Current Biology, Vol. 15, 185–191, January 26, 2005, ©2005 Elsevier Ltd All rights reserved. DOI 10.1016/j.cub.2005.01.003

A Phylogenomic Inventory of Meiotic Genes: Evidence for Sex in *Giardia* and an Early Eukaryotic Origin of Meiosis

Marilee A. Ramesh,[2,3,4] Shehre-Banoo Malik,[1,3,4] and John M. Logsdon, Jr.[1,4,*]
[1]Department of Biological Sciences
Roy J. Carver Center for Comparative Genomics
University of Iowa
Iowa City, Iowa 52242
[2]Department of Biology
Roanoke College
Salem, Virginia 24153

Summary

Sexual reproduction in eukaryotes is accomplished by meiosis, a complex and specialized process of cell division that results in haploid cells (e.g., gametes). The stereotypical reductive division in meiosis is a major evolutionary innovation in eukaryotic cells [1], and delineating its history is key to understanding the evolution of sex [2]. Meiosis arose early in eukaryotic evolution, but when and how meiosis arose and whether all eukaryotes have meiosis remain open questions [3]. The known phylogenetic distribution of meiosis comprises plants, animals, fungi, and numerous protists [4]. Diplomonads including *Giardia intestinalis* (syn. *G. lamblia*) are not known to have a sexual cycle [5]; these protists may be an early-diverging lineage [6] and could represent a premeiotic stage in eukaryotic evolution. We surveyed the ongoing *G. intestinalis* genome project data [7] and have identified, verified, and analyzed a core set of putative meiotic genes—including five meiosis-specific genes—that are widely present among sexual eukaryotes. The presence of these genes indicates that: (1) *Giardia* is capable of meiosis and, thus, sexual reproduction, (2) the evolution of meiosis occurred early in eukaryotic evolution, and (3) the conserved meiotic machinery comprises a large set of genes that encode a variety of component proteins, including those involved in meiotic recombination.

Results and Discussion

Eukaryotic Evolution and Meiotic Origins

Our central goal was to determine whether genes encoding meiotic proteins were present in *Giardia*, both as indicators of *Giardia*'s potential to undergo meiosis and sexual reproduction and as markers for the evolution of meiosis itself. We have taken an inventory of *Giardia* genes that are clear homologs of genes with known roles in meiosis in other eukaryotes. Considering *Giardia* as an exemplar protist, a more complete picture of the phylogenetic distribution of both meiotic genes, and thus meiosis itself, emerges. *Giardia* has unambiguous homologs of well-known meiotic genes. The direct implications are that *Giardia* is, or was recently, capable of sexual reproduction and, thus, does not represent an ancient eukaryotic lineage that diverged before meiosis arose. Instead, the origin of meiosis predates the divergence of *Giardia* in eukaryotic evolution.

In defining a "core meiotic recombination machinery" and providing the first extensive review of meiotic genes in an explicitly comparative context, Villeneuve and Hillers [2] recently argued that "the very essence of sex is meiotic recombination." Molecular mechanisms of meiosis are being elucidated from diverse eukaryotic model systems, providing insight into the origin and evolution of eukaryotic sex through conserved meiotic components. However, such systematic genetic and/or genomic approaches have only been applied to a phylogenetically-restricted set of eukaryotes—animals, fungi, and plants (herein, AFP)—precluding protists with only scant information on meiosis [8]. Phylogenetic analyses of rRNA and proteins show that eukaryotic diversity is largely represented by protists [6, 9–12], and a comprehensive comparative analysis of meiosis must consider protists [13] to determine if meiosis is universal among eukaryotes. Protists reveal evidence of both evolutionary conservation and variation in the meiotic machinery; both unconventional meiosis and no meiosis have been reported among protists [14]. Indeed, sexuality is variant—even deviant—among the protists [8, 15]; whether they possess meiotic machinery homologous to that in AFP is unknown because underlying molecular mechanisms have generally not been determined because of the lack of genetic tools. Recent protist genome sequencing projects have allowed access to such genetic data, making comparative genomic studies feasible. Among known meiotic genes, a list of genes central to meiosis can be compiled and used to identify homologs of these genes in a more diverse set of eukaryotes, including the earliest branches on the eukaryotic tree of life [13]. Cavalier-Smith [3] recently offered a synopsis of the origin of meiotic recombination machinery with available, albeit limited, information.

A genome project for *G. intestinalis* is nearly complete [7, 16]. This pathogen causes giardiasis, a diarrheal disease contracted in mammalian hosts by ingesting contaminated water [5]. Phylogenetic analyses of SSU rRNA and of other genes indicate that diplomonads are among the deepest divergences in the eukaryotic lineage [17–19]. Trees from different genes have not always placed *Giardia* on the deepest branch, but they generally put *Giardia* among the most early-diverging eukaryotes [20, 21]. The deep placement of diplomonads depends on the root position, and most analyses employ a prokaryotic outgroup root. Recent analyses that define the root by gene fusions suggest that a grouping of animals, fungi, and some amoebae may represent the deepest eukaryotic split, with plants and most protists (including *Giardia*) placed on the other side of this ur-divide [22, 23]. If true, comparisons between AFP would be largely sufficient to diagnose the generalities and ancestral

*Correspondence: john-logsdon@uiowa.edu
[3]These authors contributed equally to this work.
[4]The research for this report was begun in the Department of Biology at Emory University, Atlanta, Georgia 30322.

states, including the presence of meiosis, early in eukaryotic evolution. Although this scenario is possible [12], current evidence is simply not sufficient to preclude *Giardia* from among the earliest diverging lineages of eukaryotes; recent multigene analyses support diplomonads with parabasalids (together, two major groups of the "Excavate" protists [24]) as the deepest eukaryotic lineage [21].

If basal, *Giardia* could provide a phylogenetic key to understanding the origin and evolution of meiosis [25], and it has been suggested [26] that *Giardia* may represent an older group of "ancient asexuals" than the famous bdelloid rotifers. *Giardia* has many characteristics common among eukaryotic cells (e.g., a nucleus with a nuclear membrane, a cytoskeleton, and an endomembrane system), but it lacks features in most eukaryotes (e.g., nucleoli, peroxisomes, and mitochondria) [13, 27]. A relict organelle in *Giardia* appears to be mitochondrial [28], and mitochondrial genes are present in its nucleus [29]. *Giardia* has long been considered to be asexual because there is no evidence to date for either a meiotic division cycle or for genetic recombination [5, 30]. However, DNA content studies reveal that "during excystation, the recently excysted cell divides twice without DNA replication … (which) is therefore reminiscent of meiosis" [31], and some genetic evidence consistent with (but not unambiguously attributable to) sex in *Giardia* has been reported [32]. Clonal (asexual) population structure for *Giardia* has been inferred from fixed heterozygosity, deviation from Hardy-Weinberg expectations, and absence of recombinant genotypes [33]. Finally, a previous inspection of the *Giardia* genome data recently concluded "it is likely that sexual reproduction, if any, did not play a major role in shaping its genome" [34]. One possible interpretation of these observations is that diplomonads such as *Giardia* represent an "intermediate" (i.e., premeiotic) stage in eukaryotic evolution [35]. We have evaluated this compelling hypothesis and can now reject it.

Genes and Their Functions at the Core of Meiosis

We used a phylogenomic approach [36] to identify and validate a core group of meiotic genes as indicators for the presence or absence of meiosis. Our selection of meiotic genes was based on previous studies in eukaryotic model systems—mostly AFP. We chose genes having major meiotic functions in AFP and conservation in sequence and function [2]. Gene products with similar important roles in meiosis among known models have the greatest potential as meiotic indicators in diverse eukaryotes. Readily detectable protein sequence conservation among these genes (~15% amino acid identity) was needed to identify potential homologs in evolutionarily distant species. We included some "meiosis specific" genes—null mutations in all of these genes (in *S. cerevisiae,* at least) are defective *only* in meiosis, and most genes do not function outside of meiosis, with some possible exceptions in mammals (Hop2 [37] and Mnd1 [38]). The genes included in this survey, along with brief descriptions of their functions in meiosis, are given in Table 1.

The "core meiotic recombination machinery" [2] is comprised of *Spo11, Rad50/Mre11, Dmc1, Rad51, Msh4/Msh5*, and *Mlh1*. We expanded a list of "core meiotic genes" by including additional single genes and meiotic members of multigene families. The added genes encode a synaptonemal complex protein (Hop1), recombination proteins (Hop2, Mnd1, and Rad52), and additional members (i.e., paralogs) of the *Msh* and *Mlh* gene families, homologs of bacterial *mutS* and *mutL* [39]. Meiotic roles for these genes have been characterized in some AFP, but usually not in protists. Only some of the proteins are meiosis specific—Spo11, Hop1, Hop2, Mnd1, Dmc1, Msh4, and Msh5—but they all play key roles in meiosis. These genes would be expected in organisms capable of meiotic sex. Proteins encoded by these additional genes perform critical functions during the early stages of meiosis and are conserved among AFP and several protists.

Meiotic Genes in *Giardia* Suggest Sexuality

These meiotic genes (Table 1) were used to survey the raw sequence data at the *Giardia* genome project database (http://www.mbl.edu/giardia; [7]). Ours is not a comprehensive list of every potential meiotic gene in *Giardia*. Instead, we sought homologs of genes shown to have crucial roles in meiosis conserved in AFP. All genes were identified with a rigorous bioinformatic procedure, and their presence in the *Giardia* genome validated by cloning and sequencing of PCR products from genomic DNA. We discovered only minor sequence differences between our clones and draft sequences (see the Supplemental Data available with this article online). The cloned genes, with multiple sequence alignments and rigorous phylogenetic analyses (see below and Supplemental Data), all indicate that these genes are complete, bona fide homologs of meiotic genes.

Table 2 shows that the *Giardia* genome encodes homologs of key genes required for meiosis. Homologs of five meiosis-specific genes are present in *Giardia*: *Dmc1*, *Spo11*, *Mnd1*, *Hop1*, and *Hop2*. Proteins encoded by these genes are not known to function outside of meiosis among AFP species. Of meiosis-specific genes surveyed, *Giardia* does not appear to encode *mutS* homologs *Msh4* and *Msh5*. Whether this is due to the loss of these genes or later evolutionary origins is unclear; the sparse distribution of *Msh4/5* genes indicates some losses. Five meiosis-specific genes in *Giardia* in conjunction with other known meiotic genes provide strong evolutionary evidence for meiosis in the putatively asexual diplomonads. The presence of meiotic genes in *Giardia* and in the other protists surveyed (see below) extends the diversity of eukaryotes known to contain these genes. The collective presence of multiple genes that work together in meiotic recombination strongly suggests that *Giardia* has or very recently had the capacity to undergo meiosis and, thus, sexual reproduction. The possibility that *Giardia* undergoes meiosis warrants further investigation; experiments will be required to ascertain whether *Giardia* is capable of meiotic sexual reproduction, but the comparative genomic data presented here clearly support this testable hypothesis. Alternative hypotheses for the retention of meiosis genes in *Giardia* require that they function in nonmeiotic

Table 1. Core Meiotic Genes and Some Key Functions of Their Encoded Proteins in Meiosis

Gene	Protein Function(s)
*Spo11**	Transesterase; creates DNA double-strand breaks (DSBs) in meiosis I
Mre11	3′-5′ dsDNA exonuclease and ssDNA endonuclease; forms complex with Rad50 and Xrs2/Nbs1
Rad50	ATPase, DNA binding protein; in a complex with Mre11/Xrs2, holds broken DNA ends together while Mre11 trims
*Hop1**	Synaptonemal complex protein; binds DSBs and oligomerizes during meiotic prophase I
*Hop2**	Forms a complex with Mnd1 to ensure accurate and efficient homology searching during pachytene of meiotic prophase I
*Mnd1**	With Hop2, functions after meiotic DSB formation and is required for stable heteroduplex DNA formation and interhomolog repair
Rad52	Binds DSBs and initiates assembly of meiotic recombination complexes
*Dmc1**	Homolog of strand exchange protein Rad51; promotes interhomolog recombination
Rad51	With Dmc1, catalyzes homologous DNA pairing and strand exchange
*Msh4**	Forms heterodimer with Msh5; interacts with Mlh1/Mlh3; recombination crossover control
*Msh5**	Forms heterodimer with Msh4; interacts with Mlh1/Mlh3; recombination crossover control
Msh2	Forms a heterodimer with Msh3 or Msh6
Msh6	Forms a heterodimer with Msh2; binds base mismatches
Mlh1	Mismatch repair and promotion of meiotic crossing over; interacts with Msh2/Msh6 and Msh4/Msh5; forms heterodimers with Mlh2, Mlh3, and Pms1
Mlh2	Forms a heterodimer with Mlh1; interacts with Msh2/Msh3 and Msh2/Msh6
Mlh3	Forms a heterodimer with Mlh1; interacts with Msh2/Msh3 and Msh2/Msh6 for mismatch repair or with Msh4/Msh5 to promote meiotic crossovers
Pms1	Forms heterodimer with Mlh1 for repair of heteroduplex DNA; interacts with Msh2/Msh3

*denotes meiosis-specific genes. See Supplemental Data for exemplar references for each gene.

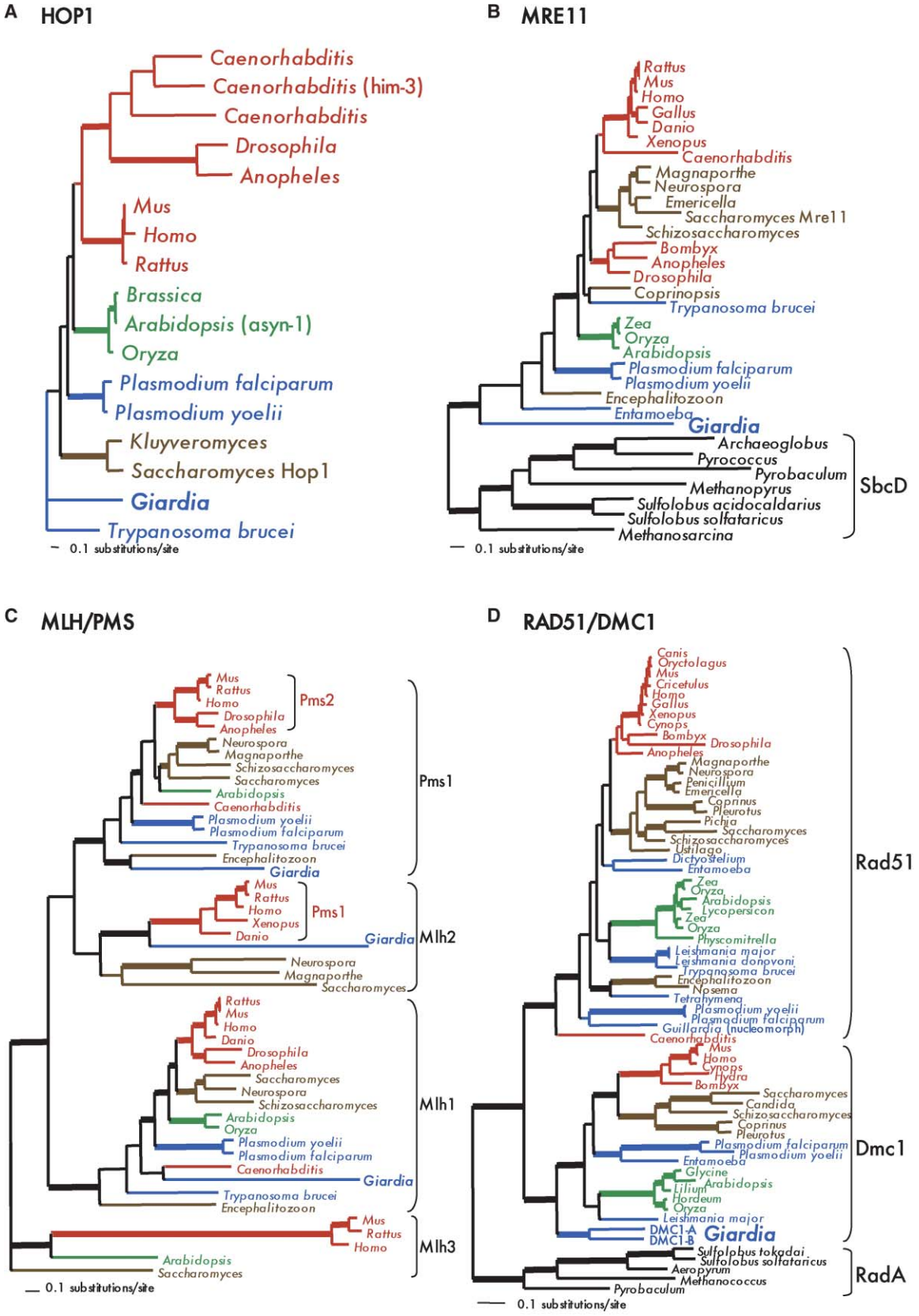
processes, present either from the loss of meiosis (derived asexuality) or prior to its invention (ancestral asexuality). As the numbers of meiotic genes mount, these arguments are increasingly problematic; each meiosis-specific gene present in *Giardia* requires independent loss or gain of meiotic functions in all other eukaryotes known.

The strong inference of meiosis being present is

Table 2. Phylogenetic Distribution among Eukaryotes of Core Meiotic Genes and the Identities of Their Prokaryotic Homologs

	EUKARYOTES	Spo11	Mre11	Rad50	Hop1	Hop2	Mnd1	Rad52	Dmc1	Rad51	Msh4, Msh5	Msh2, Msh6	Mlh1	Mlh2	Mlh3	Pms1
PROTISTS	***Giardia***	**+S**	+S	+S	**+S**	**+P**	**+S**	+S	**+S(2)**	−	**−, −**	+S, +S	+S	+S	−	+S
	Trypanosoma/ Leishmania	**+P**	+P	+n	**+P**	**+P**	**+n**		**+P**	+P		+P, +P	+P		[+P]	+P
	Entamoeba	**+P**	+P	+n		**+n**	**+n**	+n	**+P**	+P	**+n, +n**	+n, +n	+n		+n	+n(2)
	Plasmodium	**+P**	+P	+P	**+P**	**+P**	**+P**	−	**+P**	+P	**−, −**	+P(2), +P	+P	−	−	+P
ANIMALS	*Homo*	**+P**	+P	+P	**+P**	**+P**	**+P**	+P	**+P**	+P	**+P, +P**	+P, +P	+P	+P	+P	+P
	Mus	**+P**	+P	+P	**+P**	**+P**	**+P**	+P	**+P**	+P	**+P, +P**	+P, +P	+P	+P	+P	+P
	Drosophila	**+P**	+P	+P	**+P**	**−**	**−**	−	**−**	+P	**−, −**	+P, +P	+P	−	−	+P
	Anopheles	**+P**	+P	+P	**+P**	**−**	**−**	−	**−**	+P	**+P, −**	+P, +P	+P	−	−	+P
	Caenorhabditis	**+P**	+P	+P	**+P**	**−**	**−**	−	**−**	+P	**+P, +P**	+P, +P	+P	−	−	+P
FUNGI	*Saccharomyces*	**+P**	+P	+P	**+P**	**+P**	**+P**	+P	**+P**	+P	**+P, +P**	+P, +P	+P	+P	+P	+P
	Schizosaccharo.	**+P**	+P	+P	**−**	**+P**	**+P**	+P	**+P**	+P	**−, −**	+P, +P	+P	−	−	+P
	Neurospora	**+P**	+P	+P	**+P**	**−**	**−**	+P	**−**	+P	**(+n), +P**	+P, +P	+P	+P	(+n)	+P
	Encephalitozoon	**+P**	+P	+P	**−**	**+P**	**+P**	+P	**−**	+P	**−, −**	+P, +P	+P	−	−	+P
PLANTS	*Arabidopsis*	**+P(3)**	+P	+P	**+P**	**+P**	**+P**	−	**+P**	+P	**+P, +P**	+P, +P(2)	+P	−	+P	+P
	Oryza/Zea	**+P(3)**	+P	+P	**+P**	**+P**	**+P**	−	**+P**	+P(2)		+P, +P(2)	+P			+n
	ARCHAEA	TopoVI	SbcD	SbcC	−	−	−	−	RadA		MutS		MutL			
	BACTERIA	−			−	−	−	−	RecA							

Genes encoding meiosis-specific proteins are highlighted in grey and bold. The presence (+) of homologous genes is based on data obtained by BLAST searches of protein (P) and nucleotide (n) sequence databases at NCBI (and TIGR for *Entamoeba*) and by cloning/sequencing (S) of selected genes (from *Giardia*). If more than one gene is present, copy number is indicated in parentheses. Genes apparently absent from completed genome sequencing projects are indicated by minus (−), and putatively missing data (from unfinished genomes) are left blank. Protein homology was inferred by multiple sequence alignment and Bayesian phylogenetic analyses (see Figure 1; Supplemental Data). Putatively homologous nucleotide sequences (from unnannotated contigs or unassembled reads) were initially identified by tBLASTn and then verified by alignments and preliminary phylogenies of their inferred translations (trees not shown). The uncertain phylogenetic placement of a *Leishmania Mlh3* homolog is indicated by brackets, and putative *Msh4* and *Mlh3* genes in *Neurospora* [53] are given in parentheses.


A
HOP1
Caenorhabditis
Caenorhabditis (him-3)
Caenorhabditis
Drosophila
Anopheles
Mus
Homo
Rattus
Brassica
Arabidopsis (asyn-1)
Oryza
Plasmodium falciparum
Plasmodium yoelii
Kluyveromyces
Saccharomyces Hop1
Giardia
Trypanosoma brucei
0.1 substitutions/site
B
MRE11
Rattus
Mus
Homo
Gallus
Danio
Xenopus
Caenorhabditis
Magnaporthe
Neurospora
Emericella
Saccharomyces Mre11
Schizosaccharomyces
Bombyx
Anopheles
Drosophila
Coprinopsis
Trypanosoma brucei
Zea
Oryza
Arabidopsis
Plasmodium falciparum
Plasmodium yoelii
Encephalitozoon
Entamoeba
Giardia
Archaeoglobus
Pyrococcus
Pyrobaculum
Methanopyrus
Sulfolobus acidocaldarius
Sulfolobus solfataricus
Methanosarcina
SbcD
0.1 substitutions/site
C
MLH/PMS
Mus
Rattus
Homo
Drosophila
Anopheles
Pms2
Neurospora
Magnaporthe
Schizosaccharomyces
Saccharomyces
Arabidopsis
Caenorhabditis
Plasmodium yoelii
Plasmodium falciparum
Trypanosoma brucei
Encephalitozoon
Giardia
Pms1
Mlh2
Xenopus
Danio
Mlh1
Mlh3
0.1 substitutions/site
D
RAD51/DMC1
Canis
Oryctolagus
Cricetulus
Gallus
Cynops
Xenopus
Bombyx
Drosophila
Anopheles
Magnaporthe
Neurospora
Penicillium
Emericella
Coprinus
Pleurotus
Pichia
Saccharomyces
Schizosaccharomyces
Ustilago
Dictyostelium
Entamoeba
Zea
Oryza
Arabidopsis
Lycopersicon
Physcomitrella
Leishmania major
Leishmania donovani
Trypanosoma brucei
Encephalitozoon
Nosema
Tetrahymena
Plasmodium yoelii
Plasmodium falciparum
Guillardia (nucleomorph)
Caenorhabditis
Rad51
Hydra
Candida
Glycine
Lilium
Hordeum
DMC1-A
DMC1-B
Giardia
Dmc1
Sulfolobus tokodai
Sulfolobus solfataricus
Aeropyrum
Methanococcus
Pyrobaculum
RadA
0.1 substitutions/site

based on the evolutionary principle of "use it or lose it": functions, and the genes that encode them, will be lost by mutation and drift if not maintained by selection. This process is remarkably rapid for genes, with the half-life for newly duplicated genes at ~3–7 million years [40]. Yet, the apparent absence of any one gene (or even a few) by evolutionary loss or by inability to find it does not imply loss of meiosis. Indeed, some meiotic genes included in our inventory have clearly been lost in species in which meiosis is known. Most striking is the shared absence of a group of genes (*Hop2*, *Mnd1*, *Rad52*, *Dmc1*, *Mlh1*, and *Mlh3*) in *Drosophila*, *Anopheles*, and *Caenorhabditis*, as well as the absence of a subset of these genes in *Neurospora*. We note that at least some these proteins are known to function together [37, 38, 41, 42] and are present widely among other eukaryotes (Table 2). Thus, together, these genes usually play an important meiotic role in most eukaryotes. The patterns observed in Table 2 for this set (by extension, other genes in other species) suggest that their presence is a strong positive indicator of meiosis, but absence is uninformative for evaluating the loss or absence of meiosis in a putatively asexual organism.

Phylogenetic Analyses of Meiotic Genes and Gene Families

Our analyses go beyond documenting the presence of meiotic genes in *Giardia* and expanding the inventory of meiotic genes among eukaryotes: We have confirmed our evolutionary inferences by constructing phylogenies for the proteins encoded by each of these genes with a Bayesian likelihood method [43]. It is only by consideration of phylogenetic trees, instead of pairwise database searches (e.g., with BLAST), that one can assess the specific evolutionary relationships of genes to each other. Trees are needed to specify orthologs from paralogs and to discern some cases of gene loss. A clear example is a previous BLAST-based analysis of the *Giardia* genome that erroneously indicated the presence of *Rad51* [34] that is a *Dmc1* ortholog instead.

For the set of core meiosis genes in *Giardia*, our phylogenetic analyses provide strong evidence that all of the genes listed in Table 2 are bona fide homologs of meiotic genes in other eukaryotes (Figure 1; Supplemental Data). The observed relationships among organisms usually follow other well-studied phylogenetic markers (e.g., [10]), placing *Giardia* as an early branch among eukaryotes; in some of the trees, the observed relationships do not, or only weakly, support it. Many of these genes are relatively short and may have rapid evolutionary rates, making them inadequate to reconstruct evolutionary relationships across the depth of eukaryotic phylogeny. However, strong conclusions in all cases for the orthology of individual *Giardia* genes to other meiotic genes can be made: The *Giardia* proteins group unequivocally with other eukaryotic homologs of the proteins in question, usually as a deep branch (see Figure 1; Supplemental Data). For gene families, the ascertainment of orthology requires careful evaluation of the phylogeny. For example, Figure 1D demonstrates that both *recA* homologs present in *Giardia* are clearly Dmc1 orthologs. This strongly supported result demands that a *Rad51* ortholog be actually missing from *Giardia* by either gene loss or as a result of incomplete genome data. An alternative hypothesis that the *Giardia* lineage diverged prior to the *Rad51*/*Dmc1* duplication is rejected because of the robust phylogenetic placement of *Giardia Dmc1* orthologs closely amongst other *Dmc1*s.

Representative examples of phylogenetic trees of various meiotic genes surveyed are in Figure 1 (all trees are in the Supplemental Data); those shown here represent proteins encoded by single genes (*Hop1* and *Mre11*) and gene families (*recA* and *Mlh*), both rooted (*Mre11* and *recA*) and unrooted (*Hop1* and *Mlh*) analyses, and both meiosis-specific (*Hop1* and *Dmc1*) and general meiotic genes (*Mre11* and *Mlh*). In general, all of the assignments of orthology among gene family members were ascertained by inspecting the phylogeny and including support for particular branches in question.

Protists as Keys to Expanding the "Core Machinery"

The potential of *Giardia* to perform meiosis—and the direct inference for the presence of sex—is certainly the most surprising result of this work. Nonetheless, a more fundamental consequence is using *Giardia* as an evolutionarily distant exemplar for the presence and conservation of the core meiotic genes across a wide diversity of eukaryotes. We are extending our census of these meiotic genes to other protists. Included here is our initial survey (Table 2) of additional protist lineages for which substantial data exist from genome sequencing projects, both complete (*Plasmodium*) and ongoing (*Entamoeba* and the Kinetoplastids *Trypanosoma* and *Leishmania*). These additional data also supply clear evidence for the maintenance of meiosis across eukary-

Figure 1. Phylogenetic Trees for Representative Meiotic Genes

All trees shown are the consensus tree topologies determined from ≥900 best trees (i.e., those with the highest posterior probabilities) inferred by Bayesian analysis with alignments of inferred proteins. Major eukaryotic groups are indicated in color, with animals red, fungi brown, plants green, protists blue, and Archaea shown in black. Branches with the best support—that is, those with 0.95–1.00 Bayesian posterior probabilities—are shown thicker. Scale bars represent 0.1 amino acid substitutions per site. Details for each tree, the analytical methods used, and accession numbers for all sequences are provided in the Supplemental Data. Meiosis-specific genes shown are *Hop1* and *Dmc1*.

(A) HOP1 homologs, unrooted. 179 aligned amino acid sites were analyzed; this consensus topology derived from 980 trees; $\alpha = 6.48$ ($3.81 < \alpha < 11.38$), $pI = 0.03$ ($0.006 < pI < 0.08$), and $lnL = -5624.11$.

(B) MRE11 homologs rooted with the archaeal SbcD homolog outgroup. 264 aligned amino acid sites were analyzed; this consensus topology derived from 900 trees; $\alpha = 2.07$ ($1.66 < \alpha < 2.57$), $pI = 0.07$ ($0.03 < pI < 0.11$), and $lnL = -12525.57$.

(C) MutL homologs MLH and PMS rooted with MLH3 paralogs. 437 aligned amino acid sites were analyzed; this consensus topology derived from 970 trees; $\alpha = 2.53$ ($2.21 < \alpha < 2.93$), $pI = 0.005$ ($0.0001 < pI < 0.02$), and $lnL = -36730.34$.

(D) RecA homologs RAD51 and DMC1 rooted with the archaeal *RadA* homolog outgroup. 301 aligned amino acids were analyzed; this consensus topology derived from 970 trees; $\alpha = 1.39$ ($1.08 < \alpha < 1.71$), $pI = 0.07$ ($0.03 < pI < 0.11$), and $lnL = -14317.88$.

otic evolution and argue for its ancestral presence in all extant eukaryotes. Meiotic recombination is known in *Plasmodium* and Kinetoplastids [44, 45]. We note that *Entamoeba* and the microspordian *Encephalitozoon* are not known to undergo meiosis [46, 47]; the fact that they contain some of the same meiosis-specific genes we find in *Giardia* (Table 2) suggests that undiscovered sexual cycles may be present in these species as well.

Although the discovery of meiotic genes in organisms provides evidence for meiosis, there is no clear case in which the converse can yet be evaluated: These genes should be missing from eukaryotic organisms that have indisputably lost meiotic sexual reproduction. It is unclear whether *any* eukaryotic lineage is truly an ancient asexual; a recent review [48] admits that "the corroboration of claims of ancient asexuality remains extremely difficult." Bdelloid rotifers provide the most compelling case for loss of sex [49, 50], but the presence or absence of meiotic genes has not yet been determined in these organisms.

Compared to genes isolated from AFP alone, those of an expanded set of eukaryotic species provide a more meaningful basis for inferences about the pattern of meiotic gene evolution. Protists add both depth and breadth to the questions of conservation and universality of meiosis throughout eukaryotes. Further identification of meiosis genes in protists and continued expansion of a core set will include protists not basal among eukaryotes and those with incomplete genome sequences. This approach will provide a more complete understanding of how meiosis is or is not conserved among *all* eukaryotes, when compared to the few model AFP eukaryotes on which most functional studies are focused.

Meiosis—Early in Eukaryotes

This comparative analysis was motivated by the possibility that meiosis arose during the course of eukaryotic evolution after the divergence of some early-branching eukaryotic lineage(s). If meiosis arose concomitantly with the eukaryotic cell, there would be fewer clues to understanding its origin (from processes in our closest prokaryotic relatives, Archaea). This is the same dilemma faced in understanding all such "major transitions" in evolution [51]: Such massive transformations generally leave scant clues with regard to the original selective pressures and raison d'être for particular features.

Giardia was an excellent candidate for being a premeiotic organism, but our results discount this hypothesis. However, the full diversity of protists, including a clear delineation of the most early-branching lineages, remains unknown, leaving the real possibility that a bona fide ancestrally ameiotic protist will be discovered—and the analyses and tools developed here will be critical for its elucidation. However, the data currently available—especially those from *Giardia*—lead us to favor the more straightforward hypothesis that meiosis is simply a process central to and present in all extant eukaryotes, except for some putative losses from previously meiotic ancestors.

Experimental Procedures

BLAST searches [52] were used to find homologs of meiotic proteins in eukaryotes for which genome sequence data is publicly available. Incomplete sequence data from *Giardia* (http://www.mbl.edu/giardia; [7]) was searched, and genes found were PCR-amplified, cloned, and sequenced to completion. Homology inferences and evolutionary relationships of meiotic gene sequences were validated by phylogenetic analyses with a Bayesian likelihood method (MrBayes, version 3.0b4; [43]). See the Supplemental Data for details.

Supplemental Data

Detailed Experimental Procedures, as well as several supplemental figures and tables, are available online at http://www.current-biology.com/cgi/content/full/15/2/185/DC1/.

Acknowledgments

The authors are grateful to numerous members of the Logsdon lab for advice and suggestions throughout this project. Steven Thomas is thanked for sequencing. We also thank Anne Villeneuve and Kenneth Hillers for helpful comments on a previous version of the manuscript. Bob Malone, Mimi Zolan, Veronique Perrot, Linda Demma, Diane Genereux, Bill Lanier, Arthur Pightling, and Andrew Schurko are thanked for their comments and suggestions. We are especially indebted to Mitch Sogin and the *Giardia* Genome Project, Marine Biological Laboratory at Woods Hole, funded by the National Institute of Allergy and Infectious Diseases/National Institutes of Health under cooperative agreement AI 043273 and to the genome projects for *Entamoeba histolytica* (Institute for Genomic Research), *T. brucei, T. cruzi, L. major*, and *L. donovoni* for making sequence data available. This work was supported in part by start-up funds from Emory University and by a grant from the National Science Foundation (MCB-0216702) to J.M.L.

Received: September 13, 2004
Revised: November 19, 2004
Accepted: November 22, 2004
Published: January 26, 2005

References

1. Kleckner, N. (1996). Meiosis: How could it work? Proc. Natl. Acad. Sci. USA *93*, 8167–8174.
2. Villeneuve, A.M., and Hillers, K.J. (2001). Whence meiosis? Cell *106*, 647–650.
3. Cavalier-Smith, T. (2002). Origins of the machinery of recombination and sex. Heredity *88*, 125–141.
4. Dacks, J., and Roger, A.J. (1999). The first sexual lineage and the relevance of facultative sex. J. Mol. Evol. *48*, 779–783.
5. Adam, R.D. (2001). Biology of *Giardia lamblia*. Clin. Microbiol. Rev. *14*, 447–475.
6. Baldauf, S.L. (2003). The deep roots of eukaryotes. Science *300*, 1703–1706.
7. McArthur, A.G., Morrison, H.G., Nixon, J.E., Passamaneck, N.Q., Kim, U., Hinkle, G., Crocker, M.K., Holder, M.E., Farr, R., Reich, C.I., et al. (2000). The *Giardia* genome project database. FEMS Microbiol. Lett. *189*, 271–273.
8. Raikov, I.B. (1995). Meiosis in protists: Recent advances and persisting problems. Eur. J. Protistol. *31*, 1–17.
9. Patterson, D.J. (1999). The diversity of Eukaryotes. Am. Nat. *154*, S96–S124.
10. 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.
11. Cavalier-Smith, T. (2004). Only six kingdoms of life. Proc. R. Soc. Lond. B. Biol. Sci. *271*, 1251–1262.
12. Simpson, A.G., and Roger, A.J. (2004). The real "kingdoms" of eukaryotes. Curr. Biol. *14*, R693–R696.
13. Dacks, J.B., and Doolittle, W.F. (2001). Reconstructing/deconstructing the earliest eukaryotes: How comparative genomics can help. Cell *107*, 419–425.

14. Haig, D. (1993). Alternatives to meiosis: The unusual genetics of red algae, microsporidia, and others. J. Theor. Biol. *163*, 15–31.
15. Heywood, P., and Magee, P.T. (1976). Meiosis in protists. Some structural and physiological aspects of meiosis in algae, fungi, and protozoa. Bacteriol. Rev. *40*, 190–240.
16. Adam, R.D. (2000). The *Giardia lamblia* genome. Int. J. Parasitol. *30*, 475–484.
17. Sogin, M.L., Gunderson, J.H., Elwood, H.J., Alonso, R.A., and Peattie, D.A. (1989). Phylogenetic meaning of the kingdom concept: An unusual ribosomal RNA from *Giardia lamblia*. Science *243*, 75–77.
18. Hirt, R.P., Logsdon, J.M., Jr., Healy, B., Dorey, M.W., Doolittle, W.F., and Embley, T.M. (1999). Microsporidia are related to Fungi: Evidence from the largest subunit of RNA polymerase II and other proteins. Proc. Natl. Acad. Sci. USA *96*, 580–585.
19. Silberman, J.D., Simpson, A.G., Kulda, J., Cepicka, I., Hampl, V., Johnson, P.J., and Roger, A.J. (2002). Retortamonad flagellates are closely related to diplomonads—implications for the history of mitochondrial function in eukaryote evolution. Mol. Biol. Evol. *19*, 777–786.
20. Embley, T.M., and Hirt, R.P. (1998). Early branching eukaryotes? Curr. Opin. Genet. Dev. *8*, 624–629.
21. Arisue, N., Hasegawa, M., and Hashimoto, T. (2004). Root of the Eukaryota tree as inferred from combined maximum likelihood analyses of multiple molecular sequence data. Mol. Biol. Evol., in press. Published online October 20, 2004. 10.1093/molbev/msi023
22. Stechmann, A., and Cavalier-Smith, T. (2003). The root of the eukaryote tree pinpointed. Curr. Biol. *13*, R665–R666.
23. Stechmann, A., and Cavalier-Smith, T. (2002). Rooting the eukaryote tree by using a derived gene fusion. Science *297*, 89–91.
24. Simpson, A.G. (2003). Cytoskeletal organization, phylogenetic affinities and systematics in the contentious taxon Excavata (Eukaryota). Int. J. Syst. Evol. Microbiol. *53*, 1759–1777.
25. Nasmyth, K. (1996). A homage to *Giardia*. Curr. Biol. *6*, 1042.
26. Otto, S.P., and Lenormand, T. (2002). Resolving the paradox of sex and recombination. Nat. Rev. Genet. *3*, 252–261.
27. Lloyd, D., and Harris, J.C. (2002). *Giardia*: Highly evolved parasite or early branching eukaryote? Trends Microbiol. *10*, 122–127.
28. Tovar, J., Leon-Avila, G., Sanchez, L.B., Sutak, R., Tachezy, J., van der Giezen, M., Hernandez, M., Muller, M., and Lucocq, J.M. (2003). Mitochondrial remnant organelles of *Giardia* function in iron-sulphur protein maturation. Nature *426*, 172–176.
29. Roger, A.J., Svard, S.G., Tovar, J., Clark, C.G., Smith, M.W., Gillin, F.D., and Sogin, M.L. (1998). A mitochondrial-like chaperonin 60 gene in *Giardia lamblia*: Evidence that diplomonads once harbored an endosymbiont related to the progenitor of mitochondria. Proc. Natl. Acad. Sci. USA *95*, 229–234.
30. Upcroft, P., and Upcroft, J.A. (1999). Organization and structure of the *Giardia* genome. Protist *150*, 17–23.
31. Bernander, R., Palm, J.E., and Svard, S.G. (2001). Genome ploidy in different stages of the *Giardia lamblia* life cycle. Cell. Microbiol. *3*, 55–62.
32. Baruch, A.C., Isaac-Renton, J., and Adam, R.D. (1996). The molecular epidemiology of *Giardia lamblia*: A sequence-based approach. J. Infect. Dis. *174*, 233–236.
33. Tibayrenc, M., Kjellberg, F., and Ayala, F.J. (1990). A clonal theory of parasitic protozoa: The population structures of *Entamoeba, Giardia, Leishmania, Naegleria, Plasmodium, Trichomonas*, and *Trypanosoma* and their medical and taxonomical consequences. Proc. Natl. Acad. Sci. USA *87*, 2414–2418.
34. Arkhipova, I.R., and Morrison, H.G. (2001). Three retrotransposon families in the genome of *Giardia lamblia*: Two telomeric, one dead. Proc. Natl. Acad. Sci. USA *98*, 14497–14502.
35. Samuelson, J. (2002). What *Entamoeba histolytica* and *Giardia lamblia* tell us about the evolution of eukaryotic diversity. J. Biosci. *27*, 559–565.
36. Eisen, J.A., and Fraser, C.M. (2003). Phylogenomics: Intersection of evolution and genomics. Science *300*, 1706–1707.
37. Petukhova, G.V., Romanienko, P.J., and Camerini-Otero, R.D. (2003). The Hop2 protein has a direct role in promoting interhomolog interactions during mouse meiosis. Dev. Cell *5*, 927–936.
38. Zierhut, C., Berlinger, M., Rupp, C., Shinohara, A., and Klein, F. (2004). Mnd1 is required for meiotic interhomolog repair. Curr. Biol. *14*, 752–762.
39. Culligan, K.M., Meyer-Gauen, G., Lyons-Weiler, J., and Hays, J.B. (2000). Evolutionary origin, diversification and specialization of eukaryotic MutS homolog mismatch repair proteins. Nucleic Acids Res. *28*, 463–471.
40. Lynch, M., and Conery, J.S. (2000). The evolutionary fate and consequences of duplicate genes. Science *290*, 1151–1155.
41. Tsubouchi, H., and Roeder, G.S. (2002). The Mnd1 protein forms a complex with Hop2 to promote homologous chromosome pairing and meiotic double-strand break repair. Mol. Cell. Biol. *22*, 3078–3088.
42. Tsubouchi, H., and Roeder, G.S. (2003). The importance of genetic recombination for fidelity of chromosome pairing in meiosis. Dev. Cell *5*, 915–925.
43. Huelsenbeck, J.P., and Ronquist, F. (2001). MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics *17*, 754–755.
44. Conway, D.J., Roper, C., Oduola, A.M., Arnot, D.E., Kremsner, P.G., Grobusch, M.P., Curtis, C.F., and Greenwood, B.M. (1999). High recombination rate in natural populations of *Plasmodium falciparum*. Proc. Natl. Acad. Sci. USA *96*, 4506–4511.
45. Gibson, W. (2001). Sex and evolution in trypanosomes. Int. J. Parasitol. *31*, 643–647.
46. Banerjee, S., Das, S., and Lohia, A. (2002). Eukaryotic checkpoints are absent in the cell division cycle of *Entamoeba histolytica*. J. Biosci. *27*, 567–572.
47. Biderre, C., Mathis, A., Deplazes, P., Weber, R., Metenier, G., and Vivares, C.P. (1999). Molecular karyotype diversity in the microsporidian *Encephalitozoon cuniculi*. Parasitology *118*, 439–445.
48. Normark, B.B., Judson, O.P., and Moran, N.A. (2003). Genomic signatures of ancient asexual lineages. Biological Journal of the Linnean Society *79*, 69–84.
49. Birky, C.W., Jr. (2004). Bdelloid rotifers revisited. Proc. Natl. Acad. Sci. USA *101*, 2651–2652.
50. Mark Welch, J.L., Mark Welch, D.B., and Meselson, M. (2004). Cytogenetic evidence for asexual evolution of bdelloid rotifers. Proc. Natl. Acad. Sci. USA *101*, 1618–1621.
51. Szathmary, E., and Smith, J.M. (1995). The major evolutionary transitions. Nature *374*, 227–232.
52. Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. Nucleic Acids Res. *25*, 3389–3402.
53. Borkovich, K.A., Alex, L.A., Yarden, O., Freitag, M., Turner, G.E., Read, N.D., Seiler, S., Bell-Pedersen, D., Paietta, J., Plesofsky, N., et al. (2004). Lessons from the genome sequence of *Neurospora crassa*: Tracing the path from genomic blueprint to multicellular organism. Microbiol. Mol. Biol. Rev. *68*, 1–108.

Accession Numbers

All sequences that we have determined have been deposited in GenBank under accession numbers AY295088 to AY295102 (see Table S3 for details).