A genetically diverse but distinct North American population of *Sarcocystis neurona* includes an overrepresented clone described by 12 microsatellite alleles

Ingrid M. Asmundsson, J.P. Dubey, Benjamin M. Rosenthal *

*Animal Parasitic Disease Laboratory, Animal and Natural Resources Institute, Agricultural Research Service, United States Department of Agriculture, Beltsville, MD 20705, USA*

Received 18 October 2005; received in revised form 30 December 2005; accepted 2 January 2006

Available online 20 February 2006

Abstract

The population genetics and systematics of most coccidians remain poorly defined despite their impact on human and veterinary health. Non-recombinant parasite clones characterized by distinct transmission and pathogenesis traits persist in the coccidian *Toxoplasma gondii* despite opportunities for sexual recombination. In order to determine whether this may be generally true for tissue-cyst forming coccidia, and to address evolutionary and taxonomic problems within the genus *Sarcocystis*, we characterized polymorphic microsatellite markers in *Sarcocystis neurona*, the major causative agent of equine protozoal myeloencephalitis (EPM). Bayesian statistical modeling, phylogenetic reconstruction based on genotypic chord distances, and analyses of linkage disequilibrium were employed to examine the population structure within *S. neurona* and closely related *Sarcocystis falcatula* isolates from North and South America. North American *S. neurona* were clearly differentiated from those of *S. falcatula* and also from isolates of *S. falcatula*. Although *S. neurona* is characterized by substantial allelic and genotypic diversity typical of interbreeding populations, one genotype occurs with significantly excessive frequency; thus, some degree of asexual propagation of *S. neurona* clones may naturally occur. Finally, *S. neurona* isolated from disparate North American localities and diverse hosts (opossums, a Southern sea otter, and horses) comprise a single genetic population. Isolates associated with clinical neurological disease bear no obvious distinction as measured by these presumably neutral genetic markers.

Published by Elsevier B.V.

Keywords: *Sarcocystis neurona*; *Sarcocystis falcatula*; Microsatellites; Population structure; Population genetics; *Toxoplasma gondii*; Coccidia; *Didelphis*; Equine protozoal myeloencephalitis

1. Introduction

The coccidia comprise a diverse group of parasitic apicomplexan protists. Among these include the large group of species belonging to the genus *Sarcocystis*. Accurately enumerating, differentiating, and understanding the interrelationships among these species requires not only stable diagnostic attributes, but also means to verify the nature and limits of gene flow within (and perhaps between) such taxa. The practical interest in developing a stable and accurate nomenclature for these veterinary (and potentially zoonotic) parasites may be undermined by the formation of genetically stable and phenotypically distinct strains. If genes only infrequently flow among such potentially interfertile strains, delineating microbial species can become difficult because such species will lack overall genetic cohesion.

*Sarcocystis neurona* and *Sarcocystis falcatula* are morphologically similar coccidian species (Apicomplexa: Sarcocystidae) excreted in feces of opossum definitive hosts belonging to the genus *Didelphis*. These environmentally resistant sporocysts are infective to a range of vertebrate hosts and differentiating among them has traditionally been accomplished by bioassays (Marsh et al., 1997a,b; Dubey and Lindsay, 1998). Although sequence differences at a handful of genetic loci also discriminate parasites infective to avian intermediates (presumed to represent *S. falcatula*) from those that induce neurological disease in interferon-gamma deficient...
mice (presumed to represent *S. neurona*) (Tanhauser et al., 1999; Rosenthal et al., 2001), we presently lack good information about the genetic diversity of parasites manifesting either of these phenotypes. We know neither the extent of gene flow occurring among them, nor their relationship to other morphologically and epidemiologically similar taxa (Dubey and Lindsay, 1999; Dubey et al., 2001).

*S. neurona*, the major causative agent of equine protozoal myeloencephalitis (EPM), was first isolated and described from the neural tissue of a horse in 1990 (Dubey et al., 1991), although this serious disease, was recognized earlier by Rooney et al. (1970). EPM occurs in the Western Hemisphere where horses are exposed to sporocysts shed by new world opossums, in the genus *Didelphis*. Although clinical EPM manifests itself in only 0.5-1% of horses (MacKay et al., 2000), serological surveys identify exposure in as many as 54% of horses (Vardeleon et al., 2001) and 66.6% of ponies and other equids (Tillotson et al., 1999). The clinical signs are caused by schizogenic multiplication of *S. neurona* in neural tissues of a variety of animals including cats, raccoons, sea otters and other marine and land mammals. In true intermediate hosts (e.g., raccoons, and cats) the parasites released from schizonts can encyst, particularly in the muscles, forming sarcocysts. Opoosums become infected when they ingest sarcocysts (Dubey et al., 2001c).

Although *S. falcata* Stiles, 1893 was named as a parasite maintained through transmission among birds and opossums over 100 years ago, recent molecular evidence indicates that isolates designated as *S. falcata* may be more accurately understood as an assemblage of related species (Dubey et al., 2001b). Because the type specimen for *S. falcata* is no longer available (Box et al., 1984) we cannot determine to which of any currently recognized genotype it may have corresponded. For the purpose of this work we will provisionally refer to all such avian-infective isolates as *S. falcata*.

Highly polymorphic microsatellite markers have been used to gain important insights into the demographic history, population structure, and epidemiology of various parasitic apicomplexans (Ajzenberg et al., 2002, 2004; Leclerc et al., 2004; Machado et al., 2004). Among the coccidia, especially rewarding insights have been derived by studying the extent and pattern of microsatellite variation in *Toxoplasma gondii*, the agent of human and animal toxoplasmosis (Blackston et al., 2001; Ajzenberg et al., 2002). A close relative of *Sarcocystis* spp., *T. gondii* infects an exceptionally broad range of vertebrates and as many as 85% of people in some locales (Sibley and Boothroyd, 1992). Genetic analyses of *T. gondii* have revealed an unusual population structure comprising three distinct, but highly similar genotypes, which account for a preponderance of infections in temperate regions of the Northern Hemisphere. Of these, only one genotype is highly virulent in mice, and may confer especially great risk to human reproductive health (Sibley and Boothroyd, 1992; Ajzenberg et al., 2002; Ferreira et al., 2004). These three pandemic strains persist despite well documented sexual reproduction in cats, the definitive host, which excrete myriad recombinant genotypes contracted by diverse wildlife hosts, especially evident in studies of tropical French Guyana and Brazil (Ajzenberg et al., 2004; Lehmann et al., 2004; Ferreira et al., 2006). Asexual dissemination, by circumventing the definitive host, may have promoted widespread dispersal of particularly well-adapted clones (Su et al., 2003). Understanding whether these characteristics are determined by features specific to *T. gondii* history and epidemiology, or whether they instead typify a larger group of evolutionarily related parasites, will require that the tools of modern population genetics be applied to other members of the coccidia.

Here we analyzed 12 new microsatellite markers (Asmundson and Rosenthal, 2006) as a basis to better understand the extent and pattern of neutral genetic variation among haploid specimens of *S. neurona* and *S. falcata*. In particular, we assayed a large group of North American *S. neurona* derived from a broad geographic range and derived from various hosts, and compared these to a smaller number South American isolates of *S. neurona* as well as to a small number of *S. falcata* of diverse geographic origins. By estimating the extent of inbreeding and linkage disequilibrium among pre-defined isolate assemblages and by assigning individuals to populations-based only on their genotypes using Bayesian probabilistic methods, we established that North American isolates of *S. neurona* are especially closely interrelated, verified that the neurological disease in a Southern sea otter was caused by infection with *S. neurona* typical of that occurring in terrestrial North American vertebrates, differentiated North from South American *S. neurona* isolates, and found preliminary evidence that a certain multilocus genotype of *S. neurona* may occur with excessive frequency.

2. Materials and methods

2.1. Parasite isolates and microsatellite markers

We genotyped and analyzed 34 previously described *Sarcocystis* isolates (Table 1). The 25 *S. neurona* isolates collected in North America (Panama and the USA) represent diverse geographical origins and hosts (opossum, horse and sea otter). Four of these isolates were cultured from horses with clinical EPM. With the exception of the Panamanian isolate, the isolates from horse and sea otter are assumed to have been excreted by *Didelphis virginiana*. The Panamanian isolate is believed to have originated from *Didelphis marsupialis*, the sole species of *Didelphis* endemic to Panama. Two *S. neurona* isolates from South America were collected from *Didelphis albiventer* in the vicinity of São Paulo, Brazil. Seven *S. falcata* isolates originated from opossums, two from North America (*D. virginiana* from the US) and five from South America (*D. albiventer* and *D. marsupialis* in Argentina and Brazil). DNA was extracted from sporocysts, or from cell cultures, using Qiagen DNeasy columns according to manufacturer protocols. Microsatellite isolation, characterization and genotyping were performed as previously described (Asmundson and Rosenthal, 2006).
Table 1

Parasite isolates

<table>
<thead>
<tr>
<th>Isolate</th>
<th>Host from which isolated</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>North American <em>S. neurona</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SN19-OP⁹</td>
<td>Opossum, <em>D. virginiana</em></td>
<td>Mississippi, USA</td>
</tr>
<tr>
<td>SN23-OP</td>
<td>Opossum, <em>D. virginiana</em></td>
<td>Mississippi, USA</td>
</tr>
<tr>
<td>SN⁷</td>
<td>Horse</td>
<td>Oregon, USA</td>
</tr>
<tr>
<td>SN25-OP⁹</td>
<td>Opossum, <em>D. virginiana</em></td>
<td>Mississippi, USA</td>
</tr>
<tr>
<td>Sarco57,</td>
<td>Opossum, <em>D. virginiana</em></td>
<td>Cornell University</td>
</tr>
<tr>
<td>batch 98-8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SN20-OP⁹ (not characterized but presumed to be <em>S. neurona</em>)</td>
<td>Opossum, <em>D. virginiana</em></td>
<td>Mississippi, USA</td>
</tr>
<tr>
<td>SN18-OP⁴</td>
<td>Opossum, <em>D. virginiana</em></td>
<td>Mississippi, USA</td>
</tr>
<tr>
<td>SN16-OP⁴</td>
<td>Opossum, <em>D. virginiana</em></td>
<td>Mississippi, USA</td>
</tr>
<tr>
<td>SN26-OP⁴</td>
<td>Opossum, <em>D. virginiana</em></td>
<td>Mississippi, USA</td>
</tr>
<tr>
<td>SN-OT1</td>
<td>Sea otter, <em>Enhydra lutris nereis</em></td>
<td>California, USA</td>
</tr>
<tr>
<td>SN⁶</td>
<td>Horse</td>
<td>Oregon, USA</td>
</tr>
<tr>
<td>SN12-OP⁵-⁶</td>
<td>Opossum, <em>D. virginiana</em></td>
<td>Maryland, USA</td>
</tr>
<tr>
<td>SN21-OP⁶</td>
<td>Opossum, <em>D. virginiana</em></td>
<td>Mississippi, USA</td>
</tr>
<tr>
<td>SN34-OP⁶</td>
<td>Opossum, <em>D. virginiana</em></td>
<td>Mississippi, USA</td>
</tr>
<tr>
<td>SN33-OP⁶</td>
<td>Opossum, <em>D. virginiana</em></td>
<td>Mississippi, USA</td>
</tr>
<tr>
<td>SN24-OP⁶</td>
<td>Opossum, <em>D. virginiana</em></td>
<td>Mississippi, USA</td>
</tr>
<tr>
<td>SN29-OP⁶</td>
<td>Opossum, <em>D. virginiana</em></td>
<td>Mississippi, USA</td>
</tr>
<tr>
<td>IMF-005</td>
<td>Unknown</td>
<td>Unknown</td>
</tr>
<tr>
<td>SN15-OP⁶</td>
<td>Opossum, <em>D. virginiana</em></td>
<td>Virginia, USA</td>
</tr>
<tr>
<td>SN32-OP⁶</td>
<td>Opossum, <em>D. virginiana</em></td>
<td>Mississippi, USA</td>
</tr>
<tr>
<td>SN8-OP⁵-⁶</td>
<td>Opossum, <em>D. virginiana</em></td>
<td>Pennsylvania, USA</td>
</tr>
<tr>
<td>SN30-OP⁵</td>
<td>Opossum, <em>D. virginiana</em></td>
<td>Mississippi, USA</td>
</tr>
<tr>
<td>SN27-OP⁵</td>
<td>Opossum, <em>D. virginiana</em></td>
<td>Mississippi, USA</td>
</tr>
<tr>
<td>SN³</td>
<td>Horse</td>
<td>North-western Panama</td>
</tr>
<tr>
<td>SN⁴</td>
<td>Horse</td>
<td>California, USA</td>
</tr>
<tr>
<td>South American <em>S. neurona</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Opossum ³</td>
<td>Opossum, <em>D. albiventris</em></td>
<td>Areas around São Paulo, Brazil</td>
</tr>
<tr>
<td>Opossum ⁷</td>
<td>Opossum, <em>D. albiventris</em></td>
<td>Areas around São Paulo, Brazil</td>
</tr>
<tr>
<td>North American <em>S. falcatula</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Opossum 808³</td>
<td>Opossum, <em>D. virginiana</em></td>
<td>Louisiana, USA</td>
</tr>
<tr>
<td>Sarco57, batch 99-4⁴</td>
<td>Cowbird/opossum (?)</td>
<td>Cornell University</td>
</tr>
<tr>
<td>South American <em>S. falcatula</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bird 200⁴</td>
<td>Opossum, <em>D. marsupialis</em></td>
<td>São Paulo, Brazil</td>
</tr>
<tr>
<td>Opossum 2⁵</td>
<td>Opossum, <em>D. albiventris</em></td>
<td>Buenos Aires Province, Argentina</td>
</tr>
<tr>
<td>Opossum 14⁴</td>
<td>Opossum, <em>D. marsupialis</em></td>
<td>São Paulo, Brazil</td>
</tr>
<tr>
<td>Bird 100⁴</td>
<td>Opossum, <em>D. albiventris</em></td>
<td>Jaboticabal, Brazil</td>
</tr>
<tr>
<td>9087⁴</td>
<td>Opossum, <em>Didelphis sp.</em></td>
<td>Jaboticabal, Brazil</td>
</tr>
</tbody>
</table>

a Dubey et al. (2001c).
b Dubey et al. (2001d).
c Lindsay et al. (2000).
d Dubey et al. (1999b).
e Dubey (2000).
f Granstrom et al. (1992).
g Davis et al. (1991).
h Dubey et al. (2001a).
i Not characterized, presumed to be *S. falcata*.
j Dubey et al. (2001b).
k Dubey et al. (1999a).
l Dubey et al. (2000).

2.2. Clustering analysis

The model-based clustering method Structure 2.1 (Pritchard et al., 2000) was used to explore evidence for population structure using only the multilocus genotype data (but not a priori taxonomic or geographic information) as its basis. Having verified the essential robustness of our central conclusions to various model assumptions, we elected the ‘no admixture’ ancestral model and one million MCMC replications and a burn-in period of 100,000 generations. Allele frequencies among populations were assumed to be independent. The probability of the data, assuming one to ten populations (K), was estimated in five independent replicate analyses, and we printed credible regions and Q-hat, which described how confidently individual specimens could be assigned to one or more population subdivisions. These analyses were performed on all isolates and also on a subset comprising only the North American *S. neurona*.

2.3. Phylogenetic analysis

Neighbor joining trees were reconstructed from the genetic distances among the multilocus genotypes of individual parasite isolates using Population 1.2.1 of Olivier Langella (http://www.pge.cnrs-gif.fr/bioinfo/populations/index.php?page=en) based on ASD (Goldstein et al., 1995; Slatkin, 1995). As applied to our pairs of haploid individuals, ASD represents the squared difference in allele size, summed across loci. This analysis was repeated for 1000 bootstrap replicates in which loci were sampled with replacement. Phylogenies were also reconstructed using the chord distance estimator ASD (Goldstein et al., 1995; Slatkin, 1995). As applied to our pairs of haploid individuals, ASD represents the squared difference in allele size, summed across loci. This analysis was repeated for 1000 bootstrap replicates in which loci were sampled with replacement. Phylogenies were also reconstructed using the chord distance estimator Dc of Cavalli-Sforza and Edwards (1967), in this application a simple function of the proportion of shared alleles between individual pairs.

2.4. Genotypic diversity

Genotypic diversity was calculated as the number of unique multilocus genotypes divided by the total number of individuals genotyped. Diversity was also expressed as the probability that any two specimens, drawn at random, would differ genotypically.

2.5. Linkage disequilibrium tests

The extent of pairwise linkage disequilibrium was estimated using Genetic Data Analysis (GDA) and its significance assessed with Fisher’s exact tests (Lewis and Zaykin, 2001) using 3200 shufflings after discarding missing data. These tests were performed on all 34 isolates, and also limited to the following isolate subsets: *S. neurona* isolates, *S. falcata* isolates, and North American *S. neurona*.

To assess the Index of association (Iₐ) among alleles in genotypes, and determine the statistical significance of any departure from its null expectation of 0 in panmictic populations, we used MultiLocus v. 1.3 (Agapow and Burt, 2001) invoking the LD option and comparing the observed data
to 10,000 datasets into which alleles were randomized into genotypes. These analyses were performed on all individuals, and on all unique genotypes, in North American isolates of *S. neurona* and in the dataset as a whole to assess the extent of allelic correlation among individuals and genotypes.

### 3. Results

#### 3.1. Genotypic diversity

Genotypic diversity over all 34 isolates was great (0.94), with 32 distinct genotypes. Among these 34 isolates, 31 distinct multilocus genotypes occurred at a frequency of 0.0294. One genotype characterized three North American *S. neurona* isolates, two at all 12 loci and a third at 11 loci with the twelfth allele missing, and occurred at a frequency of 0.0882.

#### 3.2. Clustering analysis

Because the phylogenetic significance of taxonomic and geographic designations were unclear at the outset, we explored the utility of a clustering method (Structure 2.1) which assigns individuals, with stated probabilities, to any of *K* subpopulations. By iterating this analysis for increasing numbers of populations, and evaluating the resulting populations within a Bayesian likelihood framework, we identified the range of subpopulations into which our total sample could be credibly subdivided, as well as the confidence with which any individual could be assigned to any particular population subdivision. The approximate likelihood of the data “should be regarded as rough guides to which models are consistent with the data, rather than accurate estimates of the posterior probabilities” (Pritchard et al., 2000). Most individuals could be unambiguously assigned only when they were assumed to have derived from fewer than three populations (Fig. 1). The probability of the data, averaged over five independent replicate analyses for each *K*, assuming three groups was approximately 20 and 10 times greater than the probability assuming two or four groups, respectively, and nearly 138 and 21 times greater than the probability of one and five groups, respectively (Fig. 1). Considering any of these five possibilities, the posterior probability of three groups was nearly 100%.

Designating more populations caused most individuals to be variously and uncertainly placed into several groups, thereby...
breaking down the genetic cohesion of any group. Although the probability of the data was maximized assuming \( K = 6 \), no individuals could be placed into any one of these six groups with statistical confidence. Moreover, assuming six or more populations incurred large and increasing variance estimates. Previous experience suggests that, as occurs here for \( K \geq 5 \), one should “be skeptical about population structure inferred on the basis of small differences in \( Pr(K) \) if (1) there is no clear biological interpretation for the assignments, and (2) the assignments are roughly symmetric to all populations and no individuals are strongly assigned” (Structure documentation). In sum, the probability of these data are discernibly improved as the assumed number of distinct populations is increased from one to three, and then moderately or not at all thereafter.

As the number of assumed populations was increased, the composition of one population (exclusively comprising North American isolates of \( S. neurona \)) remained unchanged. Despite the fact that these represented 25 of the 34 individuals assayed, these were consistently and exclusively assigned to a single population subdivision. This was true despite the fact that, among these North American \( S. neurona \) isolates, 10 of 12 loci (83%) were polymorphic with approximately four alleles per locus. Importantly, this genetic group included parasites isolated from not only opossums, but also from clinical equine cases. Finally, this group also included a specimen isolated from a Southern sea otter (\( Enhydra lutia \)) that succumbed to encephalitis (Lindsay et al., 2000). We thus unequivocally establish that clinically similar neurological disease in terrestrial and marine mammals is caused by the same
etiological agent and infer that viable oocysts of terrestrial origin apparently contaminate certain coastal environments.

How did increasing the number of assumed population subdivisions affect the distribution of the remaining specimens? These were all assigned to a second group when two populations were assumed. When three populations were assumed, South American \textit{S. falcata} were generally assigned to a second group, whereas South American \textit{S. neurona} and North American \textit{S. falcata} were generally assigned to a third. Unlike the North American \textit{S. neurona}, assigning these remaining isolates could not be accomplished with statistical certainty, a phenomenon whose magnitude increased as \textit{K} was increased further.

In order to further explore any evidence for genetic structure among the 25 specimens of North American \textit{S. neurona}, we constrained a second set of cluster analyses to this sub-sample. Doing so resulted in all isolates being placed, with intermediate probabilities, into each of \textit{K} populations as \textit{K} was varied from two to six. Thus, the available data provided no suggestion of further genetic subdivision or strain differentiation among North American \textit{S. neurona} isolates, irrespective of their geographic or host origins.

### 3.3. Phylogenetic analysis

In order to visualize the extent of genotypic differentiation among individual parasites and evaluate the strength of any phylogenetic signal intrinsic to microsatellite variation, we reconstructed Neighbor Joining trees from bootstrap replicates of inter-individual distances. The squared difference in allele size among pairs of individuals, summed across loci, was minimal among North American isolates of \textit{S. neurona}, but greater in the overall specimen collection (Fig. 2). An especially extensive distance, defining the single node receiving extensive support when the data were bootstrapped across loci, differentiated the two North American isolates of \textit{S. falcata}. Further phylogenetic resolution was lacking in either these analyses, or those based on the proportion of shared alleles among pairs of loci (Fig. 2). A poorly supported monophyly was indicated among pan-American \textit{S. neurona} only if three South American \textit{S. falcata} isolates were also included. It is unclear whether additional phylogenetic resolution might be achieved from more sophisticated models of genotype evolution, or from characters less prone than microsatellites to homoplasy. Notwithstanding the inherent difficulty in interpreting the absence of phylogenetic resolution, the integrity of pan-American \textit{S. neurona} as a monophyletic taxon seems undermined by these observations.

### 3.4. Linkage disequilibrium tests

Significant linkage disequilibrium (\(p < 0.01\)) was detected among 39 of 66 locus pairs when all 34 isolates were considered together (11 of these remain significant even when accepting a 1% risk of type I error among any of 66 multiple comparisons, individual \(p < 0.00015\)). When this dataset was reduced by randomly removing nine North American \textit{S. neurona} samples (\(n = 25\)), statistical power remained adequate to reject the null hypothesis (in no fewer than 25 of 66 locus pairs, \(p < 0.01\) or in no fewer than six \(p < 0.00015\)). When only pan-American \textit{S. neurona} isolates were considered, linkage disequilibrium was still detected among almost one fifth (13/66) of the locus pairs. By contrast, no significant linkage disequilibrium was detected among the 25 North American isolates of \textit{S. neurona}. Taken together, these results suggest that alleles do not flow freely among the sample overall, or among the pan-American sample of \textit{S. neurona}.

### 3.5. Index of association

The index of association was used to evaluate the extent to which alleles were correlated among individuals or among unique genotypes. This null hypothesis is strongly rejected when all individuals (\(I_A = 1.67\)) or all unique genotypes (\(I_A = 1.56\)) are considered (\(p < 0.0001\) in both cases). Among all North American isolates of \textit{S. neurona}, \(I_A = 0.31, p < 0.025\); restricting this test to the unique genotypes in this sub-sample slightly reduced both the estimate and statistical significance of this departure \(I_A = 0.23, p < 0.073\).

### 4. Discussion

Describing a pathogen’s genetic structure requires an adequate sampling of sufficiently polymorphic genes across a representative sampling across that taxon’s ecogeographic range (Gauthier and Tibayrenc, 2005). We have successfully applied an informative array of variable genetic markers to the widest available sampling of parasites previously designated as either \textit{S. neurona} and \textit{S. falcata}. Although these data permit new and important biological inferences to be drawn, especially concerning the genetic structure of \textit{S. neurona} in North America from which we have most broadly sampled, additional sampling will be required in order to establish a complete understanding of inter-regional and interspecific relationships.

Our clustering and phylogenetic analyses each resolved partitions among the total sample of 34 isolates, most designated as either \textit{S. neurona} or \textit{S. falcata} according to their pathogenicity in interferon-gamma deficient mice or birds (\textit{Melopsittacus undulatus}). These analyses suggested a natural partitioning into three populations: (1) North American \textit{S. neurona}, (2) South American \textit{S. neurona} and North American \textit{S. falcata}, and (3) South American \textit{S. falcata}. Although additional sampling would be likely to further refine the genetic cohesiveness or diversity of these groups, even the available data raise doubts as to whether specimens identified as \textit{S. neurona} from throughout the Americas comprise a monophyletic group that excludes \textit{S. falcata}.

We then used more traditional approaches to estimate the significance of population subdivision by provisionally accepting this “three-population” configuration. Extensive linkage disequilibrium characterizes the sample as a whole, but not the constituent populations inferred above. Moreover, about half of the overall genotypic variance can be attributed to differences among these three sub-samples. Similarly, highly
significant departures from panmixia resulted from our analyses of the index of association among individuals and among genotypes, overall. Departures of considerably less magnitude and statistical significance resulted from the application of this test to the North American S. neurona sample, especially after removing redundant genotypes. Taken together, these results suggest that alleles do not flow freely among the sample overall, or among the pan-American sample of S. neurona, but that the North American isolates of S. neurona resemble what would be expected from a panmictic, interbreeding population.

North American S. neurona comprise a genetically variable yet cohesive assemblage. All but two of 12 microsatellite loci were polymorphic among North American S. neurona (characterized by 2–10 alleles). In these specimens, though they vastly outnumber others in our study, only half of all observed alleles occur. The percentage of a locus’ alleles observed in North American S. neurona ranges from 0.17 to 0.77 (mean 0.50; variance 0.21). Although we lack direct estimates of the rate of microsatellite diversification in Sarcocystis, these data suggest that North American S. neurona are characterized by appreciable yet discernibly reduced microsatellite variation when viewed from a pan-American perspective.

With notable exceptions (discussed below), most isolates of S. neurona are characterized by a unique multilocus genotype. This resembles T. gondii, in which certain populations are characterized by clonal propagation of particularly well-adapted genotypes. A reported 87% of 83 European and American T. gondii isolates bore a unique microsatellite genotype (Ajzenberg et al., 2002), as did 69% of 13 French isolates and all of nine from French Guyana (Ajzenberg et al., 2004). We observe among North American specimens the gamut of presumably recombinant multilocus genotypes. If this generally characterizes S. neurona populations (as future sampling should either confirm or refute), it would less resemble the population structure first proposed for temperate adapted genotypes. A reported 87% of 83 European and South American isolates are infective to South American opossums. This resembles T. gondii (formerly all considered S. neurona spp. (Lemos and Cerqueira, 2002) established themselves in North America after formation of the Panamanian isthmus, approximately three million years ago (Ma). Didelphid fossils in Florida date to the Irvingtonian age, less than 1.8 Ma (Marshall et al., 1982). Whereas D. virginiana occupies a broad and expanding range from Costa Rica north into Canada, including major portions of Mexico and the United States, D. marsupialis is known north of Central America only in southeastern Mexico. Correspondingly little overlap characterizes the South American distributions of D. marsupialis with the distribution of D. albiventris, Didelphis imperfecta and Didelphis pernigra (formerly all considered D. albiventris). The former are prevalent in interior Amazonia and throughout northern and western regions of the continent, whereas the latter occupy a more southeasterly distribution—with some overlap in Andean Peru, Ecuador, Columbia and Venezuela. These largely disjunct host distributions may limit migration and gene flow between North and South America. At present, we do not know whether the hosts endemic to particular regions vary in their susceptibility to such infections, or whether parasite fecundity varies among hosts. Knowing if North American S. neurona isolates are infective to South American opossum species would be helpful in understanding the degree of potential gene flow between the continents.

Notably little differentiation characterizes S. neurona across the vast North American continent. An evolutionarily recent and rapid expansion may explain the lack of differentiation among isolates collected several thousand miles apart and from diverse host species (opossums, horses, sea otter), and may also explain the repeated isolation of a particular genotype. Exclusively asexual S. neurona transmission is not known, unlike those coccidians capable of placental transmission or those, like T. gondii, whose bradyzoites and tachyzoites are in the future would constitute very strong evidence that S. neurona is capable of perpetuating clonally.

Although alleles are distributed among North American isolates of S. neurona as would be expected in a freely interbreeding population, a “Wahlund” effect appears to significantly constrain gene flow across the Americas as a whole. Among North American S. neurona, no evidence for substructure is evident when they are subjected to likelihood-based clustering analysis, used to reconstruct phylogenies based on inter-individual allelic distance, or evaluated for linkage disequilibrium. This contrasts markedly with the evident allelic and genotypic distinction between these and South American isolates. Understanding whether genes fail to freely flow among these groups primarily owing to their allopatry, a historic founder’s effect, or to reproductive incompatibility (Maynard Smith et al., 1993) would be of considerable biological interest. Diminishing returns have probably been reached in the sampling of S. neurona in North America in establishing the general extent and partitioning of their genetic diversity. By contrast, further sampling from South America would undoubtedly promote our understanding of their diversity and population structure.

How may the natural history of S. neurona have been constrained by that of its definitive hosts? Only two of six Didelphis spp. (Lemos and Cerqueira, 2002) established themselves in North America after formation of the Panamanian isthmus, approximately three million years ago (Ma). Didelphid fossils in Florida date to the Irvingtonian age, less than 1.8 Ma (Marshall et al., 1982). Whereas D. virginiana occupies a broad and expanding range from Costa Rica north into Canada, including major portions of Mexico and the United States, D. marsupialis is known north of Central America only in southeastern Mexico. Correspondingly little overlap characterizes the South American distributions of D. marsupialis with the distribution of D. albiventris, Didelphis imperfecta and Didelphis pernigra (formerly all considered D. albiventris). The former are prevalent in interior Amazonia and throughout northern and western regions of the continent, whereas the latter occupy a more southeasterly distribution—with some overlap in Andean Peru, Ecuador, Columbia and Venezuela. These largely disjunct host distributions may limit migration and gene flow between North and South America. At present, we do not know whether the hosts endemic to particular regions vary in their susceptibility to such infections, or whether parasite fecundity varies among hosts. Knowing if North American S. neurona isolates are infective to South American opossum species would be helpful in understanding the degree of potential gene flow between the continents.

Notably little differentiation characterizes S. neurona across the vast North American continent. An evolutionarily recent and rapid expansion may explain the lack of differentiation among isolates collected several thousand miles apart and from diverse host species (opossums, horses, sea otter), and may also explain the repeated isolation of a particular genotype. Exclusively asexual S. neurona transmission is not known, unlike those coccidians capable of placental transmission or those, like T. gondii, whose bradyzoites and tachyzoites are
infectious not only to cats but also to other intermediate hosts. Transmission of *S. neurona* may, instead, require the sexual cycle completed in opossums. It is clear that *S. neurona* populations encompass many recombinant genotypes, and any particular clone will comprise a comparatively small proportion of the total. Nonetheless, it is notable that a particular multilocus genotype, against great odds, occurs three times in our North American sample of *S. neurona*. Should that genotype again be observed in future sampling, it would constitute strong evidence for the existence of asexual transmission alternatives.

No discernable genetic signatures typify the isolates that induced symptomatic neurological disease in horses or in a Southern sea otter. Rather, these samples were genotypically “typical” of North American *S. neurona*. Thus, we found no predisposition of certain *S. neurona* genotypes to induce neurological disease. If natural variation in virulence does occur, the underlying cause(s) appear not to be embedded within a particular genomic background.

**Acknowledgements**

We thank Daniel Howe for providing us with two *S. neurona* isolates and Lakeia Brown for her work in the laboratory. We are also grateful to Tad Sonstegard and Tina Sphos for all their assistance with the sequencing and fragment analysis.

**References**


*falciparum* populations in the Brazilian Amazon region. J. Infect. Dis. 190, 1547–1555.


