



Fusion of the subunits α and β of succinyl-CoA synthetase as a phylogenetic marker for Pezizomycotina fungi

Amanda M. Koire and Andre R.O. Cavalcanti

Biology Department, Pomona College, Claremont, CA, USA.

Abstract

Gene fusions, yielding the formation of multidomain proteins, are evolutionary events that can be utilized as phylogenetic markers. Here we describe a fusion gene comprising the α and β subunits of succinyl-coA synthetase, an enzyme of the TCA cycle, in Pezizomycotina fungi. This fusion is present in all Pezizomycotina with complete genome sequences and absent from all other organisms. Phylogenetic analysis of the α and β subunits of succinyl-CoA synthetase suggests that both subunits were duplicated and retained in Pezizomycotina while one copy was lost from other fungi. One of the duplicated copies was then fused in Pezizomycotina. Our results suggest that the fusion of the α and β subunits of succinyl-CoA synthetase can be used as a molecular marker for membership in the Pezizomycotina subphylum. If a species has the fusion it can be reliably classified as Pezizomycotina, while the absence of the fusion is suggestive that the species is not a member of this subphylum.

Key words: fused genes, fungi phylogeny, evolutionary marker.

Received: March 19, 2011; Accepted: June 23, 2011.

Introduction

Amongst eukaryotes, Fungi is the kingdom with most sequenced genomes. The sequenced genomes cover fungal species broadly and include agriculturally, medically, and industrially significant species with very diverse genomes. The fungal kingdom is divided in Chytridiomycota, Neocallimastigomycota, Glomeromycota, Ascomycota, Basidiomycota and Blastocladiomycota, and Microsporidia (Hibbett *et al.*, 2007). Ascomycota, which diverged from Basidiomycota in the region of 741-1195 million years ago (Scannel *et al.*, 2007), is the largest phylum of the Fungi kingdom and is again divided into Saccharomycotina, Taphrinomycotina, and Pezizomycotina (Lumbsch and Huhndorf, 2007). Pezizomycotina is the largest subphylum of Ascomycota and contains more than 33,000 described species (Spatafora *et al.*, 2006).

Pezizomycotina, often referred to as filamentous fungi, grow primarily in a filamentous form and degrade biomass to free sugars that are used as source of energy and carbon (Arvas *et al.*, 2007). Most species exhibit a dominant hyphal growth form, with almost all of the sexually reproducing forms possessing ascospores (Spatafora *et al.*, 2006). Filamentous ascomycetes have been shown to be a monophyletic group distinguished from Taphrinomycotina and Saccharomycotina on the basis of ribosomal DNA se-

quence analysis (Alexopoulos *et al.*, 1996; Spatafora *et al.*, 2006).

In recent years, comparative genomics rather than gross morphology has been used to analyze specific subgroups of fungi. Notably, analysis of the phylum Saccharomycotina has revealed major events of evolution such as recent whole genome duplication and subsequent gene loss or subspecialization (Arvas *et al.*, 2007; Dietrich *et al.*, 2004). In Pezizomycotina, on the other hand, only a subset of protein families related to plant biomass degradation and secondary metabolism revealed signs of recent expansion, likely because of a repeat-induced point mutation mechanism for the active removal of duplicated sequences that has been identified in varying levels in Pezizomycotina species (Arvas *et al.*, 2007). Gene duplication significantly influences evolution, as duplicated genes may be retained if they provide an advantage, either because of an increased gene dosage effect or because of increased functional diversification, or may alternatively be lost (Wagner, 1998; Li and Graur, 2000; Cornell *et al.*, 2007). Multiple copies of the genes coding for the α and β subunits of succinyl-CoA synthetase have been shown to exist in several species of *Aspergillus* and in *Neosartorya fisheri*, placing succinyl-CoA synthetase among the subset of proteins showing gene expansion in Pezizomycotina (Flippi *et al.*, 2009).

Succinyl-CoA synthetase, also known as succinate-CoA ligase and succinate thiokinase, is an acid-thiol ligase that forms carbon-sulfur bonds. Succinyl-CoA synthetase is an enzyme of the citrate cycle, a major metabolic path-

way in which acetyl-CoA derived from the catabolism of sugars and lipids is oxidized to carbon dioxide and energy in the form of GTP, NADH, and FADH₂ (Hanson, 2008). This enzyme catalyzes the reversible transformation of succinyl-CoA, orthophosphate, and GDP to succinate, GTP, and the four-carbon metabolite coenzyme A in a three-step substrate-level phosphorylation reaction (Kanehisa and Goto, 2000; Hamblin *et al.*, 2008). Succinyl-CoA synthetase is composed of two subunits, α and β , which generally form dimers in eukaryotes and tetramers in prokaryotes. Nucleotide specificity of succinyl-CoA synthetase is thought to be determined by the N-terminal domain of the beta subunit (Hamblin *et al.*, 2008), and the alpha subunit may contain a targeting sequence for the enzyme (Birney and Klein, 1995). While prokaryotes are capable of using both types of nucleotides, eukaryotic succinyl-CoA synthetase enzymes usually occur as isoforms specific for either GTP or ATP (Hamblin *et al.*, 2008). Multiple copies of the succinyl-coA synthetase genes tallies with the finding that duplications are most common in Pezizomycotina when the proteins produce secondary metabolites (Arvas *et al.*, 2007), and with the common belief that natural selection on metabolic phenotypes mainly targets enzymes that control the rate of turnover of molecules through a metabolic pathway, such as succinyl-CoA synthetase (Flipphi *et al.*, 2009).

In Pezizomycotina, one set of the duplicated α and β gene copies for succinyl-CoA synthetase has become fused. Gene fusion, yielding formation of multidomain proteins, is a major driving force in protein evolution and is a rare evolutionary event in comparison to the frequency of point mutations. Consequently, gene fusion events were recently utilized as an evolutionary marker to pinpoint the root of the eukaryotic tree (Nara *et al.*, 2000; Stechmann and Cavalier-Smith, 2002, 2003). Here we examine how the presence of the succinyl-CoA synthetase fusion gene can be used as a simple phylogenetic marker for Pezizomycotina, and attempt to trace the origin and nature of this fusion in Fungi.

Materials and Methods

GenBank sequence data

We used the fungi genome BLAST tool from NCBI to search all complete fungal genomes with predicted proteins in GenBank. We started with the sequence of a Pezizomycotina succinyl-CoA synthetase α and β fusion protein, and used blastp to identify all other succinyl-CoA synthetase proteins in different fungal genomes.

With the exception of Microsporidia species, we were able to identify at least one α and one β subunit in each fungal genome available in the Genomic BLAST page (Table S1). The searches returned both fused and non-fused genes. For aligning the fused proteins with non-fused α and α cop-

ies of the protein, the fusions were artificially split into α' and β' subunits. To increase the diversity of fungal species analyzed, we also included the sequences of the α and β succinyl-CoA synthetase subunits of two Neocallimastigomycota fungi. As an outgroup we used the α and β subunits of succinyl-CoA synthetase from three mammals, *Homo sapiens*, *Mus musculus* and *Canis familiaris*.

Phylogenetics analyses

An alignment of the α subunits of succinyl-CoA synthetase for 53 organisms, including 25 Pezizomycotina, 14 Saccharomycotina, 2 Taphrinomycotina, 7 Basidiomycota, 2 Neocallimastigomycota as well as 3 metazoans to provide an outgroup, was created in ClustalX v.2.0.11 (Larkin *et al.*, 2007) using default settings. A total of 80 sequences were aligned, including the region corresponding to the α subunit in the fusion genes in Pezizomycotina. A separate alignment of β and β' subunits of the same organisms was created in the same manner. In total, 83 sequences were used in the β subunit alignment, including the region corresponding to the β subunit in the fusion genes in Pezizomycotina.

Each set of aligned sequences corresponding to the α and β subunits was used as input to infer phylogenetic trees using maximum likelihood and Bayesian analysis. For the maximum likelihood analysis we first used the program ProtTest (Abascal *et al.*, 2005) to select the amino acid substitution model that best fits the protein alignments. In both datasets, the best model of amino acid substitution was the LG model (Le and Gascuel, 2008), with invariable sites and across site rate variation. We used PhyML 3.0 (Guindon and Gascuel, 2003), as implemented in SeaView (Galtier *et al.*, 1996) to create maximum likelihood trees using the LG model of amino acid substitution (Le and Gascuel, 2008) with optimized number of invariable sites and optimized across site rate variation. Bootstrap support was calculated using the aLRT model (Anisimova and Gascuel, 2006). The Bayesian analysis was performed using MrBayes (Ronquist and Huelsenbeck, 2003) with four Monte Carlo Markov Chains and variable rates of substitution. Each chain was run for 1,000,000 generations and trees were sampled every 1,000 generations. The first 100 trees were excluded as burn-in and we used the remaining 900 trees to calculate bootstrap values. The trees were visualized and prepared for publication using Dendroscope (Huson *et al.*, 2007) and FigTree.

Results

The fusion of the α and β subunits of succinyl-CoA can serve as a phylogenetic marker for Pezizomycotina fungi

We used the sequence of the fused α and α subunits of succinyl-CoA synthetase from *Pyrenophora tritici-*

repentis, a Pezizomycotina, to search all the complete fungal genomes with protein predictions available in the genomic blast webpage from NCBI.

As of August 1st 2009, the genomic blast webpage allowed the search of the protein sequences of 51 fungal genomes: 41 Ascomycota, 7 Basidiomycota and 3 Microsporidia. The 41 Ascomycota are divided in 25 Pezizomycotina, 14 Saccharomycotina, 2 Taphrinomycotina. All genomes analyzed had both the α and β subunits of succinyl-CoA synthetase, with the exception of the three Microsporidial genomes that completely lack both subunits of the protein and were not analyzed further. The genome of the Basidiomycota *Postia placenta* also seems to be lacking both subunits of the gene; however, an additional search of the genome page for this organism in the Joint Genome Institute (JGI), revealed that the genes are present, and we used the sequences obtained from JGI in the analyses for this organism.

In addition to the sequences obtained from the genomic blast searches we included the sequences of 2 Neocallimastigomycota fungi in the analysis to increase the diversity of the fungal species. The Ascomycota, together with Basidiomycota, form the subkingdom Dikarya, and are more closely related to each other than to Neocallimastigomycota or Microsporidia. Finally, we used the sequences of the α and β subunits of succinyl-CoA synthetase from three Metazoans as outgroup.

Out of the 25 Pezizomycotina fungi with complete genomes available in NCBI, we immediately detected the fusion gene in 22 genomes. According to the annotation in NCBI the fusion appears to be absent from *Aspergillus niger*, *Coccidioides immitis* and *Gibberella zeae*. Interestingly, each of these species has at least two α and two β subunits of succinyl-CoA synthetase. Thus, we further analyzed these genomes to determine if the fusions are really absent or if they are just not annotated.

In *Aspergillus niger*, there are two subunits α and two subunits β annotated in the genome; a tblastn search of the fusion protein against *A. niger* genome showed that a subunit β gene immediately follows a subunit α gene in the same strand, indicating that the genes could be fused in this organism. Furthermore, a blastp search of the gene predictions in this organism's genome webpage in JGI revealed that a fusion gene is present (gene model number jgi|Aspni1|176118). To maintain consistency, we used the proteins in GenBank in the analyses. The proteins involved in the putative fusion protein in *A. niger* group with the fusion proteins in the other Pezizomycotina in the phylogenetic trees, further confirming that these two proteins are indeed fused in this genome (Figures 1 and 2, Table S1).

A similar problem occurs in *Coccidioides immitis*. A tblastn search revealed that two consecutive genes in the genome hit the fusion gene. In this case, however, the first gene only hits part of subunit α while the second gene hits

the remainder of subunit α and the full length of subunit β . We aligned both genes with the fusion genes and determined the boundaries between the α and β subunits assuming the proteins are fused and used these sequences in the analyses. Like in *A. niger*, the putative fusion proteins groups with the fusion proteins in the other Pezizomycotina in the phylogenetic trees (Figures 1 and 2, Table S1).

Finally, while the fusion is not annotated in *Gibberella zeae*, in this organism a subunit β gene follows a subunit α gene. Both subunits are in the same strand but are 3 kb apart and there seems to be a small intervening gene in the opposite strand between the subunits. A large intron between the subunits could have confounded the gene prediction software, and it is not unknown for a gene to be nested inside an intron of another (Kumar, 2009). It is worth noting that the version of the genome available in GenBank is based on a preliminary whole-genome-shotgun that has been superseded by a new assembly of the genome. Unfortunately, the new assembly doesn't seem to be available at the time of this writing. Interestingly, these two proteins group with the fusion proteins in the phylogenetic trees (Figures 1 and 2, Table S1), indicating that they could be fused. The alternative explanation would be that the proteins were once fused and that the insertion of a gene in between the α and β subunits broke the fusion into two working copies of the subunits; while plausible, this second explanation seems to be unlikely as the insertion would have to have occurred right at the junction between the two subunits. This organism also has an extra copy of the subunit α that groups with the fusions, suggesting that the fusion gene was duplicated in this organism and the subunit β was eliminated in one of the copies of the fusion.

These results indicate that, even though it might not have been annotated as such, the fusion of the α and β subunits of succinyl-CoA synthetase is present in all 25 Pezizomycotina genomes available at GenBank, with the possible exception of that of *G. zeae* in which the fusion protein seem to at least have once been present. In addition, the fusion gene was not detected in any other fungal genome and a blast search of the entire NR database revealed that the fusion is only present in Pezizomycotina fungi.

The apparent presence of the fusion gene in most Pezizomycotina and its absence in all other organisms suggest that the fused succinyl-CoA synthetase gene can serve as a phylogenetic marker for Pezizomycotina. If a species has the fusion it can be reliably classified as Pezizomycotina; if a species does not have the fusion it most likely does not belong to this subphylum.

Evolution of the α and β subunits of succinyl-CoA synthetase in fungi

We used the sequences of the succinyl-CoA synthetase of 53 species to build phylogenetic trees for both the α and β subunits of the protein. The data set contained 48

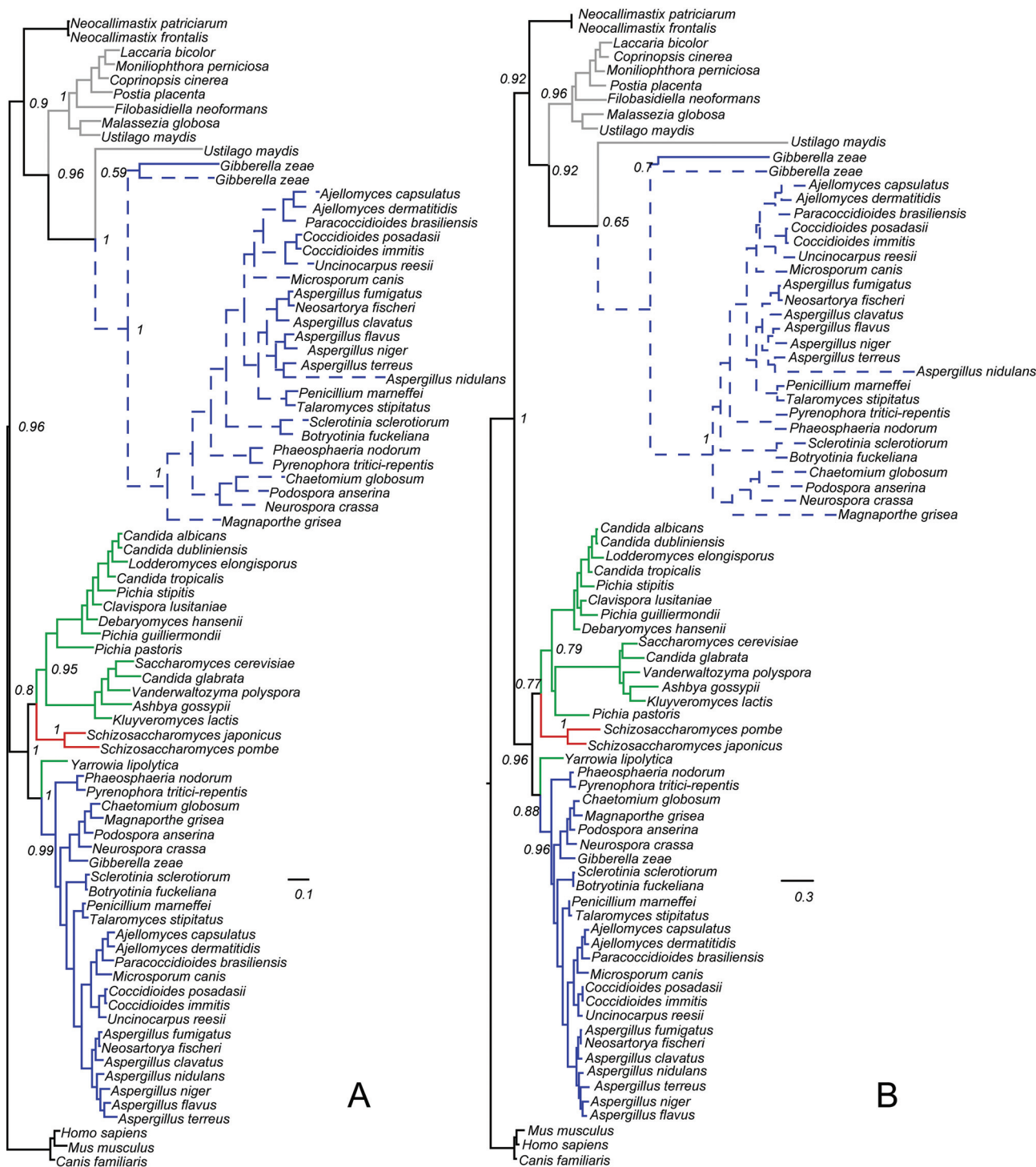


Figure 1 - Phylogenetic trees of the α subunit of succinyl-CoA synthetase in fungi. Pezizomycotina sequences are in blue, Saccharomycotina in green, Taphrinomycotina in red and Basidiomycota in gray. Neocallimastigomycotina and Metazoans (outgroup) are in black. Dashed lines represent fused genes. Bootstrap support values are given in relevant nodes. (a) Mr Bayes tree; (b) PhyML Maximum likelihood tree.

Dikarya, 2 Neocallimastigomycota and 3 metazoans. The Dikarya are divided in 41 Ascomycota (25 Pezizomycotina, 14 Saccharomycotina, 2 Taphrinomycotina) and 7 Basidiomycota (Table S1). The sequences of the fusion genes in the 25 Pezizomycotina were split into their α and β constituents.

All sequences were aligned in ClustalW using default parameters and trees were built using MrBayes and PhyML. The resulting trees are shown in Figures 1 and 2. In all trees the fused genes form a monophyletic group (with strong bootstrap support) and the non-fused copies in Pezi-

zomycotina form another monophyletic group (with strong bootstrap support).

The phylogenetic trees of the α subunits (Figure 1) suggest that this subunit duplicated before the divergence of Neocallimastigomycota from Dikarya. Saccharomycotina and Taphrinomycotina lost the copy that would later become fused in Pezizomycotina, while Neocallimastigomycota and Basidiomycota lost the other copy. The α

subunit present in Neocallimastigomycota and Basidiomycota groups with the fused copy in Pezizomycotina, while the α subunit present in Saccharomycotina and Taphrinomycotina groups with the unfused copies of the subunit in Pezizomycotina. *G. zeae* has an extra copy of subunit α that groups with the fused genes, indicating that it probably resulted from a recent duplication of the fused gene followed by the loss of the β subunit. This loss should



Figure 2 - Phylogenetic tree of the β subunit of succinyl-CoA synthetase in fungi. Pezyzomycotina sequences are in blue, Saccharomycotina in green, Taphrinomycotina in red and Basidiomycota in gray. Neocallimastigomycotina and Metazoans (outgroup) are in black. Dashed lines represent fused genes. Bootstrap support values are given in relevant nodes. (a) Mr Bayes tree; (b) PhyML Maximum likelihood tree.

not have a large phenotypic effect, as another copy of the fusion gene would be present in the genome.

The trees of the β subunit (Figure 2) indicate a different evolutionary story for this subunit. The β subunit duplicated in Dikarya after their divergence from Neocallimastigomycota, and after the divergence of basidiomycetes from ascomycetes, but before the divergence of Pezizomycotina from Saccharomycotina and Taphrinomycotina. The copy that would become fused in Pezizomycotina was lost from Saccharomycotina and Taphrinomycotina.

A second unfused copy of the β subunit is present in three *Aspergillus* species (*A. clavatus*, *A. flavus* and *A. fumigatus*) and in *Neosartorya fischeri* (Figure 2), all members of Trichocomaceae, and groups with the fusion proteins. This suggests that the β subunit duplicated a second time before becoming fused in Pezizomycotina and was lost from most Pezizomycotina with the exception of some Trichocomaceae.

Interestingly, one Basidiomycota, *Ustilago maydis*, also has two copies of each subunit of succinyl-CoA synthetase. One of the α and one of the β subunits group with the fusion gene in Pezizomycotina, which would suggest that this gene could have been the result of a lateral gene transfer between a Pezizomycotina and *U. maydis*. This would not be the first gene shown to have been transferred from Pezizomycotina to *U. maydis* (Shen *et al.*, 2009). We checked the location of these genes in the genome of *U. maydis* and found that they are consecutive but in opposite strands, and therefore cannot be not fused in this species. A closer look at the trees reveals that the sequence of the *U. maydis* subunits that group with the fusions has a long branch, and its grouping with the fusions could be explained by long branch attraction. The presence of two copies of each of the subunits in the genome of *U. maydis* might have left one copy of each subunit free to accumulate mutations, explaining why they would be in such a long branch. However, it is also possible that the fusion gene was horizontally transferred to *U. maydis* and a genome rearrangement broke the gene in two parts and reversed one of the subunits in relation to the other, although this explanation is less likely considering that such rearrangement would need to have happened in such a way as to generate working versions of the genes.

With the exception of *U. maydis*, the Basidiomycota form a monophyletic group in the trees. The Neocallimastigomycota, Taphrinomycotina and Pezizomycotina also form monophyletic groups. The Saccharomycotina only form a monophyletic group in the maximum likelihood tree of the β subunit (Figure 2b); in all other trees *Yarrowia lipolytica* groups with the unfused copy of the proteins in Pezizomycotina and not with Saccharomycotina. This unexpected placement of *Yarrowia lipolytica* has been seen in other analyses as well (Arvas *et al.*, 2007, Cornell *et al.*, 2007) and has been attributed to *Y. lipolytica*

ORFs bearing more similarity to other fungi than to Saccharomycotina (Arvas *et al.*, 2007).

Discussion

Our results show that the sub-phylum Pezizomycotina is characterized by the presence of a fusion of the α and β subunits of succinyl-CoA synthetase that can be used as a phylogenetic marker to determine if an organism belongs in Pezizomycotina. If a species has the fusion it can be reliably classified as Pezizomycotina, while the absence of the fusion is suggestive that the species is not a member of this subphylum.

Phylogenetic analyses show that the two subunits of succinyl-CoA synthetase duplicated at different times in fungal evolution and then became fused in an ancestor of Pezizomycotina. The gene for the α subunit of succinyl-CoA synthetase duplicated in an ancestor of the Neocallimastigomycota and Dikarya. One copy of the gene was then lost in all fungi with the exception of Pezizomycotina. The gene for the subunit β of succinyl-CoA synthetase most likely duplicated in an ancestor of the Ascomycota after the divergence of this group from Basidiomycota. One copy of the gene was then lost from Saccharomycotina and Taphrinomycotina, and kept in Pezizomycotina.

In Pezizomycotina, one copy of the α subunit fused to a copy of the β subunit, generating a fusion protein while retaining the unfused copies as well. As this fusion gene involves two subunits of the same enzyme it is possible that it provides a selective advantage by leading to greater catalytic activity or more efficient co-regulation of the expression of the two succinyl-CoA subunits.

Even though the presence of this fused gene could lead to a more efficient enzyme, either by greater catalytic activity or by more efficient co-regulation, Pezizomycotina genomes maintained two copies (one fused one unfused) of each subunit of succinyl-CoA synthetase. This might indicate that this enzyme is needed in large quantities in Pezizomycotina or, more interestingly, it is possible that one of the copies of succinyl-CoA synthetase in Pezizomycotina has acquired an as yet unreported new function, as suggested by Flipphi *et al.* (2009) for other duplicated genes in the primary carbon metabolism of *Aspergillus* species.

Acknowledgments

This material is based upon work supported by the National Science Foundation under Award No. 0920697. AMK was funded with a grant of the HHMI to Pomona College's Biology Department. The authors thank Daniel Martinez (Pomona College) for help generating and interpreting the phylogenetic trees.

References

- Abascal F, Zardoya R and Posada D (2005) ProtTest: Selection of best-fit models of protein evolution. *Bioinformatics* 21:2104-2105.
- Alexopoulos CJ, Mims CW and Blackwell M (1996) *Introductory Mycology*. 4th edition. John Wiley & Sons, New York, 880 pp.
- Anisimova M and Gascuel O (2006) Approximate likelihood ratio test for branches: A fast, accurate and powerful alternative. *Syst Biol* 55:539-552.
- Arvas M, Kivioja T, Mitchell A, Saloheimo M, Ussery D, Penttila M and Oliver S (2007) Comparison of protein coding gene contents of the fungal phyla Pezizomycotina and Saccharomycotina. *BMC Genomics* 8:e325.
- Birney MA and Klein C (1995) Cloning and expression of the α subunit of succinyl-CoA synthetase from *Dictyostelium discoideum*. *Arch Biochem Biophys* 319:93-101.
- Cornell MJ, Alam I, Soanes DM, Wong HM, Hedeler C, Paton NW, Rattray M, Hubbard SJ, Talbot NJ and Oliver SG (2007) Comparative genome analysis across a kingdom of eukaryotic organisms: Specialization and diversification in the fungi. *Genome Res* 17:1809-1822.
- Dietrich FS, Voegeli S, Brachat S, Lerch A, Gates K, Steiner S, Mohr C, Pohlmann R, Luedi P, Choi S, *et al.* (2004) The *Ashyba gossypii* genome as a tool for mapping the ancient *Saccharomyces cerevisiae* genome. *Science* 304:304-307.
- Flippi M, Sun J, Robellet X, Karaffa L, Fekete E, Zeng A and Kubicek C (2009) Biodiversity and evolution of primary carbon metabolism in *Aspergillus nidulans* and other *Aspergillus spp.* *Fungal Genet Biol* 46:S19-S44.
- Guindon S and Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* 52:696-704.
- Galtier N, Gouy M and Gautier C (1996) SeaView and Phylo_win, two graphic tools for sequence alignment and molecular phylogeny. *Comp Appl Biosci* 12:543-548.
- Hamblin K, Standley D, Rogers M, Stechmann A, Roger A, Maytum R and van der Giezen M (2008) Localization and nucleotide specificity of *Blastocystis* succinyl-CoA synthetase. *Mol Microbiol* 68:1395-1405.
- Hanson J (2008) *The Chemistry of Fungi*. Royal Society of Chemistry Publishing, Cambridge, 244 pp
- Hibbett DS, Binder M, Bischoff JF, Blackwell M, Cannon PF, Eriksson OE, Huhndorf S, James T, Kirk PM, Lücking R, *et al.* (2007) A higher-level phylogenetic classification of the Fungi. *Mycol Res* 111:509-547.
- Huson D, Richter D, Rausch C, DeZulian T, Franz M and Rupp R (2007) Dendroscope: An interactive viewer for large phylogenetic trees. *BMC Bioinformatics* 8:e460.
- Kanehisa M and Goto S (2000) KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res* 28:27-30.
- Kumar A (2009) An overview of nested genes in eukaryotic genomes. *Eukaryotic Cell* 8:1321-1329.
- Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, *et al.* (2007) Clustal W and Clustal X v. 2.0. *Bioinformatics* 23:2947-2948.
- Le SQ and Gascuel O (2008) An improved general amino-acid replacement matrix. *Mol Biol Evol* 25:1307-1320.
- Li WH and Graur D (2000) *Fundamentals of Molecular Evolution*. 2nd edition. Sinauer Associates Inc, Sunderland, 439 pp.
- Lumbsch HT and Huhndorf SM (2007) Outline of Ascomycota. *Mycotax* 13:1-58.
- Nara T, Hshimoto T and Aoki T (2000) Evolutionary implications of the mosaic pyrimidine-biosynthetic pathway in eukaryotes. *Gene* 257:209-222.
- Ronquist F and Huelsenbeck JP (2003) MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19:1572-1574.
- Scannell D, Butler G and Wolfe K (2007) Yeast genome evolution – The origin of the species. *Yeast* 24:11:929.
- Shen Y, Lang B and Burger G (2009) Diversity and dispersal of a ubiquitous protein family: Acyl-CoA dehydrogenases. *Nucleic Acids Res* 37:5619-5631.
- Spatafora J, Sung G, Johnson D, Hesse C, O'Rourke B, Serdani M, Spotts R, Lutzoni F, Hofstetter V, Miadlikowska J, *et al.* (2006) A five-gene phylogeny of Pezizomycotina. *Mycologia* 98:1018-1028.
- Stechmann A and Cavalier-Smith T (2002) Rooting the eukaryote tree by using a derived gene fusion. *Science* 297:89-91.
- Stechmann A and Cavalier-Smith T (2003) The root of the eukaryote tree pinpointed. *Curr Biol* 13:R665-R666.
- Wagner A (1998) The fate of duplicated genes: Loss or new function? *Bioessays* 20:785-788.

Internet Resources

- Fungi genome BLAST tool, http://www.ncbi.nlm.nih.gov/sutils/genom_table.cgi?organism=fungi (August 2009).
- FigTree Software, <http://tree.bio.ed.ac.uk/software/figtree/> (July 2010).
- Postia placenta genome page at Joint Genome Institute (JGI), <http://genomeportal.jgi-psf.org/Pospl1/Pospl1.home.html> (June 2010).
- Aspergillus niger genome page at Joint Genome Institute (JGI), <http://genome.jgi-psf.org/Aspni5/Aspni5.home.html> (June 2010).

Supplementary Material

The following online material is available for this article:

Table S1 - Organisms and accession numbers of the proteins used in this study.

This material is available as part of the online version of this article from <http://www.scielo.br/gmb>.

Associate Editor: Célia Maria de Almeida Soares

License information: This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Table S1: Organisms and accession numbers of the proteins used in this study.

Species	Accession numbers		
	Fusion	α	β
Pezizomycotina			
<i>Ajellomyces capsulatus</i>	gbIEEH08294.1	gbIEEH06102.1	gbIEEH03354.1
<i>Ajellomyces dermatitidis</i>	gbIEEQ69414.1	gbIEEQ77311.1	gbIEEQ76061.1
<i>Aspergillus clavatus</i>	gbIEAW07752.1	gbIEAW12523.1	gbIEAW13338.1 gbIEAW09110.1
<i>Aspergillus flavus</i>	gbIEED50054.1	gbIEED46808.1	gbIEED44815.1 gbIEED53391.1
<i>Aspergillus fumigatus</i>	gbIEAL90410.2	gbIEAL91981.1	gbIEAL84600.1 gbIEAL84389.2
<i>Aspergillus nidulans</i>	gbIEAA65786.1	gbIEAA64406.1	gbIEAA61646.1
<i>Aspergillus niger</i>	reflXP_001392414.1 reflXP_001392415.1 1	reflXP_001398396. 1	reflXP_001400660.1
<i>Aspergillus terreus</i>	gbIEAU38921.1	gbIEAU30552.1	gbIEAU37706.1 gbIEAU31032.1
<i>Botryotinia fuckeliana</i>	gbIEDN22743.1	gbIEDN24969.1	gbIEDN23758.1
<i>Chaetomium globosum</i>	gbIEAQ92530.1	gbIEAQ86176.1	gbIEAQ86581.1
<i>Coccidioides immitis</i>	gbIEAS31442.1 gbIEAS31443.1 ¹	gbIEAS30565.1	gbIEAS29800.1
<i>Coccidioides posadasii</i>	gbIEER27972.1	gbIEER27280.1	gbIEER28332.1

<i>Gibberella zeae</i>	reflXP_388074.1 reflXP_388076.1 ¹	reflXP_382206.1 reflXP_391661.1	reflXP_391061.1
<i>Magnaporthe grisea</i>	gbLEDK06465.1	gbLEDK06157.1	gbLEDJ98088.1
<i>Microsporum canis</i>	gbLEEQ28748.1	gbLEEQ34080.1	gbLEEQ29517.1
<i>Neosartorya fischeri</i>	gbEAW22789.1	gbEAW24125.1	gbEAW20453.1 gbEAW25715.1
<i>Neurospora crassa</i>	gbEAA34834.1	gbEAA32357.1	gbEAA34077.1
<i>Paracoccidioides brasiliensis</i>	gbEEH17930.1	gbEEH16608.1	gbEEH19675.1
<i>Penicillium marneffei</i>	gbEEA20408.1	gbEEA24495.1	gbEEA24589.1
<i>Phaeosphaeria nodorum</i>	gbEAT83662.2	gbEAT77317.1	gbEAT84267.1
<i>Podospora anserina</i>	reflXP_001908936.1	reflXP_001911226. 1	reflXP_001909573.1
<i>Pyrenophora tritici-repentis</i>	gbLEDU41918.1	gbLEDU40695.1	gbLEDU51437.1
<i>Sclerotinia sclerotiorum</i>	gbEDO04371.1	gbEDO05039.1	gbEDN91995.1
<i>Talaromyces stipitatus</i>	gbEED24140.1	gbEED18241.1	gbEED18353.1
<i>Uncinocarpus reesii</i>	gbEEP81085.1	gbEEP79486.1	gbEEP77760.1
Saccharomycotina			
<i>Ashbya gossypii</i>		reflNP_985681.1	reflNP_984650.1
<i>Candida albicans</i>		gbLEEQ42982.1	gbLEEQ43728.1
<i>Candida dubliniensis</i>		reflXP_002416832. 1	reflXP_002422190.1
<i>Candida glabrata</i>		reflXP_447560.1	reflXP_446321.1

<i>Candida tropicalis</i>	gblEER31682.1	gblEER35932.1
<i>Clavispora lusitaniae</i>	gblEEQ39475.1	gblEEQ38347.1
<i>Debaryomyces hansenii</i>	reflXP_461997.1	reflXP_462565.1
<i>Kluyveromyces lactis</i>	reflXP_454091.1	reflXP_451242.1
<i>Lodderomyces elongisporus</i>	gblEDK43263.1	gblEDK45048.1
<i>Pichia guilliermondii</i>	gblEDK38689.2	gblEDK39982.2
<i>Pichia pastoris</i>	reflXP_002493061. 1	reflXP_002491782.1
<i>Pichia stipitis</i>	reflXP_001387359. 1	reflXP_001382488.1
<i>Saccharomyces cerevisiae</i>	reflNP_014785.1	reflNP_011760.1
<i>Vanderwaltozyma polyspora</i>	gblEDO15166.1	gblEDO19215.1
<i>Yarrowia lipolytica</i>	reflXP_504333.2	reflXP_502415.2
Taphrinomycotina		
<i>Schizosaccharomyces japonicus</i>	gblEEB06483.1	gblEEB08651.1
<i>Schizosaccharomyces pombe</i>	reflNP_594230.1	reflNP_588466.1
Basidiomycota		
<i>Coprinopsis cinerea</i>	gblEAU93204.1	gblEAU93446.1
<i>Filobasidiella neoformans</i>	reflXP_571914.1	reflXP_567043.1
<i>Laccaria bicolor</i>	gblEDR15553.1	gblEDR15438.1
<i>Malassezia globosa</i>	gblEDP43341.1	gblEDP42407.1

<i>Moniliophthora perniciosa</i>	gblEEB89615.1	gblEEB90675.1
<i>Postia placenta</i> ²	jgilPosp11 128778	jgilPosp11 120437
<i>Ustilago maydis</i>	gblEAK82784.1	gblEAK81186.1
	gblEAK82289.1	gblEAK82785.1
Neocallimastigomycota		
<i>Neocallmastix frontalis</i>	gblAAP13544.1	splP53587.2
<i>Neocallimastix patriciarum</i>	gblAAP83350.1	gblAAP83351.1
Metazoa - Outgroup		
<i>Homo sapiens</i>	gblAAH00504.1	dbjlBAA91939.1
<i>Mus musculus</i>	dbjlBAB22331.1	dbjlBAE29687.1
<i>Canis familiaris</i>	reflXP_532985.2	reflXP_542566.2

¹ Even though these two genes are annotated as independent genes, they are consecutive and in the same strand in the genome and group with the fusion genes in the phylogenetic trees, indicating that they might be fused. See text for a discussion.

² Although both genes seem to be absent from this genome in NCBI, further search in the organism's genome webpage at JGI (<http://genomeportal.jgi-psf.org/Posp11/Posp11.home.html>) revealed that the genes are present. The accession numbers given in the table correspond to the JGI accession numbers.