Zygomycetes, Microsporidia, and the Evolutionary Ancestry of Sex Determination

Tina Koestler, and Ingo Ebersberger*

Center for Integrative Bioinformatics Vienna, University of Vienna, Medical University of Vienna, University of Veterinary Medicine Vienna, Vienna, Austria

*Corresponding author: E-mail: ingo.ebersberger@univie.ac.at.

Accepted: 3 February 2011

Abstract

Zygomycetes and their alleged sister taxon, the microsporidia, exclusively share the presence of a cluster of three genes encoding a sugar transporter, a high mobility group (HMG)-type transcription factor, and an RNA helicase. In zygomycetes, the HMG-type transcription factor acts as the sole sex determinant. This intimately ties the evolutionary history of this gene cluster to the evolution of sex determination. Here, we have unraveled the relationships of the two gene clusters by vicariously analyzing the sugar transporters and the RNA helicases. We show that if the two gene clusters share a common ancestry, it dates back to the early days of eukaryotic evolution. As a consequence, the zygomycete MAT locus would be old enough to represent the archetype of fungal and animal sex determination. However, the evolutionary scenario that has to be invoked is complex. An independent assembly of the two clusters deserves therefore consideration. In either case, shared ancestry or convergent evolution, the presence of the gene cluster in microsporidia and in zygomycetes represents at best a plesiomorphy. Hence, it is not phylogenetically informative. A further genome-wide reanalysis of gene order conservation reveals that gene order is not significantly more similar between microsporidia and zygomycetes than between microsporidia and any other fungal taxon or even humans. Consequently, the phylogenetic placement of microsporidia as sister to the zygomycetes needs to be reconsidered.

Key words: sex determination, gene cluster, HMG, plesiomorphy, MAT locus, shared synteny.

Introduction

In zygomycetes, an early branching fungal lineage, the mating type is determined by a single gene (Idnurm et al. 2008) encoding a high mobility group (HMG)-type transcription factor (Thomas and Travers 2001). The sex-determining locus is flanked by two genes, one encoding a triosephosphate transporter (TPT) and the other encoding an RNA helicase. Initial analyses suggested that the arrangement of these three genes is unique to the zygomycetes. However, recently, it was reported that a similar cluster encompassing also genes for a TPT, an HMG protein, and an RNA helicase is present in microsporidia (fig. 1). It was concluded that the two gene clusters have a common evolutionary origin, that is, are shared syntenic (Lee et al. 2008). The consequences of this conclusion are of relevance for two open questions.

First, is sex determination via HMG-type transcription factors evolutionarily ancient? Fungal sex is determined by mating-type specific genes organized in so-called MAT loci. A number of MAT loci have been described in ascomycetes and basidiomycetes (e.g., Haber 1998; Lengeler et al. 2002; Butler et al. 2004, reviewed in Lee, Ni, et al. 2010). Based on the transcription factors present, the MAT loci are classified into three major groups: 1) HMG type, 2) homeodomain type, and 3) alpha-domain type. The identification of the zygomycete MAT locus (Idnurm et al. 2008) revealed that in the earliest branching fungal lineage characterized so far, an HMG-type transcription factor determines sex. This laid the odds on an HMG-type MAT locus having determined sex in the last common ancestor of all fungi. The simplicity of the zygomycete MAT locus further suggested that it could resemble the archetype of fungal sex determination (Dyer 2008a). Interestingly, also in mammals, sex is determined by a single HMG-type transcription factor (Haqq et al. 1993). This coincidence was taken as an indication that both fungal and mammalian sex-determining systems descended from the same HMG-type MAT locus in the last common ancestor of fungi and animals (Dyer 2008a; Idnurm et al. 2008). However, this scenario is

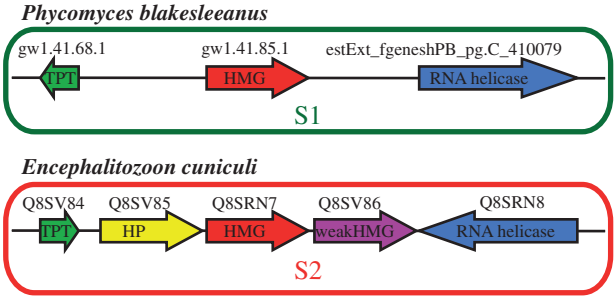

Fig. 1.—Gene arrangements in the MAT locus of *Phycomyces blakesleeanus* and in the corresponding gene cluster of *Encephalitozoon cuniculi*. In the sex-related region of *P. blakesleeanus* (S1), the gene encoding the sex-determining transcription factor (HMG) is flanked by two genes coding for a sugar transporter (TPT) and an RNA helicase, respectively. The corresponding cluster in *E. cuniculi* (S2) contains also genes for a TPT, an HMG-type transcription factor, and an RNA helicase. The two additional genes in S2 encode a hypothetical protein (HP) and a protein with a weak similarity to a HMG domain protein.

speculative. The high evolutionary rate of sex-determining genes (Swanson and Vacquier 2002) prevents a reconstruction of their evolutionary relationships already within the fungi (e.g., Lee, Corradi, et al. 2010). Thus, protein sequence data provide no information about whether sex-determining HMG-type transcription factors in fungi and animals are derived from a single ancestral gene, or whether they are a product of convergent evolution.

To still establish homology for highly diverged genes, gene order has proven helpful (Dietrich et al. 2004). The sex determining gene in zygomycetes and the gene for the microsporidian HMG protein identified by Lee et al. (2008) are both flanked by genes encoding a TPT and an RNA helicase. They concluded that this results from shared synteny, ergo the microsporidian HMG-type transcription factor is the first homolog to the zygomycete sex determinant identified in a nonzygomycete taxon. Consequently, we can now investigate the evolutionary history of HMG-driven sex determination by reconstructing the evolutionary history of the two gene clusters in zygomycetes and microsporidia.

Second, what is the exact position of microsporidia in the eukaryotic tree of life (reviewed in Corradi and Keeling 2009)? Initially, microsporidia were considered an early branching eukaryotic lineage (e.g., Cavalier-Smith 1986; Vossbrinck et al. 1987). Later findings, however, were not consistent with this view (e.g., Thomarat et al. 2004; Brinkmann et al. 2005; Gill and Fast 2006; James et al. 2006; Keeling 2009). After several taxonomic revisions, it is now widely accepted that microsporidia are associated with the fungi (Corradi and Keeling 2009). Alas, so far protein phylogenies failed to resolve whether microsporidia are sister to the fungi, or whether they fall within the true fungi. Some studies suggested a grouping of microsporidia with various fungal lineages, such as the ascomycetes, the basidiomycetes, the zygomycetes, or *Rozella* a chytridiomycete (e.g., Keeling et al. 2000; Keeling 2003; Thomarat et al. 2004; Gill and Fast 2006; James et al. 2006). However, their position as a sister taxon to all fungi could not be rejected (James et al. 2006). Only recently comparative genome structure analyses provided complementary information about the phylogenetic position of the microsporidia. Among all tested fungal and nonfungal species, only the zygomycetes are reported to have a gene order that is more similar to that of the microsporidia than it is expected by chance (Lee et al. 2008). The microsporidian gene cluster that is shared syntenic to the zygomycete sex-related locus was the most prominent example of conserved gene order. Its presence, together with the finding that microsporidia contain several genes required for meiosis, implies that microsporidia actually may have sex (Lee, Ni, et al. 2010). In summary, the analysis of gene order indicated that microsporidia share an exclusive common ancestry with the zygomycetes, and it was concluded that microsporidia evolved from ancient sexual fungi (Dyer 2008b; Lee et al. 2008; reviewed in Corradi and Keeling 2009).

Studies of both the evolutionary origins of HMG-driven sex determination and the phylogenetic position of the microsporidia hinge on the microsporidian gene cluster. It is therefore unfortunate that the evolutionary history of this gene cluster itself is not clear (Lee, Corradi, et al. 2010). Here, we perform a comprehensive analysis to unravel the phylogenetic relationships of the genes linked to the sex-determining transcription factor in zygomycetes and their counterparts in the microsporidia. Based on the results, we discuss the evolutionary history of the microsporidian gene clusters, as well as the implications for both the evolutionary ancestry of sex determination and the phylogenetic placement of the microsporidia. In a subsequent genome-wide analysis of gene order conservation, we carefully readdress the proposed sister-group relationship of microsporidia and zygomycetes.

Materials and Methods

Ortholog Search and Phylogeny Reconstruction

We predicted orthologs to the RNA helicases and the TPTs using InParanoid v. 3.0 (Berglund et al. 2008). For a relaxed stringency ortholog search, a standard reciprocal Blast search using NCBI Blast v. 2.2.13 was performed. The gene IDs for the identified orthologs together with the corresponding data sources are summarized in supplementary table 1 (Supplementary Material online). RNA helicase and TPT alignments were generated with MAFFT v. 6.833b (Katoh et al. 2005). Throughout all analyses, MAFFT was used with the options "maxiterate 1000 and localpair." The resulting multiple sequence alignments were then each used for tree reconstruction. Maximum likelihood (ML) tree reconstruction was performed with RAxML v. 7.2.2 (Stamatakis 2006), and branch support was assessed with 100 bootstrap

replicates. Bayesian tree reconstruction was performed with Phylobayes v. 2.3 (Lartillot and Philippe 2004) running two independent chains per data set. The chains were stopped after 84,000 generations (TPT) and 120,000 generations (RNA helicase), respectively, and we discarded the first 10,000 generations as burn-in. Convergence was confirmed with bpcomp from the Phylobayes package sampling every 10th tree (maxdiff: TPT: 0.02; RNA helicase: 0.08).

Analysis of Characteristic Sites in the Multiple Sequence Alignments

For the analysis of characteristic sites, we pursued the following strategy: We aligned the sequences individually for the four ortholog groups, TPT-S1 and TPT-S2 and RNA helicase-S1 and RNA helicase-S2. The corresponding S1 and S2 alignments were then combined with MAFFT using the option "addprofile." We then called a site characteristic if in the combined S1-S2 alignment the majority of sequences from one group share an amino acid or an insertion/deletion that is not seen in the respective other group. To assess whether the microsporidian sequences share more characteristic sites with the S1 or the S2 sequences, we added them individually to the appropriate S1-S2 alignment using MAFFT and the option "add."

For the analysis of characteristic sites in the DEXDc domain (SM00487), we downloaded the alignment for this domain from the SMART database (Letunic et al. 2009). The alignment was converted into a profile Hidden Markov Model with *hmmbuild* from the HMMER package v.3 (http://hmmer.janelia.org/). The Logo (Schuster-Bockler et al. 2004) for the pHMM was generated with the tool provided at http://www.sanger.ac.uk/cgi-bin/software/analysis/logomat-m.cgi (supplementary fig. 1, Supplementary Material online). The subsequences in the RNA helicases corresponding to the DEXDc domain were extracted and aligned with *hmmalign* using the option "trim." In the resulting pHMM alignment, the analysis of characteristic sites was performed as described above.

Identification of RNA Helicases, TPTs, and HMG Box Proteins

We identified putative RNA helicases, TPTs, and HMG-type transcription factors in *Batrachochytrium dendrobatidis*, *Phycomyces blakesleeanus*, and *Saccharomyces cerevisiae* by searching for proteins harboring the characteristic conserved functional domains. For the RNA helicase, we used the DEXDc SMART domain (SM00487), for the HMG type proteins the HMG_box PFAM domain (PF00505), and for the TPTs the TPT PFAM domain (PF03151). All three domains are present both in the proteins encoded in the sex-related cluster of *P. blakesleeanus* and in the microsporidian counterparts. Domain annotations of the proteins were performed as described in Koestler et al. (2010).

Analysis of Gene Order Conservation

The extent of gene order conservation to *Encephalitozoon cuniculi* was determined in two zygomycete taxa *P. blakesleeanus* and *Rhizopus oryzae*, as well as in the following species: *B. dendrobatidis* (Fungi; *Chytridiomycota*; *Chytridiomycetes*), *Sporobolomyces roseus* (Fungi; *Dikarya*; *Basidiomycota*; *Pucciniomycotina*; *Microbotryomycetes*), *Laccaria bicolor* (*Dikarya*; *Basidiomycota*; *Agaricomycotina*; *Agaricomycetes*), *Aspergillus niger* (*Dikarya*; *Ascomycota*; *Pezizomycotina*; *Eurotiomycetes*), *Homo sapiens* (*Metazoa*). The nonzygomycete species were chose to complement the fungal lineages whose gene order conservation with respect to *E. cuniculi* were already found to be not conserved (Lee et al. 2008), that is, *S. cerevisiae* (*Dikarya*; *Ascomycota*; *Saccharomycotina*; *Saccharomycetes*), *Ashbya gossypii* (*Dikarya*; *Ascomycota*; *Saccharomycotina*; *Saccharomycetes*), *Schizosaccharomyces pombe* (*Dikarya*; *Ascomycota*; *Taphrinomycotina*; *Schizosaccharomycetes*), *Neurospora crassa* (*Fungi*, *Dikarya*, *Ascomycota*, *Pezizomycotina*; *Sordariomycetes*), *Cryptococcus neoformans* (*Fungi*; *Dikarya*; *Basidiomycota*; *Agaricomycotina*; *Tremellomycetes*). Genome assemblies and annotated gene sets were downloaded from: Joint Genome Institute (http://genome.jgi-psf.org/euk_cur1.html): *A. niger*, *B. dendrobatidis*, *L. bicolor*, *P. blakesleeanus*, *S. roseus*; Broad Institute (http://www.broad.mit.edu/): *R. oryzae*; EBI (http://www.ebi.ac.uk/integr8): *E. cuniculi*; ENSEMBL (http://www.ensembl.org): *H. sapiens*. For taxa for which the order of the annotated genes was not readily provided for download, we mapped the coding sequences for the predicted genes to the genome assembly using BLAT (Kent 2002). The position of the best BLAT hit was taken as the gene position.

Results and Discussion

The Evolutionary History of the TPTs and the RNA Helicases

In syntenic gene clusters, each gene shares the evolutionary history of the entire gene cluster. Thus, the split between the sex-related region in zygomycetes and its counterpart in the microsporidia can be vicariously dated by analyzing the evolutionary relationships of the TPTs and RNA helicases, respectively. The HMG-type transcription factors were omitted from this analysis because they lack any phylogenetic information (Lee, Corradi, et al. 2010). In the following, we refer to the zygomycete sex-related gene cluster as "syntenic region 1" (S1) and to the microsporidia gene cluster as "syntenic region 2" (S2). Correspondingly, we refer to the respective genes as *TPT-S1/RNA helicase-S1* and as *TPT-S2/RNA helicase-S2* (fig. 1).

To start our analyses, we used the *P. blakesleeanus* (zygomycetes) S1 proteins to identify orthologs in *E. cuniculi* (microsporidia). Similarly, we searched for orthologs to the

Table 1
Ortholog Pairings for the *Phycomyces blakesleeanus* and *Encephalitozoon cuniculi* S1 and S2 Genes

	P. blakesleeanus[a] (Genomic Location)	*E. cuniculi* (Genomic Location)
TPT-S1	11516[b] (Scaffold 41)	—
RNA helicase-S1	80075[c] (Scaffold 41)	—
TPT-S2	4053 (Scaffold 4)	Q8SV84 (Chr VI)
	19565 (Scaffold 8)	
RNA helicase-S2	14395 (Scaffold 1)	Q8SRN8 (Chr VI)

Note.—Genes located in the sex-related gene cluster of zygomycetes and in its microsporidian counterpart are underlined.

[a] JGI Gene Id.

[b] Accession number ABX27908.1.

[c] Accession number ABX27910.1.

E. cuniculi S2 proteins in *P. blakesleeanus*. We chose the two species in which the S1 and S2 gene cluster had been initially described (Idnurm et al. 2008; Lee et al. 2008). InParanoid (Remm et al. 2001), one of the most reliable orthology prediction programs (Chen et al. 2007), was used for this purpose. No orthologs to the S1 proteins were found in *E. cuniculi*. In contrast, both S2 proteins have orthologs in *P. blakesleeanus*. The corresponding genes are, however, not located in the sex-determining region but resided on different scaffolds in the *P. blakesleeanus* genome assembly (table 1). Note that the results did not change when we reduced the stringency of the ortholog search by performing only a reciprocal best Blast hit search and omitted the additional filtering steps invoked by InParanoid (Remm et al. 2001). Thus, neither TPT-S1/TPT-S2 nor RNA helicase-S1/RNA helicase-S2 were identified as ortholog pairs.

We assessed next when during evolution the corresponding genes in the S1 and S2 clusters of zygomycetes and microsporidia have separated. A screen in 15 plant, animal, and fungal species for orthologs to each of the four genes resulted in four disjoint ortholog groups (c.f. supplementary table 1, Supplementary Material online). We combined all RNA helicases and all TPTs, respectively, and conducted ML tree reconstructions for both data sets. The resulting trees are shown in figure 2 (RNA helicases) and supplementary figure 2 (Supplementary Material online) (TPTs). In both trees, the S1 orthologs and the S2 orthologs are placed into two well-supported clades (RNA helicase: BS = 100, TPT: BS = 100). A complementary Bayesian analysis corroborated the results (RNA helicase: BPP = 1; TPT: BPP = 1; trees not shown). All four clades, corresponding to the four ortholog groups, contain sequences from animals, fungi, and plants. This indicates that the genes in the zygomycete sex-related region have separated from their microsporidian homologs already before the three eukaryotic kingdoms emerged.

To further corroborate that the zygomycete S1 genes and the microsporidian S2 genes are evolutionarily only very distantly related, we analyzed the protein sequence alignments. We first removed all microsporidian RNA helicases. We then aligned the S1 RNA helicases and the S2 RNA helicases separately and, subsequently, combined them using a profile-to-profile alignment. In the resulting alignment, we searched for evolutionarily conserved sites that characterize the S1 and the S2 ortholog groups. We called a site characteristic if in the combined alignment of two ortholog groups, the majority of sequences from one group share an amino acid or an insertion/deletion that is not seen in the respective other group. Three hundred and ninety characteristic sites distinguish the S1 RNA helicases from the S2 RNA helicases. We applied the same procedure to the TPTs and identified 58 characteristic sites. To assess whether the microsporidian sequences display any marked similarity with either the S1 or the S2 sequences, we aligned each of them to the corresponding combined S1-S2 alignment. This revealed that the *E. cuniculi* RNA helicase shares 197 characteristic sites with the S2 RNA helicases and only 24 with the S1 RNA helicases. Similarly, the *E. cuniculi* TPT shares 22 of the 58 characteristic sites with the S2 TPTs and 0 with the S1 TPTs. The same results were obtained with the other microsporidian sequences (table 2). Thus, the proteins encoded in the microsporidian gene cluster share a substantial extent of sequence conservation with the other S2 sequences. In contrast, they have virtually nothing in common with the S1 sequences.

We pursued the analysis of characteristic sites in greater depth exemplarily for the RNA helicases. We extracted the subsequences matching to the DEXDc SMART domain (Letunic et al. 2009), the functional domain of DEAD and DEAH box helicases, and performed a pHMM alignment. A section from this alignment is shown in figure 3, and the full alignment of the DEXDc SMART domain is provided as supplementary information (Supplementary Material online). Figure 3 shows clearly that the marked sequence conservation between the microsporidian RNA helicase and the S1 RNA helicases is present also in the functional domain of the proteins. The helicase domain of the RNA helicase encoded in the zygomycete sex-related locus contains two short sequence motifs IQGPPGTGKT and NHALDQF that are almost perfectly conserved among all sequences in the RNA helicase S1 group (green stars in fig. 3). A parsimony argument implies that these motifs were already present in the most recent common ancestor of these sequences. At the same alignment positions, two evolutionarily highly conserved motifs are also seen in the helicase domains from the S2 group, which however are distinct from the S1 motifs (*P. blakesleeanus*: GMARTGSGKT and RELALQT; red stars in fig. 3). The microsporidian sequences display slight variants of the S2 motifs.

The evolutionary history of the genes in the zygomycete sex-related region (gene cluster S1) and its shared syntenic-counterpart in the microsporidia (gene cluster S2) has been investigated before (Lee, Corradi, et al. 2010). However, the authors did not decisively conclude whether the corresponding genes in the two clusters are paralogs or extremely

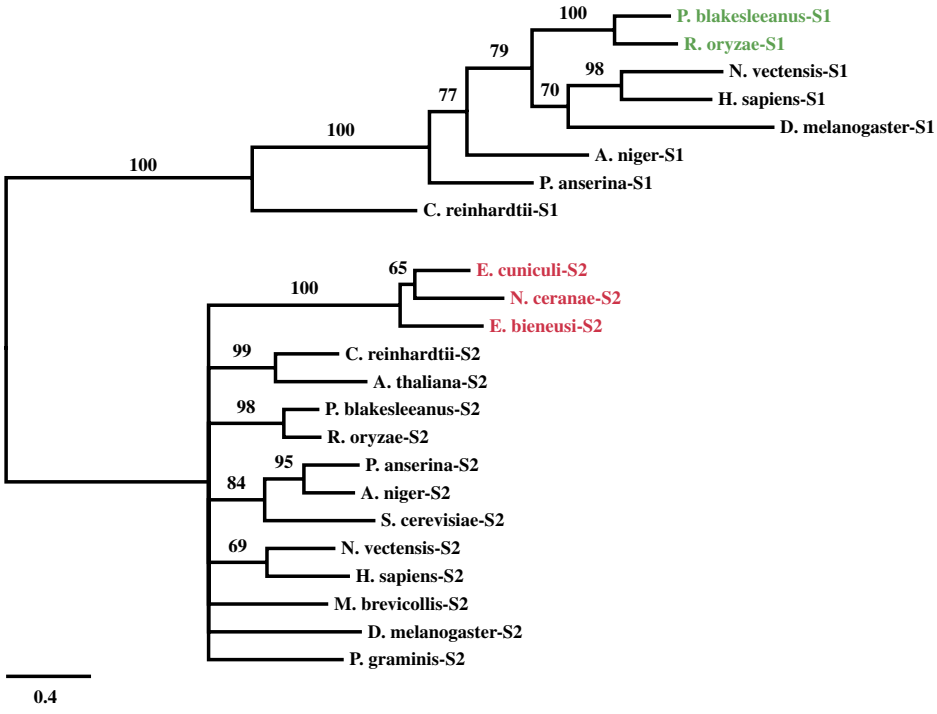

Fig. 2.—ML tree of the S1 and S2 RNA helicases. Sequences in the zygomycete sex-related region are labeled in red, and sequences in the corresponding region of the microsporidia are labeled in green. Branch labels denote bootstrap support values.

diverged orthologs. We followed a tripartite approach to solve this issue. Orthology predictions using both InParanoid (Berglund et al. 2008) and a less stringent reciprocal best Blast hit search failed to recognize the S1 RNA helicase of *P. blakesleeanus* and the S2 RNA helicase of *E. cuniculi* as orthologs, and the same applies to the TPTs. A phylogenetic tree reconstruction placed S1 and S2 sequences in distinct clades where each clade contained sequences from fungi, animals, and plants. This already suggests an early separation of the S1 and S2 genes that predates the split of microsporidia and fungi. However, the validity of conclusions drawn from both orthology assignment and phylogenetic tree reconstruction can be compromised by the high evolutionary rate particularly of microsporidian proteins (e.g., Brinkmann et al. 2005). Hence, we added the analysis of evolutionarily conserved characteristic sites as a third line of evidence. We found that the microsporidian sequences share substantially more characteristic sites with the S2 sequences then with the S1 sequences. This finding seamlessly integrates with the results from the ortholog search and the tree reconstruction. Thus, all evidences point toward a common ancestry of the microsporidian genes and the respective other S2 genes to the exclusion of the S1 genes. On the contrary, they are not compatible with the hypothesis that the microsporidian genes are extremely diverged orthologs of the genes in the zygomycete sex-related cluster, as it has been suggested before (Lee, Corradi, et al. 2010).

The Implications of Shared Synteny

Our analyses have revealed that both gene pairs, *RNA helicase-S1/S2* and *TPT-S1/S2*, separated early in eukaryotic evolution and long before the split of zygomycetes and microsporidia. How can this result be reconciled with the proposed shared synteny of the genomic regions the genes reside in? To do so, we need to assume that an ancestral TPT-HMG-RNA helicase gene cluster existed already in the

Table 2
Number of Characteristic Sites Conserved in the Microsporidian RNA Helicases and TPTs

	Number of Characteristic Sites[a]	Ortholog Group	*Encephalitozoon cuniculi*	*Antonospora locustae*	*Nosema ceranae*	*Enterocytozoon bieneusi*
RNA helicase	390	S1	24	—	22	17
		S2	197	—	194	206
TPT	58	S1	0	0	1	3
		S2	22	23	21	15

[a] Total number of characteristic sites distinguishing the S1 from the S2 sequences.

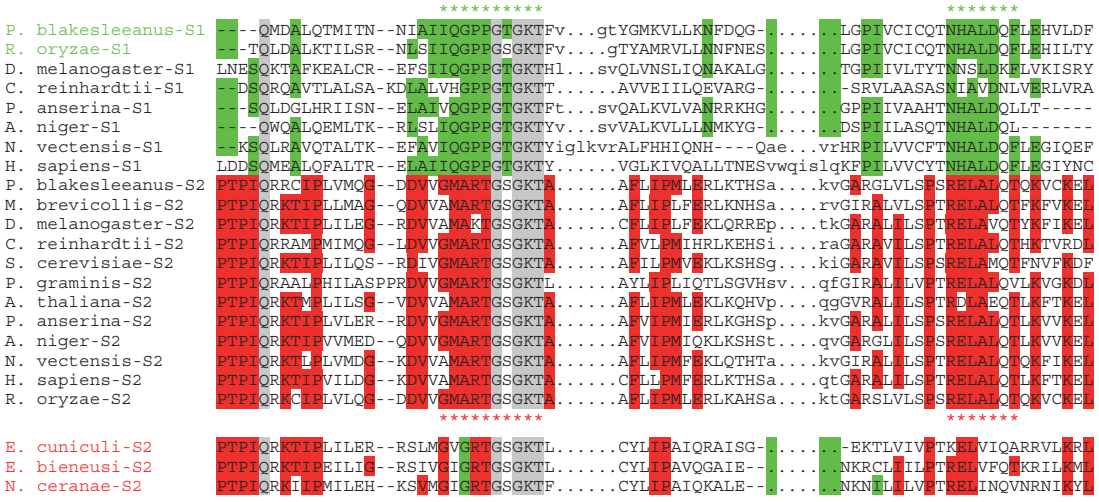

Fig. 3.—Section of the pHMM-guided multiple sequence alignment of the DEXDc domain in the S1 and S2 RNA helicases. Characteristic sites for the S1 RNA helicases are labeled in green, and for the S2, RNA helicases are labeled in red. The microsporidian sequences display almost exclusively characteristic sites of the S2-type RNA helicases. Amino acids in the gray shaded columns are conserved in all sequences and are specific for the DEXDc domain (c.f. supplementary fig. 1, Supplementary Material online). Dashes denote delete states in the pHMM alignment, and lower case letters opposed to dots denote insert states. Green and red sequence labels denote the sequences in the zygomycete sex-related cluster and its microsporidian counterpart, respectively. The green stars denote two evolutionarily conserved motifs of the RNA helicase in the sex-related region of *Phycomyces blakesleeanus*. The red stars denote the corresponding motif in the S2 RNA helicases. The full alignment is provided as supplementary information (Supplementary Material online).

common ancestor of plants, animals, and fungi. This gene cluster gave rise to two independently evolving copies in this primordial species. One copy each was retained with conserved gene order in microsporidia and zygomycetes, and the second copy was reciprocally lost in the two lineages. In all other taxa analyzed so far, both copies of the ancestral gene cluster were lost. Supplementary figure 3 (Supplementary Material online) summarizes one possible evolutionary scenario (see also Lee, Ni, et al. 2010). Thus, conditioned on the shared synteny assumption, we provide now for the first time evidence that the evolutionary history of the zygomycete MAT locus can be traced back to the early days of eukaryote evolution. It would be therefore old enough to represent the archetype of fungal and animal sex determination.

However, a considerable number of evolutionary events need to be assumed to uphold the initial assumption of shared synteny (c.f. supplementary fig. 3, Supplementary Material online). It is, therefore, worthwhile to consider an alternative hypothesis. The two gene clusters in zygomycetes and microsporidia may have been assembled twice and independently during evolution and are not shared syntenic. In this case, the presence of a microsporidian HMG-type transcription factor flanked by a TPT, and an RNA helicase allows no conclusions about the evolutionary history of HMG-driven sex determination. Although convergent evolution appears on the first sight unlikely, we will now show that it is not implausible. *Phycomyces blakesleeanus* has 132 different proteins with a DEXDc smart domain (Letunic et al. 2009), the characteristic feature of the RNA helicase-S1. Further 30 proteins contain a HMG box and 17 TPTs exist. The numbers for *B. dendrobatidis* (chytridiomycetes) and *S. cerevisiae* (ascomycetes) are similar (RNA helicases: 115/123; HMG: 10/9; TPT: 8/11). This indicates that these genes were as abundant in the last common ancestor of all fungi. Microsporidia evolved from the ancestor shared with the fungi by undergoing a massive genome compaction and an associated loss of genes (Katinka et al. 2001). Still *E. cuniculi* has retained 48 helicases, 3 TPTs, and 2 HMG type proteins. It can be easily imagined that the reorganization of the microsporidian genome during its evolution has just by chance placed any of the genes encoding RNA helicases, TPTs, and HMG-type transcription factors next to each other. By that a gene cluster emerged resembling that of the zygomycete sex-related region.

At the moment, it is impossible to decide which of the two scenarios, ancient relationships or convergent evolution, applies to the two gene clusters. Presumably, only an in-depth functional analysis of the individual proteins in microsporidia, with a focus on the HMG-type transcription factor, will help to shed further light on this matter.

Are Microsporidia and Zygomycetes Monophyletic?

How do our findings relate to the debate about the phylogenetic position of microsporidia? The conservation of gene order and, in particular, the presence of the microsporidian gene cluster resembling the zygomycete sex-related locus

Table 3
Number of *Encephalitozoon cuniculi* Gene Pairings Recovered in Six Fungi and Humans

		Number of Intervening Genes[a]								
		0	1	2	3	4	5	6	7	8	9
E. cuniculi: 674[b]	*Batrachochytrium dendrobatidis*	3		1	1						
	Aspergillus niger	1			1		1				
	Laccaria bicolor	1						1			
	Sporobolomyces roseus	1			1			1			
	Rhizopus oryzae	2	1	1							
	Phycomyces blakesleeanus						1				
	Homo sapiens	1				1					1

[a] Number of intervening genes between two orthologs to an *E. cuniculi* gene pair.
[b] Gene pairs separated by no more than three genes with orthologs in six fungal taxa and *H. sapiens*.

has served as argument to place the microsporidia next to the zygomycetes in the fungal tree of life (Lee et al. 2008; Lee, Ni, et al. 2010). However, we have shown that the suggested shared synteny traces the two gene clusters back to an ancient gene cluster in the common ancestor of plants, animals, and fungi. In cladistic terms, they represent a shared ancestral character or a plesiomorphy. Plesiomorphies are phylogenetically not informative (Hennig 1966). Hence, they cannot serve as supporting evidence for the proposed monophyly of microsporidia and zygomycetes (but see Lee et al. 2008; Lee, Corradi, et al. 2010). This emphasizes that phylogenetic inferences based on gene order conservation are problematic when the exact evolutionary relationships of the genes remain uninvestigated. Unfortunately, the only quantitative analysis of gene order conservation to determine the phylogenetic position of microsporidia used a unidirectional BlastP search (*e*-value cutoff 10×10^{-5}) for homology inference (Lee et al. 2008). A comparison between several orthology prediction methods has shown that orthology assignments based on unidirectional BlastP searches are wrong in 50% of the cases (Chen et al. 2007). This bears the risk that a considerable fraction of the identified zygomycete–microsporidia gene pairs comprise paralogs. To assess whether this has any consequences for the conclusions of this study, we reinvestigated the extent of gene order conservation between microsporidia and zygomycetes. In brief, we used InParanoid (Remm et al. 2001) for orthology prediction. In contrast to the unidirectional BlastP search, InParanoid has a reported false positive rate of only 7% (Chen et al. 2007). We established the evolutionary relationships between the genes of *E. cuniculi* and two zygomycete species, *P. blakesleeanus* and *R. oryzae*. Four further fungal species (*B. dendrobatidis*, *S. roseus*, *L. bicolor*, *A. niger*) and human were analyzed to investigate whether the extent of gene order conservation to *E. cuniculi* varies between species. Note, that we chose the nonzygomycete fungi to represent major fungal lineages whose gene order conservation with respect to *E. cuniculi* has so far not been investigated (c.f. Lee et al. 2008). For 461 *E. cuniculi* genes, an ortholog was present in all

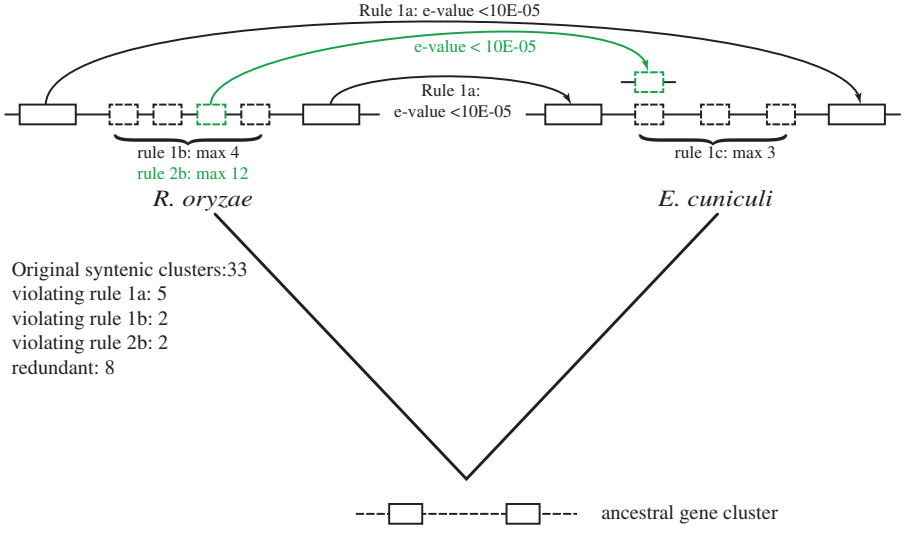

Fig. 4.—Reanalysis of the proposed shared syntenic gene clusters between *Rhizopus oryzae* and *Encephalitozoon cuniculi*. Two decision rules were used to assign shared synteny (c.f. fig. S5 in Lee et al. 2008). Rule 1 requires that (1a) two *R. oryzae* genes have a Blast hit with an e-value $< 10 \times 10^{-05}$ in *E. cuniculi*, (1b) the two *R. oryzae* genes must be separated by no more than four genes, and (1c) the corresponding *E. cuniculi* genes must be separated by no more than three genes. Rule 2 extends rule 1 if one of the intervening *R. oryzae* genes also has a Blast hit (*e*-value $< 10 \times 10^{-05}$) in *E. cuniculi*. In this case, the *R. oryzae* genes must be separated by no more than 12 genes (2b). Please note that in the original publication, the decision rules where described for a search in the opposite direction, that is, with the *E. cuniculi* proteins as query. However, the data presented in table S1 of Lee et al. 2008 (c.f. also supplementary table 2, Supplementary Material online) are not compatible with this direction of the search. Hence, we adjusted the decision rules to fit the data.

seven taxa. In *E. cuniculi*, these 461 genes are arranged in 674 gene pairs with no more than three genes in-between. We then considered a microsporidian gene pair as conserved if its orthologs in the nonmicrosporidian species are separated by no more than nine intervening genes. The results are summarized in table 3. From the 674 *E. cuniculi* gene pairs, only 4 were recovered in *R. oryzae*, 5 were present in *B. dendrobatidis*, a chytridiomycete, and 3 in humans. In essence, no marked conservation of gene order between *E. cuniculi* and zygomycetes is seen. However, again in the light that orthology prediction for the fast evolving microsporidia is hard, our approach bears the risk of being overly stringent. We thus may lack the sensitivity for a meaningful analysis of gene order conservation. To address this point, we reanalyzed the existing data that were obtained with the unidirectional Blast searches (table S1 in Lee et al. 2008). The results are summarized in figure 4 and in supplementary table 2 (Supplementary Material online). In their "relaxed stringency" analysis, Lee et al. (2008) found 33 clusters with conserved gene order in *E. cuniculi* and *R. oryzae*. Of these, clusters 5 do not fulfill the *e*-value cutoff of 10×10^{-5} in the original data. Further 4 clusters exceed the maximally allowed number of intervening genes in *R. oryzae*. Of the remaining 24 clusters, 13 *R. oryzae* clusters point to only five clusters in *E. cuniculi*. The corresponding clusters must have duplicated on the *R. oryzae* lineage after the split from the microsporidia. Hence, they can be counted only once each. This reduces the number of independent shared syntenic regions between the two species to 16 what would be expected by chance (Lee et al. 2008). Thus, a proper analysis of Lee et al.'s (2008) data provides no evidence that the gene order is more conserved between microsporidia and zygomycetes than between microsporidia and any other fungal taxon or even humans. As a consequence, the proposed placement of microsporidia as a sister to the zygomycetes receives no support by the data. The question remains therefore open where to confidently place this enigmatic taxon in the fungal tree of life.

In summary, our study has revealed what can and what cannot be inferred from the observation that microsporidia harbor a gene cluster closely resembling the sex-related region of zygomycetes. If we take shared synteny for granted, our results trace the zygomycete sex-related region back to the early days of eukaryote evolution. It may therefore indeed comprise the archetype of animal and fungal sex determination. However, the evolutionary scenario that has to be invoked is complex. Thus, sacrificing the shared synteny assumption may lead to a more parsimonious hypothesis, that is, that the two gene clusters arose independently through convergent evolution. Independent of the true evolutionary relationships of the two gene clusters, however, one observation stands out. Their presence in zygomycetes and microsporidia represent, at best, a plesiomorphy and provide no information about the phylogenetic relationships of zygomycetes and microsporidia. As there is no further evidence for a significant conservation of gene order between the two taxa, the proposed alliance of microsporidia and zygomycetes remains speculative.

Supplementary Material

Supplementary figures 1–3 and tables 1–2 are available at *Genome Biology and Evolution* online (http://www.gbe.oxfordjournals.org/).

Acknowledgments

We acknowledge financial support by the Wiener Wissenschafts-, Forschungs- und Technologie Fonds and from the DFG priority program SPP 1174 Deep Metazoan Phylogeny. We thank Arndt von Haeseler for helpful discussion and for critically reading the manuscript.

Literature Cited

Berglund AC, Sjolund E, Ostlund G, Sonnhammer EL. 2008. InParanoid 6: eukaryotic ortholog clusters with inparalogs. Nucleic Acids Res. 36:D263–D266.

Brinkmann H, van der Giezen M, Zhou Y, Poncelin de Raucourt G, Philippe H. 2005. An empirical assessment of long-branch attraction artefacts in deep eukaryotic phylogenomics. Syst Biol. 54:743–757.

Butler G, et al. 2004. Evolution of the MAT locus and its Ho endonuclease in yeast species. Proc Natl Acad Sci U S A. 101: 1632–1637.

Cavalier-Smith T. 1986. The kingdoms of organisms. Nature 324: 416–417.

Chen F, Mackey AJ, Vermunt JK, Roos DS. 2007. Assessing performance of orthology detection strategies applied to eukaryotic genomes. PLoS One. 2:e383.

Corradi N, Keeling PJ. 2009. Microsporidia: a journey through radical taxonomical revisions. Fungal Biol Rev. 23:1–8.

Dietrich FS, et al. 2004. The Ashbya gossypii genome as a tool for mapping the ancient Saccharomyces cerevisiae genome. Science 304:304–307.

Dyer PS. 2008a. Evolutionary biology: genomic clues to original sex in fungi. Curr Biol. 18:R207–R209.

Dyer PS. 2008b. Evolutionary biology: microsporidia sex—a missing link to fungi. Curr Biol. 18:R1012–R1014.

Gill EE, Fast NM. 2006. Assessing the microsporidia–fungi relationship: combined phylogenetic analysis of eight genes. Gene 375:103–109.

Haber JE. 1998. Mating-type gene switching in Saccharomyces cerevisiae. Annu Rev Genet. 32:561–599.

Haqq CM, King CY, Donahoe PK, Weiss MA. 1993. SRY recognizes conserved DNA sites in sex-specific promoters. Proc Natl Acad Sci U S A. 90:1097–1101.

Hennig W. 1966. Phylogenetic systematics. Urbana (IL): University of Illinois Press.

Idnurm A, Walton FJ, Floyd A, Heitman J. 2008. Identification of the sex genes in an early diverged fungus. Nature 451:193–196.

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

Katinka MD, et al. 2001. Genome sequence and gene compaction of the eukaryote parasite Encephalitozoon cuniculi. Nature 414:450–453.

Katoh K, Kuma K, Toh H, Miyata T. 2005. MAFFT version 5: improvement in accuracy of multiple sequence alignment. Nucleic Acids Res. 33:511–518.
Keeling P. 2009. Five questions about microsporidia. PLoS Pathog. 5:e1000489.
Keeling PJ. 2003. Congruent evidence from alpha-tubulin and beta-tubulin gene phylogenies for a zygomycete origin of microsporidia. Fungal Genet Biol. 38:298–309.
Keeling PJ, Luker MA, Palmer JD. 2000. Evidence from beta-tubulin phylogeny that microsporidia evolved from within the fungi. Mol Biol Evol. 17:23–31.
Kent WJ. 2002. BLAT-the BLAST-like alignment tool. Genome Res. 12:656–664.
Koestler T, von Haeseler A, Ebersberger I. 2010. FACT: functional annotation transfer between proteins with similar feature architectures. BMC Bioinformatics. 11:417.
Lartillot N, Philippe H. 2004. A Bayesian mixture model for across-site heterogeneities in the amino-acid replacement process. Mol Biol Evol. 21:1095–1109.
Lee SC, et al. 2008. Microsporidia evolved from ancestral sexual fungi. Curr Biol. 18:1675–1679.
Lee SC, et al. 2010. Evolution of the sex-related locus and genomic features shared in microsporidia and fungi. PLoS One. 5:e10539.
Lee SC, Ni M, Li W, Shertz C, Heitman J. 2010. The evolution of sex: a perspective from the fungal kingdom. Microbiol Mol Biol Rev. 74:298–340.
Lengeler KB, et al. 2002. Mating-type locus of Cryptococcus neoformans: a step in the evolution of sex chromosomes. Eukaryot Cell. 1:704–718.
Letunic I, Doerks T, Bork P. 2009. SMART 6: recent updates and new developments. Nucleic Acids Res. 37:D229–D232.
Remm M, Storm CE, Sonnhammer EL. 2001. Automatic clustering of orthologs and in-paralogs from pairwise species comparisons. J Mol Biol. 314:1041–1052.
Schuster-Bockler B, Schultz J, Rahmann S. 2004. HMM Logos for visualization of protein families. BMC Bioinformatics. 5:7.
Stamatakis A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics. 22:2688–2690.
Swanson WJ, Vacquier VD. 2002. The rapid evolution of reproductive proteins. Nat Rev Genet. 3:137–144.
Thomarat F, Vivares CP, Gouy M. 2004. Phylogenetic analysis of the complete genome sequence of Encephalitozoon cuniculi supports the fungal origin of microsporidia and reveals a high frequency of fast-evolving genes. J Mol Evol. 59:780–791.
Thomas JO, Travers AA. 2001. HMG1 and 2, and related "architectural" DNA-binding proteins. Trends Biochem Sci. 26:167–174.
Vossbrinck CR, Maddox JV, Friedman S, Debrunner-Vossbrinck BA, Woese CR. 1987. Ribosomal RNA sequence suggests microsporidia are extremely ancient eukaryotes. Nature. 326:411–414.

Associate editor: Bill Martin