Common inheritance of chromosome Ia associated with clonal expansion of Toxoplasma gondii

Asis Khan  
Washington University School of Medicine in St. Louis

L. David Sibley  
Washington University School of Medicine in St. Louis

et al

Follow this and additional works at: http://digitalcommons.wustl.edu/open_access_pubs

Recommended Citation
http://digitalcommons.wustl.edu/open_access_pubs/2011

This Open Access Publication is brought to you for free and open access by Digital Commons@Becker. It has been accepted for inclusion in Open Access Publications by an authorized administrator of Digital Commons@Becker. For more information, please contact engeszer@wustl.edu.
Common inheritance of chromosome Ia associated with clonal expansion of *Toxoplasma gondii*

Asis Khan, Ulrike Böhme, Krystyna A. Kelly, et al.

*Genome Res.* 2006 16: 1119-1125
Access the most recent version at doi:10.1101/gr.5318106

---

**Supplemental Material**

[http://genome.cshlp.org/content/suppl/2006/08/10/gr.5318106.DC1.html](http://genome.cshlp.org/content/suppl/2006/08/10/gr.5318106.DC1.html)

**References**

This article cites 28 articles, 16 of which can be accessed free at: [http://genome.cshlp.org/content/16/9/1119.full.html#ref-list-1](http://genome.cshlp.org/content/16/9/1119.full.html#ref-list-1)

**Creative Commons License**

This article is distributed exclusively by Cold Spring Harbor Laboratory Press for the first six months after the full-issue publication date (see [http://genome.cshlp.org/site/misc/terms.xhtml](http://genome.cshlp.org/site/misc/terms.xhtml)). After six months, it is available under a Creative Commons License (Attribution-NonCommercial 3.0 Unported License), as described at [http://creativecommons.org/licenses/by-nc/3.0/](http://creativecommons.org/licenses/by-nc/3.0/).

**Email Alerting Service**

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](http://genome.cshlp.org/subscriptions).

---

To subscribe to *Genome Research* go to: [http://genome.cshlp.org/subscriptions](http://genome.cshlp.org/subscriptions)

Copyright © 2006, Cold Spring Harbor Laboratory Press
Common inheritance of chromosome Ia associated with clonal expansion of *Toxoplasma gondii*

Asis Khan,1,10 Ulrike Böhme,2,10 Krystyna A. Kelly,3,4,10 Ellen Adlem,2 Karen Brooks,2 Mark Simmonds,2 Karen Mungall,2 Michael A. Quail,2 Claire Arrowsmith,2 Tracey Chillingworth,2 Carol Churcher,2 David Harris,2 Matthew Collins,2 Nigel Fosker,2 Audrey Fraser,2 Zahra Hance,2 Kay Jagels,2 Sharon Moule,2 Lee Murphy,2 Susan O’Neil,2 Marie-Adele Rajandream,2 David Saunders,2 Kathy Seeger,2 Sally Whitehead,2 Thomas Mayr,3 Xuenan Xuan,5 Junichi Watanabe,6 Yutaka Suzuki,7 Hiroyuki Wakaguri,7 Sumio Sugano,7 Chihiro Sugimoto,5 Ian Paulsen,8 Aaron J. Mackey,9 David S. Roos,9 Neil Hall,2,8 Matthew Berriman,2 Bart Barrell,2 L. David Sibley,1 and James W. Ajioka3,11

1Department of Molecular Microbiology, Washington University School of Medicine, St. Louis, Missouri 63110, USA; 2The Wellcome Trust Sanger Institute, The Wellcome Trust Genome Campus, Hinxton, CB10 1SA, United Kingdom; 3Department of Pathology, University of Cambridge, Cambridge, CB2 1QP, United Kingdom; 4MRC Biostatistics Unit, Institute of Public Health, Cambridge, CB2 2SR, United Kingdom; 5National Research Center for Protozoan Diseases, Obihiro University of Agriculture and Veterinary Medicine, Inada-cho, Obihiro, Hokkaido, 080-8555, Japan; 6Department of Parasitology, Institute of Medical Science, The University of Tokyo, 4-6-1, Shirokanedai, Minato-ku, Tokyo, 108-8639, Japan; 7Department of Medical Genomics, Graduate School of Frontier Sciences, The University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa, Chiba 277-8562, Japan; 8The Institute for Genomic Research, Rockville, Maryland 20850, USA; 9Department of Biology, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA

*Toxoplasma gondii* is a globally distributed protozoan parasite that can infect virtually all warm-blooded animals and humans. Despite the existence of a sexual phase in the life cycle, *T. gondii* has an unusual population structure dominated by three clonal lineages that predominate in North America and Europe, (Types I, II, and III). These lineages were founded by common ancestors ∼10,000 yr ago. The recent origin and widespread distribution of the clonal lineages is attributed to the circumvention of the sexual cycle by a new mode of transmission—asexual transmission between intermediate hosts. Asexual transmission appears to be multigenic and although the specific genes mediating this trait are unknown, it is predicted that all members of the clonal lineages should share the same alleles. Genetic mapping studies suggested that chromosome Ia was unusually monomorphic compared with the rest of the genome. To investigate this further, we sequenced chromosome Ia and chromosome Ib in the Type I strain, RH, and the Type II strain, ME49. Comparative genome analyses of the two chromosomal sequences revealed that the same copy of chromosome Ia was inherited in each lineage, whereas chromosome Ib maintained the same high frequency of between-strain polymorphism as the rest of the genome. Sampling of chromosome Ia sequence in seven additional representative strains from the three clonal lineages supports a monomorphic inheritance, which is unique within the genome. Taken together, our observations implicate a specific combination of alleles on chromosome Ia in the recent origin and widespread success of the clonal lineages of *T. gondii*.

Supplemental material is available online at www.genome.org. The sequence data from this study have been submitted to EMBL under accession nos. AM055942 and AM055943. Accession numbers for full-length cDNAs associated with chromosomes Ia and Ib are AK223637–AK223774.

*Toxoplasma gondii* is a globally distributed protozoan parasite that infects ∼25% of the human population and a similar proportion of birds and mammals (Dubey and Beattie 1988; Hall et al. 2001). Acute infection during pregnancy can result in birth defects and the parasite may cause fatal encephalitis in immunocompromised individuals (Hall et al. 2001). However, the vast majority of infections go unnoticed or cause mild flu-like symptoms, resulting in an asymptomatic, chronic infection lasting the lifetime of the host. Importantly, the relatively quiescent tissue cyst form of the parasite in the chronic infection is readily transmitted between intermediate hosts by carnivorous or omnivorous feeding (Dubey and Beattie 1988). This is thought to be the driving force behind the recent and rapid expansion of *T. gondii* (Su et al. 2003).

By their very nature, parasitic organisms have evolved com-
plex life cycles that exploit their intimate association with their hosts. For apicomplexan species, including the causative agents of malaria (Plasmodium spp.) and avian coccidiosis (Eimeria spp.), the maintenance of a sexual cycle may be vital for generating the diversity necessary to adapt to varying environments by evading host immunity and optimizing transmission and virulence (Kyes et al. 2001; Mackinnon and Read 2004). The cost of this diversity comes at the expense of being reliant on a single definitive host and the more general problem of breaking up advantageous combinations of alleles (Maynard-Smith 1978). T. gondii, arguably the most successful species in the phylum, is a striking exception because, along with the sexual cycle, the parasite can transmit asexually by ingestion of tissue cysts by intermediate hosts (Dubey and Beattie 1988). This property, termed asexual oral transmission, underpins its highly clonal population structure with only three predominant lineages (Types I, II, and III) widespread in North America and Europe, where members within each lineage are virtually identical (Darde et al. 1992; Sibele and Boothroyd 1992; Howe and Sibley 1995). Equally remarkable, when comparing between lineages, there exist only two alleles at each locus. Together, these patterns indicate that the clonal lineages arose very recently from a common set of parental strains (referred to as Adam and Eve or A&E) and subsequently underwent only limited genetic exchange (Grigg et al. 2001a; Su et al. 2003). Thus, the need to define the mechanisms and common genetic elements that underlie asexual transmission and the ubiquitous distribution of T. gondii provides a compelling rationale for genetic and comparative genomic analyses.

Experimental genetic crosses between prototypic members of the three common lineages demonstrate a Mendelian pattern of inheritance (Su et al. 2002; Khan et al. 2005). There are 14 distinct linkage groups that comprise the ~65 Mb haploid genome. The recently compiled genome map relies on the presence of frequently occurring A&E polymorphisms for 229 dispersed genetic markers (Su et al. 2002; Khan et al. 2005); polymorphisms occur at a frequency of ~1 in 100 bp on all chromosomes with the exception of chromosome Ia (chrIa). To our surprise, and unlike any other chromosome, genetic polymorphisms proved exceedingly difficult to identify on chrIa. Moreover, the density of polymorphic markers elsewhere in the genome (spaced approximately every 300 kb) suggests that chrIa is unique in its lack of polymorphism compared with the rest of the genome. This lack of polymorphism has been confirmed by a recent global analysis of single nucleotide polymorphisms (SNPs) undertaken at TIGR (available at http://www.tigr.org and http://www.toxodb.org/ToxoDB.shmIt). Genomic sequence alignments between the Type I RH and Type II ME49 clone B7 for both chromosomes show that each chromosome pair has the same gene content and structural elements. However, there is a very dramatic difference in the levels of polymorphism between the two chromosomes (Fig. 1; Supplemental Table 2): chrIa has 12,925 SNPs compared with only 332 SNPs on chrIb (Table 1). An analysis of the SNPs in the intron sequences of a sample of genes from different chromosomal locations estimates that about seven SNPs per kilobase are true biallelic polymorphisms (i.e., A&E) due to sequence differences between the original parents (Su et al. 2003).

<table>
<thead>
<tr>
<th>Feature</th>
<th>chrIa</th>
<th>chrIb</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total sequence length (bp)</td>
<td>1,923,819</td>
<td>2,013,089</td>
</tr>
<tr>
<td>% GC of genome</td>
<td>52.63</td>
<td>52.30</td>
</tr>
<tr>
<td>% GC of CDS</td>
<td>57.97</td>
<td>57.75</td>
</tr>
<tr>
<td>Number of predicted CDSs*</td>
<td>211</td>
<td>252</td>
</tr>
<tr>
<td>Number of pseudogenes</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>Expression confirmed by ESTs*</td>
<td>102 (47%)</td>
<td>117 (46%)</td>
</tr>
<tr>
<td>Average CDS length (bp)</td>
<td>2682</td>
<td>2185</td>
</tr>
<tr>
<td>Number of spliced genes</td>
<td>152 (69%)</td>
<td>192 (75%)</td>
</tr>
<tr>
<td>Average exon length (bp)</td>
<td>463</td>
<td>386</td>
</tr>
<tr>
<td>Average intron length (bp)</td>
<td>519</td>
<td>528</td>
</tr>
<tr>
<td>Average number of exons per gene</td>
<td>5.79</td>
<td>5.65</td>
</tr>
<tr>
<td>Gene density</td>
<td>1 per 9117 bp</td>
<td>1 per 7988 bp</td>
</tr>
<tr>
<td>Coding percentage</td>
<td>56.7%</td>
<td>58.1%</td>
</tr>
<tr>
<td>Number of tRNA genes</td>
<td>3</td>
<td>10</td>
</tr>
<tr>
<td>Number of SNPs**</td>
<td>332</td>
<td>12,925</td>
</tr>
</tbody>
</table>

*Excluding pseudogenes.

**Alignment to one or more ESTs.

*Based on a comparison between RH and ME49.

Table 1. Features of T. gondii chromosome Ia and Ib DNA sequence for the RH strain

Results

Comparison of chrIa and chrIb in the Type I strain RH reveals similar features

To provide a basis for initial comparative genomic analysis, we undertook complete sequencing of chrIa and chrIb genomic DNA from the RH strain of T. gondii. The genomic sequence for the RH strain reveals that the 1.92-Mb chrIa and the 2.01-Mb chrIb have an average G+C content of just over 52% with 211 protein-coding sequences (CDSs) predicted for chrIa and 252 for chrIb (Table 1). Expression of 47% of the total predicted genes was confirmed by comparison with an extensive set of Expressed Sequence Tag (EST) data available at http://www.cblif.upenn.edu/apidots/downloads/ (Wan et al. 1996; Ajoka et al. 1998), of which 35% were identified completely. Comparative genome analyses revealed that both chrIa and chrIb have similar levels of coding and noncoding regions and contain conventional telomeres (Table 1; Supplemental material). Crucially, chrIa does not appear to be unusual in structure or content compared with chrIb, which we have used to be representative of the rest of the T. gondii genome.
A single ancestral chromosome that is extremely recent in origin. Overall low rate of polymorphism argues that these regions that would explain the lower polymorphism, also identical between RH and ME49, having only 0.17 SNPs per kilobase. This is likely to be an overestimate, because more than half of the apparent polymorphism is likely to be due to sequencing errors, even with the low sequencing error rate of ~1 in 10^5 bp achieved in this study. The 332 SNPs on chrIa are not distributed at random across the chromosome (P~0), largely due to the paucity of SNPs between positions 150,000 and 750,000 and at the right-hand end of chrIa (Fig. 1). The genes in the relatively polymorphism-free region in the center of the chromosome can be identified from the start positions given in Supplemental Table 3. There is no obvious bias in the functional categories of genes in these regions that would explain the lower polymorphism, although many of the genes are annotated as “unknown.” The overall low rate of polymorphism argues that chrIa is derived from a single ancestral chromosome that is extremely recent in origin.

chrIa is shared in all lineages

In order to assess the generality of these observations, we analyzed a Type III strain called VEG, which represents the third predominant lineage. Since whole-genome sequence is not yet available, the EST database (Li 2003) was screened and 26 scattered regions of 800–900 bp that included at least three ESTs were selected for sequencing. There were only two SNPs in 23,654 bp of chrIa sequenced from VEG (Fig. 2). This low rate of polymorphism argues that the Type III lineage also inherited the same version of chrIa without apparent crossover. In order to confirm the uniparental inheritance of chrIa among the three clonal lineages, we sequenced 12 of the 26 regions from chrIa in two additional strains from each clonal type. These strains were selected from human or animal isolates that are representative of the three clonal lineages that predominate in the USA and France (Table 2). No SNPs were found in the sequence from these additional strains (a total of 9646 bp for each strain; see Fig. 2). These data are consistent with a common inheritance of chrIa for all members of the clonal lineages, which predominate in North America and Europe.

Time of origin of the common chrIa is coincident with origin of the clonal lineages

The combined data show two SNPs in 128,146 total bp, where 86,814 bp represents sequence from the 12 regions available for all nine strains (Fig. 2; Table 2). This very low rate of polymorphism of 2/86,814 bp for chrIa is strikingly similar to the overall rate of unique SNPs between the three lineages. We have previously shown that while A&E polymorphisms occur across the genome at about 1/100 bp, when these biallelic polymorphisms are removed, the remaining unique polymorphism rate is on the order of 2/90,000 bp (Su et al. 2003). Based on the assumption that each strain evolves independently and these polymorphisms represent random (Poisson) neutral mutations that have occurred since their common ancestry (Rich et al. 1998), we have previously estimated that the three lineages originated ~10,000 yr ago (Su et al. 2003). Using this same model, the simplest explanation of our observations is that chrIa was inherited in common by all three lineages at a time coincident with their recent origin.

Repression of meiotic recombination does not explain the monomorphic chrIa

How might a virtually monomorphic chrIa be established and maintained within an otherwise polymorphic genome? A monomorphic chrIa is currently maintained in the population because the frequency of a sexual encounter with anything other than a member of a clonal lineage is very low (i.e., when crosses do occur, chrIa is effectively “selfing” in the vast majority of crosses). However, this does not explain how the current monomorphic chrIa was generated and initially propagated in a sexually reproducing polymorphic population. In sexually reproducing populations, population genetic theory suggests that chromosomes/chromosomal regions where meiotic recombination is suppressed may have reduced levels of polymorphism (Maynard-Smith and Haigh 1974; Charlesworth et al. 1993). This could be due to a selective sweep or purifying selection as proposed for the small fourth chromosome of Drosophila melanogaster (Berry et al. 1991; Jensen et al. 2002; Bartolome and Maside 2004), and similarly, it is possible that suppression of recombination could have led to the establishment of a monomorphic chrIa in T. gondii. To
test whether a similar process might occur in *T. gondii*, we analyzed the segregation of SNPs on *chla* in the progeny of a genetic cross between the Type I and Type III lineages as previously reported (Khan et al. 2005). Our analysis showed that 12 of the 31 progeny showed one or more crossing-over events in a single meiosis on *chla* (Fig. 3) and, in experimental crosses, it is observed to have the highest frequency of recombination of any *T. gondii* chromosome (Khan et al. 2005). Thus, there is no physical suppression of recombination on *chla*. Nevertheless, in experimental crosses of *T. gondii*, individual chromosomes can be inherited unparentally (i.e., without crossing over). This raises the possibility that it is simply the result of chance that *chla* is the only chromosome/chromosomal region common between the three clonal lineages, particularly if they are related by only a few crosses.

**Chance inheritance of *chla* is highly unlikely**

A study of EST-identified SNPs by Boyle et al. (2006) revealed a surprising pattern of SNP distribution across the genome—haplotype blocks are conserved across large regions of the genome. Further analysis of these patterns indicates that the relationship between the clonal lineages might be explained by as few as two separate crosses. They propose that the existence of three clonal lineages were generated by two crosses, in which one lineage (Type II) is a common parent of the other two, and that the subsequent lineages have undergone little mixing since this recent origin. Given the two-cross model, it is expected that some chromosomes will not recombine and will be shared in their entirety between the three lineages, some chromosomes will have shared regions, and some will have no shared regions. Experimental crosses with *T. gondii* predict a recombination rate of 0.4 (Khan et al. 2005); hence, the expected frequencies of these three classes are 0.09, 0.36, and 0.55, respectively. The probability of observing a single intact chromosome with no other chromosome having shared regions can be calculated using the trinomial distribution:

\[
P = \frac{N!}{n_1!n_2!n_3!} p_1^{n_1} p_2^{n_2} p_3^{n_3}
\]

where *N* is the number of chromosomes, *n*<sub>1</sub>, *n*<sub>2</sub>, *n*<sub>3</sub> are the observed numbers of shared chromosomes, chromosomes with shared sections, and chromosomes with no shared sections, and *p*<sub>1</sub>, *p*<sub>2</sub>, *p*<sub>3</sub> are the probabilities of observing these three types of chromosomes on the two-cross model.

Therefore:

\[
P = \frac{14!}{1!0!13!} 0.09^1 0.36^0 0.55^{13}
\]

\[
= 0.0005
\]

Collapsing the first two classes to give a minimum expected number greater than five, the expected numbers of chromosomes with shared material and with no shared material are 7.7 and 6.3 compared with observations of 1 and 13, giving \(\chi^2 = 8.1\) (*P* = 0.004). Given the two-cross model, the observed pattern is highly unlikely, with fewer shared chromosome/chromosomal regions than expected by chance. If more than two crosses led to the extant clonal lineages, the probability is further reduced, so it is extremely unlikely that this pattern would occur by chance.

**Discussion**

Comparative genome analysis of the two smallest chromosomes of *T. gondii*, *chla* and *chlb* reveals a similar chromosome organization and content of genes and repeats between two of the three common lineages of *T. gondii*. Strikingly, analysis of the frequency of polymorphisms within these chromosomes reveals a markedly different pattern. *chlb* contains the expected frequency of SNPs seen in other regions of the genome of approximately one change per 100 bp. In contrast, the extremely low polymorphism rate on *chla* argues that the Type I and Type II lineages acquired this entire chromosome from a common source. While our sampling of the Type III lineages and other representative strains is more limited, given the relatively low rate of recombination in *T. gondii*, it is unlikely that the intervening regions differ from this pattern. Hence, our data suggest that all members of the three clonal lineages that predominate in North America and Europe possess the identical *chla*. Moreover, the estimated time of the establishment of this monoclonal chromosome is coincident with the origin of these clonal lineages ~10,000 yr ago (Su et al. 2003). Although the clonal lineages may be related by as few as two crosses, probability esti-
mates indicate that the common inheritance of chrIa alone is extremely unlikely to have arisen by chance. Rather, our analyses support a model where natural selection acting on the products of a recombination event in the wild led to the unusual fixation of chrIa.

Despite the results from the experimental cross that argue that chrIa monomorphism is not due to chromosome-specific repression of meiotic recombination, variations of a positive selective sweep model remain a possibility. Examination of a cross between relatively avirulent Type II and Type III T. gondii strains demonstrates that recombination can generate progeny far more virulent than either parent (Grigg et al. 2001a; Grigg and Suzuki 2003); not surprisingly, virulence involves multiple genes, as shown by linkage mapping studies (Su et al. 2002). The monomorphic chrIa could be the result of just such a recombination process, in which the generation of a particular combination of alleles from different parents led to successful expansion of the clonal lineages bearing this unique chromosome. While the precise mechanism at play is uncertain, one plausible model suggests that this event first allowed or greatly enhanced asexual oral transmission between intermediate hosts (Su et al. 2002). Parasites capable of such asexual transmission would be expected to expand rapidly via clonal propagation, thus limiting, but not excluding exposure to sexual, meiotic recombination. The special combination of alleles on chrIa would subsequently be maintained together by asexual growth, while still allowing the rare opportunity for recombination to establish related but distinct lineages (Fig. 4). At present, we cannot predict which gene or genes on chrIa might be responsible for the success of the clonal lineages, due to the fact that this chromosome has been maintained in entirety in all of the strains we have examined (all regions are equally associated with the phenotype; annotated genes are listed in Supplemental Table 3). Successful testing of this model will rely on identifying ancestral or exotic strains that express a wider range of asexual oral transmission phenotypes that we predict will have more variable chrIa genotypes.

Asexual transmission now affords T. gondii with population dynamics more akin to zoonotic viruses, where rare recombinants can spread via direct horizontal transmission (Worobey and Holmes 1999). This has reached a culmination in the example of the three clonal lineages that predominate in North America and Europe, where widely different geographic regions share identical multilocus genotypes. However, recent studies have revealed that the population structure of T. gondii in South America differs from this pattern and appears genetically more divergent (Ajzenberg et al. 2004; Lehmann et al. 2004; Khan et al. 2006; Su et al. 2006). It is presently uncertain whether a clonal population structure also predominates in these regions, albeit dominated by different genotypes, or whether they are truly panmictic. More severe forms of toxoplasmosis in humans have been associated with so-called atypical strains from such regions (Darde et al. 1998; Grigg et al. 2001b; Khan et al. 2006) and it is

![Figure 3. Progeny from genetic crosses between the Type I and Type III lineage were analyzed using six polymorphic markers on each of chrIa and chrIb. The resulting allele patterns are shown for each of 31 progeny at each locus (solid boxes, Type III lineage; open boxes, Type I lineage). Progeny where recombination has occurred are indicated at the bottom by an asterisk. The maps at the top represent the linear genetic maps for each chromosome with the markers shown on the right and the genetic distances shown on the left in centiMorgans.](image-url)
uncertain to what extent these strains share the monomorphic chrIa described here. Despite a recently derived and highly clonal population structure, the continued recombination of T. gondii in the wild, while infrequent, may lead to strains that acquire new pathogenic mechanisms and can readily expand in the population. Further investigations into the genetic basis of this and other mechanisms that profoundly alter transmissibility and virulence will provide a greater understanding of the evolution of toxoplasmosis and global parasitic disease in general.

Methods

Parasite culture for DNA preparation
Tachyzoites of the RH (American Type Culture Collection [ATCC] # 50174) strain were propagated in monolayers of human foreskin fibroblasts as described previously (Su et al. 2003). Parasite harvesting, DNA isolation, and shotgun library preparation for chrIa and chrIb are described in the Supplemental material.

Whole-chromosome sequencing and annotation
Whole-chromosome shotgun sequencing (Hall et al. 2003) was used to determine the 1.92 and 2.01 Mb DNA sequences of chrIa and chrIb for T. gondii strain RH. The methods used in annotating the RH sequence are provided in the Supplemental material.

Comparison of strain RH with strain ME49
The manually finished sequence of chrIa and chrIb from the RH strain (http://www.sanger.ac.uk/Projects/T_gondii/) was compared with a whole-genome shotgun assembly of the Type II ME49 (B7 clone) strain generated by TIGR (preliminary sequence data was obtained from The Institute for Genomic Research Web site at http://www.tigr.org). The polymorphisms were identified and classified as SNPs, insertion/deletions, and replacements (Supplemental material). Su et al. (2003) analyzed the frequency of SNPs in non-coding regions (including 11 introns and the ITS1 region) in four Type I, three Type II, and three Type III strains. The frequency of SNPs between Type I and Type II strains was estimated from Table S3 of their Supplemental online materials as 28 SNPs per 4067 bp or 6.9 SNPs per kilobase compared.

The SNPs are expected to follow a Poisson distribution if they are randomly distributed across chrIa. The chromosome was divided into “bins” and the number of SNPs in each bin were counted. Agreement with the Poisson distribution was assessed using the Poisson dispersion test. The result was not sensitive to the bin size chosen.

Recombination parameters for chrIa and chrIb
To identify genetic markers for chrIa, we analyzed 10-kb sequences spaced ~500 kb apart along the chromosome. We compared the genomes of Type I and Type II strains using CLUSTALX (Higgins et al. 1996) to identify potential SNPs. Once SNPs had been identified, we amplified and sequenced genomic regions flanking SNPs from all three lineages using the type strains—RH (Type I), ME49 (Type II), and VEG (Type III) (Supplemental material). CLUSTALX was used to align the sequences and identify strain-specific polymorphisms. Further analysis of these markers is described in the Supplemental material. In total, nine polymorphic markers were developed on chrIa. However, due to the biallelic pattern of T. gondii, only six of these detect differences between the Type I and Type III lineages, while three of them detect differences between the Type II and Type III lineages. Due to the larger number of markers and progeny available for the IxIII cross, we analyzed the segregation of markers in the 31 progeny from a previous cross (Su et al. 2002). Progeny were analyzed for segregation of polymorphisms on chrIa and chrIb by PCR amplification, restriction enzyme digestion, and gel electrophoresis to score alleles. The resulting genotypes were analyzed using MapMAKER (Lander et al. 1987) to generate genetic linkage maps of the chromosomes.

Acknowledgments
We acknowledge the support of the Wellcome Trust, the National Institutes of Health, the Wellcome Trust Sanger Institute for Human Genetics, the National Institutes of Health, the National Institutes of Health, the National Institutes of Health, the National Institutes of Health, the National Institutes of Health.
core sequencing and informatics groups, the Washington University Genome Sequencing Center, and The Institute for Genomic Research. We thank Andrew Berry, Wally Gilks, John Boothroyd, and Jon Boyle for discussion and review of the manuscript. Preliminary sequence data for the Type II ME49 (B7 clone) strain of Toxoplasma gondii was obtained from The Institute for Genomic Research Web site at http://www.tigr.org. Sequencing of Toxoplasma gondii was funded by the National Institute of Allergy and Infectious Disease.

References


Received March 21, 2006; accepted in revised form June 6, 2006.