

FungiDB: Secondary Metabolites and clusters

Learning objectives:

- Explore InterPro search in FungiDB
- Cross-reference the results with MycoCosm data

1. Finding secondary metabolites and gene clusters

There are several main classes of secondary metabolites in fungi: the non-ribosomal peptides and amino acid-derived compounds, polyketides and fatty acid-derived compounds, and terpenes (Hoffmeister and Keller, 2007). The secondary metabolites can be segregated into groups based on the first step of their biosynthesis, more specifically, the “key enzymes” that are required: Non-ribosomal peptide synthetases (NRPSs), NRPS-like, Polyketide synthases (PKSs), PKS-like, Hybrid PKS – NRPS, Prenyltransferases (DMAT), Terpene cyclases/synthase (TC) (Read more: <https://www.ncbi.nlm.nih.gov/pubmed/24692239>)

- InterPro search to identify NRPS genes in all *Aspergilli*. NRPS genes have at least the three domains: AMP-binding (PF00501), PP-binding (PF00550), and Condensation (PF00668).

Search for...

InterP

Genes

Protein features and properties

[InterPro Domain](#)

NRPS: Genes that have at least the three domains AMP-binding (PF00501), PP-binding (PF00550), and Condensation (PF00668). *

Step 1: InterPro Dom PF00501 (1,910 Genes)

Step 2: InterPro Dom PF00550 (1,561 Genes)

Step 3: InterPro Dom PF00668 (618 Genes)

926 Genes

557 Genes

+ Add a step

- How many genes were identified in *Aspergillus fumigatus* Af293?

Hint: use the organism filter on the left to limit your search results to Af293 genes only.

15 Genes (3 ortholog groups)

Gene Results | Genome View | Analyze Results

Rows per page: 1000

Download | Add to Basket | Add Columns

| Gene ID | Transcript ID | Genomic Location (Gene) | Product Description | # Transcripts |
|-------------|---------------|--|---|---------------|
| Afu4g11240 | Afu4g11240-T | Chr4_A_fumigatus_Af293:2,938,272..2,942,727(+) | Putative alpha-aminoadipate reductase, large subunit, involved in lysine metabolism | 1 |
| Afu8R00054N | Afu8R00054N-T | Chr8_A_fumigatus_Af293:116,803..120,273(+) | PKS-NRPS hybrid synthase nonA | 1 |

- Create a search for NRPS genes in MycoCosm. Access the *A. fumigatus* Af293 portal (<https://mycocosm.jgi.doe.gov/Aspfu1>) and navigate to the secondary metabolism clusters page (under the ‘annotations’ tab). How many genes did you get?

SEARCH

BLAST

BROWSE

ANNOTATIONS ▾

MCL CLUSTERS

SYNTENY

DOWNLOAD

INFO

HOME

HELP!

Genomes

Cluster Type

Scale

Clusters Per Page

Aspergillus fumigatus Af293 from AspGD

all

DMAT

HYBRID

NRPS

NRPS-Like

Per Cluster ▾

50 ▾

Refresh

| Total 9 cluster(s) found. 1 | | | | |
|-----------------------------|--------------|---|-----------|-------|
| Cluster Id | Cluster Type | Scaffold | Size (bp) | Genes |
| Aspfu1.5 | NRPS | Chr_3_A_fumigatus_Af293:876157-937897 | 61,740 | |
| Aspfu1.7 | NRPS | Chr_3_A_fumigatus_Af293:3423866-3446129 | 22,263 | |
| Aspfu1.10 | NRPS | Chr_3_A_fumigatus_Af293:4007787-4023468 | 15,681 | |
| Aspfu1.15 | NRPS | Chr_1_A_fumigatus_Af293:2655644-2694887 | 39,243 | |
| Aspfu1.16 | NRPS | Chr_1_A_fumigatus_Af293:4662924-4713331 | 50,407 | |
| Aspfu1.18 | NRPS | Chr_8_A_fumigatus_Af293:20854-49410 | 28,556 | |
| Aspfu1.28 | NRPS | Chr_5_A_fumigatus_Af293:3307809-3342792 | 34,983 | |
| Aspfu1.31 | NRPS | Chr_6_A_fumigatus_Af293:2334637-2372302 | 37,665 | |
| Aspfu1.32 | NRPS | Chr_6_A_fumigatus_Af293:3004871-3035305 | 30,434 | |
| Cluster Id | Cluster Type | Scaffold | Size (bp) | Genes |

- What do you think may be causing the difference in the predicted gene number?
- This view on MycoCosm allows you to analyze backbone and auxiliary proteins across the entire predicted secondary metabolism cluster. How conserved are these secondary metabolite clusters across related *Aspergilli*? Click on the scaffold coordinates for AspFu1.5 and analyze the Vista curve tracks. How many related *Aspergilli* show some synteny with this region? Repeat this exercise for the next cluster, AspFu1.7.
 - Answer: Synteny is observed across most *Aspergilli* for AspFu1.5, raising the possibility that this SM cluster is widespread across the genus. However, AspFu1.7 shows no synteny except for at a couple auxiliary genes in *Aspergillus wentii*, suggesting that it is possibly lineage specific.

Aspfu1.5:





The screenshot displays the NCBI SRA Run Viewer interface for the SRR1991661.1 dataset. The top navigation bar includes links for Search, BLAST, RefSeq, Annotations, NCBI Clusters, Download, Help, Home, and Help. The main content area shows a detailed view of the sequence alignment, with a large block of mismatches highlighted in red at the bottom. The alignment tracks include Base Position (1-100,000,000), Read ID (SRR1991661.1), and various alignment metrics (CIGAR, MAPQ, etc.). The alignment shows a high degree of similarity between the read and the reference genome, with a large block of mismatches highlighted in red at the bottom.

