Globalization and the population structure of *Toxoplasma gondii*

Tovi Lehmann*†‡, Paula L. Marcet*, Doug H. Graham*§¶, Erica R. Dahl*¶‖, and J. P. Dubey**

*Division of Parasitic Diseases, Centers for Disease Control and Prevention, 4770 Buford Highway, Chamblee, GA 30341; †Laboratory of Malaria and Vector Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health, 12735 Twinbrook Parkway, Rockville, MD 20852; and **Animal Parasitic Diseases Laboratory, Animal and Natural Resources Institute, Agricultural Research Service, U.S. Department of Agriculture, Beltsville, MD 20705

Edited by Francisco J. Ayala, University of California, Irvine, CA, and approved May 30, 2006 (received for review February 21, 2006)

***Toxoplasma gondii* is a protozoan parasite that infects nearly all mammal and bird species worldwide. Usually asymptomatic, toxoplasmosis can be severe and even fatal to many hosts, including people. Elucidating the contribution of genetic variation among parasites to patterns of disease transmission and manifestations has been the goal of many studies. Focusing on the geographic component of this variation, we show that most genotypes are locale-specific, but some are found across continents and are closely related to each other, indicating a recent radiation of a pandemic genotype. Furthermore, we show that the geographic structure of *T. gondii* is extraordinary in having one population that is found in all continents except South America, whereas other populations are generally confined to South America, and yet another population is found worldwide. Our evidence suggests that South American and Eurasian populations have evolved separately until recently, when ships populated by rats, mice, and cats provided *T. gondii* with unprecedented migration opportunities, probably during the transatlantic slave trade. Our results explain several enigmatic features of the population structure of *T. gondii* and demonstrate how pervasive, prompt, and elusive the impact of human globalization is on nature.**

evolutionary history | migration | pandemic genotype | protozoan parasite | trade

Upon infection, *Toxoplasma gondii* multiplies asexually and develops into a persistent stage that initiates a new infection when ingested by a predator or scavenger. Sexual reproduction occurs only in domestic or wild cats and results in shedding (by defecation) millions of oocysts that can survive for months (1, 2) until ingested by a new host. Toxoplasmosis is usually subclinical, but it can cause mental retardation, blindness, and death. The disease is especially severe in congenitally transmitted cases and is also an important opportunistic infection in AIDS. Motivated to evaluate the contribution of genetic variation in parasites to patterns of toxoplasmosis transmission and manifestation, studies on *T. gondii*'s population structure were undertaken, and a perplexing picture has emerged. Despite its worldwide distribution and capacity to infect virtually all mammal and bird species (3), the genetic diversity of *T. gondii* was found to be remarkably low (4–7), and geographic variation was virtually absent (5, 8–10; but see ref. 11). Despite the essential sexual cycle, *T. gondii*'s population structure was found to consist of three clonal lineages (named I, II, and III), and recombinants among them were rare (<5%) (4, 5, 7–10, 12, 13). Recent studies, however, showed that sexual recombination, especially in South America, plays an important role in shaping *T. gondii*'s genetic structure (6, 11, 14). Most polymorphic loci revealed only two divergent allele classes, suggesting that recombination event(s) between these ancestral populations led to the emergence of the third lineage (6, 8, 15). Unlike studies that used isolates from an assortment of hosts and mostly from clinical human cases from Europe and North America, we sought to elucidate the population structure of *T. gondii* using previously unanalyzed, asymptomatic infections in a single host species. Our sample included 275 independent isolates (of >450) from free-ranging domestic chickens collected around the world (Table 2, which is published as supporting information on the PNAS web site, and Fig. 1). We analyzed genetic diversity at seven polymorphic loci, including a minisatellite and five short tandem repeat (STR) loci (16) as well as the *SAG2* (17).

Results

Moderate to high diversity was observed within all populations (Table 2, Fig. 1). High multilocus genotype (haplotype) diversity within populations (90–100%) revealed that local populations comprised low-frequency haplotypes, consistent with stable rather than epidemic transmission. Principal components analysis (PCA) was used to summarize the information contained in 19 measures of per-locus diversity of each population (expected heterozygosity and allele richness for seven loci and variance in repeat number for the five STR loci) (Table 2). The PCA revealed three clusters along the first principal component (Fig. 2). Five South American populations represented the highest diversity, whereas all other populations, except Grenada, represented moderate diversity. The low diversity of Grenada is probably attributed to genetic drift and founder effect typical of small islands. Heterogeneity among South American populations was also apparent, with southern populations possessing the highest diversity.

A phylogenetic network summarizing relationships among five STR haplotypes was derived by using the median-joining methods implemented by the program NETWORK 4.1 (18, 19). The network, incorporating variation in the number of repeats between alleles, revealed multiple clusters and branches with no clear discontinuity separating them (Fig. 3). Members of each lineage concentrated in different regions of the network: lineage II at its lower right area, lineage III at its upper left area, and lineage I in the center, but many haplotypes were misplaced with respect to their lineage membership. Loci were weighted inversely to their diversity (mutation rate) to reduce the effect of homoplasy (same allele length derived from independent origins) on misplacement of haplotypes. Furthermore, many haplotypes are located at a large distance from other members of their lineage, indicating either homoplasy events in multiple loci, which is rare, or recombination. Considering genotypes found on a single continent, Eurasian genotypes concentrated in the lower right area, whereas those from South America concentrated in the center, suggesting a geographical division between the new and

POPULATION BIOLOGY

Conflict of interest statement: No conflicts declared.

This paper was submitted directly (Track II) to the PNAS office.

Freely available online through the PNAS open access option.

Abbreviation: STR, short tandem repeat.

‡To whom correspondence should be addressed: at the present address: Laboratory of Malaria and Vector Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health, 12735 Twinbrook Parkway, Room 2W13A, Rockville, MD 20852. E-mail: tlehmann@niaid.nih.gov.

§Present address: Department of Biomedical Sciences, Grand Valley State University, 218 Padnos Hall, 1 Campus Drive, Allendale, MI 49401.

¶D.H.G. and E.R.D. contributed equally to this work.

‖Present address: Department of Microbiology and Immunology, University of Texas Medical Branch, 301 University Boulevard, Galveston, TX 77555-1019.

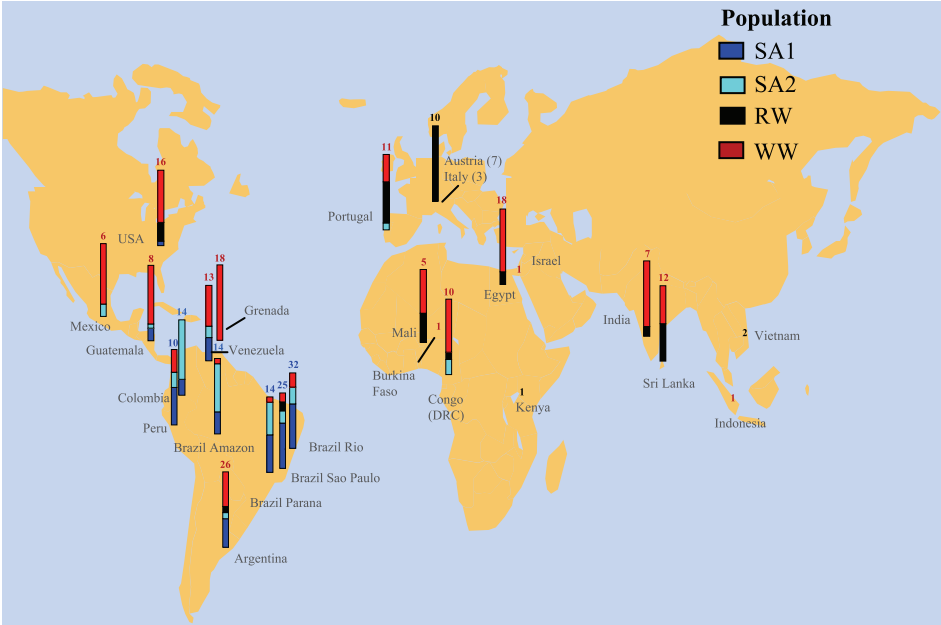

Fig. 1. A schematic showing the geographic origin of the samples ($n = 275$). Vertical bars over sites (where $n \geq 5$) depict sample composition with respect to the four populations identified by the program STRUCTURE (SA1 and SA2 predominate in South America; RW is common in all continents but is rare in South America, WW is cosmopolitan, see *Results* for details). Numbers indicate sample size from each locale, and their colors correspond to that of the dominant population in the sample.

the old world. This division was not sharp, however, because of the presence of four North American isolates among those from the old world, the distribution of African genotypes almost evenly between old and new world clusters, and because all sections of the network included at least a few South American genotypes.

Most haplotypes were sampled one to three times and were found on one continent. A notable cluster (arrow in Fig. 3) consisted of high-frequency haplotypes, including two, each representing 12 isolates from five continents, and two others representing 7 isolates from four continents. The distribution of shared-alleles distance among the multicontinent haplotypes was dramatically shorter than that among haplotypes found on one continent (Fig. 4; Wilcoxon two-sample test statistic, 1,635.5; $Z_{[2\text{-sided}]} = -6.46$; $P < 0.0001$). The extremely close genetic relationship between the multicontinent haplotypes suggests a recent "radiation" of a single genotype with exceptional capacity for long-distance migration.

A different approach to evaluate population structure relies on a Bayesian statistical model for clustering genotypes into populations without information on their origin. Recently implemented in the program STRUCTURE (20), this approach uses iterative computation process to simultaneously assign multilocus genotypes into populations and estimate the probability of observing the data, given the number of populations and their estimated allele frequencies.

By using the admixture model with independent allele frequencies, the likelihood of the data increased substantially from one population to five (Fig. 5 *Inset*), indicating that the gene pool was subdivided. Identical assignment of isolates among independent simulations (with the same number of populations) was achieved

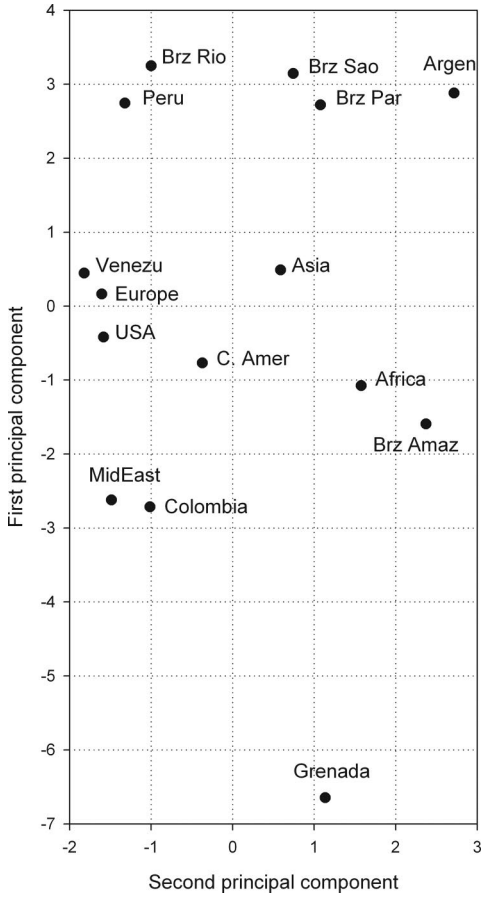

Fig. 2. Clustering of populations based on their genetic diversity was performed by using standard principal component (PC) analysis using the correlation matrix of the original variables. Per-locus estimates of expected heterozygosity (27), allele richness (28), and variance in allele size for the STR loci were used for each population (Table 2). Coordinates are the first (vertical) and second (horizontal) PCs. The first PC represented overall diversity because its eigenvector's loadings were positive and similar in magnitude (data not shown). The first PC alone accounted for 40% of the total variation, and together with the second PC, 53% of the total variation was captured. Brz Rio, Brazil Rio; Brz Sao, Brazil Sao Paulo; Argen, Argentina; Brz Par, Brazil Parana; Venezu, Venezuela; C. Amer, Central America; Brz Amaz, Brazil Amazon; MidEast, Middle East.

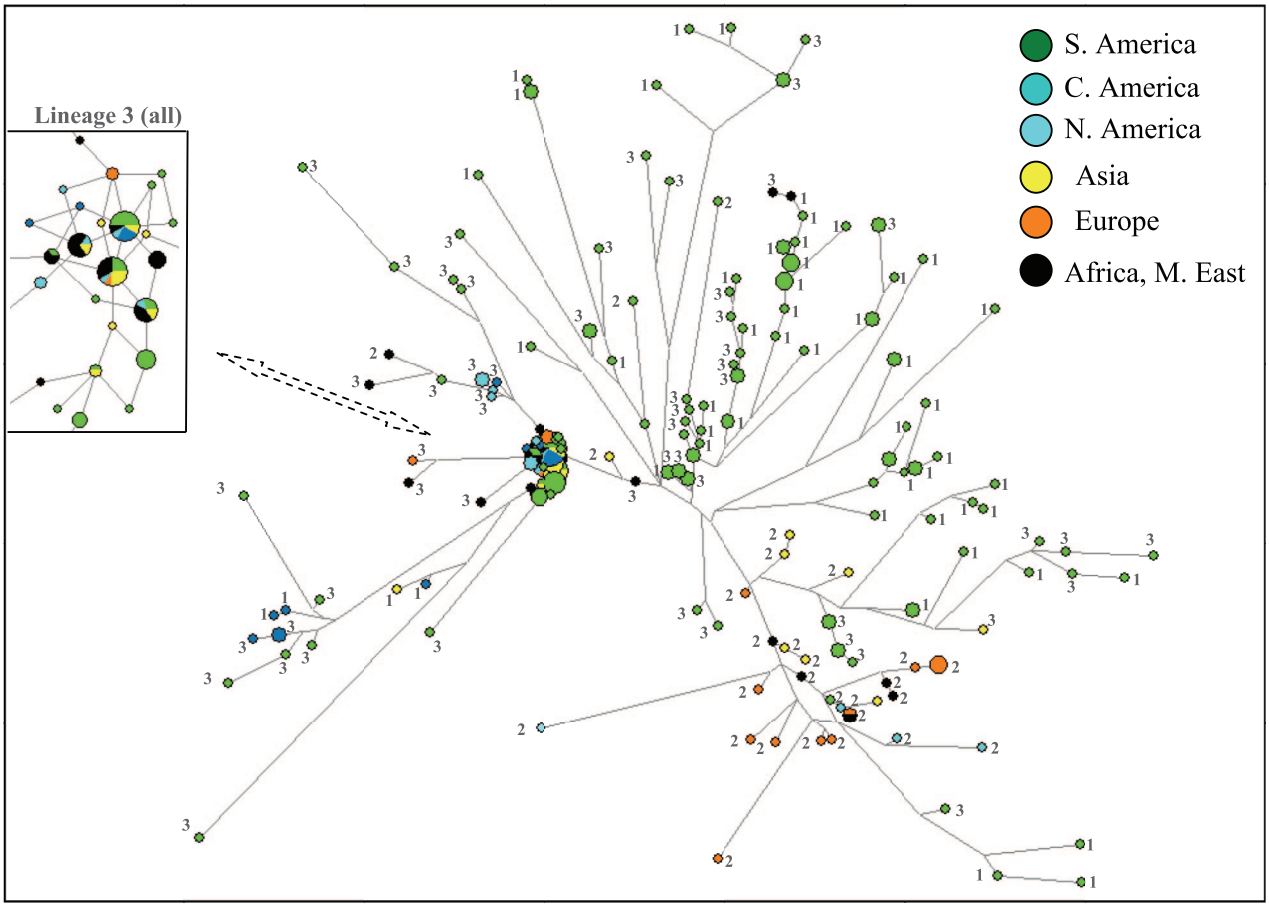

Fig. 3. Phylogenetic network showing relationships among five STR haplotypes in relation to their geographic origin and lineage (marked by 1, 2, and 3, respectively). (*Inset*) Magnification of the area showing the tight cluster of lineage III haplotypes (arrow). The network was derived by using the median-joining algorithm (18) ($\varepsilon = 0$) after processing the data with the reduced median method (19) as implemented by NETWORK 4.1. The network incorporated variation in the number of repeats (assuming the step-wise mutation model) of STR loci that were weighted inversely to their variance (M33, 9; M6, M48, and M102, 4; and M163, 3). S. America, South America; C. America, Central America; N. America, North America; M. East, Middle East.

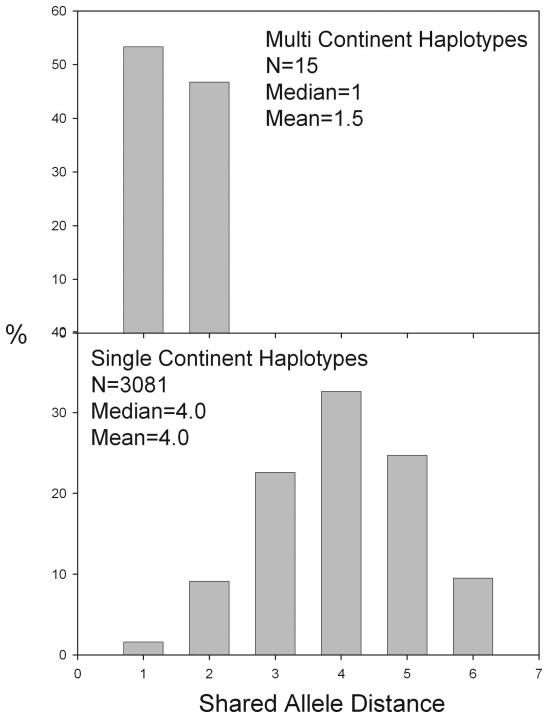

Fig. 4. Within-group divergence between haplotypes observed on two or more continents (*Upper*) and those observed on one continent (*Lower*). Divergence was measured by the distribution of allele-sharing distance across seven loci (including the minisatellite M95, which was not included in the STR network). To avoid sampling bias, we excluded identical haplotypes. Because all (but one) multicontinent haplotypes were of lineage III, the comparison included only haplotypes of this lineage; otherwise, the difference between the distributions was even larger.

with four, but not with five or more, populations. Further, the frequency of individuals that were assigned into any population with probability >75% sharply dropped if the number of populations was increased over four (Fig. 5 *Inset*). These results indicate that the gene pool is divided into four populations. The populations formed by the program STRUCTURE corresponded well to different sections of a neighbor-joining tree based on shared-allele distance (Fig. 5). Although the populations were correlated with the lineages (Table 1), every population comprised individuals from all lineages, and showed fewer misplaced haplotypes compared with the lineages (Fig. 5).

The four populations formed by STRUCTURE fell into three geographical divisions (Table 1). Two populations (hereafter called SA) were generally confined to South and Central America, and one population (hereafter called RW) was found in Europe, Asia, Africa, and North America, but was virtually absent from South and Central America. The fourth population (hereafter called WW),

POPULATION BIOLOGY

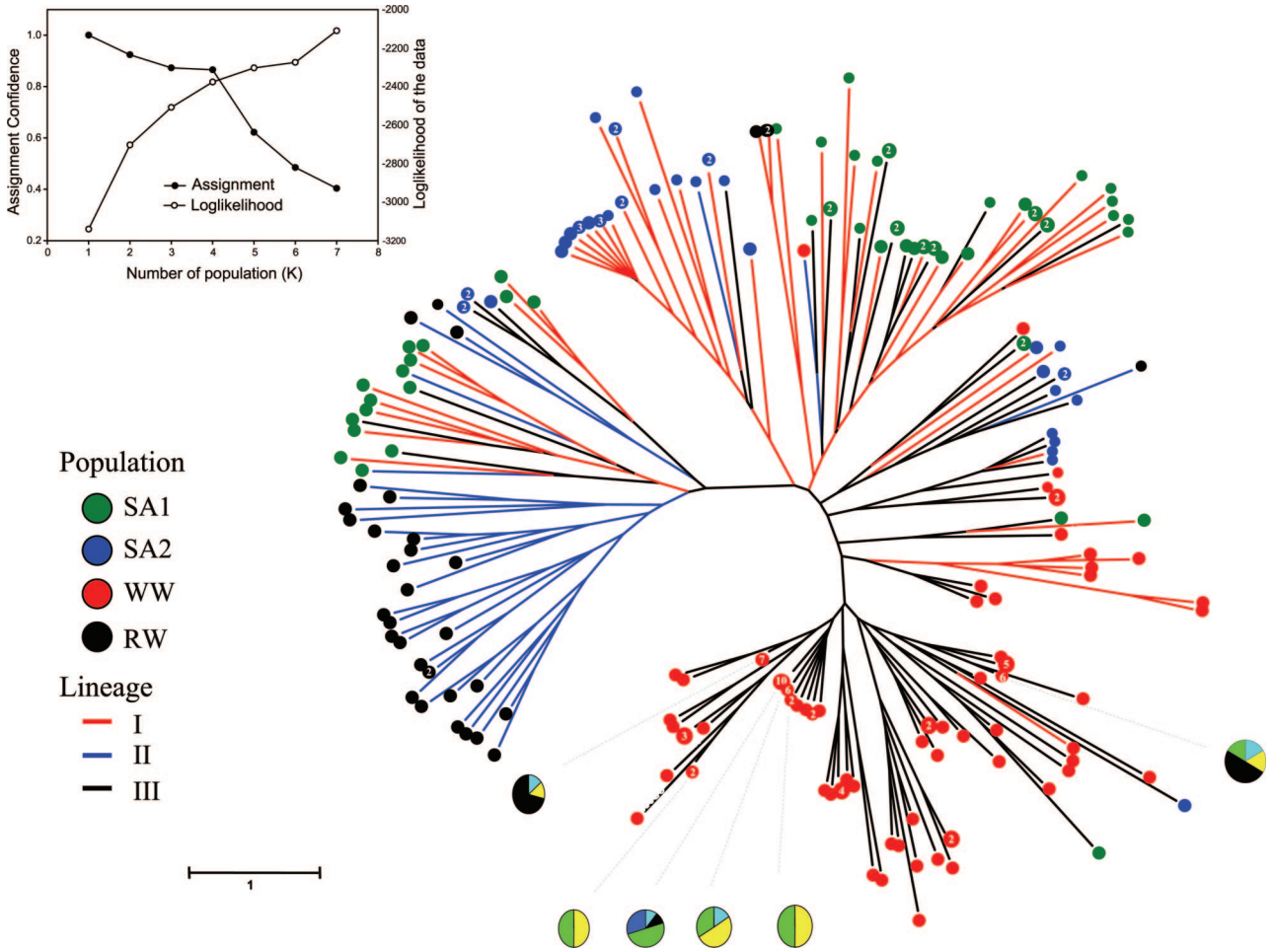

Fig. 5. Neighbor-joining tree of haplotypes based on the shared-allele distance across six loci, showing lineage (branch color) and populations identified by STRUCTURE (tip color). The observed frequency of each haplotype is shown if >1. Pie charts show geographical distribution of haplotypes found on two or more continents (color key as in Fig. 3). The distribution of multicontinent haplotypes differs slightly from the network (Fig. 3) because an additional locus (M95) is included in generating the shared-allele distance tree. (*Inset*) Determination of the number of populations (K) in the *T. gondii* gene pool using the admixture model (with independent allele frequencies) implemented by STRUCTURE, based on the likelihood of observing the data and the assignment certainty (the fraction of isolates assigned into any population with probability >75%). The results are averages across three independent simulations (with 10^5 burn-in iterations, followed by 10^6 MCMC iterations).

however, was cosmopolitan. This organization implies that long-term isolation and extensive migration acted simultaneously between continents. Understanding how this unusual structure was shaped starts with exploring the genetic interrelationships among these divisions.

By using the admixture model with the correlated allele frequencies model (F model) in STRUCTURE (21), posterior F_{ST} values were computed, measuring the divergence of each population from their common "ancestral" population. This model provided nearly identical clustering to the admixture model, with only 13 of 275 individuals clustered differently (4.7%, data not shown). F_{ST} values of the SA populations (0.06 and 0.10) were substantially smaller than that of the RW (0.16) and WW (0.27) populations (Fig. 6), indicating that the SA populations diverged little from the ancestral population of *T. gondii*. A high F_{ST} value indicates that the population associated with it had smaller effective population size or, in other words, that it experienced a strong genetic drift (21). Within-population divergence, measured by the shared-allele distance, was greatest in one of the SA populations (Fig. 6; Wilcoxon two-sample test statistic, >187,500; $Z_{[2\text{-sided}]} < -9$; $P < 0.0001$) and was significantly smaller in the WW population (Fig. 6; Wilcoxon two-sample test statistic, 755,497.5; $Z_{[2\text{-sided}]} = 9$; $P < 0.0001$). Together with highest locus-specific diversity in SA populations (Fig. 2 and Table 2), these complementary results indicated that *T. gondii* evolved in SA longer than in any other continent; in other words, that South America was the "birthplace" of *T. gondii*. Despite its widest geographical range, the WW population possessed the lowest within-population divergence and the highest divergence from the ancestral population (Fig. 6), indicating that it has evolved most recently. Both between-population and within-population divergence were intermediate in the RW population (Fig. 6).

Discussion

Hypervariable loci captured variation generated in *T. gondii*'s most recent evolutionary history as well as its deeper history, as evidenced by the correlations with *SAG2*. The results revealed small genetic differences among *T. gondii* populations from Eurasia, Africa, and North America but large differences between them and South American populations. Notably, North American populations were similar to those from Eurasia and Africa, despite vast oceans separating them. Furthermore, certain *T. gondii* genotypes spread globally very recently, as evidenced by the short mutational distances among them. This organization implies that long-term isolation and extensive migration acted simultaneously between continents. Thus, the geographical structure of *T. gondii* is more complex than previously realized, and understanding its evolutionary history must explain these unusual findings.

Table 1. The geographical and lineage composition (percentage and the number of isolates) of the populations identified by STRUCTURE

Region	SA1, % (*n*)	SA2, % (*n*)	RW, % (*n*)	WW, % (*n*)	Total (*n*)
Asia	0 (0)	0 (0)	41 (9)	59 (13)	22
Europe	0 (0)	5 (1)	76 (16)	19 (4)	21
Mid East	0 (0)	0 (0)	16 (3)	84 (16)	19
Africa	0 (0)	12 (2)	23 (4)	65 (11)	17
N. America	6 (1)	0 (0)	25 (4)	69 (11)	16
C. America	14 (2)	7 (1)	0 (0)	79 (11)	14
S. America	40 (67)	26 (43)	3 (5)	31 (51)	166
Total (*n*)	70	47	41	117	275
Lineage*					
I	51 (40)	33 (26)	4 (3)	12 (9)	78
II	7 (3)	7 (3)	84 (36)	2 (1)	43
III	17 (25)	12 (18)	0 (0)	71 (107)	150
Total (*n*)	68	47	39	117	271

Mid East, Middle East; N. America, North America; C. America, Central America; S. America, South America. The admixture model was used with independent allele frequencies and four populations. To avoid bias by the lineage classification, *SAG2* was excluded. Identical assignment of individuals was achieved in three independent runs (with 10^5 burn-in iterations, followed by 10^6 MCMC iterations). Excluding 121 isolates from South America and retaining samples from extreme geographical points of South America, i.e., Argentina ($n = 17$), Brazil ($n = 14$), and Colombia ($n = 14$), of comparable sample size to those from other continents produced highly consistent assignments: 100% for RW, 96% for SA2, 95% for WW, and 83% for SA1, with the overall mismatch rate of 3% (5 of 154).

*Lineage of four isolates was not determined (see *Materials and Methods*), reducing the total sample size to 271.

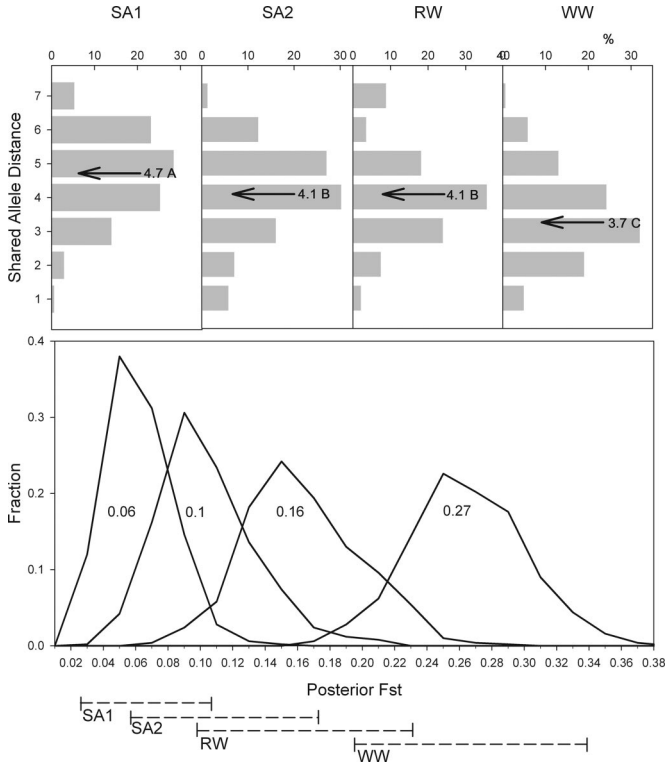

Fig. 6. Within-population divergence measured by the shared-allele distance (*Upper*) and posterior F_{ST} distributions measuring divergence of the populations identified by STRUCTURE from the "ancestral" population (*Lower*), with lines depicting the central 95% of the values of each posterior distribution. The two models identified nearly the same populations (only 13 of 275 individuals were clustered differently, comprising 4.7%). Mean F_{ST} values are shown in the center of each distribution ($n = 500$). Mean shared-allele distance of each population is shown with an arrow. Statistically distinct groups ($P < 0.001$), by using Wilcoxon two-sample test accommodating multiple comparisons, are indicated by letters.

The results, based on the most extensive geographical sampling of *T. gondii* to date, point to a South American origin of the species. Notably, only wild felids of relatively low abundance could serve as definite hosts to *T. gondii* in South America until the introduction of the domestic cat during the 16th century (22), suggesting that *T. gondii* was a far less ubiquitous parasite throughout most of its evolutionary history. The global structure entails at least two migration events from South America. The first migration, probably into Eurasia (Fig. 1), was probably a rare event, such as a stray migratory bird infected with *T. gondii*. Aided by abundant domestic cats, the organism evolved separately into the RW population. *T. gondii*'s widespread distribution, however, required means for long-distance migration. Long-distance dispersal of *T. gondii* could be achieved by infected migratory birds, but several features of *T. gondii* population structure are incompatible with patterns of bird migration. Bird migration between North America and Eurasia is very rare (23) and cannot solely account for the similar genotype composition in these regions. If rare bird migration generates extensive gene flow of *T. gondii*, why did the massive bird migration between North and South America (23) fail to do the same? Because bird migration has been ongoing for millions of years in specific routes, this failure is incompatible with the very recent spread of the multicontinent genotypes, the global distribution, and the small fraction of genotypes involved. Thus, dispersal of *T. gondii* by migratory birds probably played a secondary role. Alternatively, long-distance dispersal of *T. gondii* can be mediated by man. We propose that ships populated by rats, mice, and domestic cats provided unprecedented migration opportunity for *T. gondii*, especially during the transatlantic slave trade. Once introduced on board a ship inhabited by rodents and cats, explosive transmission of *T. gondii* would ensue, because a single infected cat will shed millions of oocysts into the ship's cramped space, resulting in repeated amplification cycles on board and creating a hyperendemic focus (24). When such ships arrived in port, they could easily introduce *T. gondii* to the port area and to other ships by unloading cargo infested with oocysts as well as with infected rats and mice. The global maritime trade beginning in the 16th century probably transported the RW population from Eurasia to other continents. The introduced genotypes became predominant where there was no resident *T. gondii* population, such as in North America and Southeast Asia, but could not predominate in areas, such as South America, already inhabited by *T. gondii*. During the 18th and 19th centuries, large quantities of agricultural produce, such as sugar cane, rice, tobacco, and cotton, were loaded on ships in South and Central America. Unlike the cargo loaded in Europe (e.g., textiles and guns) and Africa (slaves), agricultural goods must have contained soil and, possibly, soil contaminated with oocysts, sharply raising the probability that South and Central American isolates would colonize ships. Thus, one or few such introductions led rapidly to the global spread of the WW population. The dramatic increase in the volume of maritime trade correspondingly increased the prospects of the genotypes on board to establish themselves even in areas already inhabited by *T. gondii* (such as Europe). Modern ships, however, with their large size, sealed cargo containers, high sanitation, including rodent control and restrictions on cats, are far less hospitable for *T. gondii*; thus, its global spread has probably slowed considerably during the past century, despite the spectacular increase in volume and speed of transoceanic travel.

This explanation leads to the formulation of unique predictions, including that genotype composition and abundance near ports active during early transatlantic trade will differ markedly from that in regions distant from such ports. A closer examination of our data supports this prediction (Fig. 1), showing high WW proportion near major ports (Grenada) compared with inland regions (e.g., Austria

POPULATION BIOLOGY

and Quinidio of Colombia). Human prevalence in the eastern United States is twice that in the western United States, (25) in accordance with *T. gondii*'s colonization of North America through the main eastern ports. Independent studies will be needed for testing these predictions. The main strength of this explanation, however, is in resolving the main enigmatic features of *T. gondii* previously reported (see Introduction) and in being consistent with previous findings. Accordingly, the low genetic diversity in a ubiquitous parasite (6, 9, 17) is attributed to its long history as a relatively sparse parasite confined to South American felids. Our findings support a recent expansion, as described (10), although we attribute this expansion to the early maritime trade rather than to the adaptation of direct oral infection. The geographic isolation between South America and Eurasia led to the division into two ancestral allele lineages previously reported but not explained (6, 8, 15). The lower frequency of recombinants in North American and European isolates (5, 7, 13) compared with South America (11, 14) is explained by the shorter time for accumulation of recombinants between (the relatively uncommon) distinct genotypes following colonization in a species with a high selfing rate. Our study demonstrates how geographical analysis of population structure illuminates seemingly unrelated problems and how important human influence is, even in shaping the genetic structure of a zoonotic disease agent.

Materials and Methods

Isolates. To avoid confounding geographic structure with possible association between host species and *T. gondii* genotype, a single host species was used throughout. Free-range ("backyard") chickens were used as sentinels for parasite isolation because of their worldwide distribution and their efficiency in detecting *T. gondii* oocysts in the ground because of their foraging habits. To obtain a representative sample of isolates within a location, only independent isolates, i.e., individual chickens from farms or households at least 200 m apart in each location were included in the analysis. Parasite isolation was performed in a single lab (J.P.D.) following the same procedures (26). Briefly, tissues of serologically positive chickens consisting of brain, heart, and/or breast muscle were individually inoculated into out-bred Swiss–Webster mice (Taconic Farms). Tissues from serologically negative chickens were pooled and fed to *T. gondii*-free cats (26). Feces of cats were examined for shedding of oocysts 3–14 days after ingesting chicken tissues. Oocysts obtained from cat feces were bioassayed in mice, and the brains of all mice were examined for tissue cysts. Tissues of mice infected directly with parasites from chickens or indirectly with oocysts shed by cats that were infected with parasites from chickens were used for parasite DNA extraction to avoid multiple parasite passages. DNA was extracted as described (6). Isolates were genotyped at seven loci (Table 2). Lineage determination was performed by using the nested PCR assay on *SAG2* (17). Genotyping at the other six loci was performed as described (11, 16). Five of the seven loci were genetically mapped onto different linkage groups; loci M95 and M102 were 21 cM apart, whereas M6 and M163 were 65 cM apart (27) (Table 2). Complete genotyping results were obtained for 238 isolates, whereas one or more loci were missing for the remaining 37 isolates, even after repeated PCR amplification attempts. Low DNA abundance in the extract is thought to explain the failures, but null alleles (i.e., mutations in the primer regions) cannot be excluded. No multigenotype infections were detected.

Data Analysis. Genetic diversity in each population was measured by per-locus expected heterozygosity, also known as gene diversity (28) and allele richness (29). For STR loci, the variance in allele size was also computed. Because lineage was determined based on the restriction fragment length polymorphism in *SAG2* (17), it was excluded from analyses aimed at evaluating concordance of lineage with genome-wide variation. A phylogenetic network of STR haplotypes was derived by using the median-joining algorithm (18) ($\varepsilon = 0$) after processing the data with the reduced-median method (19) as implemented by NETWORK 4.1 (Flexus Engineering). The network included 238 isolates that were genotyped across all seven loci, but STR loci were used and weighted inversely to their variance (M33, 9; M6, M48, and M102, 4; and M163, 3). Locus M95 was excluded because it is not a STR (16). The relationship between haplotypes incorporated variation in the number of repeats between STR alleles. The neighbor-joining tree, based on shared-allele distance (defined for all pairs of isolates as the number of loci with different alleles), was drawn by using the program MEGA 3.0 (30). The sequential Bonferroni procedure (31) was used to detect a single significant test when multiple tests were used. Calculations not available by STRUCTURE, and NETWORK were carried out by using programs written by T.L. in SAS language (32).

We thank S. K. Shen, O. C. H. Kwok, M. C. B. Vianna, D. E. Hill, S. M. Gennari, A. M. A. Ragozo, S. M. Nishi, D. S. Silva, D. Seipel da Silva, L. M. G. Bahia-Oliveira, M. Hilali, A. El-Ghaysh, C. Sreekumar, M. F. Davis, T. Y. Morishita, I. T. Navarro, R. L. Freire, L. B. Prudencio, M. C. Venturini, L. Venturini, M. Piscopo, M. Levy, E. S. Morales, H. Salant, D. Spira, J. Hamburger, S. Karhemere, A, Diabaté, K. R. Dabiré, M. I. Bhaiyat, C. de Allie, C. N. L. Macpherson, R. N. Sharma, R. Edelhofer, B. Lopez, M. Alveraz, C. Mendoza, J. E. Gomez-Marin, A. Bedoya, F. Lora, R. P. V. J. Rajapakse, D. K. Ekanayake, A. Lenhart, C. E. Castillo, L. Alvarez, M. B. Labruna, L. M. A. Camargo, S. Sousa, N. Canada, C. S. Meireles, J. M. Correia da Costa, M. L. Dardé, P. Thulliez, M. Raman, and D. P. Bhalerao for help with obtaining samples and José M. C. Ribeiro, Randy Dejong, Robert Gwadz, Lou Miller, Su Xinzhuan [Laboratory of Malaria and Vector Research, National Institutes of Health (NIH)], Jeff Jones [Centers for Disease Control and Prevention (CDC)], and, especially, Ben Rosenthal (U.S. Department of Agriculture) and three anonymous reviewers for critical comments and discussions on earlier versions of the manuscript. This research was supported, in part, by the Food Safety Initiative (CDC) and by the Intramural Research Program of the NIH, National Institute of Allergy and Infectious Diseases.

1. Lindsay, D. S., Blagburn, B. L. & Dubey, J. P. (2002) *Vet. Parasitol.* **103,** 309–313.
2. Dubey, J. P. (1998) *J. Parasitol.* **84,** 862–865.
3. Dubey, J. P. & Beattie, C. P. (1988) *Toxoplasmosis of Animals and Man* (CRC, Boca Raton, FL).
4. Dardé, M. L., Bouteille, B. & Pestre-Alexandre, M. (1992) *J. Parasitol.* **78,** 786–794.
5. Howe, D. K. & Sibley, L. D. (1995) *J. Infect. Dis.* **172,** 1561–1566.
6. Lehmann, T., Blackston, C. R., Parmley, S. F., Remington, J. S. & Dubey, J. P. (2000) *J. Parasitol.* **86,** 960–971.
7. Sibley, L. D. & Boothroyd, J. C. (1992) *Nature* **359,** 82–85.
8. Ajzenberg, D., Banuls, A. L., Tibayrenc, M. & Dardé, M. L. (2002) *Int. J. Parasitol.* **32,** 27–38.
9. Dardé, M. L. (1996) *Curr. Top. Microbiol. Immunol.* **219,** 27–41.
10. Su, C., Evans, D., Cole, R. H., Kissinger, J. C., Ajioka, J. W. & Sibley, L. D. (2003) *Science* **299,** 414–416.
11. Lehmann, T., Graham, D. H., Dahl, E. R., Bahia-Oliveira, L. M., Gennari, S. M. & Dubey, J. P. (2004) *Infect. Genet. Evol.* **4,** 107–114.
12. Tibayrenc, M. & Ayala, F. (2002) *Trends Parasitol.* **18,** 405.
13. Tibayrenc, M., Kjellberg, F., Arnaud, J., Oury, B., Breniere, S. F., Dardé, M. L. & Ayala, F. J. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 5129–5133.
14. Ajzenberg, D., Banuls, A. L., Su, C., Dumetre, A., Demar, M., Carmé, B. & Dardé, M. L. (2004) *Int. J. Parasitol.* **34,** 1185–1196.
15. Grigg, M. E., Bonnefoy, S., Hehl, A. B., Suzuki, Y. & Boothroyd, J. C. (2001) *Science* **294,** 161–165.
16. Blackston, C. R., Dubey, J. P., Dotson, E., Su, C., Thulliez, P., Sibley, D. & Lehmann, T. (2001) *J. Parasitol.* **87,** 1472–1475.
17. Howe, D. K., Honoré, S., Derouin, F. & Sibley, L. D. (1997) *J. Clin. Microbiol.* **35,** 1411–1414.
18. Bandelt, H. J., Forster, P. & Rohl, A. (1999) *Mol. Biol. Evol.* **16,** 37–48.
19. Bandelt, H. J., Forster, P., Sykes, B. C. & Richards, M. B. (1995) *Genetics* **141,** 743–753.
20. Pritchard, J. K., Stephens, M. & Donnelly, P. (2000) *Genetics* **155,** 945–959.
21. Falush, D., Stephens, M. & Pritchard, J. K. (2003) *Genetics* **164,** 1567–1587.
22. Todd, N. B. (1977) *Sci. Am.* **237(5),** 100–107.
23. Elphick, J. (1995) *The Atlas of Bird Migration* (Random House, New York).
24. Lehmann, T., Graham, D. H., Dahl, E., Sreekumar, C., Launer, F., Corn, J. L., Gamble, H. R. & Dubey, J. P. (2003) *Infect. Genet. Evol.* **3,** 135–141.
25. Jones, J. L., Kruszon-Moran, D., Wilson, M., McQuillan, G., Navin, T. & McAuley, J. B. (2001) *Am. J. Epidemiol.* **154,** 357–365.
26. Dubey, J. P., Graham, D. H., Blackston, C. R., Lehmann, T., Gennari, S. M., Ragozo, A. M., Nishi, S. M., Shen, S. K., Kwok, O. C., Hill, D. E. & Thulliez, P. (2002) *Int. J. Parasitol.* **32,** 99–105.
27. Khan, A., Taylor, S., Su, C., Mackey, A. J., Boyle, J., Cole, R., Glover, D., Tang, K., Paulsen, I. T., Berriman, M., *et al.* (2005) *Nucleic Acids Res.* **33,** 2980–2992.
28. Nei, M. (1987) *Molecular Evolutionary Genetics.* (Columbia Univ. Press, New York), p. 176–186.
29. Petit, R. J., Mousadik, A. & Pons, O. (1998) *Conserv. Biol.* **12,** 844–855.
30. Kumar, S., Tamura, K. & Nei, M. (2004) *Brief. Bioinform.* **5,** 150–163.
31. Holm, S. (1979) *Scand. J. Stat.* **6,** 65–70.
32. SAS Institute (2002) SAS *for Windows Version 9.0* (SAS Institute, Cary, NC).