Variation in the structure of Toxoplasma gondii and the roles of selfing, drift, and epistatic selection in maintaining linkage disequilibria

Tovi Lehmann a,∗, Douglas H. Graham b,1, Erica R. Dahl a, Lilian M.G. Bahia-Oliveira b, S.M. Gennari c, J.P. Dubey d

a Division of Parasitic Diseases, MS: F22, Centers for Disease Control and Prevention, 4770 Buford Hwy, Chamblee, GA 30041, USA
b Laboratorio de Biologia do Reconhecer, Centro de Biotecnologicas, Universidade Estadual do Norte Fluminense, Campos dos Goytacazes, Goytacazes 28015, Brazil
c Departamento de Medicina Veterinaria Preventiva e Saude Animal, Universidade de Sao Paulo, Cidade Universitaria, Sao Paulo SP-05508-080, Brazil
d Animal Parasitic Diseases, USDA, Agricultural Research Service, Annual and Natural Resources Institute, Beltsville, MD, USA

Received 7 October 2003; received in revised form 20 January 2004; accepted 20 January 2004

Available online 11 March 2004

Abstract

Previous studies of Toxoplasma gondii, based on samples dominated by clinical isolates, have concluded that its population structure is clonal, despite the sexual reproduction that occurs in cats. To determine whether this applies to non-clinical isolates, we compared patterns of linkage disequilibrium (LD) among seven loci in samples of T. gondii from Brazil and the US. LD was detected in both locations, but it was substantially lower in Brazil. The lower LD in Brazil can be explained by a higher rate of sexual reproduction between different genotypes (outcrossing) because of a higher rate of transmission. The extent of LD between pairs of physically unlinked loci varied significantly in each location. Moreover, the magnitude of LD between corresponding locus pairs in Brazil and the US was correlated, despite minimal gene exchange between the continents (mean Fs = 0.19). The heterogeneity among locus pairs and the correlation in LD between physically unlinked locus pairs from different continents suggests that locus-specific factors, such as epistatic selection are involved in maintaining LD in T. gondii. Possibly, the unique life cycle of T. gondii with its unpredictable transmission among diverse host species and distinct ecological habitats requires specific combinations of alleles from multiple loci. The usefulness of typing isolates based on physically unlinked loci is questioned not only by the geographic variation in the reproductive population structure, but mainly by the low overall predictability of the genotype of one locus based on the genotype in another (unlinked) locus. This predictability ranged between 23 and 45%, but was close to nil for a considerable fraction of locus pairs.

© 2004 Elsevier B.V. All rights reserved.

Keywords: Clonal structure; Linkage disequilibrium; Microsatellite; Molecular epidemiology; Sex; Strain typing; Toxoplasma; Zoonosis

1. Introduction

The population breeding structure of pathogens is key to understanding epidemiological patterns such as heterogeneity in disease manifestation, the emergence of new pathogens, and the evolution of drug resistance (Tibayrenc et al., 1991; Lenski, 1993; Smith et al., 1993; Grigg et al., 2001; Tibayrenc and Ayala, 2002). The perception that population structure is either sexual or clonal and that it is fixed for a given species has been challenged by the revelations that these states are merely extreme points in a continuum and that the population structure can vary considerably even in a single species (Souza et al., 1992; Anderson et al., 2000; Spratt et al., 2001; Tibayrenc and Ayala, 2002). Linkage disequilibrium (LD) analysis based on multi-locus genotyping has been the main tool to elucidate the population structure of many pathogens (cited above). LD measures the departure of the association of alleles of different loci from that expected by chance in a randomly mating population. High LD throughout the genome signifies that the whole genome is transmitted between generations as one unit, as in clonal reproduction, whereas linkage equilibrium or low LD signifies that different parts of the genome are
transmitted independently, as in sexual reproduction. The primary force shaping the population structure in the above studies was the frequency of sexual exchanges between distinct parasite genotypes within and between populations.

Toxoplasmosis is a common zoonotic disease worldwide (Dubey and Beattie, 1988). Usually sub-clinical, it can be severe and even fatal when infection is acquired congenitally (Jones et al., 2001) or when it occurs in immunocompromised persons (Belanger et al., 1999). Toxoplasma gondii is considered as an extreme example of a generalist organism because it infects virtually all mammal and bird species. In addition to infecting literally thousands of host species, the oocyst stage survives in such diverse environments as the frozen Alaskan soil, the mud of the Amazonian rain forest, and sea water (Zarnke et al., 2000; Carme et al., 2002; Miller et al., 2002). Upon infection, the haploid T. gondii multiplies asexually and then develops into a "chronic" stage (bradyzoite) that persists until the death of the host, but can initiate a new infection upon ingestion by a predator or a scavenger. Sexual reproduction occurs only in felids (domestic and wild cats), where male and female gametes are formed and fuse into a diploid zygote, which develops into an oocyst. The oocyst undergoes meiosis after being excreted to the environment in the cat's feces and produces two sporocysts, each containing four haploid sporozoites. Sexual reproduction in cats lasts 3–7 days and results in the shedding of millions of oocysts that persist in the environment for many months (Dubey, 1998; Lindsay et al., 2002). Infection results in life-long immunity of the host, and typically, a single parasite genotype is recovered from individual hosts (Howe et al., 1997, but see Aspinall et al., 2002; Aspinall et al., 2003). Experimental infection of cats with two distinct strains resulted in the expected 1:1 ratio of recombinants and parental types under random uniformization of gametes (Pfefferkorn and Pfefferkorn, 1980), and the US seeking to determine if the reproductive structure of T. gondii varied between and within countries, and evaluating factors that shape these structures.

2. Materials and methods

2.1. Isolates and genotyping

Isolates of T. gondii from domestic animals in Brazil and the US were included in this study. In Brazil, a total of 53 isolates were obtained from one collection of chickens near Sao Paulo (n = 23) and another collection near Campos dos Goytacazes (n = 30). These samples were described in detail previously (Dubey et al., 2002; da Silva et al., 2003). In the US, a total of 54 isolates were obtained from pigs (n = 33), chickens (n = 11), domestic cats (n = 8), a sheep and a goat. Most of the pig isolates were described previously (Dubey et al., 1995; Mondragon et al., 1998) as were most chicken isolates (Dubey et al., 2003a). Because the pig isolates were sampled from sows in an abattoir, their origin could not be ascertained and we suppose that the area encompassing their origin covers several states. Altogether, the isolates from the US represent a geographically more widespread sample than Brazil. Except for three isolates from the USA (from a chicken, a sheep, and a goat) that were maintained in the laboratory for years, the DNA of all other isolates represent the primary isolation of the parasite in mice or cats or the subsequent passage. Parasite isolation protocols were described in detail previously (see above citations).

DNA extraction was performed following Lehmann et al. (2000) and genotyping was performed as described in detail previously (Blackston et al., 2001). The genotype at the SAG2 locus was determined by the PCR-RFLP assay (Howe et al., 1997). The genotype at each of five dinucleotide microsatellite loci and a minisatellite locus (M95) were determined by PCR followed by fragment size determination as described by Blackston et al. (2001). Six of the seven loci were genetically mapped onto different linkage groups, except loci M95 and M102 which were 21 cM apart (Blackston et al., 2001; Su et al., 2002). Locus M48 was not mapped because both parental genotypes had the same allele (Table 1).

2.2. Data analysis

Genetic diversity was measured by the number of alleles per locus and the unbiased expected heterozygosity (\(H_u\), also known as gene diversity (Nei, 1987). Expected heterozygosity measures the probability of randomly selecting two different alleles in the population and can be used with diploid and haploid organisms. Differentiation between populations was estimated using GENEPop 3.3 (Raymond and Rousset, 1995a). Two locus linkage disequilibrium (LD) statistics, \(D^f\) (Lewontin, 1964) and \(R^2\) (Nei, 1987) were estimated based on available genotype data for the haploid genomes. These measures of the magnitude of LD better accommodate the effect of different allele frequencies on the extent of LD, and their properties have been extensively studied (Lewontin, 1964; Nei, 1987; Zapata, 2000; Zapata, 2000).
The overall R (“unweighted” or equally weighted) mean R values in parenthesis refer to the epidemic model (see text).

Locus a Brazil US All Pooled Differentiation

Table 1

Genetic diversity and differentiation of T. gondii from Brazil and the US

<table>
<thead>
<tr>
<th>Locus</th>
<th>N*</th>
<th>A*</th>
<th>H* (%)</th>
<th>N</th>
<th>A</th>
<th>H* (%)</th>
<th>A</th>
<th>H* (%)</th>
<th>FST</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAG2 (8)</td>
<td>53 (41)</td>
<td>2</td>
<td>41 (42)</td>
<td>53 (39)</td>
<td>3</td>
<td>51 (52)</td>
<td>3</td>
<td>67 (67)</td>
<td>0.46** (0.47***)</td>
</tr>
<tr>
<td>M6 (u2)</td>
<td>49 (30)</td>
<td>12</td>
<td>87 (89)</td>
<td>50 (37)</td>
<td>10</td>
<td>40 (43)</td>
<td>14</td>
<td>84 (87)</td>
<td>0.05** (0.02**)</td>
</tr>
<tr>
<td>M35 (4)</td>
<td>53 (41)</td>
<td>2</td>
<td>48 (51)</td>
<td>54 (40)</td>
<td>3</td>
<td>47 (47)</td>
<td>3</td>
<td>51 (51)</td>
<td>0.15** (0.10**)</td>
</tr>
<tr>
<td>M48 (7)</td>
<td>51 (41)</td>
<td>12</td>
<td>90 (90)</td>
<td>54 (40)</td>
<td>3</td>
<td>69 (71)</td>
<td>15</td>
<td>84 (83)</td>
<td>0.11** (0.09***)</td>
</tr>
<tr>
<td>M102 (7)</td>
<td>53 (41)</td>
<td>7</td>
<td>66 (61)</td>
<td>54 (40)</td>
<td>6</td>
<td>68 (60)</td>
<td>10</td>
<td>82 (81)</td>
<td>0.30*** (0.53***)</td>
</tr>
<tr>
<td>M163 (u1)</td>
<td>51 (40)</td>
<td>12</td>
<td>89 (90)</td>
<td>52 (39)</td>
<td>15</td>
<td>91 (91)</td>
<td>22</td>
<td>94 (95)</td>
<td>0.08*** (0.08***)</td>
</tr>
<tr>
<td>M95 (7)</td>
<td>52 (40)</td>
<td>5</td>
<td>71 (67)</td>
<td>53 (39)</td>
<td>3</td>
<td>63 (62)</td>
<td>6</td>
<td>76 (74)</td>
<td>0.18** (0.22***)</td>
</tr>
<tr>
<td>Mean</td>
<td>51.7 (40.4)</td>
<td>7.4</td>
<td>70.5 (69.9)</td>
<td>52.9 (39.3)</td>
<td>7.4</td>
<td>67.3 (67.4)</td>
<td>10.4</td>
<td>77 (77)</td>
<td>0.19** (0.19***)</td>
</tr>
</tbody>
</table>

Values in parentheses refer to the epidemic model (see test).

Values in parentheses refer to the epidemic model (see test).

* Number of isolates genotyped.
* Observed number of alleles.
* Unbiased expected heterozygosity (Nei, 1987).
* Based on (Weir and Cockerham, 1984).
* Significance level: P < 0.05.
* Significance level: P < 0.01.
* Significance level: P < 0.001.

The possible values of both statistics range between 0, at equilibrium, and 1, at extreme disequilibrium. These statistics were calculated for all combinations of allele pairs (allele level). The overall averages represent the LD at the locus level rather than the allele level. Two-locus genotypes with a singleton allele (an allele appearing only once in our sample) were excluded from the disequilibrium analysis. The significance of LD for each pair of alleles was tested by the 1df, $\chi^2 = nD_{ij}^2 /[p_i(1 - p_i)q_j(1 - q_j)]$, where $n$ is the number of individuals sampled and $p_i$ and $q_j$ denote the frequency allele $i$ from one locus and allele $j$ from the other locus.

The significance of the LD for each pair of loci was tested between North and South America was high (mean and median $F_{ST}$ were 0.19 and 0.15, respectively, Table 1), consistent with very limited migration between North and South America. Notably, the magnitude of differentiation varied greatly between loci (range 0.02–0.47, Table 1), suggesting that different parts of the genome evolve under different selective pressures as has been previously observed (Lehmann et al., 2000). Even after excluding loci with the least (M6) and greatest (SAG2) $F_{ST}$ values, which might
Table 2
Significance of LD based on exact tests on the single and multiple (employing the sequential Bonferroni adjustment for the total number of tests in each location and \( P < 0.05 \) tests levels)

<table>
<thead>
<tr>
<th>Population model</th>
<th>Pairs</th>
<th>( N^a )</th>
<th>Brazil (% significant tests) (N)</th>
<th>US (% significant tests) (N)</th>
<th>Brazil vs. US</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Single</td>
<td>Multiple</td>
<td>Single</td>
</tr>
<tr>
<td>Stable Alleles</td>
<td>53/54</td>
<td>13.8 (828)</td>
<td>3.1 (828)</td>
<td>30.9 (528)</td>
<td>8.0 (528)</td>
</tr>
<tr>
<td>Stable Loci</td>
<td>53/54</td>
<td>76 (21)</td>
<td>71 (21)</td>
<td>91 (21)</td>
<td>91 (21)</td>
</tr>
<tr>
<td>Epidemic Alleles</td>
<td>41/40</td>
<td>11 (654)</td>
<td>1 (654)</td>
<td>22 (450)</td>
<td>6 (450)</td>
</tr>
<tr>
<td>Epidemic Loci</td>
<td>41/40</td>
<td>62 (21)</td>
<td>33 (21)</td>
<td>76 (21)</td>
<td>71 (21)</td>
</tr>
</tbody>
</table>

The fraction of significant tests is used as a measure of the extent of LD in each population.

* Number of isolates.

...disproportionately bias estimates of genomic differentiation, high mean \( F_{ST} (0.164) \) persisted. A number of private alleles (with frequency larger than 5%) were observed in each continent in most loci (not shown); together with similar genetic diversity they indicate that the high differentiation was created by prolonged isolation rather than by a bottleneck-like event. There was no evidence for differentiation between the two Brazilian populations (mean \( F_{ST} \) was \(-0.001\) and \(-0.005\) for the stable and epidemic population models, respectively) despite being located nearly 600 km apart.

3.2. Linkage disequilibrium between loci

Our analyses addressed (i) whether the association of alleles from different loci exceeded the departures expected under equilibrium due to sample size, i.e., the significance of LD indices, and (ii) the intensity of the departure from equilibrium indicated by the magnitude of the LD indices. Between-population differences in the extent of LD was evaluated by comparing the fraction of significant LD tests in each population (given similar sample sizes), as well as the values of \( |D'| \) and \( R^2 \).

Significant departures from linkage equilibrium between pairs of loci were detected in the US and Brazil (Table 2). In both locations, significant LD was less frequently observed under the epidemic population model than under the stable model, although this difference was significant only in Brazil (\( P < 0.008 \), <0.03, for the alleles and the loci tests, respectively, Fisher exact test). Thus, certain genotypes may be over represented in the Brazilian sample, either as a consequence of “epidemic expansion” or sampling bias. Therefore, subsequent analyses were performed assuming the epidemic model, unless otherwise specified. The fraction of significant LD tests was higher in the US but the difference was not always significant (Table 2).

Unlike the physically linked loci M95 and M102 (21 cM apart, Su et al., 2002) that had high \( |D'| \) values, all pairs of loci with locus M48 (for which no mapping information was available) had moderate or low values of \( |D'| \) in both populations (Fig. 1). Thus, we supposed that M48 was not physically linked with the other loci.

\( |D'| \) varied considerably among locus pairs in both Brazil and the US (0.35–0.78 and 0.39–0.96, respectively under the epidemic model, Fig. 1). In both populations, smaller
variation in $|D'|$ was found among alleles within a locus-pair than between locus pairs ($P < 0.002$, d.f. = 20, $x^2 > 44$, Wilcoxon Rank Sum Test of $|D'|$), suggesting heterogeneity in LD between loci. Heterogeneity remained ($P < 0.005$, d.f. = 18, $x^2 > 37$) after excluding SAG2-M33 (with only two alleles per locus in Brazil) and M102-M95 (21 cM apart), and even after excluding rarely expected allele pairs (expected frequency < 2%). At the locus pair level, heterogeneity tested by bootstrapping over isolates was significant only in the US ($P < 0.027$, one side bootstrap test). To test our hypothesis ($|D'|_{ij} - |D'|_{kl} = 0$, subscripts denote any combination of different locus pairs with $j \neq l$), we calculated $10^4$ weighted $|D'|$ for each locus pair by bootstrapping over isolates (haplotypes). All values were centered on zero by subtracting the corresponding locus-pair mean, and $10^4$ random samples of 21 bootstrap values (one from each locus pair) were drawn to calculate the range (maximum–minimum). The bootstrap distribution of the range was compared with the observed range. This analysis was separately performed on the Brazilian isolates and on those for the US.

The magnitude of LD, measured by $R^2$ and the weighted $|D'|$ between locus pairs, was lower in Brazil than in the US when the range was compared with the observed range. This analysis was separately performed on the Brazilian isolates and on those for the US.

To investigate whether the extent of LD between a given pair of loci was similar in the Brazilian and US samples, we evaluated the correlations in the $R^2$ and $|D'|$ between locations. Although LD for a given pair of loci was consistently smaller in Brazil than in the US, the extent of LD was positively correlated in both geographic samples (Fig. 2 and Table 3). Significant correlation, however, was only found for the $R^2$ (Fig. 2 and Table 3). To minimize the possible bias stemming from variation in the level of polymorphism among loci, we also calculated these correlations after excluding allele pairs if either allele was observed fewer than four times and found significant correlations with both, $R^2$ and $|D'|$ (Table 3).

Table 3: Magnitude of LD in Brazil and the US and the correlation between locations

<table>
<thead>
<tr>
<th>Model</th>
<th>Index</th>
<th>Brazil (mean)</th>
<th>US (mean)</th>
<th>Difference (US–Brazil)</th>
<th>$r$ (US–Brazil)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stable</td>
<td>$</td>
<td>D'</td>
<td>$</td>
<td>0.60</td>
<td>0.81</td>
</tr>
<tr>
<td></td>
<td>$R^2$</td>
<td>0.03</td>
<td>0.14</td>
<td>0.11</td>
<td>0.09** (0.15**)</td>
</tr>
<tr>
<td>Epidemic</td>
<td>$</td>
<td>D'</td>
<td>$</td>
<td>0.57</td>
<td>0.77</td>
</tr>
<tr>
<td></td>
<td>$R^2$</td>
<td>0.07</td>
<td>0.18</td>
<td>0.10</td>
<td>0.08** (0.12**)</td>
</tr>
</tbody>
</table>

---

* $*$: $x^2$ (Reynolds, 1984) measures the standardized gain in predictability of the genotype in one locus provided information on the individual genotype by the other locus.

* Mean paired difference. Significance is based on Wilcoxon Sign Rank Test.

* Pearson correlation coefficient between corresponding loci pairs from Brazil and the US. Significance is based on Mantel test. Values in parentheses refer to some statistics after excluding the physically linked locus pair M102-M95 (21 cM apart, Su et al., 2002) and allele pairs if the absolute observed frequency of any allele was smaller than four (see text for detail).

* Significance level: $P < 0.05$.

** Significance level: $P < 0.01$.

*** Significance level: $P < 0.001$. 

---

Fig. 2: The relationship between LD indices ($|D'|$ and $R^2$) in Brazil and the US under the epidemic population model. All data are included. The dotted line denotes equal values of LD in Brazil and the US.
4. Discussion

Our main findings include (i) high differentiation between *T. gondii* populations from Brazil and the US, (ii) LD was detected in both locations, but was substantially lower in Brazil, (iii) the extent of LD within populations varied among locus pairs, and LD of corresponding locus pairs in Brazil and the US was correlated, and (iv) predictability of the genotype in one locus given the genotype in another locus was not high. Below we evaluate these and other findings and consider the role of other factors in maintaining this structure, in addition to high rate of selfing.

Remarkably high geographic differentiation was found between *T. gondii* from Brazil and the US, whereas no differentiation was detected between the Brazilian subpopulations (600 km apart). Together with finding many private alleles in each continent and similar genetic diversity in both populations, these results indicate prolonged evolutionary separation between Brazil and the US (unlike differentiation produced by a bottleneck). The barrier limiting gene exchange between these regions has yet to be identified, and the plausible role of migratory birds in spreading *T. gondii* merits further attention because bird migration along the north–south axis is extensive (Elphick, 1995). Importantly, the different genotype composition in Brazil may be a contributing factor in the unusually high frequency of severe ocular toxoplasmosis reported there (Silveira et al., 2001).

The ability to predict the phenotype from the genotype is of special interest in pathogens because it may allow us to predict disease manifestations (Lan and Reeves, 2001). The perception that the reproductive structure of *T. gondii* is clonal formed the basis for strain typing, i.e., predicting the genotype throughout the genome from knowledge of the genotype in a single locus (Howe and Sibley, 1995). To evaluate the extent of this predictability in our data, we calculated the symmetric coefficient $\lambda$ (Reynolds, 1984), which varies between 0 and 1, and in this case measures the standardized gain in predictability of the genotype in one locus provided information on the individual genotype in another (physically unlinked) locus. The range of the average $\lambda$ was 23–45% (Table 3). The $\lambda$ values were not high even for loci with very few alleles. For example, in Brazil $\lambda$ was 0.0 for loci SAG2 (two alleles) and M33 (three alleles). Other evidence against clonal population structure was provided by the heterogeneity in *P* values (Table 1 and Lehmann et al., 2000), because locus-specific selection implies dissociation between the loci. Therefore, strain typing to predict a phenotype of *T. gondii* based on loci that are not linked to genes conferring a particular phenotype should be approached with great caution. For instance, mouse virulence is conventionally attributed exclusively to all lineage I isolates (e.g., Su et al., 2002), but in South America eight out of 63 lineage I isolates were avirulent (Dubey et al., 2001, 2003b, and unpublished); corroborating our inference based on LD above.

Although the significant LD detected in all sampled locations was consistent with previous studies (Sibley and Boothroyd, 1992; Howe and Sibley, 1995; Darde, 1996; Ajzenberg et al., 2002; Tibayrenc and Ayala, 2002), we found substantial geographic variation in its magnitude. Overall, LD was considerably greater in the US than in Brazil, indicating that the rate of selfing is higher in the US. A higher rate of toxoplasmosis transmission in Brazil would facilitate a higher rate of outcrossing because higher proportion of cats would be infected with two different genotypes in 1 or 2 days. Human prevalence is higher in Brazil than in the US (e.g., Jones et al., 2001; Bahia-Oliveira et al., 2003, and references cited therein), consistent with higher rate of transmission in Brazil.

Selfing, clonal reproduction (in non-felid hosts), and drift undoubtedly contribute to maintenance of LD in populations of *T. gondii*. However, does their combined effect fully account for the observed LD? Importantly, all these forces affect LD across the whole genome uniformly, whereas selection for or against certain combinations of alleles from different loci (epistasis) affects only the loci involved and physically linked loci. Therefore, heterogeneous distribution of LD among (physically unlinked) locus pairs is evidence for epistasis. The homogeneity of LD among pairs of loci was rejected by the analyses at the allele-pair level (even after excluding rare alleles and M102-M95 and M33-SAG2 locus pairs). Additionally, homogeneity was rejected by the bootstrapping analysis at the locus pair level in the US sample. Complementary results were based on the correlation in the magnitude of LD in corresponding locus pairs between Brazil and the US. Because selfing, clonal reproduction, outcrossing, and drift are random processes with respect to their effect on any (physically unlinked) locus pair, they cannot explain this correlation. Gene flow could produce such correlation, but we found evidence that gene flow between these populations is negligible (above). These results suggest that epistasis plays a role in maintaining LD in *T. gondii*. Accordingly, the frequency of outcrossing events is enough to dissociate different loci, but many recombinants are removed by selection. Had epistasis alone maintained complete LD (|D'| = 1) between two independent loci, 50% of the progeny of double heterozygotes would die every generation, and a higher fraction would die if more than two loci were epistatically linked. Actually, the burden expected on *T. gondii* is not as heavy because (i) expected |D'| is considerably lower than 1 (e.g., 0.56 in Brazil), and this magnitude of LD is maintained not only by epistasis but also by selfing, clonal propagation in non-felid hosts, and drift, implying that many genotype combinations are viable, (ii) the number of oocysts typically produced is greater than $10^7$ (Dubey and Frenkel, 1976), even so the loss of half of the oocysts does not impact the reproductive potential the way it would if the number of progeny were small, and (iii) since meiosis occurs within the oocysts, each recombinant oocyst contains four different genotypes of sporozoites because of the assortment of maternal and paternal
alleles cannot be observed, resulting in a small size, many allele combinations involving low frequency alleles have many alleles at low frequency. Using a moderate sample size and many allele combinations, we can observe frequent and unpredictable change in extremely different environments (diverse host species and markedly varied external conditions). Such a life cycle probably requires an extraordinary degree of compatibility between various sets of genes.

We cannot entirely rule out, however, that the heterogeneity among locus pairs in LD and the correlation discussed above are a result of different polymorphism and different mutation rates among loci. Highly polymorphic loci tend to show frequent and unpredictable change in extremely different environments (diverse host species and markedly varied external conditions). Such a life cycle probably requires an extraordinary degree of compatibility between various sets of genes.

Although most studies in natural populations of sexually reproducing species have not found evidence for epistasis (e.g., Langley et al., 1974; Wallace, 1991; Lehrmann et al., 1998), it fits especially well with T. gondii ability to ensure frequent and unpredictable change in extremely different environments (diverse host species and markedly varied external conditions). Such a life cycle probably requires an extraordinary degree of compatibility between various sets of genes.

References


