Current Biology 17, 1984–1988, November 20, 2007 ©2007 Elsevier Ltd All rights reserved DOI 10.1016/j.cub.2007.10.020

Report

Population Genetics Provides Evidence for Recombination in *Giardia*

Margarethe A. Cooper,[1] Rodney D. Adam,[2] Michael Worobey,[3] and Charles R. Sterling[1,*]
[1]Department of Veterinary Science and Microbiology
[2]Departments of Medicine and Immunobiology
[3]Department of Ecology and Evolutionary Biology
University of Arizona
Tucson, Arizona 85721

Summary

***Giardia lamblia* (syn. *Giardia intestinalis, Giardia duodenalis*) is an enteric protozoan parasite with two nuclei, and it might be one of the earliest branching eukaryotes [1]. However, the discovery of at least rudimentary forms of certain features, such as Golgi and mitochondria, has refuted the proposal that its emergence from the eukaryotic lineage predated the development of certain eukaryotic features. The recent recognition of many of the genes known to be required for meiosis in the genome has also cast doubt on the idea that *Giardia* is primitively asexual, but so far there has been no direct evidence of sexual reproduction in *Giardia*, and population data have suggested clonal reproduction. We did a multilocus sequence evaluation of the genotype A2 reference strain, JH, and five genotype A2 isolates from a highly endemic area in Peru. Loci from different chromosomes yielded significantly different phylogenetic trees, indicating that they do not share the same evolutionary history; within individual loci, tests for recombination yielded significant statistical support for meiotic recombination. These observations provide genetic data supportive of sexual reproduction in *Giardia*.**

Results and Discussion

Giardia lamblia (syn *G. intestinalis, G. duodenalis*) is specific to mammalian hosts and has been divided into a number of genotypes or assemblages (A to G), which have some degree of host specificity [2]. These genotypes might represent different species, but the recognition of new species within *G. lamblia* has remained controversial. All human infections are caused by genotypes A and B. Genotype A1 consists of a very homogeneous group of organisms and includes the genome isolate WB, which has been sequenced to 10× coverage [1] (www.mbl.edu/Giardia). The relatively few A2 nucleotide sequences have been about 98–99% identical to WB.

Comparison of JH and Five *Giardia* Field Isolates

As part of an investigation of the epidemiology of *Giardia* infections, we obtained *Giardia* human fecal isolates from a population in a region of Lima that is highly endemic for human giardiasis [3]. All isolates belonged to genotypes A2 or B. We chose five A2 isolates from four households and used PCR primers based on the WB sequence to compare these isolates to the reference isolate, JH. For the 6.0 kb chromosome 3 region there were 45 single-nucleotide polymorphisms (SNPs), representing a 0.8% SNP density (Figure 1). The 9.5 kb chromosome 5 locus had a total of 82 single nucleotide polymorphisms (SNPs) for a 0.9% SNP density (Figure 2). The chromosome 4 *gdh* coding region had no SNPs for any of the samples, but the chromosome 4 β-*giardin* region had a total of 14 SNPs for a 1% SNP density. The majority of the SNPs were synonymous (Table S1 in the Supplemental Data available online).

Evidence for Chromosome Reassortment

Phylogenetic alignments of the three loci with SNPs demonstrated distinctly different histories for each loci (Figure 3). For example, JH, 55, and 335 were nearly identical at the chromosome 5 locus, but all three were substantially different at the chromosome 3 locus. However, at the chromosome 4 locus, 55 and 335 were identical to each other but substantially different from JH. The apparently different inheritance pattern for loci on different chromosomes suggests the possibility of sexual or other recombination because clonal propagation should result in identical inheritance for loci on different chromosomes.

Evidence for Recombination within Individual Chromosomes

Among eukaryotes, recombination most often occurs by meiotic pairing of chromosomes during sexual reproduction. Our analysis has allowed us to test the prediction that population genetics would provide evidence of meiotic recombination whenever meiotic genes were present. A visual inspection of the SNP alignments (Figures 1 and 2) suggests the possibility of sites for meiotic crossing over, even within individual loci. We designed new sets of primers to sequence across these breakpoints in order to rule out the possibility that these apparent crossing-over sites represented amplification from mixed infections. In each case, the accuracy of the original sequence was confirmed. The genes amplified were single-copy genes, and there were no examples of allelic sequence heterozygosity.

In order to supplement the qualitative evidence for recombination, we used GENECONV to perform a quantitative test to determine whether the within-chromosome regions showed evidence of recombination [4]. GENECONV tests whether certain sequences contain regions that are so similar to the homologous region of other sequences that a history of recombination is implied. For example, if there is a stretch of 2 kb where two sequences are identical despite being highly divergent in the flanking regions, the 2 kb fragment might be detected by GENECONV as a putative mosaic region. A simple follow-up analysis with phylogenetic trees inferred from the different regions detected by GENECONV

*Correspondence: csterlin@u.arizona.edu

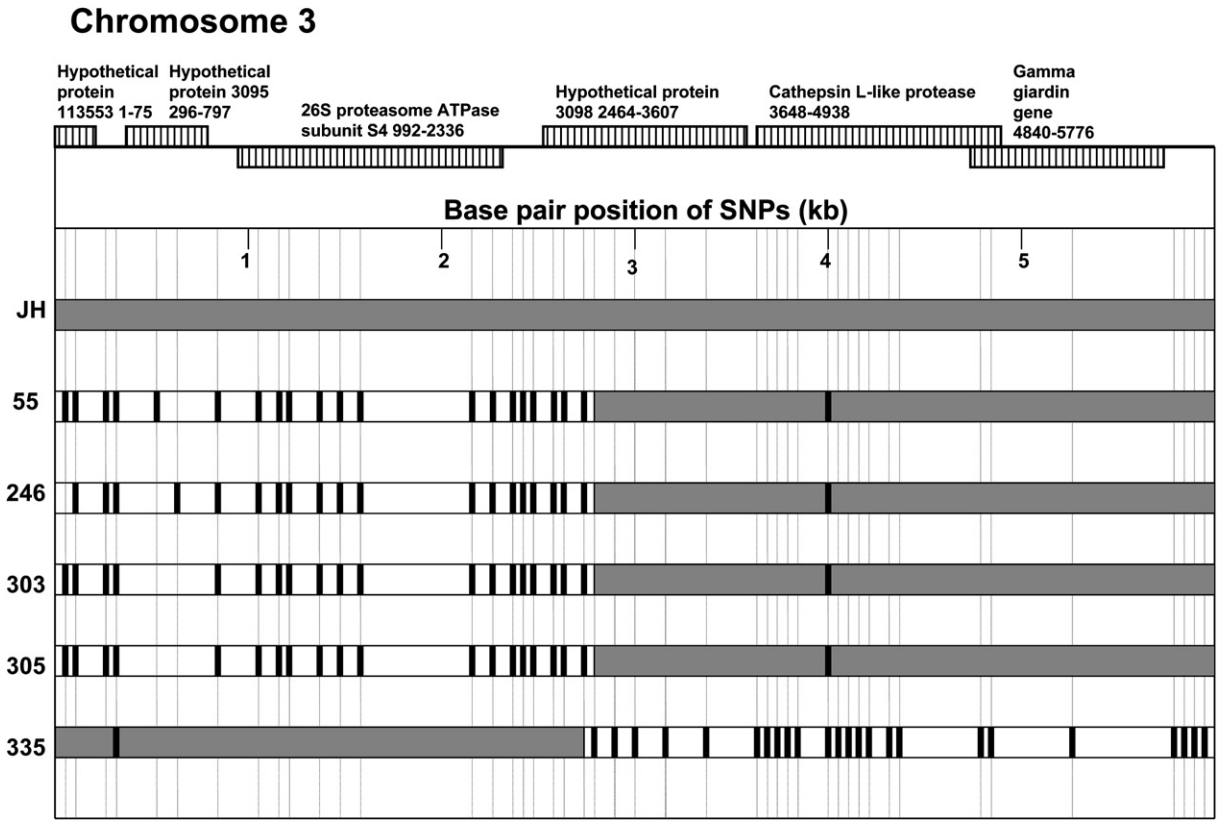

Figure 1. SNP Diagrams for Six Genotype A2 *G. lamblia* Isolates

Chromosome 3 open reading frames and SNPs across 6 kb. ORFs are depicted in vertical striped boxes; top boxes are forward, bottom boxes are reverse. ORF names are written above the box, and if the ORF is a hypothetical protein, its corresponding GiardiaDB ORF number is provided. JH is the reference genotype A2 isolate, and 55, 246, 303 and 305, and 335 are genotype A2 field isolates. Isolates 303 and 305 are from siblings and are identical. Hash marks along each horizontal line represent SNPs in the field isolates, compared to JH.

can then confirm whether these sequences contain regions with significantly conflicting evolutionary histories indicative of recombination. For each chromosomal region there was significant evidence of non-clonal evolution (Table S2 and Figures S1 and S2).

Sexual reproduction is widespread in eukaryotes and is generally considered the norm. However, for many of the protists, documentation of sexual reproduction has been difficult to achieve, and these organisms have been assumed to be asexual. For *Giardia*, the identification of many of the genes required for meiosis in the *Giardia* genome cast doubt on this assumption [5]. However, the complete significance of this observation was not clear because some of the meiosis-specific genes could potentially have other functions, such as other aspects of DNA repair. Even if sexual reproduction occurred, it remained possible that it was rare enough that it would have been difficult to identify.

There are at least three potential reasons why the current work provides evidence favoring sexual reproduction when prior studies did not do so. First, the studies that led to the proposal of clonal reproduction utilized enzyme electrophoresis, which depends on size- or charge-related changes in amino acid sequences [6–8]. The use of nucleotide sequence data in the current study made it possible to identify silent substitutions (which constituted the majority of SNPs), rather than only mutations that change amino acid charge in the relatively conserved metabolic enzymes. Second, we also limited our analysis to members of a single genotype (A2) in order to maximize the possibility of identifying recombination that is limited to closely related organisms. It appears that the A2 genotype is much more heterogeneous than the widespread A1 genotype to which WB belongs, making it easier to identify SNPs. Third, with the exception of the reference isolate, the isolates used in the current work were obtained from a highly endemic host population in which infection with multiple *Giardia* isolates at the same time is common. Perhaps recombination is more frequent in this region of high endemicity because higher endemicity increases the odds of dual infection of a single mammalian host and allows for recombination within the host. For example, some data suggest that the recombination rates for the malaria parasite, *Plasmodium falciparum*, are much higher in regions of high transmission, where there are numerous opportunities for dual infection [9]. However, even in areas of high transmission, the genetic structure of the population of *P. falciparum* can be partially clonal [10]. For *Giardia*, the degree of clonality versus panmixis in different settings remains to be determined. Whether dual infection is necessary and/or sufficient for recombination of *Giardia* is unknown because the life-cycle stage and the location of nuclear fusion in the host is not known.

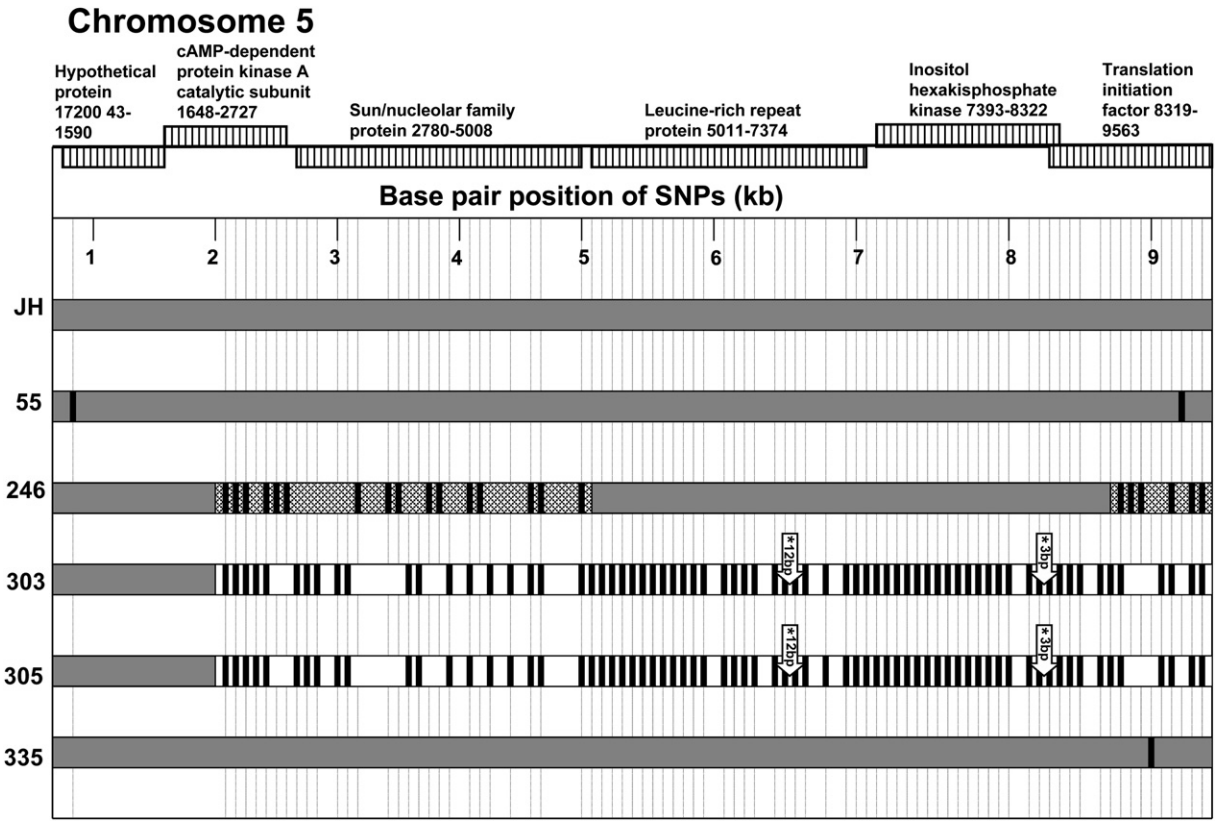

Figure 2. SNP Diagrams for Six Genotype A2 *G. lamblia* Isolates

Chromosome 5 open reading frames and SNPs across 9.5 kb. Open arrows with an asterisk depict an insertion or deletion (indel). See legend for Figure 1.

The identification of potential meiotic recombination in *Giardia* also helps to address another dilemma, that of the low level of allelic heterozygosity. When asexual organisms have a ploidy of greater than one their chromosome homologs diverge, as has happened with bdelloid rotifers that have reproduced asexually for more than 40 million years [11, 12]. In addition, *Giardia* has the requirement for homogenizing its two nuclei. Trophozoites have a ploidy of approximately four; two sets of five chromosomes [13] are divided approximately equally between the nuclei [14, 15]. Replication is equational rather than reductional [15–17], which means that nuclear asymmetry is maintained throughout the replication cycle. Therefore, if the nuclei were asexual then they would be expected to diverge from one another. However, the observed allelic sequence heterozygosity (ASH) is less than 0.002% in the genome isolate WB [18]. This low ASH means that alleles within nuclei as well as those from opposite nuclei are nearly identical. The existence of meiotic recombination could provide an explanation for this level of homogeneity if it provides for reassortment and recombination within and between nuclei.

Although evidence from genomic and population genetics suggests that meiosis occurs in *Giardia*, sexual reproduction has never been directly observed. Sex in *Giardia* could be infrequent, furtive, or cryptic [19], and it could occur during trophozoite replication or during encystation or excystation. The answers to these questions will probably require experimental protocols with mixed infections or with selectable drug markers. It will also be important to determine whether recombination occurs only within certain genotypes. Genotype B is markedly different from genotypes A1 and A2 and has already been proposed as a separate species [20, 21]. If it is found that recombination occurs only within genotype A2 isolates, and not between A and B, this would be strong evidence that A and B are separate species. A manuscript published while this work was under review sequenced 20 clones at each of ten loci from three axenic cultures of A1, three of A2, and three of B isolates [22]. At several loci, there were rare genotype B haplotypes that grouped with A1, which raised the possibility of past recombination between A1 and B. Interestingly, the A2 isolates formed a single clade with no evidence of recombination between the other genotypes. Further investigation will be required to determine the precise species structure of these *G. lamblia* lineages. Because analysis of the *Giardia* genome supports that organism's role as one of the earliest diverging eukaryotes [1], the evidence for sexual reproduction in *Giardia* would seem to support the proposals that sexual reproduction is a fundamental eukaryotic property [5, 23] and that eukaryotes without sexual reproduction have reached that state by secondary loss. Therefore, it is possible that *Giardia* will provide important clues to the origin of meiosis and sexual reproduction.

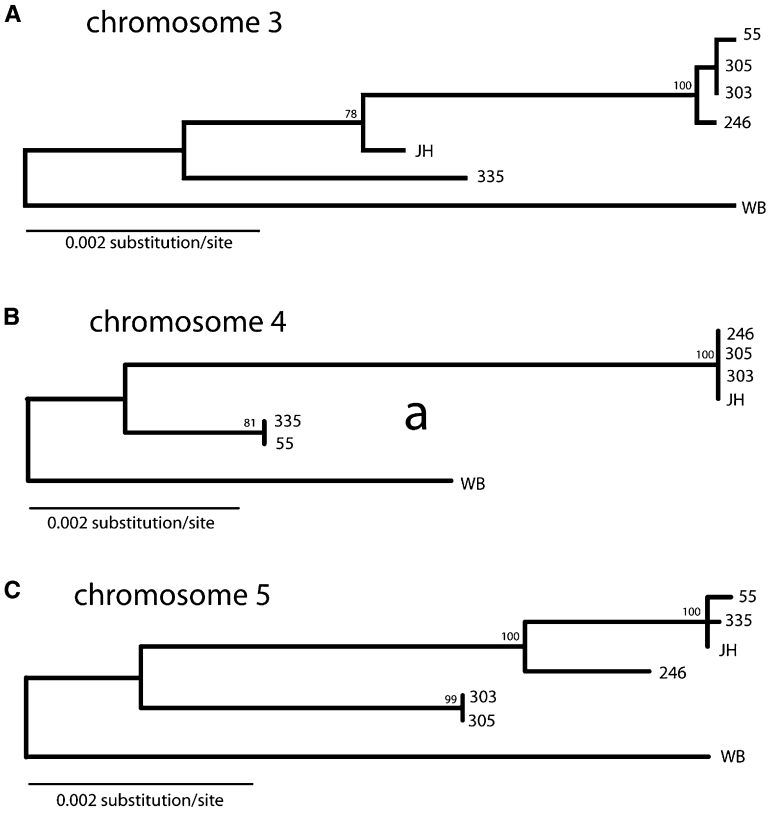

Figure 3. Maximum-Likelihood Trees, with Bootstrap Support, of Sequences from Chromosomes 3, 4, and 5 of *Giardia*

WB is the genotype A1 outgroup.
(A) Chromosome 3 6 kb region including γ-*giardin*.
(B) Chromosome 4 1.2 kb region including β-*giardin*.
(C) Chromosome 5 9.6 kb region adjacent to *tpi*.

Experimental Procedures

Acquisition of *G. lamblia* Samples

Fecal samples with Giardia were obtained from an epidemiological study in Las Pampas de San Juan de Miraflores, a desert shantytown 25 km to the south of Lima, Peru. More details about the demographics and environment of this shantytown have been described [24–26]. The Institutional Review Board of PRISMA, the non-government organization in Peru, and the Human Subjects Approval Committee at the University of Arizona approved the protocol used for the study.

Five isolates were randomly chosen from four households (the two samples from the same household were from siblings). Fecal samples were purified by sedimentation via methods already described [27], with the exception that purified double distilled water was used in lieu of formalin. Fecal concentrate was purified via the QIAmp DNA Stool Kit (QIAgen, Valencia, CA). The reference isolate, JH, has been previously described as a group 2 (genotype A2) isolate [20, 21, 28].

Sequencing Strategy

The sequences adjacent to the *triose phosphate isomerase* (*tpi*) gene on chromosome 5 (GenBank #L02120, GiardiaDB ORF: 93938), γ-*giardin* gene (GenBank #X55287; GiardiaDB ORF: 17230) from chromosome 3, and the *glutamate dehydrogenase* (*gdh*) (GenBank #M84604; GiardiaDB ORF: 21942) and β-*giardin* (GenBank #X85958; GiardiaDB ORF: 4812) sequences from chromosome 4 (the *Giardia lamblia* genome, http://www.mbl.edu/Giardia) were used for designing PCR primers that would amplify 500 bp regions of genomic DNA with at least 50 bp of overlapping DNA so that any SNPs located within the chosen primer sequences could be identified. Non-acetylated bovine serum albumin (BSA; 0.5%) and dimethyl sulfoxide (DMSO; 10%) were included in the reactions. PCR products were purified with the microcentrifuge-based Gel Extraction protocol of the QIAquick PCR purification kit (QIAgen, Valencia, CA). Nested amplification primers and DMSO were used for sequencing purified PCR products. Sequenced products were deposited into the GenBank database under accession numbers EU188624–EU188641.

Phylogenetic and Recombination Breakpoint Analyses

Exploratory tree analysis was used for determining whether there was evidence of conflicting phylogenetic signal [29]. Neighbor-joining trees were reconstructed on 600 nucleotide (nt) windows of the original alignment shifted in 300 nt increments, as well as for the full-length data set, with distances estimated with the HKY85 [30] model of DNA substitution. Conflicting phylogenetic signal contained within the sequences of putative signals was confirmed by bootstrap phylogenetic trees. To complement the phylogenetic approach, Sawyer's [4] runs test, in the package GENECONV (S. Sawyer, Department of Mathematics, Washington University, St. Louis, MO), was performed. This test searches for unusually long fragments within an alignment over which a pair of sequences are identical or nearly identical, then assesses the significance of the hypothesis that the similar fragments arose by recombination by using randomly permuted data sets derived from the real alignment.

Supplemental Data

Two figures and two tables are available with this article online at http://current-biology.com/cgi/content/full/17/22/1984/DC1/.

Acknowledgments

We thank R. Gilman for sponsorship of the project in Peru and L. Cabrera for sample collection at the study site. We thank B.H. Jost, D. Wolk, L. Joens, and C.W. Birky, Jr. for helpful comments. This study was made possible by PRISMA (Asociación Benéfica Proyectos en Informática, Salud, Medicina y Agricultura), by funding from National Institutes of Health Minority International Research Training grant TW00036 and Achievement Rewards for College Scientists to M.A.C., and by grant 136034-H-02-124 from the United States Department of Agriculture, Agricultural Experiment Station, University of Arizona to C.R.S.

Received: July 5, 2007
Revised: October 4, 2007
Accepted: October 5, 2007
Published online: November 1, 2007

References

1. Morrison, H.G., McArthur, A.G., Gillin, F.D., Aley, S.B., Adam, R.D., Olsen, G.J., Reiner, D.S., Svard, S.G., Elmendorf, H.G., Hehl, A.B., et al. (2007). Genomic minimalism in the early diverging, intestinal parasite, Giardia lamblia. Science *17*, 1921–1926.
2. Adam, R.D. (2001). Biology of Giardia lamblia. Clin. Microbiol. Rev. *14*, 447–475.
3. Gilman, R.H., Marquis, G.S., Miranda, E., Vestegui, M., and Martinez, H. (1988). Rapid reinfection by Giardia lamblia after treatment in a hyperendemic Third World community. Lancet *1*, 343–345.
4. Sawyer, S. (1989). Statistical tests for detecting gene conversion. Mol. Biol. Evol. *6*, 526–538.
5. Ramesh, M.A., Malik, S.B., and Logsdon, J.M., Jr. (2005). A phylogenomic inventory of meiotic genes; evidence for sex in *Giardia* and an early eukaryotic origin of meiosis. Curr. Biol. *15*, 185–191.
6. Meloni, B.P., Lymbery, A.J., and Thompson, R.C. (1995). Genetic characterization of isolates of Giardia duodenalis by enzyme electrophoresis: Implications for reproductive biology, population structure, taxonomy, and epidemiology. J. Parasitol. *81*, 368–383.
7. Tibayrenc, M., Neubauer, K., Barnabe, C., Guerrini, F., Skarecky, D., and Ayala, F.J. (1993). Genetic characterization of six parasitic protozoa: Parity between random-primer DNA typing and multilocus enzyme electrophoresis. Proc. Natl. Acad. Sci. USA *90*, 1335–1339.
8. 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.
9. Urdaneta, L., Lal, A., Barnabe, C., Oury, B., Goldman, I., Ayala, F.J., and Tibayrenc, M. (2001). Evidence for clonal propagation in natural isolates of Plasmodium falciparum from Venezuela. Proc. Natl. Acad. Sci. USA *98*, 6725–6729.
10. Razakandrainibe, F.G., Durand, P., Koella, J.C., De, M.T., Rousset, F., Ayala, F.J., and Renaud, F. (2005). "Clonal" population structure of the malaria agent Plasmodium falciparum in high-infection regions. Proc. Natl. Acad. Sci. USA *102*, 17388–17393.
11. Welch, D.M., and Meselson, M. (2000). Evidence for the evolution of bdelloid rotifers without sexual reproduction or genetic exchange. Science *288*, 1211–1215.
12. Welch, D.B., and Meselson, M.S. (2001). Rates of nucleotide substitution in sexual and anciently asexual rotifers. Proc. Natl. Acad. Sci. USA *98*, 6720–6724.
13. Adam, R.D., Nash, T.E., and Wellems, T.E. (1988). The Giardia lamblia trophozoite contains sets of closely related chromosomes. Nucleic Acids Res. *16*, 4555–4567.
14. 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.
15. Yu, L.Z., Birky, C.W., Jr., and Adam, R.D. (2002). The two nuclei of Giardia each have complete copies of the genome and are partitioned equationally at cytokinesis. Eukaryot. Cell *1*, 191–199.
16. Ghosh, S., Frisardi, M., Rogers, R., and Samuelson, J. (2001). How Giardia swim and divide. Infect. Immun. *69*, 7866–7872.
17. Dawson, S.C., Sagolla, M.S., and Cande, W.Z. (2007). The cenH3 histone variant defines centromeres in Giardia intestinalis. Chromosoma *116*, 175–184.
18. Prabhu, A., Morrison, H.G., Martinez, C.R., III, and Adam, R.D. (2007). Characterisation of the subtelomeric regions of Giardia lamblia genome isolate WBC6. Int. J. Parasitol. *37*, 503–513.
19. Birky, C.W., Jr. (2005). Sex: Is *Giardia* doing it in the dark? Curr. Biol. *15*, R56–R58.
20. Nash, T.E., McCutchan, T., Keister, D., Dame, J.B., Conrad, J.D., and Gillin, F.D. (1985). Restriction-endonuclease analysis of DNA from 15 Giardia isolates obtained from humans and animals. J. Infect. Dis. *152*, 64–73.
21. Nash, T.E., and Keister, D.B. (1985). Differences in excretory-secretory products and surface antigens among 19 isolates of Giardia. J. Infect. Dis. *152*, 1166–1171.
22. Teodorovic, S., Braverman, J.M., and Elmendorf, H.G. (2007). Unusually low levels of genetic variation among Giardia lamblia isolates. Eukaryot. Cell *6*, 1421–1430.
23. Cavalier-Smith, T. (2002). Origins of the machinery of recombination and sex. Heredity *88*, 125–141.
24. Bern, C., Ortega, Y., Checkley, W., Roberts, J.M., Lescano, A.G., Cabrera, L., Verastegui, M., Black, R.E., Sterling, C., and Gilman, R.H. (2002). Epidemiologic differences between cyclosporiasis and cryptosporidiosis in Peruvian children. Emerg. Infect. Dis. *8*, 581–585.
25. Checkley, W., Epstein, L.D., Gilman, R.H., Black, R.E., Cabrera, L., and Sterling, C.R. (1998). Effects of Cryptosporidium parvum infection in Peruvian children: Growth faltering and subsequent catch-up growth. Am. J. Epidemiol. *148*, 497–506.
26. Checkley, W., Gilman, R.H., Black, R.E., Lescano, A.G., Cabrera, L., Taylor, D.N., and Moulton, L.H. (2002). Effects of nutritional status on diarrhea in Peruvian children. J. Pediatr. *140*, 210–218.
27. Young, K.H., Bullock, S.L., Melvin, D.M., and Spruill, C.L. (1979). Ethyl acetate as a substitute for diethyl ether in the formalin-ether sedimentation technique. J. Clin. Microbiol. *10*, 852–853.
28. 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.
29. Worobey, M., and Holmes, E.C. (2001). Homologous recombination in GB virus C/hepatitis G virus. Mol. Biol. Evol. *18*, 254–261.
30. Hasegawa, M., Kishino, H., and Yano, T. (1985). Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. J. Mol. Evol. *22*, 160–174.