Current Biology 21, 134–139, January 25, 2011 ©2011 Elsevier Ltd All rights reserved DOI 10.1016/j.cub.2010.12.020

Report

Horizontal Transfer of a Large and Highly Toxic Secondary Metabolic Gene Cluster between Fungi

Jason C. Slot[1] and Antonis Rokas[1,*]
[1]Department of Biological Sciences, Vanderbilt University, Nashville, TN 37235, USA

Summary

Genes involved in intermediary and secondary metabolism in fungi are frequently physically linked or clustered. For example, in *Aspergillus nidulans* the entire pathway for the production of sterigmatocystin (ST), a highly toxic secondary metabolite and a precursor to the aflatoxins (AF), is located in a ~54 kb, 23 gene cluster. We discovered that a complete *ST* gene cluster in *Podospora anserina* was horizontally transferred from *Aspergillus*. Phylogenetic analysis shows that most *Podospora* cluster genes are adjacent to or nested within *Aspergillus* cluster genes, although the two genera belong to different taxonomic classes. Furthermore, the *Podospora* cluster is highly conserved in content, sequence, and microsynteny with the *Aspergillus ST/AF* clusters and its intergenic regions contain 14 putative binding sites for AflR, the transcription factor required for activation of the *ST/AF* biosynthetic genes. Examination of ~52,000 *Podospora* expressed sequence tags identified transcripts for 14 genes in the cluster, with several expressed at multiple life cycle stages. The presence of putative AflR-binding sites and the expression evidence for several cluster genes, coupled with the recent independent discovery of ST production in *Podospora* [1], suggest that this HGT event probably resulted in a functional cluster. Given the abundance of metabolic gene clusters in fungi, our finding that one of the largest known metabolic gene clusters moved intact between species suggests that such transfers might have significantly contributed to fungal metabolic diversity.

Results and Discussion

The overwhelming majority of documented horizontal gene transfer (HGT) events in eukaryotes concerns the transfer of single or a few genes from bacterial donors [2–4]. In contrast, there are far fewer reports of eukaryote-to-eukaryote HGT events. Typically, such events involve the transfer of a single or a few genes between very distantly related organisms belonging to different kingdoms [5, 6]. The small amount of genetic material typically transferred between eukaryotes and the large phylogenetic distances separating eukaryotic donors and recipients raise questions as to whether larger genetic fragments undergo HGT as well as whether such HGT events also take place between more closely related species.

Genes involved in intermediary and secondary metabolism make good candidates for HGT events between fungi. This is so because genes participating in fungal metabolic pathways are frequently physically clustered, raising the possibility of wholesale pathway transfers in single events [7]. Several fungal metabolic clusters contain >15 genes and span tens of kilobases [8–13]. One such large fungal gene cluster is comprised of the genes required for the production of sterigmatocystin (ST), a highly toxic, mutagenic and carcinogenic secondary metabolite and the precursor to aflatoxins (AF) [11, 14]. In *Aspergillus nidulans*, the *ST* gene cluster contains 23 genes dispersed across ~54 kb [11], whereas the related *AF* gene cluster in *Aspergillus flavus* is a ~67 kb, 26 gene cluster [13, 15]. Gene content is highly conserved between the two clusters, whereas the gene order, orientation, and sequence similarity are less conserved (Figure 1) [11, 16].

In a survey of 94 fungal genomes (see Experimental Procedures), we noted the existence of an intact 24 gene, ~57 kb cluster in *Podospora anserina* with striking similarity to the *A. nidulans ST* cluster (Figure 1). The *P. anserina* cluster is nested within a very large contig and is thus unlikely to be the product of contamination [17]. Furthermore, we did not find any other similar clusters outside the *Aspergillus* lineage (Figure S1, available online). Comparison of the two species' clusters against the *A. flavus AF* cluster [13, 15] showed that the gene order and orientation of the *P. anserina* and *A. nidulans* clusters is extremely well-conserved, well above what would be expected based on their genome-wide evolutionary divergence, whereas the *A. flavus AF* cluster is more divergent (Figure 1).

Sequence analysis of the *P. anserina* cluster genes showed that all have intact open reading frames and appear similar in length to *A. nidulans ST* cluster genes. The average amino acid identity between *P. anserina* and *A. nidulans* gene cluster orthologs is 63%, a value very similar to the 62% identity observed between *A. nidulans* and *A. flavus*. This similarity in identity percentages is surprising because the two genera belong to distinct taxonomic classes; *Podospora* is a member of the class Sordariomycetes, *Aspergillus* a member of the class Eurotiomycetes, both within the phylum Ascomycota [18, 19]. In contrast, the average amino acid identity in a data set of 4915 orthologs inferred from a comparison of the three species' proteomes is 75% for *A. nidulans-A. flavus*, but only 53% for *A. nidulans-P. anserina*. The remarkable conservation in microsynteny and sequence raised the possibility that the *P. anserina ST* cluster originated via HGT from *Aspergillus*.

To test the hypothesis that the *P. anserina ST* cluster originated via HGT from *Aspergillus*, we conducted phylogenetic analyses on all the homologs of *A. nidulans ST* cluster genes identified across fungi (see Experimental Procedures). Under the HGT hypothesis, the expectation is that *P. anserina* genes would nest within *Aspergillus* or Eurotiomycetes in cluster gene phylogenies and be in stark contrast to the species phylogeny (Figures 2A and 2B). Phylogenetic analysis of 23 cluster genes revealed six distinct topological patterns (Figure 2C and Figure S2):

(1) six gene phylogenies support an adjacent-group relationship between the *A. nidulans* clan (*A. nidulans* and *Aspergillus. ochraceoroseus*) and *P. anserina* (pattern 1 in Figure 2C),
(2) seven gene phylogenies support an adjacent-group relationship between other *Aspergillus* species (i.e.,

*Correspondence: antonis.rokas@vanderbilt.edu

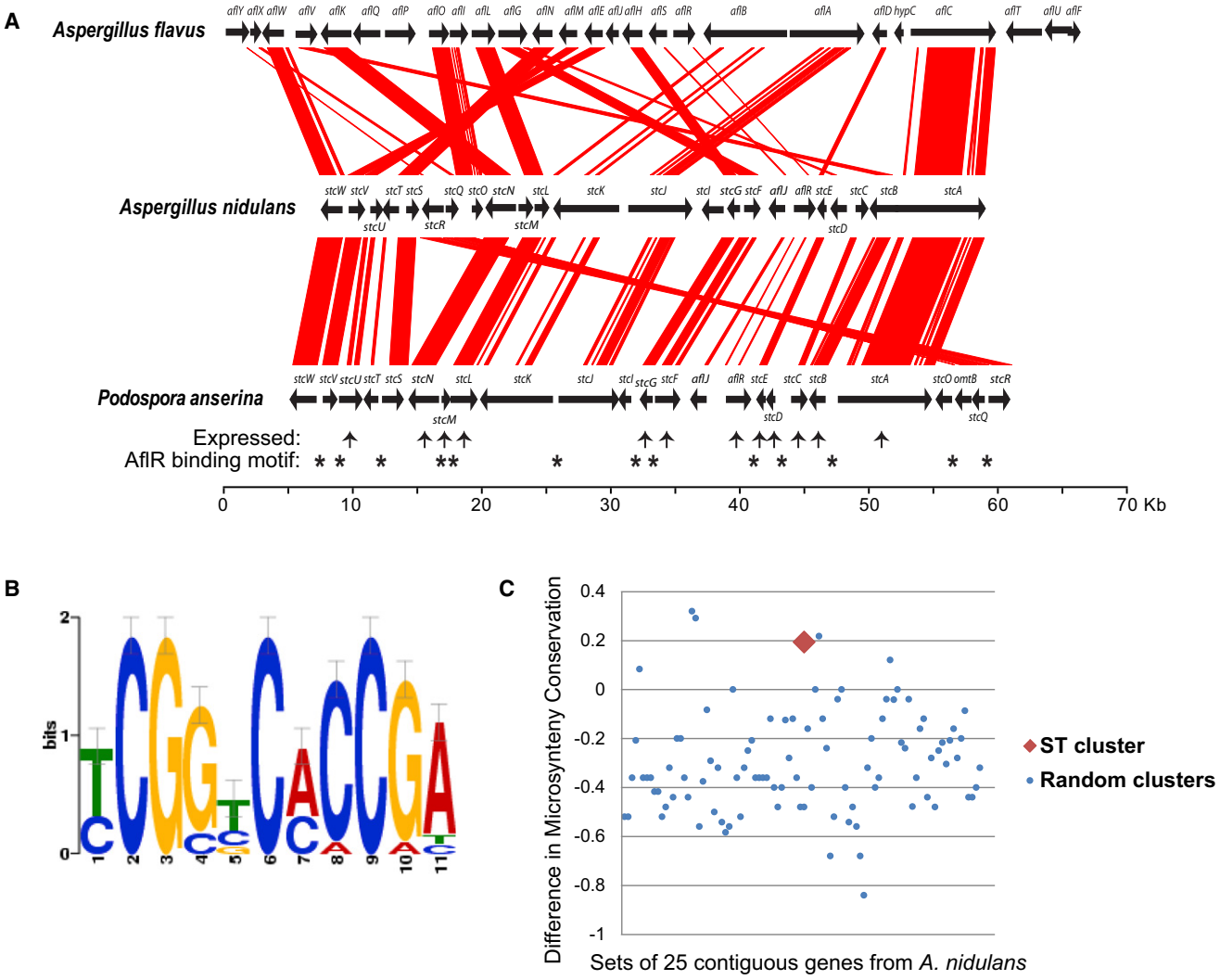

Figure 1. The *P. anserina ST* Gene Cluster Is Remarkably Similar to the *A. nidulans ST* Gene Cluster

(A) Microsynteny and sequence conservation between the *ST/AF* gene clusters from *A. flavus*, *A. nidulans*, and *P. anserina*. Alignment blocks correspond to DNA fragments exhibiting significant similarity when the genomic regions comprising the gene clusters are compared with the discontinuous megablast algorithm [30]. *P. anserina* genes with evidence of expression are indicated by arrows, whereas intergenic regions containing the consensus AflR-binding motif identified in a de novo search by the MEME algorithm [54] are shown by asterisks. All the genomes used in this study are shown in Table S1 and all gene clusters detected in Figure S1.

(B) The predicted consensus-binding motif for the transcription factor AflR identified by examination of the *P. anserina* intergenic regions. Error bars are equal to twice the small sample correction *e*(*n*) for sequence logos [55].

(C) Plot of the difference in microsynteny conservation between *P. anserina*:*A. nidulans* and *A. flavus*:*A. nidulans* (shown on the y axis) for the *ST* cluster and for each of the 100 randomly selected sets of 25 contiguous genes from the *A. nidulans* genome (shown on the x axis). See also Figure S1 and Table S1.

excluding species in the *A. nidulans* clan) and *P. anserina* (pattern 2 in Figure 2C),

(3) one gene phylogeny supports an adjacent-group relationship between *Aspergillus* and *P. anserina* (pattern 3 in Figure 2C),

(4) five gene phylogenies support the monophyly of the *Aspergillus-Podospora* clan but fail to resolve relationships within the clan (pattern 4 in Figure 2C),

(5) three gene phylogenies support an adjacent-group relationship between the *A. nidulans* clan (*A. nidulans* and *A. ochraceoroseus*) and *P. anserina*, but lack sequences from other *Aspergillus* species (pattern 5 in Figure 2C), and

(6) two gene phylogenies contain sequences uniquely present in *Aspergillus* and *P. anserina* (pattern 6 in Figure 2C).

All gene phylogenies are consistent with the hypothesis that *P. anserina* acquired its *ST* cluster via HGT from *Aspergillus* (Figure 2). Furthermore, 9 out of 23 gene phylogenies rejected a monophyletic *Aspergillus* clan (the remaining 14 genes were uninformative; Table 1). Precisely determining the donor lineage is challenging, given that single gene phylogenies are often unreliable [20, 21]. For example, whereas six gene phylogenies support an origin of the *P. anserina ST* cluster from an *A. nidulans* relative after its divergence from other

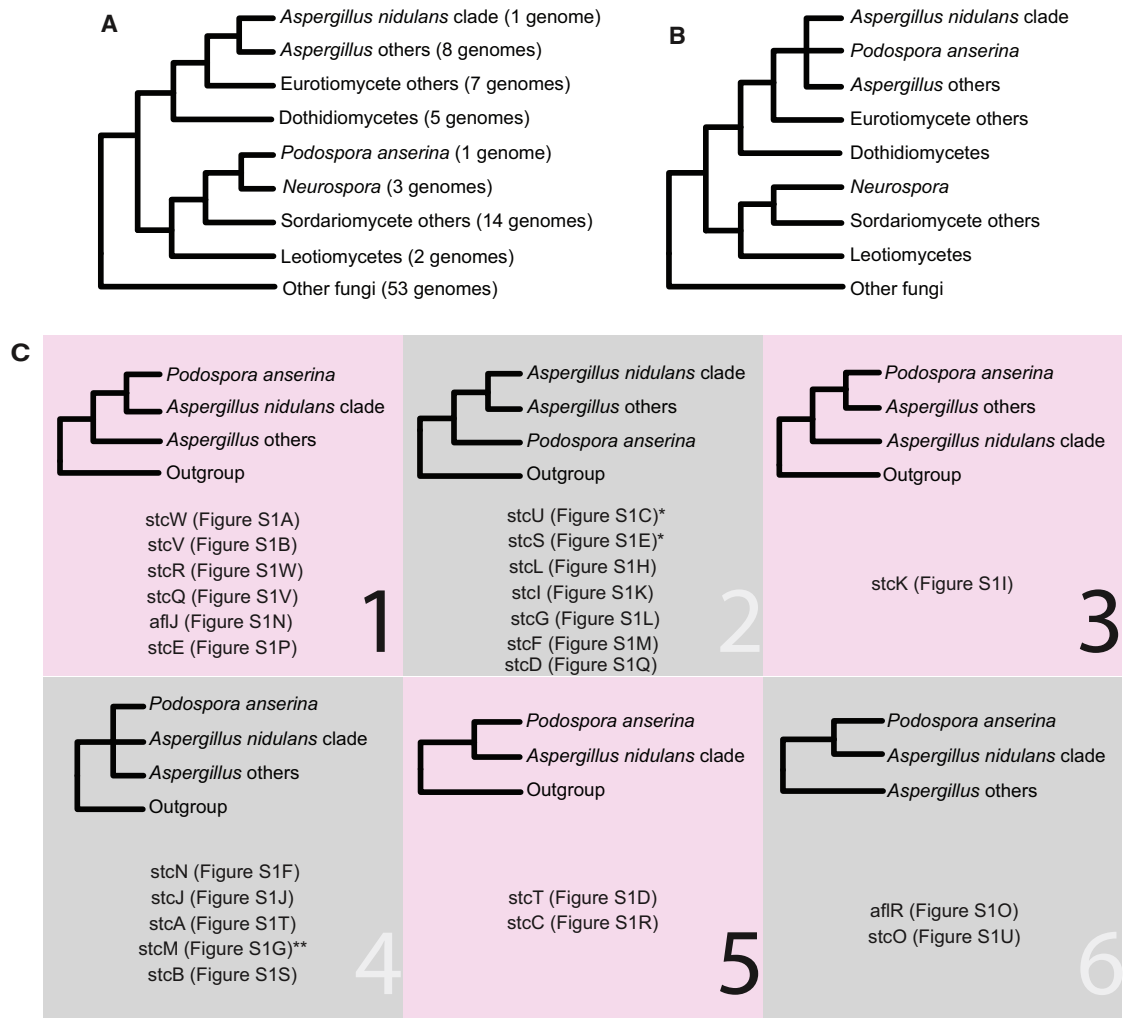

Figure 2. Cluster Gene Phylogenies Indicate that the *P. anserina ST* Cluster Was Acquired via HGT from *Aspergillus*

(A) Fungal species phylogeny [19]. Numbers of genomes from each lineage analyzed in this study are shown in parentheses.

(B) Phylogeny expected under the hypothesis that the *P. anserina ST* gene cluster originated via HGT from *Aspergillus*.

(C) The 23 cluster gene phylogenies correspond to six different patterns. The genes supporting each of the patterns are listed below each pattern. Single asterisks (*) denote gene phylogenies where *A. nidulans* and *A. ochraceoroseus* do not form a clan. Double asterisks (**) denote gene phylogenies that contain additional *A. flavus* and *A. oryzae* paralogs. The full single gene phylogenies, indicating sequences from *ST/AF* and related clusters [16, 56, 57], are shown in Figure S2.

Aspergillus species (pattern 1 in Figure 2), several other gene phylogenies suggest that the cluster transferred from the *Aspergillus* ancestor or from other *Aspergillus* lineages (patterns 2 and 3; Figure 2).

Interestingly, our gene phylogenies also indicate that six *Mycosphaerella pini* (anamorph: *Dothistroma septosporum*) genes from three gene clusters required for the production of the AF-related compound dothiostromin [22], and which are homologs to *ST* cluster genes, were obtained via HGT. The phylogenetic placement of all six genes is in conflict with the expected position of *M. pini* genes based on the species phylogeny and breaks up the *Aspergillus* sequence clan (Figure S2). This placement suggests that the *M. pini* genes probably originated from the same donor lineage as the *P. anserina* cluster and that HGT events might have been frequent in the evolution of *ST* cluster-related genes.

An alternative to the HGT hypothesis is that the ancestor of *Aspergillus* and *P. anserina* contained a *ST* cluster that was conserved in *A. nidulans* and *P. anserina* but evolved into an *AF* cluster in the *A. flavus* lineage. This "vertical descent" hypothesis is highly unlikely for several reasons. For example, it requires massive losses of clusters and their constituent genes across Sordariomycetes (the 10,564 species-rich class to which *Podospora* belongs [23], including 18 with draft genome sequences; Table S1) and Eurotiomycetes (the 3,401 species-rich class to which *Aspergillus* belongs [23], including >20 with draft genome sequences; Table S1) to explain the presence of clusters only in *Aspergillus* and *Podospora*. Furthermore, the hypothesis has difficulty explaining the remarkable microsynteny and sequence conservation (Figure 1 and Figure S1) between the *P. anserina* and *A. nidulans ST* clusters and the uniquely shared genes between Eurotiomycetes and *P. anserina*, as well as the several gene phylogenies that support an origin of the *P. anserina ST* cluster from within *Aspergillus* (patterns 1 and 3 in Figure 2).

The *A. nidulans ST* cluster contains two regulatory genes, *AflR* and *AflJ* (*AflS* in *A. flavus*) [24], and both of them appear

Table 1. Comparative Topology Tests Provide Support for the HGT of the Podospora *ST* Gene Cluster from *Aspergillus*

Gene	LnL of Best Tree	LnL of Pattern 1	LnL of Pattern 2	LnL of Pattern 3	LnL of Pattern 4
stcW[a]	−11466.24	−11466.24	−11480.12	−11480.13	**−11545.01**
stcV[a]	−21246.35	−21252.62	−21257.02	−21260.71	**−21383.64**
stcR[a,b]	−32492.58	−32494.70	**−32525.28**	**−32522.89**	**−32522.89**
stcE	−23927.73	−23950.01	−23946.01	−23945.90	−23946.33
stcQ	−3338.08	−3349.30	−3338.08	**−3350.06**	−3338.08
aflJ	−31989.07	−31990.44	−31995.39	−31993.92	−31995.43
stcG[a,b]	−16810.37	−16811.97	−16810.37	−16813.50	**−16844.76**
stcI[a]	−11370.46	−11377.12	−11370.46	−11377.23	**−11409.22**
stcS[a,b]	−30703.50	**−30727.49**	−30703.97	**−30733.69**	**−30844.03**
stcL[b]	−23210.18	**−23255.82**	−23210.18	**−23253.95**	−23210.18
stcF[b]	−27471.52	**−27513.84**	−27471.57	**−27513.84**	−27471.57
stcU	−8394.52	**−8432.60**	−8394.52	−8396.83	−8394.51
stcD	−2734.30	−2745.03	−2734.80	−2745.34	−2735.85
stcK	−20214.74	−20231.88	**−20238.69**	−20214.74	−20214.74
stcB[a]	−39230.39	**−39423.43**	**−39422.95**	**−39428.10**	**−39449.51**
stcN[a,b]	−47195.68	**−47316.03**	**−47493.21**	**−47568.19**	**−47486.07**
stcJ[a,b]	−46692.31	−46699.85	−46693.51	−46692.31	**−46922.38**
stcM	−7773.96	−7789.53	−7782.63	−7776.18	−7786.17
stcC	−16353.87	−16354.61	−16355.07	−16354.07	−16370.58
stcT	−16203.34	−16203.34	−16206.21	−16208.20	−16213.14

Topology patterns 1–3 correspond to topology patterns 1–3 in Figure 2 and support HGT, whereas topology pattern 4 is the topology consistent with the species phylogeny and assumes no HGT. Log-likelihood (LnL) scores in bold indicate topology patterns that are significantly worse than the optimal topology. See also Table S2.

[a] Genes that reject the null hypothesis of no HGT.

[b] Genes that recovered the *Aspergillus* + *Podospora* clan in protein Log-Det phylogenetic analyses.

conserved in *P. anserina*. Although relatively little is known about the function of the AflJ protein, AflR is a well-characterized Zn_2Cys_6 transcription factor required for transcriptional activation of the *ST/AF* biosynthetic genes [24–26]. In both *A. nidulans* and *A. flavus*, AflR directly regulates the expression of several pathway genes through binding to the consensus sequence 5′–TCG(N_5)CGA–3′, an 11 bp motif found upstream of several *ST/AF* cluster genes [27, 28](Figure 1B). To test for the presence of putative motifs in the *P. anserina* cluster, we searched all intergenic regions with a de novo DNA-binding motif prediction algorithm (see Experimental Procedures). The most significant palindromic motif identified in our search was 14 instances of the same motif previously identified in *A. nidulans* and *A. flavus* [27, 28], suggesting that AflR might also participate in the regulation of the *P. anserina* pathway. This conservation is surprising given that *cis*-regulatory elements are typically divergent in sequence between fungi belonging to different classes [29] and offers independent support for the HGT hypothesis.

To test whether genes in the *P. anserina ST* cluster are expressed, we searched a collection of 51,862 expressed sequence tags (ESTs) generated from the sequencing of cDNA libraries constructed at seven different stages of the *P. anserina* life cycle [17]. Similarity searches [30] of the cluster proteins against the EST collection identified transcripts for 14 of the 24 cluster genes (including for the *AflR* gene), with genes showing evidence of expression at five of the seven life cycle stages. It is not surprising that we did not detect expression of all *ST* cluster genes in *P. anserina* because different genes in secondary metabolic pathways, including in *ST/AF* pathways, are expressed at different levels [11, 26]—in the *AF* pathway variation in gene expression is also known between different fungal strains [31]—and the small numbers of reads typically sequenced in EST collections mean that they frequently miss detecting genes expressed at lower levels. Thus, the expression of several members of the gene cluster as well as one of their regulators throughout the *Podospora* lifecycle, coupled with the recent observation of ST production in *Podospora* [1], suggests that this horizontally acquired cluster is most likely responsible for ST production in *Podospora*.

Our data suggest that an ~57 kb genomic region containing the entire *ST* gene cluster horizontally transferred from *Aspergillus* to *Podospora* (Figures 1 and 2). Several recent studies have reported transfers of gene clusters across fungi, but all cases so far involve transfers of clusters composed from fewer than half a dozen genes [32–36]. Novo and coworkers [37] recently reported the transfer of a 65 kb region from *Zygosaccharomyces bailii* to a *Saccharomyces cerevisiae* commercial wine strain, whereas Ma and coworkers [38] experimentally showed the transfer of entire chromosomes between *Fusarium* strains, but in both cases the transferred fragments were not metabolic gene clusters.

The ecological settings and molecular mechanisms underlying this HGT event remain obscure. In fungi, coincubation of different isolates can result in the transfer of entire chromosomes [38], suggesting that niche overlap might be sufficient to facilitate HGT events. Like *Aspergillus* [19, 39], *Podospora* species occupy a similar opportunistic saprotroph niche [17, 19], and secondary metabolites appear to be involved in antagonistic interactions (including colonization of *Aspergillus* sclerotia) between these lineages [40, 41], thus providing putative means and motive for HGT of metabolic gene clusters. Transposable elements or other mobile genetic elements such as plasmids and viruses could facilitate chromosomal rearrangement and integration of foreign genetic material [38, 42, 43]. Unfortunately, any evidence for the molecular mechanism responsible for this HGT in the currently available genomes is limited to an *A. nidulans* transposable element sequence (AN7830) found five genes away from the *stcA* end of the *ST* cluster.

Fungi are remarkably diverse in their metabolism, capable of catabolizing a wide variety of substrates as well as producing a wide variety of secondary metabolites. Several of the pathways responsible for these activities form gene clusters. Our

finding that one of the largest metabolic gene clusters moved by HGT between fungi suggests that nonvertical transmission of the numerous metabolic gene clusters present in fungal genomes might have significantly contributed to the remarkable metabolic diversity of fungi, including their ability to produce highly toxic compounds. Finally, the increasing number of reported HGT events between fungi adds support to the notion that HGT-acquired DNA is a significant contributor to fungal genome remodeling [37, 38].

Experimental Procedures

Gene Cluster Identification

We analyzed the distribution, clustering, and phylogenies of all genes found the *A. nidulans* sterigmatocystin (*ST*) gene cluster and the *A. flavus* aflatoxin (*AF*) gene cluster. In microsynteny searches, we defined putative *ST/AF* gene clusters as physically linked groups of genes, no one member of which is more than seven genes away from any gene that is also a member of the *ST/AF* pathway [34]. We inferred homologous *ST/AF* clusters in a sample of 94 fungal genomes when greater than ten constituent genes were microsyntenic (Table S1 and Figure S1), using a series of custom Perl scripts that rely on BLAST similarity searches [30] and gene order data from publicly available genome project assemblies or from http://fungalgenomes.org (maintained by Jason Stajich). We detected the background distribution of microsynteny conservation between *A. nidulans* and *A. flavus/P. anserina* by analyzing one hundred randomly selected sets of 25 contiguous genes (25 gene set) from *A. nidulans*. *A. flavus/P. anserina* orthologs of genes belonging to these 25 gene sets were considered to have conserved microsynteny if they were found clustered with other *A. flavus / P. anserina* orthologs from the same 25 gene set.

Phylogenetic Analysis

We combined fungal protein sequences with all similar sequences from all domains of life retrieved from the GenBank nonredundant database using BLASTP. We identified groups of homologous genes using OrthoMCL [44], treating all sequences as within genome. We then aligned each homologous gene group using MAFFT version 6.624 [45] and manually curated the resulting alignment. We inferred gene phylogenies using maximum likelihood (ML) in RAxML version 7.2 [46] under a Jones-Taylor-Thorton plus GAMMA model and Bayesian analysis in MrBayes version 3.1.2 [47, 48] under mixed protein models. For the ML analyses, we assessed robustness of inference by running 500 bootstrap replicates. For the Bayesian analyses, we ran two independent analyses using four Markov Chain Monte Carlo chains (one cold and three hot) for 1,000,000 generations. We sampled trees every 100 generations and discarded the first 2,000 sampled trees as burn-in, by which point the posterior distribution had already reached stationarity. We performed comparative topology analyses using the Shimodaira-Hasegawa test [49], as implemented in RAxML [46] (Table 1). To exclude the possibility that our inference of HGT was due to compositional artifacts [50], we also performed protein log-determinant (Log-Det) analyses (Table 1 and Table S2). We generated 100 bootstrap replicates and converted them into distance matrices using the LDDist ver. 1.4 alpha 10 software [51], after excluding 25% of the most rapidly evolving sites and 50% of invariant sites and using four rate heterogeneity categories. The resulting 100 distance matrices were analyzed with the neighbor-joining algorithm as implemented in PAUP* 4.0b10 [52]. We then constructed 70% majority rule consensus trees and manually inspected them for retention of clans of interest. Because all our phylogenetic trees are unrooted, we adopted the terminology proposed by Wilkinson and coworkers [53] when discussing tree topology (e.g., clade ~ clan, sister group ~ adjacent group, more closely related ~ split from).

DNA-Binding Motif Prediction

For DNA-binding motif predictions, we used the MEME Suite (version 4.3.0) motif-based sequence analysis tools [54]. We conducted de novo MEME searches for palindromes on all intergenic regions immediately adjacent to *ST/AF* gene open reading frames using motif position-specific scoring matrices to search for similar sequences not found in the initial search using the MAST algorithm (in the MEME Suite).

Expression Analysis

We retrieved the 51,862 *P. anserina* ESTs generated by Espagne et al. [17] from GenBank. The ESTs were generated from seven different lifecycle stages: 26,211 ESTs from mycelium grown for 48 hr, 8,323 from young perithecia of less than 48 hr, 8,007 from perithecia older than 48 hr, 5,785 from ascospores 20 hr after germination trigger, 1,230 from senescent mycelium, 1,171 from incompatible mycelium, and 1,135 from rapamycin-induced mycelium. We considered all cluster proteins that had 100% identity hits in similarity searches against the EST collection expressed at the lifecycle stage the EST hit was generated from.

Supplemental Information

Supplemental Information includes Supplemental Experimental Procedures, two figures, and two tables and can be found with this article online at doi:10.1016/j.cub.2010.12.020.

Acknowledgments

The authors thank all anonymous reviewers for helpful suggestions and Mikael Thollesson for making an updated LDDist script available. 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.).

Received: August 25, 2010
Revised: November 23, 2010
Accepted: December 9, 2010
Published online: December 30, 2010

References

1. Matasyoh, J.C., Dittrich, B., Schueffler, A., and Laatsch, H. (2010). Larvicidal activity of metabolites from the endophytic *Podospora* sp. against the malaria vector *Anopheles gambiae*. Parasitol. Res, in press. Published online October 5, 2010. 10.1007/s00436-010-2098-1.
2. Keeling, P.J. (2009). Functional and ecological impacts of horizontal gene transfer in eukaryotes. Curr. Opin. Genet. Dev. *19*, 613–619.
3. Rolland, T., Neuvéglise, C., Sacerdot, C., and Dujon, B. (2009). Insertion of horizontally transferred genes within conserved syntenic regions of yeast genomes. PLoS ONE *4*, e6515.
4. Marcet-Houben, M., and Gabaldón, T. (2010). Acquisition of prokaryotic genes by fungal genomes. Trends Genet. *26*, 5–8.
5. Richards, T.A., Dacks, J.B., Jenkinson, J.M., Thornton, C.R., and Talbot, N.J. (2006). Evolution of filamentous plant pathogens: Gene exchange across eukaryotic kingdoms. Curr. Biol. *16*, 1857–1864.
6. Reyes-Prieto, A., Moustafa, A., and Bhattacharya, D. (2008). Multiple genes of apparent algal origin suggest ciliates may once have been photosynthetic. Curr. Biol. *18*, 956–962.
7. Walton, J.D. (2000). Horizontal gene transfer and the evolution of secondary metabolite gene clusters in fungi: An hypothesis. Fungal Genet. Biol. *30*, 167–171.
8. Gardiner, D.M., Cozijnsen, A.J., Wilson, L.M., Pedras, M.S., and Howlett, B.J. (2004). The sirodesmin biosynthetic gene cluster of the plant pathogenic fungus *Leptosphaeria maculans*. Mol. Microbiol. *53*, 1307–1318.
9. Kennedy, J., Auclair, K., Kendrew, S.G., Park, C., Vederas, J.C., and Hutchinson, C.R. (1999). Modulation of polyketide synthase activity by accessory proteins during lovastatin biosynthesis. Science *284*, 1368–1372.
10. Proctor, R.H., Brown, D.W., Plattner, R.D., and Desjardins, A.E. (2003). Co-expression of 15 contiguous genes delineates a fumonisin biosynthetic gene cluster in *Gibberella moniliformis*. Fungal Genet. Biol. *38*, 237–249.
11. Brown, D.W., Yu, J.H., Kelkar, H.S., Fernandes, M., Nesbitt, T.C., Keller, N.P., Adams, T.H., and Leonard, T.J. (1996). Twenty-five coregulated transcripts define a sterigmatocystin gene cluster in *Aspergillus nidulans*. Proc. Natl. Acad. Sci. USA *93*, 1418–1422.
12. Keller, N.P., Turner, G., and Bennett, J.W. (2005). Fungal secondary metabolism - from biochemistry to genomics. Nat. Rev. Microbiol. *3*, 937–947.
13. Yu, J., Chang, P.K., Ehrlich, K.C., Cary, J.W., Bhatnagar, D., Cleveland, T.E., Payne, G.A., Linz, J.E., Woloshuk, C.P., and Bennett, J.W. (2004). Clustered pathway genes in aflatoxin biosynthesis. Appl. Environ. Microbiol. *70*, 1253–1262.

14. Bennett, J.W., and Klich, M. (2003). Mycotoxins. Clin. Microbiol. Rev. *16*, 497–516.
15. Yu, J., Bhatnagar, D., and Cleveland, T.E. (2004). Completed sequence of aflatoxin pathway gene cluster in *Aspergillus parasiticus*. FEBS Lett. *564*, 126–130.
16. Carbone, I., Ramirez-Prado, J.H., Jakobek, J.L., and Horn, B.W. (2007). Gene duplication, modularity and adaptation in the evolution of the aflatoxin gene cluster. BMC Evol. Biol. *7*, 111.
17. Espagne, E., Lespinet, O., Malagnac, F., Da Silva, C., Jaillon, O., Porcel, B.M., Couloux, A., Aury, J.M., Ségurens, B., Poulain, J., et al. (2008). The genome sequence of the model ascomycete fungus *Podospora anserina*. Genome Biol. *9*, R77.
18. Hibbett, D.S., Binder, M., Bischoff, J.F., Blackwell, M., Cannon, P.F., Eriksson, O.E., Huhndorf, S., James, T., Kirk, P.M., Lücking, R., et al. (2007). A higher-level phylogenetic classification of the Fungi. Mycol. Res. *111*, 509–547.
19. James, T.Y., Kauff, F., Schoch, C.L., Matheny, P.B., Hofstetter, V., Cox, C.J., Celio, G., Gueidan, C., Fraker, E., Miadlikowska, J., et al. (2006). Reconstructing the early evolution of Fungi using a six-gene phylogeny. Nature *443*, 818–822.
20. Rokas, A., Williams, B.L., King, N., and Carroll, S.B. (2003). Genome-scale approaches to resolving incongruence in molecular phylogenies. Nature *425*, 798–804.
21. Cummings, M.P., Otto, S.P., and Wakeley, J. (1995). Sampling properties of DNA sequence data in phylogenetic analysis. Mol. Biol. Evol. *12*, 814–822.
22. Zhang, S., Schwelm, A., Jin, H.P., Collins, L.J., and Bradshaw, R.E. (2007). A fragmented aflatoxin-like gene cluster in the forest pathogen *Dothistroma septosporum*. Fungal Genet. Biol. *44*, 1342–1354.
23. Kirk, P.M., Cannon, P.F., Minter, D.W., and Stalpers, J.A. (2008). Dictionary of the Fungi, Tenth Edition (Oxon, UK: CABI).
24. Georgianna, D.R., and Payne, G.A. (2009). Genetic regulation of aflatoxin biosynthesis: From gene to genome. Fungal Genet. Biol. *46*, 113–125.
25. Yu, J.H., Butchko, R.A., Fernandes, M., Keller, N.P., Leonard, T.J., and Adams, T.H. (1996). Conservation of structure and function of the aflatoxin regulatory gene aflR from *Aspergillus nidulans* and *A. flavus*. Curr. Genet. *29*, 549–555.
26. Price, M.S., Yu, J., Nierman, W.C., Kim, H.S., Pritchard, B., Jacobus, C.A., Bhatnagar, D., Cleveland, T.E., and Payne, G.A. (2006). The aflatoxin pathway regulator AflR induces gene transcription inside and outside of the aflatoxin biosynthetic cluster. FEMS Microbiol. Lett. *255*, 275–279.
27. Ehrlich, K.C., Montalbano, B.G., and Cary, J.W. (1999). Binding of the C6-zinc cluster protein, AFLR, to the promoters of aflatoxin pathway biosynthesis genes in *Aspergillus parasiticus*. Gene *230*, 249–257.
28. Fernandes, M., Keller, N.P., and Adams, T.H. (1998). Sequence-specific binding by *Aspergillus nidulans* AflR, a C6 zinc cluster protein regulating mycotoxin biosynthesis. Mol. Microbiol. *28*, 1355–1365.
29. Gasch, A.P., Moses, A.M., Chiang, D.Y., Fraser, H.B., Berardini, M., and Eisen, M.B. (2004). Conservation and evolution of cis-regulatory systems in ascomycete fungi. PLoS Biol. *2*, e398.
30. Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. Nucleic Acids Res. *25*, 3389–3402.
31. Rodrigues, P., Venâncio, A., Kozakiewicz, Z., and Lima, N. (2009). A polyphasic approach to the identification of aflatoxigenic and non-aflatoxigenic strains of *Aspergillus* Section Flavi isolated from Portuguese almonds. Int. J. Food Microbiol. *129*, 187–193.
32. Khaldi, N., and Wolfe, K.H. (2008). Elusive origins of the extra genes in *Aspergillus oryzae*. PLoS ONE *3*, e3036.
33. Khaldi, N., Collemare, J., Lebrun, M.H., and Wolfe, K.H. (2008). Evidence for horizontal transfer of a secondary metabolite gene cluster between fungi. Genome Biol. *9*, R18.
34. Slot, J.C., and Rokas, A. (2010). Multiple *GAL* pathway gene clusters evolved independently and by different mechanisms in fungi. Proc. Natl. Acad. Sci. USA *107*, 10136–10141.
35. Slot, J.C., and Hibbett, D.S. (2007). Horizontal transfer of a nitrate assimilation gene cluster and ecological transitions in fungi: A phylogenetic study. PLoS ONE *2*, e1097.
36. Patron, N.J., Waller, R.F., Cozijnsen, A.J., Straney, D.C., Gardiner, D.M., Nierman, W.C., and Howlett, B.J. (2007). Origin and distribution of epipolythiodioxopiperazine (ETP) gene clusters in filamentous ascomycetes. BMC Evol. Biol. *7*, 174.
37. Novo, M., Bigey, F., Beyne, E., Galeote, V., Gavory, F., Mallet, S., Cambon, B., Legras, J.L., Wincker, P., Casaregola, S., and Dequin, S. (2009). Eukaryote-to-eukaryote gene transfer events revealed by the genome sequence of the wine yeast *Saccharomyces cerevisiae* EC1118. Proc. Natl. Acad. Sci. USA *106*, 16333–16338.
38. Ma, L.J., van der Does, H.C., Borkovich, K.A., Coleman, J.J., Daboussi, M.J., Di Pietro, A., Dufresne, M., Freitag, M., Grabherr, M., Henrissat, B., et al. (2010). Comparative genomics reveals mobile pathogenicity chromosomes in *Fusarium*. Nature *464*, 367–373.
39. Wilson, D.M., Mubatanhema, W., and Jurjevic, Z. (2002). Biology and ecology of mycotoxigenic *Aspergillus* species as related to economic and health concerns. Adv. Exp. Med. Biol. *504*, 3–17.
40. Che, Y., Gloer, J.B., and Wicklow, D.T. (2004). Curvicollides A-C: New polyketide-derived lactones from a sclerotium-colonizing isolate of *Podospora curvicolla* (NRRL 25778). Org. Lett. *6*, 1249–1252.
41. Che, Y., Araujo, A.R., Gloer, J.B., Scott, J.A., and Malloch, D. (2005). Communiols E-H: New polyketide metabolites from the coprophilous fungus *Podospora communis*. J. Nat. Prod. *68*, 435–438.
42. Schaack, S., Gilbert, C., and Feschotte, C. (2010). Promiscuous DNA: Horizontal transfer of transposable elements and why it matters for eukaryotic evolution. Trends Ecol. Evol. (Amst.) *25*, 537–546.
43. Liu, H., Fu, Y., Jiang, D., Li, G., Xie, J., Cheng, J., Peng, Y., Ghabrial, S.A., and Yi, X. (2010). Widespread horizontal gene transfer from double-stranded RNA viruses to eukaryotic nuclear genomes. J. Virol. *84*, 11876–11887.
44. Li, L., Stoeckert, C.J., Jr., and Roos, D.S. (2003). OrthoMCL: Identification of ortholog groups for eukaryotic genomes. Genome Res. *13*, 2178–2189.
45. Katoh, K., and Toh, H. (2008). Recent developments in the MAFFT multiple sequence alignment program. Brief. Bioinform. *9*, 286–298.
46. Stamatakis, A. (2006). RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics *22*, 2688–2690.
47. Altekar, G., Dwarkadas, S., Huelsenbeck, J.P., and Ronquist, F. (2004). Parallel Metropolis coupled Markov chain Monte Carlo for Bayesian phylogenetic inference. Bioinformatics *20*, 407–415.
48. Ronquist, F., and Huelsenbeck, J.P. (2003). MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics *19*, 1572–1574.
49. Shimodaira, H., and Hasegawa, M. (1999). Multiple comparisons of log-likelihoods with applications to phylogenetic inference. Mol. Biol. Evol. *16*, 1114–1116.
50. Foster, P.G., and Hickey, D.A. (1999). Compositional bias may affect both DNA-based and protein-based phylogenetic reconstructions. J. Mol. Evol. *48*, 284–290.
51. Thollesson, M. (2004). LDDist: A Perl module for calculating LogDet pairwise distances for protein and nucleotide sequences. Bioinformatics *20*, 416–418.
52. Swofford, D.L. (2002). PAUP*: Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4.0b10 Edition (Sunderland, MA: Sinauer).
53. Wilkinson, M., McInerney, J.O., Hirt, R.P., Foster, P.G., and Embley, T.M. (2007). Of clades and clans: Terms for phylogenetic relationships in unrooted trees. Trends Ecol. Evol. *22*, 114–115.
54. Bailey, T.L., and Elkan, C. (1994). Fitting a mixture model by expectation maximization to discover motifs in biopolymers. Proc. Int. Conf. Intell. Syst. Mol. Biol. *2*, 28–36.
55. Schneider, T.D., and Stephens, R.M. (1990). Sequence logos: A new way to display consensus sequences. Nucleic Acids Res. *18*, 6097–6100.
56. Chiang, Y.M., Szewczyk, E., Davidson, A.D., Entwistle, R., Keller, N.P., Wang, C.C., and Oakley, B.R. (2010). Characterization of the *Aspergillus nidulans* monodictyphenone gene cluster. Appl. Environ. Microbiol. *76*, 2067–2074.
57. Bok, J.W., Chiang, Y.M., Szewczyk, E., Reyes-Dominguez, Y., Davidson, A.D., Sanchez, J.F., Lo, H.C., Watanabe, K., Strauss, J., Oakley, B.R., et al. (2009). Chromatin-level regulation of biosynthetic gene clusters. Nat. Chem. Biol. *5*, 462–464.