Syst. Biol. 59(2):162–184, 2010

DOI:10.1093/sysbio/syp089
Advance Access publication on December 14, 2009

Species Delimitation: A Case Study in a Problematic Ant Taxon

KENNETH G. ROSS[1,*], DIETRICH GOTZEK[2], MARINA S. ASCUNCE[3], AND D. DEWAYNE SHOEMAKER[3]

[1]*Department of Entomology, University of Georgia, Athens, GA 30602, USA;*
[2]*Department of Ecology and Evolution, University of Lausanne, 1015 Lausanne, Switzerland; and*
[3]*USDA-ARS Center for Medical, Agricultural and Veterinary Entomology, 1600 S.W. 23rd Drive, Gainesville, FL 32608, USA;*
**Correspondence to be sent to: Department of Entomology, University of Georgia, Athens, GA 30602, USA; E-mail: kenross@uga.edu.*

Received 22 November 2008; reviews returned 13 April 2009; accepted 5 November 2009
Associate Editor: Kelly Zamudio

Abstract.—Species delimitation has been invigorated as a discipline in systematics by an influx of new character sets, analytical methods, and conceptual advances. We use genetic data from 68 markers, combined with distributional, bioclimatic, and coloration information, to hypothesize boundaries of evolutionarily independent lineages (species) within the widespread and highly variable nominal fire ant species *Solenopsis saevissima*, a member of a species group containing invasive pests as well as species that are models for ecological and evolutionary research. Our integrated approach uses diverse methods of analysis to sequentially test whether populations meet specific operational criteria (contingent properties) for candidacy as morphologically cryptic species, including genetic clustering, monophyly, reproductive isolation, and occupation of distinctive niche space. We hypothesize that nominal *S. saevissima* comprises at least 4–6 previously unrecognized species, including several pairs whose parapatric distributions implicate the development of intrinsic premating or postmating barriers to gene flow. Our genetic data further suggest that regional genetic differentiation in *S. saevissima* has been influenced by hybridization with other nominal species occurring in sympatry or parapatry, including the quite distantly related *Solenopsis geminata*. The results of this study illustrate the importance of employing different classes of genetic data (coding and noncoding regions and nuclear and mitochondrial DNA [mtDNA] markers), different methods of genetic data analysis (tree-based and non–tree based methods), and different sources of data (genetic, morphological, and ecological data) to explicitly test various operational criteria for species boundaries in clades of recently diverged lineages, while warning against over reliance on any single data type (e.g., mtDNA sequence variation) when drawing inferences. [Cryptic species; cuticular coloration; ecological niche modeling; fire ants; genetic markers; *Solenopsis*; species delimitation.]

A major function of systematics is to discover and describe species, a task dependent on judicious delimitation of species boundaries. Species delimitation currently is receiving increased attention for several reasons. Species serve as fundamental units of analysis in an array of basic ecological, biogeographical, and evolutionary studies (Isaac et al. 2004; Balakrishnan 2005; Wiens 2007a, 2007b; Renema et al. 2008), and increasingly, species determinations and records have taken on practical importance in fields such as conservation biology, public health, and pest management (Van Driesche and Bellows 1996; Gentile et al. 2002; Agapow et al. 2004; Ross and Shoemaker 2005; Light et al. 2008). Moreover, interest in species delimitation has been invigorated by an influx of new genomic-scale character sets (Turner et al. 2005; Shaffer and Thomson 2007; Minder and Widmer 2008), the introduction of diverse analytical and diagnostic methods (Marshall et al. 2006; Knowles and Carstens 2007; Raxworthy et al. 2007), and conceptual advances relating to species concepts and properties (de Queiroz 2007). Although the empirical task of delimiting species draws on a rich theoretical foundation grounded in models of species formation, the data generated in turn yield unique insights into the genetic underpinnings of lineage diversification and thus fuel the further development of speciation theory (Hey 2006a; Shaffer and Thomson 2007; Nosil 2008; Minder and Widmer 2008; Mullen et al. 2008; Strasburg and Rieseberg 2008; Xie et al. 2008).

Issues of species delimitation inevitably are shaped by our conceptual definitions of the species category and the relation of these to operational criteria for hypothesizing the existence of species taxa. Several authors have argued persuasively that species concepts and species delimitation criteria should be held as separate epistemological constructs that serve either to define the necessary and sufficient properties of the species category or to specify the types of evidence required to assign particular entities to this category (Mayden 1999; Hey 2006b; de Queiroz 2007). One advantage of a clear separation of these constructs is that species delimitation studies are effectively liberated from the controversies surrounding the introduction and acceptance of myriad species concepts. Instead, the task of such studies is clarified as marshalling and synthesizing all evidence relevant to the recognition of evolutionarily independent metapopulation lineages, an appropriate universal conceptual definition of the species category (de Queiroz 2007). Such operational criteria for species delimitation comprise the fundamental properties upon which most earlier "species concepts" were based.

The task of determining whether a population constitutes an independently evolving lineage that is likely to have reached, or is on a trajectory to reach, irreversible evolutionary separation from other populations is particularly challenging in actively radiating groups composed of recently diverged lineages. The difficulty lies in the fact that recently separated species are less likely to possess all or even many of the contingent properties

that constitute operational criteria for their delimitation, such as phenetic distinctiveness, intrinsic reproductive incompatibility, ecological uniqueness, deficits of genetic intermediates, or monophyly (de Queiroz 2007; Shaffer and Thomson 2007). Moreover, the order and timing of appearance of these properties are not generally predictable. An emerging consensus is that the best approach to dealing with such problematic groups is to employ diverse data types and analytical approaches to extract the most meaningful information relevant to acquisition of different properties (Morando et al. 2003; Sites and Marshall 2004; Dayrat 2005; Wiens and Graham 2005; Marshall et al. 2006; Knowles and Carstens 2007; Sei and Porter 2007; Shaffer and Thomson 2007; Strasburg and Rieseberg 2008; Leaché et al. 2009; Menkis et al. 2009; Schlick-Steiner et al. 2010). Even with a comprehensive synthesis of such data, however, some ambiguity in hypotheses of the boundaries of recently diverged species can remain due to incomplete separation, secondary introgression, sampling deficiencies, or disagreement over the importance of various operational criteria.

The *Solenopsis saevissima* species group is an assemblage of apparently recently diverged fire ant species native to South America that, historically, has posed substantial challenges to taxonomists (Creighton 1930; Wilson 1952; Buren 1972; Trager 1991). Despite recent morphology-based studies employing extensive samples and novel character sets (Trager 1991; Pitts 2002; Pitts et al. 2005), distinguishing species remains exceedingly difficult because of a dearth of reliable diagnostic morphological characters coupled with apparently common intraspecific variation that transcends interspecific variation. Ongoing uncertainty over the number and composition of species in this important group may become an impediment to their continued development as premier subjects of ecological and evolutionary studies (Tschinkel 2006) and also has direct practical implications with respect to establishing the identity and source of new invasive pest populations. A recent genetic study of 2 such invasives, *Solenopsis richteri* and the widespread pest *Solenopsis invicta*, revealed strong regional genetic differentiation within the native ranges of each, commensurate with long-term lineage independence (Ross and Shoemaker 2005). A more comprehensive follow-up study of *S. invicta* (Ross et al. 2007) confirmed marked geographical genetic structure across its large native range and documented the occurrence of genetically distinct forms at one locality that likely comprise morphologically cryptic species.

The discovery of previously unknown presumably independent lineages within nominal *S. invicta* sets the context for further detailed investigations of species limits in this group of ants. A promising candidate for such study is *S. saevissima*, the only member of the *S. saevissima* species group that appears to surpass *S. invicta* in terms of both size of its natural range and magnitude of its phenotypic variability (Pitts 2002). Indeed, the observed range of variation led Trager (1991) to state "one is easily convinced they [regional variants of *S. saevissima*] represent separate species" but for the existence of "an array of samples including every possible intermediate condition" of size and coloration. In this study, we use data from 68 genetic markers for 94 individuals of nominal *S. saevissima* and related species, combined with distributional, bioclimatic, and coloration information, in an integrated approach to learn whether the apparent continuum of morphological variation may be superimposed on distinct evolutionary lineages. In this approach, we sequentially test whether populations meet specific operational criteria for candidacy as cryptic species and then we synthesize the results to formulate new hypotheses of species boundaries. The results are intended to contribute to resolution of some persistent problems in the alpha-taxonomy of South American fire ants while illustrating the general power of combining diverse data and analytical methods with a formal testing framework in the task of delimiting closely related species.

Materials and Methods

Samples

The recorded range of *S. saevissima* covers a vast area in Brazil, with the main part extending along the Atlantic coast inland several hundred kilometers and reaching from Rio Grande do Sul State in the south to the mouth of the Amazon River in the north (Fig. 1). A western tongue of the range extends along the Amazon drainage in the north, with records as far west as western Amazonas State and Peru (Trager 1991; Pitts 2002). Fire ants of the *S. saevissima* species group typically nest in open disturbed habitats in areas with sufficient rainfall; thus, these ants do not occur uniformly throughout all parts of their recorded ranges. We collected samples of *S. saevissima* from 68 nests located across much of its range. Most samples (from 59 nests) were obtained by traveling major highways along the eastern and western periphery of the main range, stopping every 30–50 km to collect fire ants during a 30-min period (nests commonly are found along roadsides and in nearby disturbed habitats). Thus, our sampling generally is expected to reflect the occurrence and abundance of this species in these surveyed areas. Additional samples of 9 nests were obtained from a single locality in Envira, Amazonas, in the far western part of the range (Fig. 1). Several other fire ant species have ranges that abut or slightly overlap that of *S. saevissima*, including the common species *S. invicta* and *S. richteri*, and the much rarer *Solenopsis pythia* and *Solenopsis megergates*, whereas the common but quite distantly related fire ant *Solenopsis geminata* co-occurs extensively with *S. saevissima* along Brazil's Atlantic coast (Fig. 1).

The different life stages, sexes, and castes present in each nest were collected to facilitate correct species identification (Pitts 2002) as well as to provide appropriate material for scoring the allozyme loci. Adult major workers were obtained from 100% of sampled nests, larval workers from 78%, adult queens from 62%, and

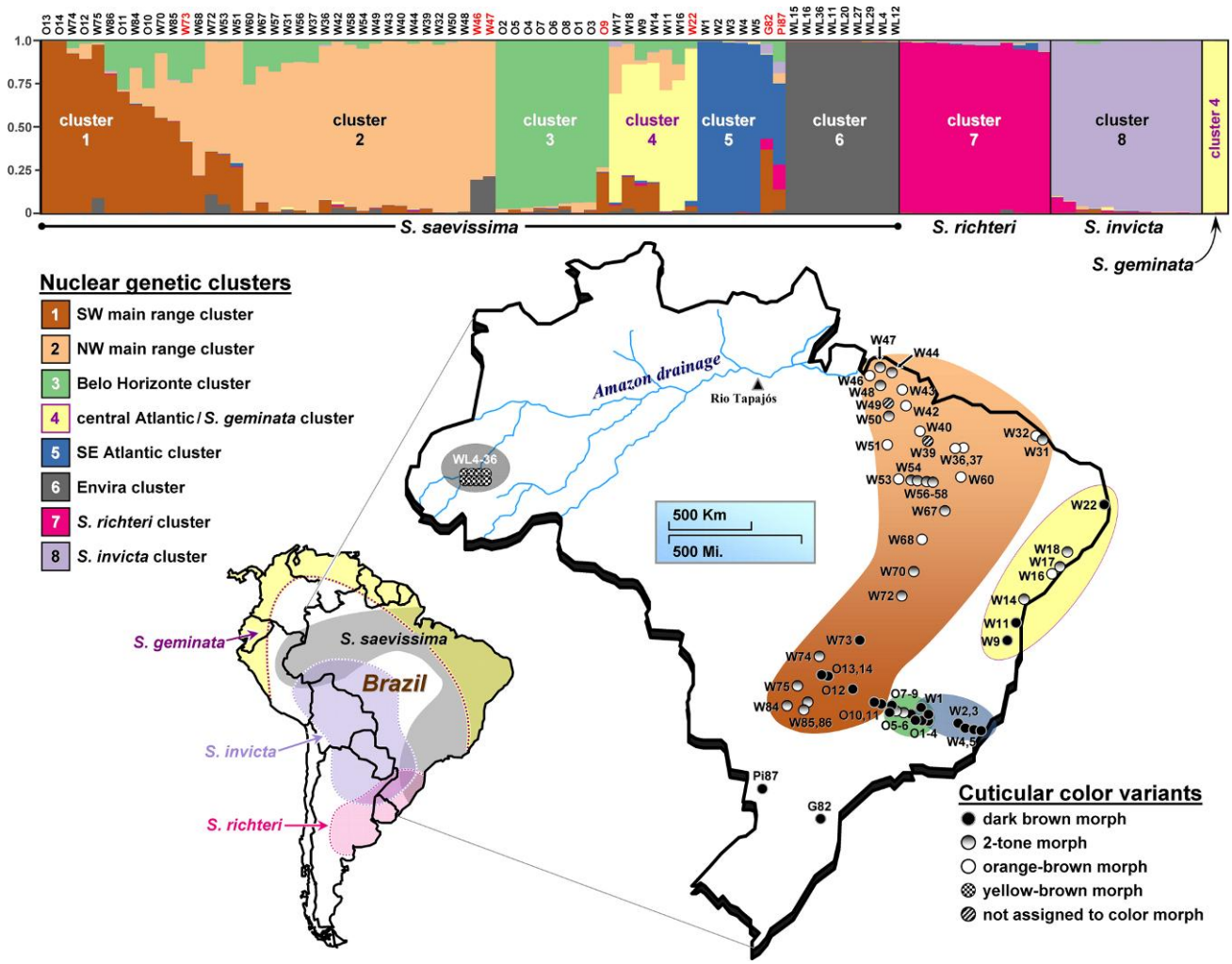

FIGURE 1. Membership coefficients from STRUCTURE analyses for 94 individual fire ants from South America identified as *Solenopsis saevissima*, *Solenopsis richteri*, *Solenopsis invicta*, or *Solenopsis geminata* (top). Each individual is represented by a vertical bar divided into parts proportional to the individual's nuclear ancestry in each of 8 proposed genetic clusters (*S. saevissima* specimens are identified, those of special note in red font). The geographical distribution of the 6 clusters predominantly represented among nominal *S. saevissima* is shown on the map of Brazil, with the sample localities for this species superimposed. Individuals were assigned to a STRUCTURE cluster if the majority of their nuclear ancestry fell in that cluster (G82 and Pi87 had no majority cluster ancestry). The cuticular color morph represented by each *S. saevissima* specimen also is indicated on the locality map, as is the type locality for the species (Rio Tapajós, Brazil). The recorded range of *S. saevissima* is shaded grey in the map of South America, whereas the ranges of *S. invicta*, *S. richteri*, and *S. geminata* are shaded violet, pink, and yellow, respectively (distributional data from Trager 1991; Pitts 2002; Ross and Shoemaker 2005; Ross et al. 2007).

adult males from 38% of nests. Based on this material, study colonies from the main range were identified as *S. saevissima* by J. P. Pitts using existing diagnostic morphological characters of adult workers, winged queens, and males (diagnostic characters include setal form and number in the larval workers; shape and sculpturing of the mesonotum, propodeum, and postpetiole of the adult workers and sexuals; and coloration of the adult males; Pitts 2002). The remaining 9 sampled colonies (those from Envira, Amazonas, for which only adult workers were obtained) were identified as *S. saevissima* by D. F. Williams. Voucher specimens from each sampled nest of *S. saevissima* are deposited in the National Museum of Natural History (NMNH), Washington, D.C., USA.

Genetic Data

Genotypes of a single female from each of the sampled *S. saevissima* colonies were determined at 10 allozyme and 57 microsatellite loci. All the allozyme loci were developed previously, and methods for scoring their genotypes are detailed in Shoemaker et al. (1992). A subset of 7 of the microsatellite loci were developed previously (see Krieger and Keller 1997; Chen et al. 2003), whereas the remaining 50 loci were newly developed using 3 different approaches (see Supplementary Appendix 1, available from http://www.sysbio.oxfordjournals.org/). First, we developed 18 loci from a fire ant genomic library enriched for dinucleotide and trinucleotide repeats. Second, we developed 6 loci by searching for microsatellite motifs in an *S. invicta* expressed

sequence tag database available in GenBank. Finally, we developed 26 loci by searching for microsatellite motifs in a large *S. invicta* sequence data set generated by shotgun-sequencing a cDNA library using 454 Life Sciences (Roche, Indianapolis, IN) pyrosequencing technology. Additional information on the development of the first class of new microsatellite loci appears in Ascunce et al. (2009).

Total genomic DNA for scoring the microsatellite loci was isolated from single individuals using the Puregene DNA Isolation Kit (Gentra-Qiagen, Valencia, CA). One primer of each locus primer pair was labeled at the 5′ end with 1 of 4 fluorescent dyes (FAM, PET, NED, or VIC; Applied Biosystems, Foster City, CA). Primer pairs were combined in multiplex reactions by taking into account polymerase chain reaction (PCR) thermal cycling profiles, dye labels, and expected size and yield of the PCR products. The complete set of 57 loci was amplified in 22 separate 12-μL multiplex PCR reactions, each containing 2–5 pairs of primers, Hot-Start *Taq* 2X Mastermix (Denville Scientific, Metuchen, NJ), 0.06–0.4 μM of each primer, 1–2 μL of total genomic DNA (10–20 ng), and water. Thermal cycling profiles were as follows: one cycle at 94 °C (1 min); followed by 35 cycles at 94 °C (30 s), primer-specific annealing temperature (45 s), and 72 °C (60 s); followed by a single final extension of 72 °C (40 min). Resulting PCR amplicons were diluted (100:1 to 200:1) and pooled into a total of 6 plates, each one representing a different injection on the capillary sequencer. GeneScan 600 LIZ size standard (GeneScan, Metairie, LA; 0.1 μL) was added to all pool-plex dilutions, which subsequently were run on an ABI-3730XL 96-capillary sequencer (Applied Biosystems) at the University of Florida biotechnology facility (Interdisciplinary Center for Biotechnology Research). Microsatellite genotypes were scored using GENEMARKER software (SoftGenetics, State College, PA).

Genomic DNA isolated from our *S. saevissima* specimens was used also to PCR amplify a approximately 785-bp portion of the mitochondrial genome that includes a part of the cytochrome oxidase I (COI) gene. The primers C1-J-2195 and JerryGarcia-C1 were employed in these amplifications (Ahrens et al. 2005; Shoemaker et al. 2006). PCR reaction conditions, thermal cycling profiles, and sequencing reaction protocols are reported in Ahrens et al. (2005). All mitochondrial DNA (mtDNA) amplicons were sequenced in both directions. Comparison of amplicon sequences with complete mtDNA genome sequences from exemplars of *S. invicta*, *S. richteri*, and *S. geminata* indicated that the amplified fragments reside in the mitochondrial genome rather than representing mtDNA homologs inserted in the nuclear genome (Numts).

For comparative purposes, genetic data from all 68 markers were also obtained for 12 individuals (from as many colonies) of *S. invicta* and 12 individuals (colonies) of *S. richteri*. Each species was sampled from 2 widely separated, genetically distinct native populations, with half of the nests sampled at Arroio dos Ratos, Brazil, and half at Rosario, Argentina (Ross and Shoemaker 2005). Genetic data were obtained as well for 2 individuals (colonies) of *S. geminata* from Tallahassee, FL, USA. *Solenopsis invicta* and *S. richteri* are members of the *S. saevissima* species group, whereas *S. geminata* is placed in a different species group (Pitts et al. 2005). For *S. invicta* and *S. richteri*, data for 7 of the allozymes and the mtDNA are from Ross and Shoemaker (2005), data for the allozymes *Ddh-1* and *Lap-1* are from Ross et al. (1997) and Shoemaker et al. (1992), respectively, whereas the allozyme marker *Acy1* was not scored. For *S. geminata*, allozyme data and mtDNA sequences are from Ross et al. (2003) or from unpublished data associated with that study (8 of the allozymes). Data for all 57 microsatellite loci were newly generated for this study for the nonfocal species. Voucher specimens from each sampled nest of these species are deposited in the NMNH.

Genetic Data Analyses

Estimates of gene diversity for the nuclear markers were obtained using equation 2.18a in Hedrick (2000) based on allele frequencies computed by the program MSA (Dieringer and Schlötterer 2003). The same program was used to obtain nuclear allele counts (richness). Estimates of mtDNA sequence haplotype richness and diversity were obtained using the program DNASP (Rozas et al. 2003).

A useful operational criterion for formulating initial hypotheses of species boundaries in problematic groups is the presence of genetic clusters with few genetically intermediate individuals (Mallet 1995). We determined the number of distinct nuclear genetic clusters represented by the individuals in our study using the Bayesian method of Pritchard et al. (2000) as implemented in the program STRUCTURE. The method used the individual multilocus genotypic data from all 67 nuclear loci to evaluate models assuming different numbers of genetic clusters (K) based on the posterior probabilities, given the data and model (we used no prior information). Based on the individual multilocus genotypes and the allele frequencies estimated for the reconstructed clusters, each individual's genome was probabilistically partitioned into membership fractions (ancestry) in each cluster. The models we employed assumed some level of population admixture but allowed allele frequencies to vary independently across populations. Because of the large numbers of rare alleles in our data set, we estimated the parameter λ in an initial run with $K = 1$ and then set λ to this estimated value (0.875) in all subsequent runs. All other model parameter values were the defaults for the program. All simulations used 300,000 Markov chain Monte Carlo generations in the burn-in phase and 1,000,000 generations in the data collection phase. Four independent runs were conducted for each K to ensure equilibration during burn-in and consistency in estimation of the posterior probabilities, with stability in each run ascertained by plotting key parameter values through its course. Selection of

the number of distinct clusters was based on joint evaluation of the ΔK statistic of Evanno et al. (2005) and the estimated posterior probability of the data.

Another approach to evaluating the presence of genetic clusters as an operational criterion is to employ the ordination technique termed nonmetric multidimensional scaling (NMDS). This approach can be a valuable adjunct to STRUCTURE analyses by confirming the presence of clusters as well as providing information on their distinctiveness and phenetic relationships to one another. The technique reduces multidimensional allele frequency relationships represented in a matrix of pairwise distances to a few dimensions that explain most of the original distance data (Lessa 1990; Guiller et al. 1998; see also Patterson et al. 2006; Novembre and Stephens 2008). We used the program PERMAP (Heady and Lucas 2007) to conduct NMDS analyses using D_{ps} distance values ($-$ln proportion of shared alleles; Bowcock et al. 1994), which were calculated with MSA for all pairs of study individuals based on all the nuclear loci. A high-precision convergence setting of 0.000005 was used to find a global minimum mapping solution under a ratio + bounds NMDS analysis with 5% error bounds; the stress statistic of Kruskal (1964) was used as the objective function. The analysis was done using 2 or 3 dimensions. We graphed the object coordinates from the model output to visually distinguish clusters of genetically similar individuals and to infer overall similarity of the various clusters.

Monophyly of a population constitutes another operational criterion for inferring species limits from genetic data (Donoghue 1985). Phylogenetic relationships based on the nuclear genomes of the individuals included in this study were inferred by constructing neighbor-joining (NJ) trees (Saitou and Nei 1987) using the interindividual D_{ps} values. D_{ps} distance is expected to closely reflect the genealogical or phylogenetic affinities of individuals when a large number of informative loci is available (Chakraborty and Jin 1993; Bowcock et al. 1994), and this measure has been employed in conjunction with NJ analysis of individuals and populations in an array of recent studies (Schlötterer 2001; Vila et al. 2001; Caracristi and Schlötterer 2003; Chirhart et al. 2005; Glowatzki-Mullis et al. 2005; Hughes and Hollingsworth 2008; Kanbe and Akimoto 2009). The PHYLIP program package (Felsenstein 2004) was used to construct the trees, with node support evaluated by bootstrapping across loci (1,000 replicates). Support for important nodes was further evaluated by using PHYLIP to construct 1,000 bootstrap trees using both the NJ and the Fitch and Margoliash (1967) algorithms on the Cavalli-Sforza and Edwards (1967) chord distance (D_C). All trees were rooted by specifying *S. geminata* as the outgroup. The NJ/D_{ps} analysis was conducted both with and without the allozyme loci included in the input data set.

To infer the phylogenetic relationships of the recovered mtDNA sequences, all sequences first were aligned by eye. Models of sequence evolution for the complete data set were evaluated using the program JMODELTEST (Posada 2008). Log-likelihoods of different models of nucleotide substitution under NJ tree topologies were compared using the AIC and BIC model selection criteria (Posada and Buckley 2004). Because substantial uncertainty can accompany model choice (Alfaro and Huelsenbeck 2006), we calculated the 95% confidence set of models and used model-averaged parameter estimates in the subsequent analyses. We used BIONJ topologies (Gascuel 1997) to estimate the optimal evolutionary model from 88 possibilities (11 substitution schemes + F + I + G). AIC and BIC converged on nearly identical models (TrN + I + G [Tamura and Nei 1993]). Because of our primary concern with the tree topology, we selected the closest overparameterized model that was implemented in the phylogeny construction programs used (e.g., Huelsenbeck and Rannala 2004; Lemmon and Moriarty 2004).

MtDNA phylogenetic hypotheses were constructed under various optimality criteria. We estimated the tree topology under maximum parsimony (MP) with the program PAUP* (Swofford 2002) using tree bisection-reconnection branch swapping during heuristic searches of tree space on 10 randomly constructed starting trees. Gaps were treated as missing data. The resulting 1134 trees of equal length were summarized using the strict consensus criterion. We estimated the maximum likelihood (ML) topology using the online program PHYML (Guindon and Gascuel 2003; http://www.atgc-montpellier.fr/phyml/). Again, 10 random starting trees were created, followed by subtree pruning-regrafting + nearest-neighbor interchange branch swapping. Tree topology, substitution model parameters, and branch lengths were optimized. We estimated the topology under Bayesian inference (BI) using the program MRBAYES (Ronquist and Huelsenbeck 2003). Four independent runs were conducted with 4 chains each (1 cold and 3 incrementally heated; temperature = 0.05). Chains were run for 7,500,000 generations, with sampling every 1,000th generation. All other parameter values were the program defaults. Convergence was assessed by measuring average standard deviations of split frequencies ($<$0.01), potential scale reduction factor ($\sim$1.0), and plateauing of log-likelihood values. Tree topology, branch lengths, and posterior probabilities of branch support were summarized across all 4 runs using the total 20,000 trees remaining after the burn-in from each run (2,500 trees) was discarded. We calculated the log-likelihoods for the ML, BI, and MP consensus trees by employing the model-averaged estimates of nucleotide substitution under the ML criterion using the program PAUP*. No statistical difference in log-likelihoods was found between the ML and BI trees (Shimodaira–Hasegawa test; Shimodaira and Hasegawa 1999), whereas the MP tree was judged to be significantly poorer than the other 2. Because the ML and BI trees agree in the essential features of their topologies, we report results only for the latter.

Association of the nuclear and mtDNA divergence between individuals was evaluated by means of a Mantel procedure based on 5,000 permutations of the nuclear

D_{ps} values and mtDNA ML distances (averaged substitution model). This procedure was coupled with estimation of the Spearman rank correlation coefficient as the test statistic.

The operational criteria of reproductive isolation (Mayr 1982) and its converse, reproductive cohesiveness (Lambert and Spencer 1995), were evaluated from our genetic data in several ways. Two measures of nuclear population genetic differentiation, F_{ST} and G'_{ST}, were estimated for 7 *S. saevissima* populations to help infer levels of interpopulation isolation or gene flow. The populations were defined by assigning individuals to 1 of 6 STRUCTURE-inferred clusters in which the majority of their nuclear ancestry fell. Two individuals that had no majority cluster ancestry, G82 and Pi87, were assigned to a unique seventh population because they appear to form a distinct lineage in the phylogenetic analyses and because they were collected in proximity to one another at the southern margin of the species range (Fig. 1). A third individual with no majority cluster ancestry, W73, was assigned to the population with which it shared the greatest proportion of its ancestry. Values of F_{ST} were estimated by the method of Weir and Cockerham (1984) using GENEPOP (Raymond and Rousset 1995; http://genepop.curtin.edu.au/); values of G'_{ST}, an estimator of F_{ST} that is standardized by the level of population genetic variation, were estimated by the method of Hedrick (2005) using MSA. Probabilities that G'_{ST} estimates exceeded zero were determined by permuting genotypes across populations 10,000 times (Michalakis and Excoffier 1996) in MSA. Multilocus G'_{ST} estimates involving the other species included in this study (*S. invicta*, *S. richteri*, *S. geminata*) also were obtained in order to gauge genetic divergence between recognized fire ant species.

We estimated levels of genetically effective nuclear gene flow ($N_e m$) between the same 7 *S. saevissima* populations using 2 different but related methods. The first makes use of the formula $N_e m = (1/F_{ST} - 1)/4$ (Slatkin 1987; Whitlock and McCauley 1999), whereas the second is the private alleles method of Barton and Slatkin (1986), as implemented in the online program GENEPOP. Interspecific estimates were obtained in cases where the STRUCTURE results suggested possible hybridization between species.

Recent immigration rates between populations were assessed using the Bayesian nonequilibrium approach of Wilson and Rannala (2003), as implemented in the program BAYESASS. In order to optimize the accepted numbers of proposed changes in the Markov chain, delta values for the allele frequencies, migration rates, and inbreeding values were set at 0.35, 0.33, and 0.17, respectively. The total number of generations in each run was 3,000,000, of which 1,000,000 were burn-in, with sampling every 1,000th generation. Because the program is limited to 35 loci for the input data, a resampling procedure was instituted whereby data from 35 of the 67 nuclear loci were sampled (without replacement) and the program was run; this procedure was repeated 50 times. The estimates of immigration rates, along with their 95% confidence limits, were taken as the means across all the simulations. We note that this method estimates noneffective dispersal, in that the long-term reproductive contributions of immigrant genotypes to the recipient gene pool, if any, are not considered (Broquet et al. 2009). Thus, the relevance of the results with regard to reproductive isolation or cohesiveness is not as clear as for estimates of effective gene flow.

Results of the STRUCTURE analyses indicated a possible clinal change in allele frequencies from south to north along the western segment of the main range of *S. saevissima*, a pattern relevant to the issue of gene flow over this part of the range. We examined this possibility by testing for a pattern of isolation by distance (IBD) in the nuclear composition of 34 individuals from this area, a pattern whereby genetic similarity decreases with geographical separation (Slatkin 1993; Rousset 1997). We analyzed the relationship of the genetic distance D_{ps} with the great circle geographical distance between individuals using GENEPOP. Significance of IBD relationships was determined by means of a Mantel procedure based on 5,000 data permutations coupled with estimation of the Spearman rank correlation coefficient.

Ecological Niche Modeling

Our sample locality data were used in conjunction with ecological niche modeling to evaluate the operational species delimitation criterion of occupation of distinctive niches or adaptive zones (Van Valen 1976). Niche-based distribution models were created by maximum entropy modeling using the program MAXENT (Phillips et al. 2006; Phillips and Dudík 2008). This program, which performs particularly well with small sample sizes (Guisan et al. 2007; Pearson et al. 2007; Wisz et al. 2008), yields projections of potentially suitable habitat using environmental layers and presence-only data (Elith et al. 2006; Phillips et al. 2006; Pearson et al. 2007). Environmental layers were resampled at 2.5-arc-minute resolution, which approximates the estimated error in our georeferencing efforts. The layers represent bioclimatic and landcover variables. The former include temperature and precipitation (from WorldClim's BIOCLIM data set; Hijmans et al. 2005), whereas the latter include current landcover (GLC2000 database; http://www-gem.jrc.it/glc2000/) and reconstructed undisturbed vegetation types (eco_reg layer of Phillips et al. 2006). The relative utility of each variable was assessed using MAXENT's jackknife test and response curves with all the *S. saevissima* locality data. We evaluated models by testing distributions predicted from 75% of the samples with the remaining 25% of samples using 100 replicate bootstrap tests. We then used binomial tests to calculate the statistical significance of the predictions relative to random predictions with the same proportions of predictive and test samples. To avoid overparameterization, we reduced the environmental layers to the following 9 variables for all subsequent analyses: temperature seasonality, annual temperature range, annual precipitation, precipitation in the wettest

quarter, precipitation in the driest quarter, annual cloud cover, annual vapor pressure, elevation, and historical landcover. The convergence threshold (0.00001), maximum number of iterations (500), regularization multiplier (1), and other settings were the default values for the program. We generated logistic model output, which can be regarded as giving the probability of suitable habitat (or of population presence) for a given cell. Distribution models were constructed for 5 of the 7 *S. saevissima* populations (genetic clusters) described above that were sampled at 5 or more localities (e.g., Pearson et al. 2007).

To test for significant differences between estimated niche spaces, we followed Rissler and Apodaca (2007) and Stockman and Bond (2007). We conducted principal components analysis (PCA) on the environmental values for each sample locality and then used scores of the first 4 principal components (which explained 93% of the variance) as dependent variables in a multivariate analysis of variance (MANOVA) to test for significant differences among populations. Tukey's honestly significant difference (HSD) test was used to determine which pairs of populations differed in their habitat environmental attributes. All statistical tests were conducted in the R software environment (http://www.r-project.org/) using the ADE4 package (Chessel et al. 2004).

Cuticular Color Analyses

Solenopsis saevissima is notorious for variation in adult cuticle color among colonies (Trager 1991; Pitts 2002), so this feature generally has not been relied upon in morphological diagnoses of the species. Nonetheless, it is logical to ask whether color variants cluster in accord with other indicators of species boundaries obtained in this study, that is, whether the phenetic cluster operational criterion for delimitation (Michener 1970) may be met with respect to color. Based on visual inspection, we selected a single representative mature adult worker from each sampled *S. saevissima* nest for detailed cuticular hue and brightness analysis; minor workers with potentially atypical coloration (Trager 1991) and young workers with incompletely sclerotized cuticles were avoided. Our study colonies are assumed to comprise full-sibling families in which workers are related to one another by 0.75 (e.g., Ross et al. 1988); thus, within-colony color variation of mature majors likely is small compared with among-colony variation. Each specimen was mounted on heavy white paper with the dorsal surfaces of the head and thorax clearly presented. All were scanned in a single scan in 24-bit RGB mode at a spatial resolution of 1,200 dpi with a Hewlett-Packard Scanjet (Hewlett-Packard, Palo Alto, CA) 5530 flatbed scanner. Images of single representative rectangles of cuticle (150–400 pixels) from the mesothoracic dorsum, the head posterior (vertex), and the head anterior (near the clypeal margin) of each specimen were imported into the program Corel (Corel, Mountain View, CA) PHOTO-PAINT, and the mean luminosity across all pixels of each image was determined in the R, G, and B channels. Luminosity in the B channel generally was very low and so was not considered further. Both the R and the G luminosity values were strongly correlated between the thorax and the head posterior, so means of the values for these 2 regions were used in subsequent analyses.

A matrix of Euclidean distances between each pair of specimens based on the set of 4 individual luminosity values was constructed using MINITAB (Minitab, State College, PA) statistical software. An agglomerative hierarchical method was used to cluster specimens into larger groups of similarly colored ants based on these distances using the same program; the single linkage rule and a cutoff level of 90% similarity were employed to determine the final clustering. PCA using the covariance matrix for the luminosity values also was run in MINITAB to identify features of cuticle color contributing to overall similarity of group members. Four principal components (the number of color measurements) were calculated, but the first 2 accounted for more than 99% of the total variance in the original color data.

RESULTS

Characteristics of the Genetic Markers

Nuclear genotype and mtDNA haplotype data for all study specimens are presented in online Appendix 1. The diversity observed at each marker class is

TABLE 1. Diversity statistics for the genetic markers used in this study

		Nuclear markers					
		Allozymes ($N = 10$)		Microsatellites ($N = 57$)		mtDNA	
		Gene diversity	Allele richness	Gene diversity	Allele richness	Haplotype diversity	Haplotype richness
Solenopsis saevissima	Minimum	0.097	2	0.016	2		
	Maximum	0.647	7	0.963	33		
	Mean	0.338	4.30	0.647	12.05	0.990	48
All species	Minimum	0.113	3	0.043	2		
	Maximum	0.797	13	0.971	45		
	mean	0.361	5.30	0.721	15.72	0.993	66

Notes: Entries are shown separately for the allozyme and microsatellite nuclear markers and for the mtDNA, as well as for nominal *S. saevissima* and for all the species pooled. Minima, maxima, and means across loci are shown for the nuclear markers.

summarized in Table 1. The number of alleles per microsatellite locus spans a remarkable range in both nominal *S. saevissima* and the entire collection of samples, and the allozyme allele richness, although predictably less than that for the microsatellites, also spans an impressive range for this marker type. Most study individuals possessed unique mtDNA variants over the approximately 785-bp fragment sequenced, leading to exceptionally high haplotype diversity and richness. The range and magnitude of variation found at our markers are such that, in aggregate, they are expected to be highly informative with regard to evaluating several operational criteria for species delimitation.

Nuclear Genetic Clusters

Application of the method of Evanno et al. (2005) to our STRUCTURE output revealed that the value of $\Delta K_{\max}$ for 8 clusters exceeded by an order of magnitude the next greatest ΔK value (26.9, 2.70). Moreover, our model with $K = 8$ clusters consistently returned the highest posterior probabilities for our data ($\ln \Pr[X|8] = -15,268$; $\Pr[K = 8] > 0.999$; see Pritchard et al. 2007). The genomic apportionment of each sampled individual in the 8 nuclear clusters is presented in Figure 1, along with the generalized geographical distributions of the clusters. A remarkable feature of the apportionment results is that the 3 nonfocal species, *S. richteri*, *S. invicta*, and *S. geminata*, are each represented almost exclusively by a single cluster, whereas, in contrast, 6 clusters are well represented among nominal *S. saevissima*. Cluster 1 is the dominant component in *S. saevissima* samples from the southwest part of the main range (the range east of Amazonas from which the species is best known), whereas Cluster 2 is dominant in samples from further north in the main range. A third cluster chiefly characterizes the genomes of ants collected around Belo Horizonte, Brazil, whereas Cluster 4, the *S. geminata* cluster, dominates as well among *S. saevissima* from the central Atlantic seaboard. A fifth cluster chiefly characterizes the genomes of *S. saevissima* from the southeast Atlantic seaboard, and a sixth cluster characterizes the genomes of *S. saevissima* from Envira, at the headwaters of the Amazon drainage. Based on the genetic cluster operational criterion, these 6 clusters designate candidate populations for delimitation as cryptic species.

Several other features of the individual genomic apportionments made by STRUCTURE are noteworthy. Only 3 of the 94 study individuals, all nominal *S. saevissima*, did not have a majority apportionment in a single cluster. One of these, W73, had large ancestry components (>33% each) from the 2 clusters common around its collection site (Clusters 1 and 2) as well as a substantial component (24%) from a cluster well represented in a neighboring area (Cluster 3). The other 2 individuals, G82 and Pi87, are quite remarkable in the degree of admixture of their genomes; both have >10% representation of 2 or 3 *S. saevissima* clusters as well as 5–15% representation of both the *S. richteri* and the *S. invicta* clusters. Admixture of other *S. saevissima* individuals attributable to *S. richteri* and *S. invicta* rarely approached 1%. Samples G82 and Pi87 were collected in the highlands at the extreme southern margin of the range of *S. saevissima*, an area near the range margins of the 2 congeners (see Fig. 1).

A plot of each individual's 3D coordinates derived from NMDS analysis performed on nuclear D_{ps} distances is shown in Figure 2. The points representing the 3 nonfocal species, which occupy the margins of the 3D phenetic space, cluster well by species. Even so, the 2 geographical populations of *S. invicta*, themselves candidate cryptic species (Ross et al. 2007), are clearly identifiable. Among nominal *S. saevissima*, the NMDS results indicate the existence of several distinct and well-differentiated genetic clusters. Three of these largely correspond to the regionally segregated Clusters 4, 5, and 6 detected in the STRUCTURE analyses, whereas the fourth consists of the apparently highly admixed southern highlands samples, G82 and Pi87. Samples of *S. saevissima* with majority assignments to STRUCTURE Clusters 1–3 form a group of relatively similar individuals that collectively are quite distinct from the remaining *S. saevissima* samples.

Two-dimensional NMDS analysis revealed similar patterns of overall genetic similarity, with one exception. The samples with predominant ancestry in STRUCTURE Cluster 4 were separated into 3 distinct groups, one containing *S. geminata* and *S. saevissima* sample W22; one containing *S. saevissima* samples W9, W11, W14, and W16; and one containing *S. saevissima* samples W17 and W18 (data not shown). This sundering of the central Atlantic *S. saevissima* samples and close linkage of W22 with the distantly related *S. geminata* is evident to a lesser extent in the 3D analysis (Fig. 2).

Overall, the NMDS results agree with the STRUCTURE results in suggesting that the populations corresponding to Clusters 4–6, as well as the southern highlands samples, are appropriate candidates for species delimitation. On the other hand, the NMDS results do not provide compelling evidence for separating Clusters 1–3 from one another.

Monophyly of Populations

The NJ tree derived from the nuclear interindividual D_{ps} distances is shown in Figure 3 (TreeBASE accession number M4818). According to this tree, nominal *S. saevissima* is paraphyletic with respect to *S. invicta* and *S. richteri*, a result well supported for specimen W22 and moderately so for the 2 southern highlands specimens, G82 and Pi87. Nominal *S. invicta* also appears to be paraphyletic, with 2 well-supported nonsister lineages corresponding to the sampled geographical populations (presumed cryptic species) of this taxon. There is good support for a monophyletic *S. invicta* + *S. richteri*, in keeping with the hypothesis that they are members of an apical clade of the species group (the "socially polymorphic" clade) that excludes *S. saevissima* (Pitts 2002).

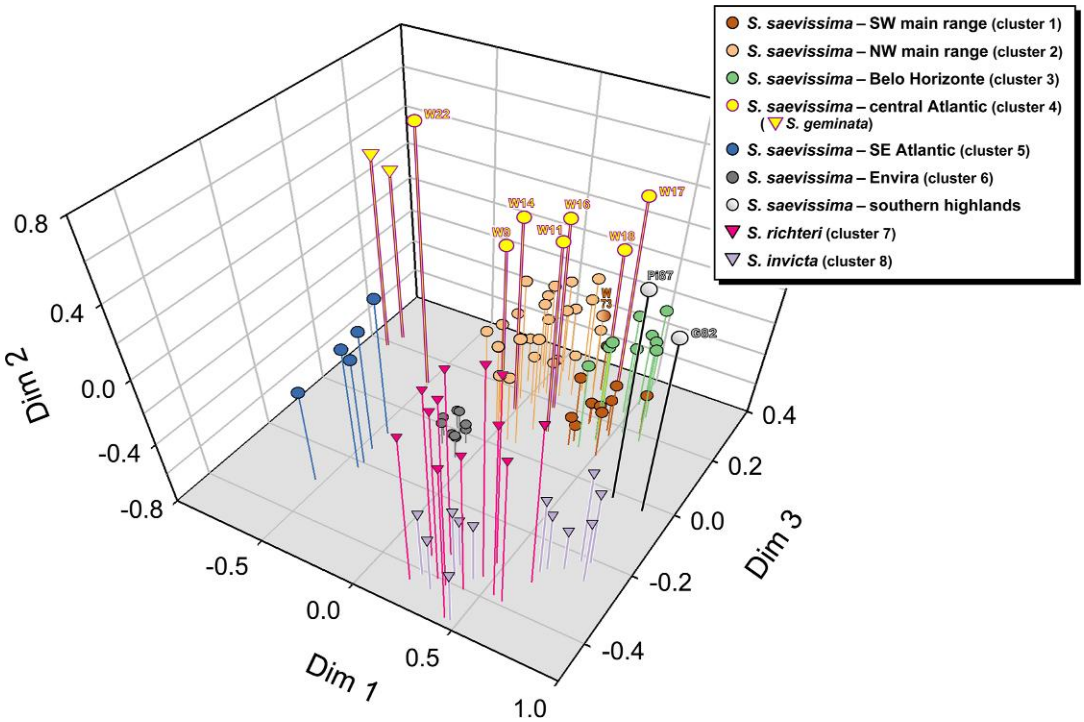

FIGURE 2. Plot of individual 3D coordinates derived from NMDS analysis performed on interindividual nuclear genetic distances (D_{ps}). Colors correspond to distinct genetic clusters detected by STRUCTURE analysis, with each individual assigned to the cluster in which the majority of its nuclear ancestry fell. Three individuals had no majority assignment, the 2 individuals from the southern highlands population (G82 and Pi87) and a single individual with largest ancestry components from the SW and NW main range clusters (W73). These and the central Atlantic *S. saevissima* specimens are identified.

Several well-supported regional clades of *S. saevissima* are recovered in the nuclear NJ tree. Notable among these are the SE Atlantic, Envira, Belo Horizonte, and southern highlands groups, which correspond well with groups identified by the STRUCTURE and/or NMDS analyses. Also, the central Atlantic cluster (minus W22) is marginally supported as a monophyletic group and is divided into well-supported groups of northern and southern samples, paralleling the NMDS results. Finally, a group corresponding to all the samples with majority ancestry in STRUCTURE Cluster 1 (except W70) is recovered but with relatively weak support. The hypothesized monophyly of several nominal *S. saevissima* populations that also correspond to distinct genetic clusters provides additional support for their delimitation as previously unrecognized species.

Several other features of the NJ tree warrant mention. The southern highlands *S. saevissima* appear to be more closely related to *S. invicta* and *S. richteri* than to other *S. saevissima*. Also, a sister relationship is evident between the *S. saevissima* samples nearest the Amazon mouth (W46 and W47) and the Envira samples collected several thousand kilometers distant in the headwaters of the Amazon drainage. A suggestion of this relationship is found as well in the STRUCTURE results (Fig. 1), with the Amazon mouth samples sharing significant ancestry (≈20%) with the Envira cluster. The combined Envira and Amazon mouth group may in turn be most closely related to another group of samples taken near the mouth of this drainage (W44, W48, W49, and W50), reinforcing the idea that historical or recent migration along the length of this enormous basin has left its genomic imprint in these samples.

The relatively divergent sample W22 is placed at the base of the ingroup in the NJ tree. According to this depiction of individual relationships, W22 is quite distantly related to the other central Atlantic samples despite the fact that they all have >63% ancestry in STRUCTURE Cluster 4, the *S. geminata* cluster (Fig. 1). A closer phenetic resemblance of W22 than the other central Atlantic samples to this distantly related species is evident in its higher Cluster 4 ancestry, its closer proximity to *S. geminata* in the NMDS plot (Fig. 2), and the relatively low D_{ps} values between W22 and *S. geminata* (see Fig. 4). The substantial shared ancestry between all the central Atlantic *S. saevissima* and *S. geminata* suggested by STRUCTURE, which is not evident for the other *S. saevissima* samples from areas of sympatry, presumably is the result of one or more bouts of historical hybridization between the species in the central Atlantic region. Large-scale retention of shared ancestral variation is a less likely explanation, given the phylogenetic separation of these taxa into different species groups (Pitts 2002) and the presumably long period since inception of reproductive isolation. The elevated shared ancestry between W22 and *S. geminata* may signal more recent or extensive hybrization between these lineages, which may also explain the position of W22

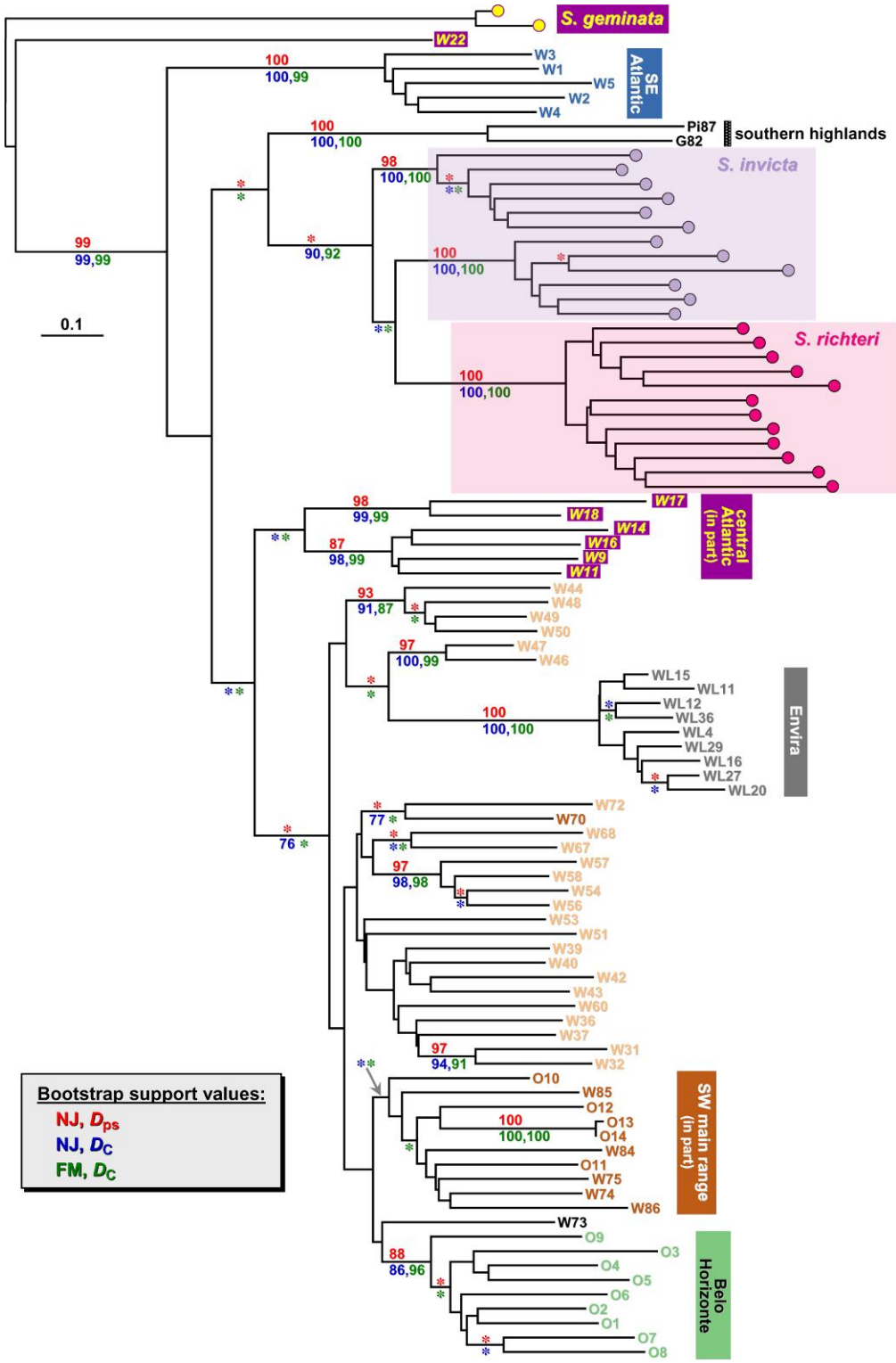

FIGURE 3. NJ tree of individual relationships derived from nuclear genetic distances (D_{ps}). Specimens of nominal *Solenopsis saevissima* are indicated by their codes at the terminals; individuals of the other species are indicated by circles. Colors of the terminals correspond to genetic clusters detected by STRUCTURE (majority assignment; see Fig. 1). Several hypothesized lineages congruent with STRUCTURE and/or NMDS clusters are identified. Bootstrap support values indicate the percentage of times a node was recovered in 1,000 bootstrap trees based on NJ analysis of D_{ps} distances, NJ analysis of D_C distances, or Fitch and Margoliash (FM) analysis of D_C distances (only values >75% reported; asterisks indicate other bootstrap values >50%). The tree was rooted by specifying *S. geminata* as the outgroup.

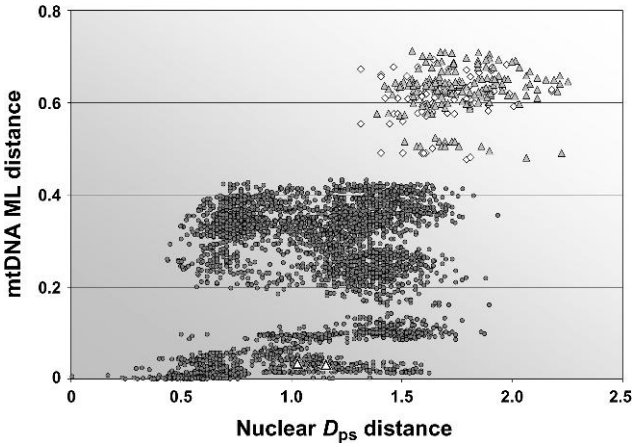

FIGURE 4. Comparison of nuclear distances (D_{ps}) and mtDNA sequence divergence (ML distances under an averaged substitution model) between pairs of individuals included in this study. Pairs comprising *Solenopsis geminata* and other samples are indicated by triangles (white triangles: *S. geminata*/W22 pairs), pairs comprising W22 and other samples are indicated by diamonds, and all other pairs are indicated by circles.

at the base of the ingroup in the NJ tree (e.g., McDade 1992).

Exclusion of the allozyme data in construction of the NJ/D_{ps} tree produced a very similar topology to that shown in Figure 3 but with lower support for several major clades. For example, bootstrap values fell from 100% to 89% for the SE Atlantic clade and from 88% to 76% for the Belo Horizonte clade. This suggests that the allozyme loci contribute some genealogical/phylogenetic signal that generally is congruent with that contained in the much larger microsatellite data set.

The BI tree depicting relationships of the mtDNA sequences from our samples is shown in Figure 5 (TreeBASE accession number M4819). Some major features of the tree are congruent with the nuclear NJ tree, such as the sequence from *S. saevissima* sample W22 lies at the base of the ingroup; the *S. invicta* sequences are paraphyletic; the SE Atlantic sequences (along with that from W16) form a well-supported clade; a clade including sequences from all the samples with majority nuclear ancestry in STRUCTURE Cluster 1 (except W70) is recovered with strong support; the 2 southern highland samples yield sister sequences that are distantly related to all other sequences; the sole sequence found in the Amazon mouth samples (W46 and W47) is sister to the Envira (Amazon headwaters) sequence clade; and sequences from most of the samples with majority membership in STRUCTURE Cluster 3 (Belo Horizonte) form a well-supported clade. On the other hand, there are also fundamental conflicts between the mtDNA and nuclear trees, such as the *S. richteri* sequences are paraphyletic; the sequence from the central Atlantic *S. saevissima* sample W16 falls within the SE Atlantic sequence clade; the sequence from Belo Horizonte samples O7 and O8 is placed in a different major branch of the tree than the clade containing the bulk of sequences from this population; the central Atlantic sequences are distributed across several highly divergent clades (even disregarding W22); and the southern highlands sequences are allied with NW main range and Envira sequences rather than heterospecific sequences from *S. invicta* and *S. richteri*.

One telling feature of the mtDNA phylogeny is the placement of several *S. saevissima* sequences within the major clade containing the *S. invicta* and *S. richteri* sequences (Fig. 5). Although most of the source *S. saevissima* samples originated in the southerly parts of its range close to the ranges of the other species, samples W16, W51, and W67 were collected many hundred kilometers from any current area of sympatry. This result suggests a highly complex evolutionary history of the mtDNA in *S. saevissima*, a recurring conclusion for members of its species group (Ross and Shoemaker 2005; Shoemaker et al. 2006). The discordance with the nuclear tree suggests that the mtDNA relationships do not accurately portray the relationships of the *S. saevissima* populations in many instances, calling into question the general relevance of the mtDNA tree to the monophyly operational criterion for species delimitation.

The mtDNA sequence of *S. saevissima* sample O9 from the Belo Horizonte population falls within a well-supported clade of sequences that otherwise come from the adjacent SW main range population. This sample was collected at the geographic boundary of the 2 nuclear clusters, and its nuclear genome has almost 25% ancestry in the SW main range cluster according to the STRUCTURE analysis (see Fig. 1), far more than any other Belo Horizonte sample. Thus, some gene exchange between the clusters at their contact zone may explain the genetic makeup of this specimen.

The position of the mtDNA sequence from *S. saevissima* sample W22 at the base of the ingroup in the BI tree adds further support to the suggestion from the nuclear data that the lineage including this specimen experienced hybrization with sympatric *S. geminata*. Again, given the presumed phylogenetic distance between the 2 nominal species, the highly similar sequences of *S. geminata* and W22 (see Fig. 4) are unlikely to have persisted since the time of lineage separation (i.e., to represent ancestral polymorphism or haplotypes that diverged from a common ancestral sequence following speciation). Instead, the W22 sequence may represent a slightly modified variant of a secondarily acquired *S. geminata* haplotype.

A comparison of the nuclear and mtDNA divergence between all pairs of individuals in this study is depicted in Figure 4. No correlation exists between the nuclear and mtDNA distances, regardless of whether or not *S. geminata* and sample W22 are included in the data set (Mantel tests; both $P > 0.99$). This lack of association between individual nuclear and mitochondrial distances parallels the substantial incongruence between the trees derived from the different genomes.

Gene Flow and Reproductive Isolation

Information relevant to the operational criteria of gene flow (reproductive cohesiveness) or its converse,

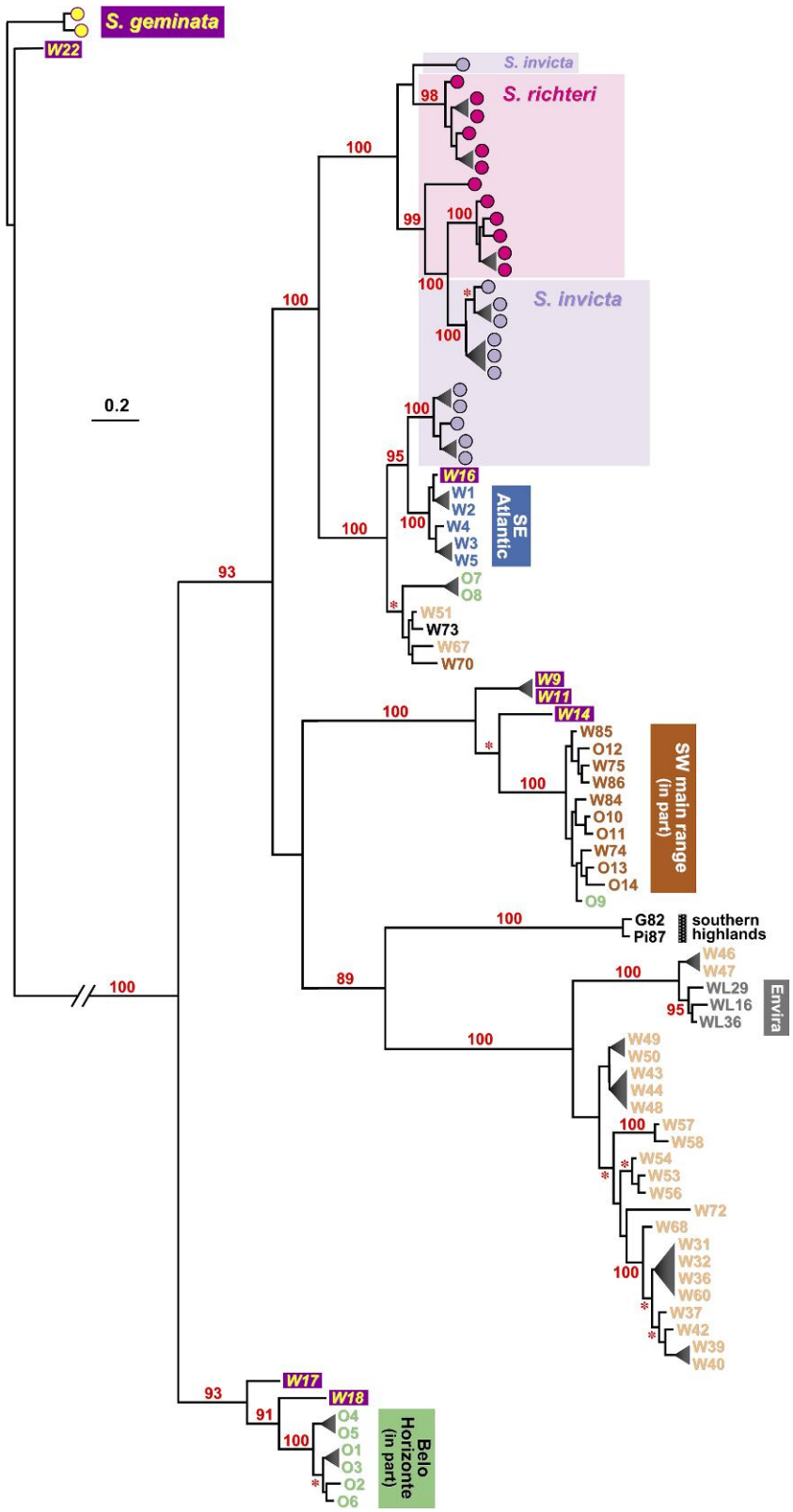

FIGURE 5. BI tree of phylogenetic relationships of mtDNA sequences from fire ants included in this study (triangles link identical sequences). Specimens of nominal *Solenopsis saevissima* are indicated by their codes at the terminals; individuals of the other species are indicated by circles. Colors of the terminals correspond to nuclear genetic clusters detected by STRUCTURE. Several hypothesized lineages largely congruent with STRUCTURE and/or NMDS clusters are identified. Posterior probability values are shown for each node (only values >75% reported; asterisks indicate other posterior probability values >50%). The tree was rooted by specifying the *S. geminata* sequences as the outgroup.

reproductive isolation, was extracted in several ways from our genetic data. First, the STRUCTURE results suggest a general clinal change in allele frequencies from south to north along our western transect in the main range of *S. saevissima*, signified by a gradual transition from Cluster 1 to Cluster 2 (Fig. 1). A test for IBD in the nuclear composition of ants from this area confirmed that genetic similarity decreases with geographical separation along this transect (Mantel test; $P < 0.0001$), supporting the existence of the proposed genetic cline and implicating spatially restricted but significant gene flow between the SW and the NW main range populations as the cause.

In a more general analysis of gene flow and reproductive isolation, we estimated 2 standard measures of nuclear population genetic differentiation, F_{ST} and G'_{ST}. As expected, given the differences in diversity between the 2 classes of nuclear loci, single-locus values of F_{ST} across all 7 *S. saevissima* populations generally are greater for the allozymes than for the microsatellites (means: 0.573 and 0.233; ranges: 0.333–0.921 and 0.001–0.786 for the respective marker types). On the other hand, values of G'_{ST}, a population differentiation measure standardized by heterozygosity level, are distributed similarly between the 2 marker classes (means: 0.608 and 0.605; ranges: 0.186–0.942 and 0.001–0.954). Importantly, comparisons of pairwise population F_{ST} and G'_{ST} values derived separately from the allozymes and microsatellites for *S. saevissima* and the other species show that the 2 marker classes display parallel patterns of population genetic differentiation (F_{ST}: Mantel test, $P = 0.009$; G'_{ST}: Mantel test, $P = 0.032$). This result reinforces our conclusion from the phylogenetic analyses that the 2 types of nuclear markers contribute congruent signal pertaining to individual and population relationships.

Estimates of G'_{ST} for each pair of nominal *S. saevissima* populations based on all 67 nuclear loci are presented in Table 2. The populations display from about one-quarter to two-thirds of the maximum divergence possible given their observed heterozygosity (Hedrick 2005). These estimates of differentiation generally are statistically significant based on a permutation test, confirming that the populations comprise genetically distinct clusters with restricted gene flow between them (the nonsignificant estimates involve the southern highlands population, for which sample size is smallest). Distributions of G'_{ST} values are compared between pairs of *S. saevissima* populations and pairs of heterospecific fire ant populations in Figure 6. Values of G'_{ST} between *S. geminata* and species of the *S. saevissima* species group typically exceed 0.70, whereas those for heterospecific populations of *S. saevissima*-group species, whether allopatric or sympatric, generally fall in the range 0.45–0.65. Among the nominal *S. saevissima* pairs, only 6 G'_{ST} estimates (29%) fall below 0.45, all of which involve the SW or NW main range populations (see Table 2). These comparisons reveal that nuclear genetic differentiation of regional populations of nominal *S. saevissima* commonly rivals or exceeds the differentiation observed between recognized species in the group that are known or assumed to be fully reproductively isolated (e.g., Ross and Shoemaker 2005).

Effective rates of nuclear gene flow ($N_e m$) across all 7 *S. saevissima* populations are estimated at 0.748 (from F_{ST}) and 0.674 (private alleles method). Distributions of $N_e m$ values obtained by the 2 methods are shown for each pair of these conspecific populations as well as for pairs of heterospecific populations of special interest in Figure 7. Estimates of the pairwise $N_e m$ values obtained by the different methods are highly correlated (Mantel test, $P = 0.0002$), but the values generated using the private alleles method consistently are somewhat lower than those calculated from F_{ST}. All but 3 of the conspecific *S. saevissima* pairs (86%) yield $N_e m$ values less than 1.0 by both methods (see also Table 2). Gene flow estimates between the SW main range, NW main range, and Belo Horizonte populations exceed 1.0, as determined by both estimation methods in the case of the former 2 populations; evidence for some admixture among all 3 of these populations is apparent also in the individual ancestry assignments made by STRUCTURE (Fig. 1), especially in the clinal transition between the 2 main range clusters. Eight of the *S. saevissima* population pairs (38%) yield $N_e m$ values less than 0.5 by both methods; all involve the SE Atlantic or southern highlands populations. Values of $N_e m$ in or below the range 0.5–1.0 are viewed with special interest for species delimitation because the range corresponds to a general upper limit at which gene flow can counteract drift to prevent population differentiation at neutral genes in a simple island model (Slatkin 1987; Porter 1990; Hedrick 2000; Shaffer and Thomson 2007). Thus, most pairs of nominal *S. saevissima* populations may be effectively fully reproductively isolated even at genes not experiencing divergent selection (e.g., Petit and Excoffier 2009).

Values of $N_e m$ between the *S. saevissima* central Atlantic population and *S. geminata*, as well as between the *S. saevissima* southern highlands population and *S. invicta* or *S. richteri*, consistently are close to or below 0.5 (Fig. 7). Thus, despite the more or less substantial ancestry these *S. saevissima* populations share with the other species according to the STRUCTURE and other analyses, recent introgression would appear to have been minimal.

Recent noneffective migration into populations detectable in the nuclear genome was quantified using BAYESASS. Based on this approach, nonimmigrant genotypes typically make up around 80% of each *S. saevissima* population (Table 3). In only 4 cases did migration from a specific source exceed 4%; these involved immigration between the SW and the NW main range populations and immigration into these from the Belo Horizonte population. Results from our other analyses indicate that the elevated migration in these cases leads to elevated effective gene flow rates (see Table 2 and Fig. 1). In no case were *S. geminata*, *S. invicta*, or *S. richteri* estimated to contribute more than 3.2% to the migrant composition of any *S. saevissima* population, and this low rate evidently rarely results in contemporary interspecific introgression (Fig. 7).

TABLE 2. Estimates of nuclear genetic differentiation (G'_{ST}, above diagonal) and effective rates of nuclear gene flow ($N_e m$, below diagonal) between *Solenopsis saevissima* study populations

	Southern highlands	SW main range	NW main range	Belo Horizonte	Central Atlantic	SE Atlantic	Envira
Southern highlands	—	0.527	0.465*	0.552	0.674	0.625	0.573
SW main range	0.472 0.342	—	0.242**	0.322**	0.428**	0.630**	0.486**
NW main range	0.534 0.347	**1.808** **1.235**	—	0.344**	0.258**	0.508**	0.439**
Belo Horizonte	0.369 0.280	**1.172** 0.920	**1.245** 0.866	—	0.467**	0.656**	0.517**
Central Atlantic	0.918 0.448	0.823 0.527	0.971 0.608	0.772 0.547	—	0.615*	0.553**
SE Atlantic	0.487 0.319	0.414 0.327	0.487 0.343	0.368 0.303	0.732 0.469	—	0.581**
Envira	0.314 0.177	0.681 0.464	0.915 0.734	0.540 0.365	0.653 0.342	0.454 0.249	—

Notes: Values of $N_e m$ were estimated by the private alleles method (shaded entries) or from F_{ST}. Values of $N_e m > 1.0$ are highlighted in bold. G'_{ST} values that significantly exceed zero (after Bonferroni correction) are indicated by *($P < 0.05$) or ** ($P < 0.01$).

Ecological Niche-Based Distribution Models

Ecological niche modeling was conducted to evaluate the operational species delimitation criterion of occupation of distinct niches. Generally, there seems to be little overprediction from our models, with most samples collected in areas of high probability of habitat suitability ($P > 0.4$). The lowest probabilities of suitability were found for 2 of the southernmost samples in the SW main range population ($P_{W85} = 0.18$; $P_{W86} = 0.20$).

Distributions of the ecological niches predicted for the 5 best sampled *S. saevissima* populations are mapped in Figure 8. Scores of the first 4 principal components derived from the environmental values differ significantly across all populations (Wilk's lambda = 0.031; $F_{6,150.34} = 19.83$; $P < 0.0001$), signaling meaningful overall niche space differentiation. Pairwise comparisons reveal that the predicted niches for the 3 southern populations, SW

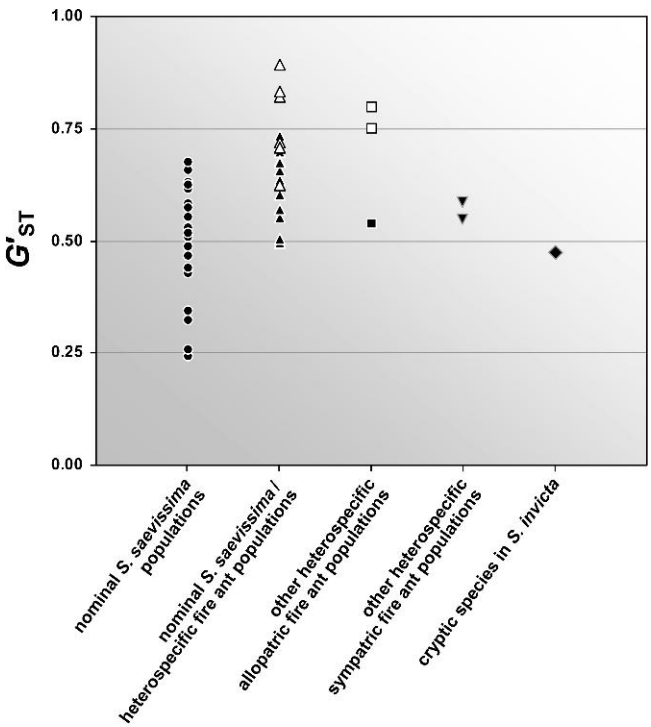

FIGURE 6. Values of G'_{ST} estimated between paired populations of nominal *Solenopsis saevissima*, between populations of nominal *S. saevissima* and heterospecific fire ants, and between other heterospecific fire ant populations (*Solenopsis geminata*, *Solenopsis invicta*, and *Solenopsis richteri*). Symbols for heterospecific pairs are white if *S. geminata*, the only study species not a member of the *S. saevissima* species group, is included.

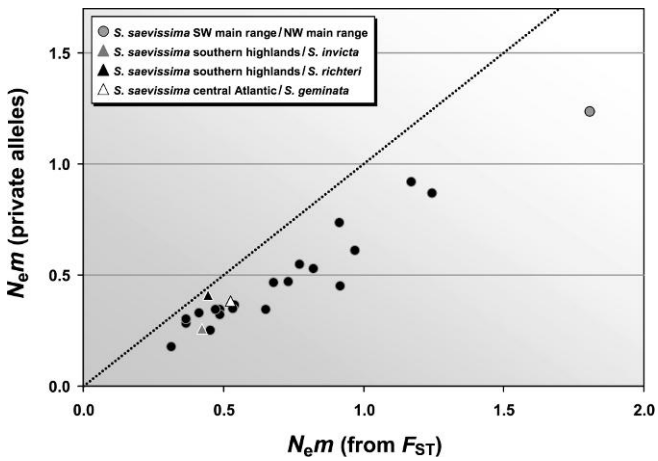

FIGURE 7. Levels of effective nuclear gene flow ($N_e m$) estimated between pairs of populations by 2 methods. Circles represent estimates for pairs of *Solenopsis saevissima* populations, whereas triangles represent estimates between selected *S. saevissima* and heterospecific populations. A legend to symbols for $N_e m$ estimates of special interest is shown. The dotted line indicates equality of estimates from the 2 methods.

TABLE 3. Rates of recent immigration into *Solenopsis saevissima* study populations inferred from the program BAYESASS

Recipient population	Source population									
	S. saevissima							*Solenopsis richteri*	*Solenopsis invicta*	*Solenopsis geminata*
	Southern highlands	SW main range	NW main range	Belo Horizonte	Central Atlantic	SE Atlantic	Envira			
S. saevissima										
Southern highlands	0.786 (0.688–0.922)	0.024 (0.000–0.127)	0.023 (0.000–0.126)	0.021 (0.000–0.120)	0.023 (0.000–0.127)	0.029 (0.001–0.138)	0.026 (0.000–0.133)	0.024 (0.000–0.126)	0.019 (0.000–0.118)	0.024 (0.000–0.128)
SW main range	0.017 (0.006–0.056)	0.751 (0.717–0.808)	**0.041** (0.020–0.086)	**0.063** (0.033–0.110)	0.017 (0.006–0.057)	0.021 (0.008–0.063)	0.027 (0.011–0.070)	0.007 (0.000–0.047)	0.031 (0.014–0.073)	0.025 (0.011–0.067)
NW main range	0.015 (0.009–0.037)	**0.044** (0.032–0.067)	0.757 (0.737–0.787)	**0.077** (0.056–0.104)	0.009 (0.005–0.030)	0.015 (0.009–0.036)	0.024 (0.014–0.047)	0.005 (0.000–0.028)	0.032 (0.022–0.053)	0.021 (0.013–0.044)
Belo Horizonte	0.016 (0.005–0.059)	0.029 (0.012–0.074)	0.029 (0.012–0.074)	0.805 (0.757–0.861)	0.022 (0.008–0.068)	0.019 (0.005–0.064)	0.024 (0.009–0.068)	0.012 (0.002–0.053)	0.029 (0.012–0.074)	0.016 (0.005–0.058)
Central Atlantic	0.023 (0.005–0.080)	0.023 (0.005–0.078)	0.038 (0.010–0.101)	0.022 (0.004–0.078)	0.806 (0.744–0.880)	0.023 (0.005–0.078)	0.015 (0.002–0.068)	0.011 (0.001–0.062)	0.021 (0.004–0.077)	0.018 (0.003–0.073)
SE Atlantic	0.030 (0.006–0.098)	0.024 (0.004–0.090)	0.017 (0.002–0.081)	0.017 (0.002–0.082)	0.030 (0.006–0.100)	0.791 (0.725–0.876)	0.015 (0.001–0.079)	0.024 (0.004–0.091)	0.024 (0.004–0.089)	0.027 (0.005–0.094)
Envira	0.030 (0.012–0.077)	0.038 (0.017–0.086)	0.034 (0.014–0.081)	0.039 (0.017–0.091)	0.025 (0.010–0.071)	0.012 (0.002–0.057)	0.773 (0.728–0.835)	0.017 (0.005–0.061)	0.016 (0.005–0.061)	0.016 (0.005–0.060)

Notes: Entries for immigration rates greater than 0.04 are indicated in bold. Entries for proportions of nonimmigrant genotypes are shaded. The 95% confidence limits for the estimates are in parentheses.

main range, Belo Horizonte, and SE Atlantic, do not differ significantly at any of the first 4 principal components (Tukey's HSD test, PC1: $F_{4,52} = 17.61$; P = 0.856–0.997; PC2: $F_{4,52} = 5.29$; P = 0.586–0.996; PC3: $F_{4,52} = 15.57$; P = 0.088–0.957; PC4: $F_{4,52} = 3.77$; P = 0.936–1.000), but all other population pairs differ significantly in their scores at 1 or 2 principal components (Tukey's HSD test, all $P < 0.015$). The environmental features at the Envira (western Amazonas) locality are not encompassed in the predicted niche envelope of any of the modeled populations, whereas features at the southern highlands localities are well within the niche envelope of the SE Atlantic population ($P_{G82} = 0.66$; $P_{Pi87} = 0.31$; see Fig. 8). Results of the niche modeling thus support the candidacy of several of the regional genetic clusters as separate newly delimited species, including the 2 main range populations.

Cuticular Color

Cuticular color analysis was conducted to learn whether color variants within nominal *S. saevissima* cluster in accord with other indicators of species boundaries, that is, whether the phenetic cluster operational criterion for delimitation is met with respect to this trait. Four major clusters (color morphs) were recovered from the hierarchical cluster analysis using Euclidean distances derived from thorax and head color, although 2 specimens (W39 and W49) were not placed in any of the groups. PCA revealed that 2 components together explain 99.2% of the total data variance. PC1, which explains 87.6% of the variance, represents the overall R and G luminosity of the ants, whereas PC2, which explains 11.6 % of the variance, contrasts the R and G luminosity of the head anterior with that of the head posterior and thorax. The plot of these 2 principal component scores reveals the 4 groups found with the cluster analysis (Fig. 9). Three of the morphs separate completely along PC1, with a large cluster of uniformly dark brown specimens from localities confined almost entirely to the southern part of the main range, a large cluster of uniformly orange-brown specimens confined almost entirely to the northern main range, and a cluster of yellow-brown specimens from the western Amazonas locality (see Fig. 1). The distributional patterns for these morphs coincide generally with those reported previously (Trager 1991; Pitts 2002). The fourth major cluster, which separates well from the

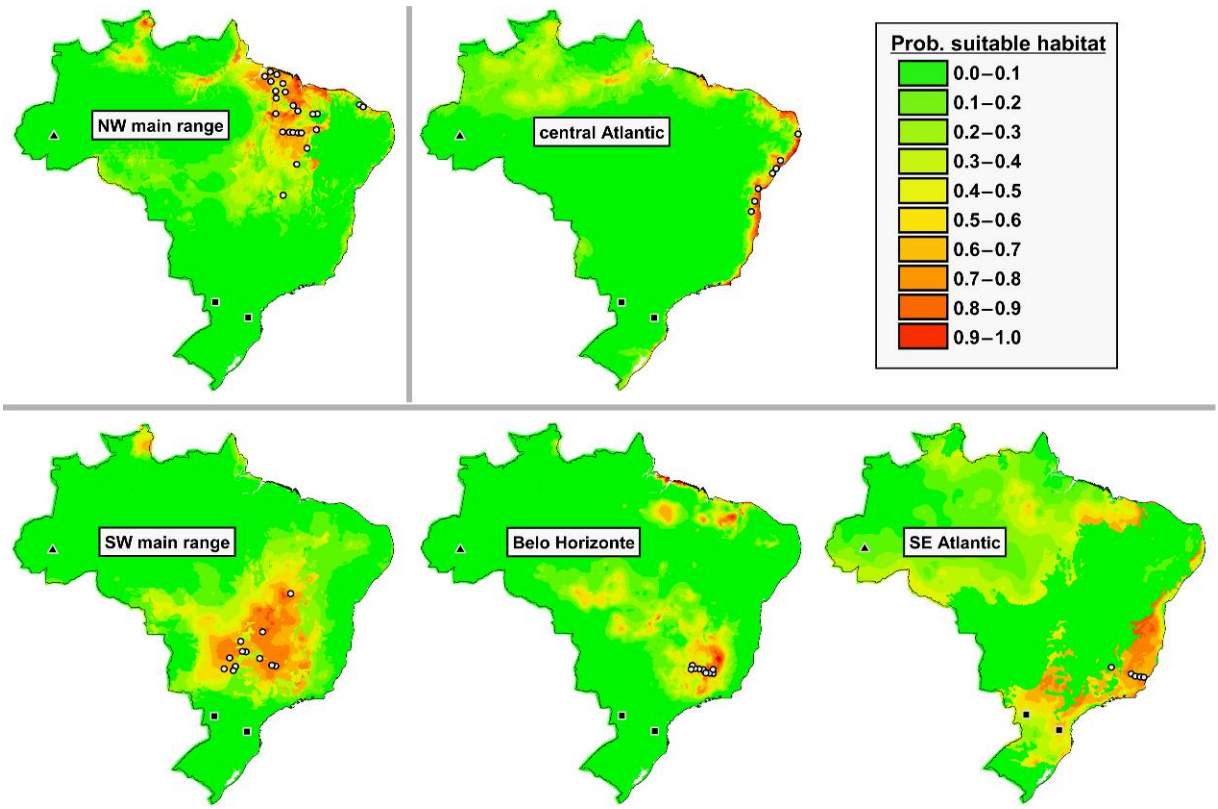

FIGURE 8. Predicted distributions of suitable ecological niches for 5 genetically distinct *Solenopsis saevissima* populations based on 9 environmental variables. Collection localities are indicated with white circles for samples with majority assignment to the listed genetic cluster based on STRUCTURE analyses. Collection localities for samples from Envira and the SE Atlantic population are indicated with a black triangle and black squares, respectively. Probabilities of suitable habitat derived from MAXENT are shown as heat maps, with warmer colors representing higher associated probabilities (typical collection localities have probability ≈0.5 of occurring in suitable habitat; Phillips et al. 2006). Grey lines demarcate predicted niche envelopes that differ statistically from one another using MANOVA on the scores of the first 4 principal components derived from PCA.

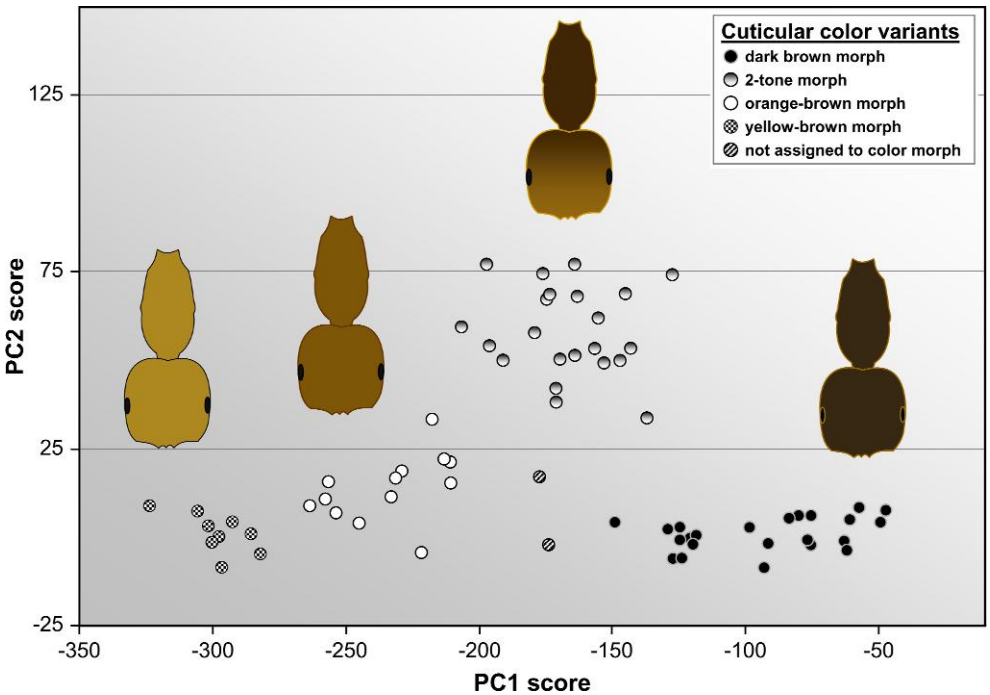

FIGURE 9. Plot of scores of individual *Solenopsis saevissima* workers at the first 2 principal components resulting from PCA of cuticular color. Points corresponding to 4 distinct morphs revealed by hierarchical cluster analysis are indicated by differently shaded circles (2 specimens did not fall into one of these clusters). Ant worker icons depicting the morph cuticular colors are reproductions based on the average measured R, G, and B luminosities of the thorax, head posterior, and head anterior (see Materials and Methods).

others along both PC1 and PC2, comprises a morph that combines overall dorsal coloration intermediate between the dark brown and the orange-brown forms with a much paler head anterior (2-tone morph; see also Trager 1991). These latter specimens were found at localities scattered throughout the main part of the range (Fig. 1). The 4 cuticular color morphs are highly significantly differentiated from one another by their PC1 scores (Kruskal–Wallis test, $P < 0.001$), with each pair of clusters also significantly differentiated (Dunn's multiple comparisons test, all $P < 0.01$).

The *S. saevissima* color morphs are not distributed randomly among the 7 genetically delineated populations (Fisher's exact test, $P < 0.0001$), a finding perhaps to be anticipated given the strong geographic structure evident in both the morphological and the genetic data sets. Nonetheless, whereas the Envira, SE Atlantic, and southern highlands populations contain only a single color variant, the remaining populations are polymorphic (Fig. 1). In fact, the central Atlantic cluster has all 3 morphs found east of Amazonas. Because cuticular coloration generally is not predictive of genetic affinity in the main range, its diagnostic taxonomic utility is likely to be negligible for most specimens of nominal *S. saevissima*.

DISCUSSION

The *S. saevissima* species group comprises a taxonomically challenging assemblage of recently diverged fire ant species that are of considerable economic and scientific importance. Among the nominal species, *S. saevissima* is especially noteworthy in terms of the size and ecological breadth of its natural range and the magnitude of phenotypic variability observed across this range. The goal of the current study was to employ data from a large and diverse set of genetic markers, along with morphological and distributional/bioclimatic information, to hypothesize the existence of evolutionarily independent lineages (cryptic species) within this widespread and variable taxon. We adopted an integrated approach in which we sequentially tested whether populations meet various operational criteria (contingent properties) for delimitation as cryptic species. One advantage of such an approach is that congruence of different independent operational criteria lends confidence to the conclusions drawn regarding species boundaries (de Queiroz 2005, 2007; Shaffer and Thomson 2007; Wiens 2007b; Leaché et al. 2009; Schlick-Steiner et al. 2010).

Evidence for Cryptic Species within Nominal S. saevissima

We formulated initial hypotheses of species boundaries in nominal *S. saevissima* by inferring the presence of distinct genetic clusters represented among our sampled individuals. STRUCTURE analyses of the nuclear genotypic data revealed several geographically localized clusters. Despite the common occurrence of individuals with some mixed ancestry, most individuals have a great preponderance of their ancestry in a single

cluster, supporting the existence of 6 discrete genetic groups (Fig. 1). The utility of these STRUCTURE analyses is augmented by inclusion of samples from other fire ant species that display pronounced regional genetic structure. Specifically, conspecific samples of both *S. invicta* and *S. richteri* were relegated to single unique clusters despite their extensive geographical divergence (Ross and Shoemaker 2005), a result attributable to the tendency of STRUCTURE analyses using the ΔK statistic to detect the highest hierarchical level of structure (Evanno et al. 2005). The discovery of regional genetic clusters within nominal *S. saevissima* despite this tendency constitutes strong initial evidence for the hypothesis of 6 morphologically cryptic, regionally localized species within the taxon.

A complementary approach employed to evaluate the presence, distinctiveness, and phenetic relationships of the genetic clusters was NMDS. Results of this analysis confirmed the distinctiveness and clear phenetic separation of 3 of the regional genetic groups detected by STRUCTURE, the central Atlantic, SE Atlantic, and Envira populations (Fig. 2), whereas only modest separation was revealed among the remaining 3, the SW main range, NW main range, and Belo Horizonte populations. The NMDS analysis also revealed a seventh distinct cluster consisting of the southern highlands samples. Finally, the NMDS results provided a first indication that some of the regional *S. saevissima* populations are as genetically divergent from one another as are heterospecific fire ant populations even, in some cases, when they are not separated by great geographical distances.

The issue of monophyly of nominal *S. saevissima* as well as its consituent genetic clusters was examined by constructing phylogenetic trees based on interindividual genetic distances derived from the 67 nuclear markers. The apparent paraphyly of *S. saevissima* with respect to *S. invicta* and *S. richteri*, attributable to 2 southern populations and a single sample from the central Atlantic cluster (Fig. 3), suggests that study populations of the focal species are not part of a single descendant lineage that does not include other nominal species; thus, nominal *S. saevissima* does not meet the monophyly criterion for recognition as a single species (Donoghue, 1985). On the other hand, the phylogenetic analyses yielded strong support for several monophyletic lineages within *S. saevissima* that largely correspond to the regional genetic clusters. The 2 southern clades that render *S. saevissima* paraphyletic (SE Atlantic and southern highlands) as well as 2 other clades within the monophyletic core of nominal *S. saevissima* (central Atlantic [minus W22] and Envira) are moderately or well supported and are relatively divergent from all other samples, a fact evident also from their marginal locations in phenetic space in the NMDS analyses. Samples with majority assignment to the Belo Horizonte cluster also comprise a well-supported monophyletic lineage, even though this population does not appear especially divergent from the SW and NW main range populations in the NMDS analyses. Even the SW main range population (with the exception of sample W70) is hypothesized to be monophyletic, despite its evident admixture with, and lack of strong divergence from, the NW main range population. On the basis of the monophyly criterion, all 7 regional genetic clusters, with the possible exception of the 2 main range populations, are candidate cryptic species.

The considerable discordance of our mtDNA haplotype phylogeny with the nuclear relationships inferred for the study individuals, and its relevance to the monophyly operational criterion, warrants comment. Widespread paraphyly and polyphyly of mtDNA sequences with respect to nominal species boundaries have been detected in previous studies of South American fire ants (Ross and Shoemaker 2005; Shoemaker et al. 2006). Some instances can be attributed to the existence of cryptic species, but other causes of mismatches between mtDNA variation and species boundaries, or between mtDNA and nuclear DNA variation, also are likely (Funk and Omland 2003; Zink and Barrowclough 2008). The mtDNA is a single cytoplasmically inherited marker, and its distribution may not reflect patterns of ancestry of the majority of nuclear genes because of the different modes of inheritance of the 2 genomes and the idiosyncratic nature of sorting of ancestral variation at any single gene (Brito and Edwards 2009; Petit and Excoffier 2009). Given the evidence for historical hybridization in fire ants, interlineage mtDNA capture and fixation in the absence of extensive nuclear introgression may also occur with some regularity (see also Funk and Omland 2003; Seifert and Goropashnaya 2004; Chan and Levin 2005; Currat et al. 2008; Feldhaar et al. 2008; Seifert 2009). Such cytoplasmic genome capture may be facilitated by selective sweeps associated with infection by the cytoplasmic endosymbiont *Wolbachia* (Raychoudhury et al. 2009), different strains of which are present at variable frequencies in many native fire ant populations (Shoemaker et al. 2003). That *Wolbachia* infection may affect the distribution of mtDNA variants in *S. saevissima* and create some of the observed mtDNA/nuclear discordance is indicated by the fact that several unexpected haplotype relationships reflect parallel infection status (data not shown). For example, the fragmenting of the central Atlantic haplotypes across several highly divergent mtDNA clades reflects shared possession of specific *Wolbachia* strains: central Atlantic sample W16 shares an identical *Wolbachia* strain with SE Atlantic samples W1 and W2, whereas central Atlantic samples W9, W11, and W14 share an identical *Wolbachia* strain with most of the SW main range samples (see Fig. 5). Many other instances of parallel transmission of the 2 cytoplasmic genomes may be obscured by the apparent frequent loss of *Wolbachia* infection or replacement of infection strain (Shoemaker et al. 2003).

Because our nuclear DNA trees are based on a large number of markers and make intuitive biogeographical sense (samples that are geographically adjacent or groups linked by conduits for gene flow [i.e., the Amazon River] show affinity to one another), we believe that they represent population relationships more

faithfully than the mtDNA tree (see also Zink and Barrowclough 2008). Thus, the relevance of mtDNA trees for *S. saevissima* and related fire ants to the monophyly operational criterion for species delimitation generally is questionable. Even so, the *S. saevissima* mtDNA haplotype phylogeny did yield some support for the existence of the southern highlands, SE Atlantic, Belo Horizonte, and Envira clades recognized in the nuclear analyses.

We addressed the important operational criteria of reproductive isolation and reproductive cohesiveness (gene flow) in several ways. One involved calculation of standardized measures of genetic differentiation (structure) for the genetic clusters within *S. saevissima* as well as for heterospecific populations known to be reproductively isolated. Estimates of G'_{ST} for pairs of *S. saevissima* populations commonly rivaled or exceeded estimates for pairs of recognized species in the *S. saevissima* species group (Fig. 6), with the exceptions generally involving the SW main range, NW main range, and Belo Horizonte populations. Such comparisons are considered to be especially relevant to the species status of allopatric populations, for which intrinsic reproductive isolation normally cannot be directly evaluated (Price 2008). In absolute terms, divergence between most of the surveyed *S. saevissima* populations exceeded 50% of the maximal divergence possible given the levels of within-population heterozygosity at our markers (Table 2). While we do not advocate the use of particular divergence benchmarks as a criterion for species delimitation (cf. Lefébure et al. 2006; Richards et al. 2009), the fact that nuclear genomic differentiation of the southern highlands, SE Atlantic, central Atlantic, and Envira *S. saevissima* populations is comparable in extent that observed between fully reproductively isolated species in the species group may be taken as support for delimiting them as separate species (see Helbig et al. 2002; Cardoso et al. 2009).

It is noteworthy that the genetic discontinuities detected between the SW main range, Belo Horizonte, and SE Atlantic groups revealed by several of our methods overlay what are parapatric or even effectively sympatric distributions. Some samples assigned to the former 2 populations were collected within 50 kilometers of one another, whereas sample W1 of the SE Atlantic population actually was collected well within the distribution of the Belo Horizonte cluster. Fire ant sexuals are capable of dispersing up to 10 km on wind currents during their mating flights (Markin et al. 1971; Ross and Keller 1995), suggesting considerable potential for gene flow over several generations among any of these populations. While the observed genetic differentiation of other populations may be attributable in some measure to IBD over huge tracts of habitable area (SW main range to NW main range), to complete geographic isolation due to gaps in suitable fire ant habitat (central Atlantic population), or to incomplete sampling (Envira), observed genetic disjunctions between the parapatric and sympatric pairs imply the existence of intrinsic reproductive isolating barriers.

We further examined reproductive isolation and cohesiveness by using population genetic models to estimate effective rates of interpopulation gene flow ($N_e m$). Again with the exception of gene flow between the SW main range, NW main range, and Belo Horizonte populations, estimates of $N_e m$ fall within or below the range at which gene flow theoretically becomes too weak to prevent divergence by drift alone (0.5–1.0). Thus, from a theoretical population genetics perspective, at least the southern highlands, SE Atlantic, central Atlantic, and Envira populations are predicted to be effectively reproductively isolated from nominal conspecific populations, even were diversifying selection related to habitat specialization, mate choice, or hybrid fitness absent.

That ecologically mediated clade-specific selection may actually enhance that the effect of drift in driving the continued divergence of regional *S. saevissima* populations receives some support from our ecological niche modeling. Although 3 populations from the southern half of the main range (SW main range, Belo Horizonte, and SE Atlantic) occupy predicted niches that are statistically indistinguishable on the basis of the environmental variables we used, the NW main range and central Atlantic niches differ statistically from one another and from all other predicted niches (Fig. 8), thus fulfulling the ecological delimitation criterion of occupation of separate adaptive zones. Moreover, the niche envelopes of the southern populations are only partially overlapping, suggesting incipient or incomplete ecological separation among them. Directional selection involving specialization to different habitats may directly (via pleiotropy) or indirectly (via linkage) promote reproductive isolation among the ecologically distinct *S. saevissima* populations (Price 2008; Nosil et al. 2009; Schluter 2009), thus contributing to their continued evolutionary divergence.

Finally, our data are relevant to 2 other operational criteria for delimiting species. We found fixed diagnostic alleles that characterize the southern highlands (3 loci) and SE Atlantic (1 locus) populations. Also, private alleles present at high frequency in just a single population were found in the SE Atlantic (1 locus) and Envira (1 locus) populations. Diagnosability of these regional populations with genetic markers means that they meet this criterion for delimitation as morphologically cryptic species (see de Queiroz 2007; Xie et al. 2008). On the other hand, we found that the notorious variation in adult cuticular color patterns observed in *S. saevissima* is of very limited use in diagnosing the distinct genetic clusters (see Fig. 1); thus, as perhaps expected given the problematic taxonomic history of the group, delimitation within nominal *S. saevissima* is not aided under the diagnosability and phenetic criteria with respect to this most obvious of morphological traits.

Hybridization and the Origin of Diversity in Nominal S. saevissima

Introgression from sympatric heterospecific populations is suggested for the southern highlands and central

Atlantic populations of *S. saevissima* by the results of our STRUCTURE and phylogenetic analyses. Most surprising is the apparent introgression of genes from *S. geminata*, a quite distantly related fire ant that is not a member of the *S. saevissima* species group. If such introgression indeed occurred, then the degree of interspecific divergence across genes should display relatively high variance (Hey 2006a). The reason is that different elements of the *S. geminata* genome would be more or less free to introgress following hybridization, depending on their linkage relationships with genes involved in reproductive barriers or hybrid unfitness. Comparison of pairwise single-gene F_{ST} estimates across all nuclear loci for the divergence between *S. geminata* and 9 heterospecific populations (7 *S. saevissima*, *S. invicta*, and *S. richteri*) indeed confirmed that the central Atlantic *S. saevissima*/*S. geminata* pair exhibits the highest coefficient of variation for these estimates (0.924 vs. 0.470–0.892), yielding additional support for the idea of hybridization between these divergent taxa.

Hybridization between well-differentiated lineages evidently has occurred with some regularity during the evolutionary diversification of South American fire ants (Ross and Shoemaker 2005; Shoemaker et al. 2006), and interspecific hybridization in other ants is being increasingly documented (Seifert and Goropashnaya 2004; Feldhaar et al. 2008). Relative uniformity in male phenotypes and habits across species due to a general relaxation of sexual selection is proposed to be one proximate factor contributing to ant hybridization (Feldhaar et al. 2008). Male phenotypic characters, including the genitalic characters so useful for taxonomic diagnoses in other insect groups, are indeed remarkably uniform across the *S. saevissima* species group (Pitts 2002). It remains for future work to determine whether instances of significant interspecific introgression in fire ants, such as those proposed for the southern highlands and central Atlantic populations of *S. saevissima*, are relatively incidental or, at the other extreme, regularly drive lineage origins through hybrid speciation (e.g., Mavárez and Linares 2008).

Conclusions

Hypotheses regarding species boundaries must rely on evidence that putative species consitute lineages that are evolutionarily independent or clearly on a trajectory to become independent. The evidence comes in the form of the various operational criteria derived as fundamental properties of the "species concepts" of the latter part of the 20th century, properties now recognized as contingent attributes that may or may not pertain to any single case (de Queiroz 2005, 2007). Although no single criterion necessarily must be met to delimit a species, fulfillment of each additional criterion adds corroboration to a hypothesis of lineage independence (existence of a separate species). Thus, we view our task in this study as assimilating evidence relevant to the various operational criteria to reach well-informed decisions about species boundaries. Even so, such decisions necessarily are hypotheses subject to revision, especially in groups of recently diverged lineages where incomplete separation or secondary introgression may be common.

We conclude from our analyses that nominal *S. saevissima* includes several evolutionarily independent lineages that constitute morphologically cryptic species. Based on evidence that they represent distinct genetic clusters, are monophyletic, and are effectively reproductively isolated, the southern highlands, SE Atlantic, and central Atlantic populations are proposed to constitute exemplars of newly recognized undescribed species. Adding to the evidence, the former 2 populations appear to be diagnosable genetically, whereas the latter is ecologically unique among the study populations. The status of the Belo Horizonte population is less clear; although it forms a distinct genetic cluster and is monophyletic, it apparently has experienced recent genetic exchange with the western main range populations and it is not ecologically distinct from one of these. At this point, we feel that the weight of the genetic evidence points toward such restricted gene flow into this population from neighboring populations that we hypothesize it to be on an independent evolutionary trajectory and thus to merit species status. We provisionally call for the remaining 3 populations, from the western main range and Amazonia (Envira), to constitute *S. saevissima sensu stricto* because the type locality (Rio Tapajós, Para, Brazil) is on the Amazon River between our collection localities for these populations (Trager 1991; see Fig. 1). However, the status of the Envira population should also be regarded as provisional; although strongly divergent, our lack of samples from throughout the Amazon drainage where the ant presumably occurs makes it impossible to rule out the presence of continuous clines in gene frequencies and morphological features between this population and the NW main range population.

The confusing panoply of morphological variation documented in *S. saevissima* and related species led Wilson (1952) to conclude that most of the observed fire ant diversity in South America represents geographical variation within a single widespread polytypic species. The newer consensus from modern morphological taxonomic studies is that over a dozen species comprise the *S. saevissima* species group (Trager 1991; Pitts 2002). Nuclear genetic studies generally have substantiated the existence of distinct reproductively isolated populations that correspond to the recognized species (Ross and Trager 1990; Ross and Shoemaker 2005), but strongly divergent genetic groups within some of these nominal species also have been detected, implying the widespread existence of cryptic species (Ross and Shoemaker 2005; Shoemaker et al. 2006; Ross et al. 2007). Results of the present study of *S. saevissima* strongly support the idea of a general underresolution in the morphological delimitation of South American fire ant species.

In addition to contributing to resolution of species boundaries in a taxonomically difficult but economically and scientifically important group, a major goal of this study was to illustrate the power of an integrative

modern approach to tackling species delimitation in taxa with recently diverged lineages that have acquired only some of the contingent properties of species. As such, this and other recent studies (e.g., Leaché et al. 2009; Schlick-Steiner et al. 2010) may serve as models for future research in other difficult groups.

Supplementary Material

Supplementary material can be found at http://www.sysbio.oxfordjournals.org/.

Funding

This study was supported by grants from the United States Department of Agriculture National Research Initiative Competitive Grants Program (036393-01 and 2006-35302-18001) and by the Georgia Agricultural Experiment Stations, University of Georgia. 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.

Acknowledgements

We thank Chris DeHeer, Laurent Keller, Mark Mescher, James Pitts, James Trager, Edward Vargo, and David Williams for their assistance in obtaining specimens; James Pitts, James Trager, and David Williams for identifying all study specimens; and Antoine Guisan, Jérôme Goudet, Steven Phillips, Robin Engler, and Dorothea Pio for their help with the ecological niche modeling. Joe McHugh, James Pitts, and 3 anonymous reviewers provided helpful comments on an earlier draft of the manuscript.

References

Agapow P.M., Bininda-Emonds O.R.P., Crandall K.A., Gittleman J.L., Mace G.M., Marshall J.C., Purvis A. 2004. The impact of species concept on biodiversity studies. Q. Rev. Biol. 79:161–179.

Ahrens M.E., Ross K.G., Shoemaker D.D. 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.

Alfaro M.E., Huelsenbeck J.P. 2006. Comparative performance of Bayesian and AIC-based measures of phylogenetic model uncertainty. Syst. Biol. 55:89–96.

Ascunce M.S., Bouwma A.M., Shoemaker D.D. 2009. Characterization of 24 microsatellite markers in 11 species of fire ants in the genus *Solenopsis* (Hymenoptera: Formicidae). Mol. Ecol. Res. 9:1475–1479.

Balakrishnan R. 2005. Species concepts, species boundaries and species identification: a view from the tropics. Syst. Biol. 54:689–693.

Barton N.H., Slatkin M. 1986. A quasi-equilibrium theory of the distribution of rare alleles in a subdivided population. Heredity. 56:409–415.

Bowcock A.M., Ruíz-Linares A., Tomfohrde J., Minch E., Kidd J.R., Cavalli-Sforza L.L. 1994. High-resolution of human evolutionary trees with polymorphic microsatellites. Nature. 368:455–457.

Brito P.H., Edwards S.V. 2009. Multilocus phylogeography and phylogenetics using sequence-based markers. Genetica. 135:439–455.

Broquet T., Yearsley J., Hirzel A.H., Goudet J., Perrin N. 2009. Inferring recent migration rates from individual genotypes. Mol. Ecol. 18:1048–1060.

Buren W.F. 1972. Revisionary studies on the taxonomy of the imported fire ants. J. Ga. Entomol. Soc. 7:1–26.

Caracristi G., Schlötterer C. 2003. Genetic differentiation between American and European *Drosophila melanogaster* populations could be attributed to admixture of African alleles. Mol. Biol. Evol. 20: 792–799.

Cardoso A., Serrano A., Vogler A.P. 2009. Morphological and molecular variation in tiger beetles of the *Cicindela hybrida* complex: is an 'integrative taxonomy' possible? Mol. Ecol. 18:648–664.

Cavalli-Sforza L.L., Edwards A.W.F. 1967. Phylogenetic analysis: models and estimation procedures. Am. J. Hum. Genet. 19:233–257.

Chakraborty R., Jin L. 1993. Determination of relatedness between individuals using DNA-fingerprinting. Hum. Biol. 65:875–895.

Chan K.M.A., Levin S.A. 2005. Leaky prezygotic isolation and porous genomes: rapid introgression of maternally inherited DNA. Evolution. 59:720–729.

Chen Y.P., Lu L.Y., Skow L.C., Vinson S.B. 2003. Relatedness among co-existing queens within polygyne colonies of a Texas population of the fire ant, *Solenopsis invicta*. Southwest. Entomol. 28:27–36.

Chessel D., Dufour A.-B., Thioulouse J. 2004. The ade4 package-I. One-table methods. R News. 4:5–10.

Chirhart S.E., Honeycutt R.L., Greenbaum I.F. 2005. Microsatellite variation and evolution in the *Peromyscus maniculatus* species group. Mol. Phylogenet. Evol. 34:408–415.

Creighton W.S. 1930. The new world species of the genus *Solenopsis* (Hymenoptera: Formicidae). Proc. Am. Acad. Arts Sci. 66:39–151.

Currat M., Ruedi M., Petit R.J., Excoffier L. 2008. The hidden side of invasions: massive introgression by local genes. Evolution. 62:1908–1920.

Dayrat B. 2005. Towards integrative taxonomy. Biol. J. Linn. Soc. 85: 407–415.

de Queiroz K. 2005. Different species problems and their resolution. BioEssays. 27:1263–1269.

de Queiroz K. 2007. Species concepts and species delimitation. Syst. Biol. 56:879–886.

Dieringer D., Schlötterer C. 2003. Microsatellite analyser (MSA): a platform independent analysis tool for large microsatellite data sets. Mol. Ecol. Notes. 3:167–169.

Donoghue M.J. 1985. A critique of the biological species concept and recommendations for a phylogenetic alternative. Bryologist. 88:172–181.

Elith J., Graham C.H., Anderson R.P., Dudik M., Ferrier S., Guisan A., Hijmans R.J., Huettmann F., Leathwick J.R., Lehmann A., Li J., Lohmann L.G., Loiselle B.A., Manion G., Moritz C., Nakamura M., Nakazawa Y., Overton J.M., Peterson A.T., Phillips S.J., Richardson K., Scachetti-Pereira R., Schapire R.E., Soberón J., Williams S.E., Wisz M.S., Zimmermann N.E. 2006. Novel methods improve prediction of species' distributions from occurrence data. Ecography. 29:129–151.

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

Feldhaar H., Foitzik S., Heinze J. 2008. Lifelong commitment to the wrong partner: hybridization in ants. Philos. Trans. R. Soc. B. 363:2891–2899.

Felsenstein J. 2004. PHYLIP (phylogeny inference package). Seattle (WA): Department of Genome Sciences, University of Washington. Available from http://evolution.genetics.washington.edu/phylip.html.

Fitch W.M., Margoliash E. 1967. Construction of phylogenetic trees. Science. 155:279–284.

Funk D.J., Omland K.E. 2003. Species-level paraphyly and polyphyly: frequency, causes, and consequences, with insights from animal mitochondrial DNA. Annu. Rev. Ecol. Evol. Syst. 34:397–423.

Gascuel O. 1997. BIONJ: an improved version of the NJ algorithm based on a simple model of sequence data. Mol. Biol. Evol. 14:685–695.

Gentile G., della Torre A., Maegga B., Powell J.R., Caccone A. 2002. Genetic differentiation in the African malaria vector, *Anopheles gambiae* s.s., and the problem of taxonomic status. Genetics. 161: 1561–1578.

Glowatzki-Mullis M.L., Muntwyler J., Pfister W., Marti E., Rieder S., Poncet P.A., Gaillard C. 2005. Genetic diversity among horse populations with a special focus on the Franches-Montagnes breed. Anim. Genet. 37:33–39.

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

Guindon S., Gascuel O. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol. 52:696–704.

Guisan A., Zimmermann N.E., Elith J., Graham C.H., Phillips S., Peterson A.T. 2007. What matters for predicting the occurrences of trees: techniques, data, or species' characteristics? Ecol. Monogr. 77:615–630.

Heady R.B., Lucas J.L. 2007. PERMAP 11.6 operation manual. Lafayette (LA): University of Louisiana at Lafayette. Available from http://www.ucs.louisiana.edu/~rbh8900/.

Hedrick P.W. 2000. Genetics of populations. 2nd ed. Sudbury (MA): Jones and Bartlett.

Hedrick P.W. 2005. A standardized genetic differentiation measure. Evolution. 59:1633–1638.

Helbig A.J., Knox A.G., Parkin D.T., Sangster G., Collinson M. 2002. Guidelines for assigning species rank. Ibis. 144:518–525.

Hey J. 2006a. On the failure of modern species concepts. Trends Ecol. Evol. 21:447–450.

Hey J. 2006b. Recent advances in assessing gene flow between diverging populations and species. Curr. Opin. Genet. Dev. 16:592–596.

Hijmans R.J., Cameron S.E., Parra J.L., Jones P.G., Jarvis A. 2005. Very high resolution interpolated climate surfaces for global land areas. Int. J. Climatol. 25:1965–1978.

Huelsenbeck J.P., Rannala B. 2004. Frequentist properties of Bayesian posterior probabilities of phylogenetic trees under simple and complex substitution models. Syst. Biol. 53:904–913.

Hughes M., Hollingsworth P.M. 2008. Population genetic divergence corresponds with species-level biodiversity patterns in the large genus *Begonia*. Mol. Ecol. 17:2643–2651.

Isaac N.J.B., Mallet J., Mace G.M. 2004. Taxonomic inflation: its influence on macroecology and conservation. Trends Ecol. Evol. 19:464–469.

Kanbe T., Akimoto S.-I. 2009. Allelic and genotypic diversity in long-term asexual populations of the pea aphid, *Acyrthosiphon pisum* in comparison with sexual populations. Mol. Ecol. 18:801–816.

Knowles L.L., Carstens B.C. 2007. Delimiting species without monophyletic gene trees. Syst. Biol. 56:887–895.

Krieger M.J.B., Keller L. 1997. Polymorphism at dinucleotide microsatellite loci in fire ant *Solenopsis invicta* populations. Mol. Ecol. 6:997–999.

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

Lambert D.M., Spencer H.G., editors. 1995. Speciation and the recognition concept: theory and application. Baltimore (MD): Johns Hopkins University Press.

Leaché A.D., Koo M.S., Spencer C.L., Papenfuss T.J., Fisher R.N., McGuire J.A. 2009. Quantifying ecological, morphological, and genetic variation to delimit species in the coast horned lizard species complex (*Phrynosoma*). Proc. Natl. Acad. Sci. USA. 106:12418–12423.

Lefébure T., Douady C.J., Gouy M., Gibert J. 2006. Relationship between morphological taxonomy and molecular divergence within Crustacea: proposal of a molecular threshold to help species delimitation. Mol. Phylogenet. Evol. 40:435–447.

Lemmon A.R., Moriarty E.C. 2004. The importance of proper model assumption in Bayesian phylogenetics. Syst. Biol. 53:265–277.

Lessa E.P. 1990. Multidimensional-analysis of geographic genetic-structure. Syst. Zool. 39:242–252.

Light J.E., Toups M.A., Reed D.L. 2008. What's in a name: the taxonomic status of human head and body lice. Mol. Phylogent. Evol. 47:1203–1216.

Mallet J. 1995. A species definition for the modern synthesis. Trends Ecol. Evol. 10:294–299.

Markin G.P., Dillier J.H., Hill S.O., Blum M.S., Hermann H.R. 1971. Nuptial flight and flight ranges of the imported fire ant, *Solenopsis saevissima richteri* (Hymenoptera: Formicidae). J. Ga. Entomol. Soc. 6:145–156.

Marshall J.C., Arévalo E., Benavides E., Sites J L., Sites J.W. 2006. Delimiting species: comparing methods for Mendelian characters using lizards of the *Sceloporus grammicus* (Squamata: Phrynosomatidae) complex. Evolution. 60:1050–1065.

Mavárez J., Linares M. 2008. Homoploid hybrid speciation in animals. Mol. Ecol. 17:4181–4185.

Mayden R.L. 1999. Consilience and a hierarchy of species concepts: advances toward closure on the species puzzle. J. Nematol. 31:95–116.

Mayr E. 1982. The growth of biological thought: diversity, evolution, and inheritance. Cambridge (MA): Harvard University Press.

McDade L.A. 1992. Hybrids and phylogenetic systematics II. The impact of hybrids on cladistic analysis. Evolution. 46:1329–1346.

Menkis A., Bastiaans E., Jacobson D.J., Johannesson H. 2009. Phylogenetic and biological species diversity within the *Neurospora tetrasperma* complex. J. Evol. Biol. 22:1923–1936.

Michalakis Y., Excoffier L. 1996. A generic estimation of population subdivision using distances between alleles, with special reference for microsatellite loci. Genetics. 142:1061–1064.

Michener C.D. 1970. Diverse approaches to systematics. Evol. Biol. 4:1–38.

Minder A.M., Widmer A. 2008. A population genomic analysis of species boundaries: neutral processes, adaptive divergence and introgression between two hybridizing plant species. Mol. Ecol. 17:1552–1563.

Morando M., Avila L.J., Sites J.W. 2003. Sampling strategies for delimiting species: genes, individuals, and populations in the *Liolaemus elongatus-kriegi* complex (Squamata: Liolaemidae) in Andean-Patagonian South America. Syst. Biol. 52:159–185.

Mullen S.P., Dopman E.B., Harrison R.G. 2008. Hybrid zone origins, species boundaries, and the evolution of wing-pattern diversity in a polytypic species complex of North American admiral butterflies (Nymphalidae: *Limenitis*). Evolution. 62:1400–1417.

Nosil P. 2008. Speciation with gene flow could be common. Mol. Ecol. 17:2103–2106.

Nosil P., Funk D.J., Ortiz-Barrientos D. 2009. Divergent selection and heterogeneous genomic divergence. Mol. Ecol. 18:375–402.

Novembre J., Stephens M. 2008. Interpreting principal component analyses of spatial population genetic variation. Nat. Genet. 40:646–649.

Patterson N., Price A.L., Reich D. 2006. Population structure and eigenanalysis. PLoS Genet. 2:2074–2093.

Pearson R.G., Raxworthy C.J., Nakamura M., Peterson A.T. 2007. Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. J. Biogeogr. 34:102–117.

Petit R.J., Excoffier L. 2009. Gene flow and species delimitation. Trends Ecol. Evol. 24:386–393.

Phillips S.J., Anderson R.P., Shapire R.E. 2006. Maximum entropy modeling of species geographic distributions. Ecol. Modell. 190:231–259.

Phillips S.J., Dudík M. 2008. Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. Ecography. 31:161–175.

Pitts J.P. 2002. A cladistic analysis of the *Solenopsis saevissima* species-group (Hymenoptera: Formicidae) [Ph.D. dissertation]. Athens (Georgia): University of Georgia.

Pitts J.P., McHugh J.V., Ross K.G. 2005. Cladistic analysis of the fire ants of the *Solenopsis saevissima* species-group (Hymenoptera: Formicidae). Zool. Scr. 34:493–505.

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

Posada D. 2008. jModelTest: phylogenetic model averaging. Mol. Biol. Evol. 25:1253–1256.

Posada D., Buckley T.R. 2004. Model selection and model averaging in phylogenetics: advantages of Akaike information criterion and Bayesian approaches over likelihood ratio tests. Syst. Biol. 53:793–808.

Price T. 2008. Speciation in birds. Greenwood Village (CO): Roberts and Company.

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

Pritchard J.K., Wen X., Fa1ush D. 2007. Documentation for STRUCTURE software. Version 2.2. Chicago (IL): Department of Human Genetics, University of Chicago. Available from http://pritch.bsd.uchicago.edu/software.

Raxworthy C.J., Ingram C.M., Rabibisoa N., Pearson R.G. 2007. Applications of ecological niche modeling for species delimitation: a review and empirical evaluation using day geckos (Phelsuma) from Madagascar. Syst. Biol. 56:907–923.

Raychoudhury R., Baldo L., Oliveira D.C.S.G., Werren J.H. 2009. Modes of acquisition of *Wolbachia*: horizontal transfer, hybrid introgression, and codivergence in the *Nasonia* species complex. Evolution. 63:165–183.

Raymond M., Rousset F. 1995. GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. J. Hered. 86:248–249.

Renema W., Bellwood D.R., Braga J.C., et al. 2008. Hopping hotspots: global shifts in marine biodiversity. Science. 321:654–657.

Richards V.P., Henning M., Witzell W., Shivji M.S. 2009. Species delineation and evolutionary history of the globally distributed Spotted Eagle Ray (*Aetobatus narinari*). J. Hered. 100:273–283.

Rissler L.J., Apodaca J.J. 2007. Adding more ecology into species delimitation: ecological niche models and phylogeography help define cryptic species in the black salamander (*Aneides flavipunctatus*). Syst. Biol. 56:924–942.

Ronquist F., Huelsenbeck J.P. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics. 19:1572–1574.

Ross K.G., Keller L. 1995. Joint influence of gene flow and selection on a reproductively important genetic polymorphism in the fire ant *Solenopsis invicta*. Am. Nat. 146:325–348.

Ross K.G., Krieger M.J.B., Keller L., Shoemaker D.D. 2007. Genetic variation and structure in native populations of the fire ant *Solenopsis invicta*: evolutionary and demographic implications. Biol. J. Linn. Soc. 92:541–560.

Ross K.G., Krieger M.J.B., Shoemaker D.D. 2003. Alternative genetic foundations for a key social polymorphism in fire ants. Genetics. 165:1853–1867.

Ross K.G., Krieger M.J.B., Shoemaker D.D., Vargo E.L., 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 K.G., Shoemaker D.D. 2005. Species delimitation in native South American fire ants. Mol. Ecol. 14:3419–3438.

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

Ross K.G., Vargo E.L., Fletcher D.J.C. 1988. Colony genetic structure and queen mating frequency in fire ants of the subgenus *Solenopsis* (Hymenoptera: Formicidae). Biol. J. Linn. Soc. 34:105–117.

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

Rozas J., Sánchez-DelBarrio J.C., Messeguer X., Rozas R. 2003. DNASP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics. 19:2496–2497.

Saitou N., Nei M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4:406–425.

Schlick-Steiner B.C., Steiner F.M., Seifert B., Stauffer C., Christian E., Crozier R.H. 2010. Integrative taxonomy: a multisource approach to exploring biodiversity. Annu. Rev. Entomol. 55:421–438.

Schlötterer C. 2001. Genealogical inference of closely related species based on microsatellites. Genet. Res. 78:209–212.

Schluter D. 2009. Evidence for ecological speciation and its alternative. Science. 323:737–741.

Sei M., Porter A.H. 2007. Delimiting species boundaries and the conservation genetics of the endangered maritime ringlet butterfly (*Coenonympha nipisiquit* McDunnough). Mol. Ecol. 16:3313–3325.

Seifert B. 2009. Cryptic species in ants (Hymenoptera: Formicidae) revisited: we need a change in the alpha-taxonomic approach. Myrmecol. News. 12:149–166.

Seifert B., Goropashnaya A.V. 2004. Ideal phenotypes and mismatching haplotypes—errors of mtDNA treeing in ants (Hymenoptera: Formicidae) detected by standardized morphometry. Org. Divers. Evol. 4:295–305.

Shaffer H.B., Thomson R.C. 2007. Delimiting species in recent radiations. Syst. Biol. 56:896–906.

Shimodaira H., Hasegawa M. 1999. Multiple comparisons of loglikelihoods with applications to phylogenetic inference. Mol. Biol. Evol. 16:1114–1116.

Shoemaker D.D., Ahrens M.E., Ross K.G. 2006. Molecular phylogeny of fire ants of the *Solenopsis saevissima* species-group based on mtDNA sequences. Mol. Phylogenet. Evol. 38:200–215.

Shoemaker D.D., Costa J.T., Ross K.G. 1992. Estimates of heterozygosity in two social insects using a large number of electrophoretic markers. Heredity. 69:573–582.

Shoemaker D.D., Keller G., Ross K.G. 2003. Effects of *Wolbachia* on mtDNA variation in two fire ant species. Mol. Ecol. 12:1757–1771.

Sites J.W., Marshall J.C. 2004. Operational criteria for delimiting species. Annu. Rev. Ecol. Evol. Syst. 35:199–227.

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.

Stockman A.K., Bond J.E. 2007. Delimiting cohesion species: extreme population structuring and the role of ecological interchangeability. Mol. Ecol. 16:3374–3392.

Strasburg J.L., Rieseberg L.H. 2008. Molecular demographic history of the annual sunflowers *Helianthus annuus* and *H. petiolaris*: large effective population sizes and rates of long-term gene flow. Evolution. 62:1936–1950.

Swofford D.L. 2002. PAUP*: phylogenetic analysis using parsimony (*and other methods). Sunderland (MA): Sinauer Associates.

Tamura K., Nei M. 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial-DNA in humans and chimpanzees. Mol. Biol. Evol. 10:512–526.

Trager J.C. 1991. A revision of the fire ants, *Solenopsis geminata* group (Hymenoptera: Formicidae: Myrmicinae). J. N. Y. Entomol. Soc. 99:141–198.

Tschinkel W.R. 2006. The fire ants. Cambridge (MA): Harvard University Press.

Turner T.L., Hahn M.W., Nuzhdin S.V. 2005. Genomic islands of speciation in *Anopheles gambiae*. PLoS Biol. 3:1572–1578.

Van Driesche R., Bellows T.S. 1996. Biological control. New York: Chapman & Hall.

Van Valen L. 1976. Ecological species, multispecies, and oaks. Taxon. 25:233–239.

Vila C., Leonard J.A., Gotherstrom A., Marklund S., Sandberg K., Liden K., Wayne R.K., Ellegren H. 2001. Widespread origins of domestic horse lineages. Science. 291:474–477.

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

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

Wiens J.J. 2007b. Global patterns of diversification and species richness in amphibians. Am. Nat. 170:S86–S106.

Wiens J.J. 2007b. Species delimitation: new approaches for discovering diversity. Syst. Biol. 56:875–878.

Wiens J.J., Graham C.H. 2005. Niche conservatism: integrating evolution, ecology, and conservation biology. Annu. Rev. Ecol. Evol. Syst. 36:519–539.

Wilson E.O. 1952. O complexo *Solenopsis saevissima* na America do Sul (Hymenoptera: Formicidae). Mem. Inst. Oswaldo Cruz. 50:49–59.

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

Wisz M.S., Hijmans R.J., Li J., Peterson A.T., Graham C.H., Guisan A., NCEAS Predicting Species Distributions Working Group. 2008. Effects of sample size on the performance of species distribution models. Divers. Distrib. 14:763–773.

Xie X., Michel A.P., Schwarz D., Rull J., Velez S., Forbes A.A., Aluja M., Feder J.L. 2008. Radiation and divergence in the *Rhagoletis pomonella* species complex: inferences from DNA sequence data. J. Evol. Biol. 21:900–913.

Zink R.M., Barrowclough G.F. 2008. Mitochondrial DNA under siege in avian phylogeography. Mol. Ecol. 17:2107–2121.