Horizontal Transfer and Death of a Fungal Secondary Metabolic Gene Cluster

Matthew A. Campbell, Antonis Rokas*, and Jason C. Slot

Department of Biological Sciences, Vanderbilt University

*Corresponding author: E-mail: antonis.rokas@vanderbilt.edu.

Accepted: 25 January 2012

Abstract

A cluster composed of four structural and two regulatory genes found in several species of the fungal genus *Fusarium* (class Sordariomycetes) is responsible for the production of the red pigment bikaverin. We discovered that the unrelated fungus *Botrytis cinerea* (class Leotiomycetes) contains a cluster of five genes that is highly similar in sequence and gene order to the *Fusarium* bikaverin cluster. Synteny conservation, nucleotide composition, and phylogenetic analyses of the cluster genes indicate that the *B. cinerea* cluster was acquired via horizontal transfer from a *Fusarium* donor. Upon or subsequent to the transfer, the *B. cinerea* gene cluster became inactivated; one of the four structural genes is missing, two others are pseudogenes, and the fourth structural gene shows an accelerated rate of nonsynonymous substitutions along the *B. cinerea* lineage, consistent with relaxation of selective constraints. Interestingly, the *bik4* regulatory gene is still intact and presumably functional, whereas *bik5*, which is a pathway-specific regulator, also shows a mild but significant acceleration of evolutionary rate along the *B. cinerea* lineage. This selective preservation of the *bik4* regulator suggests that its conservation is due to its likely involvement in other non–bikaverin-related biological processes in *B. cinerea*. Thus, in addition to novel metabolism, horizontal transfer of wholesale metabolic gene clusters might also be contributing novel regulation.

Key words: pseudogene, bikaverin, secondary metabolites, pathway degeneration, *Botrytis*, *Fusarium*.

Fungi are the primary decomposers of organic matter in many different natural ecosystems. As a result, fungi have evolved a substantial and diverse arsenal of genes involved in intermediary and secondary metabolism that enables them not only to break down and extract energy from a remarkable variety of different substrates but also to generate a cadre of toxins with which they can fend off competition and carve their ecological niches (Keller et al. 2005). Interestingly, the genes from many of the metabolic pathways that bestow fungi with such diverse physiologies are physically linked or clustered (Keller and Hohn 1997). In recent years, studies have shown that several of these metabolic gene clusters have undergone wholesale horizontal transfers (Patron et al. 2007; Slot and Hibbett 2007; Khaldi et al. 2008; Slot and Rokas 2010, 2011; Khaldi and Wolfe 2011), suggesting that they might have played a key role in the diversification of fungal metabolism. As the precise molecular mechanisms facilitating such transfers of entire metabolic gene clusters are largely unknown, an interesting and so far unanswered question is whether such transfers always result in pathways that are functional in recipient species.

To address this question, we studied the evolution of the metabolic gene cluster responsible for the production of bikaverin, a red pigment with antibacterial and antitumor activity (reviewed in Limon et al. 2010). Bikaverin production is only known from several species in the genus *Fusarium* (class Sordariomycetes, phylum Ascomycota) as well as from a single species from each of two other genera in Sordariomycetes (Limon et al. 2010). In *Fusarium fujikuroi*, where bikaverin synthesis has been best characterized, the gene cluster spans 18 kb and contains four structural genes, encoding for three biosynthetic enzymes (*bik1*, *bik2*, and *bik3*) and a transporter (*bik6*) and two regulatory genes (*bik4* and *bik5*) (Wiemann et al. 2009).

As part of a larger survey of the evolution of fungal metabolic pathways (Slot and Rokas 2010, 2011), we discovered a cluster of five genes in the genome of the unrelated necrotrophic plant pathogen *Botrytis cinerea* (phylum Ascomycota, class Leotiomycetes) that is identical in synteny to the *Fusarium* bikaverin cluster. This was surprising not only because of the large evolutionary distance separating Sordariomycetes and Leotiomycetes (James et al. 2006) but

also because these are the only two lineages among the 103 draft fungal genomes we examined that contained this gene cluster.

Materials and Methods

The *B. cinerea* bikaverin gene cluster was discovered using previously described techniques (Slot and Rokas 2010, 2011). Briefly, amino acid sequences similar to genes of a cluster were detected with BlastP (Altschul et al. 1990) in a local proteome database of 103 fungal genomes, and homologs were considered clustered when separated by no more than seven genes on a chromosome.

Conservation of synteny in the bikaverin cluster region between *F. oxysporum*, *F. verticillioides*, *B. cinerea*, and *S. sclerotiorum* was estimated with reference to sequence annotation of high scoring BlastP hits in GenBank and confirmed by an alignment of the bikaverin cluster genes and 20 kb on both sides using Mauve, version 2.3.1 (Darling et al. 2010).

Homologs from each protein in the bikaverin cluster were retrieved by BlastP from GenBank and our local database. Proteins with *e* values $<1 \times 10^{-4}$, query coverage >50%, and an amino acid sequence similarity >50% were combined and aligned using MAFFT, version 6.847 (Katoh and Toh 2008). Poorly aligned taxa were removed, and the sequences were realigned. Sites containing >30% missing data were removed using trimAl, version 1.2 (Capella-Gutierrez et al. 2009). Maximum likelihood (ML) analysis was performed using RAxML, version 7.2.0 (Stamatakis 2006), with 100 bootstrap replicates under the PROTGAMMAJTT model of amino acid substitution. Constraint analyses, where topologies constrained to have all sequences from Sordariomycetes or Eurotiomycetes monophyletic, were compared with the ML topology using the Shimodaira-Hasegawa test (Shimodaira and Hasegawa 1999) as implemented in RAxML (Stamatakis 2006) (table 1).

We also conducted phylogenetic analyses using nucleotide data. The protein alignments of *bik2–bik6* were converted to nucleotide alignments using the PAL2NAL script (Suyama et al. 2006), this time including only genes from *F. fujikuroi*, *F. oxysporum*, *F. verticillioides*, two *B. cinerea* strains (T4 and B05.10), and the nearest outgroup from the RAxML amino acid analyses. ML analysis was performed with 100 bootstrap replicates and the GTRGAMMA model of nucleotide substitution. Additionally, an ML analysis of an EF1-α amino acid alignment was used to illustrate expected species relationships in figure 1.

To calculate the evolutionary rates of the bikaverin genes, we examined variation in selection pressure along branches of the species tree and tested each gene for evidence of positive selection using the CODEML module from PAML, version 4.4 (Yang 2007). To do so, we evaluated the log likelihood of the null hypothesis H_0, under which all branches of the phylogeny exhibited the same ω ratio of nonsynonymous (d*N*) to synonymous (d*S*) substitutions against the alternative hypothesis H_1, under which the ω ratio along the *B. cinerea* branches was different from that in the rest of the branches of the phylogeny.

Results and Discussion

To examine the origin of the *B. cinerea* five-gene cluster, we retrieved the six genes comprising the bikaverin cluster from *F. fujikuroi* (Wiemann et al. 2009), *F. oxysporum*, and *F. verticillioides* (Ma et al. 2010). Sequence similarity searches of *Fusarium* bikaverin genes against *B. cinerea* genome data from two different strains (Amselem et al. 2011) showed that the *B. cinerea* cluster contained homologs of five of the six bikaverin genes (fig. 1) but lacked a *bik1* homolog encoding a polyketide synthase. Examination of the nucleotide alignments for each of the five genes from *Fusarium* and *Botrytis* showed that the structural genes *bik2* and *bik3* contained internal stop codons and indels in both *B. cinerea* strains, suggesting that they are pseudogenes (ψ). The finding that the three biosynthetic enzymes responsible for bikaverin synthesis are either missing or inactivated indicates that the *B. cinerea* bikaverin cluster is nonfunctional.

Examination of the genomic regions containing the bikaverin gene cluster in *B. cinerea* and *Fusarium* showed that the cluster genes had the same gene order and orientation in both lineages (fig. 1). The bikaverin gene cluster was not present in the genome of *Sclerotinia sclerotiorum* (class Leotiomycetes), a close relative to *B. cinerea*, or in any other of the 103 fungal genomes examined (supplementary table S1, Supplementary Material online). Even though the bikaverin gene cluster is absent from *S. sclerotiorum*, *B. cinerea* and *S. sclerotiorum* showed conservation of synteny in the regions flanking the bikaverin gene cluster. Specifically, a region of 17 annotated genes in *S. sclerotiorum* shares common order with nine annotated homologs flanking the bikaverin cluster in *B. cinerea*, including the final gene on the *B. cinerea* contig (fig. 1), suggesting that the cluster originated after the divergence of the two lineages. Overall identity and synteny conservation between these two genomes suggest their genetic divergence was recent (Amselem et al. 2011), but the poorness of the fungal fossil record makes the estimation of an exact date difficult. The flanking regions of the bikaverin gene cluster in *F. oxysporum* and *F. verticillioides* also showed synteny conservation, although they contained genes unrelated to those found in the *B. cinerea–S. sclerotiorum* flanking regions (fig. 1).

Comparison of the transcriptome sequence content and divergence between *Fusarium* and *B. cinerea* indicated that the bikaverin genes in the two lineages were much more similar than would be expected if they had been inherited vertically, suggesting that the *B. cinerea* bikaverin cluster might have been acquired via horizontal transfer.

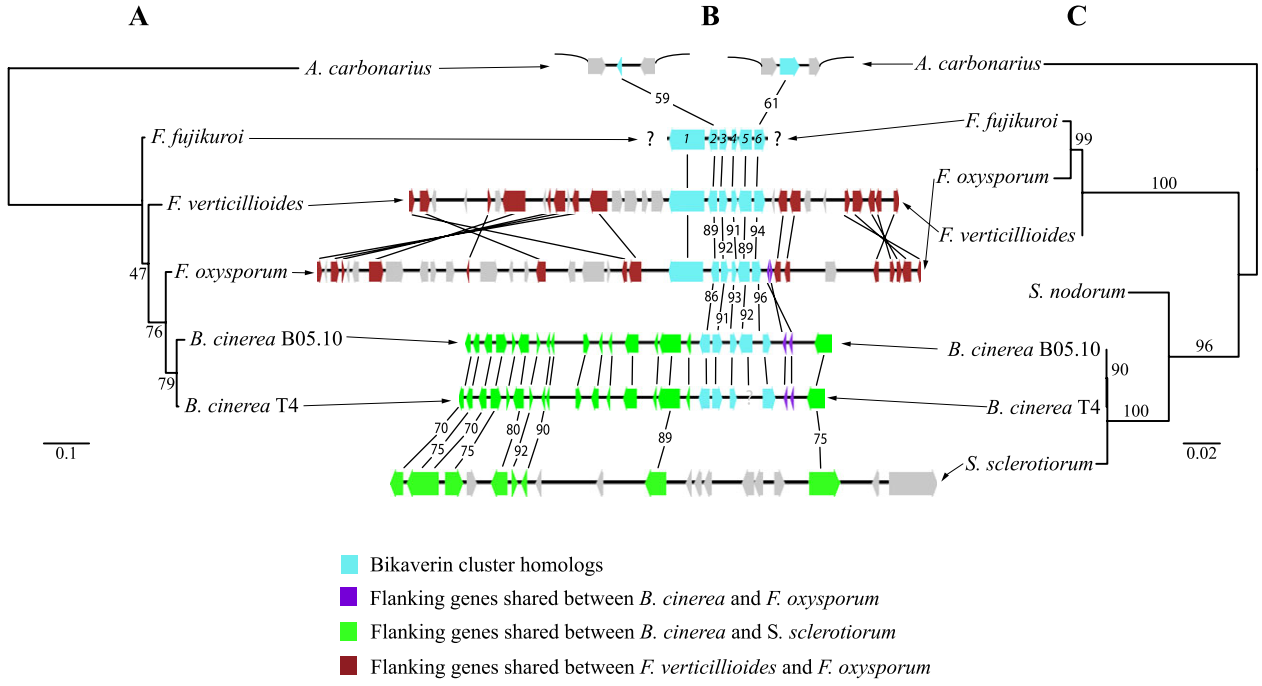

FIG. 1.—The bikaverin gene cluster was horizontally transferred from *Fusarium* to *Botrytis*. (*A*) Phylogeny of the *bik6* gene. Note that the *B. cinerea bik6* sequences nest within the *Fusarium* clade. (*B*) Conservation of synteny of the genomic region containing the bikaverin gene cluster in *Fusarium* species, two strains of *B. cinerea* (T4 and B05.10), and *Sclerotinia sclerotiorum*, a close relative of *B. cinerea*. Bikaverin gene cluster homologs are indicated by blue-colored boxes. *Fusarium fujikuroi* bikaverin genes are labeled *1–6* and correspond to the *bik1–bik6* genes. *Fusarium* flanking region homologs are indicated by red-colored boxes, and *B. cinerea–S. sclerotiorum* flanking region homologs by green-colored boxes. Genes that lack homologs are colored gray. Lines between genes indicate homolog pairs, whereas numbers indicate percentage of nucleotide identity between (select) homologs. Two lines connecting to the same gene feature result from differential annotation of the region between genomes. Note that the degree of synteny conservation in the regions flanking the gene cluster in *B. cinerea* and *S. sclerotiorum* and the absence of the gene cluster in *S. sclerotiorum*. Synteny is also conserved between *Fusarium verticillioides* and *Fusarium oxysporum*, with the exception of two inversions, one in each of the two flanking regions. The question mark indicates a gap in the assembly of the T4 strain of *B. cinerea*. (*C*) Phylogeny of *EF1-α*, a housekeeping gene, showing the established species relationships.

Specifically, the average similarity between *Fusarium* and *B. cinerea* proteins from the bikaverin gene cluster was 91%, whereas the average similarity between the two proteomes was 57%. Furthermore, the GC content of the bikaverin clusters in *Fusarium* species and *B. cinerea* was similar to the transcriptome-wide GC content in *Fusarium* species (>50%) but different from the GC content of the *B. cinerea* transcriptome (46%; table 2).

To investigate further whether the bikaverin gene cluster in *B. cinerea* was horizontally transferred from *Fusarium*, we conducted phylogenetic analyses of all proteins in and flanking the bikaverin gene cluster. All analyses supported clades composed of *B. cinerea* and *Fusarium* homologs (supplementary fig. S2, Supplementary Material online). For example, the *B. cinerea* ψ*bik2* and *bik5* genes group with the *F. verticillioides* and *F. fujikuroi* homologs (72% and 75% clade support, respectively), whereas the ψ*bik3*, *bik4*, and *bik6* genes group with *F. oxysporum* (89%, 66%, and 76% clade support, respectively). Furthermore, constraint analyses (Shimodaira and Hasegawa 1999), which forced orthologs from Sordariomycetes to be monophyletic, rejected the null hypothesis of vertical inheritance for ψ*bik3* and *bik4* ($P < 0.05$; table 1). In summary, nucleotide composition, synteny conservation, and phylogenetic analyses suggest

Table 1
Bikaverin Gene Constraint Analyses

Gene	Significantly Worse?	Likelihood of Optimal Tree	Likelihood of Constrained Tree	Alignment Length (bp)
bik2	*	−72,633.837503	*	1,450
bik3	Yes	−72,393.402090	−72,416.963878	1,730
bik4	Yes	−13,486.913547	−13,560.939705	396
bik5	No	−8,870.185847	−8,875.447740	2,961
bik6	No	−40,558.088803	−40,561.178478	2,053

NOTE.—*, No constraint was possible due to the topology of the tree.

Table 2
Average GC Content and Codon Adaptation Index Values for the Entire Transcriptome and for Bikaverin Genes from *Fusarium* Species and *Botrytis cinerea*

Taxon	Gene Set	GC Content	Codon Adaptation Index
Fusarium oxysporum	Transcriptome	51.85	0.839
Fusarium verticillioides	Transcriptome	52.03	0.796
Botrytis cinerea	Transcriptome	46.31	0.803
Mean (SD)		50.06 (3.25)	0.81 (0.023)
F. oxysporum	Bikaverin cluster	53.26	0.774
F. verticillioides	Bikaverin cluster	53.54	0.753
B. cinerea	Bikaverin cluster	51.91	0.772
Mean (SD)		52.90 (0.87)	0.77 (0.012)

NOTE.—SD, standard deviation.

that the most likely explanation for the presence of the partially inactivated five-gene bikaverin cluster in *B. cinerea* is horizontal transfer from *Fusarium*.

Four of the six gene phylogenies also showed that one or more members of the genus *Aspergillus* is the immediate outgroup to the clade formed by the clustered bikaverin genes from *Fusarium* and *Botrytis*. The presence of an *Aspergillus* species as the adjacent group in four of the six gene phylogenies (supplementary figs. S1 and S2, Supplementary Material online) suggests that the bikaverin gene cluster in *Fusarium* might have also originated, at least partially, from genes also acquired by horizontal transfer.

Given that three of the four structural genes in the *B. cinerea* bikaverin gene cluster are either missing or inactivated, we examined whether there was evidence for variation in selection pressure along the *B. cinerea* lineage. We found that the ω ratio of nonsynonymous (d*N*) to synonymous (d*S*) substitutions along the *B. cinerea* branch was significantly higher that the ω ratio in the rest of the phylogeny for four of the five bikaverin genes. Specifically, the two structural pseudogenes (ψ*bik2* and ψ*bik3*) as well as the transporter *bik6* showed strong acceleration of the ω ratio along the *B. cinerea* branch (table 3), which is likely the result of relaxation of selection in these genes. In contrast, the two regulatory genes show either no evidence for rate acceleration (*bik4*) or milder but significant relaxation of selection (*bik5*) when their ω ratio is examined along the same branch.

In summary, the available data indicate that the *B. cinerea* bikaverin gene cluster was acquired via horizontal transfer from a *Fusarium* donor. Unlike any other reported horizontal transfers of metabolic gene clusters between fungi (Patron et al. 2007; Slot and Hibbett 2007; Khaldi et al. 2008; Slot and Rokas 2010, 2011; Khaldi and Wolfe 2011), it appears that the *B. cinerea* gene cluster became inactivated upon or subsequent to the transfer. Interestingly, only the structural genes of the pathway appear to have been inactivated, whereas the two regulatory genes are still intact and presumably functional.

The nonrandom inactivation of the *B. cinerea* bikaverin gene cluster suggests functional constraints on the order in which the genes became nonfunctional. One plausible model for pathway degeneration suggests that the order of loss of pathway genes might be inversely related to their degree of pleiotropy (Hittinger et al. 2004), with more pleiotropic genes being retained longer. In the context of fungal metabolic gene clusters, this model would predict that regulatory genes decay last, as several studies have shown that the regulatory genes are more pleiotropic than their structural counterparts (e.g., Price et al. 2006). This prediction appears to hold true for the bikaverin gene cluster; for example, the available evidence from studies in *Fusarium* suggests that the *bik4* gene is the most pleiotropic (Wiemann et al. 2009; Limon et al. 2010), and our results indicate that *bik4* has not only been retained but also that it is the only cluster gene that does not show any evidence of acceleration in evolutionary rate. In contrast, the *bik5* gene, whose protein product is a pathway-specific regulator in *Fusarium* (Wiemann et al. 2009), shows a mild but significant acceleration in evolutionary rate. This acceleration is presumably because, in the absence or nonfunctionality of the structural genes, all functional constraints on *bik5* along the *B. cinerea* lineage have been removed.

The preferential retention of the regulatory genes, especially *bik4*, in the inactivated bikaverin gene cluster of *B. cinerea* is consistent with observations from other pathway degeneration events that do not involve horizontal transfer. For example, the repressor protein gal4p, which has been shown to regulate nongalactose genes in some yeasts (Martchenko et al. 2007; Rokas and Hittinger 2007), is the only one retained from the galactose pathway in *Eremothecium gossypii*, whereas the three regulatory genes of the

Table 3
Evolutionary Rates of Coding Sequences in the Bikaverin Gene Cluster

Gene	H_0 ln*L*	H_1 ln*L*	2Δ*L*	*P* Value	H_0 ω	H_1 ω (*Botrytis cinerea* branch)	H_1 ω (rest of the tree)
bik2	−3719.09	−3,694.13	49.92	<0.0001	0.171	1.301	0.092
bik3	−2,851.83	−2,837.90	27.86	<0.0001	0.058	0.225	0.023
bik4	−1,963.67	−1,961.92	3.50	0.0614	0.088	N/A (d*S* = 0)	0.081
bik5	−3,613.34	−3,608.40	9.88	0.0017	0.088	0.185	0.065
bik6	−2,621.81	−2,598.96	45.69	<0.0001	0.078	0.844	0.029

NOTE.—N/A, not applicable.

same pathway in *Saccharomyces kudriavzevii* were likely inactivated last during pathway degeneration (Hittinger et al. 2004, 2010). More indirectly, a recent study identified remnants of the genomic expression program for xylose assimilation in the yeast *Lodderomyces elongisporus*, even though the species has lost the corresponding structural genes (Wohlbach et al. 2011).

We propose that co-option to non–bikaverin-related biological processes favored the selective preservation of the regulatory genes of the *B. cinerea* bikaverin gene cluster long after the majority of structural genes was inactivated. This biased decay of genes following a recent horizontal transfer of a fungal secondary metabolism gene cluster suggests that in addition to novel metabolism, transfer of wholesale metabolic gene clusters in fungi might also be a contributor of novel regulators.

Supplementary Material

Supplementary table S1 and figures S1 and S2 are available at *Genome Biology and Evolution online* (http://gbe.oxfordjournals.org/).

Acknowledgments

This work was conducted in part with the resources of the Advanced Computing Center for Research and Education at Vanderbilt University. This work was supported by funds provided by the Searle Scholars Program (A.R.) and the National Science Foundation (DBI-0805625 to J.C.S. and DEB-0844968 to A.R.).

Literature Cited

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J Mol Biol. 215:403–410.

Amselem J, et al. 2011. Genomic analysis of the necrotrophic fungal pathogens *Sclerotinia sclerotiorum* and *Botrytis cinerea*. PLoS Genet. 7:e1002230.

Capella-Gutierrez S, Silla-Martinez JM, Gabaldon T. 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics 25:1972–1973.

Darling AE, Mau B, Perna NT. 2010. progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. PLoS One 5:e11147.

Hittinger CT, Rokas A, Carroll SB. 2004. Parallel inactivation of multiple *GAL* pathway genes and ecological diversification in yeasts. Proc Natl Acad Sci U S A. 101:14144–14149.

Hittinger CT, et al. 2010. Remarkably ancient balanced polymorphisms in a multi-locus gene network. Nature 464:54–58.

James TY, et al. 2006. Reconstructing the early evolution of Fungi using a six-gene phylogeny. Nature 443:818–822.

Katoh K, Toh H. 2008. Recent developments in the MAFFT multiple sequence alignment program. Brief Bioinform. 9:286–298.

Keller NP, Hohn TM. 1997. Metabolic pathway gene clusters in filamentous fungi. Fungal Genet Biol. 21:17–29.

Keller NP, Turner G, Bennett JW. 2005. Fungal secondary metabolism—from biochemistry to genomics. Nat Rev Microbiol. 3:937–947.

Khaldi N, Collemare J, Lebrun MH, Wolfe KH. 2008. Evidence for horizontal transfer of a secondary metabolite gene cluster between fungi. Genome Biol. 9:R18.

Khaldi N, Wolfe KH. 2011. Evolutionary origins of the fumonisin secondary metabolite gene cluster in *Fusarium verticillioides* and *Aspergillus niger*. Int J Evol Biol. 2011:423821.

Limon MC, Rodriguez-Ortiz R, Avalos J. 2010. Bikaverin production and applications. Appl Microbiol Biotechnol. 87:21–29.

Ma LJ, et al. 2010. Comparative genomics reveals mobile pathogenicity chromosomes in *Fusarium*. Nature 464:367–373.

Martchenko M, Levitin A, Hogues H, Nantel A, Whiteway M. 2007. Transcriptional rewiring of fungal galactose metabolism circuitry. Curr Biol. 17:1007–1013.

Patron NJ, et al. 2007. Origin and distribution of epipolythiodioxopiperazine (ETP) gene clusters in filamentous ascomycetes. BMC Evol Biol. 7:174.

Price MS, et al. 2006. The aflatoxin pathway regulator AflR induces gene transcription inside and outside of the aflatoxin biosynthetic cluster. FEMS Microbiol Lett. 255:275–279.

Rokas A, Hittinger CT. 2007. Transcriptional rewiring: the proof is in the eating. Curr Biol. 17:R626–R628.

Shimodaira H, Hasegawa M. 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. Mol Biol Evol. 16:1114–1116.

Slot JC, Hibbett DS. 2007. Horizontal transfer of a nitrate assimilation gene cluster and ecological transitions in fungi: a phylogenetic study. PLoS One 2:e1097.

Slot JC, Rokas A. 2010. Multiple *GAL* pathway gene clusters evolved independently and by different mechanisms in fungi. Proc Natl Acad Sci U S A. 107:10136–10141.

Slot JC, Rokas A. 2011. Horizontal transfer of a large and highly toxic secondary metabolic gene cluster between fungi. Curr Biol. 21:134–139.

Stamatakis A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22:2688–2690.

Suyama M, Torrents D, Bork P. 2006. PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. Nucleic Acids Res. 34:W609–W612.

Wiemann P, et al. 2009. Biosynthesis of the red pigment bikaverin in *Fusarium fujikuroi*: genes, their function and regulation. Mol Microbiol. 72:931–946.

Wohlbach DJ, et al. 2011. Comparative genomics of xylose-fermenting fungi for enhanced biofuel production. Proc Natl Acad Sci U S A. 108:13212–13217.

Yang Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. Mol Biol Evol. 24:1586–1591.

Associate editor: Kenneth Wolfe