

MycoCosm Secondary Metabolism Clusters Browser

In fungi, Secondary metabolite (SM) genes are often organized in chromosomal clusters dedicated to that metabolite's biosynthetic pathway. Each portal's SM Clusters Browser facilitates display and discovery of MycoCosm's SM-annotated genes.

Scenario: You have identified a toxic SM produced by *Septoria musiva*, a pathogenic fungus that induces cankers in the poplar tree, but not produced by *Septoria populicola*, which infects a different species of poplar and does not induce cankers. The SM's structure suggests that its biosynthetic pathway may have as its core enzyme a hybrid PKS-NRPS (polyketide synthase-nonribosomal peptide synthetase). Use MycoCosm to find candidate gene clusters for this pathway.

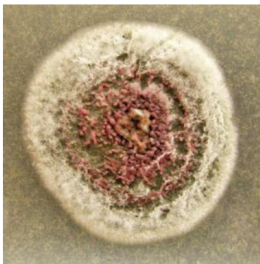
- 1) Go to the MycoCosm Septoria PhyloGroup at genome.jgi.doe.gov/Septoria. Both species are represented in the group:

Info • Septoria					
SEARCH	BLAST	ANNOTATIONS	MCL CLUSTERS	DOWNLOAD	INFO
HELP!					
Group Name: Septoria					
##	Name	Assembly Length	# Genes	Published	
1	Septoria musiva SO2202 v1.0	29,352,103	10,233	Ohm RA et al., 2012	
2	Septoria populicola v1.0	33,188,813	9,739	Ohm RA et al., 2012	

- 2) Click on '*Septoria musiva*' to go to its genome portal:

Home • Septoria musiva SO2202 v1.0

SEARCH BLAST BROWSE ANNOTATIONS MCL CLUSTERS SYNTENY DOWNLOAD INFO HOME STATUS HELP!



Septoria musiva (sexual stage: *Mycosphaerella populorum*) causes leaf spots and cankers on poplars (*Populus* spp. and hybrids). On native North American poplars the pathogen mainly causes leaf spots that can lead to defoliation but generally do not kill the host. But *S. musiva* can also cause cankers on branches and primary stems. These can be lethal and are particularly severe on hybrid poplars in plantations. They often develop on the primary shoots of 2- to 3-year-old trees, leading to restrictions in the movement of water and nutrients and weakening the wood within a few feet of ground level. The weakened trunks collapse easily, greatly reducing the production of biomass. Cankers caused by *S. musiva* can greatly hamper the production of hybrid poplars in the eastern United States and Canada and threaten poplars in western North America.

A major concern with *S. musiva* is with migration to new areas. The pathogen is endemic and appears to have originated on poplars in eastern North America, where it occurs commonly on leaves of the eastern cottonwood, *P. deltoides*. During the past 20 years *S. musiva* has appeared in South America and western Canada, where it is spreading rapidly on native and hybrid poplars causing economic damage as well as threatening native poplars in important riparian zones. It is not yet known in Europe or Asia but has the potential to cause extensive damage if introduced to those areas. Global warming and trade may facilitate the spread of the disease by making northern poplar-growing areas more favorable to growth of the fungus.

Availability of a genome sequence for *S. musiva* will help with designing strategies to effectively manage this destructive disease. The genome of the black cottonwood poplar *P. trichocarpa*, another host for *S. musiva*, has been sequenced and provides a rare opportunity to analyze host-pathogen interactions when both host and pathogen have been sequenced. Opportunities for comparative genomics also are available. Other members of the genus *Mycosphaerella* with sequenced genomes include the pine pathogen *Dothistroma septosporum* (aka *Mycosphaerella pini*), the banana pathogen *M. fijiensis* and the wheat pathogen *M. graminicola*. Comparative genomics of these sequences with that of *S. musiva* will help identify the differences involved in pathogenicity to woody vs. herbaceous hosts and could help in our understanding of why *S. musiva* causes cankers on hybrids but not native poplars. Understanding the mechanisms of fungal pathogenicity and of resistance in the host will be essential for protecting our forests and ensuring a stable supply of renewable biomass as the climate warms in the future.

Photo credit: [Glen Stanosz, Ph.D., University of Wisconsin-Madison](#)

- 3) Click on "ANNOTATIONS => SECONDARY METABOLISM CLUSTERS" to go to the portal's SM clusters browser:

Secondary Metabolism Clusters • Septoria musiva SO2202 v1.0						
SEARCH	BLAST	BROWSE	ANNOTATIONS	MCL CLUSTERS	SYNTENY	DOWNLOAD
INFO HOME STATUS HELP!						
Genomes			GENE ONTOLOGY	Refresh		
Alternaria brassicicola			PFAM DOMAINS			
Baudoinia compniacensis UAMH 1			KEGG			
Cochliobolus heterostrophus C5			KOG			
Dothistroma septosporum NZE10			SECONDARY METABOLISM CLUSTERS			
Hysterium pulicre			CAZYMES			
Genome			PEPTIDASES			
Alternaria brassicicola			TRANSPORTERS	NRPS	NRPS-Like	PKS
Baudoinia compniacensis UAMH 1			TRANSCRIPTION FACTORS	4	6	6
				2	6	2
						3
						5
						24
						13

- 4) Scroll through the ‘Genomes’ list box and select both ‘*Septoria musiva*’ and ‘*Septoria populicola*’, and only those 2 species. Click the ‘Refresh’ button. Only the SM cluster core gene counts of the 2 *Septoria* sp. are shown, and may be directly compared. *S. musiva* has 2 hybrid core genes (PKS-NRPS genes) while *S. populicola* has none:

Secondary Metabolism Clusters • *Septoria musiva* SO2202 v1.0

SEARCH BLAST BROWSE ANNOTATIONS MCL CLUSTERS SYNTENY DOWNLOAD INFO HOME STATUS HELP!

Genomes: *Septoria musiva* SO2202 v1.0, *Septoria populicola* v1.0

Cluster Type: all, DMAT, HYBRID, NRPS, NRPS-Like

Genome	DMAT	HYBRID	NRPS	NRPS-Like	PKS	PKS-Like	TC	Total
<i>Septoria musiva</i> SO2202 v1.0	0	2	7	8	9	2	2	30
<i>Septoria populicola</i> v1.0	0	0	8	7	9	2	3	29
Total	0	2	15	15	18	4	5	59

- 5) There is a total of 2 genes in the Hybrid column. Click on the number to show a graphical representation of the 2 *S. musiva* gene clusters. The ‘Size’ column displays each cluster’s length, and the ‘Genes’ column displays each cluster’s core PKS-NRPS gene (in color) and its accessory, decorator, and other genes (in gray). A core hybrid gene is typically very large, but the total cluster size can be highly variable. To resize the 2 clusters to scale to each other, go to the ‘Scale’ pull-down menu, select ‘Across All Clusters’, and click on the ‘Refresh’ button:

Secondary Metabolism Clusters • *Septoria musiva* SO2202 v1.0

SEARCH BLAST BROWSE ANNOTATIONS MCL CLUSTERS SYNTENY DOWNLOAD INFO HOME STATUS HELP!

Genomes: *Septoria musiva* SO2202 v1.0

Cluster Type: all, DMAT, HYBRID, NRPS, NRPS-Like

Scale: Per Cluster, **Per Cluster**, Per Cluster No Gaps, Across All Clusters

Clusters Per Page: 50

Total 2 cluster(s) found. 1

Cluster Id	Cluster Type	Scaffold	Size (bp)	Genes
Sepmu1.24	HYBRID	scaffold_6.1522811-1553990	31,179	
Sepmu1.25	HYBRID	scaffold_6.1977373-2004431	27,058	

- 6) Each gene in the clusters is represented by an arrow with a single pair of fletching that indicates the gene’s 5’ → 3’ direction. Mouse-over the top cluster’s core gene to get more information about the PKS-NRPS hybrid. The listed domains are typical of a hybrid enzyme:

Secondary Metabolism Clusters • *Septoria musiva* SO2202 v1.0

SEARCH BLAST BROWSE ANNOTATIONS MCL CLUSTERS SYNTENY DOWNLOAD INFO HOME STATUS HELP!

Genomes: *Septoria musiva* SO2202 v1.0

Cluster Type: all, DMAT, HYBRID, NRPS, NRPS-Like

Scale: Across All Clusters

Clusters Per Page: 50

Total 2 cluster(s) found. 1

Cluster Id	Cluster Type	Scaffold	Size (bp)	Genes
Sepmu1.24	HYBRID	scaffold_6.1522811-1553990	31,179	
Sepmu1.25	HYBRID	scaffold_6.1977373-2004431	27,058	

Type: HYBRID
ProteinId: 134219
Domains: PF00109 ketoacyl-synt 1, PF00501 AMP-binding 1, PF00550 PP-binding 1, PF00668 Condensation 1, PF00698 Acyl_transf_1 1, PF02801 Ketoacyl-synt_C 1, PF07993 NAD_binding_4 1, PF08242 Methyltransf_12 1, PF08659 KR 1

- 7) To get domain information about the other genes in the SM cluster, mouse-over them too. The next gene 3' to the core gene has a p450 domain:
- 8) To get more detailed information about a gene, click on it directly. Click on the gene with the p450 domain to see its 'protein page'. Examination of the protein page reveals that:
 - a) The gene is expressed. The blue bars represent UTRs, which can be inferred only from transcriptomic data.
 - b) The protein has p450 Pfam and other annotations indicative of a cytochrome p450 monooxygenase.
 - c) The best Blast hit in nr is a cytochrome p450 monooxygenase from *Aspergillus nidulans*, which belongs to a different class of fungi (Eurotiomycetes) from *Septoria* (Dothideomycetes).

Description: [gij67902848\[ref|XP_681680.1| hypothetical protein AN8411.2 \[Aspergillus nidulans FGSC A4\] >gij40747877\[jgi|EAA67033.1| hypothetical protein AN8411.2 \[Aspergillus nidulans FGSC A4\] >gij259484346\[tpe|CBF80485.1| TPA: Cytochrome P450 monooxygenase \(Eurofung\).\[Aspergillus nidulans FGSC A4\]\(model%: 91, hit%: 90, score: 1905, %id: 71\) \[Aspergillus nidulans FGSC A4\]](#)

Best Hit: [AN8411.2 \[Aspergillus nidulans FGSC A4\] >gij259484346\[tpe|CBF80485.1| TPA: Cytochrome P450 monooxygenase \(Eurofung\).\[Aspergillus nidulans FGSC A4\]\(model%: 91, hit%: 90, score: 1905, %id: 71\) \[Aspergillus nidulans FGSC A4\]](#)

total hits(shown) 683 (10)

ASPECT	GO Id	GO Desc	Interpro Id	Interpro Desc
Molecular Function	0016712	oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen, reduced flavin or flavoprotein as one donor, and incorporation of one atom of oxygen	IPR002974	Cytochrome P450, E-class, CYP52
	0004497	monooxygenase activity	IPR002402	Cytochrome P450, E-class, group II
	0020037	heme binding	IPR001128	Cytochrome P450
	0005506	iron ion binding	IPR002402	Cytochrome P450, E-class, group II
	0005506	iron ion binding	IPR002974	Cytochrome P450, E-class, CYP52
Biological Process	0006118	electron transport	IPR001128	Cytochrome P450
KOG GROUP	KOG Id	KOG Class	KOG Desc	
Metabolism	KOG0158	Secondary metabolites biosynthesis, transport and catabolism	Cytochrome P450 CYP3/CYP5/CYP6/CYP9 subfamilies	

[View/modify manual annotation](#)
[View nucleotide and 3-frame translation](#) [To Genome Browser](#)
[NCBI blastp](#) [Predicted number of transmembrane domains: 1](#)

estExt_Genewise1.C_6_t30338 To Genome Browser

1 1171 523 1792

1 53 106 158 210 263 315 367 419 472 523

Signalp: secreted by HMM (0.535) by NN 3/5 at 26/27

PR00464: EP450II [FPrintScan]

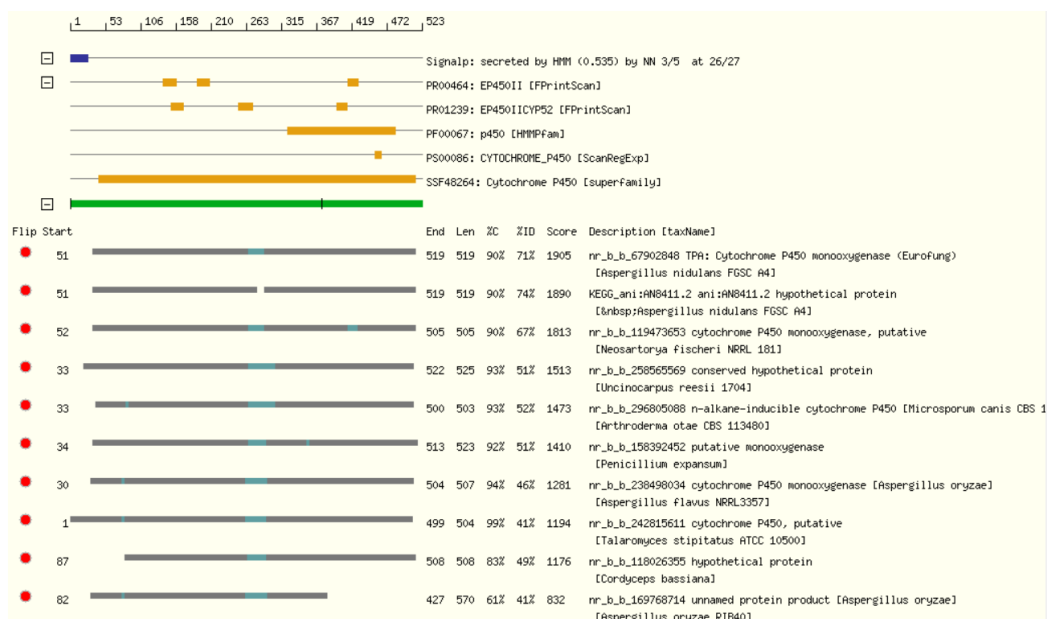
PR01239: EP450IICYP52 [FPrintScan]

PF00067: p450 [HMMFam]

PS00086: CYTOCHROME_P450 [ScanRegExp]

SSF48264: Cytochrome P450 [superfamily]

- 9) So this really is a cytochrome p450 monooxygenase, a class of enzymes that often modify core structures of SM biosynthetic pathways. Similar perusal of the other genes of the cluster says that this cluster is an excellent candidate for synthesis of your SM.



10) One explanation for *S. musiva* having this cluster and the congeneric *S. populica* not is that the former acquired the cluster by horizontal gene transfer from a phylogenetically distant source. The ‘best Blast hit’ of the cytochrome p450 enzyme supports this hypothesis. To see if the core enzyme can shed some light, click the web browser back button to go back to the SM CLUSTERS graphic, and click on the same PKS-NRPS core gene we moused over earlier. The protein page is rich in details, including domains and the top 10 hits. All of the hits are high quality and are from Eurotiomycetes. This cluster is an excellent candidate for horizontal gene transfer from the Eurotiomycetes!

References

- Dhillon B, Feau N, Aerts AL, Beauseigle S, Bernier L, Copeland A, Foster A, Gill N, Henrissat B, Herath P, LaButti KM, Levasseur A, Lindquist EA, Majoor E, Ohm RA, Pangilinan JL, Pribovo A, Saddler JN, Sakalidis ML, de Vries RP, Grigoriev IV, Goodwin SB, Tanguay P, Hamelin RC. Horizontal gene transfer and gene dosage drives adaptation to wood colonization in a tree pathogen. *Proc Natl Acad Sci U S A*. 2015 Mar 17;112(11):3451-6. doi: 10.1073/pnas.1424293112. Epub 2015 Mar 2. PubMed PMID: 25733908
- Schumann J, Hertweck C. Molecular basis of cytochalasan biosynthesis in fungi: gene cluster analysis and evidence for the involvement of a PKS-NRPS hybrid synthase by RNA silencing. *J Am Chem Soc*. 2007 Aug 8;129(31):9564-5. Epub 2007 Jul 18. PubMed PMID: 17636916.