Biological Journal of the Linnean Society, 2007, **92**, 541–560. With 8 figures

Genetic variation and structure in native populations of the fire ant *Solenopsis invicta*: evolutionary and demographic implications

KENNETH G. ROSS[1]*, MICHAEL J. B. KRIEGER[2], LAURENT KELLER[3] and D. DEWAYNE SHOEMAKER[4]

[1]*Department of Entomology, University of Georgia, Athens, GA 30602, USA*
[2]*Swiss Federal Institute of Technology Zurich, 8092 Zurich, Switzerland*
[3]*Department of Ecology and Evolution, University of Lausanne, 1015 Lausanne, Switzerland*
[4]*CMAVE, USDA-ARS, Gainesville, FL 32608, USA*

Received 12 May 2006; accepted for publication 10 January 2007

We studied population genetic variation and structure in the fire ant *Solenopsis invicta* using nuclear genotypic and mitochondrial DNA (mtDNA) sequence data obtained from samples collected throughout its native range. Geographic populations are strongly differentiated at both genomes, with such structure more pronounced in Brazil than in Argentina. Higher-level regional structure is evident from the occurrence of isolation-by-distance patterns among populations, the recognition of clusters of genetically similar, geographically adjacent populations by ordination analysis, and the detection of an mtDNA discontinuity between Argentina and Brazil coinciding with a previously identified landform of biogeographical relevance. Multiple lines of evidence from both genomes suggest that the ancestors of the ants we studied resembled extant northern Argentine *S. invicta*, and that existing Brazilian populations were established more recently by serial long-distance colonizations and/or range expansions. The most compelling evidence for this is the corresponding increase in F_K (a measure of divergence from a hypothetical ancestor) and decrease in genetic diversity with distance from the Corrientes population in northern Argentina. Relatively deep sequence divergence among several mtDNA clades, coupled with geographical partitioning of many of them, suggests prolonged occupation of South America by *S. invicta* in more-or-less isolated regional populations. Such populations appear, in some cases, to have come into secondary contact without regaining the capacity to freely interbreed. We conclude that nominal *S. invicta* in its native range comprises multiple entities that are sufficiently genetically isolated and diverged to have embarked on independent evolutionary paths.

ADDITIONAL KEYWORDS: allozymes – colonization – dispersal – gene flow – microsatellites – migration – mtDNA – population differentiation – range expansion.

INTRODUCTION

Crucial elements in unraveling a species' evolutionary history are the description of the distribution of population genetic variation and the inference of the historical demography and gene flow regimes from this description (Avise, 2000, 2004). The task of converting the patterns etched in population genetic variation into evolutionary and natural history narratives has become increasingly relevant for organisms that are pests, vectors of human disease, endangered species, model organisms for research, or otherwise of special concern. This is because knowledge gleaned from population genetic studies can aid in deciphering the paths of adaptation and diversification of such organisms, thus providing a necessary context for understanding the characteristics that make them special (Bohonak *et al*., 2001; Frankham, Briscoe & Ballou,

*Corresponding author. E-mail: kenross@uga.edu

2002; Lee, 2002; Shimizu, 2002; Sivasundara & Hey, 2003; Veuille *et al*., 2004). With the advent of multiple types of molecular markers of nuclear and organellar genomes (Zhang & Hewitt, 2003; Avise, 2004), as well as the appearance of advanced analytical techniques to evaluate data generated from them (Pritchard, Stephens & Donnelly, 2000; Beerli & Felsenstein, 2001; Hey & Machado, 2003; Beaumont, 2004), there is now unprecedented power to extract useful information from population genetic data for synthesis with other relevant results to enrich our evolutionarily knowledge of such organisms.

The fire ant *Solenopsis invicta* Buren is a highly social insect of considerable concern to biologists. The species occupies a vast native range in South America, where it is an ecologically significant component of the ant fauna, and it recently has become a serious invasive pest in the USA and elsewhere (Lofgren, 1986; Callcott & Collins, 1996; Henshaw *et al*., 2005). Largely because of the enormous amount of research devoted to it in newly-colonized areas, *S. invicta* has emerged as a useful model for an array of biological problems (Tschinkel, 2006). Nonetheless, the population genetics of native *S. invicta* and its nearest relatives has received relatively little attention. Despite this gap, it is clear that these ants comprise a group of closely-related species that display modest morphological and genetic differences between them and exhibit, in some cases, substantial intraspecific population differentiation (Ross *et al*., 1997; Ahrens, Ross & Shoemaker, 2005; Ross & Shoemaker, 2005; Shoemaker, Ahrens & Ross, 2006a; Pitts, McHugh & Ross, 2007). Compounding the complexity of relationships within the group is the apparent occurrence of interspecific hybridization in some circumstances (Ross & Shoemaker, 2005). This complexity poses several important challenges that must be addressed to enhance the value of *S. invicta* as a model species and promote research intended to underpin management strategies for invasive populations (Goodisman & Hahn, 2005; Ross & Shoemaker, 2005). Foremost among these is the need to generate a detailed picture of the distribution of genetic variation in nominal *S. invicta* based on appropriate markers and samples. With such data, the origin and diversification of this species, the emergence of gene flow patterns within it, and the phylogenetic and breeding relationships to its closest relatives can begin to be resolved.

Earlier genetic studies of native *S. invicta* used nuclear and/or mitochondrial DNA (mtDNA) data, often derived from limited sampling, to generate preliminary views of the population genetics and phylogeography of this ant (Ross & Trager, 1990; Ross *et al*., 1997; Ahrens *et al*., 2005; Ross & Shoemaker, 2005), but no comprehensive effort to integrate data for both genomes from samples collected widely over the native range has yet been attempted. In the present study, we remedy this shortcoming by analysing data for diverse classes of markers from appropriate exemplar samples using a suite of traditional and newer methods.

MATERIAL AND METHODS

Samples

A total of 568 nests were sampled from 13 populations of nominal *S. invicta* located at ten localities in Brazil and Argentina (Table 1, Fig. 1, inset). At two of the Argentine localities, Corrientes and Formosa, nests of both social forms of this species were abundant, so the sympatric forms are here considered as separate populations (Ross *et al*., 1997). The monogyne (**M**) social form is characterized by nests with a single egg-laying queen, whereas the polygyne (**P**) form is characterized by nests with multiple such queens; the two forms differ also in a number of other important reproductive and life-history traits (Ross & Keller, 1995). Polygyne nests were rare or (usually) absent among nests collected at the remaining sites (Mescher *et al*., 2003). At the Arroio dos Ratos site in Brazil, two apparently fully reproductively isolated entities referred to as *S. invicta* coexist (Ross & Shoemaker, 2005); these are also treated as separate populations in this study (designated hereafter as the 'Arroio **X**' and 'Arroio **Y**' populations). All specimens were identified as nominal *S. invicta* by Dr James C. Trager or Dr James P. Pitts (Trager, 1991; Pitts *et al*., 2007).

Genetic markers

A single specimen per nest was selected for genetic analysis. Genotypes were scored at seven allozyme loci (Shoemaker, Costa & Ross, 1992; Ross *et al*., 1997) and seven microsatellite loci (Krieger & Keller, 1997; Shoemaker *et al*., 2006b), yielding a total of 14 polymorphic nuclear loci surveyed (Table 1). The number of individuals per population for which nuclear genetic data were obtained were in the range 9–83 (mean = 43.7). Sequence data for a 920-bp fragment of the mtDNA that includes portions of the cytochrome oxidase subunit *I* and *II* genes (Ahrens *et al*., 2005) were obtained from a subset of the same individuals used to generate the nuclear data, with the following exceptions: identical sets of individuals from the Arroio **Y** and Rosario populations were used to generate both the nuclear and mtDNA data, and additional individuals not available for nuclear analyses were sequenced for the mtDNA in the São Gabriel do Oeste population. GenBank accession numbers for most of the mtDNA sequences are provided in Appendix II of Ahrens *et al*. (2005); accession numbers for

Table 1. Variation at genetic markers surveyed in native *Solenopsis invicta* populations

	Arroio dos Ratos **X**	Arroio dos Ratos **Y**	Ceu Azul	Campo Grande	Corrientes **M** form	Corrientes **P** form	Formosa **M** form	Formosa **P** form	Pedra Preta	Pontes E Lacerda	Rinco dos Cabrais	Rosario	São Gabriel do Oeste
Nuclear loci	**35**	**9**	**83**	**45**	**36**	**43**	**35**	**35**	**62**	**30**	**81**	**43**	**31**
Allozymes													
Aat-2	2	1	1	1	3	1	2	2	1	1	2	2	1
	(0.084)	(0)	(0)	(0)	(0.160)	(0)	(0.472)	(0.502)	(0)	(0)	(0.128)	(0.023)	(0)
Acoh-1	2	1	2	1	2	2	5	2	1	2	1	1	1
	(0.029)	(0)	(0.167)	(0)	(0.024)	(0.021)	(0.214)	(0.134)	(0)	(0.066)	(0)	(0)	(0)
Acoh-5	2	1	3	3	2	2	3	3	2	2	3	2	3
	(0.029)	(0)	(0.279)	(0.628)	(0.347)	(0.241)	(0.446)	(0.407)	(0.428)	(0.283)	(0.223)	(0.238)	(0.668)
Est-2	3	1	3	2	2	7	3	1	1	1	6	4	2
	(0.057)	(0)	(0.443)	(0.023)	(0.447)	(0.539)	(0.084)	(0)	(0)	(0)	(0.713)	(0.192)	(0.125)
G3pdh-1	1	1	1	1	1	1	2	2	1	1	1	1	1
	(0)	(0)	(0)	(0)	(0)	(0)	(0.228)	(0.283)	(0)	(0)	(0)	(0)	(0)
Gpi	1	2	1	2	1	1	1	1	2	1	2	1	1
	(0)	(0.208)	(0)	(0.046)	(0)	(0)	(0)	(0)	(0.063)	(0)	(0.048)	(0)	(0)
Pgm-1	2	2	2	2	4	5	2	4	3	2	3	3	3
	(0.476)	(0.321)	(0.059)	(0.131)	(0.182)	(0.377)	(0.287)	(0.256)	(0.480)	(0.033)	(0.153)	(0.529)	(0.235)
Microsatellites													
Sol-6	4	5	6	6	10	9	11	9	4	9	6	8	5
	(0.535)	(0.764)	(0.732)	(0.741)	(0.819)	(0.843)	(0.758)	(0.595)	(0.735)	(0.672)	(0.730)	(0.786)	(0.480)
Sol-11	5	6	7	6	11	12	10	11	9	7	8	9	6
	(0.591)	(0.813)	(0.530)	(0.401)	(0.718)	(0.633)	(0.845)	(0.758)	(0.754)	(0.666)	(0.694)	(0.791)	(0.555)
Sol-18	3	3	3	2	4	7	4	4	3	6	4	3	4
	(0.435)	(0.486)	(0.272)	(0.497)	(0.473)	(0.429)	(0.189)	(0.189)	(0.357)	(0.673)	(0.216)	(0.257)	(0.531)
Sol-20	7	4	6	4	14	13	10	12	4	6	14	7	4
	(0.755)	(0.472)	(0.421)	(0.746)	(0.830)	(0.795)	(0.800)	(0.773)	(0.442)	(0.701)	(0.829)	(0.788)	(0.682)
Sol-42	12	10	14	10	20	18	20	21	8	15	14	15	7
	(0.883)	(0.931)	(0.853)	(0.863)	(0.915)	(0.898)	(0.932)	(0.919)	(0.686)	(0.899)	(0.889)	(0.900)	(0.783)
Sol-49	12	6	11	9	14	15	16	16	7	5	14	11	13
	(0.835)	(0.854)	(0.866)	(0.817)	(0.910)	(0.909)	(0.929)	(0.912)	(0.611)	(0.219)	(0.900)	(0.825)	(0.765)
Sol-55	9	5	9	8	11	11	15	12	9	9	11	10	4
	(0.661)	(0.611)	(0.779)	(0.678)	(0.872)	(0.857)	(0.890)	(0.861)	(0.639)	(0.744)	(0.837)	(0.799)	(0.548)
All nuclear loci *	4.64	3.43	4.66†	4.07	7.07	7.43	7.43	7.14	3.93	4.79	5.61†	5.50	3.93
	(0.383)	(0.425)	(0.387)	(0.404)	(0.458)	(0.454)	(0.505)	(0.469)	(0.372)	(0.354)	(0.455)	(0.432)	(0.307)
mtDNA	**33**	**9**	**66**	**29**	**31**	**25**	**21**	**17**	**47**	**28**	**56**	**43**	**51**
	4	4	11	5	10	16	12	11	2	6	18	9	2
	(0.705, 0.0015)	(0.778, 0.0014)	(0.722, 0.0021)	(0.717, 0.0111)	(0.624, 0.0216)	(0.947, 0.0259)	(0.929, 0.0128)	(0.941, 0.0145)	(0.043, 0.0006)	(0.331, 0.0029)	(0.876, 0.0129)	(0.585, 0.0184)	(0.077, 0.0026)

Sample sizes [numbers of individuals (= nests)] are indicated separately for the nuclear loci and mitochondrial DNA (mtDNA) in bold. Other entries are the numbers of variants observed (alleles or haplotypes), below which are listed in parentheses the expected heterozygosity (H_{exp}) of the nuclear loci or the haplotype diversity (H) and nucleotide diversity (π), respectively, of the mtDNA.

*Allelic richness is shown in place of the number of variants for all nuclear loci considered collectively.

†Estimates of allelic richness in Ceu Azul and Rinco dos Cabrais are adjusted according to the method of Leberg (2002).

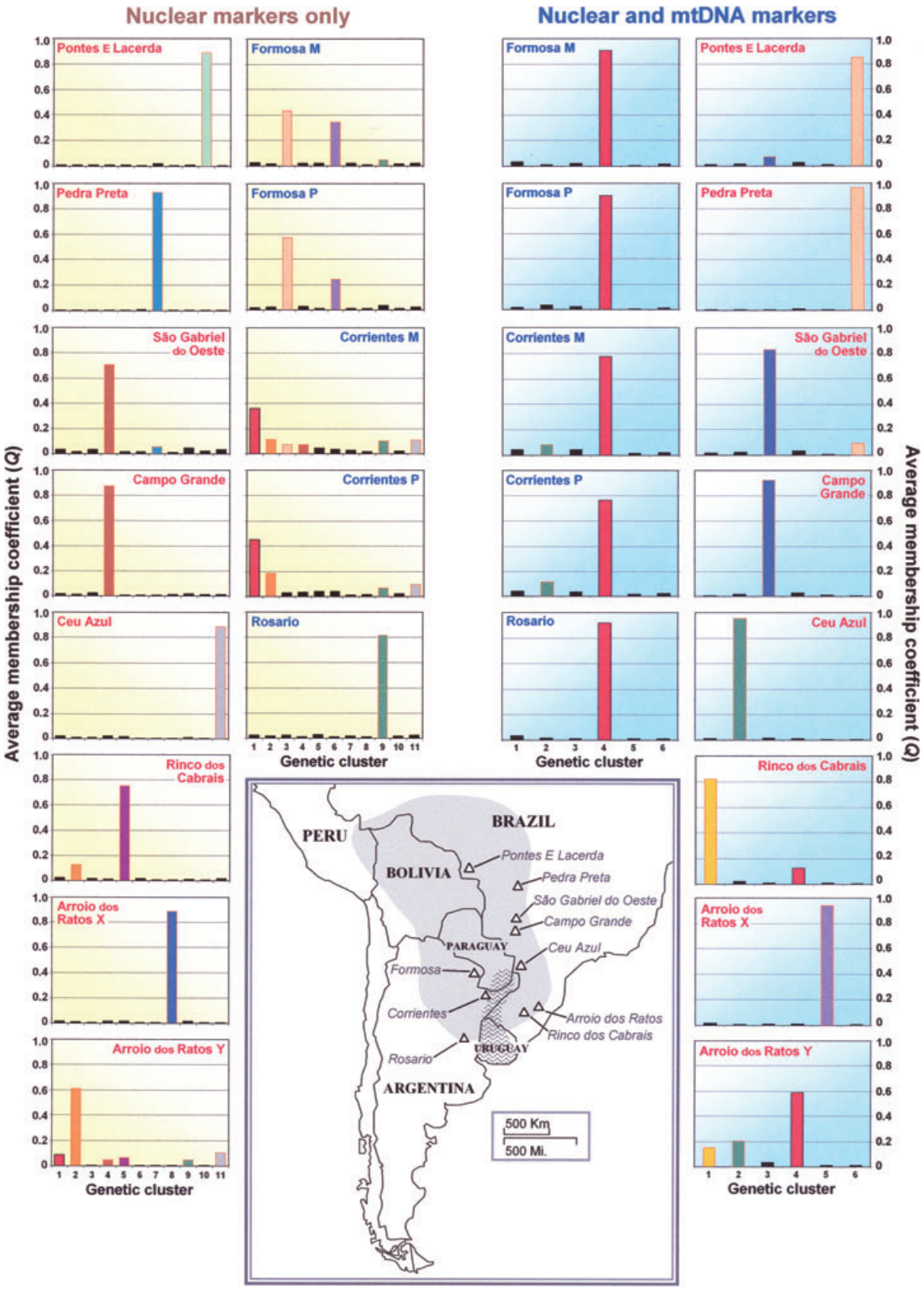
Nuclear markers only
Nuclear and mtDNA markers
Pontes E Lacerda
Formosa M
Pedra Preta
Formosa P
São Gabriel do Oeste
Corrientes M
Campo Grande
Corrientes P
Ceu Azul
Rosario
Rinco dos Cabrais
Arroio dos Ratos X
Arroio dos Ratos Y
Average membership coefficient (Q)
Genetic cluster
BRAZIL
PERU
BOLIVIA
PARAGUAY
URUGUAY
ARGENTINA
Pontes E Lacerda
Pedra Preta
São Gabriel do Oeste
Campo Grande
Ceu Azul
Formosa
Corrientes
Rosario
Arroio dos Ratos
Rinco dos Cabrais
500 Km
500 Mi.

Figure 1. Assignment of *Solenopsis invicta* from each study population to genetic clusters inferred from STRUCTURE simulations (based on average membership coefficient, Q). Results for nuclear markers only are shown in the left half of the figure (Brazilian and Argentine populations are in separate columns) and results for all markers combined are shown on the right. Bars representing clusters for which $Q < 0.05$ are coloured black. Inset: Locations of study populations. The native range of *S. invicta* is shown with grey shading. A major landform of presumed biogeographical importance, the Mesopotamia floodplain, is indicated by stippling.

sequences of the **Y** haplotype clade are AY499580, AY499581, AY499589, and AY499590. All mtDNA sequences were aligned by eye using the alignment of Ahrens *et al*. (2005) as a template. The number of individuals per population for which mtDNA sequence data were obtained were in the range 9–66 (mean = 35.1) (Table 1). A compilation of the genetic data used in the present study is available upon request (from K.G.R.).

Genetic diversity and disequilibrium analyses

Estimates of the extent of nuclear genetic diversity [allelic richness and expected heterozygosity (H_{exp})] were obtained for each population using the program GENEPOP (Raymond & Rousset, 1995a); estimates of allelic richness for the two populations with the largest sample sizes (Ceu Azul, Rinco dos Cabrais) were corrected by taking the mean values from 100 random resamplings of 35 individuals (Leberg, 2002). Estimates of mtDNA diversity [number of haplotypes, haplotype diversity (H), and nucleotide diversity (π)] were obtained using the program ARLEQUIN (Schneider, Roessli & Excoffier, 2000). The Arroio **Y** population was excluded from comparisons of diversity among populations because of the small sample size.

Departures from single-locus (Hardy–Weinberg) equilibrium at the population level were examined by estimating values of F_{IS} for the combined nuclear markers in the hierarchical F_{ST} analyses described below (conducted using the program GDA; Lewis & Zaykin, 2002). We also calculated values of F_{IS} separately for each locus and population. Significant disequilibrium was inferred in all cases where the 95% confidence limits obtained by bootstrapping over loci (10 000 replicates) did not overlap zero.

Departures from linkage equilibrium between all pairs of nuclear loci in each population were examined by conducting randomization tests to approximate the Fisher's exact test (Zaykin, Zhivotovsky & Weir, 1995) using GDA. Any departures from Hardy–Weinberg equilibrium are accounted for in this procedure.

Genetic differentiation analyses

To recognize genetically distinct clusters of individuals and infer levels of population admixture, we employed the Bayesian method of Pritchard *et al*. (2000) as implemented in the program STRUCTURE. The method makes use of individual multilocus genotypic data to evaluate models assuming different numbers of clusters based on the posterior probabilities given the data, model, and prior information. Each sampled individual is probabilistically assigned to a reconstructed genetic cluster based on its multilocus genotype and the allele frequencies estimated for each cluster. No prior information (e.g. sample location) was used in our initial runs of STRUCTURE. The models employed assume some level of population admixture but allow allele frequencies to vary independently across populations. All other model parameter values were the defaults for the program. All simulations used 100 000 Markov chain Monte Carlo iterations in the burnin phase and 300 000 iterations in the data collection phase, with four independent runs conducted on each set of data and parameter values to ensure equilibration by the end of burnin and consistency in estimation of the posterior probabilities. Selection of the number of distinct clusters was based on evaluation of the ΔK statistic of Evanno, Regnaut & Goudet (2005). The analyses were conducted on all nuclear genes as well as on the combined nuclear and mtDNA data. For the latter analyses, mtDNA haplotypes were binned into classes corresponding to seven well-supported haplotype lineages detected in a previous study of native *S. invicta* (Shoemaker *et al*., 2006a). Separate analyses of each geographical population also were conducted using the nuclear data to detect any lower-level structure (Evanno *et al*., 2005).

In a final set of STRUCTURE simulations, we incorporated the geographical localities and social form of samples as priors to calculate values of F_K for each study population. This statistic can be interpreted as an analogue of F_{ST} that measures the divergence of each extant population from a single hypothetical ancestral population (Pritchard *et al*., 2000; Falush, Stephens & Pritchard, 2003).

The magnitude of genetic differentiation between geographical populations was evaluated by estimating values of F_{ST} (Weir & Cockerham, 1984) for the two classes of nuclear markers separately and combined, ρ_{ST} (Rousset, 1996) for the microsatellites, and Φ_{ST} (Excoffier, Smouse & Quattro, 1992) for the mtDNA using the programs GENEPOP and ARLE-

QUIN. The latter two statistics take into account presumed mutational relationships of the variants, whereas the first does not. The statistical significance of interpopulation differentiation was determined by means of exact tests (Raymond & Rousset, 1995b) using GENEPOP, with probabilities combined across nuclear loci using the *Z*-transform test (Whitlock, 2005). The Arroio **Y** population was excluded from these comparisons because of the small sample size. We also estimated F_{ST} and Φ_{ST} simultaneously at two levels (region, geographical population) using GDA and ARLEQUIN, respectively, to hierarchically partition total genetic variation at the nuclear and mtDNA genomes (Weir & Cockerham, 1984; Excoffier *et al*., 1992). Statistical significance of the differentiation at each level was determined by bootstrapping over loci (10 000 replicates) for F_{ST} or permuting haplotypes across individuals (20 000 replicates) for Φ_{ST}. The two regions designated in these hierarchical analyses, Argentina and Brazil, were so conceived because a previous mtDNA study reported a pronounced phylogeographical break between ants from the two countries (Ahrens *et al*., 2005).

Isolation-by-distance (IBD) analyses were conducted to learn whether genetic differentiation between populations increases with their geographical separation. Analyses of the relationships of $F_{ST}/(1 - F_{ST})$ or $\Phi_{ST}/(1 - \Phi_{ST})$ with the natural logarithm of geographical distance were conducted for the nuclear markers and mtDNA, respectively, using GENEPOP (Slatkin, 1993; Rousset, 1997). The two social forms at Corrientes and Formosa were pooled for the nuclear but not the mtDNA analyses because of the similar nuclear compositions of sympatric forms (see below). Significance of IBD relationships was determined by means of Mantel tests based on 10 000 data permutations coupled with estimation of Spearman rank correlation coefficients. Residuals of the linear regression of F_{ST} on geographical distance were plotted against geographical distance, and significance of this relationship was tested with a Mantel test (Hutchison & Templeton, 1999).

We employed the ordination technique known as nonmetric multidimensional scaling (NMDS) to help reveal higher-level patterns of genetic relationships among populations. This technique reduces multidimensional allele or haplotype frequency relationships represented in a matrix of pairwise distances between populations to a few dimensions that explain most of the original distance data (Lessa, 1990; Guiller, Bellido & Madec, 1998). We used the program VISTA (Young, 1996) to conduct NMDS analyses based on pairwise values of Nei's genetic distance (D) for the nuclear loci and the net number of nucleotide differences (D_A) for the mtDNA obtained from the programs PHYLIP (Felsenstein, 2004) and ARLEQUIN, respectively. We determined the best dimensionality for each model by generating scree plots (Kruskal & Wish, 1978) to locate an elbow in the curve depicting the total variance in the data explained with each added dimension; any dimensions beyond the elbow explain relatively little additional variance and thus were not retained. A stress statistic measuring the discrepancy between the matrix of model distances in N-dimensional space and the original distance matrix was calculated (Kruskal, 1964), and a method of iterative approximations was applied until values of this statistic declined to an asymptote, at which point the model was accepted. We graphed projections of model output in the first three dimensions, which in all cases jointly accounted for > 75% of the total variance in the original distance data, to distinguish clusters of genetically similar populations. Individual geographical populations were subdivided for the nuclear analyses if multiple clusters were inferred from the STRUCTURE analyses.

Gene flow analyses

Patterns of gene flow between sampled populations were explicitly evaluated using the programs MIGRATE (Beerli & Felsenstein, 1999, 2001; Beerli, 2004) and BAYESASS (Wilson & Rannala, 2003). The Bayesian coalescent approach in MIGRATE was employed to infer historical rates of gene flow at both genomes by analysing the allozyme and microsatellite data, both separately and together, as well as by analysing the mtDNA data separately. Model settings for the nuclear data were an infinite alleles model of mutation, a uniform prior, random starting genealogies for each chain, and variable mutation rates among loci (the rates for the microsatellites were set at four times those of the allozymes). Results were combined across three replicate chains, each of which recorded 5000 of 110 000 sampled genealogies for parameter estimation (the first 10 000 genealogies, representing the burnin, were discarded). Other settings were the defaults for the program. Model settings for the mtDNA sequences were an F84 mutation model, a transition/tranversion ratio of 6.7 (determined empirically from previous data; Shoemaker *et al*., 2006a), a uniform prior, and random starting genealogies for each chain. Results were combined from ten replicate chains, each of which recorded 10 000 of 210 000 sampled genealogies (again, the first 10 000 represented the burnin).

The Bayesian non-equilibrium approach in BAYESASS was used to infer recent nuclear gene flow rates by analysing the allozyme and microsatellite data separately and together (the program requires diploid data). The model was run for 12 000 000 iterations (the first 2 000 000 were discarded as burnin), with

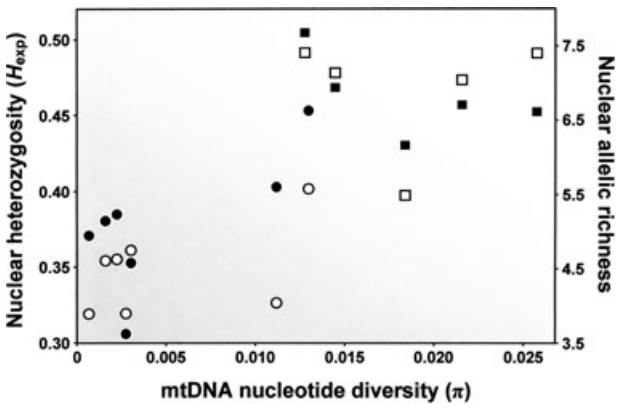

Figure 2. Association of genetic diversity at the nuclear and mitochondrial DNA genomes in the study populations. Black symbols represent nuclear heterozygosity and white symbols represent nuclear allelic richness. Brazilian populations are denoted with circles and Argentine populations with squares.

5000 of these iterations sampled for parameter estimation. Delta values were set at 0.15 to achieve accepted numbers of proposed changes to the Markov chain between 40% and 60% of the total number of iterations.

RESULTS

Population genetic diversity

Most measures of the extent of genetic diversity within the sample populations are significantly correlated, both between the different nuclear marker classes and between the nuclear markers and the mtDNA (diversity values are provided in Table 1). For example, significant correlations were found between the allozymes and microsatellites for each nuclear diversity measure (Spearman rank correlations; $P = 0.041$ for heterozygosity and $P < 0.001$ for allelic richness). Also, mtDNA nucleotide diversity is significantly correlated with both overall nuclear heterozygosity and allelic richness (Spearman rank correlations; both $P < 0.01$; Fig. 2). Importantly, estimates of the latter three statistics tend to be significantly greater for Argentine than Brazilian populations (Mann–Whitney tests; all $P < 0.02$; Fig. 2). More specifically, genetic diversity assessed for both genomes by all measures tends to decrease with geographical distance from Corrientes, Argentina (Fig. 3), the location with ants most closely resembling a hypothetical ancestral *S. invicta* population (see below). Among the Brazilian populations, Rinco dos Cabrais invariably has the highest diversity across all measures.

Single-locus and linkage disequilibrium

The lower bootstrap 95% confidence limit for the overall F_{IS} value based on all nuclear markers is greater than zero, indicating a significant general deficiency of heterozygotes relative to Hardy–Weinberg expectations within the study populations. Examination of the individual-locus values revealed two microsatellite loci with atypically high F_{IS} values in most populations (*Sol-18* and *Sol-20*), suggestive of scoring problems, null alleles, or effects of selection at these loci. Their removal results in 95% confidence limits for F_{IS} that bracket zero. Exclusion of these loci from the main analyses of the present study had minimal effects on the reported results.

Bootstrap 95% confidence limits for F_{IS} calculated separately for each population include zero for all populations except Rosario and Rinco dos Cabrais. When *Sol-18* and *Sol-20* were removed, the new limits bracket zero for the former but not the latter population. Thus, there appears to be a genuine deficiency of heterozygotes in Rinco.

Significant linkage disequilibria between pairs of nuclear genes generally occurred at frequencies < 5% in each study population. Exceptions are the Campo Grande (16.4%) and Rinco (27.3%) populations. Following removal of *Sol-18* and *Sol-20*, the frequency of significant linkage disequilibrium in the former population falls below 5%, whereas that in Rinco remains relatively high (21.8%). Thus, Rinco dos Cabrais stands out from the other populations in terms of the magnitude of its single-locus and linkage disequilibrium, as well as its diversity at both genomes.

Identification of genetically distinctive clusters

Application of the method of Evanno *et al*. (2005) to the posterior probabilities obtained from the STRUCTURE simulations showed a clear peak in ΔK values at 11 genetic clusters when just the nuclear markers were considered, and at six clusters when the mtDNA data were included as well. The average assignment of individuals from each geographical population to each of these clusters (membership coefficient, Q) is depicted in Figure 1. Considering first the results from the nuclear data, most of the Brazilian populations as well as the Rosario (Argentina) population have very high membership (> 80%) in a single unique genetic cluster. That is, these geographical populations tend to be genetically distinct and have low levels of admixture with other populations. Exceptions are: (1) the neighbouring populations São Gabriel do Oeste and Campo Grande, which share dominant membership in the same cluster (with São Gabriel also having an approximately 6% membership in the dominant cluster at Pedra Preta, the

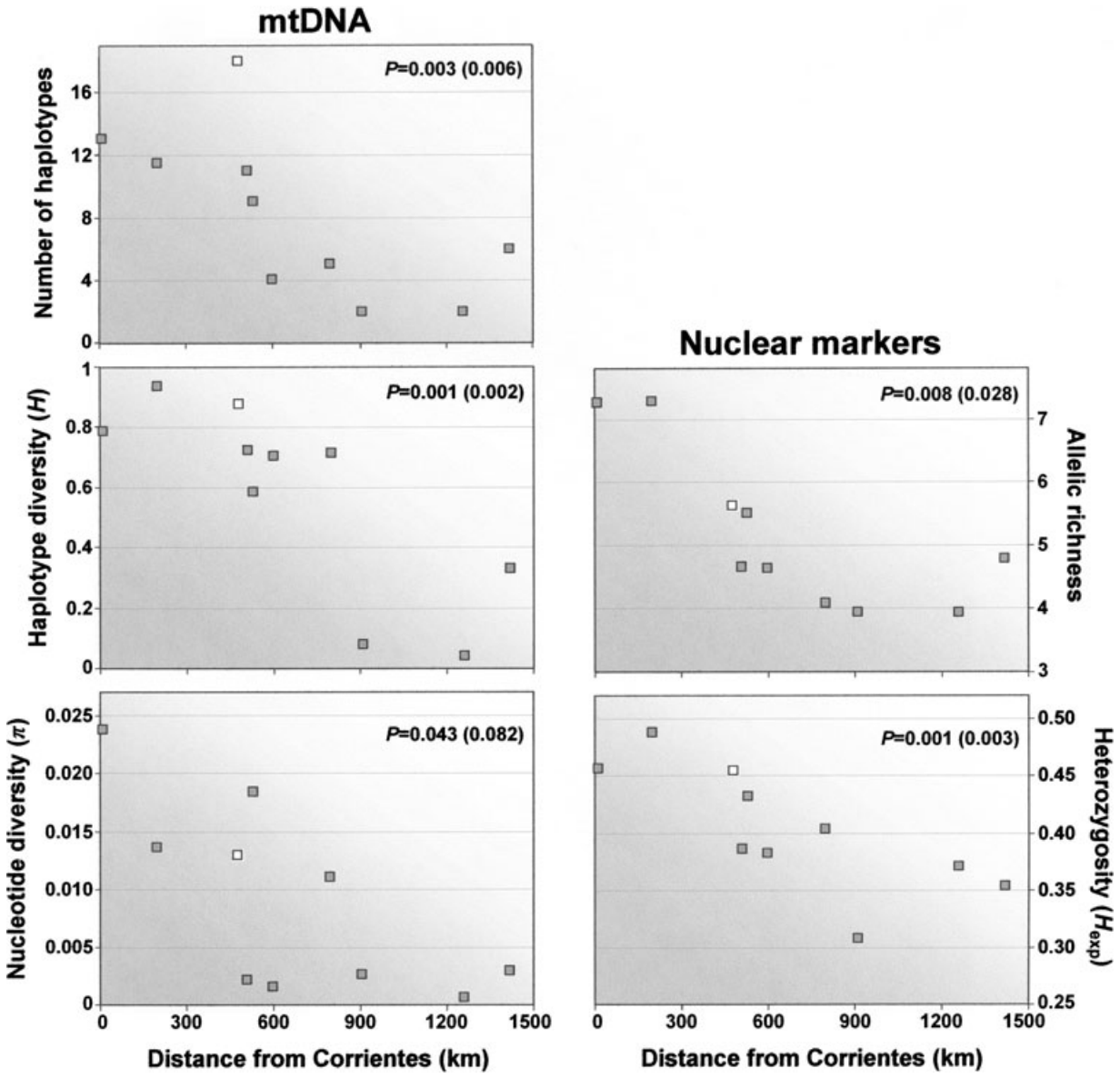

Figure 3. Association of nuclear and mitochondrial DNA diversity with geographical distances of study populations from Corrientes, Argentina. The Rinco dos Cabrais, Brazil population is indicated by the white square. Spearman rank correlation probabilities of no association between diversity and distance are shown (excluding the Rinco population in parentheses).

population immediately north); (2) the Arroio **Y** population, which in addition to majority membership in a unique cluster displays substantial minority membership in several clusters dominant in neighbouring Brazilian and Argentine populations; and (3) the Rinco population, which besides majority membership in a unique cluster displays substantial admixture (approximately 13%) with the adjacent Arroio **Y** population. Also exceptional are the Formosa and Corrientes populations from northern Argentina, which appear to be heavily admixed. In Formosa, this is due largely to the joint presence of two dominant unique clusters whereas, in Corrientes, there is one dominant unique cluster as well as substantial minority representation of clusters dominant elsewhere in Argentina and in the nearest Brazilian populations (Campo Grande, Ceu Azul, Arroio **Y**). STRUCTURE analyses conducted separately on each geographical population revealed that only in Rinco and the Formosa **P** form do the posterior probabilities implicate the presence of multiple genetic clusters (two in each case). Notably, the two clusters discerned within

each of these populations are relatively similar in their genetic composition (see NMDS analyses below).

Addition of the mtDNA data to the nuclear data results in consolidation of several of the clusters recognized by STRUCTURE (Fig. 1). Most notably, all of the Argentine populations now have overwhelmingly dominant membership in a single shared cluster although, in Corrientes, there is also substantial minority representation (approximately 10%) of a cluster characteristic of the nearest Brazilian population, Ceu Azul. Also, the two northernmost Brazilian populations have dominant membership in a unique shared cluster. Finally, with the addition of the mtDNA data, the dominant cluster in the Arroio **Y** population is now the same cluster that dominates in Argentina, indicating an affinity of this Brazilian population with the Argentine samples based on the total genetic evidence.

Magnitude of population genetic differentiation

The magnitude of differentiation between geographical populations is significantly correlated between the allozymes and microsatellites, with the highest correlation obtained when F_{ST} is used for both sets of nuclear markers (Mantel test on Spearman rank correlation coefficient; $P = 0.001$; Fig. 4A). Thus, these very different sets of nuclear markers register congruent patterns of population differentiation, although allozyme differentiation consistently exceeds microsatellite differentiation (compare axis scales in Fig. 4A). For the microsatellites, F_{ST} and ρ_{ST} values also are highly correlated ($P < 0.001$). Importantly, F_{ST} for all nuclear markers is significantly correlated with Φ_{ST} for the mtDNA ($P = 0.006$; Fig. 4B), indicating a general congruence in divergence between populations at the two genomes. Apparent outliers with respect to this intergenomic congruence involve several populations that are relatively similar at the nuclear markers but highly divergent at the mtDNA (Fig. 4B; see also below).

Results of the hierarchical analyses of F_{ST} and Φ_{ST} are shown in Table 2. The results are congruent between the allozymes and microsatellites in suggesting no meaningful differentiation between regions (Argentina and Brazil) but substantial differentiation among geographical populations within the regions. Again, the magnitude of among-population structure registered by the allozymes exceeds that for the microsatellites. A different overall pattern is observed for the distribution of mtDNA variation, with very large proportions occurring between regions as well as among populations. This pattern is expected based on the results of Ahrens *et al*. (2005), who described a major mtDNA phylogeographical discontinuity between the two regions that is superimposed on strong differentiation among the component populations. Even after accounting for the higher-level structure, the proportion of mtDNA variation residing among populations far exceeds that for the nuclear markers, a pattern evident also in the single-level analysis (compare axis scales in Fig. 4B). Thus, geographical populations of native *S. invicta* appear substantially more differentiated at their mtDNA than their nuclear genomes.

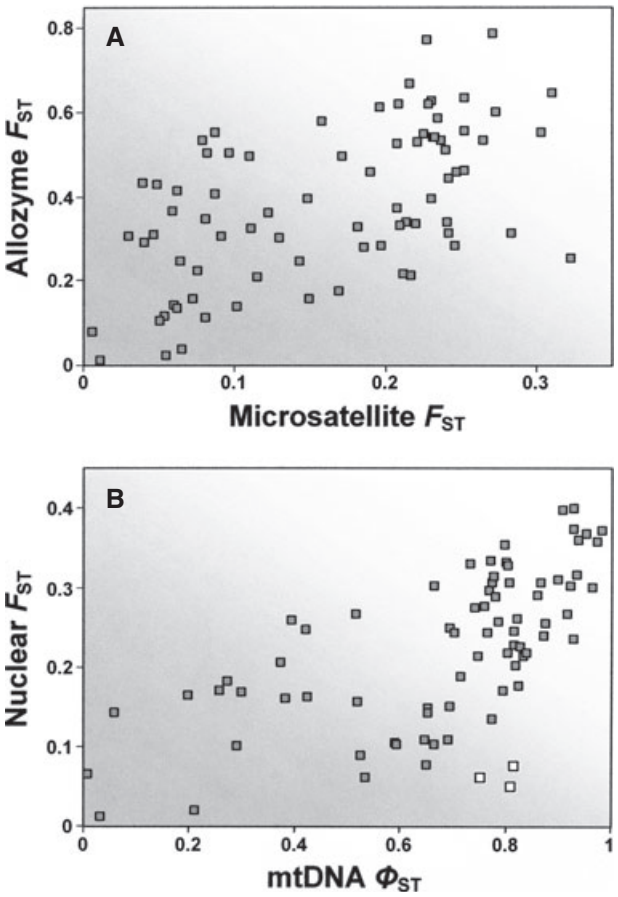

Figure 4. Association of pairwise population differentiation measures obtained for the allozymes and microsatellites (A) and for all nuclear markers and the mitochondrial DNA (B). Three apparent outliers in the bottom graph are depicted with white symbols.

Exact tests were conducted to pinpoint instances of significant pairwise population differentiation. Considering all the nuclear data, all populations are highly significantly differentiated, with only the two social forms in Formosa approaching nonsignificant differentiation ($P = 0.012$). Considering the mtDNA data, with the exception of the two forms in Formosa ($P = 0.287$) all population pairs are highly significantly differentiated.

Table 2. Proportion of total genetic variance distributed at two levels of structure [*regions* (Brazil, Argentina) and *populations*] as assessed by hierarchical analyses of F_{ST} (nuclear markers) and Φ_{ST} (mitochondrial DNA)

	Source of variance	Proportion of variance
Nuclear markers	Between regions	0.000 (0.000, 0.002)
	Among populations within regions	0.217* (0.373*, 0.172*)
	Within populations	0.783 (0.627, 0.826)
mtDNA	Between regions	0.239*
	Among populations within regions	0.540*
	Within populations	0.221

Results for the nuclear data are shown for all markers combined, as well as for the allozymes and microsatellites separately (respectively, in parentheses).
*Differentiation significant at $P < 0.05$ (nuclear markers) or $P < 0.001$ (mtDNA) at this level of structure.
mtDNA, mitochondrial DNA.

IBD analyses revealed that nuclear genetic differentiation increases significantly with geographical separation (Fig. 5A), whether all the markers are considered (Mantel test; $P = 0.002$), just the allozymes are considered ($P = 0.01$), or just the microsatellites are considered ($P = 0.004$). No significant association was found between geographical distances and the residuals from the linear regression of F_{ST} on geographical distance (Mantel test; $P = 0.464$; Fig. 5B), suggesting that the observed IBD did not arise from a long-term balance between gene flow and genetic drift (migration-drift equilibrium). Similar results for the nuclear markers were obtained when only the eight Brazilian populations were analysed. On the other hand, only a marginally significant pattern of IBD was found for the mtDNA for all populations ($P = 0.052$), and no such pattern was found for this genome for just the Brazilian populations ($P = 0.166$).

Judging from the pairwise F_{ST} values, levels of nuclear differentiation among Brazilian populations generally exceed those in Argentina (two-tail independent-sample permutation test with 1000 permutations; $P < 0.001$), a finding upheld when the allozymes and microsatellites are considered separately (both $P < 0.004$). This pattern persists even when the two northernmost or three southernmost Brazilian populations are excluded (both $P < 0.001$ with all loci). A similar pattern of greater mtDNA differentiation in Brazil than Argentina based on Φ_{ST} values also was found ($P < 0.005$), as reported previously for a subset of our samples (Ahrens *et al*., 2005). Again, this result is robust to exclusion of the geographically peripheral Brazilian populations.

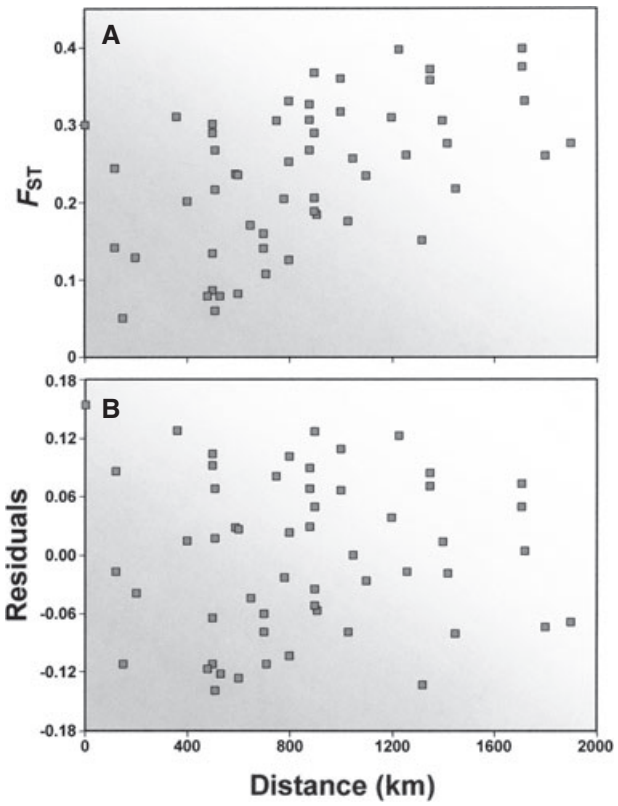

Figure 5. Results of isolation-by-distance (IBD) analyses for all nuclear markers. The relationship of pairwise population differentiation (F_{ST}) with geographical distance is depicted in A, whereas the relationship of the residuals of the linear regression of F_{ST} on geographical distance with geographical distance is depicted in B.

Higher-level genetic relationships among populations

Relationships among populations were visualized using NMDS. The genetic similarity of populations at the nuclear genome has a clear geographical component (Fig. 6). The northern populations of Pontes E Lacerda and Pedra Preta form a distinct group, as do the closely situated central-range populations of São

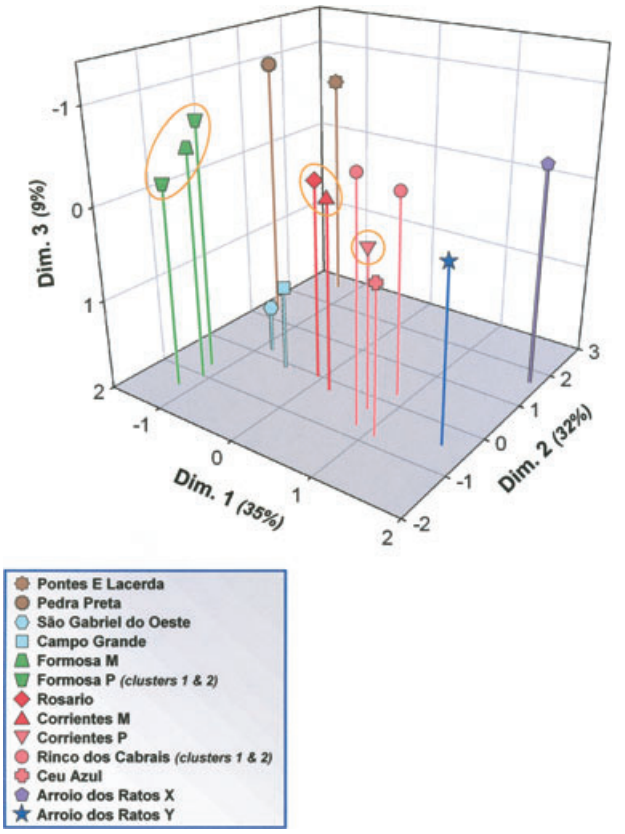

Figure 6. Projections of nonmetric multidimensional scaling (NMDS) model output in the first three dimensions for the nuclear markers. The percentage of the total variance in the original distance data explained by each dimension is shown in parentheses on the appropriate axis. Coloring of drop lines and symbols depicts groups of genetically similar populations. Argentine samples are distinguished by orange halos.

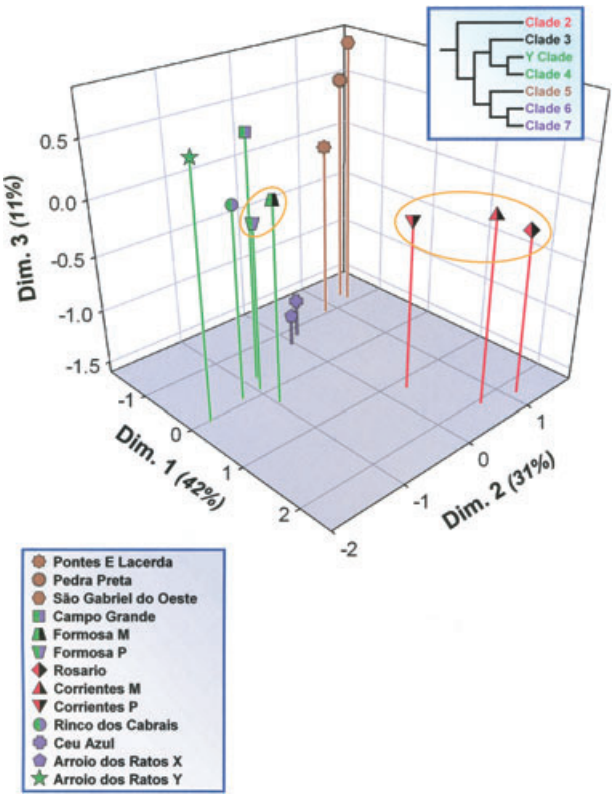

Figure 7. Projections of nonmetric multidimensional scaling (NMDS) model output in the first three dimensions for the mitochondrial DNA. The percentage of the total variance in the original distance data explained by each dimension is shown in parentheses. Coloring of drop lines depicts groups of genetically similar populations, and colouring of symbols indicates population haplotype compositions (haplotypes present at frequencies < 5% not shown). Argentine samples are distinguished by orange halos. Inset: Tree depicting relationships of major mtDNA haplotype clades in native *Solenopsis invicta* (Shoemaker *et al*., 2006a). Clade colours correspond to colours of symbols used in NMDS model output.

Gabriel and Campo Grande. The two genetic clusters in the Formosa **P** form inferred from the STRUCTURE analyses group closely with one another and with the Formosa **M** form, collectively constituting a relatively distinct assemblage. A large group comprising ants from the south-central part of the range is apparent (Ceu Azul, Rinco, Corrientes, Rosario), with the Corrientes **M** and Rosario populations somewhat distinct within it. Again, the two genetic clusters in Rinco inferred by STRUCTURE are relatively similar. The two forms from Arroio dos Ratos on the south-eastern margin of the range are relatively distinct from each other and from the remaining samples, although the **Y** form appears to have closer affinities with ants from the neighbouring populations in Ceu Azul, Rinco, and Corrientes than with sympatric conspecifics of the **X** form (see also Fig. 1). The latter form is quite divergent from all the other ants studied.

A somewhat different picture of relationships emerges from the NMDS analyses of the mtDNA (Fig. 7). Four distinct groups of sequence variants are evident, and these correspond largely to major haplotype lineages identified by Shoemaker *et al*. (2006a). In parallel with results from the nuclear markers, a geographical component to the distribution of mtDNA variation clearly exists, as shown by the restriction of clade 2 to a group comprising the central Argentina populations, of clade 3 to all three Argentina populations, of the Y clade/clade 4 lineage to the group of central range populations, of clade 5 to the northern populations, and of clades 6 and 7 predominantly to southern Brazil samples. On the the other hand, these geographical patterns of mtDNA affinity do not

always mirror closely the patterns of nuclear affinity revealed by NMDS. For example, the Ceu Azul and Arroio **X** populations are closely allied at the mtDNA but very divergent at the nuclear genes, whereas São Gabriel clusters with the more northerly populations according to the mtDNA but with its southern neighbour Campo Grande based on the nuclear genes. The São Gabriel/Campo Grande comparison represents one of the three examples of neighbouring populations with similar nuclear gene pools but dissimilar mtDNA gene pools shown as outliers in Figure 4B. The remaining two such outliers are the Corrientes **M**/Rinco and Corrientes **P**/Ceu Azul comparisons, which are of special interest because they suggest that the Mesopotamia floodplain, a landform associated with a phylogeographical discontinuity in the mtDNA sequences (Ahrens *et al*., 2005), does not act as a significant barrier to contemporary nuclear gene flow.

Historical and recent gene flow between populations

Bayesian analyses with the program MIGRATE were used to infer historical levels of gene flow among sampled populations. Rates based on the allozymes and microsatellites were found to be highly correlated (Mantel test on Spearman rank correlation coefficient; $P < 0.0001$). Overall nuclear gene flow rates are significantly lower among Brazilian than among Argentine populations, even when several of the southernmost or northernmost Brazilian populations were excluded from consideration (one-tail permutation tests with 1000 permutations; all $P < 0.0001$). Trans-Mesopotamian nuclear gene flow rates between the two countries are not significantly lower than overall rates within either country (both $P > 0.775$). This historical nuclear gene flow between the two regions seems to have occurred predominantly in the direction from Brazil to Argentina (two-tail permutation test; $P < 0.0001$). Surprisingly, rates of mtDNA gene flow inferred by MIGRATE are no lower among Brazilian than among Argentine populations (one-tail permutation test; $P = 0.471$), but trans-Mesopotamian mtDNA gene flow was found to be significantly lower than overall gene flow within either country (both $P < 0.01$). Inspection of mtDNA rates for specific population pairs revealed that many of the highest trans-Mesopotamian rates involve the Formosa population, consistent with the finding of Ahrens *et al*. (2005) that this population at the northern edge of the landform often was involved in long-distance dispersal across it. Consideration of the direction of trans-Mesopotamian mtDNA gene flow revealed a pattern opposite to that observed for nuclear gene flow, with rates from Argentina to Brazil significantly exceeding those in the other direction (two-tail permutation test; $P < 0.0001$).

Bayesian analyses with the program BAYESASS were conducted to assess recent nuclear gene flow levels. Again, rates based on the two classes of nuclear markers are highly correlated (Mantel test on Spearman rank correlation coefficient; $P < 0.0001$) and, again, overall rates are lower among Brazilian than among Argentine populations (one-tail permutation test with 1000 permutations; $P < 0.0001$). As with the estimates of historical nuclear gene flow, current trans-Mesopotamian gene flow tends not to be lower in magnitude than gene flow within either Brazil or Argentina (one-tail permutation tests; both $P > 0.999$). This ongoing interregional migration appears asymmetrical, with Argentine populations (especially the northernmost Corrientes and Formosa populations) more likely than southern Brazilian populations to receive high levels of trans-Mesopotamian immigration (two-tail permutation test; $P < 0.0001$).

Gene flow into the Rinco dos Cabrais population

The Rinco population is unique in several respects. It has the greatest genetic diversity of any Brazilian population, it is the only population that exhibits substantial single-locus and linkage disequilibrium, and it is one of only two populations for which STRUCTURE simulations indicate the presence of two distinct nuclear genetic clusters (see above). The comprehensive STRUCTURE analyses suggest that the minority cluster in Rinco represents individuals with nuclear ancestry in the neighbouring Arroio **Y** population (Fig. 1), a conclusion supported by the BAYESASS analyses (data not shown). Inspection of the haplotype compositions reveals that the mtDNA lineage fixed in the Arroio **Y** form (Y clade) also occurs commonly in Rinco, and results of the MIGRATE analyses implicate high mitochondrial gene flow from Arroio **Y** to Rinco as the cause of this pattern (data not shown). These results taken together indicate that the Rinco population is atypical because it experiences substantial immigration from the strongly differentiated, neighbouring Arroio **Y** population without evident amalgamation of the immigrant and resident gene pools having occurred.

Relationships of sampled populations to ancestral population

Values of F_K were calculated from the STRUCTURE simulations to infer the resemblance of each geographical population to a hypothetical ancestral population of South American *S. invicta*. Results based on

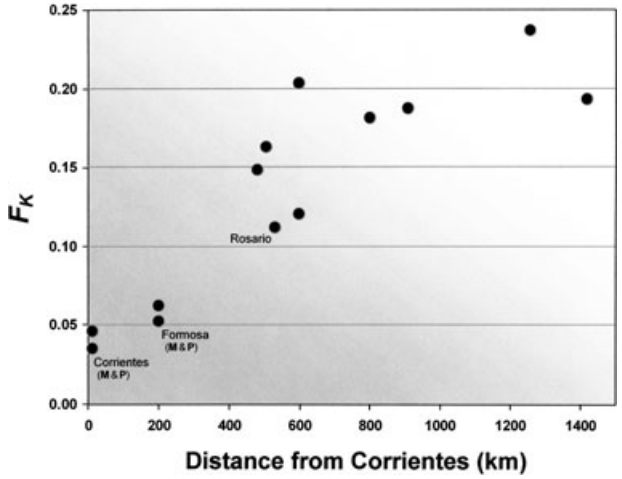

Figure 8. Association of F_K values estimated from STRUCTURE simulations conducted on the nuclear data with geographical distances from Corrientes, Argentina. All Argentine populations are labelled.

just the nuclear markers suggest that the Corrientes populations most closely resemble the ancestral population, followed by the Formosa and Rosario populations (Fig. 8). Values of F_K generally scale with distance from Corrientes (Spearman correlation; $P = 0.007$ when sympatric social forms are pooled). Essentially identical results were obtained when the mtDNA was included with the nuclear markers in the estimates of F_K. These analyses thus suggest that a population resembling extant populations in north-central Argentina is ancestral to all the *S. invicta* populations we sampled, and that the Brazilian portion of the range is likely to have been colonized later than the Argentine portion. This conclusion is consistent with the mtDNA haplotype phylogeny presented in Figure 7, in that a lineage that is sister to all remaining haplotype lineages (clade 2) is found only in Corrientes and Rosario, Argentina.

DISCUSSION

The objective of the present study was to use information from different classes of nuclear markers and the mtDNA to generate a detailed picture of the nature and distribution of genetic variation over the native range of the fire ant *S. invicta*. To that end, genotypic data from 14 nuclear loci and mtDNA sequence data were obtained from 568 nests sampled widely over the South American range of this species, then analysed with a combination of traditional and newer statistical methods. These analyses, in combination with the results of more limited earlier studies (Ross & Trager, 1990; Ross *et al*., 1997; Ahrens *et al*., 2005; Ross & Shoemaker, 2005), allow us to infer some key features of the historical and recent demography and dispersal biology of this important insect species, an essential task for reconstructing its evolution. Furthermore, the genetic information forms the basis of such important applied goals as identifying the source populations from which invasive populations around the world are derived (Van Driesche & Bellows, 1996; Tsutsui *et al*., 2001), a crucial task for focusing control-orientated research in the native range. Our major findings are that: (1) the different classes of markers yield mutually informative and complementary information; (2) geographical populations of native *S. invicta* are strongly genetically differentiated, especially in Brazil; (3) regional genetic affinities among geographical populations exist; and (4) the native *S. invicta* we studied most likely are derived from an ancestral population resembling extant ants from northern Argentina. These findings are discussed in turn below.

COMPLEMENTARINESS OF MARKERS

The two classes of nuclear markers we employed, allozymes and microsatellites, yielded remarkably congruent results with respect to estimates of the extent of diversity within and magnitude of differentiation between populations. Such congruence is not necessarily expected in view of the very different mutational mechanisms giving rise to detectable variation, the substantially different levels of polymorphism, and the different coding/noncoding status of loci in each class (Ross *et al*., 1999; Buonaccorsi, McDowell & Graves, 2001; Mcelroy *et al*., 2003; Gaudeul *et al*., 2004). This congruence shows that demographic and dispersal events have left similar imprints on these distinctive components of the nuclear genome, so that these genome-wide forces, rather than gene-specific forces such as selection and mutation pressure, can be presumed to dominate the observed structuring of nuclear variation in native *S. invicta*.

Although population genetic differentiation scales in parallel at the two classes of nuclear markers, allozyme differentiation was found to consistently exceed that registered by the microsatellites when assessed with F_{ST} estimates (Fig. 4A, Table 2; Ross *et al*., 1997). Geographically localized selection acting on the allozyme loci could play some role in this difference (Mitton, 1997; Neigel, 1997); however, a more likely explanation invokes a combination of two other factors, the statistical property that F_{ST} estimates cannot exceed homozygosity levels (Hedrick, 1999; Buonaccorsi *et al*., 2001; O'Reilly *et al*., 2004) and constraints on detectable microsatellite divergence with low migration due to the comparatively high levels of electromorph homoplasy for these

markers (Balloux *et al*., 2000; Estoup, Jarne & Cornuet, 2002). To account for the differences in levels of variation between the two marker types, we estimated values of a differentiation measure that is standardized by the observed level of variation, G'_{ST} (Hedrick, 2005), for the entire set of populations. Similar estimates for the allozymes (0.492) and microsatellites (0.546) suggest that differences in variation explain at least some portion of the disparities in F_{ST} values between these two classes of markers.

Comparison of the results from the nuclear markers and mtDNA similarly shows some common features in the patterns observed for the two genomes, while also revealing some substantive differences. Common features are the parallel variation in within-population diversity (with diversity decreasing with distance from Corrientes), the parallel variation in population differentiation (with greater differentiation between Brazilian than Argentine populations), and the co-occurring patterns of IBD registered at the two genomes. Joint detection of these patterns at both genomes adds weight to the evolutionary and natural history inferences derived from them.

Differences between the nuclear and mtDNA data also are evident and presumably can be useful for inferring demographic and gene flow patterns in native fire ant populations when interpreted in light of the different properties of these genomes (Ballard & Whitlock, 2004). Distinctive distributions of genetic variance were observed at two spatial scales, with analyses of molecular variance (AMOVA) revealing considerably greater differentiation for the mtDNA than the nuclear markers both among populations and between regions (indeed, no evidence was found for nuclear differentiation between the Argentine and Brazilian regions). Also, distinctive patterns of population affinities were recovered for the two genomes using NMDS analyses; whereas populations tend to group with their geographical neighbours using either data set, the makeup of such groups differs depending on the genome considered. This discrepancy in higher-level groupings is reflected in the strong nuclear similarity, but equally pronounced mitochondrial dissimilarity, characterizing some neighbouring populations. A good example involves São Gabriel do Oeste and Campo Grande, populations located less than 200 km apart in the east-central part of the range. Whereas their overall nuclear similarity causes them to cluster together in the NMDS analysis, their strong mtDNA differentiation leads each to group with a different set of populations when such analysis is applied to the mtDNA sequences.

These disparities in structure registered by the two genomes are likely to result predominantly from their different effective population sizes and transmission dynamics. Stronger structure may be expected for the mtDNA at all scales of analysis given the lower effective population sizes and correspondingly greater effect of drift for this genome when populations are relatively isolated (Avise, 2004). Moreover, greater vagility of males than queens can contribute to this disparity. Such sex-biased dispersal has been proposed previously for invasive *S. invicta* in the USA (Shoemaker *et al*., 2006b) and has been inferred from population genetic data in other ants as well (Gyllenstrand & Seppa, 2003; Sanetra & Crozier, 2003; Sundström, Keller & Chapuisat, 2003; Clemencet, Viginier & Doums, 2005). Finally, the occurrence of the endosymbiotic bacterium *Wolbachia* in native *S. invicta* may have altered the distribution of mtDNA variation without affecting the nuclear DNA, because cytoplasmic genomes are predicted to be influenced by indirect selection associated with this maternally transmitted microbe (Hurst & Jiggins, 2005). Although previous studies found little evidence for such an effect on native fire ant mtDNA (Ahrens *et al*., 2005), we cannot rule out the possibility that *Wolbachia* infection has played some role in driving population mtDNA divergence in *S. invicta* (Ahrens & Shoemaker, 2005).

Extent of population genetic differentiation

A major finding of this study is that geographical populations of native *S. invicta* are highly genetically distinct. The evidence comes in several forms. First, STRUCTURE analyses of the nuclear genes show that most populations represent genetically unique clusters with very low levels of admixture from other clusters. Additionally, AMOVA analyses reveal that a large proportion of the total genetic variance at both the nuclear and mtDNA genomes resides among populations (22% and 54%, respectively). Finally, BAYESASS analyses suggest that populations typically are composed of at least 95% non-immigrant nuclear genes (data not shown). Among the few exceptions, half involve migration between the sympatric social forms in Corrientes and Formosa or between the geographically adjacent populations of Campo Grande and São Gabriel do Oeste (these population pairs also do not comprise distinct clusters in the STRUCTURE analyses). The remaining exceptions in the BAYESASS analyses are the two forms in Corrientes, which appear to experience significant nuclear gene influx from neighbouring populations in Argentina and across the Mesopotamia floodplain in Brazil (this admixture is evident as well in the STRUCTURE results).

A corollary of the marked population genetic structure in native *S. invicta* is that it tends to be more pronounced in Brazil than Argentina, as indicated by significantly elevated values of F_{ST} for the nuclear

markers and Φ_{ST} for the mtDNA in Brazil. This regional effect on differentiation is apparent as well from the STRUCTURE analysis using all the genetic data, which depicts the Argentine samples as constituting a single cluster whereas the five remaining inferred clusters largely comprise single populations or pairs of adjacent populations in Brazil. This regional effect is mirrored by lower estimates of both historical and contemporary nuclear gene flow among Brazilian than among Argentine populations, as inferred from MIGRATE and BAYESASS, respectively. (Surprisingly, MIGRATE analyses of the mtDNA provide no evidence for relatively lower historical gene flow in Brazil.) The present study thus extends to the nuclear genome the conclusion of Ahrens *et al*. (2005) that Brazilian populations generally are more strongly differentiated than those in Argentina. These authors speculated that differences in the distribution of suitable habitat (open, disturbed areas) may be involved, with the more patchy distribution of such habitat in southern Brazil than northern Argentina creating greater barriers to gene flow. Expansion of the range of *S. invicta* into Brazil from Argentina could also be involved, if long distance colonizations coupled with ensuing reductions in effective population sizes created enhanced opportunities for stochastic divergence in population gene frequencies (see also below).

The remarkable degree of divergence measured for some pairs of Brazilian populations raises the question of whether any are sufficiently genetically isolated to have embarked on independent evolutionary paths. Six pairs of Brazilian populations jointly exhibit values of $F_{ST} > 0.35$ and $\Phi_{ST} > 0.925$, which correspond to biparental and maternal evolutionarily effective gene flow levels of less than 0.47 and 0.04, respectively [using $F_{ST} = 1/(4\ N_e m + 1)$ for the nuclear markers and $\Phi_{ST} = 1/(2\ N_e m + 1)$ for the mtDNA; Slatkin, 1987; Neigel, 1997; Whitlock & McCauley, 1999]. Although the meaning of such gene flow estimates can be controversial (Whitlock & McCauley, 1999), these maximal values are well below the threshold values of 1.0 and 0.5 that yield an equilibrium between gene flow and drift for neutral nuclear and mtDNA markers, respectively, in a simple island model (Slatkin, 1987). Direct estimates of levels of admixture from STRUCTURE confirm that gene exchange between these populations is rare, with the proportion of foreign nuclear genes assignable to each paired population estimated at only 0.4–1.3% (mean = 0.8%) (analyses for each population pair were conducted as described in the Material and Methods for the general analyses, with no prior information and $K = 2$ clusters). Thus, at least these several paired populations are likely to be genetically and evolutionarily independent of one another. We note that the average values of F_{ST} and Φ_{ST} for these six population pairs, 0.67 for the allozymes, 0.26 for the microsatellites, and 0.95 for the mtDNA, are in the 93rd, 84th, and 97th percentiles, respectively, of such estimates obtained in large-scale surveys of population differentiation in numerous nominal animal species (Morjan & Rieseberg, 2004).

The Arroio dos Ratos **X** form of *S. invicta* appears to be completely reproductively isolated from the sympatric **Y** form, judging from the observations that the two share no alleles at the allozyme gene *Est-2*, that they display dramatic allele frequency differences at several other nuclear loci, and that they possess non-overlapping sets of haplotypes belonging to different major clades (Fig. 7; Ross & Shoemaker, 2005). Evidently, the **X** form also has participated in little recent gene exchange with the other populations (it is a member of three of the six highly divergent pairs considered above; see also Fig. 6). The clade of South American fire ants including *S. invicta* and its closest relatives is thought to be in a phase of active radiation of species, based on the generally slight nuclear genetic differentiation between species, frequent paraphyly of their mtDNA, and subtle or inconsistent morphological differentiation between them (Ross & Trager, 1990; Ross & Shoemaker, 2005; Shoemaker *et al*., 2006a; Pitts *et al*., 2007). In this view, some nominal species with sizeable ranges, such as *S. invicta*, may be expected to comprise entities that occupy various points on the continuum from freely interbreeding, genetically indistinguishable populations to fully reproductively isolated, genetically diverged populations. The Arroio dos Ratos **X** form appears to fall at the latter end of this spectrum and, as such, probably should be considered a cryptic species.

HIGHER-LEVEL GENETIC STRUCTURE

Higher-level genetic affinities of sampled *S. invicta* are evident at scales of hundreds to thousands of kilometers across the native range. One important example is the IBD patterns observed for both genomes (marginally significant for the mtDNA). Such patterns show that neighbouring populations are genetically more similar to one another than are more distant populations, presumably owing to more recent shared ancestry and/or higher contemporary gene flow (Malécot, 1991; Slatkin, 1993; Hardy & Vekemans, 1999; Yang, 2004). The observed strong nuclear IBD pattern apparently is not the result of a long-term equilibrium between genetic drift within populations and gene flow among them because there is no general trend for increased variation in F_{ST} values with distance between populations (Hutchison & Templeton, 1999). Non-equilibrium processes such as geographical range expansion from a single point

via serial founding of peripheral populations also can produce IBD patterns (Ramachandran *et al*., 2005). The evidence that we found for expansion of *S. invicta* from an original source population in northern Argentina suggests that such a process may have contributed to the emergence of the observed IBD patterns.

Higher-level genetic structure is further evident from the fact that NMDS analyses of both the nuclear and mtDNA data recognize genetically distinctive groups of adjacent populations, even though the composition of the groups so recognized does not correspond closely between the two genomes. For the mtDNA, the tree of haplotype relationships provides additional evidence for a higher-level geographical component to the distribution of mtDNA variation. As examples, one haplotype clade arising from the basal split in the mtDNA tree (clade 2) is confined to two adjacent Argentine populations, whereas a large clade of more recently derived lineages (clades 5–7) is confined almost exclusively to Brazil.

These particular examples of geographical restriction of mtDNA lineages hint at the presence of a discontinuity in the distribution of sequence variation between Argentina and Brazil. Indeed, a 'phylogeographical break' coincident with a landform on the Argentina/Brazil border, the Mesopotamia floodplain, was formally demonstrated by Ahrens *et al*. (2005) by means of nested clade phylogeographical analysis. The situation is made more complex with our addition of new sequences, because the predominantly Brazilian haplotypes no longer constitute a monophyletic group (Fig. 7); thus, the existence of such a break cannot be upheld if this term implies the geographical segregation of a single major haplotype clade (for varying uses of this and related terms, Comes & Abbott, 2000; Riginos & Nachman, 2001; Manel *et al*., 2003; Avise, 2004: 288). Nonetheless, the results of the AMOVA analyses clearly demonstrate mtDNA differentiation between populations from the two regions that transcends differentiation within each region, and estimates of mtDNA gene flow rates from MIGRATE implicate lower rates of historical gene flow across Mesopotamia than within either region. Thus, this floodplain appears to have acted as an important historical barrier to mitochondrial gene flow in fire ants. The Río Paraná, which forms the western boundary of the floodplain, has been hypothesized to be an important biogeographical barrier in several plant groups as well (dos Santos, 1995; Romaniuc-Neto, 1998).

In contrast to the mtDNA, there is no evidence that nuclear gene flow across Mesopotamia is substantially impeded. It is unlikely that homoplasy of the microsatellites masks some signal of ancient regional divergence because the allozymes similarly reveal no differentiation at this level in the AMOVA analyses (Table 2). Rather, demographic differences between the sexes may explain the differential effect of this barrier on gene flow at the two genomes. Any colonization of new areas across a geographical barrier necessarily involves females (mated queens), yet subsequent longer-distance gene flow between established populations, including migration across such a barrier, is likely to occur predominantly by males if they possess superior dispersal abilities (Shoemaker *et al*., 2006b). A scenario in which trans-Mesopotamia range expansion occurred via queens from Argentine populations dispersing intermittently across the floodplain, but in which males subsequently disperse back at much higher rates than queens, is consistent with the different directionalities in gene flow across Mesopotamia that we observed for the nuclear and mtDNA markers.

Out of Argentina

Multiple lines of evidence from both genomes suggest that the ancestors of the native *S. invicta* we studied most closely resembled extant northern Argentina ants, and that other sampled populations were established more recently by long-distance colonization and/or range expansion starting from this source. First, estimates of F_K values from STRUCTURE suggest that the northern Argentine populations, especially Corrientes, most closely resemble such a hypothetical ancestral population and that the genetic similarity of the study populations to this ancestor falls off with distance from Corrientes. Also, there is greater among-population differentiation and lower within-population diversity in Brazil than Argentina, as expected if relatively recent founding of, and subsequent lack of gene flow among, populations east and north of Mesopotamia reduced their long-term effective sizes (McCauley, Raveill & Antonovics, 1995; Pannell & Charlesworth, 2000). Our added observation that diversity at both genomes consistently decreases with distance from Corrientes is an important signal of such range extension outward from Corrientes by means of sequential founder events (Ramachandran *et al*., 2005), a scenario that could also explain the observed IBD patterns in the absence of migration–drift equilibrium. Finally, our MIGRATE analyses of the mtDNA implicate historical queen-mediated gene flow across Mesopotamia primarily in the direction from Argentina to Brazil. Ahrens *et al*. (2005) suggested earlier on the basis of nested clade analysis of the mtDNA that northern Argentine *S. invicta* populations are characterized by long-term persistence and uninterrupted gene flow, whereas Brazilian populations are the products of more recent range extension across the floodplain.

The Rinco dos Cabrais population

This Brazilian population is notable for its relatively high genetic diversity, substantial single-locus and linkage disequilibrium, and presence of multiple genetic clusters inferred by STRUCTURE, all of which appear to be attributable to immigration from a strongly differentiated neighbouring population (the Arroio **Y** form). This immigration clearly has not led to full introgression, suggesting that it is ongoing or, perhaps more likely, that there is strong positive assortative mating or selection against recombinant genotypes in Rinco. Thus, the situation here may resemble that in the neighbouring Arroio population, where the **X** and **Y** forms were recognized and separated for analysis a *priori* based on the apparent lack of introgression between them (Ross & Shoemaker, 2005). Although the northern Argentine populations of Corrientes and Formosa were shown to be more highly admixed than Rinco based on the STRUCTURE and BAYESASS analyses, the absence of disequilibrium in these populations hints that the admixture is relatively old and that introgression is comprehensive. This is consistent with the view that 'admixture' in these Argentine populations in fact largely represents retention of ancestral polymorphisms, as expected if they represent relatively ancient populations with historically large effective sizes.

Gene flow between co-occurring social forms

The co-occurring social forms that we studied tend to be little differentiated from one another in comparison to geographically separated populations (Figs 6, 7; Ross *et al*., 1997). Indeed, exact tests for Formosa revealed only modestly significant nuclear differentiation ($F_{ST} = 0.012$) and insignificant mtDNA differentiation ($\Phi_{ST} = 0.034$) between the forms. The social forms in Corrientes display somewhat greater, highly significant, divergence at both genomes ($F_{ST} = 0.020$, $\Phi_{ST} = 0.201$), and the BAYESASS analysis suggests virtually no nuclear gene exchange between the forms there (data not shown). Differentiation is expected to develop between sympatric social forms over time because most routes of gene exchange seem to be precluded due to social incompatibilities or other factors (Ross *et al*., 1997, 1999). Shallower between-form divergence in Formosa than Corrientes may signal a more limited period of co-occurrence of the forms at the former site, as suggested by the negative association between such divergence and age of populations seen in *S. invicta* in the USA (Shoemaker *et al*., 2006b).

CONCLUSION

Solenopsis invicta displays pronounced regional population differentiation at the nuclear and mtDNA genomes in its native range. This differentiation is sufficiently developed between some Brazilian populations that they are likely to be completely genetically and evolutionarily independent; indeed, the occurrence in sympatry of one pair of clearly reproductively isolated entities demonstrates the potential for the processes driving fire ant population differentiation to culminate in speciation. Pronounced regional genetic differentiation, including the presence of cryptic species, evidently is common also in other groups of ants and can be caused by limited dispersal capabilities of sexuals, patchiness of suitable habitats, or even social barriers to gene flow (Liautard & Keller, 2001; Van der Hammen, Pedersen & Boomsma, 2002; Sanetra & Crozier, 2003; Clemencet *et al*., 2005; Goodisman & Hahn, 2005). A major challenge in such groups is to decide which entities warrant recognition as species taxa and to appreciate the practical and heuristic consequences of such decisions (Porter, 1990; Hey *et al*., 2003; Sites & Marshall, 2004). Resolution of these issues in nominal *S. invicta* will help us gain a clearer picture of the evolution of the species group of South American fire ants to which it belongs and thus facilitate research on this important group of ants. The presence of genetically unique geographical populations of native *S. invicta*, whether ultimately regarded as conspecifics or not, has great practical relevance in that tracing the source of recently established invasive populations around the globe should be feasible.

Our analyses point to the importance of serial long-distance colonizations and range expansions from a northern Argentina source as major features of the historical phylogeography of *S. invicta* responsible for the observed patterns of differentiation. The relatively deep sequence divergence reported between the major mtDNA clades (up to 5.1%; Shoemaker *et al*., 2006a) coupled with the geographical partitioning of many of these clades suggests a substantial period of occupation of South America by *S. invicta* in more-or-less isolated regional populations (Pinceel, Jordaens & Backeljau, 2005). Drift and/or local selection apparently has been sufficiently strong that, in some cases, the resulting diverged entities have come into secondary contact with minimal or no restoration of gene flow, thus illustrating the concept that reproductive isolation can develop in association with range extension (Tregenza, Pritchard & Butlin, 2000). The dynamic nature of gene flow patterns and population structure in these ants is reflected in evidence for both an initial expansion into Brazil from Argentina as well as recent nuclear gene flow primarily in the reverse direction. Future work will focus on generating data for very large numbers of nuclear markers of various classes from an expanded set of populations chosen to completely cover the native range of

nominal *S. invicta*. Such data may provide a well resolved timeline for the major demographic events that have shaped the evolution of these ants.

ACKNOWLEDGEMENTS

We thank Dietrich Gotzek for comments on an earlier draft of this manuscript. This research was supported by grants from the United States Department of Agriculture NRICGP (2003-35302-13497 and 2006-35302-16561) and the Swiss NSF. The use of trade, firm, or corporation names in this publication is for the information and convenience of the reader. Such use does not constitute an official endorsement or approval by the United States Department of Agriculture or the Agricultural Research Service of any product or service to the exclusion of others that may be suitable.

REFERENCES

Ahrens M, Ross KG, Shoemaker DD. 2005. Phylogeographic structure of the fire ant *Solenopsis invicta* in its native South American range: roles of natural barriers and habitat connectivity. *Evolution* **59:** 1733–1743.

Ahrens M, Shoemaker DD. 2005. Evolutionary history of *Wolbachia* infections in the fire ant *Solenopsis invicta*. *BMC Evolutionary Biology* **5:** 35.

Avise JC. 2000. *Phylogeographyraphy: the history and formation of species*. Cambridge, MA: Harvard University Press.

Avise JC. 2004. *Molecular markers, natural history, and evolution*, 2nd edn. Sunderland, MA: Sinauer.

Ballard JWO, Whitlock MC. 2004. The incomplete natural history of mitochondria. *Molecular Ecology* **13:** 729–744.

Balloux F, Brunner H, Lugon-Moulin N, Hausser J, Goudet J. 2000. Microsatellites can be misleading: an empirical and simulation study. *Evolution* **54:** 1414–1422.

Beaumont MA. 2004. Recent developments in genetic data analysis: what can they tell us about human demographic history? *Heredity* **92:** 365–379.

Beerli P. 2004. *MIGRATE*, *version 1.6*. Seattle, WA: Department of Genome Sciences, University of Washington.

Beerli P, Felsenstein J. 1999. Maximum-likelihood estimation of migration rates and effective population numbers in two populations using a coalescent approach. *Genetics* **152:** 763–773.

Beerli P, Felsenstein J. 2001. Maximum likelihood estimation of a migration matrix and effective population sizes in *n* subpopulations by using a coalescent approach. *Proceedings of the National Academy of Sciences of the United States of America* **98:** 4563–4568.

Bohonak AJ, Davies N, Villablanca FX, Roderick GK. 2001. Invasion genetics of new world medflies: testing alternative colonization scenarios. *Biological Invasions* **3:** 103–111.

Buonaccorsi VP, McDowell JR, Graves JE. 2001. Reconciling patterns of inter-ocean molecular variance from four classes of molecular markers in blue marlin (*Makaira nigricans*). *Molecular Ecology* **10:** 1179–1196.

Callcott AMA, Collins HL. 1996. Invasion and range expansion of imported fire ants (Hymenoptera: Formicidae) in North America from 1918–1995. *Florida Entomologist* **79:** 240–251.

Clemencet J, Viginier B, Doums C. 2005. Hierarchical analysis of population genetic structure in the monogynous ant *Cataglyphis cursor* using microsatellite and mitochondrial DNA markers. *Molecular Ecology* **14:** 3735–3744.

Comes HP, Abbott RJ. 2000. Random amplified polymorphic DNA (RAPD) and quantitative trait analyses across a major phylogeographical break in the Mediterranean ragwort *Senecio gallicus* Vill. (Asteraceae). *Molecular Ecology* **9:** 61–76.

Estoup A, Jarne P, Cornuet J-M. 2002. Homoplasy and mutation model at microsatellite loci and their consequences for population genetics analysis. *Molecular Ecology* **11:** 1591–1604.

Evanno G, Regnaut S, Goudet J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* **14:** 2611–2620.

Excoffier L, Smouse PE, Quattro JM. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* **131:** 479–491.

Falush D, Stephens M, Pritchard JK. 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* **164:** 1567–1587.

Felsenstein J. 2004. *PHYLIP (phylogeny inference package)*. Seattle, WA: Department of Genome Sciences, University of Washington.

Frankham R, Briscoe DA, Ballou JD. 2002. *Introduction to conservation genetics*. Cambridge: Cambridge University Press.

Gaudeul M, Till-Bottraud I, Barjon F, Manel S. 2004. Genetic diversity and differentiation in *Eryngium alpinum* L. (Apiaceae): comparison of AFLP and microsatellite markers. *Heredity* **92:** 508–518.

Goodisman MAD, Hahn DA. 2005. Breeding system, colony structure, and genetic differentiation in the *Camponotus festinatus* species complex of carpenter ants. *Evolution* **59:** 2185–2199.

Guiller A, Bellido A, Madec L. 1998. Genetic distances and ordination: the land snail *Helix aspersa* in North Africa as a test case. *Systematic Biology* **47:** 208–227.

Gyllenstrand N, Seppa P. 2003. Conservation genetics of the wood ant, *Formica lugubris*, in a fragmented landscape. *Molecular Ecology* **12:** 2931–2940.

Hardy OJ, Vekemans X. 1999. Isolation by distance in a continuous population: reconciliation between spatial autocorrelation analysis and population genetics models. *Heredity* **83:** 145–154.

Hedrick PW. 1999. Perspective: highly variable loci and their interpretation in evolution and conservation. *Evolution* **53:** 313–318.

Hedrick PW. 2005. A standardized genetic differentiation measure. *Evolution* **59:** 1633–1638.

Henshaw MT, Kunzmann N, Vanderwoude C, Sanetra M, Crozier RH. 2005. Population genetics and history of the introduced fire ant, *Solenopsis invicta* Buren (Hymenoptera: Formicidae), in Australia. *Australian Journal of Entomology* **44:** 37–44.

Hey J, Machado CA. 2003. The study of structured populations – new hope for a difficult and divided science. *Nature Reviews Genetics* **4:** 535–543.

Hey J, Waples RS, Arnold ML, Butlin RK, Harrison RG. 2003. Understanding and confronting species uncertainty in biology and conservation. *Trends in Ecology and Evolution* **18:** 597–603.

Hurst GDD, Jiggins FM. 2005. Problems with mitochondrial DNA as a marker in population, phylogeographic and phylogenetic studies: the effects of inherited symbionts. *Proceedings of the Royal Society of London Series B, Biological Sciences* **272:** 1525–1534.

Hutchison DW, Templeton AR. 1999. Correlation of pairwise genetic and geographic distance measures: inferring the relative influences of gene flow and drift on the distribution of genetic variability. *Evolution* **53:** 1898–1914.

Krieger MJB, Keller L. 1997. Polymorphism at dinucleotide microsatellite loci in fire ant (*Solenopsis invicta*) populations. *Molecular Ecology* **6:** 997–999.

Kruskal JB. 1964. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika* **29:** 1–17.

Kruskal JB, Wish M. 1978. *Multidimensional scaling*. Beverly Hills, CA: Sage Publications.

Leberg PL. 2002. Estimating allelic richness: effects of sample size and bottlenecks. *Molecular Ecology* **11:** 2445–2449.

Lee CE. 2002. Evolutionary genetics of invasive species. *Trends in Ecology and Evolution* **17:** 386–391.

Lessa EP. 1990. Multidimensional-analysis of geographic genetic-structure. *Systematic Zoology* **39:** 242–252.

Lewis PO, Zaykin D. 2002. *GDA (Genetic Data Analysis).* Available at http://lewis.eeb.uconn.edu/lewishome/

Liautard C, Keller L. 2001. Restricted effective queen dispersal at a microgeographic scale in polygynous populations of the ant *Formica exsecta*. *Evolution* **55:** 2484–2492.

Lofgren CS. 1986. History of imported fire ants in the United States. In: Lofgren CS, Vander Meer RK, eds. *Fire ants and leaf cutting ants: biology and management*. Boulder, CO: Westview Press, 36–49.

Malécot G. 1991. *The mathematics of heredity*. San Francisco, CA: Freeman.

Manel S, Schwartz MK, Luikart G, Taberlet P. 2003. Landscape genetics: combining landscape ecology and population genetics. *Trends in Ecology and Evolution* **18:** 189–197.

McCauley DE, Raveill J, Antonovics J. 1995. Local founding events as determinants of genetic structure in a plant metapopulation. *Heredity* **75:** 630–636.

Mcelroy TC, Kandl KL, Garcia J, Trexler JC. 2003. Extinction-colonization dynamics structure genetic variation of spotted sunfish (*Lepomis punctatus*) in the Florida Everglades. *Molecular Ecology* **12:** 355–368.

Mescher MC, Ross KG, Shoemaker DD, Keller L, Krieger MJB. 2003. Distribution of the two social forms of the fire ant *Solenopsis invicta* (Hymenoptera: Formicidae) in the native South American range. *Annals of the Entomological Society of America* **96:** 810–817.

Mitton JB. 1997. *Selection in natural populations*. New York: Oxford University Press.

Morjan CL, Rieseberg LH. 2004. How species evolve collectively: implications of gene flow and selection for the spread of advantageous alleles. *Molecular Ecology* **13:** 1341–1356.

Neigel JE. 1997. A comparison of alternative strategies for estimating gene flow from genetic markers. *Annual Review of Ecology and Systematics* **28:** 105–128.

O'Reilly PT, Canino MF, Bailey KM, Bentzen P. 2004. Inverse relationship between F_{ST} and microsatellite polymorphism in the marine fish, walleye pollock (*Theragra chalcogramma*): implications for resolving weak population structure. *Molecular Ecology* **13:** 1799–1814.

Pannell JR, Charlesworth B. 2000. Effects of metapopulation processes on measures of genetic diversity. *Philosophical Transactions of the Royal Society of London Series B, Biological Sciences* **355:** 1851–1864.

Pinceel J, Jordaens K, Backeljau T. 2005. Extreme mtDNA divergences in a terrestrial slug (Gastropoda, Pulmonata, Arionidae): accelerated evolution, allopatric divergence and secondary contact. *Journal of Evolutionary Biology* **18:** 1264–1280.

Pitts JP, McHugh JV, Ross KG. 2007. Revision of the fire ants of the. *Solenopsis saevissima* species-group (Hymenoptera: Formicidae). *Zootaxa* in press.

Porter AH. 1990. Testing nominal species boundaries using gene flow statistics: the taxonomy of two hybridizing admiral butterflies (*Limenitis*, Nymphalidae). *Systematic Zoology* **39:** 131–147.

Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. *Genetics* **155:** 945–959.

Ramachandran S, Deshpande O, Roseman CC, Rosenberg NA, Feldman MW, Cavalli-Sforza LL. 2005. Support from the relationship of genetic and geographic distance in human populations for a serial founder effect originating in Africa. *Proceedings of the National Academy of Sciences of the United States of America* **102:** 15942–15947.

Raymond M, Rousset F. 1995a. GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity* **86:** 248–249.

Raymond M, Rousset F. 1995b. An exact test for population differentiation. *Evolution* **49:** 1280–1283.

Riginos C, Nachman MW. 2001. Population subdivision in marine environments: the contributions of biogeography, geographical distance and discontinuous habitat to genetic differentiation in a blennioid fish, *Axoclinus nigricaudus*. *Molecular Ecology* **10:** 1439–1453.

Romaniuc-Neto S. 1998. Biodiversity and speciation in the south of Brazil and the basin of the Paraná river: influ-

ences in the *Sorocea* A. St-Hil. (Moraceae) genus species complex. *Comptes Rendus de L'Academie des Sciences Serie II Fascicule A – Sciences de la Terre et des Planetes* **327:** 669–675.

Ross KG, Keller L. 1995. Ecology and evolution of social organization: insights from fire ants and other highly eusocial insects. *Annual Review of Ecology and Systematics* **26:** 631–656.

Ross KG, Krieger MJB, Shoemaker DD, Vargo EL, Keller L. 1997. Hierarchical analysis of genetic structure in native fire ant populations: results from three classes of molecular markers. *Genetics* **147:** 643–655.

Ross KG, Shoemaker DD. 2005. Species delimitation in native South American fire ants. *Molecular Ecology* **14:** 3419–3438.

Ross KG, Shoemaker DD, Krieger MJB, DeHeer CJ, Keller L. 1999. Assessing genetic structure with multiple classes of molecular markers: a case study involving the introduced fire ant *Solenopsis invicta*. *Molecular Biology and Evolution* **16:** 525–543.

Ross KG, Trager JC. 1990. Systematics and population genetics of fire ants (*Solenopsis saevissima* complex) from Argentina. *Evolution* **44:** 2113–2134.

Rousset F. 1996. Equilibrium values of measures of population subdivision for stepwise mutation processes. *Genetics* **142:** 1357–1362.

Rousset F. 1997. Genetic differentiation and estimation of gene flow from *F*-statistics under isolation by distance. *Genetics* **145:** 1219–1228.

Sanetra M, Crozier RH. 2003. Patterns of population subdivision and gene flow in the ant *Nothomyrmecia macrops* reflected in microsatellite and mitochondrial DNA markers. *Molecular Ecology* **12:** 2281–2295.

dos Santos EP. 1995. Phylogeny and phytogeography of the genus *Salvia* L. section Rudes (Benth.) EpI. (Lamiacees). *Biogeographyraphica (Paris)* **71:** 15–32.

Schneider S, Roessli D, Excoffier L. 2000. *Arlequin: a software for population genetics data analysis*, version 2.000. Geneva: Genetics and Biometry Laboratory, Department of Anthropology, University of Geneva.

Shimizu KK. 2002. Ecology meets molecular genetics in *Arabidopsis*. *Population Ecology* **44:** 221–233.

Shoemaker DD, Ahrens ME, Ross KG. 2006a. Molecular phylogeny of fire ants of the *Solenopsis saevissima* species-group based on mtDNA sequences. *Molecular Phylogenetics and Evolution* **38:** 200–215.

Shoemaker DD, Costa JT, Ross KG. 1992. Estimates of heterozygosity in two social insects using a large number of electrophoretic markers. *Heredity* **69:** 573–582.

Shoemaker DD, DeHeer CJ, Krieger MJB, Ross KG. 2006b. Population genetics of the invasive fire ant *Solenopsis invicta* in the USA. *Annals of the Entomological Society of America* **99:** 1213–1233.

Sites JW, Marshall JC. 2004. Operational criteria for delimiting species. *Annual Review of Ecology, Evolution and Systematics* **35:** 199–227.

Sivasundara A, Hey J. 2003. Population genetics of *Caenorhabditis elegans*: the paradox of low polymorphism in a widespread species. *Genetics* **163:** 147–157.

Slatkin M. 1987. Gene flow and the geographic structure of natural populations. *Science* **236:** 787–792.

Slatkin M. 1993. Isolation by distance in equilibrium and non-equilibrium populations. *Evolution* **47:** 264–279.

Sundström L, Keller L, Chapuisat M. 2003. Inbreeding and sex-biased gene flow in the ant *Formica exsecta*. *Evolution* **57:** 1552–1561.

Trager JC. 1991. A revision of the fire ants, *Solenopsis geminata* group (Hymenoptera: Formicidae: Myrmicinae). *Journal of the New York Entomological Society* **99:** 141–198.

Tregenza T, Pritchard VL, Butlin RK. 2000. The origins of premating reproductive isolation: testing hypotheses in the grasshopper *Chorthippus parallelus*. *Evolution* **54:** 1687–1698.

Tschinkel WR. 2006. *The fire ants*. Cambridge, MA: Harvard University Press.

Tsutsui ND, Suarez AV, Holway DA, Case TJ. 2001. Relationships among native and introduced populations of the Argentine ant (*Linepithema humile*) and the source of introduced populations. *Molecular Ecology* **10:** 2151–2161.

Van der Hammen T, Pedersen JS, Boomsma JJ. 2002. Convergent development of low-relatedness supercolonies in *Myrmica* ants. *Heredity* **89:** 83–89.

Van Driesche R, Bellows TS. 1996. *Biological control*. New York, NY: Chapman & Hall.

Veuille M, Baudry E, Cobb M, Derome N, Gravot E. 2004. Historicity and the population genetics of *Drosophila melanogaster* and *D. simulans*. *Genetica* **120:** 61–70.

Weir BS, Cockerham CC. 1984. Estimating *F*-statistics for the analysis of population structure. *Evolution* **38:** 1358–1370.

Whitlock MC. 2005. Combining probability from independent tests: the weighted *Z*-method is superior to Fisher's approach. *Journal of Evolutionary Biology* **18:** 1368–1373.

Whitlock MC, McCauley DE. 1999. Indirect measures of gene flow and migration: $F_{ST} \neq 1/(4Nm+1)$. *Heredity* **82:** 117–125.

Wilson GA, Rannala B. 2003. Bayesian inference of recent migration rates using multilocus genotypes. *Genetics* **163:** 1177–1191.

Yang RC. 2004. A likelihood-based approach to estimating and testing for isolation by distance. *Evolution* **58:** 1839–1845.

Young FW. 1996. *Vista: the visual statistics system*. Research Memorandum 94-1(b), 2nd edn. Chapel Hill, NC: LL Thursone Psychometric Laboratory, University of North Carolina.

Zaykin D, Zhivotovsky L, Weir BS. 1995. Exact tests for association between alleles at arbitrary numbers of loci. *Genetica* **96:** 169–178.

Zhang DX, Hewitt GM. 2003. Nuclear DNA analyses in genetic studies of populations: practice, problems and prospects. *Molecular Ecology* **12:** 563–584.