Convergent Evolution of the Army Ant Syndrome and Congruence in Big-Data Phylogenetics

Marek L. Borowiec[1,2,3,*]

[1]*Department of Entomology, Plant Pathology and Nematology, 875 Perimeter Drive, University of Idaho, Moscow, ID 83844, USA;* [2]*School of Life Sciences, Social Insect Research Group, Arizona State University, Tempe, AZ 85287, USA; and* [3]*Department of Entomology and Nematology, One Shields Avenue, University of California at Davis, Davis, CA 95616, USA*
[*]*Correspondence to be sent to: Department of Entomology, Plant Pathology and Nematology, 875 Perimeter Drive, University of Idaho, Moscow, ID 83844, USA;*
E-mail: mborowiec@uidaho.edu.

Abstract.—Army ants are a charismatic group of organisms characterized by a suite of morphological and behavioral adaptations that includes obligate collective foraging, frequent colony relocation, and highly specialized wingless queens. This army ant syndrome underlies the ecological success of army ants and its evolution has been the subject of considerable debate. It has been argued to have arisen once or multiple times within the ant subfamily Dorylinae. To address this question in a phylogenetic framework I generated data from 2166 loci and a comprehensive taxon sampling representing all 27 genera and 155 or approximately 22% of doryline species. Most analyses show strong support for convergent evolution of the army ant syndrome in the Old and New World but certain relationships are sensitive to analytics. I examine the signal present in this data set and find that conflict is diminished when only loci less likely to violate common phylogenetic model assumptions are considered. I also provide a temporal and spatial context for doryline evolution with time-calibrated, biogeographic, and diversification rate shift analyses. This study shows that the army ant syndrome is both an example of remarkable convergence of a complex set of traits and a case of long-term evolutionary stasis. The sensitivity of some of the phylogenetic results underscores the need for cautious analysis of phylogenomic data and calls for more efficient algorithms employing better-fitting models of molecular evolution. Congruence among results obtained using different analytics may be used to assess robustness in phylogenomics. [Dorylinae; Formicidae; k-means; partitioning; phylogenomics; ultraconserved elements.]

Army ants (Fig. 1) are a charismatic group of organisms that inspire research in such disparate fields as behavioral ecology (Schöning et al., 2005), biodiversity conservation (Peters et al., 2008), and computational biology (Garnier et al., 2013). These ants are distributed throughout warm temperate and tropical regions of the world and belong to a more inclusive clade known as the subfamily Dorylinae (Brady et al., 2014). They are characterized by a suite of morphological and behavioral adaptations, together dubbed the army ant syndrome (Brady, 2003). This syndrome includes obligate collective foraging, frequent colony relocation, and highly specialized wingless queens. In contrast to many other ant species, army ants never forage individually. Nests of army ants are temporary and in some species colonies undergo cycles of stationary and nomadic phases (Schneirla, 1945). Unable to fly, the queens must disperse on foot and are adapted to producing enormous quantities of brood (Raignier et al., 1955). Other peculiarities of army ants include colony reproduction by fission and highly derived male morphology. No army ant species are known to lack any of the components of the syndrome (Gotwald, 1995). Because of its antiquity and persistence, the army ant syndrome has been cited as an example of remarkable long-term evolutionary stasis (Brady, 2003). Several distantly related lineages of ants evolved one or more of the components of the army ant syndrome but did not reach the degree of social complexity or ecological

dominance of the "true army ants" in the Dorylinae (Kronauer, 2009).

A major question of army ant biology is whether the army ant syndrome originated only once or if it arose independently in New World and Old World army ants (Kronauer, 2009). The answer requires a robust phylogenetic framework, but Dorylinae are an example of an ancient rapid radiation and elucidating its phylogeny has proven difficult (Brady et al., 2014). Before the advent of quantitative phylogenetic methods, the army ants were thought to have arisen more than once within the subfamily (Brown, 1975; Gotwald, 1979; Bolton, 1990). This view was based on the observation that army ants are poor at dispersal and on the assumption that they evolved after the breakup of Gondwana around 100 Ma ago. More recent studies (Baroni Urbani et al., 1992; Brady, 2003), however, reported monophyly of army ant lineages, even though statistical support for this grouping was often low (Brady et al., 2014). Furthermore, the most recent divergence time estimates suggest that army ants indeed originated after the breakup of Gondwana, implying either independent origins or long-distance dispersal (Kronauer, 2009).

Here I reassess the phylogeny of the Dorylinae including New World and Old World army ants. I use a large phylogenomic data set and taxon sampling increased two-fold compared with a previous phylogeny (Brady et al., 2014), including 155 taxa representing more than 22% of described doryline species diversity and
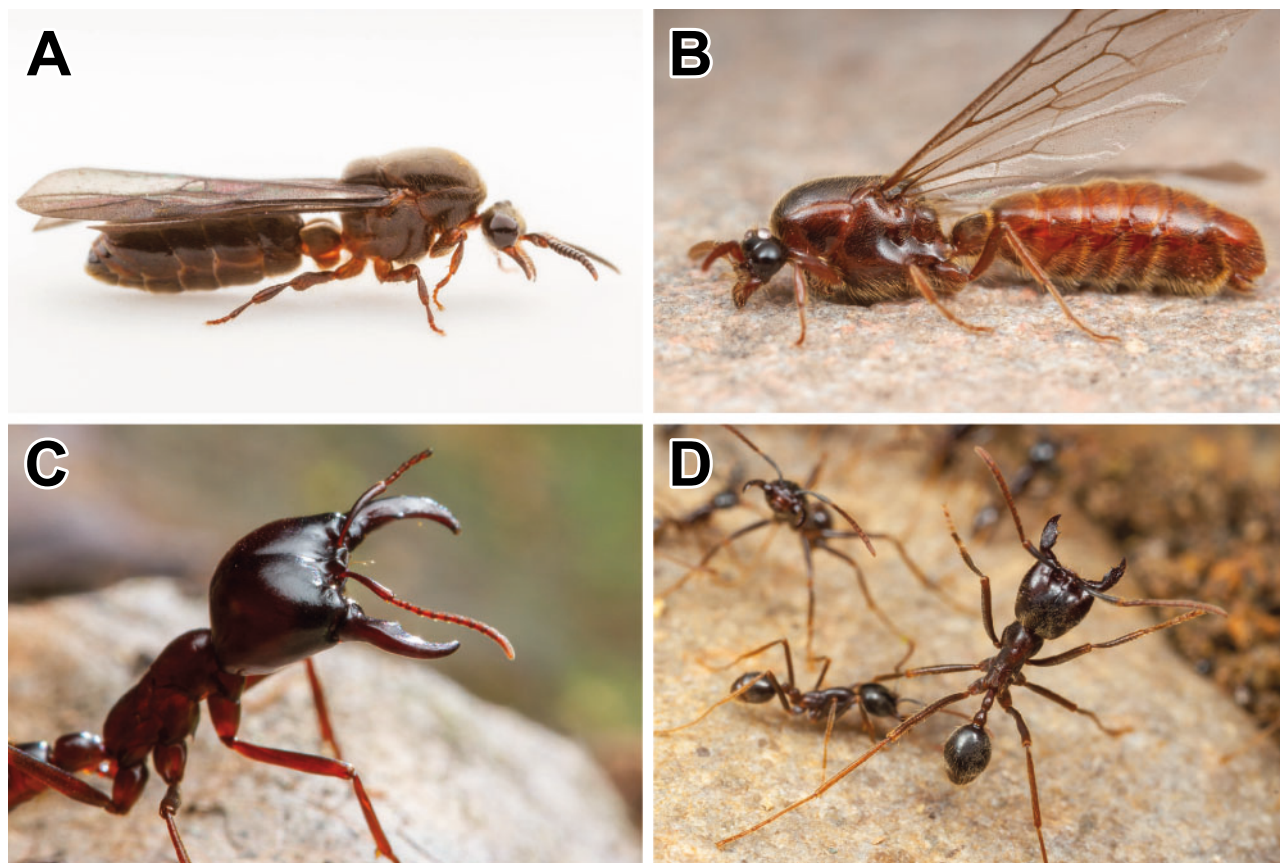
FIGURE 1.   Striking convergence of Old World (left) and New World (right) army ants. A) *Aenictus* male, B) *Neivamyrmex* male, C) *Dorylus* soldier, and D) *Labidus* soldier. Photographs courtesy of Alex Wild (http://www.alexanderwild.com/).

all 27 currently recognized extant genera (Borowiec, 2016b). The sequence data comes from a total of 2166 loci centered around ultraconserved elements or UCEs (Faircloth et al., 2012; Branstetter et al., 2017) distributed throughout the ant genomes, comprising 892,761 nucleotide sites with only 15% of missing data and gaps.

Analyses of the complete data set and loci with high average bootstrap support or "high phylogenetic signal" (Salichos and Rokas, 2013) produce results that are sensitive to partitioning scheme and inference method. The problematic yet until recently often-used k-means partitioning strategy (Frandsen et al., 2015; Baca et al., 2017) results in topologies sensitive to other analytics when all data are included but produces phylogeny congruent with other partitioning strategies when only slow-evolving or compositionally homogeneous data are considered. Analyses of data less prone to systematic bias, such as slow-evolving (Rodríguez-Ezpeleta et al., 2007; Betancur-R. et al., 2013; Goremykin et al., 2015) or compositionally homogeneous loci (Jermiin et al., 2004) and amino acid sequences (Hasegawa and Hashimoto, 1993), consistently support the hypothesis of independent origins of the army ant syndrome in the Old and New World.

MATERIALS AND METHODS

Taxon Sampling

Taxon sampling included 154 newly sequenced ingroup species from all 27 currently recognized genera of Dorylinae ants and was guided by previous taxonomic and phylogenetic studies (Borgmeier, 1955; MacKay, 1996; De Andrade, 1998; Brady, 2003; Brady and Ward, 2005; Borowiec and Longino, 2011; Jaitrong and Yamane, 2011; Bolton and Fisher, 2012; Brady et al., 2014) and expertise acquired preparing a global genus-level taxonomic revision of the group (Borowiec, 2016b). In addition to sampling all doryline genera, species from all biogeographic regions (as defined in Cox (2001) but treating the Malagasy region separate from Afrotropical) were included for each major lineage. Nine outgroup and one ingroup species (*Ooceraea biroi*) were also included based on publicly available ant genomes: *Atta cephalotes* (Suen et al., 2011), *Camponotus floridanus* (Bonasio et al., 2010), *Cardiocondyla obscurior* (Schrader et al., 2014), *Harpegnathos saltator* (Bonasio et al., 2010), *Linepithema humile* (Smith et al., 2011a), *Ooceraea biroi* (Oxley et al., 2014), *Pogonomyrmex barbatus* (Smith et al., 2011b), *Solenopsis invicta* (Wurm et al., 2011), and *Vollenhovia emeryi* (Smith et al., 2015).

Molecular Data Collection and Sequencing

I extracted DNA from all newly sequenced specimens (Supplementary Table S1) using a DNeasy Blood and Tissue Kit (Qiagen, Valencia, CA, USA). Most specimens were extracted non-destructively, with extraction voucher retained. For several extractions the DNA collection was done destructively and a voucher specimen from the same colony was kept. I quantified DNA for each extraction using a Qubit fluorometer (Thermo Fisher Scientific, Waltham, MA, USA) and sheared <5–50 ng of DNA to a target size of approximately 400–600 bp. The shearing was done by sonication on a Bioruptor instrument (Diagenode Inc., Philadelphia, PA, USA).

The library preparation protocol that follows was slightly modified from Blaimer et al. (2015). I used a KAPA Hyper Prep Library Kit (Kapa Biosystems, Inc., Wilmington, MA, USA) with magnetic bead cleanup (Fisher et al., 2011) and a SPRI substitute (Rohland and Reich, 2012) as described in (Faircloth et al., 2014). I used TruSeq adapters (Faircloth and Glenn, 2012) for ligation followed by PCR amplification of the library using a mix of HiFi HotStart polymerase reaction mix (Kapa Biosystems), Illumina TruSeq primers, and nuclease-free water. The following thermal cycler program was used for the PCR: 98°C for 45 s; 13 or 14 cycles of 98°C for 15 s, 6°C for 30 s, 72°C for 60 s, and final extension at 72°C for 5 m. After rehydrating in 23 µL pH 8 Elution Buffer (EB hereafter) and purifying reactions using 1.1–1.2× speedbeads, I pooled nine to eleven libraries at equimolar ratios for final concentrations of 132–212 n/µL.

I enriched each pool with 9446 custom-designed probes (MYcroarray, Inc.) targeting 2524 UCE loci in Hymenoptera (Branstetter et al., 2017). I followed library enrichment procedures for the MYcroarray MYBaits kit (Blumenstiel et al., 2010), except I used a 0.1× of the standard MYBaits concentration, and added 0.7 µL of 500 µM custom blocking oligos designed against the custom sequence tags. I ran the hybridization reaction for 24 h at 65°C, subsequently bound all pools to streptavidin beads (MyOne C1; Life Technologies), and washed bound libraries according to a standard target enrichment protocol (Blumenstiel et al., 2010). I used the with-bead approach for PCR recovery of enriched libraries as described in (Faircloth, 2015). I combined 15 µL of streptavidin bead-bound, enriched library with 25 µL HiFi HotStart Taq (Kapa Biosystems), 5 µL of Illumina TruSeq primer mix (5 µM each), and 5 µL of nuclease-free water. Post-enrichment PCR had the following profile: 98°C for 45 s; 18 cycles of 98°C for 15 s, 60°C for 30 s, 72°C for 60 s; and a final extension of 72°C for 5 m. I purified resulting reactions using 1.1–1.2× speedbeads, and rehydrated the enriched pools in 22 µL EB.

Following enrichment I quantified 2 µL of each pool using a Qubit fluorometer (broad range kit). I verified if the enrichment was successful by amplifying four UCE loci targeted by the probe set. I set up a relative qPCR by amplifying two replicates of 1 ng of enriched DNA from each pool for the four loci and comparing those results to two replicates of 1 ng unenriched DNA from each pool. I performed qPCR using a SYBR FAST qPCR kit (Kapa Biosystems) on CFX Connect Real-Time PCR Detection System (Bio-Rad). Following data collection, I calculated fold-enrichment values, assuming an efficiency of 1.78 and using the formula $1.78 \times abs(enrichedCp - unenrichedCp)$. I then performed qPCR quantification by creating dilutions of each pool (1:200,000, 1:800,000, 1:1.000,000, 1:10.000,000) and assuming an average library fragment length of 600 bp. Based on the size-adjusted concentrations estimated by qPCR, I pooled libraries at equimolar concentrations.

The pooled libraries were then subjected to further quality control on a Bioanalyzer and sequenced using one full and one partial lane of a HiSeq 125 Cycle Paired-End Sequencing v4 run. QC and sequencing were performed at the University of Utah High Throughput Genomics Core Facility. Quality-trimmed sequence reads generated as part of this study are available from the NCBI Sequence Read Archive: BioProject ID PRJNA504894.

Processing of UCE Data

Read cleaning, assembly, matching of contigs to probes, construction of the unaligned data set, and alignment trimming were done using the Phyluce pipeline scripts (Faircloth, 2015). I trimmed the FASTQ data using `illumiprocessor`, a wrapper around Trimmomatic (Bolger et al., 2014), with default settings (`LEADING:5`, `TRAILING:15`, `SLIDINGWINDOW:4:15`, `MINLEN:40`). Assemblies were done using Trinity v20140717 (Grabherr et al., 2011) with the `phyluce_assembly_assemblo_trinity` wrapper. The orthology assessment was then done by matching the assembled contigs to enrichment probe sequences with `phyluce_assembly_match_contigs_to_probes` (min_coverage=50, min_identity=80). This step generated a `sqlite` database, which was then used to build FASTA files for the 2524 orthologous loci with

```
phyluce_assembly_get_match_counts,
phyluce_assembly_get_fastas_from_match_
   counts,
and phyluce_assembly_explode_get_fastas_
   file
```

Extraction of Protein-Coding Sequences

For the purpose of extracting protein-coding data from the sequenced UCE loci I developed a custom bioinformatics workflow (see Supplementary Material for link to source code) that consists of three major components: 1) using NCBI BLASTX (Camacho et al., 2009) to match unaligned UCE sequences to reference proteins and 2) choosing the best hit for each sequence followed by 3) extraction of protein queries and their nucleotide equivalents from those hits. Using

makeblastdb program of the BLAST package I prepared a database from three publicly available collections of protein sequences of *Acromyrmex echinatior* (Nygaard et al., 2011), *Harpegnathos saltator* (Bonasio et al., 2010), and *Ooceraea biroi* (Oxley et al., 2014). Using BLASTX against this database resulted in one BLAST output file per UCE locus, containing multiple matches (hits) for each UCE sequence (taxon). Each hit, in turn, may be composed of one or more ranges that correspond to protein fragments (exons). I used BLAST scores for those ranges to identify best hits for each taxon and UCE. This was done with custom Python code using Biopython's (Cock et al., 2009) module for parsing BLAST XML output and the following logic:

For each sequence (taxon) and hit within, both total and maximum scores are tallied. If a hit's total score is equal to the maximum score of its ranges and it corresponds to the maximum score for the taxon, such a hit is considered best and is kept. This means that this hit was composed of a single range and its score was not exceeded by any other hits, composed of one or multiple ranges. If a hit's total score is higher than the maximum score of any one of its ranges, and that hit's total score is the best hit score for a species, this hit is kept unless it contains overlapping ranges. Finally, if the total hit score is equal to its maximum score but not to the best hit score for a species, its total score is checked to see if it corresponds to the highest individual range score for a species. If this is true, the hit is kept. Such a hit would be composed of single range and considered best even if hits with higher total scores but overlapping ranges are present. If composed of multiple ranges, a best hit is concatenated into a query in the order based on its coordinates and its presence on either forward or reverse strand. These protein queries are then matched to corresponding input nucleotide sequence or its reverse complement.

If introns that do not change reading frame are present, translations of the query sequence may span across them. Because of this I performed additional trimming if long (4 sites or more) gaps in the subject protein sequence were found. All sites corresponding to those long gaps were trimmed from the protein query and its nucleotide equivalent. If at this point there is still a stop codon in the protein query, this record was discarded.

For each record the resulting protein queries and their corresponding nucleotides were considered ready for downstream analyses.

Alignment and Trimming

Assembly and contig matching resulted in sequences of varying lengths across taxa, as evidenced by summaries produced with phyluce_assembly_get_fasta_lengths. Because of this I used UPP (Nguyen et al., 2015), a phylogeny-aware alignment tool designed to align fragmentary sequences to a backbone of longer sequences. Based on the performance (recovered final post-trimming alignment length) of different settings, for the backbone I chose the cut-off of 30% of the longest sequences present in each locus. Because the UPP version used did not have an option to specify a fixed number of longest sequences in the backbone, a locus-specific command was printed for each locus based on its fragment size distribution with UPP's -M option set to longest sequence and threshold (option -T) calculated to encompass 30% of taxa. UPP was also set to filter sequences from the backbone if their branches were five times longer than the median for all backbone sequences (-l 5 argument):

```
run_upp.py -s [input_alignment] -M
 [longest_sequence_length]
-T [locus_threshold] -d
 [alignment_output_directory] -o [output]
-p [temporary_output_directory] -l 5
```

These custom commands were printed with print_upp_command.py, a custom script utilizing code from phyluce_assembly_get_fasta_lengths.

Although alignment trimming has been recently criticized (Tan et al., 2015), the untrimmed alignments contained on average more than 75% of gaps and missing data. Because of the substantial computational burden of gap-rich data analysis, I trimmed the alignments using Gblocks (Talavera and Castresana, 2007) under settings relaxed from the default (b1=0.5 b2=0.5 b3=12 b4=7). I calculated alignment statistics and manipulated the files using AMAS v0.98 (Borowiec, 2016a). All loci with fewer than 114 taxa (less than 70%) were discarded, resulting in 2166 out of 2524 loci for downstream analyses. These loci had on average 151.7 (92.5%) taxa, were 412.2 nucleotides long, and had 7.7% missing data (gaps). Due to computing time constraints, loci with protein-coding sequences extracted were aligned using MAFFT v7.300b with --leavegappyout setting turned on. These alignments were trimmed using Gblocks with settings as above and further trimmed for outlier taxa using a custom R script and AMAS, removing any ingroup sequences whose uncorrected p-distance was more than 3σ from the mean for a locus.

Partitioning

PartitionFinder 2 (Lanfear et al., 2017) was used to partition concatenated alignments using the k-means clustering of sites based on evolutionary rates (Frandsen et al., 2015). A starting tree for model fit and site clustering algorithm was generated with RAxML Pthreads v8.2.3 (Stamatakis, 2014) using the fast tree search algorithm (-f E flag). Because of recent criticism of the k-means algorithm (Baca et al., 2017), I performed additional maximum likelihood analyses for combined data set under unpartitioned and partitioned by locus schemes. Protein-coding sequences were analyzed as partitioned by locus. K-means tends to result in relatively low number of partitions and partitioning by locus scheme was not computationally feasible for Bayesian inference.

Phylogenetic Analyses Using Maximum Likelihood

For each locus I estimated a gene tree with RAxML Pthreads v8.2.3 under a general time-reversible model of sequence evolution with rate modeled using a gamma distribution discretized into four bins (GTR+4Γ model). Two-hundred rapid bootstraps were followed by a thorough search of the maximum likelihood tree:

```
raxml -T [no_cores] -m GTRGAMMA -f a -#
  200 -p 12345 -x 12345
-s [input_alignment] -n
  [output_file_name]
```

The same mode of inference was applied to supergenes created by the statistical binning pipeline (see Species Tree analyses section below) but each supergene was partitioned by constituent loci using a partitions file (-q) and the -M flag was added for a fully partitioned analysis with branch lengths optimized separately for each partition (Bayzid et al., 2015):

```
raxml -T [no_cores] -m GTRGAMMA -f a -#
   200 -p 12345 -x 12345 -M
-s [input_alignment] -q
  [partitions_file] -n [output_file_name]
```

I ran the analyses of concatenated matrices with RAxML Hybrid v8.2.4 and v8.2.9 on CIPRES Portal (Miller et al., 2010) using the same model and bootstrap settings but with a pre-defined partitioning scheme (see above), no -M option due to the high number of partitions, and either 100 (amino acid analysis and combined data partitioned by locus) or 500 bootstrap replicates (all other analyses). The amino acid analysis was ran under the JTT+4Γ model.

Phylogenetic Analyses Using Bayesian Inference

I used ExaBayes v1.4.1 (Aberer et al., 2014) to perform analysis on k-means partitioned matrix of combined data set under GTR+4Γ. The analysis was run with two runs, four Markov Chain Monte Carlo (MCMC) chains each for 5 million generations. I disabled the default of parsimony starting tree such that analysis was initiated with a random topology. Convergence and mixing of the MCMC were determined by monitoring average standard deviation of split frequencies, which are considered acceptable below 5% (final value 1.75% for combined data set) and effective sample sizes (ESS) for all parameters, considered acceptable above 200 (minimum ESS was 900). I used 25% burn-in to construct consensus trees.

Species Tree Analyses

In addition to phylogenetic inference on concatenated loci I performed species tree analyses that attempt to reconcile gene tree incongruences arising due to incomplete lineage sorting (Edwards, 2009). I used Accurate Species TRee ALgorithm, ASTRAL v4.10.12 (Mirarab and Warnow, 2015; Sayyari and Mirarab, 2016). I used a weighted statistical binning pipeline

(Bayzid et al., 2015) to create supergene alignments and trees. Locus trees used as input for the pipeline were considered under 75 bootstrap support threshold. Summary methods for species tree inference such as those used here have been shown to be negatively impacted by error in estimated input gene trees (Roch and Warnow, 2015). The binning approach was devised to alleviate this (Mirarab et al., 2014). The data sets analyzed with species tree approaches included binned "supergenes" of all loci (514 supergenes), supergenes identified in the high bootstrap loci (147 supergenes), and supergenes from the slow-evolving loci (100 supergenes).

Measures of Compositional Heterogeneity

Most models commonly used for phylogenetic inference, including the partitioned GTR+4Γ model used here, assume that alignments are compositionally homogeneous among taxa (Moran et al., 2015). To quantify among-taxon compositional heterogeneity of the data, I used two approaches: 1) statistical tests of compositional heterogeneity and 2) a continuous measure of relative composition frequency variability (RCFV) (Zhong et al., 2011). The former included a phylogeny-corrected statistical test that compares compositions in data sets simulated under a model (the null distribution) to the compositions in the observed alignment (Foster, 2004). For the purpose of this study, the test was done on 200 simulated alignments for each observed alignment, assuming a GTR+4Γ model and a neighbor joining tree calculated using BioNJ (Gascuel, 1997). The often used but less appropriate χ^2 test for compositional heterogeneity was also performed for comparison. The two tests were carried out using the p4 program for phylogenetic inference (Foster, 2004). RCFV is the other measure used here to compare compositional heterogeneity among data (Zhong et al., 2011). RCFV is the sum of absolute values of differences observed among frequencies of all four nucleotides, divided by the number of taxa. The differences are calculated by subtracting the overall frequency of the character in a matrix from the frequency of that character in an individual sequence (taxon). The sum of these differences is then divided by the number of taxa and this number in turn is summed for each sequence/taxon in the alignment:

$$RCFV = \sum_{i=1}^{m} \sum_{j=1}^{n} \frac{|A_i j - A_i|}{n}$$

where m is the number of distinct character states (four for nucleotide sequences), $A_i j$ is the frequency of nucleotide i in taxon j, and A_i is the frequency of character (nucleotide) i across n taxa. RCFV thus gives a relative measure of compositional heterogeneity for a data set, and as the sum of differences between frequencies is calculated for each sequence, it also allows for comparison among taxa within an alignment.

Tree-Based Locus Statistics

Following maximum likelihood estimation of gene trees, I computed average branch length for each of the 2166 loci. The average branch length is a proxy for the rate of evolution of each locus. Saturation is another property that is potentially correlated with poor model fit. This I calculated by plotting p-distances of an alignment against distances on the tree from model-based maximum likelihood inference (Philippe and Forterre, 1999). These distances would show perfect fit to simple linear regression in the absence of saturation. When there is a need of correction for multiple substitutions, however, the curve will depart from linearity. I sorted each locus by slope of regression for a relative measure of saturation.

I computed all the tree-based measures with a custom R (v3.0.2 (2013-09-25)) script leveraging packages ape v3.1-1 (Paradis et al., 2004) and seqinr v3.1-2 (Charif and Lobry, 2007), modified from code originally developed for Borowiec et al. (2015).

Divergence Time Estimation

To build a time-calibrated chronogram of the Dorylinae, I used the R (v3.2.3 (Team, 2014) package ape v3.4 and its function chronos (Paradis et al., 2004; Paradis, 2013). Chronos uses the penalized likelihood (PL) method (Sanderson, 2002) and allows selection of the molecular clock model best fitting the data using an information criterion introduced in (Paradis, 2013). I used the maximum likelihood tree with branch lengths, rooted with *Harpegnathos saltator* estimated from slow-evolving loci partitioned under k-means as input. The method requires that nodes are calibrated with hard bounds of minimum and maximum ages. The calibration scheme is given in Supplementary Table S3. The information criterion implemented in chronos identified the strict molecular clock as the best fitting. Because unknown dates are initialized with a random algorithm it is possible to assess the robustness of the node ages to these initial ages by running multiple independent analyses. I ran 100 chronos replicates and summarized output trees using the sumtrees.py script distributed as a part of Denropy package v4.0.3, (Sukumaran and Holder, 2010). The summarized tree has mean branch lengths mean node ages, as well as uncertainty captured as node age ranges obtained across the 100 replicates. I visualized the tree with mean node ages and age ranges using FigTree v1.4.2. Because the information criterion implemented was shown to be poor at distinguishing the strict clock from a model with a small fixed number of rates (Paradis, 2013), I repeated the procedure for a discrete clock model with 10 categories. The results are presented in Supplementary Figure S19.

Because calibrations requiring hard minimum and maximum ages may be considered prone to bias and because PL does not utilize sequence data, I also performed Bayesian divergence dating using the recently developed birth-death process (Heath et al., 2014). This method assumes no prior belief on calibrated node ages, instead relying on a single recovery age of a fossil that is assumed to be a descendant of the calibrated node. The method simulates tree topology via a birth-death process and treats fossils as a part of the diversification process with variable attachment points on the tree and fixed recovery ages. In addition to assuming no prior beliefs on calibrated node ages, this method is not limited to using only the oldest fossils known for a given node. For these analyses I used BEAST v2.3.2 (Bouckaert et al., 2014) with Sampled Ancestors package (Gavryushkina et al., 2014). Because Bayesian divergence time estimation is computationally expensive with 150+ taxa, these analyses were limited to 109 loci (5% of all loci) sampled at even intervals throughout the rate spectrum. I set up the BEAST analyses under unpartitioned GTR+4Γ model of sequence evolution and uncorrelated clock sampling rates from a lognormal distribution. The analysis was set up with four independent runs for >300 million generations. I determined convergence and adequate sample size (all parameters sampled at ESS >200) using Tracer v1.5. The calibration scheme included seven fossils with fixed sampling times obtained by drawing a random number from a uniform distribution (runif function in R base) bounded by minimum and maximum ages of the deposit in which the fossil is found (deposit ages follow the fossilworks.org website (Alroy, 2016); Supplementary Table S4). I used conditioning on the root age with a prior.

Diversification Analyses

BAMM v2.5 (Rabosky, 2014) analyses used the consensus BEAST chronogram and a table of sampling probabilities based on extant species diversity estimates for each genus (Borowiec, 2016b). The sampling proportions were set as follows: *Acanthostichus*: 0.17, *Aenictogiton*: 0.2, *Aenictus*: 0.09, *Cerapachys*: 0.5, *Cheliomyrmex*: 0.6, *Chrysapace*: 0.8, *Cylindromyrmex*: 0.4, *Dorylus*: 0.2, *Eburopone*: 0.12, *Eciton*: 0.58, *Eusphinctus*: 0.5, *Labidus*: 0.6, *Leptanilloides*: 0.17, *Lioponera*: 0.08, *Lividopone*: 0.13, *Neivamyrmex*: 0.11, *Neocerapachys*: 0.4, *Nomamyrmex*: 1, *Ooceraea*: 0.2, *Parasyscia*: 0.08, *Simopone*: 0.18, *Sphinctomyrmex*: 0.4, *Syscia*: 0.13, *Tanipone*: 0.4, *Vicinopone*: 1, *Yunodorylus*: 0.4, *Zasphinctus*: 0.15. I used the setBAMMpriors function in the R package BAMMtools (Rabosky et al., 2014) to create priors used for the analysis. I ran the MCMC for 20 million generations, sampling every 2000 generations. I checked the convergence and plotted the analysis results using BAMMtools and CODA (Plummer et al., 2006).

Biogeographic History Estimation

I used the maximum likelihood functions available in BioGeoBEARS R package (Matzke, 2013, 2014) to compare fit and select from among models

commonly used for estimation of biogeographic histories. I discretized species distributions into six biogeographic regions following (Cox, 2001) and treating Malagasy as the seventh separate region. The regions included Neotropical, Nearctic, Palearctic, Afrotropical, Malagasy, Indomalayan, and Australasian. The boundary between Indomalayan and Australasian regions is the Wallace Line. I set up a time-stratified analysis that assumed different region adjacency and dispersal probabilities between 110–50 Ma and 50 Ma present (M3_stratified-type model of Matzke (2014)) and compared likelihoods and AICc scores under six models: DEC, DEC+J, DIVALIKE, DIVALIKE+J, BAYAREALIKE, and BAYAREALIKE+J.

RESULTS AND DISCUSSION

Independent Origins of the Army Ant Syndrome Despite Incongruence

The concatenated dataset produces different topologies depending on the partitioning scheme and method used to infer the phylogeny (Fig. 2A–D, P), although the true army ants are not monophyletic in any of the concatenated analyses. The maximum likelihood tree under partitioning by locus shows a clade that includes almost all New World dorylines, including New World army ants (hereafter called "New World Clade"; Fig. 2A; Supplementary Fig. S1). The latter include the genera *Cheliomyrmex*, *Eciton*, *Labidus*, *Neivamyrmex*, and *Nomamyrmex*. Apart from New World army ants, the New World Clade unites all exclusively New World lineages of the Dorylinae and includes *Acanthostichus*, *Cylindromyrmex*, *Leptanilloides*, *Neocerapachys*, and *Sphinctomyrmex*. Old World driver ants *Dorylus* and the poorly known genus *Aenictogiton* are sister to *Aenictus*. Other well-supported clades comprising multiple genera include a clade of South-East Asian and Malagasy lineages *Cerapachys*, *Chrysapace*, and *Yunodorylus*, a well-resolved clade of several Old World genera (*Lioponera*, *Lividopone*, *Parasyscia*, *Zasphinctus*) that was also recovered in (Brady et al., 2014). Another South-East Asian clade consists of the mostly subterranean ants in genera *Eusphinctus*, *Ooceraea*, and *Syscia*. The backbone of the tree shows multiple short internodes and generally lower resolution, consistent with a previous study (Brady et al., 2014). The maximum likelihood tree inferred from the same data set but using k-means partitioning (Frandsen et al., 2015) shows a strongly-supported Old World *Aenictus* and New World army ants clade (Fig. 2B; Supplementary Fig. S2). Army ants are not monophyletic in this tree either, because *Dorylus* and *Aenictogiton* are not sister to the *Aenictus* plus New World army ants clade. Unpartitioned analysis results in a similar topology (Fig. 2C; Supplementary Fig. S3).

Coalescent species tree inferred from all loci is similar to the concatenation results with respect to poor resolution at deep nodes (Fig. 2D; Supplementary Fig. S4). The species tree resembles the unpartitioned and k-means partitioned concatenated trees because of moderate support for the *Aenictus* plus New World army ants clade, and contrasts most with both concatenated trees by supporting army ant monophyly, the hypothesis favored by previous studies (Brady, 2003; Brady and Ward, 2005; Brady et al., 2014). Overall support for the deep nodes is lower in species trees and this approach does not recover some of the clades present across all concatenated analyses, namely the New World or South-East Asian/Malagasy groups mentioned above.

Maximum Likelihood Analyses of All and "High Signal" Loci Show Evidence of Bias

The conflict among maximum likelihood topologies derived from different partitioning schemes and between concatenated and species tree warranted further investigation. I constructed five data sets in addition to the combined data matrix: matrix of "high signal" loci (Salichos and Rokas, 2013) equivalent in length to one-fifth of the sites in combined data set matrix, matrix composed of slow-evolving loci equivalent in length to one-fifth of the sites in combined data set matrix, matrix composed of only compositionally homogeneous loci, and matrix that excluded the long-branched *Aenictus* species but is otherwise identical to combined data matrix (See Supplementary Table S2 for more information on these matrices). Additionally, I developed a workflow to extract putatively protein-coding regions from the UCE data set (see SI Materials and Methods section) and analyzed those regions separately, coded as amino acids.

Analyses of loci whose trees have highest mean bootstrap support (high phylogenetic signal sensu (Salichos and Rokas, 2013)) shows even more discordance among analyses than that of the combined data set (Fig. 2E–H; Supplementary Figs. S5–S8). Army ant lineages show different topologies in each of the three partitioning schemes.

Using only the slowest evolving one-fifth of the data ("slow-evolving" hereafter), there is universal support for a clade that unites all exclusively New World lineages, including the New World army ants under all partitioning schemes (Fig. 2I–K; Supplementary Figs. S9–S11). Species tree analyses of slow-evolving loci also support New World Clade and the *Aenictus* plus (*Aenictogiton+Dorylus*) clade (Fig. 2L; Supplementary Fig. S12).

Examining locus properties in different rate bins reveals that more rapidly evolving sequences and "high signal" loci exhibit qualities previously associated with systematic bias, namely higher compositional heterogeneity (Lockhart et al., 1994; Jermin et al., 2004) and saturation (Philippe and Forterre, 1999). The most slowly evolving loci have lower overall among-taxon sequence heterogeneity than more rapidly evolving and high bootstrap loci (Supplementary Fig. S28A; RCFV two-sample t-test slow-evolving vs. all other loci: $t=$
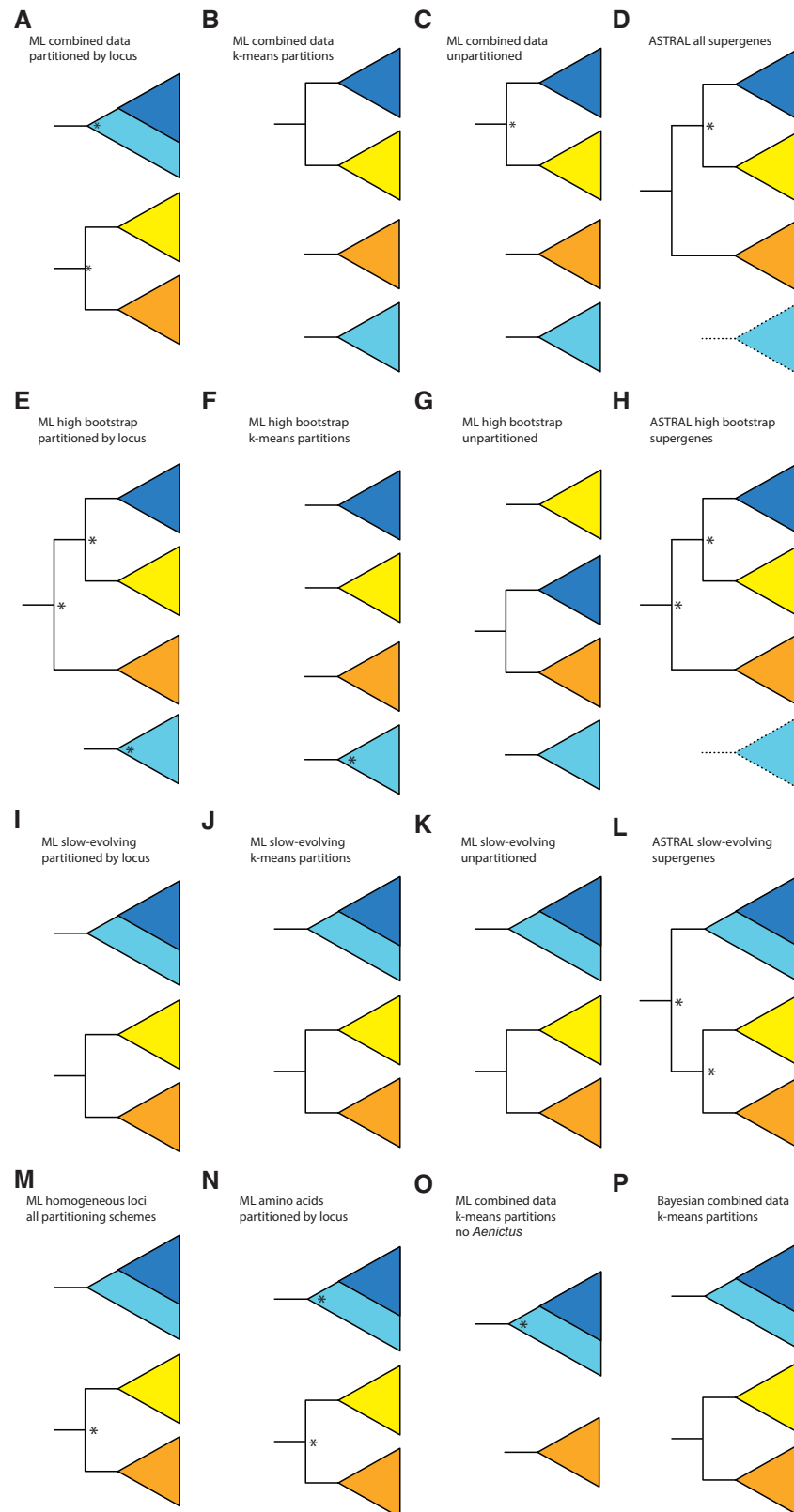
FIGURE 2.   Summary of phylogenetic position of army ant lineages recovered in different analyses. In each figure yellow and orange colors signify Old World army ants (yellow signifies genus *Aenictus* and orange the clade of *Aenictogiton + Dorylus*), dark blue signifies New World army ants, and light blue signify non-army ant New World dorylines. Note increased incongruence in analysis of combined data set (A–D) and "high signal" data set (E–H) compared with slow-evolving (I–L) and compositionally homogeneous (M) loci. Maximum likelihood analyses of amino acid data (N) and combined data without *Aenictus* (O), as well as Bayesian analysis of combined data support New World Clade monophyly. Asterisk indicates support ≤ 99%. Dashed lines signify polyphyly (see Figures S1–S18 available on Zenodo for complete trees). See on-line version of this article for color figures.

28.564, slow-evolving *mean* = 0.0358, all other loci *mean* = 0.0575, $p \ll 0.01$; slow-evolving vs. high bootstrap: $t = 27.538$ high bootstrap *mean* = 0.0647, $p \ll 0.01$). Saturation is also greater in more rapidly evolving and high bootstrap loci (Supplementary Fig. S28B; slope of regression two-sample *t*-test slow-evolving vs. all other loci: $t = 21.682$, slow-evolving *mean* = 0.474, all other loci *mean* = 0.361, $p \ll 0.01$; slow-evolving vs. high bootstrap: $t = 10.129$, high bootstrap *mean* = 0.393, $p \ll 0.01$). Of the 379 loci that pass the homogeneity test (Foster, 2004), 244 are found among slow-evolving loci and zero are present among the high bootstrap loci.

Analysis of loci that pass the phylogeny-corrected compositional homogeneity test (Foster, 2004) shows universal support for the New World Clade, regardless of partitioning scheme (Fig. 2M; Supplementary Figs. S13–S15). Among-taxon compositional heterogeneity is a relatively well-understood source of bias in phylogenetic analysis (Jermiin et al., 2004) and is not accounted for by the general time-reversible model of sequence evolution used in RAxML, the program used here for maximum likelihood inference.

Amino acid alignments are also known to be more robust against compositional bias (Hasegawa and Hashimoto, 1993) and saturation, although not free from it when distantly related lineages are considered (Foster and Hickey, 1999). In this study, the amino acid matrix of protein-coding sequences is highly conserved compared with the nucleotide matrix of combined data with 14% proportion of parsimony informative sites in the former compared with 48% in the latter. The amino acid matrix analyzed under maximum likelihood recovers monophyletic New World Clade and Old World army ants (Fig. 2N; Supplementary Fig. S16).

More evidence for the artifactual nature of the grouping of Old World and New World army ants comes from an analysis where *Aenictus* is removed from the combined data matrix. The long-branched Old World genus *Aenictus* clusters with New World army ants in analyses of unpartitioned and k-means partitioned full data set, while the rest of the New World Clade is more distantly related. This is not the case when the combined loci are analyzed partitioned by locus (Fig. 2A–D). If *Aenictus* was not affecting the position of New World army ants on the tree, one would expect to see no change of the position of the latter if the former is removed. This is not the case here; the phylogeny recovered from the alignment without *Aenictus* has New World army ants nested within the New World Clade (Fig. 2O; Supplementary Fig. S17).

Bayesian analysis of k-means partitioned combined data matrix strongly supports the New World Clade and monophyly of *Aenictogiton + Dorylus*, unlike the maximum likelihood analysis of the same data set under the same partitioning scheme and sequence evolution model (Fig. 2P; Supplementary Fig. S18).

Interestingly, the discredited k-means partitioning strategy (Frandsen et al., 2015; Baca et al., 2017) produces highly unusual and incongruent results when all loci are included, but not when the slow-evolving or homogeneous subsets are analyzed, or when the long-branched *Aenictus* is removed.

In summary, although concatenated maximum likelihood and species tree analysis of combined data matrix and high-bootstrap loci support grouping of Old World army ants *Aenictus* with New World army ants under some conditions, analyses using more reliable data are congruent in their support for independent origins of the army ant syndrome in New World and Old World.

Timeline of Doryline Evolution and Diversification Suggests Ancient Rapid Radiation

Fossilized birth–death (FBD) process divergence dating (Heath et al., 2014) employed here shows that crown doryline ants started diversifying in the Cretaceous, around 74 Ma (53–101 Ma 95% highest probability density interval or HPD) ago according to Bayesian inference (97 Ma under PL; Supplementary Fig. S19) (Fig. 3). The FBD results near the root are characterized by high uncertainty but the mean age suggests a younger crown age than 87 Ma recovered in a previous study (Brady et al., 2014). The difference is likely at least in part due to a different calibration approach used in the present study (FBD vs. node dating) and a revised, younger age estimate of Baltic amber (Aleksandrova and Zaporozhets, 2008). Similar to concatenation, divergence dating shows a tree highly compressed during early evolution, with 16 splits occurring within the first 20 Ma. According to PL this period of early diversification lasted longer, accounting for about 35 Ma. The most recent common ancestor of the New World Clade is resolved at 60 Ma (43–83 Ma 95% HPD interval; 60 Ma under PL). The split of Old World army ant lineages of *Aenictus* and *Dorylus+Aenictogiton* is apparently very ancient at 58 Ma (41–79 Ma 95% HPD; 54 Ma under PL) and occurred during the initial diversification period. The old age of this node and long branch subtending extant *Aenictus* help explain why this relationship is difficult to recover. The five currently recognized genera of New World army ants share an ancestor at about 28 Ma (20–37 Ma 95% HPD; 27 Ma under PL) ago. These dates are younger than those inferred for several of the non-army ant doryline genera such as *Eburopone* or *Leptanilloides*. The conspicuous above-ground foraging seen in some *Dorylus* driver ants and in New World *Eciton* is remarkably young, as it appears that these groups diversified within the last 5–6 Ma, a result robust across different dating analyses (see Supplementary Figs. S19–S21). The driver ant genus *Dorylus* is shown to be young relative to earlier analyses (Brady, 2003; Kronauer et al., 2007) at ca. 16 Ma (11–22 95% HPD; 15 Ma under PL).

The Bayesian divergence time estimates, especially those early in the tree and outside the New World Clade, are associated with considerable uncertainty. This is likely due to both topological uncertainty and the fact that only seven fossil calibrations were available for the
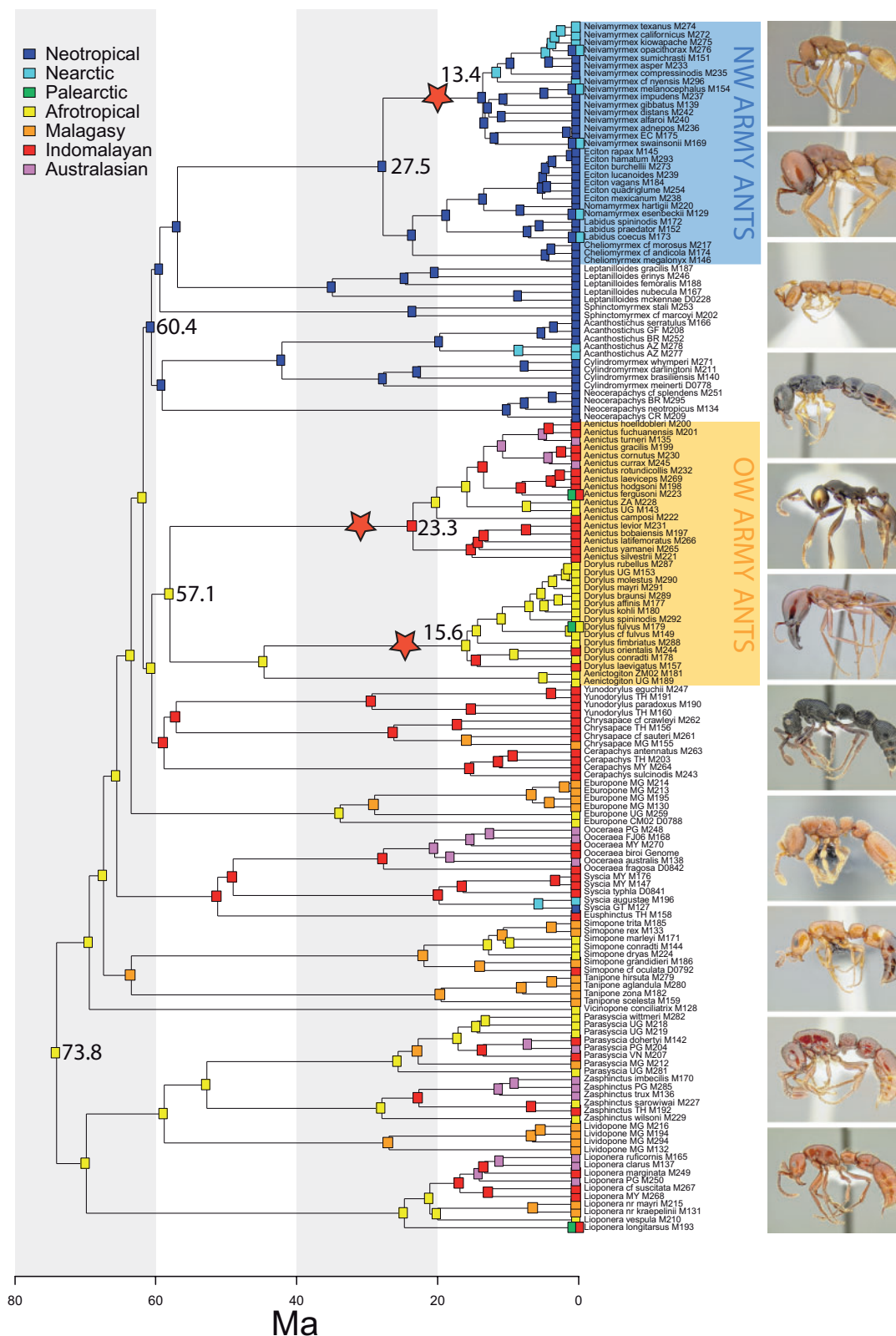
FIGURE 3.   Timeline of doryline evolution and biogeographic history. The highest relative probabilities of ancestral ranges are shown as inferred using BioGeoBEARS under DEC+J, averaged over 100 trees from the BEAST posterior. The tree is the BEAST consensus. Selected mean divergence time estimates are given at nodes. Stars signify rate shifts. All dates in Ma. See Supplementary Figures S20, S21, and S26 available on Dryad for average node dates, posterior probabilities from the BEAST analysis, and pie charts of relative probabilities of all possible ancestral ranges, respectively. See on-line version of this article for color figures.

dorylines, all except one located within the New World Clade (Supplementary Table S4, Fig. S20).

Diversification shift analyses in BAMM (Rabosky, 2014) identified a three shift-scenario for dorylines, with shifts occurring separately on the branches subtending *Aenictus*, *Dorylus*, and *Neivamyrmex* (Fig. 3). Rate shift configurations in which only two shifts were identified, however, were also common in the posterior sampling (Supplementary Fig. S25), either on branches leading to *Aenictus* and *Neivamyrmex* only, or on the branches subtending the clade of *Aenictus*, *Aenictogiton*, and *Dorylus* and *Neivamyrmex* (see Supplementary Figs. S22–S25). Because reliability and validity of diversification rate estimates obtained using BAMM is being vigorously debated (Moore et al., 2016; Meyer and Wiens, 2017; Rabosky et al., 2017; Rabosky, 2018; Meyer et al., 2018), these results should be treated with caution.

Biogeographic History Reveals Geographic Fidelity of Major Clades

Strong geographic affinities are apparent within the doryline phylogeny. Large clades are mostly confined to only one or two adjacent realms, although movement within both the Old and New World appears to have been common (Fig. 3). There is strong evidence for Old World origins of the dorylines and the analyses summarized across a sample of trees from the posterior suggest an Afrotropical ancestral range (Fig. 3; Supplementary Fig. S26), although the analysis under BEAST consensus only results in high uncertainty and combined Afrotropical-Malagasy-Oriental as the most likely ancestral range (Supplementary Fig. S27). Two lineages are confined to the New World. One gave rise to the radiation of almost all extant New World forms, including New World army ants and their kin. The dates estimated for the origin of this New World Clade coincide with warm climatic conditions and multiple land bridges connecting the Old and New World in northern latitudes (Brikiatis, 2014). The other New World group is much younger and appeared to arrive some time after 28 Ma ago. The presumably SE Asian or Palearctic ancestor of the New World species of *Syscia* either crossed the Beringian land bridge or dispersed across the Pacific further south, since by that time North Atlantic connections were closed (Sanmartín et al., 2001). The Old World *Aenictus* and New World genus *Neivamyrmex*, although superficially similar in appearance and biology, illustrate different scenarios of biogeographic history for generic lineages within Dorylinae. Crown *Aenictus* is older at 23 Ma (34 Ma under PL) and originated in the Indomalayan region. It then subsequently moved into the Afrotropics, dispersed back into Indomalaya, and moved into Australasia with possible movement back into Indomalaya. Some species also range into the Palearctic. In contrast, *Neivamyrmex* remained largely confined to the Neotropics where it originated around 13 Ma (20 Ma under PL) ago, with at least one clade

moving into and diversifying in the southern Nearctic and with some species returning to the Neotropics or straddling the boundary of the two adjacent regions. Madagascar is a center of doryline diversity with seven overall and two endemic genera but no true army ants (Fig. 3).

CONCLUSIONS

Densely Sampled Phylogenomic Data Sets Are Not Robust to Artifacts and Bias

Genome-scale data offers powerful tools for reconstructing phylogenies. This study, however, demonstrates that caution is necessary when evaluating hypotheses generated by these new data sets, even when taxon sampling is comprehensive. Empirical studies that emphasize the importance of bias often recommend improving taxon sampling but usually also deal with cases where sampling could be increased relative to first attempts, such as in broad-scale phylogenies of eukaryotes or Metazoa (Delsuc et al., 2005). The doryline phylogeny represents a case where the scope for improvement by additional sampling of currently known lineages is limited. This is because adding more species is likely to fail at significantly shortening long branches in the tree (Borowiec, 2016b). Researchers should be thus wary of systematic bias whenever combination of very short and very long branches is encountered, regardless of taxon sampling. Two general strategies are available for exploring sensitivity to model mis-specifications: using only data less prone to bias or applying better-fitting heterogeneous models (Rodríguez-Ezpeleta et al., 2007). The latter approach is ideal but computationally expensive, becoming prohibitive with data sets such as the one used here, including over 150 taxa and thousands of loci.

In this study, reduction of phylogenetic noise resulting from compositional heterogeneity and saturation increased congruence among topologies obtained using different analytics. In the case of the complete data set, mutually exclusive and strongly supported results were be obtained depending on the statistical framework (e.g., maximum likelihood vs. Bayesian) or sequence evolution model (e.g., k-means partitioning vs. partitioning by locus) chosen for analysis. Using only "high signal" loci (Salichos and Rokas, 2013) with highest average bootstrap exacerbated incongruence among analyses. These loci also exhibited undesirable properties such as higher potential for saturation and violation of the among-taxon compositional heterogeneity. This indicates that using additional measures of data quality is needed in phylogenomics, such as direct and indirect measures of model mis-specification (Brown, 2014). Other recent research suggests that analysis results can be strongly affected by a tiny proportion of highly biased loci or sites (Shen et al., 2017). In conclusion, phylogenomic studies should always perform sensitivity analyses to test the robustness of the result to different analytics.

Doryline Biology and Evolution Need Further Study

The new phylogeny presented here reveals that the army ant syndrome can be viewed as both an example of long-term evolutionary stasis and a remarkable case of convergent evolution. Brady (2003) argued that the army ant syndrome originated only once around 100 Ma ago and has since persisted without loss in any descendant lineages. While the present study suggests that this set of behavioral and morphological traits evolved at least twice in the Dorylinae, it also shows that the syndrome has been conserved within Old and New World army ants. The alternative scenario of single origin requires multiple losses on lineages leading to both Old and New World army ants (Fig. 3), an unlikely proposition given that no species are known to have lost any of the syndrome components in the large and diverse genera such as *Aenictus*, *Dorylus*, or *Neivamyrmex*.

Despite the improved resolution of the army ant tree, much work remains to be done with the regard to doryline evolution. A particularly vexing matter is poor knowledge of the Afrotropical genus *Aenictogiton*. Although based on male morphology and its phylogenetic affinity to *Dorylus* it has been assumed that it is a subterranean army ant, there is no direct evidence of army ant behavior or queen morphology (Borowiec, 2016b). If *Aenictogiton* is not an army ant, our views on the evolution of the army ant syndrome have to be adjusted. In general, the current knowledge of doryline ecology and behavior is mostly limited to the minority of species that are conspicuous above-ground foragers, although the clonal *Ooceraea biroi* is a notable exception (Tsuji and Yamauchi, 1995; Ravary and Jaisson, 2002; Oxley et al., 2014). Better understanding of doryline behavior and morphology will undoubtedly yield further insights into the evolution of the army ant syndrome. More specifically, convergent evolution of army ants is perhaps less surprising in the light of the fact that its individual components, including flightlessness of queens and frequent colony relocations, appear to have evolved multiple times in many non-army ant dorylines (Borowiec, 2016b). Unfortunately, the biology of many dorylines remains unknown, precluding a fuller understanding of conditions underlying the evolution of true army ants (Kronauer, 2009; Brady et al., 2014). The new phylogenetic and taxonomic (Borowiec, 2016b) framework should reinvigorate comparative work on doryline morphology and natural history.

SUPPLEMENTARY MATERIAL

Trimmed reads generated for this study are available at the NCBI Sequence Read Archive (BioProject ID PRJNA504894). Supplementary figures, tables, and references are available on Zenodo: https://doi.org/10.5281/zenodo.1482030. Sequence files, alignments, configuration files, and output of analyses, including phylogenetic trees, are available in a Dryad repository at http://dx.doi.org/10.5061/dryad.bj83h5d and a Zenodo repository at https://doi.

org/10.5281/zenodo.1482055. Custom scripts used in this study are available on GitHub: https://doi.org/10.5281/zenodo.571246. Pipeline for extracting protein-coding sequences is available in a separate GitHub repository at https://doi.org/10.5281/zenodo.571247.

REFERENCES

Aberer A.J., Kobert K., Stamatakis A. 2014. ExaBayes: massively parallel bayesian tree inference for the whole-genome era. Mol. Biol. Evol. 31:2553–2556.

Aleksandrova G.N., Zaporozhets N.I. 2008. Palynological characteristics of Upper Cretaceous and Paleogene deposits on the west of the Sambian Peninsula (Kaliningrad region), part 1. Stratigraphy Geol. Corr. 16:295–316.

Alroy J. 2016. Gateway to the paleobiology database. Available from: www.fossilworks.org.

Baca S.M., Toussaint E.F., Miller K.B., Short A.E. 2017. Molecular phylogeny of the aquatic beetle family Noteridae (Coleoptera: Adephaga) with an emphasis on data partitioning strategies. Mol. Phylogenet. Evol. 107:282–292.

Baroni Urbani C., Bolton B., Ward P.S. 1992. The internal phylogeny of ants (Hymenoptera: Formicidae). Syst. Entomol. 17(4):301–329.

Bayzid M.S., Mirarab S., Boussau B., Warnow T. 2015. Weighted statistical binning: enabling statistically consistent genome-scale phylogenetic analyses. PLoS One. 10:e0129183.

Betancur-R., R., Naylor G. J.P., Orti G. 2013. Conserved genes, sampling error, and phylogenomic inference. Syst. Biol. 63:257–262.

Blaimer B.B., Brady S.G., Schultz T.R., Lloyd M.W., Fisher B.L., Ward P.S. 2015. Phylogenomic methods outperform traditional multi-locus approaches in resolving deep evolutionary history: a case study of formicine ants. BMC Evol. Biol. 15:e271.

Blumenstiel B., Cibulskis K., Fisher S., DeFelice M., Barry A., Fennell T., Abreu J., Minie B., Costello M., Young G., et al. 2010. Targeted exon sequencing by in-solution hybrid selection. Curr Protoc Hum Genet. Chapter 18:Unit 18.4.

Bolger A.M., Lohse M., Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 30:2114–2120.

Bolton B. 1990. Army ants reassessed: the phylogeny and classification of the doryline section (Hymenoptera, Formicidae). J. Nat. Hist. 24:1339–1364.

Bolton B., Fisher B.L. 2012. Taxonomy of the cerapachyine ant genera *Simopone* Forel, *Vicinopone* gen. n. and *Tanipone* gen. n. (Hymenoptera: Formicidae). Zootaxa. 3283:1–101.

Bonasio R., Zhang G., Ye C., Mutti N.S., Fang X., Qin N., Donahue G., Yang P., Li Q., Li C., Zhang P., Huang Z., Berger S.L., Reinberg D., Wang J., Liebig J. 2010. Genomic comparison of the ants *Camponotus floridanus* and *Harpegnathos saltator*. Science. 329:1068–1071.

Borgmeier T. 1955. Die Wanderameisen der Neotropischen Region. Studia Entomol. 3:1–720.

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

Borowiec M.L. 2016b. Generic revision of the ant subfamily Dorylinae (Hymenoptera, Formicidae). ZooKeys. 608:1–280.

Borowiec M.L., Lee E.K., Chiu J.C., Plachetzki D.C. 2015. Extracting phylogenetic signal and accounting for bias in whole-genome data sets supports the Ctenophora as sister to remaining Metazoa. BMC Genomics. 16:e987.

Borowiec M.L., Longino J.T. 2011. Three new species and reassessment of the rare neotropical ant genus *Leptanilloides* (Hymenoptera, Formicidae, leptanilloidinae). ZooKeys. 133:19–48.

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

Brady S.G. 2003. Evolution of the army ant syndrome: the origin and long-term evolutionary stasis of a complex of behavioral and reproductive adaptations. Proc. Natl. Acad. Sci. USA. 100:6575–6579.

Brady S.G., Fisher B.L., Schultz T.R., Ward P.S. 2014. The rise of army ants and their relatives: diversification of specialized predatory doryline ants. BMC Evol. Biol. 14:93.

Brady S.G., Ward P.S. 2005. Morphological phylogeny of army ants and other dorylomorphs (Hymenoptera: Formicidae). Syst. Entomol. 30:593–618.

Branstetter M.G., Longino J.T., Ward P.S., Faircloth B.C. 2017. Enriching the ant tree of life: enhanced UCE bait set for genome-scale phylogenetics of ants and other Hymenoptera. Methods Ecol. Evol. 9:768–776.

Brikiatis L. 2014. The De Geer, Thulean and Beringia routes: key concepts for understanding early Cenozoic biogeography. J. Biogeogr. 41:1036–1054.

Brown J.M. 2014. Predictive approaches to assessing the fit of evolutionary models. Syst. Biol. 63:289–292.

Brown W.L. 1975. Contributions toward a reclassification of the Formicidae. v. Ponerinae, tribes Platythyreini, Cerapachyini, Cylindromyrmecini, Acanthostichini, and Aenictogitini. Search. Agriculture (Ithaca, New York). 5:1–115.

Camacho C., Coulouris G., Avagyan V., Ma N., Papadopoulos J., Bealer K., and Madden T.L. 2009. BLAST+: architecture and applications. BMC Bioinformatics. 10:421.

Charif D. and Lobry J.R. 2007. SeqinR 1.0-2: a contributed package to the R project for statistical computing devoted to biological sequences retrieval and analysis. In: Ugo Bastolla, Markus Porto, and Eduardo H. Roman. eds, Structural approaches to sequence evolution: molecules, networks, populations. Berlin, Heidelberg: Springer Science Business Media. p. 207–232.

Cock P. J.A., Antao T., Chang J.T., Chapman B.A., Cox C.J., Dalke A., Friedberg I., Hamelryck T., Kauff F., Wilczynski B., de Hoon M.J.L. 2009. Biopython: freely available python tools for computational molecular biology and bioinformatics. Bioinformatics. 25:1422–1423.

Cox B. 2001. The biogeographic regions reconsidered. J. Biogeogr. 28:511–523.

De Andrade M. 1998. Fossil and extant species of *Cylindromyrmex* (Hymenoptera: Formicidae). Revue Suisse de Zoologie. 105:581–664.

Delsuc F., Brinkmann H., Philippe H. 2005. Phylogenomics and the reconstruction of the tree of life. Nat. Rev. Genet. 6:361–375.

Edwards S.V. 2009. Is a new and general theory of molecular systematics emerging? Evolution. 63:1–19.

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

Faircloth B.C., Branstetter M.G., White N.D., Brady S.G. 2014. Target enrichment of ultraconserved elements from arthropods provides a genomic perspective on relationships among Hymenoptera. Mol. Ecol. Resour., 15:489–501.

Faircloth B.C. and Glenn T.C. 2012. Not all sequence tags are created equal: designing and validating sequence identification tags robust to indels. PLoS One. 7:e42543.

Faircloth B.C., McCormack J.E., Crawford N.G., Harvey M.G., Brumfield R.T., Glenn T.C. 2012. Ultraconserved elements anchor thousands of genetic markers spanning multiple evolutionary timescales. Syst. Biol. 61:717–726.

Fisher S., Barry A., Abreu J., Minie B., Nolan J., Delorey T.M., Young G., Fennell T.J., Allen A., Ambrogio L., Berlin A.M., Blumenstiel B., Cibulskis K., Friedrich D., Johnson R., Juhn F., Reilly B., Shammas R., Stalker J., Sykes S.M., Thompson J., Walsh J., Zimmer A., Zwirko Z., Gabriel S., Nicol R., Nusbaum C. 2011. A scalable, fully automated process for construction of sequence-ready human exome targeted capture libraries. Genome Biol. 12:R1.

Foster P. 2004. Modeling compositional heterogeneity. Syst. Biol. 53:485–495.

Foster P.G., Hickey D.A. 1999. Compositional bias may affect both DNA-based and protein-based phylogenetic reconstructions. J. Mol. Evol. 48:284–290.

Frandsen P.B., Calcott B., Mayer C., Lanfear R. 2015. Automatic selection of partitioning schemes for phylogenetic analyses using iterative k-means clustering of site rates. BMC Evol. Biol. 15:13.

Garnier S., Murphy T., Lutz M., Hurme E., Leblanc S., Couzin I.D. 2013. Stability and responsiveness in a self-organized living architecture. PLoS Comput. Biol. 9:e1002984.

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.

Gavryushkina A., Welch D., Stadler T., Drummond A.J. 2014. Bayesian inference of sampled ancestor trees for epidemiology and fossil calibration. PLoS Comput. Biol. 10:e1003919.

Goremykin V.V., Nikiforova S.V., Cavalieri D., Pindo M., Lockhart P. 2015. The root of flowering plants and total evidence. Syst. Biol. 64:879–891.

Gotwald W.H. 1995. Army ants: the biology of social predation. Ithaca (NY): Cornell University Press.

Gotwald W. H.J. 1979. Phylogenetic implications of army ant zoogeography (Hymenoptera: Formicidae). Ann. Entomol. Soc. Am. 72(4):462–467.

Grabherr M.G., Haas B.J., Yassour M., Levin J.Z., Thompson D.A., Amit I., Adiconis X., Fan L., Raychowdhury R., Zeng Q., Chen Z., Mauceli E., Hacohen N., Gnirke A., Rhind N., di Palma F., Birren B.W., Nusbaum C., Lindblad-Toh K., Friedman N., Regev A. 2011. Full-length transcriptome assembly from RNA-seq data without a reference genome. Nat. Biotechnol. 29:644–652.

Hasegawa M., Hashimoto T. 1993. Ribosomal RNA trees misleading? Nature. 361:23–23.

Heath T.A., Huelsenbeck J.P., Stadler T. 2014. The fossilized birth-death process for coherent calibration of divergence-time estimates. Proc. Natl. Acad. Sci. USA. 111:E2957–E2966.

Jaitrong W., Yamane S. 2011. Synopsis of *Aenictus* species groups and revision of the *A. currax* and *A. laeviceps* groups in the eastern Oriental, Indo-Australian, and Australasian regions (Hymenoptera: Formicidae: Aenictinae). Zootaxa. 3128:1–46.

Jermiin L., Ho S.Y., Ababneh F., Robinson J., Larkum A.W. 2004. The biasing effect of compositional heterogeneity on phylogenetic estimates may be underestimated. Syst. Biol. 53:638–643.

Kronauer D.J., Schöning C., Vilhelmsen L.B., Boomsma J.J. 2007. A molecular phylogeny of *Dorylus* army ants provides evidence for

multiple evolutionary transitions in foraging niche. BMC Evol. Biol. 7:56.

Kronauer D.J.C. 2009. Recent advances in army ant biology (Hymenoptera: Formicidae). Myrmecol. News. 12:51–65.

Lanfear R., Frandsen P.B., Wright A.M., Senfeld T., Calcott B. 2017. PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. Mol. Biol. Evol. 34:772–773.

Lockhart P.J., Steel M.A., Hendy M.D., Penny D. 1994. Recovering evolutionary trees under a more realistic model of sequence evolution. Mol. Biol. Evol. 11:605–612.

MacKay W.P. 1996. A revision of the ant genus *Acanthostichus*. Sociobiology. 27:129–179.

Matzke N.J. 2013. Probabilistic historical biogeography: new models for founder-event speciation, imperfect detection, and fossils allow improved accuracy and model-testing. Berkeley (CA): University of California.

Matzke N.J. 2014. Model selection in historical biogeography reveals that founder-event speciation is a crucial process in island clades. Syst. Biol. 63:951–970.

Meyer A.L.S., Román-Palacios C., Wiens J.J. 2018. BAMM gives misleading rate estimates in simulated and empirical datasets. Evolution. 72(10):2257–2266.

Meyer A.L.S., Wiens J.J. 2017. Estimating diversification rates for higher taxa: BAMM can give problematic estimates of rates and rate shifts. Evolution. 72:39–53.

Miller M.A., Pfeiffer W., Schwartz T. 2010. Creating the CIPRES science gateway for inference of large phylogenetic trees. In: 2010 Gateway Computing Environments Workshop (GCE). New Orleans, Louisiana:Institute of Electrical & Electronics Engineers (IEEE).

Mirarab S., Bayzid M.S., Boussau B., Warnow T. 2014. Statistical binning enables an accurate coalescent-based estimation of the avian tree. Science. 346(6215):1250463.

Mirarab S., Warnow T. 2015. ASTRAL-II: coalescent-based species tree estimation with many hundreds of taxa and thousands of genes. Bioinformatics. 31:i44–i52.

Moore B.R., Höhna S., May M.R., Rannala B., Huelsenbeck J.P. 2016. Critically evaluating the theory and performance of Bayesian analysis of macroevolutionary mixtures. Proc. Natl. Acad. Sci. USA. 113:9569–9574.

Moran R., Morgan C., O'Connell M. 2015. A guide to phylogenetic reconstruction using heterogeneous models—a case study from the root of the placental mammal tree. Computation. 3:177–196.

Nguyen N.D., Mirarab S., Kumar K., Warnow T. 2015. Ultra-large alignments using phylogeny-aware profiles. Genome Biol. 16:124.

Nygaard S., Zhang G., Schiøtt M., Li C., Wurm Y., Hu H., Zhou J., Ji L., Qiu F., Rasmussen M., Pan H., Hauser F., Krogh A., Grimmelikhuijzen C.J., Wang J., Boomsma J.J. 2011. The genome of the leaf-cutting ant acromyrmex echinatior suggests key adaptations to advanced social life and fungus farming. Genome Res. 21:1339–1348.

Oxley P.R., Ji L., Fetter-Pruneda I., McKenzie S.K., Li C., Hu H., Zhang G., Kronauer D.J. 2014. The genome of the clonal raider ant *Cerapachys biroi*. Curr. Biol. 24:451–458.

Paradis E. 2013. Molecular dating of phylogenies by likelihood methods: a comparison of models and a new information criterion. Mol. Phylogenet. Evol. 67:436–444.

Paradis E., Claude J., Strimmer K. 2004. APE: Analyses of phylogenetics and evolution in R language. Bioinformatics. 20:289–290.

Peters M.K., Likare S., Kraemer M. 2008. Effects of habitat fragmentation and degradation on flocks of African ant-following birds. Ecol. Appl. 18:847–858.

Philippe H., Forterre P. 1999. The rooting of the universal tree of life is not reliable. J. Mol. Evol. 49:509–523.

Plummer M., Best N., Cowles K., Vines K. 2006. Coda: Convergence diagnosis and output analysis for mcmc. R News. 6:7–11.

Rabosky D.L. 2014. Automatic detection of key innovations, rate shifts, and diversity-dependence on phylogenetic trees. PLoS One. 9:e89543.

Rabosky D.L. 2018. BAMM at the court of false equivalency: a response to Meyer and Wiens. Evolution 72(10):2246–2256.

Rabosky D.L., Grundler M., Anderson C., Title P., Shi J.J., Brown J.W., Huang H., Larson J.G. 2014. BAMMtools: an R package for the analysis of evolutionary dynamics on phylogenetic trees. Methods Ecol. Evol. 5:701–707.

Rabosky D.L., Mitchell J.S., Chang J. 2017. Is BAMM flawed? Theoretical and practical concerns in the analysis of multi-rate diversification models. Syst. Biol. 66:477–498.

Raignier A., Van Boven J., du Congo Belge M.R. 1955. Étude taxonomique, biologique et biométrique des *Dorylus* du sous-genre *Anomma* (Hymenoptera Formicidae). Ann. Mus. R. Congo Belge Nouv. Sér. Quarto Sci. Zool. 2:1–359.

Ravary F., Jaisson P. 2002. The reproductive cycle of thelytokous colonies of *Cerapachys biroi* Forel (Formicidae, Cerapachyinae). Insectes Sociaux. 49:114–119.

Roch S., Warnow T. 2015. On the robustness to gene tree estimation error (or lack thereof) of coalescent-based species tree methods. Syst. Biol. 64:663–676.

Rodríguez-Ezpeleta N., Brinkmann H., Roure B., Lartillot N., Lang B.F., Philippe H. 2007. Detecting and overcoming systematic errors in genome-scale phylogenies. Syst. Biol. 56:389–399.

Rohland N., Reich D. 2012. Cost-effective, high-throughput DNA sequencing libraries for multiplexed target capture. Genome Res. 22:939–946.

Salichos L., Rokas A. 2013. Inferring ancient divergences requires genes with strong phylogenetic signals. Nature. 497:327–331.

Sanderson M.J. 2002. Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. Mol. Biol. Evol. 19:101–109.

Sanmartín I., Enghoff H., Ronquist F. 2001. Patterns of animal dispersal, vicariance and diversification in the Holarctic. Biol. J. Linnean Soc. 73:345–390.

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

Schneirla T.C. 1945. The army-ant behavior pattern: nomad-statary relations in the swarmers and the problem of migration. Biol. Bull. 88:166.

Schöning C., Njagi W.M., Franks N.R. 2005. Temporal and spatial patterns in the emigrations of the army ant *Dorylus (Anomma) molestus* in the montane forest of Mt Kenya. Ecol. Entomol. 30:532–540.

Schrader L., Kim J.W., Ence D., Zimin A., Klein A., Wyschetzki K., Weichselgartner T., Kemena C., Stökl J., Schultner E., Wurm Y., Smith C.D., Yandell M., Heinze J., Gadau J., Oettler J. 2014. Transposable element islands facilitate adaptation to novel environments in an invasive species. Nat. Commun. 5:5495.

Shen X.-X., Hittinger C.T., Rokas A. 2017. Contentious relationships in phylogenomic studies can be driven by a handful of genes. Nat. Ecol. Evol. 1:0126.

Smith C.D., Zimin A., Holt C., Abouheif E., Benton R., Cash E., Croset, V., Currie C.R., Elhaik E., Elsik C.G., Fave M.-J., Fernandes V., Gadau J., Gibson J.D., Graur D., Grubbs K.J., Hagen D.E., Helmkampf M., Holley J.-A., Hu H., Viniegra A.S.I., Johnson B.R., Johnson R.M., Khila A., Kim J.W., Laird J., Mathis K.A., Moeller J.A., Munoz-Torres M.C., Murphy M.C., Nakamura R., Nigam S., Overson R.P., Placek J.E., Rajakumar R., Reese J.T., Robertson H.M., Smith C.R., Suarez A.V., Suen G., Suhr E.L., Tao S., Torres C.W., van Wilgenburg E., Viljakainen L., Walden K. K.O., Wild A.L., Yandell M., Yorke J.A., Tsutsui N.D. 2011a. Draft genome of the globally widespread and invasive Argentine ant (*Linepithema humile*). Proc. Natl. Acad. Sci. USA. 108:5673–5678.

Smith C.R., Cahan S.H., Kemena C., Brady S.G., Yang W., Bornberg-Bauer E., Eriksson T., Gadau J., Helmkampf M., Gotzek D., Miyakawa M.O., Suarez A.V., Mikheyev A. 2015. How do genomes create novel phenotypes? Insights from the loss of the worker caste in ant social parasites. Mol. Biol. Evol. 32:2919–2931.

Smith C.R., Smith C.D., Robertson H.M., Helmkampf M., Zimin A., Yandell M., Holt C., Hu H., Abouheif E., Benton R., Cash E., Croset V., Currie C.R., Elhaik E., Elsik C.G., Fave M.-J., Fernandes V., Gibson J.D., Graur D., Gronenberg W., Grubbs K.J., Hagen

D.E., Viniegra A. S.I., Johnson B.R., Johnson R.M., Khila A., Kim J.W., Mathis K.A., Munoz-Torres M.C., Murphy M.C., Mustard J.A., Nakamura R., Niehuis O., Nigam S., Overson R.P., Placek J.E., Rajakumar R., Reese J.T., Suen G., Tao S., Torres C.W., Tsutsui N.D., Viljakainen L., Wolschin F., Gadau J. 2011b. Draft genome of the red harvester ant *Pogonomyrmex barbatus*. Proc. Natl. Acad. Sci. USA. 108:5667–5672.

Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 30:1312–1313.

Suen G., Teiling C., Li L., Holt C., Abouheif E., Bornberg-Bauer E., Bouffard P., Caldera E.J., Cash E., Cavanaugh A., Denas O., Elhaik E., Favé M.-J., Gadau J., Gibson J.D., Graur D., Grubbs K.J., Hagen D.E., Harkins T.T., Helmkampf M., Hu H., Johnson B.R., Kim J., Marsh S.E., Moeller J.A., Muñoz-Torres M.C., Murphy M.C., Naughton M.C., Nigam S., Overson R., Rajakumar R., Reese J.T., Scott J.J., Smith C.R., Tao S., Tsutsui N.D., Viljakainen L., Wissler L., Yandell M.D., Zimmer F., Taylor J., Slater S.C., Clifton S.W., Warren W.C., Elsik C.G., Smith C.D., Weinstock G.M., Gerardo N.M., Currie C.R. 2011. The genome sequence of the leaf-cutter ant *Atta cephalotes* reveals insights into its obligate symbiotic lifestyle. PLoS Genet. 7:e1002007.

Sukumaran J., Holder M.T. 2010. DendroPy: a Python library for phylogenetic computing. Bioinformatics. 26:1569–1571.

Talavera G., Castresana J. 2007. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. Syst. Biol. 56:564–577.

Tan G., Muffato M., Ledergerber C., Herrero J., Goldman N., Gil M., Dessimoz C. 2015. Current methods for automated filtering of multiple sequence alignments frequently worsen single-gene phylogenetic inference. Syst. Biol. 64:778–791.

Team R.C. 2014. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing 2013.

Tsuji K., Yamauchi K. 1995. Production of females by parthenogenesis in the ant, *Cerapachys biroi*. Insectes Sociaux. 42:333–336.

Wurm Y., Wang J., Riba-Grognuz O., Corona M., Nygaard S., Hunt B.G., Ingram K.K., Falquet L., Nipitwattanaphon M., Gotzek D., Dijkstra M.B., Oettler J., Comtesse F., Shih C.-J., Wu W.-J., Yang C.-C., Thomas J., Beaudoing E., Pradervand S., Flegel V., Cook E.D., Fabbretti R., Stockinger H., Long L., Farmerie W.G., Oakey J., Boomsma J.J., Pamilo P., Yi S.V., Heinze J., Goodisman M.A.D., Farinelli L., Harshman K., Hulo N., Cerutti L., Xenarios I., Shoemaker D., Keller L. 2011. The genome of the fire ant *Solenopsis invicta*. Proc. Natl. Acad. Sci. USA. 108:5679–5684.

Zhong M., Hansen B., Nesnidal M., Golombek A., Halanych K.M., Struck T.H. 2011. Detecting the symplesiomorphy trap: a multigene phylogenetic analysis of terebelliform annelids. BMC Evol Biol. 11:369.