

Evolution of Filamentous Plant Pathogens: Gene Exchange across Eukaryotic Kingdoms

Thomas A. Richards,¹ Joel B. Dacks,²
Joanna M. Jenkinson,¹ Christopher R. Thornton,¹
and Nicholas J. Talbot^{1,*}

¹School of Biosciences

University of Exeter
Geoffrey Pope Building
Exeter EX4 4QD
United Kingdom

²Department of Biological Sciences

University of Calgary
2500 University Drive NW
Calgary, Alberta T2N 1N4
Canada

Summary

Filamentous fungi and oomycetes are eukaryotic microorganisms that grow by producing networks of thread-like hyphae, which secrete enzymes to break down complex nutrients, such as wood and plant material, and recover the resulting simple sugars and amino acids by osmotrophy. These organisms are extremely similar in both appearance and lifestyle [1] and include some of the most economically important plant pathogens [2, 3]. However, the morphological similarity of fungi and oomycetes is misleading because they represent some of the most distantly related eukaryote evolutionary groupings, and their shared osmotrophic growth habit is interpreted as being the result of convergent evolution [3–5]. The fungi branch with the animals, whereas the oomycetes branch with photosynthetic algae as part of the Chromalveolata [6–10]. In this report, we provide strong phylogenetic evidence that multiple horizontal gene transfers (HGT) have occurred from filamentous ascomycete fungi to the distantly related oomycetes. We also present evidence that a subset of the associated gene families was initially the product of prokaryote-to-fungi HGT. The predicted functions of the gene products associated with fungi-to-oomycete HGT suggest that this process has played a significant role in the evolution of the osmotrophic, filamentous lifestyle on two separate branches of the eukaryote tree.

Results and Discussion

Multiple HGTs Have Occurred between Fungi and Oomycetes

Comparative analysis of a large number of microbial genome sequences has begun to reveal the extent and evolutionary significance of HGT among prokaryotic species and between prokaryotes and eukaryotes [11–15]. The importance of HGT among eukaryotic species is, however, far less clear. We set out to explore the

evolutionary history of the 11,109 predicted genes in the genome of filamentous ascomycete plant pathogenic fungus *Magnaporthe grisea* [16], the causal agent of rice-blast disease. During our analyses, we detected 11 *M. grisea* genes that had a significantly higher level of sequence similarity (shown by BLASTp) to sequences from the oomycete genus *Phytophthora* than to any fungal sequences used in the primary genome-comparison analysis (see Table S1 in the Supplemental Data available online). These results are contrary to predicted gene similarities given the number of evolutionary branches that exists between fungi and oomycetes (Figure 1) and therefore suggest the possibility of HGT. To explore this idea, we carried out phylogenetic analysis, which revealed that for four of the 11 candidate HGT genes, the oomycete sequence was clearly within a clade of fungal gene sequences, branching with the filamentous ascomycetes (Figure 1). These specific phylogenetic relationships were consistently supported by at least one node with high posterior probabilities in Bayesian analysis and, importantly, by two distinct bootstrap methods (PHYML and ML distance, with 1000 replicates) with support values in excess of 85% (Figures 2A, 3A, 4A, and 4C; see also Figures S1A, S1B, S2A, and S2B). Because bootstrapping is generally considered to be a more conservative indicator of phylogenetic resolution [17], these values confirmed that the relationships were particularly robust. In the case of two of the potential HGTs (*AraJ* and *CodB*), we repeated the phylogenetic experiments by using alignments with distantly related genes removed and altered character sampling to exclude long-branch and outgroup attraction problems (Figures 2A and 3A), but we consistently recovered topologies where the oomycete sequences were specifically embedded within a clade of the fungi as a sister branch to the filamentous ascomycetes (fungal and oomycete paraphyly). Such a phylogenetic pattern is strongly indicative of HGT. To pinpoint the branching position of the oomycetes within the fungi radiation, we then performed additional phylogenetic analyses that focused on increased fungal sampling and reduced the outgroup being sampled. This allowed us to confirm sisterhood of the oomycete and the filamentous ascomycete sequences (Figures 2B, 3C, 4A, and 4C). Figure 1 summarizes this pattern of HGT and shows the most likely branching position of the fungi and the oomycetes in the eukaryotic tree.

Osmotrophy-Related Gene Functions Are Predicted among the Fungi-to-Oomycete HGT Candidates

Of the four strongly supported candidate HGTs, the first gene putatively encodes a sugar transporter (PFAM classification—pfam00083) of the multifacilitator superfamily. This sugar transporter possesses an *AraJ* arabinose permease-like domain (COG2814) based on interrogation of the conserved domain database (CDD) [18]. The transfer of a multifacilitator sugar-transporter-encoding gene could potentially increase the accessibility of sugar

*Correspondence: n.j.talbot@exeter.ac.uk

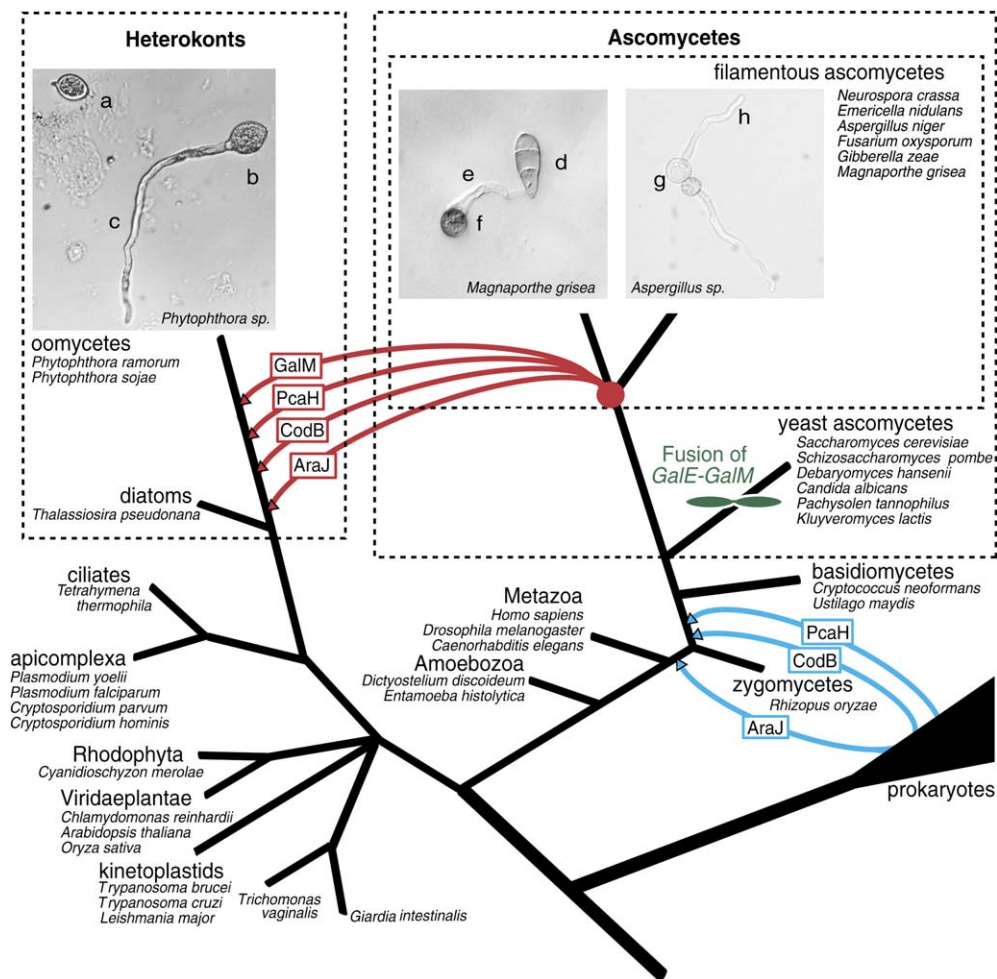


Figure 1. Schematic Representation of Eukaryotic Phylogeny, Indicating Oomycete and Fungi HGT Events and Demonstrating Superficial Morphological Similarities between the Fungi and the Oomycetes

The tree topology shown serves as a hypothesis of eukaryotic tree topology for contrasting with the gene phylogenies reported here. Only eukaryotic groups with genome project representation, those that were searched during this study, are placed within the tree. Candidate prokaryote-to-fungi HGTs are shown in blue, and candidate fungi-to-oomycete HGTs are shown in red. The *GalE-GalM* fusion gene used to root the ascomycetes is indicated (see Figure 4E). Pictures show typical and similar filamentous characters exhibited by both oomycetes and fungi: Lowercase letters a and b indicate oomycete sporangia that can germinate to produce motile zoospores or to form germ tubes as indicated by lowercase letter c. The letter d indicates asexual conidium; e indicates germ tube; f indicates specialized infection cell known as an appressorium; and g indicates asexual conidia that have germinated to produce germ tubes (marked by h).

substrates to an osmotrophic microorganism. Phylogeny of the *AraJ* HGT was supported by three nodes (1/60/90%, 1/100/98%, and 1/100/100% support), which specifically grouped the oomycetes within the fungi and the ascomycete radiation (Figure 2A). The three phylogeny support values are listed, here and subsequently, in the order Bayesian posterior probability, % PHYML bootstrap value, and % ML-distance bootstrap value.

The second HGT candidate putatively encodes a permease protein containing a *CodB* cytosine/purine, uracil/thiamine/allantoin permease domain, identified with CDD [18] (COG1457). This gene phylogeny also demonstrated a HGT event from the filamentous ascomycetes to the oomycetes (resolved with 1/100/100% and 0.88/80/78% phylogeny support values—Figure 3A and 0.99/85/87% phylogeny support values—Figure 3C). The budding yeast *Saccharomyces cerevisiae* *CodB* gene encodes a broad specificity permease for purine uptake [19]. Thus, acquisition of *CodB* represents

a potential means by which oomycetes could access nucleotide substrates.

The third gene reported putatively encodes a protocatechuate 3,4-dioxygenase β -subunit (3,4-PCD) annotated as *PcaH* (COG3485) in CDD. Phylogenetic analysis of the homologs of the *PcaH* gene family demonstrated tree topologies consistent with fungi-to-oomycete HGT and with 1/94/86% support values (Figure 4A). *PcaH* encodes an enzyme involved in degradation of aromatic compounds as part of the β -ketoadipate pathway [20].

Finally, phylogenetic analysis of the aldose-1-epimerase (*GalM*) gene family (COG2017), demonstrated *Phytophthora ramorum* and *P. sojae* sequences grouping within the fungi (1/99/100% support), specifically as a sister group to the filamentous ascomycete *GalM* homologs (1/100/100% support—Figure 4C). The *GalM*-encoded aldose-1-epimerases can demonstrate broad substrate specificity [21] and are present in an evolutionary diverse selection of eukaryotes. The *GalM* protein of

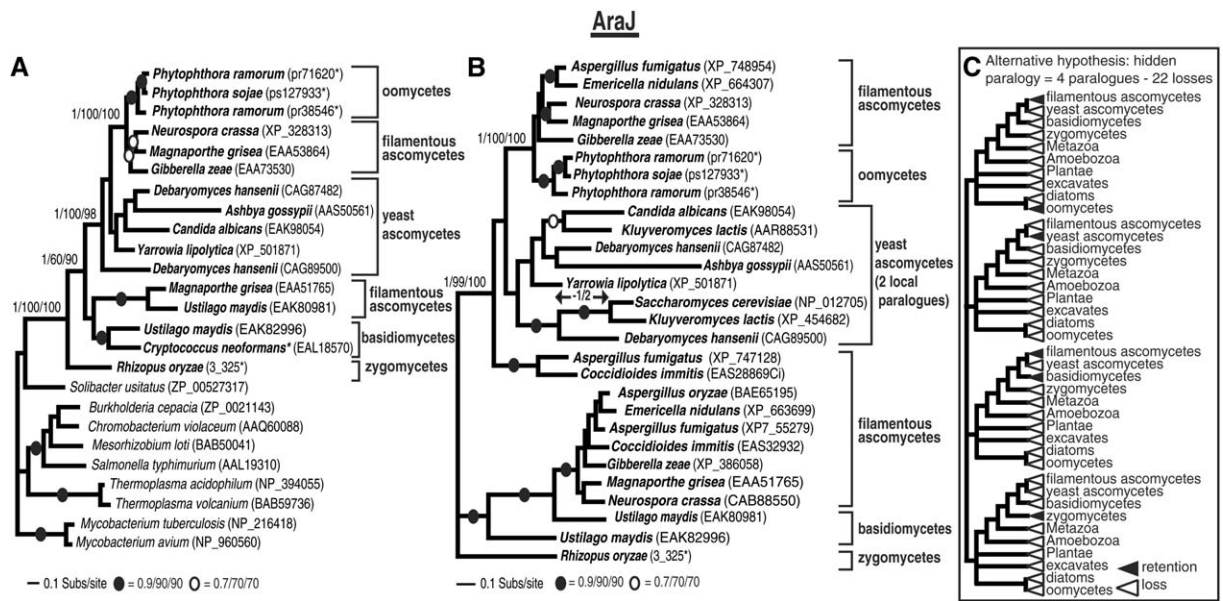


Figure 2. Phylogenies of the Putative Sugar-Transporter Gene *AraJ*

(A) Phylogeny focusing on the anciently derived ortholog set defined by previous analyses and including an outgroup of prokaryote orthologs (the analysis of the wider gene family justifying sequence and paralog exclusion is reported in Figure S1A). Sampling = 25 genes and 305 characters. This phylogeny is arbitrarily rooted on *Mycobacterium*.

(B) Phylogeny with reduced taxon sampling and increased character sampling focusing on the diversity of fungi and oomycete genes. Sampling = 28 genes and 360 characters. This phylogeny is rooted on the zygomycetes. 1/2 branch reduction is labeled on the relevant branch. Eukaryote taxa are in bold, and higher taxonomic groupings are labeled. GenBank accession numbers or genome project identifier numbers are given for each sequence included. All trees shown are Bayesian-consensus phylogenies. Support values (Bayesian posterior probability/1000 ML-FITCH-distance bootstraps/1000 PHYML bootstraps) are marked if bootstrap values are above 90% (shaded circle) or if all are above 70% (open circle). Key nodes for HGT hypothesis discussed in the manuscript are labeled with actual values.

(C) Demonstrates pattern of paralog distribution and loss required for alternative hypothesis of hidden paralogy rather than HGT.

Escherichia coli has been demonstrated to possess mutarotase activity, which converts α -aldose to the β -anomer and is a key step for efficient lactose metabolism [22].

Testing Alternative Hypotheses to HGT

Because of the surprising nature of our results, we carried out alternative topology comparison tests, which were specifically designed to test the robustness of the fungal and oomycete paraphyly. Alternative topologies with monophyly of the fungal orthologs were rejected at the 5% significance level with both the Shimodaira-Hasegawa/weighted Shimodaira-Hasegawa and approximately unbiased tests in two of the four cases (see Figure S4), again providing additional support for the HGT hypothesis in these two instances. In the other two HGTs, topology comparison tests did not reject monophyly of the specific fungal ortholog sets (Figures 3 and 4A), although in the case of the *PcaH* phylogeny, fungi monophyly was only very narrowly accepted with the more appropriate approximately unbiased test (0.051 versus the cutoff of 0.05). However, in these two cases, where the alternative hypothesis could not be rejected, however, a zygomycete ortholog was not detected. Consequently, the topology constraint required a less radical topology alteration than in the other two cases. In view of the fact that the oomycete genes nested within the fungal clade with bootstrap values in excess of 85% and because no other eukaryotic

homologs could be detected, we still favor HGT in these two cases.

Hidden paralogy—unidentified ancestral gene duplications with subsequent gene loss—is often identified as an alternative hypothesis to HGT. We therefore used our phylogenetic trees to calculate the minimum number of paralogs and losses required to explain the tree topologies (see Figures 2C, 3B, 4B, and 4D). Only nodes with bootstrap support of 85% or more were used to infer patterns of paralogy. These comparative analyses demonstrated alternative gene-evolution patterns ranging from two paralogs (one gene duplication) with 11 gene losses to four paralogs (three duplications) and 22 losses. How to compare the likelihood of hidden paralogy events to HGT events is currently unknown. However, by using strict parsimony criteria and assuming that all evolutionary events are equally likely, we find that in each case, the HGT hypothesis is much more parsimonious than that of hidden paralogy (one HGT versus one gene duplication event and 11 subsequent gene losses is the nearest scoring scenario). Hence, we favor the HGT scenario but also accept the unlikely, though plausible, alternative evolutionary history of multiple cases of hidden paralogy. Furthermore, we have limited the examples of HGT to phylogenies showing paraphyly of the fungi with strong bootstrap support (by selecting only four candidates from Table S1). This is because the number of gene-duplication and -loss events required is more complex than for HGT and because hidden paralogy requires invoking an evolutionary trend where the

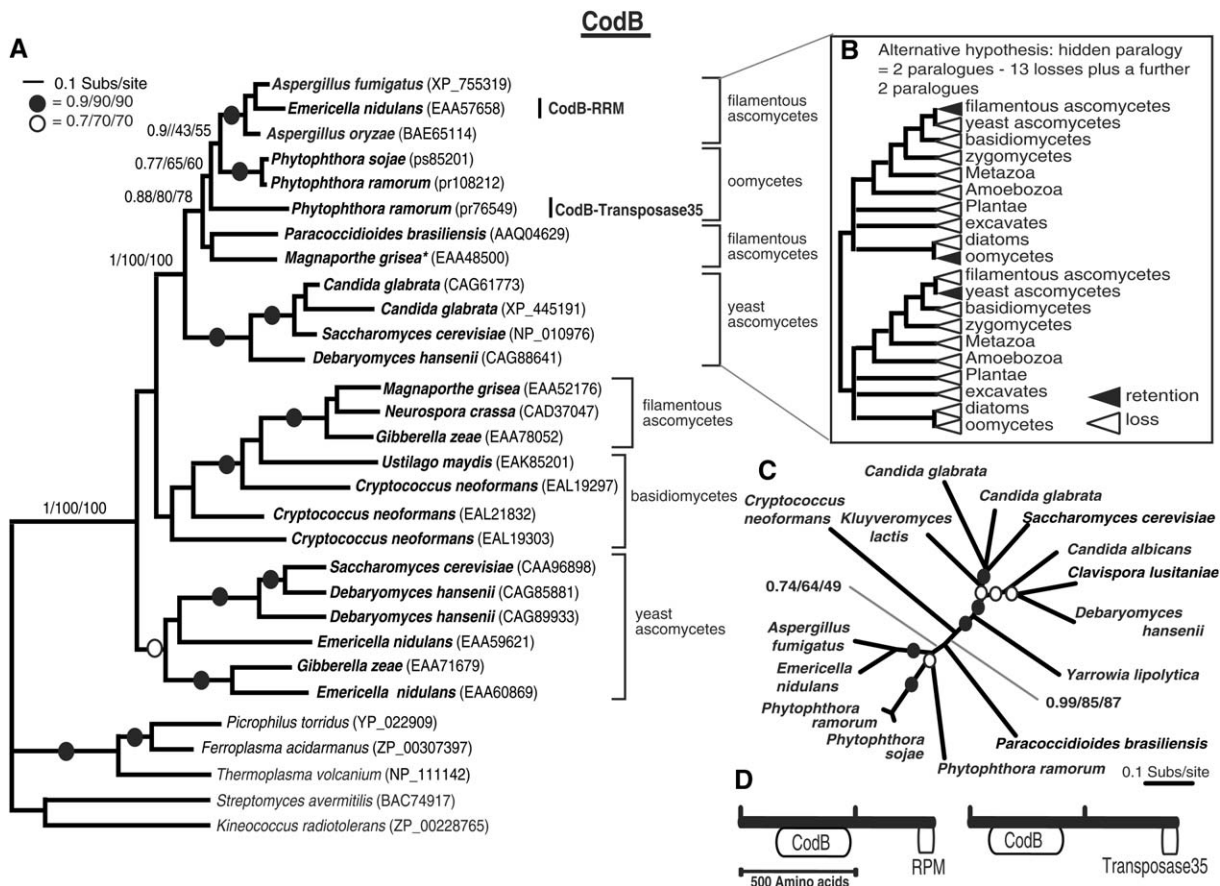


Figure 3. Phylogenies of Purine Permease, *CodB*, Gene-Family Phylogeny

(A) Phylogeny of purine permease gene. The focus is on the related orthologs, and an outgroup of prokaryotes is included (the analysis of the wider gene family justifying sequence exclusion is reported in Figure S1B). Sampling = 30 genes and 365 characters. Phylogeny is rooted arbitrarily on the prokaryotes included. Different protein domain structures are labeled.

(B) Demonstrates pattern of paralog distribution and loss required for alternative hypothesis of hidden paralogy rather than HGT.

(C) Phylogeny of purine permease gene (*CodB*) focusing on the diversity of fungi and oomycete genes. Sampling = 15 genes and 362 characters. This phylogeny is unrooted. Key nodes for HGT hypothesis discussed in the manuscript are labeled with actual values. Note also that *Magnaporthe* (Figure 3A), which is partial, is removed from this third phylogeny (Figure 3C). The phylogenies are illustrated as described in Figure 2.

(D) The predicted conserved domain structure of fusion genes is illustrated according to CDD results [18]. Distributions of fusion genes are illustrated in Figure 3A. Note the independent acquisition of gene fusions of *CodB* with an RRM domain and with Transposase35 in *Emericella nidulans* and *Phytophthora ramorum*, respectively.

eukaryotic ancestor possessed numerous paralogs, and convergent patterns of gene loss have occurred among all other eukaryotic lineages sampled for these four gene families. In addition, in the four cases reported, the paraphyly of fungi and oomycetes is interrupted in a relatively recent ascomycete evolutionary branch (Figure 1), indicating that such a pattern of gene loss would have had to occur convergently in recent evolutionary branches. This would imply a long-standing maintenance of numerous paralogs followed by recent large-scale and convergent patterns of gene loss across all four-gene families; hence, we favor the more parsimonious HGT scenario.

Investigating the Prevalence of Fungi-to-Oomycete HGT

Our original analysis revealed a total of 11 *M. grisea* genes with higher BLASTp scores to an oomycete ortholog than to other fungal relatives (Table S1). BLAST values such as these have been used on their own in

the past as evidence for HGT. In this study, we chose to use more stringent criteria based on strongly supported branching relationships that showed oomycete genes grouping within the fungal clade in phylogenetic trees. Although four cases were able to be resolved as such (Figures 2–4), five additional candidates were excluded from our analysis because four did not show fungi and oomycete paraphyly and an additional gene encoded a proline-rich-repeat protein, which was not amenable to meaningful phylogenetic analyses (Table S1). Phylogenetic analysis of two additional gene families, encoding putative esterase/lipase and aconitase enzymes, respectively, demonstrated four additional putative fungal-oomycete HGTs with weak support or with a different pattern of fungi-oomycete transfer from that discussed above (see Figures S3A and S3B). Nonetheless, although we do not strongly advocate these additional phylogenies as examples of HGT, when considered together, the eight possible transfers are suggestive of a more pervasive pattern of HGT

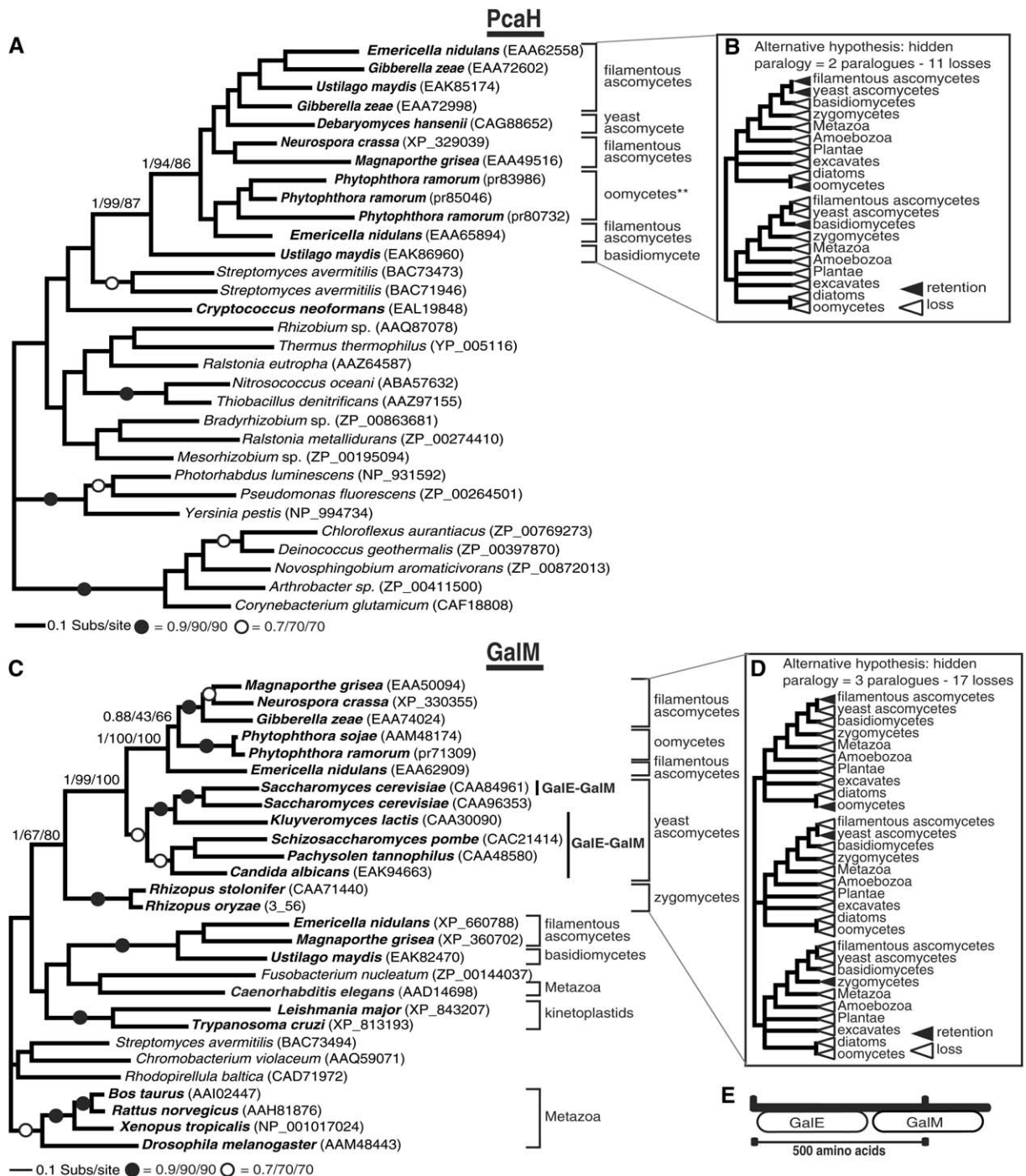


Figure 4. Phylogenies of *PcaH* and *GalM*

(A) Phylogeny of extracellular dioxygenase/Protocatechuate 3,4-dioxygenase β -subunit-encoding genes (*PcaH*) focusing on anciently derived ortholog set. Sampling = 31 genes and 142 characters. Key nodes for HGT hypothesis discussed in the manuscript are labeled with actual values. This analysis includes an outgroup of prokaryotes (the analysis of the wider gene family justifying sequence exclusions is reported in Figure S2A). Furthermore, a phylogenetic analysis focusing on the diversity of fungi and oomycete genes did not increase resolution between the fungi and oomycetes (sampling = 37 genes and 151 characters)—data not shown. The phylogenies are illustrated as described in Figure 2.

(B) Demonstrates pattern of paralog distribution and loss required for alternative hypothesis of hidden paralogy rather than HGT.

(C) Phylogeny of aldose 1-epimerase (*GalM*) genes (the analysis of the wider gene family justifying sequence exclusions is reported in Figure S2B). Sampling = 28 genes and 239 characters. Phylogeny is rooted arbitrarily on a prokaryote and metazoan polytomy. The phylogenies are illustrated, as described in Figure 2. Key nodes for HGT hypothesis discussed in the manuscript are labeled with actual values.

(D) Demonstrates pattern of paralog distribution and loss required for alternative hypothesis of hidden paralogy rather than HGT.

(E) Predicted conserved domain structure of *GalE-GalM* fusion genes is illustrated according to CDD [18] results, and distribution of this fusion gene supports the monophyly of the yeast ascomycetes including *S. pombe* (Figure 1). Distributions of fusion genes are illustrated on Figure 4C.

between the fungi and the oomycetes. As further genome sequence information from both fungi and chromalveolates becomes available, it will be interesting to test how prevalent HGT events have been between these organisms.

Determining the Source of Fungi-to-Oomycete HGTs

Phylogenetic investigation demonstrated strong bootstrap support for clustering of the filamentous ascomycetes with the oomycetes to the exclusion of all other fungal genes sampled in three of the four HGTs analyzed (Figures 2, 3, and 4C). In the remaining dataset (Figure 4A), the phylogenies also showed weak support for the same relationship. In the three investigations where the filamentous ascomycetes grouped strongly with the oomycetes, we only recovered weak to moderate support for branching relationships among the following: (1) the *Aspergillus* and *Emericella* branch, (2) the *Neurospora* and *Magnaporthe* branch, (3) the *Gibberella* branch, and (4) the *Phytophthora* branch. We interpret this reduced level of support among these branches as a situation similar to a hard polytomy, i.e., the genes transferred from the fungi to the oomycetes originated from an organism, or organisms, that branched very close to the bifurcation events between the *Aspergillus* and *Emericella*, *Gibberella*, and the *Neurospora* and *Magnaporthe* branch. These observations also raise the possibility that some or all of the four gene transfers could have originated from the same filamentous ascomycete donor lineage.

We also noted that for three of the four HGTs, we were unable to identify an ortholog in any other eukaryote genome sampled. In the *AraJ* gene family, for example, the oomycete and fungi sequences grouped with prokaryotic sequences supported by bootstrap values above 90% (Figure S1A). Similarly, in the cases of *CodB* and *PcaH*, no additional putative eukaryote homologs could be detected. This pattern of gene distribution raises the possibility that three of the fungal genes investigated were initially acquired by a previous prokaryote-to-fungi HGT. Such a scenario seems more likely than vertical inheritance, coupled with multiple polyphyletic losses, throughout the eukaryote tree. The gene homolog distribution data (summarized in Figures 2–4; see also Figures S1 and S2) also suggests that the *CodB* and the *PcaH* genes were transferred prior to the basidiomycete and ascomycete bifurcation, whereas the *AraJ* gene may have been transferred prior to the earlier bifurcation of the zygomycetes in the fungal evolutionary tree (Figure 1). These observations also suggest that HGT may have affected the evolutionary history of the fungi and therefore both groups of filamentous microbial eukaryotes (Figure 1).

HGT in the Absence of Phagotrophy

It has been suggested that frequent HGTs from prokaryotes to eukaryotes are consequences of phagotrophic lifestyles in which bacteria are engulfed and consumed by eukaryotic cells. This has been postulated as a mechanism that has led to a continual flow of genetic material under a gene-transfer ratchet from the genomes of consumed cells to phagocytic (eukaryotic) cells [23]. However, filamentous eukaryotes, such as fungi and oomycetes, feed by absorption (osmotrophy), and so the

implication of the “you-are-what-you-eat” hypothesis, described above, is that cases of HGT should be minimized in such organisms. It is, however, also clear that both fungi and oomycetes consume plant material and are often found in the same ecological niche [3]. We tentatively suggest that the gene transfers reported here provide evidence for a close ecological association that has not previously been recognized between fungi and oomycetes, one which has provided the opportunity for frequent, or more likely a large-scale, gene- or genome-transfer event. The mechanism of HGT between fungi and oomycetes is not yet clear, but anastomosis of mycelia, transduction via mycoviruses, or propagation of retrotransposons could each facilitate HGT in the absence of phagocytosis. We note that analyses of ascomycete genomes did not demonstrate high-linkage disequilibrium between the HGT candidates and transposable elements (Table S3), although a *P. ramorum* *CodB* domain is fused to a transposase domain, as shown in Figure 3D.

Conclusion

This study provides strong evidence for the occurrence of HGT between fungi and oomycetes. The transferred genes identified are likely to expand the range of growth substrates available to an osmotrophic microorganism. *PcaH*, for example, is a key enzyme in the β -ketoadipate pathway found in many soil bacteria and fungi, providing a means of utilizing aromatic compounds such as lignin derivatives, coumarate, and salicylate [20, 23]. In addition to this extracellular enzyme, the reported HGT events have led to propagation of two permeases/transporter proteins. The transfer of such genes between fungi and oomycetes is also likely to have been advantageous to an osmotrophic organism. Consistent with other analyses [14], our findings here have demonstrated that intraeukaryotic HGT does occur. This should encourage caution during interpretation of eukaryote species’ relationships based on single-gene analyses in the absence of other data [10, 24, 25]. However, the HGTs may in part explain the long-standing quandary relating to the convergent evolution of osmotrophy and the filamentous growth habit of two disparate and unrelated eukaryote lineages.

Experimental Procedures

Detection of Candidate Fungi-Oomycete HGT

Initial BLASTp searches surveyed the *Magnaporthe grisea* genome against 21 genomes, representing a diverse sampling of taxa scattered across the eukaryotic evolutionary tree (see Table S1 for genomes included in the first analyses and Figure 1 for the distribution of the taxa across the eukaryote evolutionary tree—note that for later phylogenetic analyses, every genome represented on Figure 1 was included in the analyses where putative homolog sequences were available). The genome sampling for the initial BLASTp searches also included several prokaryote genomes representing the Archaea and Eubacteria. Eubacterial genomes related to the progenitor of the plastid and the mitochondrial organelles [26] were also included (see Table S1 for more details). A putative HGT event was identified for further investigation when the BLASTp score for the *Phytophthora* candidate match was equal to or higher than the corresponding match from a basidiomycete.

Alignment and Phylogenetic Analyses

Alignments were made for all 11 candidate HGTs with T-COFFEE [27]. The alignments were refined manually with SE-AL [28]. Taxon

sampling includes homologs from a wide diversity of prokaryotic and eukaryotic taxa (minimum eukaryotic genome sampling, where homologs were present, is shown on Figure 1). Genes were sampled with BLASTp and tBLASTn searches of the GenBank nr database and from finished and unfinished genome projects listed at The Institute for Genomic Research website (<http://www.tigr.org>), The Department of Energy Joint Genome Institute (<http://www.jgi.doe.gov>), and the *C. merolae* genome website (<http://merolae.biol.s.u-tokyo.ac.jp/>). Additional BLASTp searches were conducted with divergent sequences recovered during the first search. This multiple-step BLASTp approach was used to ensure that the gene family investigated was appropriately sampled. After automated and then manual alignment, nonhomologous insertions and sequence characters that could not be aligned with confidence were removed from the alignments. Edited alignments were analyzed with MODELGENERATOR [29] to find the most appropriate model for phylogenetic analyses for each dataset (Table S2). All alignments are available from the corresponding author upon request.

The results of the MODELGENERATOR analyses were implemented in MRBAYES 3.1.2 [30] with a Γ distribution (\pm eight categories \pm one invariant; see Table S2). MRBAYES 3.1.2 [30] was run with two separate MCMCMC analyses for 1,000,000 generations at a sampling frequency of 100 generations. Each MCMCMC run had four MCMC chains (three heated and one cold; heat parameter = 0.2). Comparisons of likelihood score and model parameter values and topologies within and between the two independent runs for each of the 11 analyses confirmed that the tree log-likelihood scores and parameters had reached a plateau and converged by 150,000 generations at the latest (most analyses reached a plateau far below this value). Consequently, a maximum of 1500 samples were excluded as a burnin, the remaining generations were sampled, and a tree with branch lengths was calculated.

Maximum-likelihood-distance bootstrap values (from 1000 replicates) were obtained with Tree-Puzzle 5.1 [31] for parameter estimation (substitution model, eight multivariant + invariant sites, or only eight multivariant dependant on MODELGENERATOR analyses; Table S2) and in coordination with Puzzleboot [32] to obtain distance matrices. Programs from the PHYLIP package [33] were used to create pseudo-replicate datasets (SEQBOOT), calculate distance trees (FITCH—3 \times jumbling with global rearrangements), and assemble a bootstrap-consensus tree (CONSENSE). In addition, 1000 fast ML (PHYML) [34] bootstrap replicates were run for each alignment, with the model selected as before (Table S2). Results are shown in Figures S1, S2, and S3.

In two cases, the initial phylogenies (Figures S1A and S1B) included sampling from distantly related orthologs or paralogs, or both. To test our putative HGTs further, we repeated the phylogenetic methods as before but reduced sequence sampling to focus on a relevant subsection of the gene family that demonstrated the HGT. This was performed to reduce the possible affects of long-branch attraction problems [35], which may have been caused by inclusion of distantly related genes in the initial phylogenies. In addition, all four phylogenies were repeated with refined and reduced taxon sampling and in some cases, increased character sampling (e.g., *GalM* Figures S2B–S4B) to test the precise branching position of the oomycetes with respect to the ascomycete groups. Note that in all these cases (Figures 2–4), the diatom—theoretically the closest group sampled relative to the oomycetes—did not possess a similar gene that grouped closely to the fungi and oomycete clade and so was not included in these subanalyses. Therefore, these second-round analyses did not include the *AraJ* and *GalE* diatom sequences, which grouped separately from the oomycetes with multiple nodes supported by bootstrap support in excess of 90%. Because these datasets included significantly reduced taxa sampling, the analyses were conducted as before, but for the MRBAYES analyses, the covarian option was selected and the MCMCMC sampling was conducted for 500,000 generations and with a burnin sampling that did not exceed 500 generations for any of the four cases.

By using only nodes significantly supported with a 85% bootstrap-support value, we inferred the minimum number of paralogs and polyphyletic loss events required to explain the putative HGT by hidden paralogy, the alternative hypothesis (Figures 2C, 3B, 4B, and 4D). Figure S4 shows the methods and results of the alternative topology tests.

Supplemental Data

Supplemental Data include three tables and four figures and can be found with this article online at <http://www.current-biology.com/cgi/content/full/16/18/1857/DC1/>.

Acknowledgments

Preliminary sequence data were obtained from The Institute for Genomic Research website (<http://www.tigr.org>) and the Department of Energy Joint Genome Institute website (<http://www.jgi.doe.gov>). We thank both institutes for making data publicly available. The work was supported by a Department for Environmental Food and Rural Affairs grant to N.J.T. and C.R.T. A Wellcome Trust/Canadian Institutes of Health Research Travelling Research Fellowship supported J.B.D.

Received: May 11, 2006

Revised: July 4, 2006

Accepted: July 18, 2006

Published: September 18, 2006

References

- Berkeley, M.J. (1846). Observations, botanical and physiological, on the potato murrain. *Horticultural Society Journal* 1, 9–34.
- Hayden, K.J., Rizzo, D., Tse, J., and Garbelotto, M. (2004). Detection and quantification of *Phytophthora ramorum* from California forests using a real-time polymerase chain reaction assay. *Phytopathology* 94, 1075–1083.
- Latijnhouwers, M., de Wit, P.J.G.M., and Govers, F. (2003). Oomycetes and fungi: Similar weaponry to attack plants. *Trends Microbiol.* 11, 462–469.
- Simpson, A.G., and Roger, A.J. (2004). The real “kingdoms” of eukaryotes. *Curr. Biol.* 14, R693–R696.
- Money, N.P., Davis, C.M., and Ravishanker, J.P. (2004). Biomechanical evidence for convergent evolution of invasive growth process among fungi and oomycete water molds. *Fungal Genet. Biol.* 41, 872–876.
- Patterson, D.J. (1989). Stramenopiles: Chromophytes from a protistan perspective. In *The Chromophyte Algae, Problems and Perspectives*, J.C. Green, B.S.C. Leadbeater, and W.L. Diver, eds. (Oxford: Clarendon), pp. 357–379.
- Leedale, G.F. (1967). *Euglenoid Flagellates* (Englewood Cliffs, NJ: Prentice-Hall).
- Gundersen, J.H., Elwood, H., Ingold, A., Kindle, K., and Sogin, M.L. (1987). Phylogenetic relationships between chlorophytes, chrysophytes and oomycetes. *Proc. Natl. Acad. Sci. USA* 84, 5823–5827.
- Lang, B.F., O’Kelly, C., Nerad, T., Gray, M.W., and Burger, G. (2002). The closest unicellular relatives of animals. *Curr. Biol.* 12, 1773–1778.
- Baldauf, S.L., Roger, A.J., Wenk-Siefert, I., and Doolittle, W.F. (2000). A kingdom-level phylogeny of eukaryotes based on combined protein data. *Science* 290, 972–977.
- Richards, T.A., Hirt, R.P., Williams, B.A., and Embley, T.M. (2003). Horizontal gene transfer and the evolution of parasitic protozoa. *Protist* 154, 17–32.
- Loftus, B., Anderson, I., Davies, R., Alsmark, U.C., Samuelson, J., Amedeo, P., Roncaglia, P., Berriman, M., Hirt, R.P., Mann, B.J., et al. (2005). The genome of the protist parasite *Entamoeba histolytica*. *Nature* 433, 865–868.
- Andersson, J.O., Sjogren, A.M., Davis, L.A.M., Embley, T.M., and Roger, A.J. (2003). Phylogenetic analyses of diplomonad genes reveal frequent lateral gene transfers affecting eukaryotes. *Curr. Biol.* 13, 94–104.
- Archibald, J.M., Rogers, M.B., Toop, M., Ishida, K., and Keeling, P.J. (2003). Lateral gene transfer and the evolution of plastid-targeted proteins in the secondary plastid-containing alga *Bigeloniella natans*. *Proc. Natl. Acad. Sci. USA* 100, 7678–7683.
- Boucher, Y., Douady, C.J., Papke, R.T., Walsh, D.A., Boudreau, M.E., Nesbo, C.L., Case, R.J., and Doolittle, W.F. (2003). Lateral gene transfer and the origins of prokaryotic groups. *Annu. Rev. Genet.* 37, 283–328.

16. Dean, R.A., Talbot, N.J., Ebbole, D.J., Farman, M.L., Mitchell, T.K., Orbach, M.J., Thon, M., Kulkarni, R., Xu, J.R., Pan, H., et al. (2005). The genome sequence of the rice blast fungus *Magnaporthe grisea*. *Nature* 434, 980–986.
17. Douady, C.J., Delsuc, F., Boucher, Y., Doolittle, W.F., and Douzery, E.J. (2003). Comparison of Bayesian and maximum likelihood bootstrap measures of phylogenetic reliability. *Mol. Biol. Evol.* 20, 248–254.
18. Marchler-Bauer, A., Anderson, J.B., Cherukuri, P.F., DeWeese-Scott, C., Geer, L.Y., Gwadz, M., He, S., Hurwitz, D.I., Jackson, J.D., Ke, Z., et al. (2005). CDD: A conserved domain database for protein classification. *Nucleic Acids Res.* 33, D192–D196.
19. Wagner, R., Straub, M.L., Souciet, J.L., Potier, S., and Montigny, J. (2001). New plasmid system to select for *Saccharomyces cerevisiae* purine-cytosine permease affinity mutants. *J. Bacteriol.* 183, 4386–4388.
20. Harwood, C.S., and Parales, R.E. (1996). The β -ketoacid pathway and the biology of self-identity. *Annu. Rev. Microbiol.* 50, 553–590.
21. Majumdar, S., Ghatak, J., Mukherji, S., Bhattacharjee, A., and Bhaduri, A. (2004). UDP galactose 4-epimerase from *Saccharomyces cerevisiae*. *Eur. J. Biochem.* 271, 753–759.
22. Buchan, A., Collier, L.S., Neidle, E.L., and Moran, M.A. (2000). Key aromatic ring cleaving enzyme, protocatechuate 3,4-dioxygenase, in the ecologically important marine *Roseobacter* lineage. *Appl. Environ. Microbiol.* 66, 4662–4672.
23. Doolittle, W.F. (1998). You are what you eat: A gene transfer ratchet could account for bacterial genes in eukaryotic nuclear genomes. *Trends Genet.* 14, 307–311.
24. Rodriguez-Ezpeleta, N., Brinkmann, H., Burey, S.C., Roure, B., Burger, G., Löffelhardt, W., Bohnert, H.J., Philippe, H., and Lang, B.F. (2005). Monophyly of primary photosynthetic eukaryotes: Green plants, red algae, and glaucophytes. *Curr. Biol.* 15, 1325–1330.
25. Taylor, F.J. (1999). Ultrastructure as a control for protistan molecular phylogeny. *Am. Nat.* 154, S125–S136.
26. Viale, A.M., and Arakaki, A.K. (1994). The chaperone connection to the origins of the eukaryotic organelles. *FEBS Lett.* 341, 146–151.
27. Poirot, O., Suhre, K., Abergel, C., O'Toole, E., and Notredame, C. (2004). 3DCoffee: A web server for mixing sequences and structures into multiple alignments. *Nucleic Acids Res.* 1, 37–40.
28. Rambaut, A. (1996). Se-AL: Sequence Alignment Editor. <http://evolve.zoo.ox.ac.uk/software.html?id=seal>.
29. Keane, T.M., Creevey, C.J., Pentony, M.M., Naughton, T.J., James, O., and McInerney, J.O. (2006). Assessment of methods for amino acid matrix selection and their use on empirical data shows that ad hoc assumptions for choice of matrix are not justified. *BMC Evol. Biol.* 6, 29.
30. Ronquist, F., and Huelsenbeck, J.P. (2003). MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19, 1572–1574.
31. Schmidt, H.A., Strimmer, K., Vingron, M., and von Haeseler, A. (2002). TREE-PUZZLE: Maximum likelihood phylogenetic analysis using quartets and parallel computing. *Bioinformatics* 18, 502–504.
32. Holder, M., and Roger, A.J. PUZZLEBOOT version 1.03. <http://hades.biochem.dal.ca/Rogerlab/Software/software.html>.
33. Felsenstein, J. (1995). PHYLIP (Phylogeny Inference Package), 3.57 ed. Department of Genetics, University of Washington Seattle.
34. Guindon, S., Lethiec, F., Duroux, P., and Gascuel, O. (2005). PHYML Online- a web server for fast maximum likelihood-based phylogenetic inference. *Nucleic Acids Res.* 33, W557–W559.
35. Philippe, H. (2000). Opinion: Long branch attraction and protist phylogeny. *Protist* 151, 307–316.