

Entity Linking from GSA Journals to EuPathDB Sites

GENETICS and G3: Genes|Genomes|Genetics offer In-Article Database Linking to enrich your scientific publications and ensure that data relevant to your research can be discovered and accessed by readers. Entities such as genes that are mentioned in your article will be linked to their EuPathDB record pages if you provide GENETICS or G3 with the entity's data identifier (ID) in the following format – **database abbreviation: ID**.

For example, the article you are submitting to GENETICS mentions the *MIP family channel protein* from *Allomyces macrogynus* strain ATCC 38327. This entity is a gene which has Gene ID, AMAG_00042. The journal will link the gene ID in your paper to the FungiDB record page if you include the text **(FungiDB: AMAG_00042)** in your submission wherever you want the Gene ID to be linked.

Already know your gene's ID?

Confirm an ID by visiting the gene's FungiDB record page using the Gene ID box at the top of any FungiDB page (1). If you have a list of IDs, navigate the New Search dropdown menu to the Gene ID(s) search and click the ID in the result list after running the search. (2).

The screenshot shows the FungiDB search interface. At the top, there is a 'Gene ID:' search box containing 'AMAG_00042'. A callout box points to this box with the text: '1. Go directly to a gene page by entering the Gene ID in this box.' Below the search bar is a navigation menu with 'New Search' expanded to show 'Gene ID(s)'. A callout box points to this option with the text: '2. Navigate the dropdown menu to enter a list of gene IDs'. The search results page shows a table with columns: Gene ID, Organism, Genomic Location, and Product Description. The first row is highlighted, showing AMAG_00042 for A. macrogynus ATCC 38327. A callout box points to this row with the text: 'Identify Genes based on Gene ID(s)'. Below the table, there is a 'Gene Results' section with a 'Gene ID' column. A callout box points to the 'Gene ID' column with the text: 'FungiDB Gene Page'.

How do I find a Gene ID if I only know the gene name?

1. Enter the gene name surrounded by quotations in the Gene Text Search box at the top of any FungiDB page.
2. Use the filter table to reduce the list of returned genes to only those from your organism.
3. Choose the Gene ID from the result list.

The screenshot shows the FungiDB search interface. At the top, there is a 'Gene Text Search:' box containing '"MIP family channel"'. A callout box points to this box with the text: '1.'. Below the search bar is a navigation menu with 'New Search' expanded to show 'Gene ID(s)'. The search results page shows a table with columns: Gene ID, Organism, and Product Description. The first row is highlighted, showing AMAG_00042 for A. macrogynus ATCC 38327. A callout box points to this row with the text: 'Click ID to go to gene page.'. Below the table, there is a 'Gene Results' section with a 'Gene ID' column. A callout box points to the 'Gene ID' column with the text: '3.'. To the right of the search results, there is a filter table with columns: All Results, Ortholog Groups, Agaricomycetes, and Blastocladiomycetes. The filter table shows the following data:

All Results	Ortholog Groups	Agaricomycetes		Blastocladiomycetes
		Coprinopsis cinerea	Phanerochaete chrysosporium	Allomyces macrogynus
255	5	1	1	7

A callout box points to the '7' in the 'Allomyces macrogynus' column with the text: '2.'. Below the filter table, there is a 'Gene Results' section with a 'Gene ID' column. A callout box points to the 'Gene ID' column with the text: '3.'.

Using information from other genomics resources to find a FungiDB gene.

This example uses the gene, *MIP family channel protein* from *Coccidioides immitis* RS, which has Gene ID, CIMG_06551. The journal will link the gene ID in your paper to the FungiDB record page if you include the text (**FungiDB: CIMG_06551**) in your submission wherever you want the Gene ID to be linked.

1. The **locus tag** field often contains the FungiDB gene ID. Enter the locus tag in the FungiDB Gene ID box and the gene page will open automatically if the locus tag matches the FungiDB gene ID.
2. The **gene product name** can be used to locate a FungiDB record page. Enter the gene product name surrounded by quotation marks in the FungiDB Gene Text Search box. Then use the filter table to reduce the list of returned genes to only those from your organism. Choose a gene ID from the result list to go to the gene page.
3. **BLAST your sequence** against the FungiDB database and then explore the gene results to find your gene.

The screenshot shows the FungiDB homepage with the Gene ID field containing 'CIMG_06551' (labeled 1) and the Gene Text Search field containing '"MIP family channel"' (labeled 2). Below, the gene page for CIMG_06551 is shown, including a table of related genes in Eurotiomycetes:

Eurotiomycetes				
Illustrations	Aspergillus niger (nr Genes: 2)	Aspergillus oryzae	Aspergillus terreus	Coccidioides immitis (nr Genes: 2)
A4	ATCC 1015	CBS 513.88	RIB40	H538.4 RS
	2	1	2	1

Below the table, a 'Gene Results' section shows a list of genes with their product descriptions:

Gene ID	Product Description
CIMG_10104	MIP family channel protein
CIMG_06551	MIP family channel protein

The screenshot shows the NCBI GenBank record for the *Coccidioides immitis* RS MIP family channel protein partial mRNA. The record includes the following information:

- NCBI Reference Sequence: XM_001242654.2
- LOCUS: XM_001242654 1410 bp mRNA linear PLN 28-APR-20
- DEFINITION: *Coccidioides immitis* RS MIP family channel protein partial mRNA.
- ACCESSION: XM_001242654
- VERSION: XM_001242654.2 GI:815884906

The FEATURES section shows the following details:

- source: 1..1410 /organism="Coccidioides immitis RS" /mol_type="mRNA" /strain="RS"
- gene: 1..1410 /locus_tag="CIMG_06551" /db_xref="GeneID:4561008"
- CDS: 640..1410 /locus_tag="CIMG_06551" /codon_start=1 /product="MIP family channel protein" /protein_id="XP_001242655.2" /db_xref="GI:815884906" /db_xref="GeneID:4561008"

The translation is shown as: MWFSSGSDSPDAIRNHLIAF1GEFVGTFFMFLTFAFGTQVANEKTLGATSLARFGVAALLAQLMIAVLMIAVEKQFAPYHNYMAGPTGLSRAATGFFYLLKMKVETCNPGADSDGLERFDGDSKV

The ORIGIN section shows the nucleotide sequence with a boxed region (labeled 3) corresponding to the translation above.

The screenshot shows the FungiDB search interface. The 'Search for Genes' dropdown is expanded, and the 'BLAST' option is selected (labeled 3). The 'BLAST Program' section is visible, showing options for 'blastn', 'blastp', 'blastx', 'tblastn', and 'tblastx'. The 'Target Organism' is set to 'Coccidioides immitis'. The 'Input Sequence' field contains the sequence: MIPFSSGSDSPDAIRNHLIAF1GEFVGTFFMFLTFAFGTQVANEKTLGATSLARFGVAALLAQLMIAVLMIAVEKQFAPYHNYMAGPTGLSRAATGFFYLLKMKVETCNPGADSDGLERFDGDSKV. The 'Expectation value' is set to 10, and the 'Maximum description length' is set to 50.

The screenshot shows the 'My Strategies' page in FungiDB. A strategy named 'BLAST(2)' is selected. The results show 50 genes from Step 1. A table of results is displayed:

Gene ID	Organism
CIMG_00947	C. immitis H538.4
CIMG_06551	C. immitis RS
CPAG_03917	C. posadasii RMSCC 3488
CPCT735_03020	C. posadasii C735 delta SOW1gp
CPSG_07101	C. posadasii str. Silveira
PMAA_006710	T. marneffei ATCC-18224

A callout box (labeled 3) points to the 'Gene ID' column with the text: 'Click ID to go to gene page.'