pathogen Gene	Allele	Pathogen	Host
Gene:MedA PHI-base entry:PHI:2661 Gene ID:EAL93620 Protein ID: Q4X0J5 Sequence strain:Af293 Gene function:Developmental Regulator Essential gene:no		Pathogen species:Aspergillus fumigatus Pathogen ID: 746128 Pathogen strain:Af293 Pathogen strain ID: 330879	Host species:Galleria mellonella (related: greater wax moth) Host classification:Moths Host ID: 7137 Tissue:larva

Reference	Comments	PHI Phenotype	Pathogen Phenotype
Pmid:23185496 Ref source:Pubmed Year:2012 DOI: 10.1371/journal.pone.0049959 Author reference:Q A Abdallah 2012		Phenotype:reduced virulence Disease name:Invasive Pulmonary Aspergillosis Tissue:larva Experimental technique:Gene mutation: characterised; gene complementation	

- What species are mentioned as a host in these manually curated records?
- Is this an essential gene in *A. fumigatus*? Hint: Look in the Pathogen Gene section
- The **Genomic Location** section contains coordinates for a gene or sequence within a chromosome or contig/scaffold and a link out to GBrowse, which is centred on the gene of interest.
- The **Literature** section offers access to associated publications (internal curation of user comments, direct uploads from other resources - The Broad Institute, etc.)

- The **Orthology and synteny** section provides a table of Orthologs and Paralogs within FungiDB produced by OrthoMCL (www.orthomcl.org). This section also contains synteny graphs. The table has a search box allowing to create a custom display of orthologs and also deploy the ClustalOmega analysis:

7 Orthology and synteny

Ortholog Group OG5_139343

▼ Orthologs and Paralogs within EuPathDB data sets

Showing 9 of 125 rows ← Use search box to limit to a species of interest

Protein clustal Omega	Gene	Organism	Product	is syntenic	has comments
<input checked="" type="checkbox"/>	CIHG_03767	Coccidioides immitis H538.4	transcriptional regulator Medusa	yes	no
<input checked="" type="checkbox"/>	CIMG_05073	Coccidioides immitis RS	transcriptional regulator Medusa	yes	no
<input checked="" type="checkbox"/>	CPAG_06908	Coccidioides posadasii RMSCC 3488	medusa	yes	no
<input checked="" type="checkbox"/>	CPC735_068270	Coccidioides posadasii C735 delta SOWgp	hypothetical protein	yes	no
<input checked="" type="checkbox"/>	CPSG_08587	Coccidioides posadasii str. Silveira	transcription factor medA	yes	no
<input checked="" type="checkbox"/>	CPSG_08588	Coccidioides posadasii str. Silveira	transcriptional regulator Medusa	yes	no
<input checked="" type="checkbox"/>	PAAG_12199	Paracoccidioides lutzi Pb01	hypothetical protein	yes	no
<input checked="" type="checkbox"/>	PABG_05304	Paracoccidioides brasiliensis Pb03	hypothetical protein	yes	no
<input checked="" type="checkbox"/>	PADG_12215	Paracoccidioides brasiliensis Pb18	hypothetical protein	yes	no

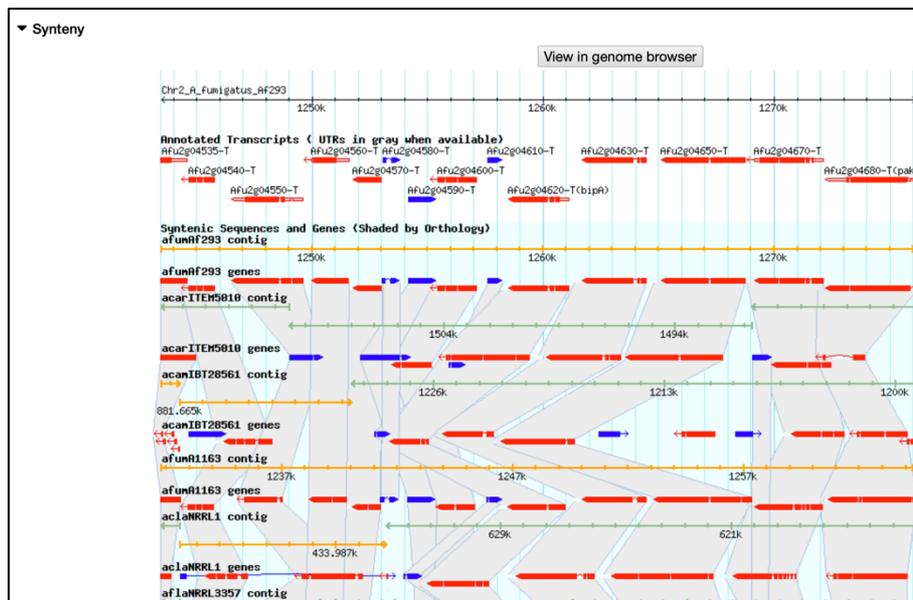
Run clustal Omega for selected genes Check All Uncheck All ← Run ClustalOmega to generate protein sequence alignments and a .dnd file for building phylogenetic trees

```

Clustal Omega 1.2.3 Multiple Sequence Alignments
PAAG_12199-t30_1-p1 1 ..... MAFPNHCAQ
PABG_05304-t30_1-p1 1 ..... MAFPNHCAQ
PADG_12215-t30_1-p1 1 ..... MAFPNHCAQ
CPSG_08587-t26_1-p1 1 ..... MAFPNHCAQ
CIMG_05073-t26_1-p1 1 ..... MAFPNHCAQ
CPC735_068270-t26_1-p1 1 ..... MAFPNHCAQ
PAAG_12199-t30_1-p1 11 LDLDFTQLI FDESDFSDP EDSEGFVGFAS DSKLAKSEGL LSMSAYEKP QVDL---LDY DATRAYHDA GFSTYAQPPY
PABG_05304-t30_1-p1 11 LDLDFTQLI FDESDFSDP EDSEGFVGFAS DSKLAKSEGL LSMSAYEKP QVDL---LDY DATRAYHDA GFSTYAQPPY
PADG_12215-t30_1-p1 11 LDLDFTQLI FDESDFSDP EDSEGFVGFAS DSKLAKSEGL LSMSAYEKP QVDL---LDY DATRAYHDA GFSTYAQPPY
AFu2g04550-T 1 ..... MSTYEQP PSRAELGKVF FSPFEEH--- --CSFGKGR QDRLRLDYP DSAKSLNGA SYSTYQGFY
CPSG_08587-t26_1-p1 67 EDLDFQSLI FDESDFSDP EESRFAGTFS DGLNAKTGGE NSMSTYEKQ PTDL---LDY DGTGALYQG ALPAYAQAPY
CPAG_06908-t26_1-p1 1 ..... MSTYEQP PTDL---LDY DGTGALYQG ALPAYAQAPY
CIHG_03767-t26_1-p1 1 ..... NSMSTYEKQ PTDL---LDY DGTGALYQG ALPAYAQAPY
CIMG_05073-t26_1-p1 72 FDLDFQSLI FDESDFSDP EDSEGFAGTFS DGLNAKTGGE NSMSTYEKQ PTDL---LDY DGTGALYQG ALPAYAQAPY
CPC735_068270-t26_1-p1 1 ..... NSMSTYEKQ PTDL---LDY DGTGALYQG ALPAYAQAPY
    
```

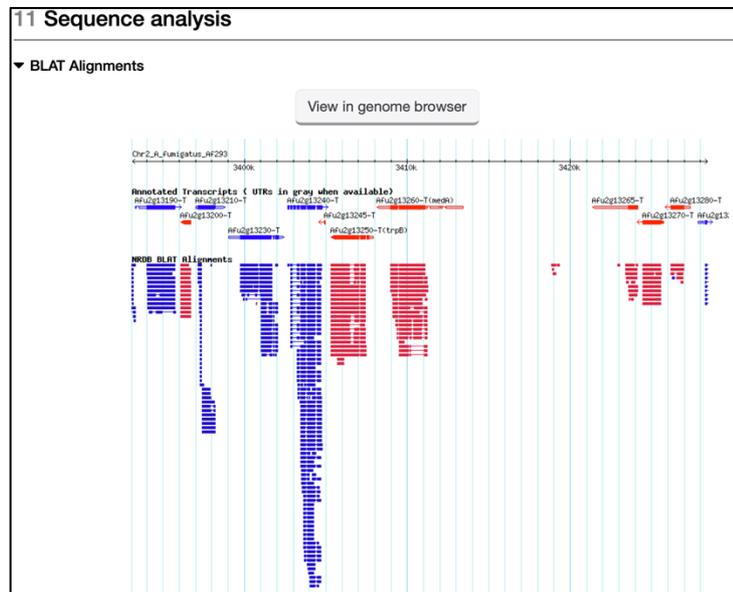
```

.dnd file
(
(
(PAAG_12199-t30_1-p1:0.0101626
,
(PABG_05304-t30_1-p1:0.00135501
,
(PADG_12215-t30_1-p1:0.00135501
):0.00880759
):0.271152
)
)
    
```



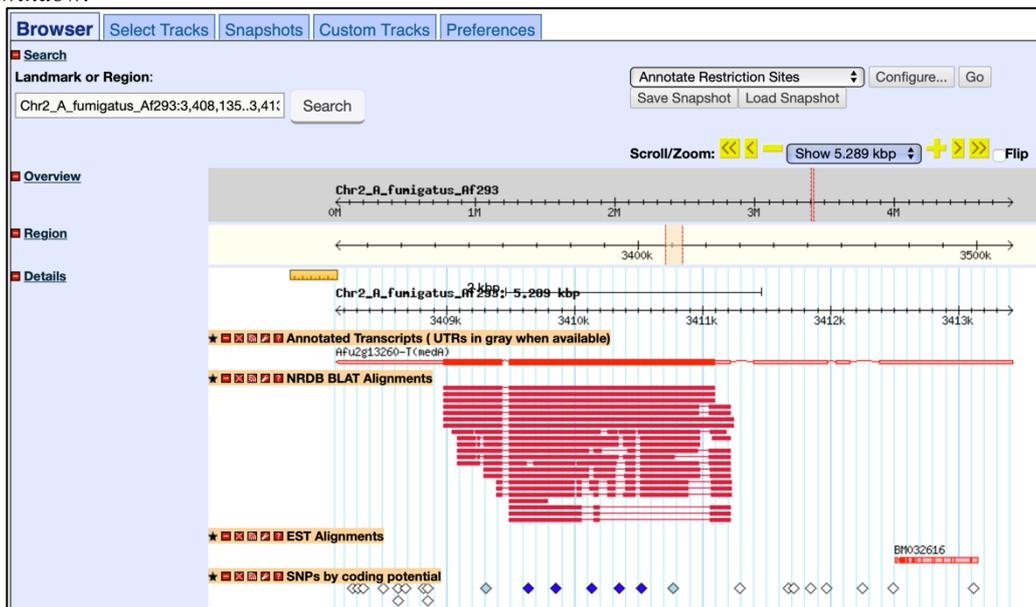
- Run ClustalOmega alignment on all sequences and copy .dnd file (without .dnd file line) to build a tree in iTOL (<https://itol.embl.de>) or other free web application you prefer.

- BLAT is a BLAST-like alignment tool. DNA BLAT database is an index of sequences derived from the assembly of the entire genome. DNA BLAT is designed to quickly find sequences of 95% and greater similarity of length 40 bases or more.



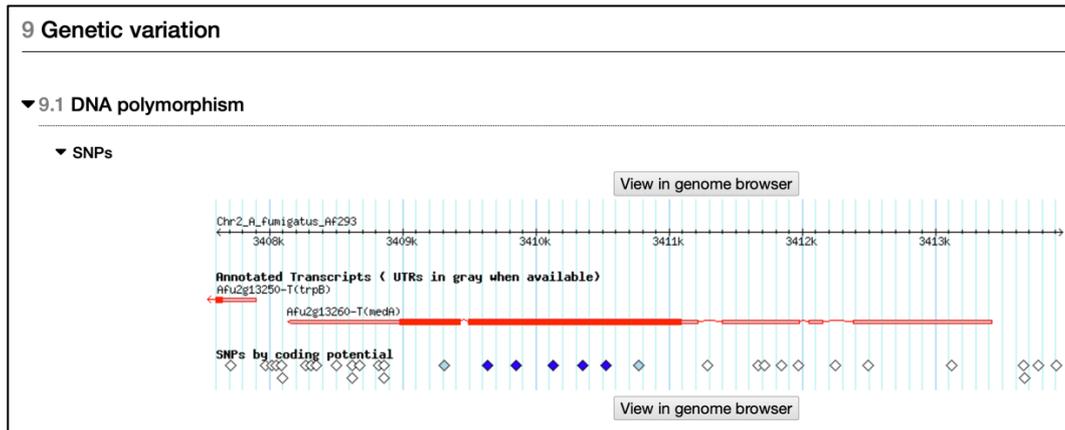
- Click on the **View in genome browser** button to view BLAT and also sequencing reads from *A. fumigatus* isolates.

Note: If you want to zoom in to viewing the B9J08_000928 only, enter the GeneID in the **Landmark or Region** search window.



Note: Rates of SNPs can reflect intra-species variation such as more SNPs can occur in less conserved regions while functional regions have fewer SNPs.

- The **Genetic Variation** section summarizes integrated SNP data for a given region and classifies SNPs into groups based on the resulting effect on gene function: noncoding (white diamonds), non-synonymous (dark blue) or synonymous (light blue), and nonsense (red) nucleotide changes.



Note: Genetic variation tracks can be explored in GBrowse and SNPs visualized by clicking on the View in genome browser button and activating appropriate tracks from the **Select Tracks** tab in GBrowse.

- Mouse over the SNP on the gene page to go directly to the SNP record.
- Can you determine the position of the first non-synonymous SNP on the left?

SNP: NGS_SNP.Chr2_A_fumigatus_Af293.3409636

Organism: Aspergillus fumigatus Af293
Location: Chr2_A_fumigatus_Af293: 3,409,636
Type: coding
Number of Strains: 48
Gene ID: Afu2g13260
Gene Strand: reverse
Major Allele: G (0.98)
Minor Allele: A (0.02)
Distinct Allele Count: 2
Reference Allele: G
Reference Product: S 485
Allele (gene strand): C
SNP context: AAGAAGTCTTCACTGTCTGCTTTGGCTTTGAGACGGTGAGGGGTCGGAAACCCCTCCAGAT
SNP context (gene strand): ATCTGGAGGGGTTCCGACCCCTCACCGTCTCAAAGCCAAAGCAGACAGTGAAGACTTCTT

- Determine the minor allele frequency and locate the matching isolate in the strains table
- Are there any geographical data associated with the minor allele isolate/strain?

▼ 2.1 DNA polymorphism

▼ Allele Summary [Download](#) [Data sets](#)

Search this table... Showing 2 rows

Allele	Allele Frequency	Allele Count	Average Coverage	Average Read Percent
G	0.979	47	77.8	100
A	0.021	1	90	98.8

▼ Country Summary [Download](#) [Data sets](#)

Search this table... Showing 5 rows

Geographic Location	#Alleles	Major Allele	Minor Allele	Other Allele
unknown	18	G (1)		
United Kingdom	13	G (.92)	A (.08)	
Kingdom of the Netherlands	8	G (1)		
India	8	G (1)		
United States of America	1	G (1)		

▼ Strains / Samples [Download](#) [Data sets](#)

Search this table... Showing 48 rows

Geographic Location	Strain	Sample	Allele	Allele (gene strand)	Product	Coverage	Read Frequency	DNA-seq reads for strain
	Af293 (reference)		G	C	S			
India	Afu_1042-09	EUSMPL0067-1-16	G	C	S	108	100	view DNA-seq reads
India	Afu_124-E11	EUSMPL0067-1-19	G	C	S	113	100	view DNA-seq reads
India	Afu_166-E11	EUSMPL0067-1-20	G	C	S	96	100	view DNA-seq reads
India	Afu_218-E11	EUSMPL0067-1-22	G	C	S	107	100	view DNA-seq reads
India	Afu_257-E11	EUSMPL0067-1-21	G	C	S	102	100	view DNA-seq reads
India	Afu_343-P-11	EUSMPL0067-1-17	G	C	S	78	100	view DNA-seq reads
India	Afu_591-12	EUSMPL0067-1-18	G	C	S	84	100	view DNA-seq reads
India	Afu_942-09	EUSMPL0067-1-15	G	C	S	128	100	view DNA-seq reads
Kingdom of the Netherlands	08-12-12-13	EUSMPL0067-1-7	G	C	S	94	100	view DNA-seq reads
Kingdom of the Netherlands	08-19-02-10	EUSMPL0067-1-14	G	C	S	118	100	view DNA-seq reads
Kingdom of the Netherlands	08-19-02-30	EUSMPL0067-1-11	G	C	S	123	100	view DNA-seq reads
Kingdom of the Netherlands	08-19-02-46	EUSMPL0067-1-13	G	C	S	126	100	view DNA-seq reads
Kingdom of the Netherlands	08-19-02-61	EUSMPL0067-1-10	G	C	S	130	100	view DNA-seq reads
Kingdom of the Netherlands	08-31-08-91	EUSMPL0067-1-9	G	C	S	94	100	view DNA-seq reads
Kingdom of the Netherlands	08-36-03-25	EUSMPL0067-1-8	G	C	S	138	100	view DNA-seq reads
Kingdom of the Netherlands	10-01-02-27	EUSMPL0067-1-12	G	C	S	117	100	view DNA-seq reads
United Kingdom	09-750806	EUSMPL0067-1-3	A	T	F	90	98.8	view DNA-seq reads
United Kingdom	12-7504462	EUSMPL0067-1-5	G	C	S	111	100	view DNA-seq reads

- The **Transcriptomics** section (RNA-Seq and microarray data) provides a searchable data table with expandable rows with tabular data, summaries, coverage, expression graphs, and more.

The Expression graph provides an overview of FPKM data, while the **Coverage** section shows uniquely mapped reads. When reads map to several genome locations and therefore could have been derived from multiple transcripts they are labeled as nonunique (shaded in grey).

Transcriptome of wild-type vs veA and mtfA deletion mutants
 RNA sequencing of *A. fumigatus* wild-type and deletion mutants of the secondary metabolite regulators, VeA and MtfA.
 Lind et al.
 RNA-seq

mtfA_fpkm - Afu2g13260

Full Dataset Description

Coverage

Non-unique mapping may be examined in the genome browser ([tutorial](#))

[View in genome browser](#)

Data table

Search this table... Showing 8 rows

Gene	Sample	Value	Standard Error	Percentile 1	Percentile 2
Afu2g13260	Delta-mtfA TTDS4.1 - mtfA - unique	12.72		34.98	
Afu2g13260	WT CEA10 mtfA matched - mtfA - unique	13.99		36.14	
Afu2g13260	Delta-mtfA TTDS4.1 - mtfA	12.58			
Afu2g13260	WT CEA10 mtfA matched - mtfA	13.84			
Afu2g13260	Delta-veA TSD1.15 - veA - unique	7.35			
Afu2g13260	WT CEA10 veA matched - veA - unique	12.57			
Afu2g13260	Delta-veA TSD1.15 - veA	7.26			
Afu2g13260	WT CEA10 veA matched - veA	12.43			

Description

X-axis
A.fumigatus wt and veA and mtfA mutants

Y-axis
 Transcript levels of fragments per kilobase of exon model per million mapped reads (FPKM). The percentile graph shows the ranking of expression for this gene compared to all others in this experiment.

Choose gene for which to display graph
 Afu2g13260

Choose graph(s) to display
 mtfA_fpkm veA_fpkm mtfA_percentile veA_percentile

Graph options
 Show log Scale (not applicable for log(ratio) graphs, percentile graphs or data tables)

Coverage plots for RNA-Seq datasets that can be viewed in greater details by clicking on the View in genome browser link

Detailed expression values data, including unique and non-unique alignment hits

Additional options for graphical representation of this transcriptomics dataset

- Take a look at the microarray data titled **Aspergillus fumigatus hypoxia response transcriptome**. Which time point corresponds to the highest expression during stress?
 - i. 2 hr after switch to 0.2% oxygen
 - ii. 6 hr after switch to 0.2% oxygen
 - iii. 12 hr after switch to 0.2% oxygen
 - iv. 24 hr after switch to 0.2% oxygen

- **Phenotypes** section include manually curated and publicly uploaded data for mutants (some but now all species).

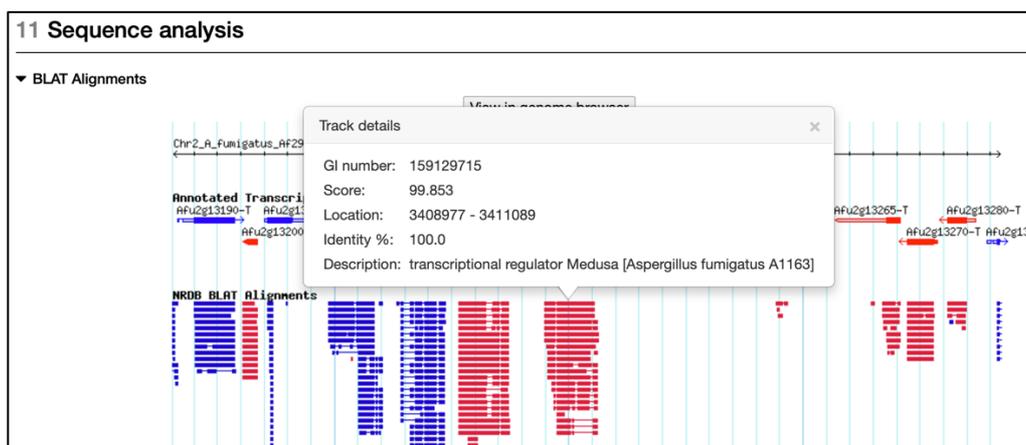
8 Phenotype

▼ Phenotype (qualities or directionality + entity or biological process) [Download](#) [Data sets](#)

Search this table... Showing 8 rows

Modification	Allele	Phenotype	Further Information	PubMed	CHEBI Annotation Extension
	medA::hygR	abnormal colony color	Details:delay in brown pigmentation	19889083	
	medA::hygR	decreased amount biological adhesion	Condition:fibronectin coated wells	19889083	
	medA::hygR	decreased amount conidium formation		19889083	
	medA::hygR	decreased amount virulence	Virulence model:immunosuppressed mouse pulmonary infection	19889083	
	medA::hygR	decreased amount virulence	Virulence model:insect infection (Galleria mellonella larvae)	19889083	
	medA::hygR	increased amount cell growth	Details:slightly larger conidia and conidiophores	19889083	
	medA::hygR	increased rate conidial germination		19889083	
	medA::hygR	viable cell		19889083	

- The **Sequences, Sequence analysis, Structure analysis** sections offer sequence information (DNA, RNA, and protein), an interactive summary of EST alignments and BLAT hits against the nonredundant protein sequence database (NRDB), 3D structure predictions and similar Protein Data Bank (PDB) chains, when available, *etc.*



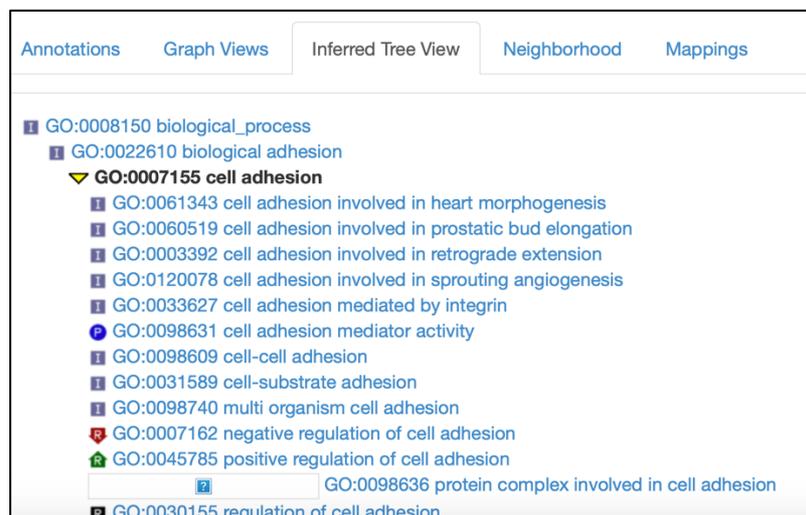
- **Protein features and properties** offer protein domains identifications (InterPro), signal and transmembrane predictions graphics, BLASTP hits, and also a selection of tools that can be deployed directly from the gene record page using the amino acid sequence of interest:
 - BLASTP (protein-protein BLAST) against 8 databases,
 - GPI anchor prediction: big-PI Predictor
 - MitoProt (prediction of mitochondrial proteins)
 - WoLF PSORT (prediction the subcellular localization sites)
- The **Function Prediction** section features Gene Ontology (GO) assignments that have been either downloaded from Gene Ontology databases and manually curated by FungiDB. GO Slim terms, Enzyme Commission (EC) numbers with links to EC number and GO terms description and relevant publications are available as well.

Transcript ID(s)	Ontology	GO ID	Is Not	GO Term Name	Source	Evidence Code	Reference
Afu2g13260-T	Biological Process	GO:0007155		cell adhesion	AspGD	IMP	AspGD_REF:ASPL0000082448 PMID:19889083
Afu2g13260-T	Biological Process	GO:0007155		cell adhesion	AspGD	IMP	AspGD_REF:ASPL0000402248 PMID:23185496
Afu2g13260-T	Biological Process	GO:0009405		pathogenesis	AspGD	IMP	AspGD_REF:ASPL0000082448 PMID:19889083
Afu2g13260-T	Biological Process	GO:0009405		pathogenesis	AspGD	IMP	AspGD_REF:ASPL0000402248 PMID:23185496
Afu2g13260-T	Biological Process	GO:0044406		adhesion of symbiont to host	AspGD	IMP	AspGD_REF:ASPL0000082448 PMID:19889083
Afu2g13260-T	Biological Process	GO:0048315		conidium formation	AspGD	IMP	AspGD_REF:ASPL0000402248 PMID:23185496
Afu2g13260-T	Biological Process	GO:0075307		positive regulation of conidium formation	AspGD	IEA	AspGD_REF:ASPL0000000005
Afu2g13260-T	Cellular Component	GO:0005634		nucleus	AspGD	IDA	AspGD_REF:ASPL0000402248 PMID:23185496
Afu2g13260-T	Molecular Function	GO:0003674		molecular_function	AspGD	ND	AspGD_REF:ASPL0000111607

- Based on the AspGD manual curation data described in this section how many unique GO terms are used to describe the **Biological process**?

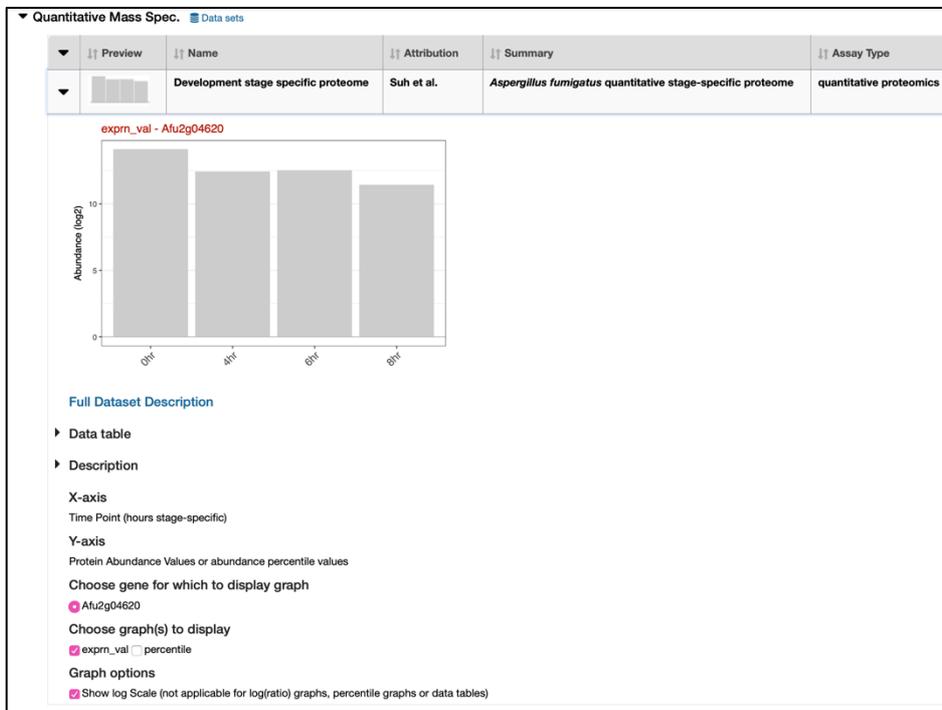
Note: Gene ontology provides statements for describing the functions of genes along three aspects: Molecular function, Biological process, and Cellular component, and it is organized in hierarchies. The GO terms table above provides GO IDs and terms associated with a particular gene and additional metadata that are available for these associations such as source, evidence code (e.g. IDA, IMP, etc.) and reference (PubMed ID). There are three classes of GO terms in FungiDB: GO terms that are automatically assigned by InterPro2GO, GO terms assigned by FungiDB curators, and GO terms obtained from external resources such as AspGD, MIPS, and others.

GO Slim is a subset of high-level terms from Gene Ontology that provides a broader overview of annotations. GO Slim table provides all annotated GO terms/IDs for a gene with the associated top-level GO Slim terms/IDs. If you click on the individual GO terms you will be redirected to AmiGO 2 site that contains detailed information for GO terms and their hierarchy. Here is an example of Inferred Tree of go terms that are “child” or component terms for cell adhesion (GO:0007155):



- **Pathways and interactions** provide descriptions of metabolic pathways loaded from the KEGG and MetaCyc repositories. Genes are linked to individual pathways by EC numbers when this data is available and clicking on any of the links redirects to an interactive metabolic pathway viewer where the user can explore individual reactions or export all known data for a given pathway.

- The **Proteomics** section is populated when Mass Spec. evidence data and phosphoproteomics datasets become available in FungiDB. This data is linked to peptide alignments against the reference genome in GBrowse and also provides a graphic summary of proteomic datasets. Here is an example for Afu2g04620:



- **Immunology** – primarily geared towards host response dataset and contains predicted epitopes from The Immune Epitope Database (IEDB).

FungiDB Genome Browser

1. Accessing Genome Browser - GBrowse

To access the Genome Browser features from the gene page click on the **View in genome browser** button:

Afu2g13260 Putative regulator of adherence, host cell interactions and virulence

Name: medA
Type: protein coding
Chromosome: 2
Location: Chr2_A_fumigatus_Af293:3,406,135..3,413,423(-)
Species: *Aspergillus fumigatus*
Strain: Af293
Status: Curated Reference Strain

Add the first user comment

Contents

- 1 Gene models
- 2 Annotation, curation and identifiers
- 3 Link outs
- 4 Genomic Location
- 5 Literature
- 6 Taxonomy
- 7 Orthology and synteny
- 8 Phenotype
- 9 Genetic variation
- 10 Transcriptomics
- 11 Sequence analysis
- 12 Sequences
- 13 Structure analysis
- 14 Protein features and properties
- 15 Function prediction
- 16 Pathways and interactions
- 17 Proteomics
- 18 Immunology

Shortcuts

- Synteny
- BLAT Alignments
- SNPs
- Transcriptomics
- Protein Features
- Proteomics

Also see Afu2g13260 in the Genome Browser or Protein Browser

1 Gene models

- # Exons in Gene 5
- # Transcripts 1
- ▼ Gene Models

View in genome browser

Chr2_A_fumigatus_Af293 3409k 3410k 3411k 3412k 3413k

Annotated Transcripts (UTRs in gray when available)

Afu2g13260-T1(trpB)

Afu2g13260-T1(orf0)

RNaseq evidence for introns (filtered)
Matches Transcript Annotation

Novel

RNaseq evidence for introns (inclusive)
Matches Transcript Annotation

Novel

View in genome browser

The Genome Browser default window offers quick access to (see figure on the next page):

1. File sharing and export options

Note: if you modified the default GBrowse window by loading custom tracks or other tracks from FungiDB and would like to share this window with a colleague, you must generate a custom URL address. Copy and paste of the default page url will display the default view only. To create a custom url, navigate to File, Generate URL.

2. Tabs

- The main **Browser** window, Select Tracks tab where you can activate custom tracks (RNA-Seq or microarray data, SNPs, etc.)
- **Snapshots** section provides a quick access to snapshots taken while accessing data in various GBrowse sessions
- **Custom tracks** - custom data files (bam or BigWig) uploaded directly from your computer or the EuPathDB Galaxy instance
- **Preferences** where you can update GBrowse appearance settings

- The four main features of the GBrowse window are **Landmark or Region** that provides information about this gene location and allows you to move to other regions by changing the location or entering a specific Gene ID. The **Overview** section displays the entire genomic sequence with a marker (red vertical lines) indication your current location. The **Region** displays the more immediate portion of the genome surrounding the region of interest. The **Details** section provides a closeup view of the region of interest. Scroll and zoom can be also done using a pull-down menu in the Scroll/Zoom section at the top of the Overview track.

Note: GBrowse can be also accessed from the Tools section on the homepage. If you do not know the exact location of the gene, you can just enter the GeneID in Landmark or Region window.

Tools

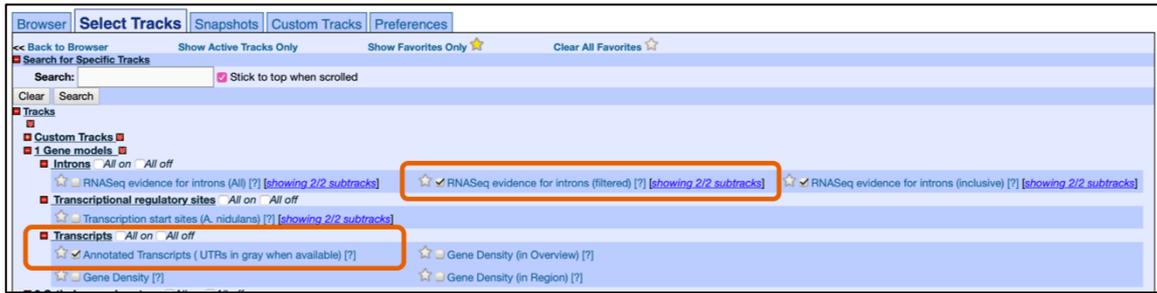
- BLAST**
Identify Sequence Similarities
- Results Analysis**
Analyze Your Strategy Results
- Sequence Retrieval**
Retrieve Specific Sequences using IDs and coordinates
- Companion**
Annotate your sequence and determine orthology, phylogeny & synteny
- EuPaGDT**
Eukaryotic Pathogen CRISPR guide RNA/DNA Design Tool
- PubMed and Entrez**
View the Latest Pubmed and Entrez Results
- Genome Browser**
View Sequences and Features in the genome browser
- Searches via Web Services**
Learn about web service access to our data

- Each track has 5 track-level controls that can be used to change the display of tracks or remove a track from the GBrowse window.

2. Understanding GBrowse display

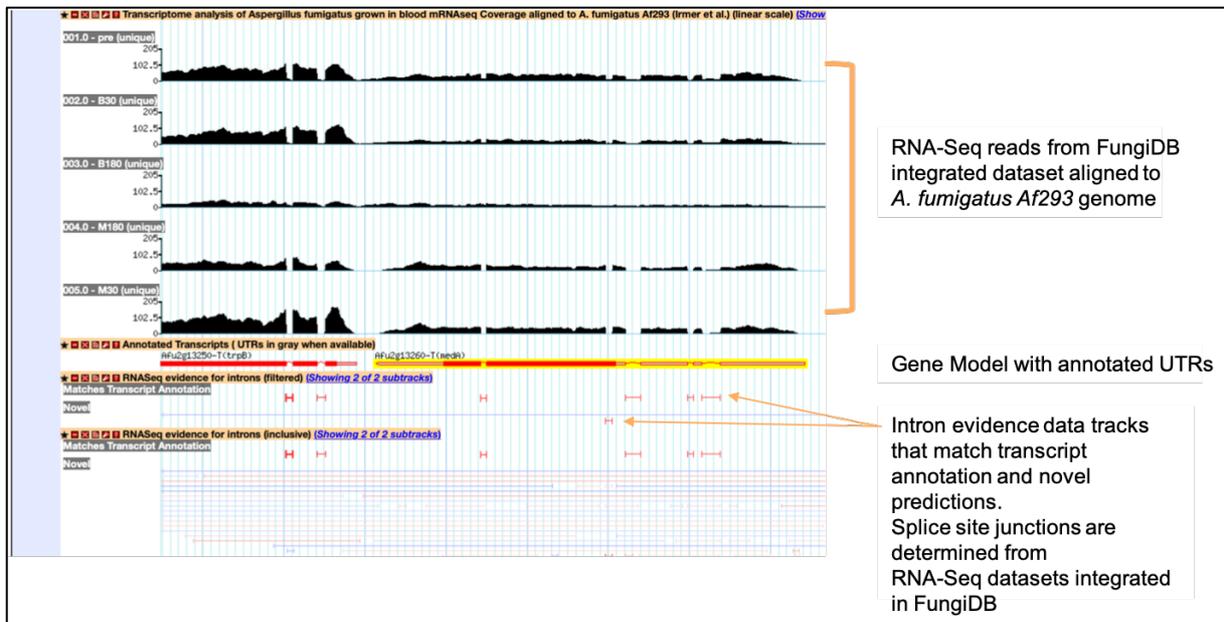
Navigate to the GBrowse view of the Afu2g13260.

To activate different tracks that are preloaded in FungiDB click on the **Select Tracks** tab and activate 1 RNA-Seq track in addition to default activated **RNASeq evidence for introns** and **Annotated Transcripts** tracks as shown in the figure below:



As soon as new tracks are selected they begin to load in the Browser tab. Go back to the Browser tab to view the newly activated tracks.

mRNA tracks can provide a comprehensive overview of changing transcriptome environment in response to various stimuli (light, sexual cycle, drug treatment, *etc.*). Visualization of RNA reads aligned against the reference genome may also shed light on transcript diversity, quality of genome annotation and provide evidence for differentially-spliced transcript isoforms and previously-unrecognized coding and non-coding RNAs.



All RNA-Seq datasets integrated in FungiDB are processed through a pipeline generating intron splice site predictions based on available RNA-Seq data. Several parameters are calculated to provide support for the intron tracks:

RNASeq Unified Splice Site Junctions (All)

[Click here to display in new window...](#)

Note that annotated introns are indicated with bold (wider) glyphs.

Intron Spanning Reads (ISR): The total number of uniquely mapped reads (all samples) which map across the junction and are on the appropriate strand. GSNAP uses splice site consensus sequences to determine strand of the mapped read.

ISR per million (ISRPM): Intron Spanning Reads Per Million intron spanning reads and thus represents a normalized count of unique reads.

% of Most Abundant Intron (MAI): The percentage (ISRPM of this junction / ISRPM of maximum junction for this gene) of this junction over the maximum for this gene.

Most abundant in: The experiment and sample that has the highest ISRPM for this gene.

ISRPM, (ISR / coverage): ISRPM from sample with highest ISRPM and the ISR/coverage for that same sample.

The table shows all experiments and samples that provide evidence for this intron junction. Note that the values for each row are based on each specific sample.

The color of the glyph changes with the Score as follows:

Reverse	Forward
less than 5	less than 5
5-15	5-15
17-64	17-64
65-256	65-256
257-1024	257-1024
1025-4096	1025-4096
4097-16000	4097-16000
greater than 16000	greater than 16000

Introns that are matching transcript annotation for which there is an abundance of supporting data from aligned RNA-Seq reads are displayed in bold colours (Blue for a forward gene shown here, and red for a gene located on the reverse strand). In the example shown here the large intron has 100% abundance in the integrated RNA datasets. A pop-up window can be initiated by hovering over the intron and will include all supporting documentation generated by in-house automated pipelines:

RNASeq evidence for introns (filtered) (Showing 2 of 2 subtracks)

Matches Transcript Annotation

Novel

RNASeq evidence for introns (inclusive) (Showing 2 of 2 subtracks)

Matches Transcript Annotation

Novel

Annotated Transcripts (UTRs in gray when available)

NCU02175-1.25.1(1nt-15)

Select Track

from File menu (above) and cut and paste resulting URL.

tings:
data was mapped.

EuPathDB

Pop-up Window:

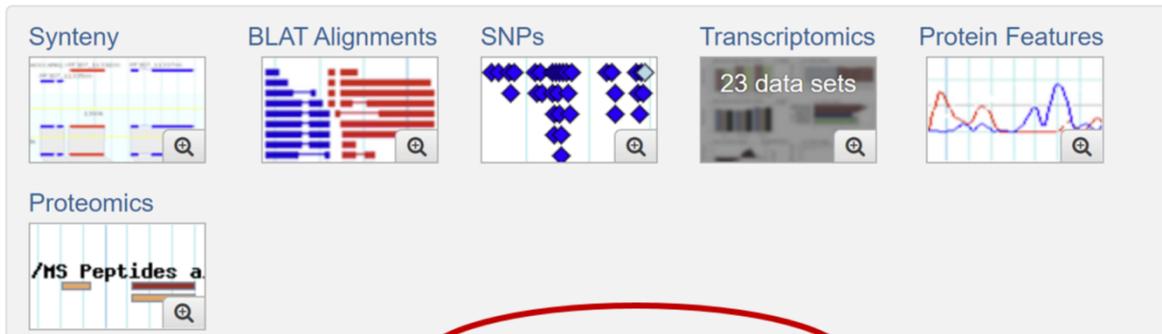
Intron Location: 986556 - 987162 (607 nt)
 Intron Spanning Reads (ISR): 8573
 ISR per million (ISRPM): 2744.88
 Gene assignment: NCU02175 - annotated intron
 % of Most Abundant Intron (MAI): 100
 Most abundant in: N. crassa transcriptomes during vegetative and sexual development stage: perithecia 5dpf stc-1
 ISRPM (ISR / gene coverage): 635.53 (.8985)

Experiment	Sample	Unique	ISRPM	ISR/Cov	% MAI
Characterization of light regulated genes	WT light 15 min	4	2.66	.4829	44.4
	WT light 60 min	9	4.06	.9948	50
	WT light 120 min	7	4.18	.5455	53.8
	WT light 240 min	7	5.74	.6561	53.8
	WT dark	12	7.74	.9134	92.3
Conserved and Essential Transcription Factors for Cellulase Gene Expression	cdr1Del Avicel 1h	50	12.95	.8081	100
	cdr1Del Avicel 4h	48	23.81	.5809	90.6
	cdr1Del Sucrose 1h	18	4.56	.7728	100
	cdr1Del Sucrose 4h	56	17.56	.7232	90.3
	cdr2Del Avicel 1h	23	7.39	.4555	57.5
	cdr2Del Avicel 4h	56	28.83	.6925	74.7
	cdr2Del Sucrose 1h	29	7.16	.991	82.9
	cdr2Del Sucrose 4h	44	15.32	.4309	61.1
	WT Avicel 1h	24	9.14	.5841	60
	WT Avicel 2h	9	12.65	.9832	81.8
	WT Avicel 30m	10	14.11	.6799	100
	WT Avicel 4h	39	21.45	.8155	88.6
	WT No Carbon 1h	26	9.67	1.0058	100
WT No Carbon 4h	127	24.45	.4336	60.8	

3. Accessing Genome Browser- JBrowse (BETA)

In addition to GBrowse, FungiDB has also implemented JBrowse that can be accessed directly from the gene records pages:

Shortcuts



Also see PF3D7_1133400 in the [JBrowse Genome Browser](#) **BETA** or Protein Browser

Click on *Select tracks* tab to view the list of available datasets. To return to the JBrowse main page click on *Back to browser* button. The pull-down tab next to individual track offer additional customization choices to create custom view of integrated data:

The screenshot displays the EuPathDB genome browser interface. At the top, the 'Select tracks' tab is active, showing a list of 45 matching tracks. The main view shows a genomic track with gene models for AN1088T, AN1089T, AN1090T, AN1237T, AN1238T, AN1239T, AN1240T, AN1241T, AN1242T, AN1243T, AN1244T, AN1245T, AN1246T, AN1247T, AN1248T, AN1249T, AN1250T, AN1251T, AN1252T, AN1253T, AN1254T, AN1255T, AN1256T, AN1257T, AN1258T, AN1259T, AN1260T, AN1261T, AN1262T, AN1263T, AN1264T, AN1265T, AN1266T, AN1267T, AN1268T, AN1269T, AN1270T, AN1271T, AN1272T, AN1273T, AN1274T, AN1275T, AN1276T, AN1277T, AN1278T, AN1279T, AN1280T, AN1281T, AN1282T, AN1283T, AN1284T, AN1285T, AN1286T, AN1287T, AN1288T, AN1289T, AN1290T, AN1291T, AN1292T, AN1293T, AN1294T, AN1295T, AN1296T, AN1297T, AN1298T, AN1299T, AN1300T.

The 'Select Tracks' panel is open, showing a table of tracks with columns for Name, Category, Subcategory, Dataset, Track, RNASeq Alignment, and RNASeq Strand. The table lists various tracks related to RNASeq and transcriptomics, including 'Delta-mtA TRVp - mtA (non-unique) Coverage', 'Delta-mtA TRVp - mtA (unique) Coverage', 'Delta-veA TXFp2.1 - veA (non-unique) Coverage', 'Delta-veA TXFp2.1 - veA (unique) Coverage', 'Inducer-Free Cellulase Secretion in Aspergillus nidulans Density', 'Inducer-Free Cellulase Secretion in Aspergillus nidulans Unique', 'Inducer-Free Cellulase Secretion in Aspergillus nidulans Intron Junctions', and 'An WT No Carbon (unique) Coverage'.

A context menu is open over the 'An WT No Carbon (unique) Coverage' track, showing options: Pin to top, Edit config, Delete track, Save track data, Change height, Log scale, and No fill.