

FungiDB: Secondary Metabolites and clusters

Learning objectives:

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

• Finding secondary metabolites and gene clusters

Fungi produce a plethora of secondary metabolites. 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).

1. Use the InterPro search to identify NRPS genes in all *Aspergilli*.

NRPS genes have at least the three domains:

- AMP-binding (PF00501)
- PP-binding (PF00550)
- Condensation (PF00668)

Search for...

Genes
Pathways and interactions
 Metabolic Pathway
 Substrates/Products
 Y2h Protein Interactions
Protein features and properties
 InterPro Domain

Find genes containing a specified protein domain from the InterPro database. The InterPro family of databases includes CATH, CCD, HAMAP, PANTHER, Pfam, PIRSF, PRINTS, PRODOM, PROFILE, PROSITE, SFLD, SMART, SUPERFAMILY, TIGRFAMs.

Identify Genes based on InterPro Domain

☒ *Aspergillus*
 ☒ *Aspergillus aculeatus* ATCC 16872 [Reference]
 ☒ *Aspergillus brasiliensis* CBS 101740 [Reference]
 ☒ *Aspergillus campestris* IBT 28561 [Reference]
 ☒ *Aspergillus carbonarius* ITEM 5010 [Reference]
 ☒ *Aspergillus clavatus* NRRL 1 [Reference]
 ☒ *Aspergillus cristatus* GZAAS20.1005 [Reference]
 ☒ *Aspergillus eucalypticola* CBS 122712 [Reference]
 ☒ *Aspergillus fijiensis* CBS 313.89 [Reference]
 ☒ *Aspergillus fischeri* NRRL 181 [Reference]
 ☒ *Aspergillus flavus*
 ☒ *Aspergillus flavus* NRRL3357 [Reference]
 ☒ *Aspergillus flavus* NRRL3357 2020
 ☒ *Aspergillus fumigatus*
 ☒ *Aspergillus fumigatus* A1163
 ☒ *Aspergillus fumigatus* AT292 [Reference]
 ☒ *Aspergillus glaucus* CBS 516.65 [Reference]
 ☒ *Aspergillus heteromorphus* CBS 117.55 [Reference]
 ☒ *Aspergillus lentulus* strain IFM 54703 [Reference]
 ☒ *Aspergillus luchuensis*
 ☒ *Aspergillus luchuensis* CBS 106.47 [Reference]
 ☒ *Aspergillus luchuensis* IFO 4308
 ☒ *Aspergillus nidulans* FGSC A4 [Reference]
 ☒ *Aspergillus niger*
 ☒ *Aspergillus niger* ATCC 1015
 ☒ *Aspergillus niger* ATCC 13496
 ☒ *Aspergillus niger* CBS 513.88 [Reference]
 ☒ *Aspergillus niger* strain N402 (ATCC64974)
 ☒ *Aspergillus novofumigatus* IBT 16806 [Reference]
 ☒ *Aspergillus ochraceoroseus* IBT 24754 [Reference]
 ☒ *Aspergillus oryzae* RIB40 [Reference]
 ☒ *Aspergillus parasiticus* CBS 117618 [Reference]
 ☒ *Aspergillus steynii* IBT 23096 [Reference]
 ☒ *Aspergillus sydowii* CBS 592.65 [Reference]
 ☒ *Aspergillus tanneri* NIH1004 [Reference]
 ☒ *Aspergillus terreus* NIH2624 [Reference]
 ☒ *Aspergillus thermomutatus* strain HMR AF 39 [Reference]
 ☒ *Aspergillus tubingensis* CBS 134.48 [Reference]
 ☒ *Aspergillus uvarum* CBS 121591 [Reference]
 ☒ *Aspergillus versicolor* CBS 583.65 [Reference]
 ☒ *Aspergillus wentii* DTO 134E9 [Reference]

Domain Database

PFAM

Specific Domain(s)
PF00501 : AMP-binding AMP-dependent synthetase/ligase

PF00550 : PP-binding Phosphopantetheine binding ACP domain

PF00668 : Condensation Condensation domain

NRPS
Step 1: PF00501 (2,639 Genes) → Step 2: 1,273 Genes → Step 3: PF00668 (766 Genes) → Add a step

Strategy URL:

<https://fungidb.org/fungidb/app/workspace/strategies/import/85a1e3a5a603efc6>

- 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.

The image shows two parts of the MycoCosm interface. On the left is the 'Organism Filter' panel. It has a search bar with 'af293' entered. Below it is a tree view of taxonomic levels: Fungi (766), Ascomycota (766), Eurotiomycetes (766), Eurotiales (766), Aspergillaceae (766), Aspergillus (766), Aspergillus fumigatus (31), and Aspergillus fumigatus Af293 (15). The 'Af293' entry is selected. On the right is a diagram of an NRPS cluster. It shows three steps: Step 1 (PF00501, 2,639 Genes), Step 2 (PF00550, 2,164 Genes), and Step 3 (PF00668, 858 Genes). The genes are represented by colored circles and arrows. An orange arrow points from the organism filter to the NRPS diagram.

- 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?

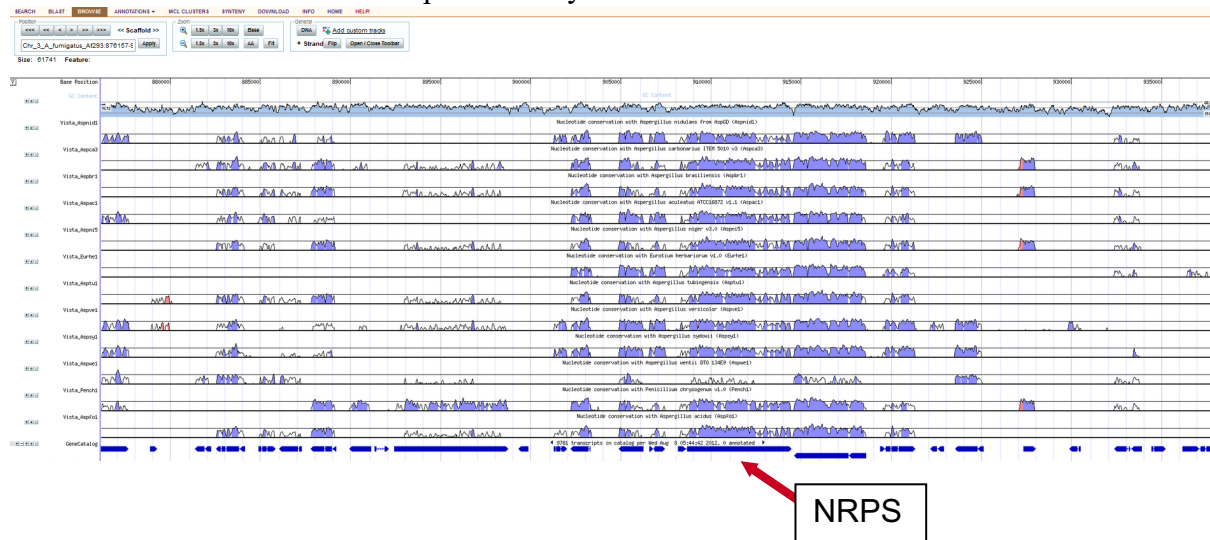
The image shows the 'Secondary Metabolism Clusters' page in MycoCosm. The 'Annotations' tab is selected. The 'Genomes' dropdown is set to 'Aspergillus fumigatus Af293 from AspGD'. The 'Cluster Type' dropdown is set to 'NRPS'. The 'Scale' dropdown is set to 'Per Cluster'. The 'Clusters Per Page' dropdown is set to '50'. The 'Refresh' button is visible. Below the filters, a table shows the results of the search.

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

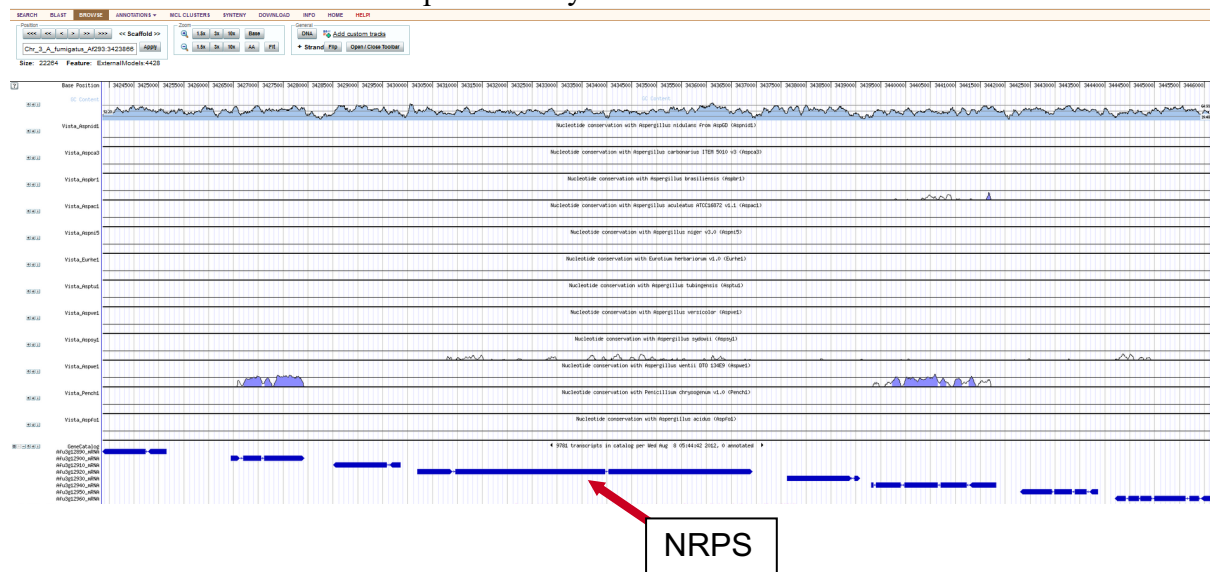
- What do you think may be causing the difference in the predicted gene number?
- This view on MycoCosm allows you to analyze backbone and auxiliary proteins across the entire predicted secondary metabolism cluster. How conserved are these secondary metabolite clusters across related Aspergilli? Click on the scaffold coordinates for Aspfu1.5 and analyze the Vista curve tracks in the genome browser. 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,

Aspf1.7 shows no synteny except for at a couple auxiliary genes in *Aspergillus wentii*, suggesting that it is possibly lineage specific.

Genome browser at locus for Aspf1.5 biosynthetic cluster:



Genome browser at locus for Aspf1.7 biosynthetic cluster:



Reference: PMID:24692239