Insect Systematics and Diversity, (2022) 6(1): 10; 1–15
https://doi.org/10.1093/isd/ixab027
Research

ENTOMOLOGICAL SOCIETY OF AMERICA
SHARING INSECT SCIENCE GLOBALLY

OXFORD

Molecular Phylogenetics, Phylogenomics, and Phylogeography

Phylogenomic Delimitation of Morphologically Cryptic Species in Globetrotting *Nylanderia* (Hymenoptera: Formicidae) Species Complexes

Jason L. Williams,[1,5] Y. Miles Zhang,[2] John S. LaPolla,[3] Ted R. Schultz,[4] and Andrea Lucky[1]

[1]Entomology & Nematology Department, University of Florida, 1881 Natural Area Drive, Gainesville, FL, 32611, USA, [2]Systematic Entomology Laboratory, ARS-USDA, Smithsonian Institution, Washington, DC, 20560, USA, [3]Department of Biological Sciences, Towson University, Towson, MD 21252, USA, [4]Department of Entomology, National Museum of Natural History, Smithsonian Institution, Washington, DC 20560, USA, and [5]Corresponding author, e-mail: jwilli81@ufl.edu

Subject Editor: Gabriela P. Camacho

Received 28 May, 2021; Editorial decision 19 October, 2021

Abstract

The ant genus *Nylanderia* Emery has a cosmopolitan distribution and includes 150 extant described species and subspecies, with potentially hundreds more undescribed. Global taxonomic revision has long been stalled by strong intra- and interspecific morphological variation, limited numbers of diagnostic characters, and dependence on infrequently collected male specimens for species description and identification. Taxonomy is further complicated by *Nylanderia* being one of the most frequently intercepted ant genera at ports of entry worldwide, and at least 15 globetrotting species have widespread and expanding ranges, making species-level diagnoses difficult. Three species complexes ('*bourbonica* complex', '*fulva* complex', and '*guatemalensis* complex') include globetrotting species. To elucidate the phylogenetic positions of these three complexes and delimit species boundaries within each, we used target enrichment of ultraconserved elements (UCEs) from 165 specimens representing 98 *Nylanderia* morphospecies worldwide. We also phased the UCEs, effectively doubling sample size and increasing population-level sampling. After recovering strong support for the monophyly of each complex, we extracted COI barcodes and SNPs from the UCE data and tested within-complex morphospecies hypotheses using three molecular delimitation methods (SODA, bPTP, and STACEY). This comparison revealed that most methods tended to over-split taxa, but results from STACEY were most consistent with our morphospecies hypotheses. Using these results, we recommend species boundaries that are conservative and most congruent across all methods. This work emphasizes the importance of integrative taxonomy for invasive species management, as globetrotting occurs independently across at least nine different lineages across *Nylanderia*.

Key words: COI barcoding, integrative taxonomy, invasive species, SNP, ultraconserved element

> "The taxonomist cannot be expected to evince enthusiasm for species which can only be distinguished by a different rate of wing beat or the structure of the salivary chromosomes."
> -William Steel Creighton (1950)

Evolutionary biologist Ernst Mayr (1942) was first to use the term 'sibling species' (also known as 'cryptic species') to describe those which are not recognizable as separate species but are reproductively isolated and coexist without interbreeding (i.e., The Biological Species Concept), one of over two dozen concepts leveraged to understand speciation and attempt to delimit species boundaries (De Queiroz 2007). Species delimitation methods have changed dramatically since the days of Mayr and his contemporaries, who were aware of cryptic diversity through keen observation of natural history but worked with smaller collections and a far more limited set of tools than are available today. A more modern and widely adopted interpretation of cryptic species can be found in Seifert (2009) who defines them as 'two or more species which are not separable by primary visual or acoustic perception of an expert.' Purely

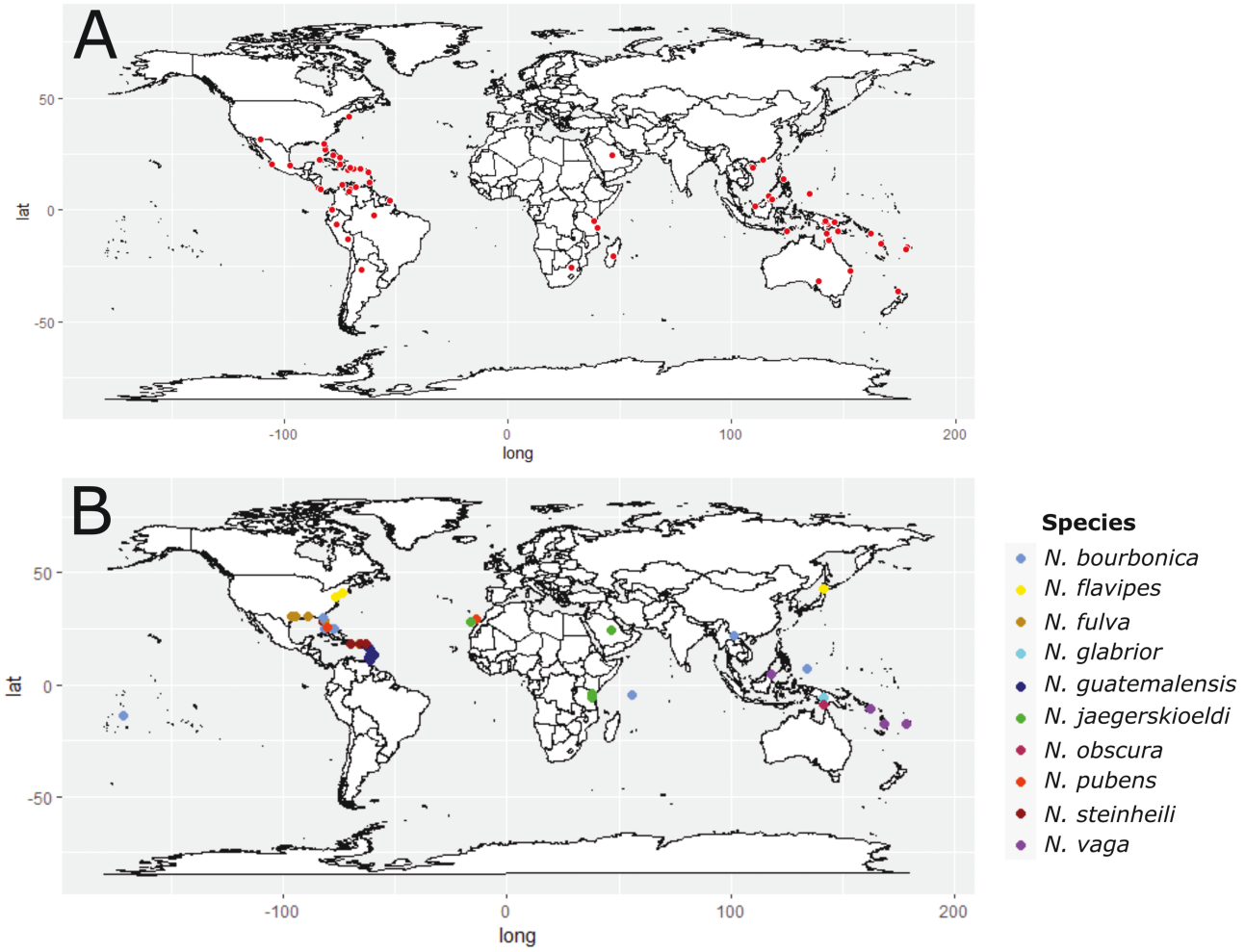

Fig. 1. Maps of sampling localities for all 165 samples used in UCE library preparation for this study. (A) (top) indicates sampling localities of all nonglobetrotting species (i.e., species in their native ranges; red points). |(B) (bottom) indicates sampling localities of all globetrotting species. See Supp Table S1 [online only] for detailed locality information for each sample.

observation-based classification often carries a kernel of truth but can be misleading when based on unreliable characters or misinterpretation of intra- and interspecific variation due to limited taxon sampling. Increasingly, studies using molecular data reveal surprising levels of genetic diversity and cryptic biodiversity in morphologically monotonous taxa (Beheregaray and Caccone 2007). A variety of molecular species delimitation methods such as bPTP (Zhang et al. 2013), BPP (Yang 2015), SODA (Rabiee and Mirarab 2020), and STACEY (Jones 2017) offer new ways to infer species boundaries, a problem that has confounded taxonomists for centuries.

Species concepts matter in any discussion of boundaries, and among the most widely accepted today is the unified species concept which states that species are separately evolving metapopulation lineages that acquire certain defining properties over the course of divergence (De Queiroz 2007). Following this concept, effective species delimitation depends on the explicit acknowledgment that a variety of biological mechanisms underlie speciation, and individual operational criteria emerge at different periods during the speciation process. Increasingly, integrative taxonomic approaches are being employed under this paradigm to delimit cryptic species based on multiple operational criteria (Barraclough 2019, Burbrink et al. 2011, Chaplin et al. 2019, Domingos et al. 2014, Fišer et al. 2018). Machine learning and model-based approaches to refine species delimitation have also been adopted (Branstetter and Longino 2019, Derkarabetian et al. 2019, Fujisawa and Barraclough 2013, Gueuning et al. 2020, Prebus 2021). In ants, thousands of genome-scale short fragment reads (400–600 bp) known as Ultraconserved Elements (UCEs) have been used to holistically resolve cryptic taxa through phylogenomic species delimitation in the context of other molecular markers (e.g., mitochondrial, SNPs), morphology, ecology, and biogeography (Ješovnik et al. 2017, Longino and Branstetter 2020, Prebus 2021, Wagner et al. 2017).

Practical Challenges to Cryptic Species

Accurate identification of invasive species in the early stages of invasion is critical for management, but cryptic invasions often go unnoticed until well beyond the point when eradication or effective quarantine would be cost-effective or even attainable (Krushelnycky et al. 2005). These 'silent invasions' afford cryptic invasive species ample time to go unnoticed and become established. For example, an already-established population of the Southeast Asian Tanezumi rat, *Rattus tanezumi* (Rodentia: Muridae) Temminck, was detected unexpectedly in South Africa through genetic monitoring of two other invasive *Rattus* species (Bastos et al. 2011). This ecologically and economically devastating species is morphologically indistinguishable from the black rat, *Rattus rattus* (L.) (Rodentia: Muridae), and both species are known as reservoirs for a wide variety of pathogens that

can cause die-offs and local extinction of larger mammals (Bastos et al. 2011). Monitoring, early detection, control, and eradication remain impractical without timely and accurate species diagnosis.

Cryptic invasions are likely more widespread than currently understood as a result of the difficulty in detecting them (Morais and Reichard 2018). A classic example of this in ants is the misidentification of the tropical fire ant, *Solenopsis geminata* (Fabricius) (Hymenoptera: Formicidae), as the red-imported fire ant, *Solenopsis invicta* Buren (Hymenoptera: Formicidae) (Wetterer 2013, 2011), in Malaysia and Singapore, and possibly elsewhere (e.g., India, the Philippines). The former species is native to southern North America, Central America, South America, and the Caribbean, while the latter is native to South America (distributions from AntMaps: Janicki et al. 2016). In this case, the inability to correctly identify *S. geminata* meant that its distribution was underestimated and the invaded range of *S. invicta* was overestimated. Both of these species are known for endangering native wildlife and domestic animals, for being costly to control, and for having a painful sting that can cause a variety of health issues, including secondary infection, anaphylactic shock, or even death (Wetterer 2013, 2011). This type of misidentification is likely in taxa with many undescribed species and multiple globetrotting species. Resolving species boundaries and making taxonomic tools available can help prevent prolonged confusion about species identities.

Nylanderia Taxonomy

Species within *Nylanderia* and six other closely-related genera—collectively called the '*Prenolepis* genus-group'—have vexed taxonomists for more than a century because most species have few to no readily distinguishable morphological characters (LaPolla et al. 2010, 2012; Williams and LaPolla 2016; Williams and Lucky 2020). In *Nylanderia*, globetrotting species co-occur alongside morphologically identical endemic species on nearly every continent, leading to difficulty distinguishing species. Historical classifications based mainly on morphology are insufficient for identifying species that concurrently exhibit high intraspecific variation and low interspecific variation. Often, the characters cited for distinguishing workers of *Nylanderia* species are subtle and include differences in color, size, cuticular sculpturing/microstructure, the presence or absence of pubescence or pilosity on parts of the body, or minor differences in macrosetal counts (Kallal and LaPolla 2012, LaPolla et al. 2011, LaPolla and Kallal 2019, Trager 1984). Male characters—especially genitalia—are often cited as useful (Gotzek et al. 2012, LaPolla and Kallal 2019, Trager 1984), though males are relatively rare in collections, not known at all for many species, and difficult to associate with worker conspecifics if not collected as part of a nest series. Queens, at least externally, appear to be the least useful among workers, queens, and males for distinguishing *Nylanderia* species.

The potential ecological and economic consequences of taxonomic confusion in this genus are significant considering the widespread co-occurrence of globetrotting *Nylanderia* species with morphologically identical endemic species. For example, three range-limited endemic Caribbean species—*Nylanderia lucayana* LaPolla and Kallal, *Nylanderia coveri* LaPolla and Kallal, and *Nylanderia esperanza* LaPolla and Kallal—are very similar morphologically to the globetrotting *N. guatemalensis* (LaPolla and Kallal 2019), rendering conservation of endemics and management of the invasive species extremely difficult. Similarly, in the southeastern United States in the early 2000s, early records of the now-economically devastating introduced species *Nylanderia fulva* (Mayr) were misidentified as *Nylanderia pubens* (Forel). The confusion resulted from identification based solely on examination of workers; currently, these two species can be distinguished only by differences in male genitalia. Males of *N. fulva* are distinguished from *N. pubens* by genitalic shape and pilosity: in profile, the secondary gonopods (formerly called parameres; see Boudinot (2018)) of *N. fulva* are triangular, less sclerotized, and have fewer macrosetae arising from the margin, while those of *N. pubens* are more rounded, well sclerotized, and have a dense, fan-like projection of long macrosetae arising from the margin (Gotzek et al. 2012, Trager 1984).

Species boundaries—along with the full extent of intraspecific and interspecific morphological variation of all castes and sexes in the *fulva* complex—remain unclear and the taxonomic utility of male genitalia in *Nylanderia* remains untested. Because the molecular phylogeny of Gotzek et al. (2012) did not provide conclusive evidence that *N. pubens* is distinct from *N. fulva*, Deyrup (2016) asserted that *N. pubens* could be a more 'hairy-tailed' form of *N. fulva*, perhaps with a different geographic origin. Investigation of genitalic variation in *N. pubens* has been particularly difficult because male specimens are relatively rare in collections.

As the specific example of *N. fulva* and *N. pubens* illustrates, the interpretation of intra- versus interspecific morphological variation in the context of an inclusive molecular phylogeny is necessary for integrative delimitation of cryptic *Nylanderia* species. Previous studies have sequenced five to six nuclear markers and COI barcodes to reconstruct the phylogeny of *Nylanderia* (Gotzek et al. 2012, LaPolla et al. 2010, Matos-Maraví et al. 2018), while Williams et al. (2020) were the first to employ UCE sequences from *Nylanderia* species representing all major biogeographic regions. However, methods of molecular species delimitation have never been employed for resolving some of the most difficult questions involving globetrotting and invasive *Nylanderia* species. Some species as currently defined, such as *N. bourbonica*, appear to be highly morphologically variable and to have broad distributions. In these cases, it is necessary to determine whether populations sampled across the full geographic range are most closely related to each other (i.e., truly constitute a single, monophyletic species) and to document the limits of morphological and genetic variation expressed across the species. Simultaneously, it is also necessary to identify cases in which morphological variation overlaps in distantly related species.

Main Objectives

The main goal of this study is to delimit species boundaries within three *Nylanderia* species complexes: (i) *bourbonica* complex (including *N. bourbonica* and *N. vaga*), (ii) *fulva* complex (including *N. fulva* and *N. pubens*), and (iii) *guatemalensis* complex (including *N. guatemalensis* and *N. steinheili*). UCEs are versatile and can also be leveraged for species delimitation by using an allelic phasing approach, outperforming mitochondrial DNA barcoding for this purpose (Gueuning et al. 2020). By phasing the UCE data into haplotypes and performing a multispecies coalescent (MSC) analysis for each species complex, the population and species-level relationships within and among globetrotting species and their closest relatives can be assessed. We compare the utility of three different molecular delimitation methods (SODA, bPTP, and STACEY), as well as COI barcodes extracted from the UCE dataset and random SNPs extracted from phased UCEs. We expected that methods using single-locus, COI barcodes to delimit taxa would over-split taxa the most, and that the methods using the full UCE dataset and genome-wide SNPs sampled from phased UCEs would delimit species boundaries

more consistent with our a priori species hypotheses than methods using COI.

We test each of the three globetrotting species complexes for monophyly. In the *bourbonica* complex, the working hypothesis is that *N. bourbonica* and *N. vaga*, despite overlapping intraspecific variation in worker mesosoma shape, color, size, and pilosity across their broadest geographical distributions, are reciprocally monophyletic sister species. In the *fulva* complex, previous studies attempted to demonstrate the independence of *N. fulva* and *N. pubens* as distinct species, diagnosable using male genitalic characters, but sampling has been limited and molecular phylogenetic results have been inconclusive (Gotzek et al. 2012, LaPolla and Kallal 2019). We, therefore, test the hypothesis that *N. fulva* and *N. pubens* represent two genetically distinct, monophyletic species with limited variation in worker color, size, and pilosity. In the *guatemalensis* complex, the phylogenetic relationship between *N. guatemalensis* and *N. steinheili* is evaluated: each of these species appears to be monophyletic, but limited sampling of Neotropical species has made it difficult to determine whether they are sister species (Williams et al. 2020). Because the biogeographical origin of each of these species is also unknown, we assess whether each species originated from islands in the Caribbean or from mainland South or Central America.

Materials and Methods

Taxon Sampling for Molecular Data

In total, our analyses included UCE sequences from 165 ant specimens representing 98 *Nylanderia* morphospecies and two outgroups (*Paratrechina longicornis* (Latreille) and *Paraparetrechina* sp. ems01) (Supp Table S1 [online only]). Sixty-five sequences were newly generated for this study; previously generated data included 96 *Nylanderia* sequences (Williams et al. 2020) and four other *Nylanderia* sequences (Blaimer et al. 2015) that were downloaded from the NCBI Sequence Read Archive (SRR2184158, SRR2184159, SRR2184160, and SRA2184179). Results from Williams et al. (2020) were used to assess phylogenetic diversity and identify gaps in taxon sampling. We attempted to maximize the representation of genus-wide biogeographic, ecological, and phylogenetic diversity, with a particular focus on globetrotting species and species native to the Neotropics and Australasia. Diversity was also assessed a priori by sorting all samples into morphospecies and selecting representatives of each across their broadest distributions and distinct habitat types. This sampling scheme included a focused sampling of taxa in the three complexes of interest in this research: *N. bourbonica* (n = 13), *N. vaga* (n = 13), *N. fulva* (n = 6), *N.* cf. *fulva* (n = 2), *N. pubens* (n = 1), *N. guatemalensis* (n = 9), and *N. steinheili* (n = 6). A map of all sampling localities included in this study can be found in Fig. 1.

DNA Extraction

DNA extractions for 23 samples from Matos-Maraví et al. (2018) were donated by Milan Janda. For the remaining 42 samples, DNA was destructively extracted from workers using DNeasy Blood and Tissue Kits (Qiagen, Valencia, CA) as in Williams et al. (2020). For three samples (Ny165: *N. bourbonica*, Ny167: *N.* sp. Ven5, and Ny174: *N.* sp. NZ01), DNA was destructively extracted from single pinned specimens for which multiple other worker representatives from the same colony were retained as vouchers. For all remaining samples, DNA was extracted from a single whole worker stored in 70% to 95% ethanol, with additional specimens pinned as vouchers. Specimen collection year ranged from 1976 to 2018. Voucher specimens are deposited in the Smithsonian Institution National Museum of Natural History Insect Collection (USNM). See Supp Table S1 [online only] for specimen data.

UCE Sequence Data Collection

All UCE laboratory work was conducted in the Laboratories of Analytical Biology (LAB) at the Smithsonian Institution's National Museum of Natural History (NMNH, Washington, DC, USA). The protocol of Branstetter et al. (2017c) was followed to capture and enrich up to 2,590 UCE loci from each sample. See Williams et al. (2020) for a full description of wet lab work that was performed during this study. The final pool-of-pools was sent to Novogene (Sacramento, CA) and sequenced on an Illumina NovaSeq 6000 (150-bp paired-end, Illumina Inc., San Diego, CA). The mean DNA concentration of all samples was 0.666 ng/µl (0.05–2.67 ng/µl) after extraction and 17.01 ng/µl (1–64.5 ng/µl) post-PCR. See Table 1 for a summary of UCE processing and sequencing statistics. A full summary of statistics for each specimen can be found in Supp Table S2 [online only].

UCE Processing and Alignment

UCE sequences were processed using the PHYLUCE v1.6.6 pipeline (Faircloth 2015) using mostly the default settings, except with the minimum sequence identity set to 85% (min_identity: 85) and with the cleaned reads assembled using SPAdes (Bankevich et al. 2012). Bossert and Danforth (2018) found that a minimum identity set between 82% and 85% similarity helps filter contamination while retaining a sufficient capture rate of target sequences. Alignments were trimmed using a GBLOCKS (Castresana 2000) wrapper script (settings: b1 = 0.5, b2 = 0.5, b3 = 12, b4 = 7). An 80% complete matrix retaining 885 loci was selected for the final dataset ('Ny165_80p'; Supp File S1 [online only]), as very few loci were available in a 90 or 100% complete matrix. Summary statistics for this matrix were calculated using the Alignment Manipulations and Summary (AMAS) script (Borowiec 2016) and can be found in Table 2. Quality-trimmed reads for all generated sequences in this study are available

Table 1. Ultraconserved element sequencing statistics for all 165 samples

	Extract conc.(ng/µl)	Post-PCR conc.(ng/µl)	Raw reads	Contigs	Total bp	Mean length	Min length	Max length	Median length
Mean	0.666	17.01	2,883,150	1743.491	1,097,319	579.669	111.612	2582.855	544.733
Min	0.05	1	43,794	64	15,301	239.078	55	394	237
Max	2.67	64.5	2,934,3045	2,249	2,696,882	1226.413	239	36,809	1,126
SD	0.436	16.008	3,687,573	590.474	599629.8	218.092	56.789	3225.14	204.118
95% CI	0.067	2.473	573,949	90.096	91493.3	33.277	8.665	492.102	31.145

CI, confidence interval; SC, standard deviation.

from the National Center for Biotechnology Sequence Read Archive (http://www.ncbi.nlm.nih.gov/sra; PRJNA728185, PRJNA553590).

The PHYLUCE (Faircloth 2015) allelic phasing workflow from Andermann et al. (2019) was used to phase the UCE data from an edge-trimmed alignment. Allelic phasing separates heterozygous sites into two allele sequences per sample, effectively doubling sample size, and has been shown to outperform contig sequences when estimating shallow-level phylogeny under multispecies coalescent models (Andermann et al. 2019, Zhang et al. 2019). Two subsets of taxa from the full dataset were phased representing: (i) the *bourbonica* complex and close relatives (50 taxa); and (ii) the *fulva* + *guatemalensis* complex (38 taxa). For the first subset an unphased 75% complete matrix retaining 1,988 loci ('bourb50-unphased_75p'; Supp File 2 [online only]) and a phased 75% complete matrix retaining 2,000 loci ('bourb50-phased_75p'; Supp File 3 [online only]) were selected. For the second subset an unphased 75% complete matrix retaining 1,253 loci ('fg37-unphased_75p; Supp File 4 [online only]) and a phased 75% complete matrix retaining 1,699 loci ('fg37-phased_75p'; Supp File 5 [online only]) were selected. Unphased and phased 90% and 100% matrices were also generated for each of the two subsets (Table 2). Summary statistics for all matrices were calculated using AMAS and can be found in Table 2.

Phylogenomic Analyses

For the full dataset (all 165 taxa), phylogenies were constructed under a maximum-likelihood (ML) criterion in IQ-TREE 2 (Minh et al. 2020) using two partitioning strategies: (i) partition by locus; and (ii) a scheme determined by PARTITIONFINDER2 (n = 483; Lanfear et al. 2017) with number of partitions determined by Sliding-Window Site Characteristics (SWSC; Tagliacollo and Lanfear 2018) based on site entropy. We assessed nodal support with a Shimodaira–Hasegawa approximate likelihood-rate test (SH-aLRT; Guindon et al. 2010) using 1,000 replicates with the -alrt command in IQ-TREE 2. One thousand ultrafast bootstrap (UFBoot; Hoang et al. 2017) replicates were run using '-bb' and '-bnni'. Only nodes with values of SH-aLRT ≥0.90 and UFBoot ≥0.95 were considered strongly supported.

Species trees for the unphased and phased subsets of data representing the *bourbonica* complex and the *fulva* + *guatemalensis* complex were reconstructed under the multispecies coalescent (MSC) model using ASTRAL-III v5.6.2 (Zhang et al. 2018) after using IQ-TREE 2 to reconstruct a gene tree for each locus using the '-S' flag. We collapsed all branches in the gene trees with <10% bootstrap support using the 'nw_ed' command included in Newick Utilities (Junier and Zdobnov 2010), as this step has been shown to improve accuracy in MSC analysis (Zhang et al. 2018). Loci were not partitioned, and optimal nucleotide substitution models were selected using the Bayesian Information Criterion (BIC). Nodal support was assessed using local posterior probabilities (Sayyari and Mirarab 2016), with nodes having a value of ≥0.95 considered to be strongly supported.

Molecular Species Delimitation

The 'hym-v2-ants' UCE probe set (Branstetter et al. 2017) includes markers for capturing COI data, and mitochondrial barcodes can be extracted from UCE enriched samples using the '*phyluce_assembly_match_contigs_to_barcodes*' script in PHYLUCE (Faircloth 2015). We used this script to obtain COI barcodes for all UCE samples included in the two subsets of taxa used to represent the *bourbonica* complex and *fulva* + *guatemalensis* complex. For each subset, the COI barcodes were aligned using MAFFT and then the alignments were used to reconstruct a ML phylogeny in IQ-TREE 2 (Minh et al. 2020) with partitioning by codon position and 1,000 Ultrafast bootstrap replicates (UFBoot; Hoang et al. 2017). The aligned FASTA files and tree files were used as inputs to delimit species using Bayesian implementation of the Poisson tree processes model for species delimitation (bPTP; Zhang et al. 2013). For the bPTP analysis we ran 100,000 MCMC generations with a thinning value of 100 and 10% burn-in, removed the outgroup taxon to improve species delimitation, and checked stationarity plots to ensure that convergence of the MCMC chain was reached.

We delimited species boundaries in the unphased and phased UCE data subsets representing the *bourbonica* complex ('bourb50-unphased_75p' and 'bourb50-phased_75p') and the *fulva* + *guatemalensis* complexes ('fg37-unphased_75p and 'fg37-phased_75p') using SODA (Rabiee and Mirarab 2020) with a guide tree built using ASTRAL-III. SODA delimits species boundaries using quartet frequencies. We also extracted one random SNP per UCE locus from phased UCE subsets ('bourb27-phased_90p' and 'fg37-phased_90p') using the '*phyluce_snp_screen_phased_alignments*' script in PHYLUCE and then concatenated the outputs using AMAS. The concatenated nexus files were used as inputs for the STACEY (Jones 2017) package in BEAST2 (Bouckaert et al. 2014). BEAUti was used to configure the input files using the STACEY template. In the 'Taxon sets' tab, phased SNP partitions were assigned to

Table 2. Summary statistics for alignment matrices, calculated using the Alignment Manipulations and Summary script (AMAS; Borowiec 2016)

Alignment	Loci	Length	Total matrix cells	Missing %	Proportion variable	Parsimony informative
Ny165_80p	885	268,331	44,274,615	20.8	0.421	0.176
bourb50-unphased_75p	1,988	1,194,451	58,528,099	38.376	0.096	0.025
bourb50-unphased_90p	964	632,672	31,000,928	38.139	0.095	0.024
bourb50-unphased_100p	14	10,111	495,439	35.978	0.101	0.021
bourb50-phased_75p	2,000	469,223	45,983,854	61.464	0.059	0.028
bourb27-phased_90p	703	314,903	1,700,4762	54.307	0.034	0.017
bourb50-phased_90p	375	171,777	16,834,146	55.892	0.064	0.032
bourb50-phased_100p	165	76,259	7,473,382	52.078	0.066	0.034
fg37-unphased_75p	1,253	945,618	34,987,866	40.821	0.097	0.023
fg37-unphased_90p	327	183,186	6,777,882	37.828	0.098	0.023
fg37-unphased_100p	0	0	0	0	0	0
fg37-phased_75p	1,699	568,145	42,042,730	59.158	0.052	0.024
fg37-phased_90p	789	323,206	23,917,244	55.206	0.055	0.026
fg37-phased_100p	434	177,304	13,120,496	52.214	0.059	0.03

one species/population representing each sample. We used 'BEAST Model Test' for the site model, a relaxed clock lognormal clock model with default settings, and set the prop prior scale to lognormal. Monophyletic constraints were set for each of the three complexes based on results from the previous MSC analyses. MCMC chain length was set to 100 million generations, storing every 1 million. Two independent analyses were run for each dataset and then and loaded into Tracer v1.7.2 (Rambaut et al. 2018) to check ESS values and verify that convergence of the MCMC chain was reached. Runs were combined using logcombiner with a burnin of 15%.

Results

Global Phylogenetic Reconstruction

As in previous studies (Blaimer et al. 2015, LaPolla et al. 2010, Matos-Maraví et al. 2018, Williams et al. 2020), the monophyly of the genus *Nylanderia* is strongly supported. Within the genus, the ML analysis (Fig. 2) recovered four Australasian lineages (represented by *N.* spp. MJ08, MJ12, 06, and Mal2) diverging as a grade with respect to the rest of the genus, with strong support. The rest of the *Nylanderia* species sampled belong to one of two major clades: (i) an Australasian-Indomalayan-Holarctic clade consisting of an Australasian subclade, a Nearctic subclade, two isolated Indomalayan lineages (represented by *N. emmae* and *N. picta*), and a Palearctic lineage (represented by *N. flavipes*); and (ii) a clade consisting of all other taxa, including species from Australasia, Indomalaya, the Afrotropics, Madagascar, and all species from the Neotropics.

The monophyly and single origin of Neotropical *Nylanderia* is strongly supported. This single Neotropical lineage was recovered as sister to a clade constituted of taxa from the Eastern Hemisphere, including: (i) an Afrotropical/Malagasy clade; (ii) an Indomalaya/Australasia clade that includes the *bourbonica* complex; (iii) a single Palearctic species (*N.* sp. 07) most closely related to the Indomalaya/Australasia clade; and (iv) an additional lineage from the Malagasy region. In contrast, the Australasian *Nylanderia* did not emerge as a single clade; at least six major lineages included Australasian species, including four lineages that were recovered as a sister grade to the rest of the genus, one in the Indomalayan–Holarctic clade, and one including the *bourbonica* complex in the Eastern Hemisphere clade. Across the entire genus, globetrotting species belong to at least nine independent lineages: one from the Afrotropics (*N. jaegerskioeldi*), four from Australasia and Indomalaya (*N. bourbonica, N. obscura, N. vaga*, and *N. glabrior*), three from the Neotropics (*N. steinheili, N. fulva/pubens, N. guatemalensis*), and one from the Palearctic (*N. flavipes*).

Nylanderia bourbonica Complex

Monophyly of the *bourbonica* complex is strongly supported by the global ML phylogeny (Fig. 2), the unphased MSC phylogeny (Fig. 3A), and the phased MSC phylogeny (Fig. 3B). Five morphospecies of those sampled here belong to the *bourbonica* complex: *N. bourbonica*, *N. vaga*, *N. obscura*, *N.* Fij4, and *N.* sp. MJ10; the last two are undescribed species from Fiji and The Philippines, respectively. Both MSC analyses of the *bourbonica* complex strongly support *N. bourbonica* and *N. vaga* as independent, monophyletic species and inclusion of previously unstudied species refutes their previously hypothesized status as sister taxa. *Nylanderia bourbonica* emerged as sister to *N.* sp. Fij4 from Fiji. *Nylanderia vaga* is most closely related to a clade comprised of *N.* sp. MJ10 and *N. obscura*. Seven specimens previously misidentified as *N. bourbonica* instead belong to two other morphospecies closely related to the *bourbonica* complex. This strongly supported clade of morphospecies resembling *N. bourbonica* includes *Nylanderia* spp. 05 and 08 from southern China and Hong Kong (Fig. 3A and B).

Nylanderia fulva Complex

Monophyly of the *fulva* complex is strongly supported by the global ML phylogeny (Fig. 2), the unphased MSC phylogeny (Fig. 4A), and the phased MSC phylogeny (Fig. 4B). In all three analyses, a single *N. pubens* sample from Florida (Ny087), the identity of which was based on male morphological diagnostic characters (Gotzek et al. 2012), is nested within the *N. fulva* supercolony (see below) from the United States with strong support. This specimen is most closely related to specimens identified as *N.* cf. *fulva* from St. Martin in the Caribbean and Lanzarote Taniche in The Canary Islands. In total, the *fulva* complex includes at least four morphospecies sampled for this study: *N. fulva*, *N. pubens*, *N.* sp. Per2 from Peru (undescribed), and *N.* sp. FG05 from French Guiana (undescribed).

Nylanderia guatemalensis Complex

The *guatemalensis* complex, thought to be comprised of *N. guatemalensis* and *N. steinheili*, is nonmonophyletic in the global ML phylogeny (Fig. 2). A strongly supported clade including *N. guatemalensis* and as many as eight morphospecies from South America (*N.* spp. FG02, FG07, FG03, Ven4, Col2, Per3, Bra1, and FG01) is more closely related to the *fulva* complex than to *N. steinheili*. A clade comprised of *N. steinheili* and three morphospecies native/endemic to parts of the Caribbean (*N. coveri*, *N. myops*, and *N.* JKW1) is more distantly related to the rest of the *guatemalensis* complex.

In contrast, the unphased (Fig. 4A) and phased (Fig. 4B) MSC phylogenies both provide strong support for the monophyly of the *guatemalensis* complex, with *N. guatemalensis* most closely related to South American species. *Nylanderia steinheili* is strongly supported as belonging to a clade with two Caribbean morphospecies (*N. coveri* and *N.* sp. JKW1). Collectively, the *guatemalensis* complex is strongly supported as sister to the *fulva* complex. Across all analyses, *N. guatemalensis* and *N. steinheili* are strongly supported as independent, monophyletic species that are not sister taxa.

Molecular Species Delimitation

Results of all molecular species delimitation methods in comparison with a priori morphospecies hypotheses are summarized in Table 3 and Figs. 5 and 6. In most cases, molecular methods split a single morphospecies into multiple species. Splitting was most frequent in the SODA (phased) analysis, followed by SODA (unphased), bPTP, and then STACEY, which delimited species most consistently with our morphospecies hypotheses.

The *bourbonica* complex was estimated to include anywhere between 5 and 25 species (Table 3; Fig. 5). *Nylanderia bourbonica* (morphospecies 'G') was split into two species by both bPTP analyses, eight species by SODA (unphased), 13 species by SODA (phased), and three species by STACEY. *Nylanderia vaga* (morphospecies 'E') was split into two species by both bPTP analyses, split into four species by both SODA analyses, and recovered as a single species by STACEY. Across all methods, *N.* sp. Fij4 (morphospecies 'F') was recovered as a distinct species. *Nylanderia obscura* (morphospecies 'C') and *N.* sp. MJ10 (morphospecies 'D') were supported as two distinct species by PTPh and STACEY but grouped as a single species by PTPml and both SODA analyses. *Nylanderia* sp. 08 (morphospecies 'A') and *N.* sp. 05 (morphospecies 'B') were supported as two distinct

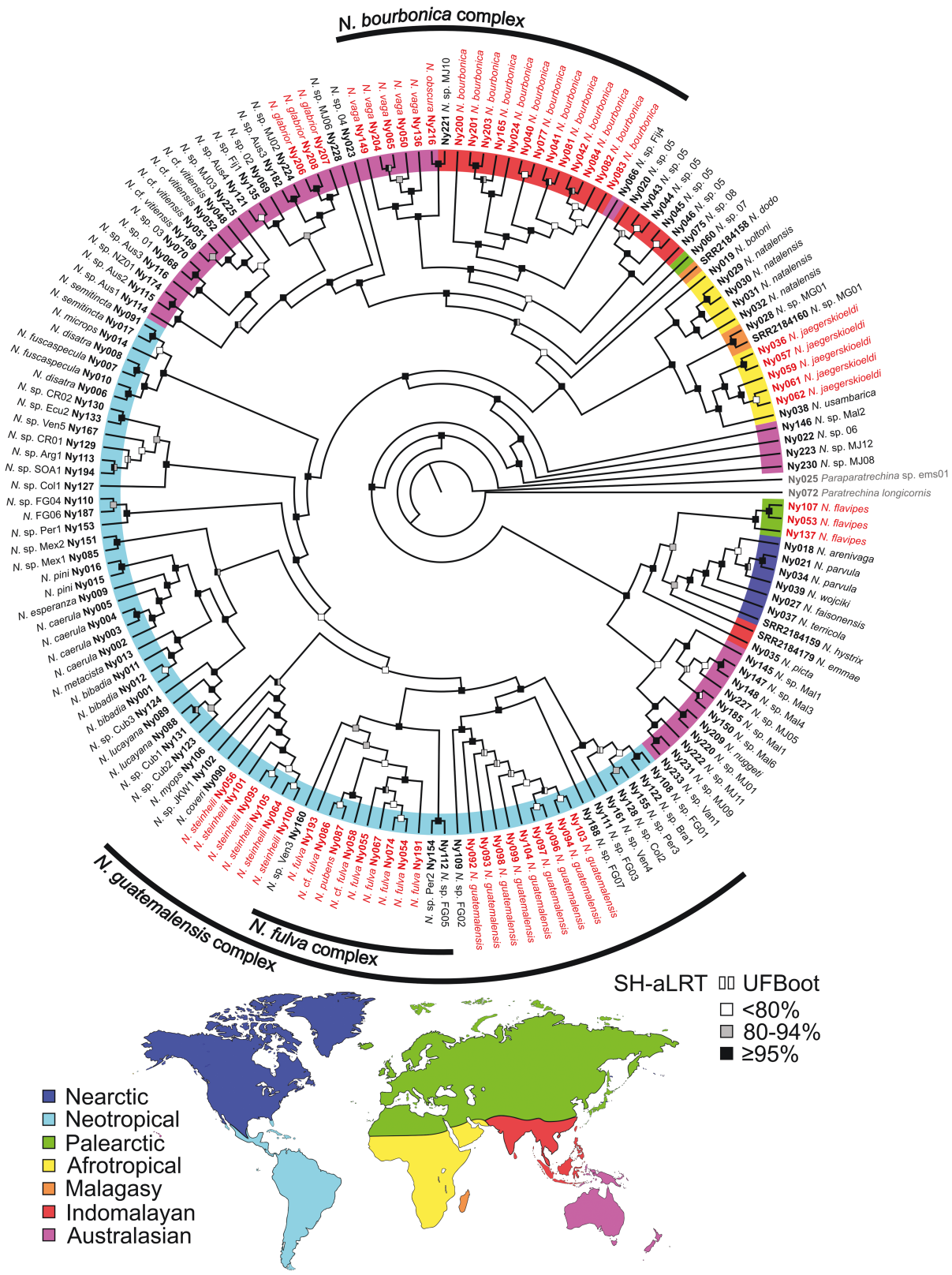

Fig. 2. Circular maximum likelihood phylogeny of *Nylanderia* generated in IQ-TREE 2 from the 80% complete SWSC-partitioned UCE matrix. Biogeography of terminal taxa is based on inferred native ranges and may not indicate collecting locality for known non-native species. Samples indicated with red text are globetrotting species. Nodal support is provided in SH-aLRT (Shimodaira–Hasegawa approximate likelihood ratio test) values on the left side of the node, and UFBoot (ultrafast bootstrap) values on the right side of the node. For both support values, less than 80% is considered weak support, between 80 and 94.9% is considered moderate support, and greater than or equal to 95% is considered strong support.

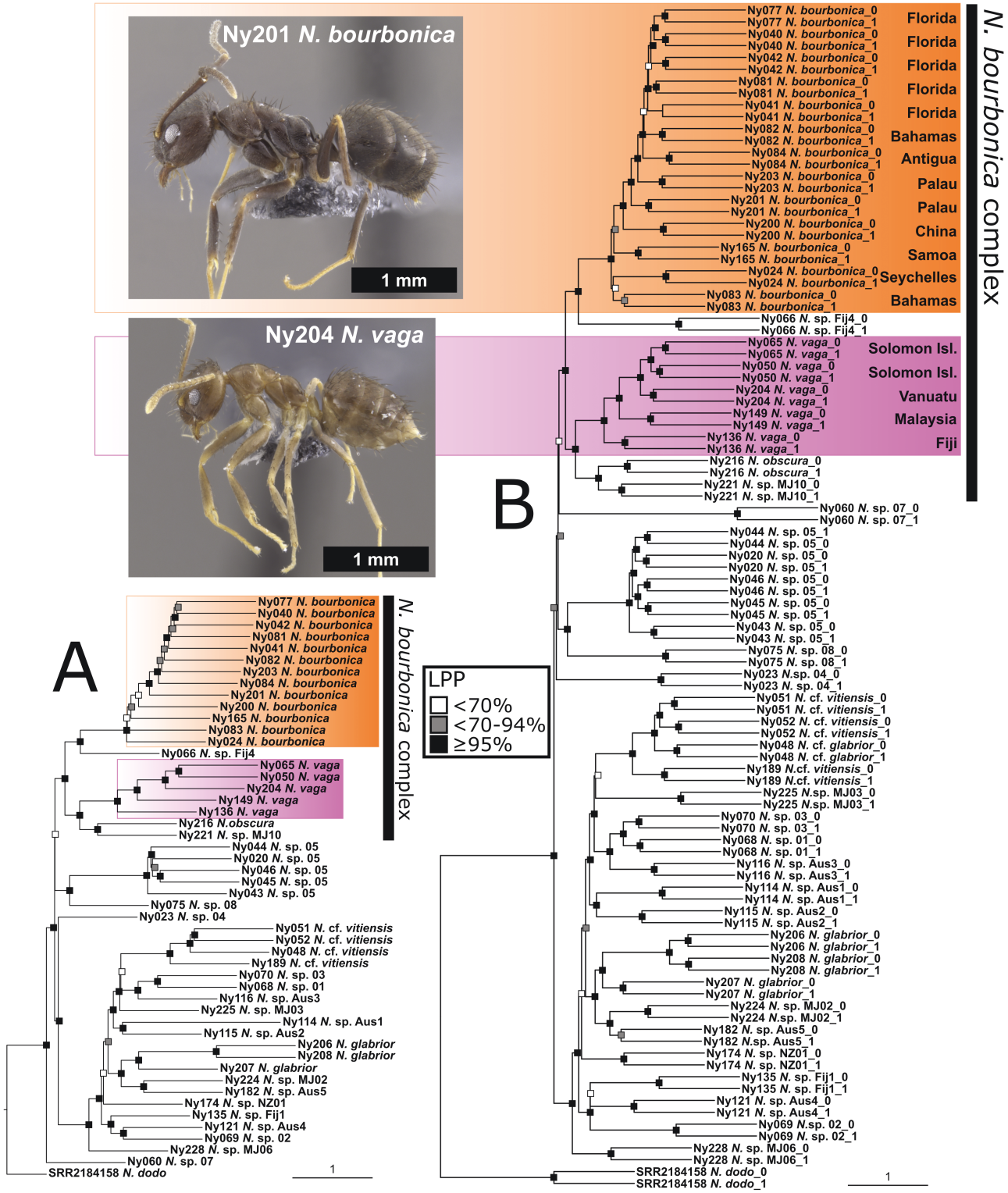

Fig. 3. Multi-species coalescent (MSC) phylogenies of the Australasian/Indomalayan clade including the *bourbonica* complex, generated in ASTRAL-III using 75% complete alignment matrices. (A) (left) represents the unphased MSC analysis and (B) (right) represents the phased MSC analysis. Values on the internal nodes represent local posterior probabilities (LPP), with ≥0.95 indicating strong support and ≤0.75 indicating weak support. Scale bars under phylogenetic trees indicate the number of substitutions per site. Photos of *N. bourbonica* (Ny201) and *N. vaga* (Ny204) workers in profile view are to scale and were taken by Milan Janda.

species by both bPTP analyses and STACEY, but SODA (unphased) and SODA (phased) split the latter into four and five distinct species respectively.

The *fulva* complex was estimated to include anywhere between 4 and 11 species (Table 3; Fig. 6). *Nylanderia fulva* (morphospecies 'N') was supported as a single species by PTPml and STACEY, but

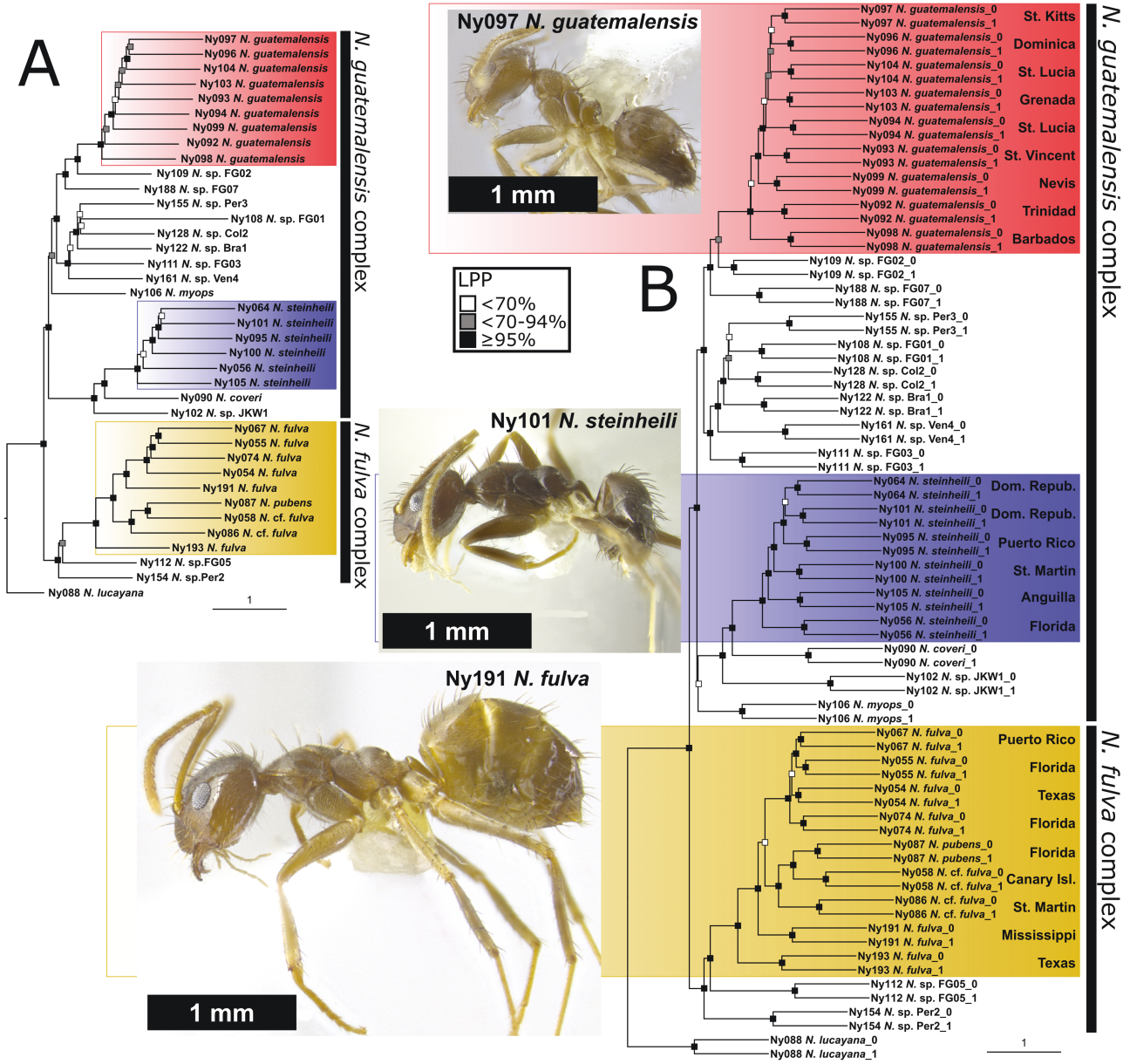

Fig. 4. Multi-species coalescent phylogenies of the *guatemalensis* and *fulva* complexes, generated in ASTRAL-III using 75% complete alignment matrices. (A) (left) represents the unphased MSC analysis and (B) (right) represents the phased MSC analysis. Values on the internal nodes represent local posterior probabilities (LPP), with ≥0.95 indicating strong support and ≤0.75 indicating weak support. Scale bars under phylogenetic trees indicate number of substitutions per site. Photos of *N. guatemalensis* (Ny097), *N. steinheili* (Ny101), and *N. fulva* (Ny191) workers in profile view are to scale and were taken by Brandon Mai.

Table 3. Number of species in each species complex under each delimitation hypothesis considered in this study

Species complex	Morphospecies	PTPh (COI)	PTPml (COI)	SODA (unphased UCEs)	SODA (phased UCEs)	STACEY (phased SNPs)	Decision
bourbonica complex	5	7	8	19	25	9	6
fulva complex	4	8	6	9	11	4	4
guatemalensis complex	12	15	13	18	24	11	11
Total	21	30	27	46	60	24	21

split into four, five, and six species respectively by PTPh, SODA (unphased), and SODA (phased). *Nylanderia pubens* and *N.* cf. *fulva* (together designated as morphospecies 'M') were supported as a single species by STACEY but were split into two species by PTPh and SODA (unphased) and three species by PTPml and SODA (phased). *Nylanderia* sp. Per2 (morphospecies 'O') and *N.* sp. FG05 (morphospecies 'P') were each supported as two distinct species by all analyses.

The *guatemalensis* complex was estimated to include anywhere between 11 and 24 species (Table 3; Fig. 6). *Nylanderia steinheili*

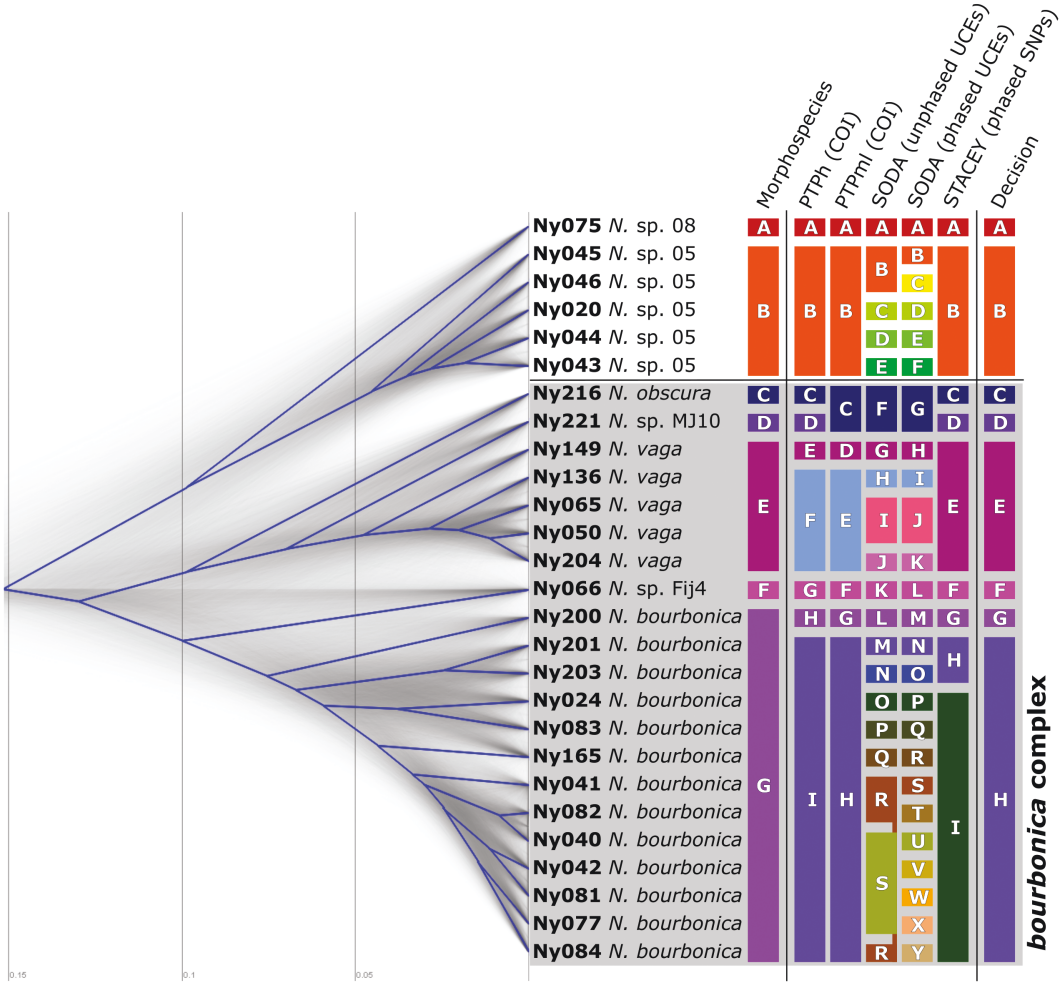

Fig. 5. Species delimitation hypotheses of the *bourbonica* complex. (Left) Cloudogram based on the STACEY analysis using 565 SNPs extracted from the 'bourb27-phased_90p' dataset, with the 'root canal' (blue) summarizing the main features of the tree set. (Right) Summary of species delimitation schemes based on morphospecies sorting and for each analysis performed in this study, with each different colored bar (also labeled with different letters) representing a single species.

(morphospecies 'J') was supported as a single species by STACEY, but split into four species by PTPml, five species by PTPh and SODA (unphased), and six species by SODA (phased). *Nylanderia guatemalensis* (morphospecies 'G') was supported as a single species by STACEY, split into two species by both bPTP analyses, and split into eight and nine species respectively by SODA (unphased) and SODA (phased). Both bPTP methods grouped *N.* sp. FG07 (morphospecies 'I') with *N. guatemalensis*, but all other analyses supported *N.* sp. FG07 as a single species. All analyses supported *N.* sp. FG02 (morphospecies 'H'), *N. coveri* (morphospecies 'K'), and *N.* sp. JKW1 (morphospecies 'L') as distinct species. *Nylanderia* sp. Ven 4 (morphospecies 'A'), *N.* sp. Bra1 (morphospecies 'B'), *N.* sp. FG01 (morphospecies 'C'), *N.* sp. FG03 (morphospecies 'D'), *N.* sp. Per3 (morphospecies 'E'), and *N.* sp. Col2 (morphospecies 'F') were all grouped as a single species by SODA (unphased). Both bPTP analyses grouped *N.* sp. Bra1 and *N.* sp. Per3 as a single species, and SODA (phased) and STACEY both grouped *N.* sp. Per3 and *N.* sp. Col2 as a single species.

Discussion

We reconstructed the largest and most comprehensive molecular phylogeny of *Nylanderia* to date (Fig. 2), with focused taxon sampling from Australasia and the Neotropics (primarily Caribbean) to establish the phylogenetic position of three globetrotting species complexes: (i) *bourbonica* complex, (ii) *fulva* complex, and (iii) *guatemalensis* complex. In total, our study includes approximately 60 undefined or undescribed morphospecies worldwide, most of which likely represent undescribed species: Of these, 26 are from the Neotropics and 29 are from Australasia. This study also includes a sampling of 10 of the 15 known globetrotting *Nylanderia* species (Williams and Lucky 2020), increasing the number of independent globetrotting lineages in the genus from five (Williams et al. 2020) to nine. We established membership of species within each of the three complexes and tested our morphospecies hypotheses against three different molecular methods of species delimitation: bPTP, SODA, and STACEY. These methods were applied across three different types of molecular markers: (i) UCEs before and after allelic phasing, (ii) COI barcodes, and (iii) SNPs randomly subsampled from phased UCE loci.

Molecular Species Delimitation

Across all three species complexes, STACEY analysis of SNPs subsampled from phased UCE loci provided the most conservative species counts (i.e., lowest number of species) and delimited species

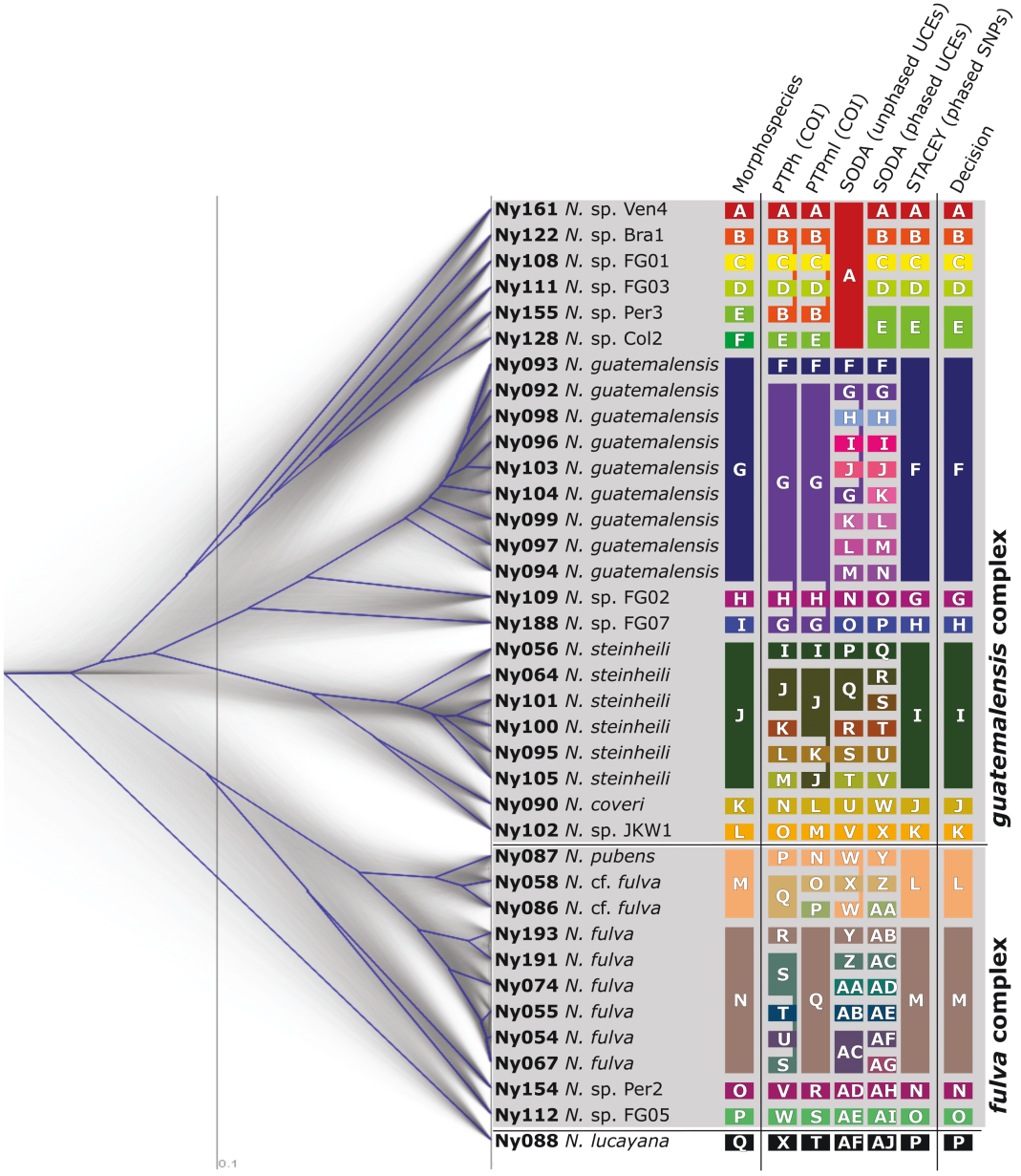

Fig. 6. Species delimitation hypotheses of the *fulva* + *guatemalensis* complexes. (Left) Cloudogram based on the STACEY analysis using 609 SNPs extracted from the 'fg37-phased_90p' dataset, with the 'root canal' (blue) summarizing the main features of the tree set. (Right) Summary of species delimitation schemes based on morphospecies sorting and for each analysis performed in this study, with each different colored bar (also labeled with different letters) representing a single species.

boundaries most consistent with our morphospecies hypotheses (Table 3; Figs. 5 and 6). We expected the bPTP analyses using single-gene COI data to over-split species the most but, surprisingly, the SODA analyses yielded the least conservative species counts (i.e., highest number of species). Moreover, SODA analyses of the phased UCE datasets split species more than analyses of the unphased UCE datasets. Species hypotheses within each complex are discussed in more detail below.

Much like assessing the practical utility of morphological characters, challenges arise in determining which genes or markers are the most useful for delimiting species; e.g., in the controversy surrounding the exclusive use of COI barcodes for species delimitation in a single-gene approach (Carstens et al. 2013, Moritz and Cicero 2004, Zamani et al. 2020). In addition to the choice of markers, careful consideration must be given to the appropriateness of models and statistical software packages for a given set of taxon-specific assumptions and circumstances (Luo et al. 2018). Delimitation of species based exclusively on the multispecies coalescent can result in the over-splitting of species, especially in metapopulations with high geographic structure (Chambers and Hillis 2020, Sukumaran and Knowles 2017).

These issues are apparent in the present study, as using single-gene COI (bPTP) and even the entire UCE datasets (SODA) over-split our morphospecies far more than expected. Interestingly, this over-splitting was even more pronounced when samples were phased and delimited using SODA; probably because each pair of heterozygous sites is treated as a population and the multispecies coalescent interprets each population as a species without accounting for

gene flow across those species (Rabiee and Mirarab 2020). Given these challenges, we take the hypotheses presented by all delimitation methods into account and interpret them conservatively when redrawing the boundaries between species in the *bourbonica*, *fulva*, and *guatemalensis* complexes (discussed below).

Phylogeny and Species Delimitation of *bourbonica* Complex

The *bourbonica* complex was identified as a subclade within a primarily Australasian–Indomalayan clade (Fig. 2), in large part because of increased taxon sampling from Australasia. This Australasian–Indomalayan clade is most closely related to species from Saudi Arabia, the Afrotropics, and Madagascar. Strong support for the monophyly of the *bourbonica* complex is corroborated by both the unphased (Fig. 3A) and phased (3B) MSC analyses. In total, we found that at least five species belong to the *bourbonica* complex: *N. bourbonica*, *N. vaga*, *N. obscura*, *N.* sp. Fij4, and *N.* sp. MJ10 (Fig. 5). However, molecular delimitation methods (Fig. 5) collectively suggest that one *N. bourbonica* sample (Ny200) from Xishuangbanna, China, may be split from *N. bourbonica* as an additional morphologically cryptic species. *Nylanderia bourbonica* (including Ny200) is most closely related to *N.* sp. Fij4, which appears to be an undescribed endemic species from Fiji. Much like *N. vaga*, *N.* sp. Fij4 bears a superficial resemblance to *N. bourbonica* but is much smaller and lighter in overall color. *Nylanderia vaga* is most closely related to *N. obscura* and *N.* sp. MJ10, an undescribed species from The Philippines. Given the number of undescribed taxa in this region, we suspect that further sampling may reveal additional taxa belonging to this group.

Nylanderia bourbonica was historically thought to originate from Southeast Asia (reviewed by Williams and Lucky 2020). However, all other species within the *bourbonica* complex are native to Australasia or, in the case of *N.* sp. MJ10, insular Indomalaya (The Philippines). This suggests that *N. bourbonica* may be native to islands across the Malay Archipelago (spanning islands across Indomalaya and Australasia), as opposed to mainland Southeast Asia. In total, three globetrotting species (*N. bourbonica, N. vaga,* and *N. obscura*) belong to the *bourbonica* complex. A fourth globetrotting species native to Australasia—*N. glabrior*—belongs to the same larger Australasian clade but is not part of the *bourbonica* complex. Collectively, these results lend support to the hypothesis that adaptation to natural, frequently disturbed insular and coastal habitats across the Indo-Pacific has facilitated the spread of these four globetrotting species since the advent of human trade (Matos-Maraví et al. 2018).

Close relatives to the *bourbonica* complex include a clade of species from mainland Southeast Asia, specifically from southern China and Hong Kong. This clade includes *N.* spp. 05 and 08, which were initially identified by local collectors as *N. bourbonica* or *N.* cf. *bourbonica*. These two species bear a strong resemblance in color and size to *N. bourbonica* with some notable distinctions. *Nylanderia* spp. 05 and 08 are more gracile in profile view and display varying degrees of a powder blue cuticular microstructural iridescence. These two species may be undescribed species, but this cannot be confirmed until specimens are compared to Asian type material in the context of a broader monographic revision of the Southeast Asian fauna. The gracile appearance and cuticular iridescence also cause these two species to strongly resemble another globetrotting species, *Paratrechina longicornis* (Latreille), which can further complicate field identification. The prevalence of these species in southern China and Hong Kong may have caused ant collectors to overestimate the prevalence of *N. bourbonica* in Southeast Asia, which may be why *N. bourbonica* has historically been considered a primarily Indomalayan species.

Phylogeny and Species Delimitation of *fulva* Complex

The relationships among species in the *fulva* complex have been the subject of uncertainty since *N. fulva* was introduced into North America where the species, regardless of its name, has become an economically costly invasive species. Here we provide the strongest support to date for the monophyly of the *fulva* complex, corroborated by the ML analysis (Fig. 2), as well as both the unphased (Fig. 4A) and phased (Fig. 4B) MSC analyses. In total, at least four species sampled in this study belong to the *fulva* complex, two of which are potentially undescribed: *N. fulva*, *N. pubens*, *N.* sp. Per2, and *N.* sp. FG05 (Fig. 6). Other species, yet unsampled, may also be part of this complex.

Confusion over the identity of *N. fulva* and *N. pubens* was brought to light in the wake of a newly introduced and problematic population of *Nylanderia* ants around Houston, Texas in 2002. The identity of *N. fulva* populations in the United States was first confirmed by Gotzek et al. (2012), who cited diagnostic male genitalic characters (first discovered by Trager (1984)) useful for distinguishing *N. fulva* from *N. pubens*: A male collected by John Mangold in 1994 from Miami Lakes, Florida, was imaged and used to illustrate these differences. LaPolla and Kallal (2019) also explored differences in male genitalia between the two species via dissection, providing illustrations to compare the penial sclerites and nineth sternites of ten *N. fulva* specimens and one *N. pubens* specimen: They found that the latter bears more anteroposteriorly elongate penial sclerites and a longer, more shield-like nineth sternite than the former. The authors used these morphological differences to support the identification of this newly introduced population of *N. fulva* and distinguish it from the previously established (and behaviorally less problematic) *N. pubens*. However clear the distinction in the specimens examined, the extent of intra- and interspecific variation expressed within and between males of these two species may have been underestimated due to limited taxon sampling.

One of the most widely accepted paradigms regarding sexually reproducing organisms is that sexually selected characters evolve rapidly in comparison to other traits and thus that genitalic morphology tends to vary discretely between species (Langerhans et al. 2016). Though this may be a broadly observable pattern, the evidence for its supposed utility in specific taxa may depend entirely on an inherently biased taxonomic literature in which genitalic shape characters are assumed a priori to be species-specific (Huber 2003). Moving forward, the utility of male genitalic characters for *Nylanderia* species delimitation—especially those of *N. pubens* and *N. fulva*—must be rigorously tested by quantitatively evaluating the full extent of expressed intra- and interspecific variation.

This study is the first to date to include a genetic representation of a *N. pubens* specimen (Ny087) from the same nest series as the male collected from Miami Lakes (see Gotzek et al. 2012)—a specimen that is the morphological linchpin to the current identity of *N. pubens*—alongside a broad representation of introduced *N. fulva* populations across the United States ranging between Texas and Florida. In determining how these species should be delimited, we considered the topologies recovered in our ML and MSC analyses, as well as results from five different species delimitation schemes. The genus-wide phylogeny (Fig. 2) and the species tree phylogenies (Fig.

4) all strongly indicate that this series of *N. pubens* specimens, along with *N.* cf. *fulva* specimens from the Caribbean and the Canary Islands, are nested within the United States *N. fulva* supercolony and would therefore render *N. fulva* as paraphyletic.

These results initially appear to suggest that *N. pubens* should therefore be synonymized with *N. fulva*, lest the latter be considered a paraphyletic species. The implications of synonymy in this case would entail a profound rewriting of the introduction history of this species: The earliest specimen records identified as *N. pubens* in Florida date from 1953 (Deyrup et al. 2000, Trager 1984, Williams and Lucky 2020). If these are indeed *N. fulva*, they document an introduction long before the establishment of *N. fulva* in Texas in 2002, as the full record of specimens identified as *N. pubens* in Florida would represent historical introductions of *N. fulva* long before the establishment of *N. fulva* in Texas in 2002. A large series of *N. pubens* specimens collected by Theodore Pergande in the late 1800s from greenhouses in Washington, D.C.—specimens studied by Emery (1893), who first described *N. pubens* from St. Vincent as a subspecies of *N. fulva*, and later examined by Trager (1984), who elevated *N. pubens* to full species status—would further suggest that *N. fulva* may have been an accidental traveler in human cargo at least a century prior to what has been considered the earliest introduction in North America.

In contrast, molecular species delimitation methods present the strongest evidence to date that *N. fulva* and *N. pubens* are not synonymous, but indeed two distinct species, albeit perhaps with some degree of introgression. ML analysis of single-gene COI barcodes (Supp Fig. S2 [online only]) and species delimitation schemes using all different markers tested in this study (COI barcodes, SNPs, unphased UCEs, and phased UCEs; Fig. 6) all support *N. pubens* and *N. fulva* as two distinct, monophyletic sister species. Given these results and the differences previously described in male genitalia, we maintain that the conservative approach to delimitation would be to recognize *N. pubens* and *N. fulva* as two distinct species. Given that our sampling primarily included representatives of these species from the United States and the Caribbean, more sampling from South America is necessary to determine how many species belong to the *N. fulva* complex, in addition to how five subspecies of *N. fulva* (reviewed in Williams and Lucky 2020) should be treated.

Phylogeny and Species Delimitation of *guatemalensis* Complex

Monophyly of the *guatemalensis* complex is strongly supported by the unphased (Fig. 4A) and phased (Fig. 4B) MSC analyses. In total, we found that at least 11 species belong to the *guatemalensis* complex: *N. guatemalensis, N. steinheili, N. coveri, N.* JKW1, *N.* sp. FG01, *N.* sp. FG02 *N.* sp. FG03, *N.* sp. FG07, *N.* sp. Ven4, *N.* sp. Bra1, and *N.* sp. Per3/Col2 (Fig. 6). Interestingly, *N. guatemalensis* and *N. steinheili*, despite their morphological similarity, are not sister species and appear to be relatively distantly related within the complex: *N. guatemalensis* is most closely related to mainland Neotropical species, while *N. steinheili* is most closely related to insular, Caribbean species.

The unphased and phased MSC analyses provide strong support for a clade that includes *N. guatemalensis* and eight South American morphospecies: *N.* spp. FG01, FG02, FG03, FG07, Per3, Col2, Bra1, and Ven4. *Nylanderia guatemalensis* is strongly supported as most closely related to *N.* sp. FG02 from French Guiana. *Nylanderia steinheili* is most closely related to *N. coveri*, and the two together are most closely related to *N.* sp. JKW1. These results have implications for the native ranges of *N. guatemalensis* and *N. steinheili*. Based on the above relationships, *N. guatemalensis* is likely of mainland South American origin and *N. steinheili* is likely of Caribbean origin.

Challenges in Species-Level Taxonomy

The genus *Nylanderia* presents an opportunity to test new species delimitation approaches—especially a combination of those using different types of molecular data—because it has long challenged taxonomists due to its high diversity, cosmopolitan distribution, and few distinguishing morphological characters at the species level. Though the toolkit available to study speciation and species delimitation has grown over the past few centuries, some of the same challenges faced by early taxonomists persist today. One of the greatest obstacles is limited taxon sampling, especially of rarer species. Limited sampling can lead to 'red herring' morphological diagnostic characters: those which initially appear as discrete forms but become dubious as taxon sampling increases and continuous variation is observed. An account of intra- and interspecific limits of trait variation—gathered through broad, population-level sampling—is a necessary step to overcome the impediments imposed by superficially identical species.

Given the above challenges, species delimitation of morphologically cryptic taxa requires holistic interpretation of the phenome and the genome. We argue that the optimal approach for resolving taxonomic impediments includes: (i) taxon sampling that is broad with regard to biogeographical range, genetics, morphology, behavior, and ecology; (ii) integrative and quantitative methods that are capable of capturing and analyzing the maximal extent of expressed intra- and interspecific variation; and (iii) meticulous evaluation of characters to determine those that offer the most utility for species delimitation, and which are most practical for species diagnosis. Furthermore, this approach must be accentuated with a robust model-comparison approach that considers the assumptions made under each model—especially those that promise to delimit species using molecular data—and interprets the results conservatively.

The results of this study inform alpha-taxonomic decisions for species belonging to three major globetrotting *Nylanderia* species complexes, which should be taken in full context through regional monographic revisions of the Neotropical, Southeast Asian, and Australasian faunas. Potentially hundreds more species belonging to this genus still await description, especially in these three biogeographic regions of the world. Overcoming the taxonomic impediment in morphologically challenging species—both in *Nylanderia* and in other taxa—is necessary now more than ever to recognize and manage the growing number of human-dispersed invasive species and decelerate their increasing negative global influence. Simultaneously, formal delimitation of previously undescribed species must remain a priority, as recognition of biodiversity is the first essential step in conservation.

Supplementary Data

Supplementary data are available at *Insect Systematics and Diversity* online.

Acknowledgments

We express our gratitude to David Blackburn, Akito Kawahara, Faith Oi, and three anonymous reviewers for feedback and advice in the preparation of this manuscript. We also thank Karen Neves and Bernardo Santos for their technical support during UCE library preparation. The following institutions and individuals loaned or donated specimens used in this study: Archbold Biological Station (Mark Deyrup), Florida State Collection of Arthropods (Elijah Talamas), Ward Lab at UC Davis (Phil Ward), David Cross, Ben Gouchnour, Benoit Guénard, Roger Ho Lee, Milan Janda, David Oi, Eric Roldan, Eli Sarnat, and Dan Suiter. Computational resources were provided by University

of Florida Research Computing (http://researchcomputing.ufl.edu). J.L.W. was funded by a University of Florida College of Agricultural and Life Sciences (UF CALS) Graduate Fellowship and in part by National Science Foundation (NSF) grant Division of Environmental Biology (DEB) 2026772. A.L. was supported in part by NSF grant DEB 2026772 and in part by USDA National Institute of Food and Agriculture, McIntire Stennis project 1011529. T.R.S. was supported by NSF grants DEB 1654829 and DEB 1927161. J.S.L. was supported in part by NSF DEB 0743542.

Author Contributions

J.L.W.: Conceptualization; Data curation; Formal analysis; Investigation; Methodology; Validation; Visualization; Writing—original draft; Writing—review & editing. Y.M.Z.: Formal analysis; Methodology; Validation; Writing—review & editing. J.S.L.: Resources; Validation; Writing—review & editing. T.R.S.: Resources; Supervision; Validation; Writing—review & editing. A.L.: Resources; Supervision; Validation; Writing—review & editing.

References Cited

Andermann, T., A. M. Fernandes, U. Olsson, M. Töpel, B. Pfeil, B. Oxelman, A. Aleixo, B. C. Faircloth, and A. Antonelli. 2019. Allele phasing greatly improves the phylogenetic utility of ultraconserved elements. Syst. Biol. 68: 32–46.

Bankevich, A., S. Nurk, D. Antipov, A. A. Gurevich, M. Dvorkin, A. S. Kulikov, V. M. Lesin, S. I. Nikolenko, S. Pham, A. D. Prjibelski, et al. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J. Comput. Biol. 19: 455–477.

Barraclough, T. G. 2019. The evolutionary biology of species. Oxford University Press.

Bastos, A. D., D. Nair, P. J. Taylor, H. Brettschneider, F. Kirsten, E. Mostert, E. von Maltitz, J. M. Lamb, P. van Hooft, S. R. Belmain, et al. 2011. Genetic monitoring detects an overlooked cryptic species and reveals the diversity and distribution of three invasive Rattus congeners in South Africa. BMC Genet. 12: 26.

Beheregaray, L. B., and A. Caccone. 2007. Cryptic biodiversity in a changing world. J. Biol. 6: 9.

Blaimer, B. B., S. G. Brady, T. R. Schultz, M. W. Lloyd, B. L. Fisher, and P. S. Ward. 2015. Phylogenomic methods outperform traditional multilocus approaches in resolving deep evolutionary history: a case study of formicine ants. BMC Evol. Biol. 15: 271.

Borowiec, M. L. 2016. AMAS: a fast tool for alignment manipulation and computing of summary statistics. PeerJ. 4: e1660.

Bossert, S., and B. N. Danforth. 2018. On the universality of target-enrichment baits for phylogenomic research. Methods Ecol. Evol. 2018: 1–8.

Bouckaert, R., J. Heled, D. Kühnert, T. Vaughan, C. H. Wu, D. Xie, M. A. Suchard, A. Rambaut, and A. J. Drummond. 2014. BEAST 2: a software platform for Bayesian evolutionary analysis. PLoS Comput. Biol. 10: e1003537.

Boudinot, B. E. 2018. A general theory of genital homologies for the Hexapoda (Pancrustacea) derived from skeletomuscular correspondences, with emphasis on the Endopterygota. Arthropod Struct. Dev. 47: 563–613.

Branstetter, M. G., and J. T. Longino. 2019. Ultra-conserved element phylogenomics of New World Ponera (Hymenoptera: Formicidae) illuminates the origin and phylogeographic history of the endemic exotic ant *Ponera exotica*. Insect Syst. Divers. 3: 1–13.

Branstetter, M. G., J. T. Longino, P. S. Ward, and B. C. Faircloth. 2017. Enriching the ant tree of life: enhanced UCE bait set for genome-scale phylogenetics of ants and other Hymenoptera. Methods Ecol. Evol. 8: 768–776.

Burbrink, F. T., H. Yao, M. Ingrasci, R. W. Bryson, Jr, T. J. Guiher, and S. Ruane. 2011. Speciation at the Mogollon Rim in the Arizona Mountain Kingsnake (*Lampropeltis pyromelana*). Mol. Phylogenet. Evol. 60: 445–454.

Carstens, B. C., T. A. Pelletier, N. M. Reid, and J. D. Satler. 2013. How to fail at species delimitation. Mol. Ecol. 22: 4369–4383.

Castresana, J. 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol. Biol. Evol. 17: 540–552.

Chambers, E. A., and D. M. Hillis. 2020. The multispecies coalescent oversplits species in the case of geographically widespread taxa. Syst. Biol. 69: 184–193.

Chaplin, K., J. Sumner. C. A. Hipsley, and J. Melville. 2019. An integrative approach using phylogenomics and high-resolution X-ray computed tomography for species delimitation in cryptic taxa. Syst. Biol. 0: 1–14.

Creighton, W. S. 1950. The ants of North America. Bull. Museum Comp. Zool. 104: 1–585.

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

Derkarabetian, S., S. Castillo, P. K. Koo, S. Ovchinnikov, and M. Hedin. 2019. A demonstration of unsupervised machine learning in species delimitation. Mol. Phylogenet. Evol. 139: 106562.

Deyrup, M. 2016. Ants of Florida: identification and natural history. CRC Press, Boca Raton, Florida, USA.

Deyrup, M., L. Davis, and S. Cover. 2000. Exotic ants in Florida. Trans. Am. Entomol. Soc. 126: 293–326.

Domingos, F. M., R. J. Bosque, J. Cassimiro, G. R. Colli, M. T. Rodrigues, M. G. Santos, and L. B. Beheregaray. 2014. Out of the deep: cryptic speciation in a Neotropical gecko (Squamata, Phyllodactylidae) revealed by species delimitation methods. Mol. Phylogenet. Evol. 80: 113–124.

Emery, C. 1893. Beiträge zur Kenntniss der nordamerikanischen Ameisenfauna. Zoologische Jahrbücher. Abteilung für Systematik, Geographie und Biologie der Tiere. 7: 633–682.

Faircloth, B. C. 2015. PHYLUCE is a software package for the analysis of conserved genomic loci. Bioinformatics. 32: 786–788.

Fišer, C., C. T. Robinson, and F. Malard. 2018. Cryptic species as a window into the paradigm shift of the species concept. Mol. Ecol. 27: 613–635.

Fujisawa, T., and T. G. Barraclough. 2013. Delimiting species using single-locus data and the Generalized Mixed Yule Coalescent approach: a revised method and evaluation on simulated data sets. Syst. Biol. 62: 707–724.

Gotzek, D., S. G. Brady, R. J. Kallal, and J. S. LaPolla. 2012. The importance of using multiple approaches for identifying emerging invasive species: the case of the Rasberry Crazy Ant in the United States. PLoS One. 7: e45314.

Gueuning, M., J. Frey., and C. Praz. 2020. Ultraconserved yet informative for species delimitation: UCEs resolve long-standing systematic enigma in Central European bees. Mol. Ecol. 1–22.

Guindon, S., J. F. Dufayard, V. Lefort, M. Anisimova, W. Hordijk, and O. Gascuel. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst. Biol. 59: 307–321.

Hoang, D. T., O. Chernomor, A. Von. Haeseler, B. Q. Minh, and L. S. Vinh. 2017. UFBoot2: Improving the ultrafast bootstrap approximation. Mol. Biol. Evol. 35: 518–522.

Huber, B. A. 2003. Rapid evolution and species-specificity of arthropod genitalia: Fact or artifact?. Org. Divers. Evol. 3: 63–71.

Janicki, J., N. Narula, M. Ziegler, B. Guénard, E. P. Economo. 2016. Visualizing and interacting with large-volume biodiversity data using client-server web-mapping applications: the design and implementation of antmaps.org. Ecol. Inform. 32: 185–193.

Ješovnik, A., J. Sosa-Calvo, M. W. Lloyd, M. G. Branstetter, F. Fernández, and T. R. Schultz. 2017. Phylogenomic species delimitation and host-symbiont coevolution in the fungus-farming ant genus Sericomyrmex Mayr (Hymenoptera: Formicidae): ultraconserved elements (UCEs) resolve a recent radiation. Syst. Entomol. 42: 523–542.

Jones, G. 2017. Algorithmic improvements to species delimitation and phylogeny estimation under the multispecies coalescent. J. Math. Biol. 74: 447–467.

Junier, T., and E. M. Zdobnov. 2010. The Newick utilities: high-throughput phylogenetic tree processing in the UNIX shell. Bioinformatics 26: 1669–1670.

Kallal, R. J., and J. S. LaPolla. 2012. Monograph of Nylanderia (Hymenoptera: Formicidae) of the world, part II: Nylanderia in the Nearctic. Zootaxa 64: 1–64.

Krushelnycky, P. D., L. L. Loope, N. J. Reimer. 2005. The ecology, policy, and management of ants in Hawaii. Proc. Hawaiian Entomol. Soc. 37: 1–25.

Lanfear, R., P. B. Frandsen, A. M. Wright, T. Senfeld, and B. Calcott. 2017. PartitionFinder 2: new methods for selecting partitioned models of

evolution for molecular and morphological phylogenetic analyses. Mol. Biol. Evol. 34: 772–773.

Langerhans, R. B., C. M. Anderson, and J. L. Heinen-Kay. 2016. Causes and consequences of genital evolution. Integr. Comp. Biol. 56: 741–751.

LaPolla, J. S., and R. J. Kallal. 2019. Nylanderia of the world part III: Nylanderia in the West Indies. Zootaxa. 4658: zootaxa.4658.3.1.

LaPolla, J. S., Brady, S. G., and S. O. Shattuck. 2010. Phylogeny and taxonomy of the Prenolepis genus-group of ants (Hymenoptera: Formicidae). Syst. Entomol. 35: 118–131.

LaPolla, J. S., P. G. Hawkes, B. L. Fisher. 2011. Monograph of Nylanderia (Hymenoptera: Formicidae) of the world, part I: Nylanderia in the Afrotropics. Zootaxa. 3110: 10–36.

LaPolla, J. S., R. J. Kallal, and S. G. Brady. 2012. A new ant genus from the Greater Antilles and Central America, Zatania (Hymenoptera: Formicidae), exemplifies the utility of male and molecular character systems. Syst. Entomol. 37: 200–214.

Longino, J. T., and M. G. Branstetter. 2020. Phylogenomic species delimitation, taxonomy, and 'bird guide' identification for the Neotropical ant genus Rasopone (Hymenoptera: Formicidae). Mol. Phylogenet. Phylogenom. Phylogeogr. 4: 1–33.

Luo, A., C. Ling, S. Y. W. Ho, and C. D. Zhu. 2018. Comparison of methods for molecular species delimitation across a range of speciation scenarios. Syst. Biol. 67: 830–846.

Matos-Maraví, P., R. M. Clouse, E. M. Sarnat, E. P. Economo, J. S. LaPolla, M. Borovanska, C. Rabeling, J. E. Czekanski-Moir, F. Latumahina, E. O. Wilson, and M. Janda. 2018. An ant genus-group (Prenolepis) illuminates the biogeography and drivers of insect diversification in the Indo-Pacific. Mol. Phylogenet. Evol. 123: 16–25.

Mayr, E. 1942. Systematics and the origin of species, from the viewpoint of a zoologist. Harvard University Press.

Minh, B. Q., H. A. Schmidt, O. Chernomor, D. Schrempf, M. D. Woodhams, A. von Haeseler, and R. Lanfear. 2020. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. Mol. Biol. Evol. 37: 1530–1534.

Morais, P., and M. Reichard. 2018. Cryptic invasions: a review. Sci. Total Environ. 613–614: 1438–1448.

Moritz, C., and C. Cicero. 2004. DNA barcoding: promise and pitfalls. PLoS Biol. 2: e354.

Prebus, M. M. 2021. Phylogenomic species delimitation in the ants of the *Temnothorax salvini* group (Hymenoptera: Formicidae): an integrative approach. Syst. Entomol. https://doi.org/10.1111/syen.12463

Rabiee, M., and S. Mirarab. 2020. Forcing external constraints on tree inference using ASTRAL. BMC Genom. 21: 218.

Rambaut, A., A. J. Drummond, D. Xie, G. Baele, and M. A. Suchard. 2018. Posterior summarization in Bayesian phylogenetics using tracer 1.7. Syst. Biol. 67: 901–904.

Sayyari, E., and S. Mirarab. 2016. Fast coalescent-based computation of local branch support from quartet frequencies. Mol. Biol. Evol. 33: 1654–1668.

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

Sukumaran, J., and L. L. Knowles. 2017. Multispecies coalescent delimits structure, not species. Proc. Natl. Acad. Sci. U. S. A. 114: 1607–1612.

Tagliacollo, V. A., and R. Lanfear. 2018. Estimating improved partitioning schemes for ultraconserved elements. Mol. Biol. Evol. 35: 1798–1811.

Trager, J. 1984. A revision of the genus Paratrechina (Hymenoptera: Formicidae) of the continental United States. Sociobiology. 9: 51–162.

Wagner, H. C., W. Arthofer, B. Seifert, C. Muster, F. M. Steiner, and B. C. Schlick-Steiner. 2017. Light at the end of the tunnel: Integrative taxonomy delimits cryptic species in the *Tetramorium caespitum* complex (Hymenoptera: Formicidae). Myrmecol. News. 25: 95–129.

Wetterer, J. K. 2011. Worldwide spread of the tropical fire ant, *Solenopsis geminata* (Hymenoptera: Formicidae). Myrmecol. News. 14: 21–35.

Wetterer, J. K. 2013. Exotic spread of *Solenopsis invicta* Buren (Hymenoptera:Formicidae) beyond North America. Sociobiology. 60: 50–55.

Williams, J. L., and J. S. Lapolla. 2016. Taxonomic revision and phylogeny of the ant genus Prenolepis (Hymenoptera: Formicidae). Zootaxa. 4200: 201–258.

Williams, J. L., and A. Lucky. 2020. Non-native and invasive Nylanderia crazy ants (Hymenoptera: Formicidae) of the world: Integrating genomics to enhance taxonomic preparedness. Ann. Entomol. Soc. Am. 113: 318–336.

Williams, J. L., Y. M. Zhang, M. W. Lloyd, J. S. LaPolla, T. R. Schultz, and A. Lucky. 2020. Global domination by crazy ants: phylogenomics reveals biogeographical history and invasive species relationships in the genus Nylanderia (Hymenoptera: Formicidae). Syst. Entomol. 730–744.

Yang, Z. 2015. The BPP program for species tree estimation and species delimitation. Curr. Zool. 61: 854–865.

Zamani, A., V. Vahtera, I. E. Sääksjärvi, and M. D. Scherz. 2020. The omission of critical data in the pursuit of 'revolutionary' methods to accelerate the description of species. Syst. Entomol. https://doi.org/10.1111/syen.12444

Zhang, J., P. Kapli, P. Pavlidis, and A. Stamatakis. 2013. A general species delimitation method with applications to phylogenetic placements. Bioinformatics. 29: 2869–2876. doi:10.1093/bioinformatics/btt499.

Zhang, C., M. Rabiee, E. Sayyari, and S. Mirarab. 2018. ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. BMC Bioinform. 19: 153.

Zhang, Y. M., J. L. Williams, and A. Lucky. 2019. Understanding UCEs: a comprehensive primer on using ultraconserved elements for arthropod phylogenomics. Insect Syst. Divers. 3. https://doi.org/10.1093/isd/ixz016