Draft *Aphaenogaster* genomes expand our view of ant variation across climate gradients

Matthew K. Lau[1], Aaron M. Ellison[1], Andrew Nguyen[2,3], Clint Penick[4,5], Bernice DeMarco[6], Nicholas J. Gotelli[2], Nathan J. Sanders[7], Robert R. Dunn[4], and Sara Helms Cahan[2]

[1]**Harvard Forest, Harvard University, Petersham, MA, USA**
[2]**Department of Biology, University of Vermont, Burlington, VT, USA**
[3]**Department of Entomology and Nematology, University of Florida, Gainesville, FL, USA**
[4]**Department of Applied Ecology, North Carolina State University, Raleigh, NC, USA**
[5]**The Biomimicry Center, Arizona State University, Tempe, AZ, USA**
[6]**Smithsonian Institution, Washington, DC, USA**
[7]**Environmental Program, Rubenstein School of Environment and Natural Resources, University of Vermont, Burlington, VT, USA**

Corresponding author:
Matthew K. Lau

Email address: matthewklau@fas.harvard.edu

ABSTRACT

Given the abundance, broad distribution, and diversity of roles that ants play in many ecosystems, they are an ideal group to serve as ecosystem indicators of climatic change. At present, only a few whole-genome sequences of ants are available (19 of $>$ 16,000 species), mostly from tropical and sub-tropical species. To address this limited sampling, we sequenced genomes of temperate-latitude species from the genus *Aphaenogaster*, a genus with important seed dispersers. In total, we sampled seven colonies of six species: *A. ashmeadi*, *A. floridana*, *A. fulva*, *A. miamiana*, *A. picea*, and *A. rudis*. The geographic ranges of these species collectively span eastern North America from southern Florida to southern Canada, which encompasses a latitudinal gradient in which many climatic variables are changing rapidly. For the six genomes, we assembled an average of 271,039 contigs into 47,337 scaffolds. The mean genome size was 370.45 Mb, which was comparable to that of other sequenced ant genomes (212.83 to 396.03 Mb). In an analysis of currently sequenced ant genomes and the new *Aphaenogaster* sequences, we observed relationships between biogeographic variables and genome similarity and size. The highest correlations were between genomic similarity and two major groups of climatic variables related to air temperature and precipitation. Although much more ant genomic sampling remains to be done, results for *Aphaenogaster* are consistent with earlier hypotheses relating climate and genomic variation in ants.

INTRODUCTION

Understanding how terrestrial ecosystems will respond to ongoing shifts in climatic variables, such as temperature and precipitation, will improve our ability to manage communities and mitigate impacts of climatic change. The mean global temperature is currently on track to meet or exceed that predicted by the most extreme forecasting models (Brown and Caldeira, 2017). Climatic change also is pushing local conditions outside the boundaries of historic ranges, potentially leading to combinations of species or

entire ecosystems that have no contemporary analogs (Burrows et al., 2014). As climate-driven impacts on evolutionary responses are likely to occur over contemporary time scales, with the potential for ecological and evolutionary dynamics to affect communities and ecosystem processes (Rowntree et al., 2011; Des Roches et al., 2017), there is a need for a comprehensive study of the genetic basis of species' responses to climate (Parmesan, 2006).

The biodiversity of most terrestrial systems is great enough to be intractable to study in its entirety. To deal with this, researchers often study 'indicator' species whose responses to environmental change are broadly representative of a much wider range of taxa (Siddig et al., 2016). Ants (Formicidae), in particular, are widely used as indicator taxa (Agosti et al., 2000) as they play key roles in community dynamics and ecosystem processes, including key interactions, such as seed dispersal and the movement of soil via colony construction (Del Toro et al., 2012). Ants also are responsive to changes in temperature and other climatic variables via individual responses, changes in colony structure and community assembly (Kaspari et al., 2015; Spicer et al., 2017; Diamond et al., 2017; Diamond and Chick, 2018).

Multiple studies support the perspective that a more complete knowledge of ant genetics will increase our understanding of ant responses to environmental change (Diamond et al., 2012; Nygaard and Wurm, 2015; Stanton-Geddes et al., 2016; Boomsma et al., 2017; Penick et al., 2017). If evolutionary dynamics in ants have been influenced by environmental conditions, then ant genomes from more similar conditions would tend to have more similar genomes (Thompson, 1999). Also, empirical studies of have reported biogeographic patterns in Gerome size in arthropod taxa, e.g. Crustacea (Hultgren et al., 2018). Also, recent studies of insect genomes (Alfsnes et al., 2017) have demonstrated support for climatic constraints to genome size, including the the potential for cold temperatures to select for smaller genomes (Mousseau, 1997; Petrov, 2001; Alfsnes et al., 2017).

At present relatively few ant species have been sequenced —20 in total, of which 19 are currently available in the NCBI Genome Database (accessed Aug 8 2018, see Supplementary Materials Table 1). Of these, most are from tropical and subtropical assemblages (Fig 1), and all but five represent unique genera (the exceptions being two species of *Atta* and three of *Trachymyrmex*). No species of *Aphaenogaster* have yet been sequenced, even though previous studies have shown that *Aphaenogaster* species' rapid responses to climatic change appear to depend both on species identity and on the geographic region in which climatic change occurs.

To increase the number of genomes of temperate-zone ant species for other genetic studies (e.g. short-read and target sequences or transcriptomics), we sequenced the genomes of *Aphaenogaster* species, which are abundant ants that play key roles in the dispersal of understory plant species in North America and temperate Asia. We conducted whole genome sequencing for six species: *A. ashmeadi*, *A.*

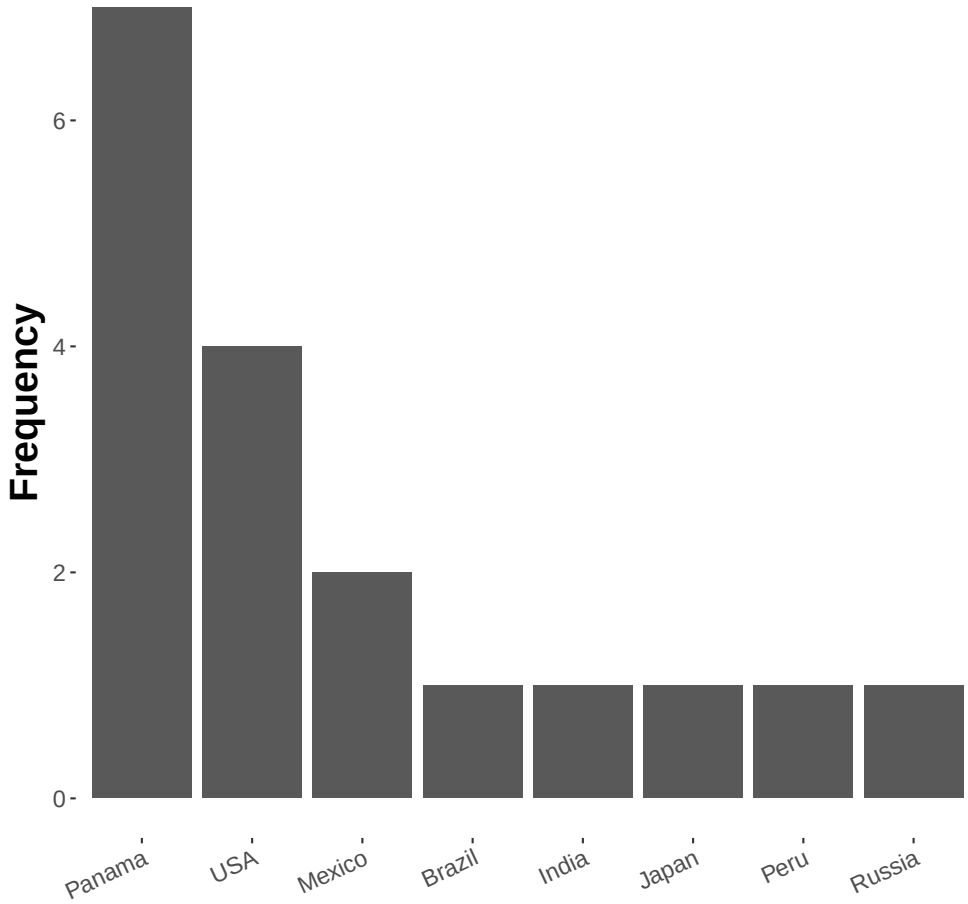

Figure 1. Number of whole-genome sequences available in NCBI by country (accessed April 2018).

floridana, *A. fulva*, *A. miamiana*, *A. picea* and *A. rudis*. These species were collected from across a broad biogeographic gradient spanning 10 degrees of longitude and 12 degrees of latitude. We also conducted an initial exploration of the potential support for the hypothesis that genomes from more similar climates would be more similar to each other in terms of sequence similarity and size. To do this we analyzed the newly collected *Aphaenogaster* sequences together with other publicly available ant whole genome sequences. We present the new genome sequences as a tool for ecological genomics and discuss current patterns of ant genome characteristics in the context of all currently sequenced ant genomes.

RESULTS

Whole-genome Sequencing

Entire colonies of the six *Aphaenogaster* species were collected by A. Nguyen and C. Penick from field sites in eastern North America (Fig 2). Ants were identified to species and voucher specimens have been deposited at the Museum of Comparative Zoology, Harvard University. Individuals from each colony were

isolated from nest material and debris, weighed, placed in 50 ml Falcon centrifuge tubes, and immediately flash frozen in a -80° C freezer. Colony weights were: 794 mg (*A. ashmeadi*), 652 mg (*A. floridana*), 520 mg (*A. fulva*), 749 mg (*A. picea*), 862 mg (*A. miamiana*), 280 mg (*A. rudis* 1) and 236 mg (*A. rudis* 2).

Figure 2. We sampled seven colonies representing six species of *Aphaenogaster* from across eastern North America (see Table 1). All photos by April Noble (available from http://www.antweb.org).

Whole colony DNA was used to have sufficient concentrations for sequencing. DNA was then extracted from each colony using methods developed previously for genomic sequencing of whole colonies of colonial mosquitos (*Anopheles* spp.) (Neafsey et al., 2010) and sequenced using an Illumina HiSeq 2500 at the Broad Institute (Cambridge, MA, USA). A combination of fragment and jump sequences were used to generate higher quality, long sequence reads. Raw sequences were processed to remove chimeric and contaminant sequences, screened for contaminants by BLAST searches (using *blastn*) to identify sequences with likely matches to non-target species (primarily *Wolbachia* and *Mycoplasma*), and assembled using ALLPATHS-LG (version r48559) (Gnerre et al., 2011). Additional assembly processing using PILON (version 1.13) (Walker et al., 2014) was applied to reduce base-call errors and gaps in coverage. On average, across all seven genomes, PILON reduced coverage gaps by 3.1% or 3.9 Mb. GAEMR (http://www.broadinstitute.org/software/gaemr/) software produced summary statistics of the final assembled genomes.

Genome Quality and Composition

DNA extractions yielded substantial amounts of high quality DNA with concentrations and quality scores ranging from 3.45–5.39 ngμL^{-1} and 4.05–4.27 ngμL^{-1}, respectively. All genome assemblies displayed good coverage, with an average of 70% of fragments mapped (Table 2). Across all species, the length of the shortest contig at 50% of the genome (i.e. N50) was 18,864 bases; average assembly GC content was

	Lat	Lon	Tmin (C)	Tmax (C)	Precip (mm)
Aphaenogaster ashmeadi	29.79	-82.03	5.80	32.70	1314
Aphaenogaster floridana	29.79	-82.03	5.80	32.70	1314
Aphaenogaster fulva	32.69	-82.51	1.30	33.30	1155
Aphaenogaster miamiana	29.66	-82.30	5.90	32.80	1322
Aphaenogaster picea	42.60	-72.58	-12.40	28.30	1122
Aphaenogaster rudis1	36.02	-78.98	-2.70	31.50	1164
Aphaenogaster rudis2	36.02	-78.98	-2.70	31.50	1164

Table 1. Climate variables for colony sample sites. Climate are 30 year normal values (1970-2000) for January minimum temperature (Tmin), July maximum temperature (Tmax) and total precipitation (Precip) from the WorldClim database accessed on 08 August 2018.

38.18%; and average genome size was 370.45 Mb. Using GAEMR's BLAST feature to conduct a search of the contigs against the NCBI's nucleotide sequence database, we discovered that 38.98% and 22.04% of the top hits were "ant" and *Aphaenogaster*, respectively. The *Aphaenogaster* genomes compared well with other ant genome sequences. The sizes of the *Aphaenogaster* genomes were within the range of other ant genomes based on size from both flow-cell cytometry (Tsutsui et al., 2008) and the previously sequenced ant genomes available in NCBI (Fig 3). The size of scaffolds size was within the size range that is theoretically large enough to contain genes, as suggested by Efron and Tibshirani (2007).

	A. ashmeadi	*A. floridana*	*A. fulva*	*A. miamiana*	*A. picea*	*A. rudis1*	*A. rudis2*
Total Scaffold Length (Mb)	310.33	382.86	346.13	342.64	386.04	395.41	429.70
Coverage (%)	81.46	71.88	70.70	77.40	67.47	66.49	65.59
Scaffold N50 (bp)	336807.00	439114.00	255328.00	351517.00	322984.00	300103.00	269776.00
Scaffolds	5087.00	6422.00	7031.00	6920.00	6808.00	7404.00	7665.00
Max Gap (bp)	13070.00	15108.00	12104.00	11453.00	14952.00	18586.00	24564.00
Captured Gaps	26350.00	30858.00	32881.00	28801.00	36417.00	34062.00	34313.00
Total Gap Length (Mb)	57.69	107.89	101.40	77.64	125.15	131.71	148.75
Total Contig Length (Mb)	252.64	274.96	244.73	265.00	260.90	263.70	280.95
Contig N50 (bp)	21677.00	23448.00	15753.00	20738.00	15440.00	15622.00	18941.00
Contigs	31437.00	37280.00	39912.00	35721.00	43225.00	41466.00	41978.00
Assembly GC (%)	38.27	38.03	38.39	38.21	38.32	38.25	37.88
Contaminants (%)	0.30	0.24	0.02	0.26	1.14	1.25	0.61

Table 2. Sequencing statistics for the genomes of the sequenced colonies of *Aphaenogaster*.

We observed patterns in genomic composition that generally were consistent with expectations based on phylogenetic relatedness. After detecting and masking repeat regions in the *Aphaenogaster* genomes using *Repeatmasker* (version 4.0.5 Institute for Systems Biology), we applied MASH distance (Ondov et al., 2016) to measure pairwise dissimilarity of genomic sequences. The MASH method extends a data compression and dimensionality-reduction algorithm to generate estimates of sequence similarity with low computational overhead. Briefly, the pairs of genomic sequences were pre-processed into sets of k-mers of size 21 with the size of the non-redundant hashes retained set to 1,000. These settings have been demonstrated to provide good representation of genomic similarity with minimal computational costs (Ondov et al., 2016). These sets were then used to estimate the Jaccard similarity coefficient (the ratio of shared k-mers to total k-mers) of subsampled k-mer pairs of genomes. This unbiased estimate of the Jaccard similarity (J) was then used to calculate the dissimilarity of the two genomes (D) as $D = 1 - J$.

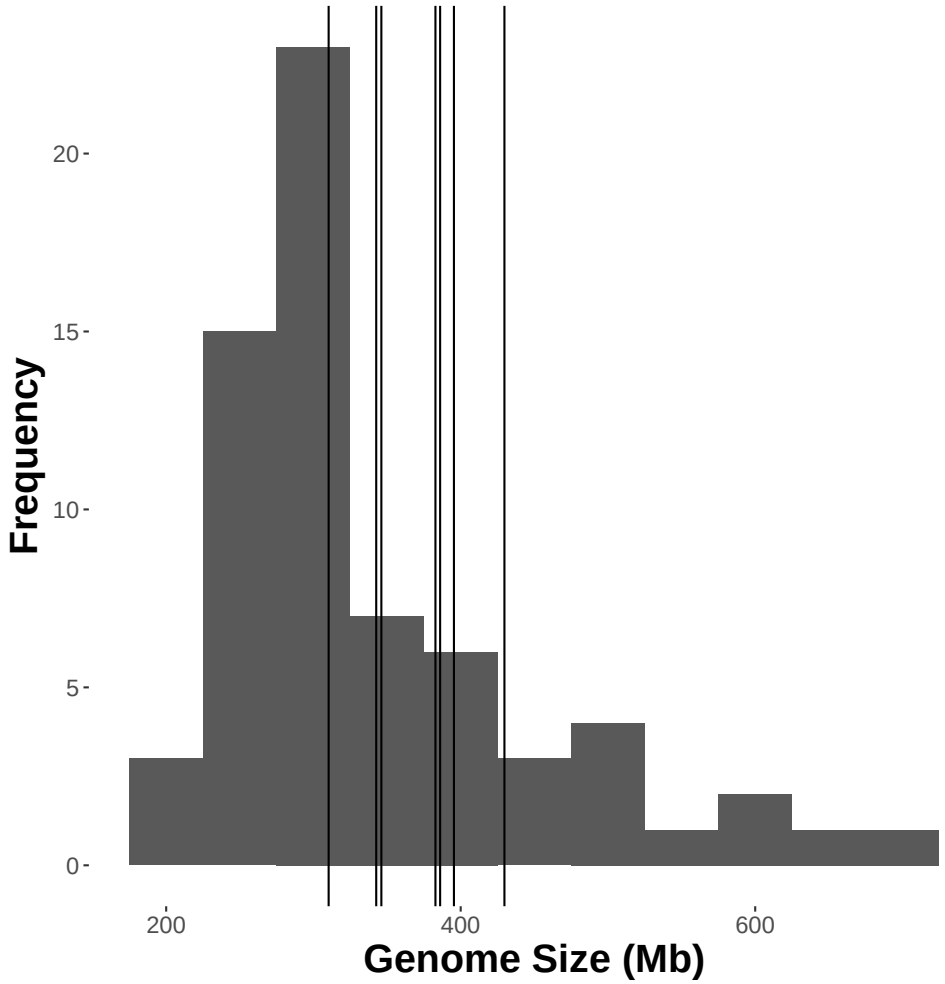

Figure 3. The size of sequenced *Aphaenogaster* genomes were within the size range of previously published observed or estimated genomes of ants. Frequency distribution of previously published genome size estimates using flow cytometry from Tsutsui et al. (2008) and those available via NCBI (accessed April 2018). Vertical lines identify the sizes of the *Aphaenogaster* assemblies (see Table 2).

All Jaccard similarity estimates had p-values less than 10^{-14}, which is below the recommended 10^{-3} probability of observing values of J due to chance.

Using the MASH genomic distances, we observed patterns of genomic similarity that were in line with expectations from established ant phylogenetics. Sequences formed groups that corresponded with subfamilies (Fig 4). *Aphaenogaster* clustered with other genera from the Myrmicinae and, in general, subfamily level clustering tended to follow previously observed patterns of subfamily relatedness (Bolton, 2006; Moreau, 2006; Ward, 2014). The *Aphaenogaster* sequences formed a single cluster containing only *Aphaenogaster* species and displayed intra-generic levels of genomic variance comparable to other genera s(e.g., *Trachymyrmex* spp.). The separation of the two *A. rudis* species was initially surprising, as these two samples were collected at the same site (Duke Forest, NC USA) and were identified as the same species based on their morphological characteristics (Ellison, 2012; DeMarco and Cognato,

2016). However, two recent studies of targeted gene regions have demonstrated the polyphyletic nature of *Aphaenogaster rudis*. One study of the evolution of the subfamily Myrmicinae observed that the genus as a whole could be split into at least four different lineages (Ward et al., 2015). Another, more detailed study of the genus in North America found that multiple individuals of *A. rudis* separated out into distinct groupings, each with other species, specifically, individuals of *A. rudis* from North Carolina (USA) were observed to form distinct clusters with individuals of *A. carolinensis*, *A. miamiana*, *A. lamellidens* and *A. texana* (DeMarco and Cognato, 2016).

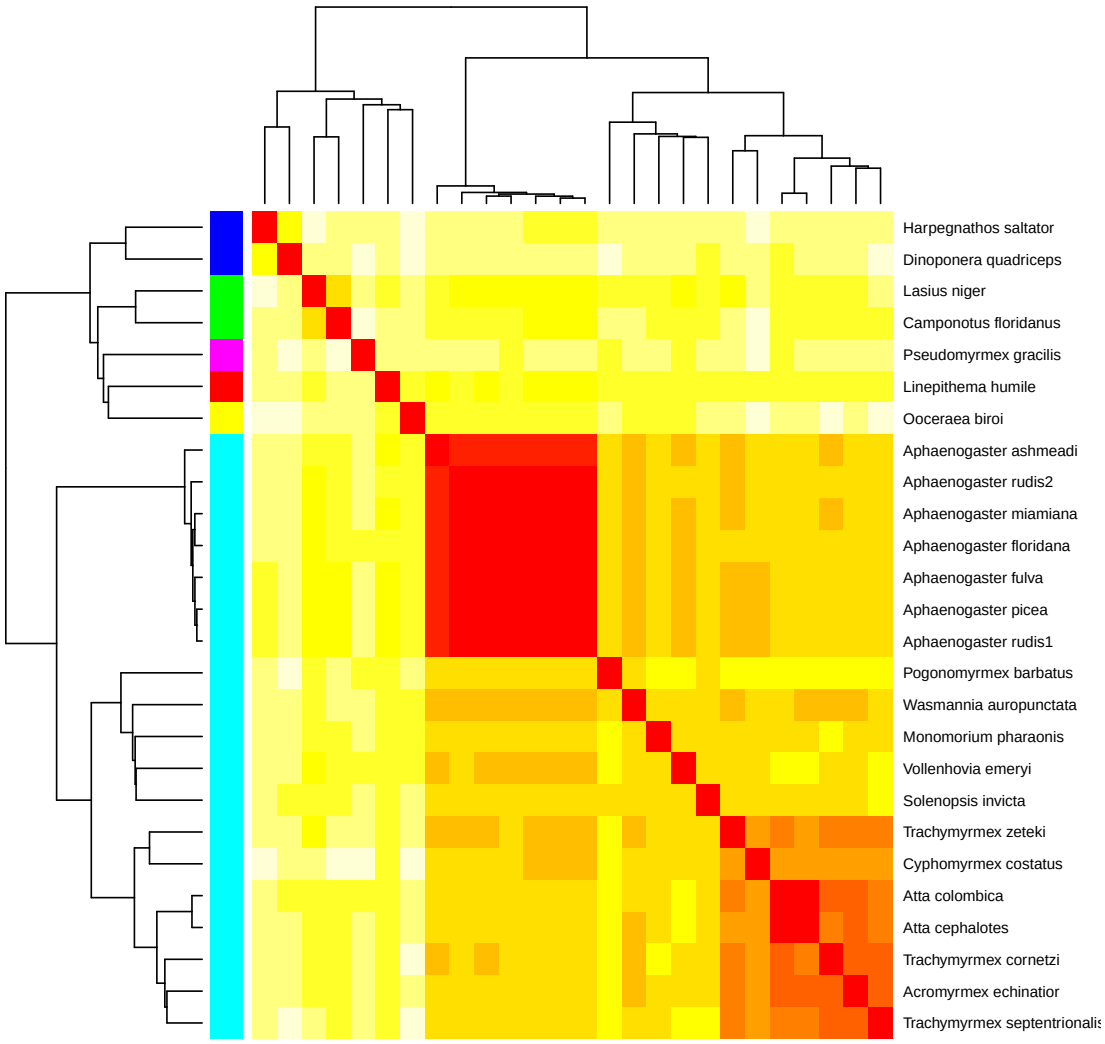

Figure 4. Heatmap of the MASH genomic distances of the *Aphaenogaster* species that we sampled together with other ant species in NCBIs. Heat colors shown in the central matrix range from high (white = 1) through moderate (orange = 0.5) to low (red = 0) genomic distance; the diagonal is entirely red because it illustrates the distance of each sequence to itself. The cladograms on the left and top show hierarchical clustering of the genomes. Colors shown to the left of the matrix indicate ant subfamilies: *Ponerinae* (dark blue), *Formicinae* (green), *Pseudomyrmecinae* (pink), *Dolichoderinae* (red), *Dorylinae* (yellow), *Myrmicinae* (light blue).

Biogeographic Patterns of Ant Genomes

To examine biogeographic relationships of *Aphaenogaster* genomes, we used multivariate correlation analyses (Mantel Tests) of interspecies whole-genome size similarity on the Euclidean distance of whole-genome length (total base pairs) and genomic similarity (MASH distance) with the Euclidean distances of standardized climatic variables. More specifically, we used directional ($H_\circ$: Mantel $r \leq 0$) partial mantel tests to control for spatial autocorrelation by including geodesic distance as a term (Goslee and Urban, 2007). Two species, (*W. auropunctata* and *M. pharaonis*), which did not have published location information, were excluded from biogeographic analyses (see Supplementary Materials Table 1). Data for climatic variables for each sampling location from the WorldClim database (version 2.0) at a 2.5 arc minute spatial resolution from the years 1970 to 2000 (Fick and Hijmans, 2017)

Here, it is important to note that we are using assembly size as an indicator of genome size. As genome size estimates, using methods such as flow cytometry (Tsutsui et al., 2008), are generally used to set assembly size targets for whole genome sequencing efforts (Hare and Johnston, 2011), we expect there to be a high degree of correlation between assembly size and genome size. Also, as a test of the potential correlation between assembly and true genome size, we examined the correlation between average assembly sizes of ant genera in the NCBI genome database that overlapped with flow cytometry estimates of ant genera published in Tsutsui et al. (2008). We found a marginally significant correlation assembly size and flow cytometry at the genus level (Spearman's $\rho = 0.57$, *p-value* = 0.076), which supports the use of assembly size as a useful indicator of genome size.

WorldClim Variable	BIO Number
Annual Mean Temperature (MAT)	BIO1
Mean Diurnal Range (MDR)	BIO2
Isothermality (Iso)	BIO3
Temperature Seasonality (TS)	BIO4
Max Temperature of Warmest Month (Tmax)	BIO5
Min Temperature of Coldest Month (Tmin)	BIO6
Temperature Annual Range (ATR)	BIO7
Mean Temperature of Wettest Quarter (MTWeQ)	BIO8
Mean Temperature of Driest Quarter (MTDQ)	BIO9
Mean Temperature of Warmest Quarter (MTWaQ)	BIO10
Mean Temperature of Coldest Quarter (MTCQ)	BIO11
Annual Precipitation (PA)	BIO12
Precipitation of Wettest Month (PWM)	BIO13
Precipitation of Driest Month (PDM)	BIO14
Precipitation Seasonality (PS)	BIO15
Precipitation of Wettest Quarter (PWeQ)	BIO16
Precipitation of Driest Quarter (PDQ)	BIO17
Precipitation of Warmest Quarter (PWaQ)	BIO18
Precipitation of Coldest Quarter (PCQ)	BIO19

Table 3. WorldClim variables, abbreviations and numbers for the climate variables used in the analysis of size and MASH similiarity of ant genomes.

Using a permutational multivariate analysis of variance (PerMANOVA) procedure, we parsed the individual variables that were correlated with both genome size and MASH similarity. PerMANOVA is a flexible multivariate analog of ANOVA that permits the use of a wider set of similarity metrics to be used for the response matrix (Anderson, 2001), such as the MASH distance. We ran a total of 10,000 permutations of the original distance matrices for each statistical permutation procedure. We chose a subset of all possible climate variables available via WorldClim for this analysis. A visual inspection of the sampled climate variable correlations indicated that the primary climate variables, mean annual temperature (MAT), annual minimum temperature, annual maximum temperature, annual precipitation and summer precipitation, represented the majority of climate variation (Fig 5). Based on this, we only included these variables, along with latitude and longitude coordinates, as factors in the PerMANOVAs.

To visualize the patterns of genomic similarity and spatio-climate variation, we applied non-metric multidimensional scaling (NMDS) ordination to the MASH genomic distances using 500 iterations, which yielded a two-dimensional lowest stress solution for all genomes and only the *Aphaenogaster* genomes, respectively. The R^2 and stress of the final solutions were 0.80 and 15%. The geographic (latitude and longitude) and WorldClim climate variables were then correlated with both sets of MASH genomic distances (i.e. all and just *Aphaenogaster*) using a vector analyses (Oksanen et al., 2016).

We found significant global, biogeographic patterns of ant species genomes. Across all whole-genome ant sequences (both the NCBI and the newly sequenced *Aphaenogaster* species), ants from climatically similar locations tended to have similar genomes (Fig 6). We also observed that collection-location climate similarity was significantly correlated with genome-size similarity (Mantel r = 0.19, p-value = 0.021) and whole-genome similarity (MASH distance) (Mantel r = 0.3248169, p-value = 0.001).

Both space and climate were associated with the size and genomic similarity of the ant genomes. Longitude but not latitude was a significant predictor of genome size (Table 4). Temperature of the coldest (Tmin) and hottest (Tmax) month and total annual precipitation (PA), were all significant predictors of genomic size similarity, but neither mean annual temperature (MAT) nor summer precipitation (PS) were significant predictors of genome size. Overall, Tmin had the highest correlation with an R^2 of 0.23. Latitude and longitude were both correlated with MASH genome distance; and all climate variables examined were significant predictors of whole-genome similarity with Tmin (R^2 = 0.10) also being the predictor with the largest correlation. When the newly sequenced *Aphaenogaster* genomes were excluded from the analysis, climate was not correlated with genome size similarity (Mantel r = 0.13, p-value = 0.190) and only annual precipitation (PA) was a significant predictor of genome size similarity, and longitude and mean annual temperature (MAT) were significant predictors of MASH genomic similarity (see Supplementary Materials Table 2).

	df	*SS*	*MS*	*Pseudo-F*	*R2*	p-*value*
Size Distance						
Lat	1	3360.41	3360.41	2.34	0.04	0.1433
Lon	1	9238.80	9238.80	6.43	0.11	0.0181
MAT	1	267.49	267.49	0.19	0.00	0.6767
Tmin	1	20413.36	20413.36	14.21	0.23	0.0025
Tmax	1	9081.67	9081.67	6.32	0.10	0.0217
PA	1	17564.07	17564.07	12.23	0.20	0.0034
PS	1	4368.07	4368.07	3.04	0.05	0.0978
Residuals	16	22985.41	1436.59		0.26	
Total	23	87279.28			1.00	
MASH Distance						
Lat	1	0.02	0.02	3.56	0.10	0.0002
Lon	1	0.02	0.02	3.26	0.10	0.0017
MAT	1	0.01	0.01	1.97	0.06	0.0341
Tmin	1	0.02	0.02	3.30	0.10	0.0004
Tmax	1	0.01	0.01	1.89	0.06	0.0382
PA	1	0.01	0.01	1.97	0.06	0.0276
PS	1	0.01	0.01	2.14	0.06	0.0159
Residuals	16	0.11	0.01		0.47	
Total	23	0.24			1.00	

Table 4. PerMANOVA pseudo-F table for the analysis of the factors correlated with ant genome size and MASH distance.

Data, Computation, and Statistics

The raw and assembled genome sequences are currently stored at Harvard Forest (Petersham, MA, USA) and NCBI's genome database (Genome Accessions NJRK00000000-NJRQ00000000 and BioSample Accessions SAMN06892346-SAMN06892352). Genomic distance (MASH) computations were run on the Odyssey cluster supported by the FAS Division of Science, Research Computing Group at Harvard University. All analyses were conducted in **R** (R Core Team, 2017). Analytical scripts for the project have been versioned and archived (DOI: 10.5281/zenodo.1341982) and are available online at https://zenodo.org/record/1341982. We used the *vegan* (Oksanen et al., 2016) and *ecodist* (Goslee and Urban, 2007) packages in R for multivariate analyses.

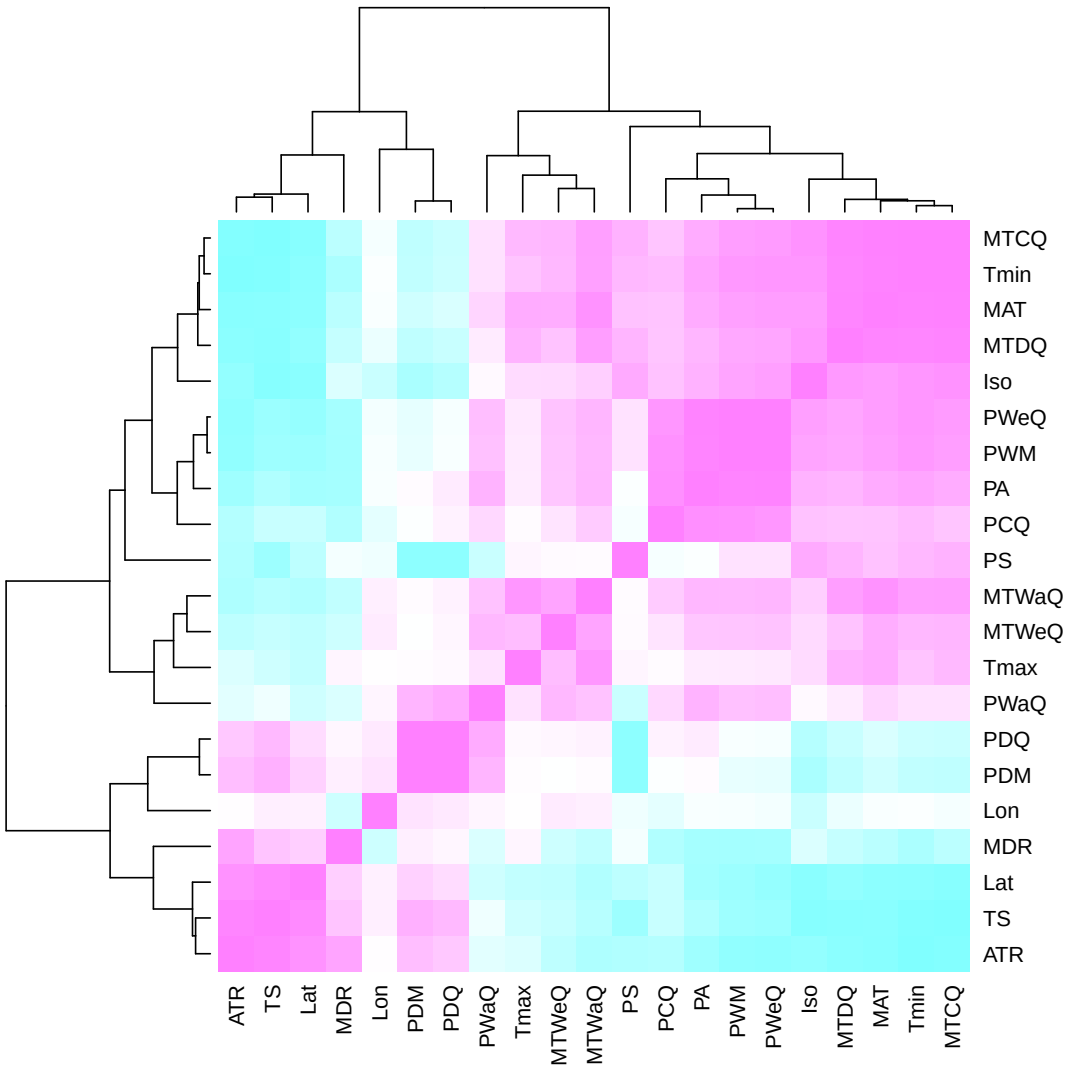

Figure 5. Heatmap of Pearson correlations among climate variables. Cells in the heatmap are colored by the correlation between the two variables that intersect at that location ranging from blue = -1 to white = 0 to pink = 1. The variables are arrayed by hierarchical clustering of the correlations, as shown by the dendrograms on the top and left side. For variable descriptions see Table 3.

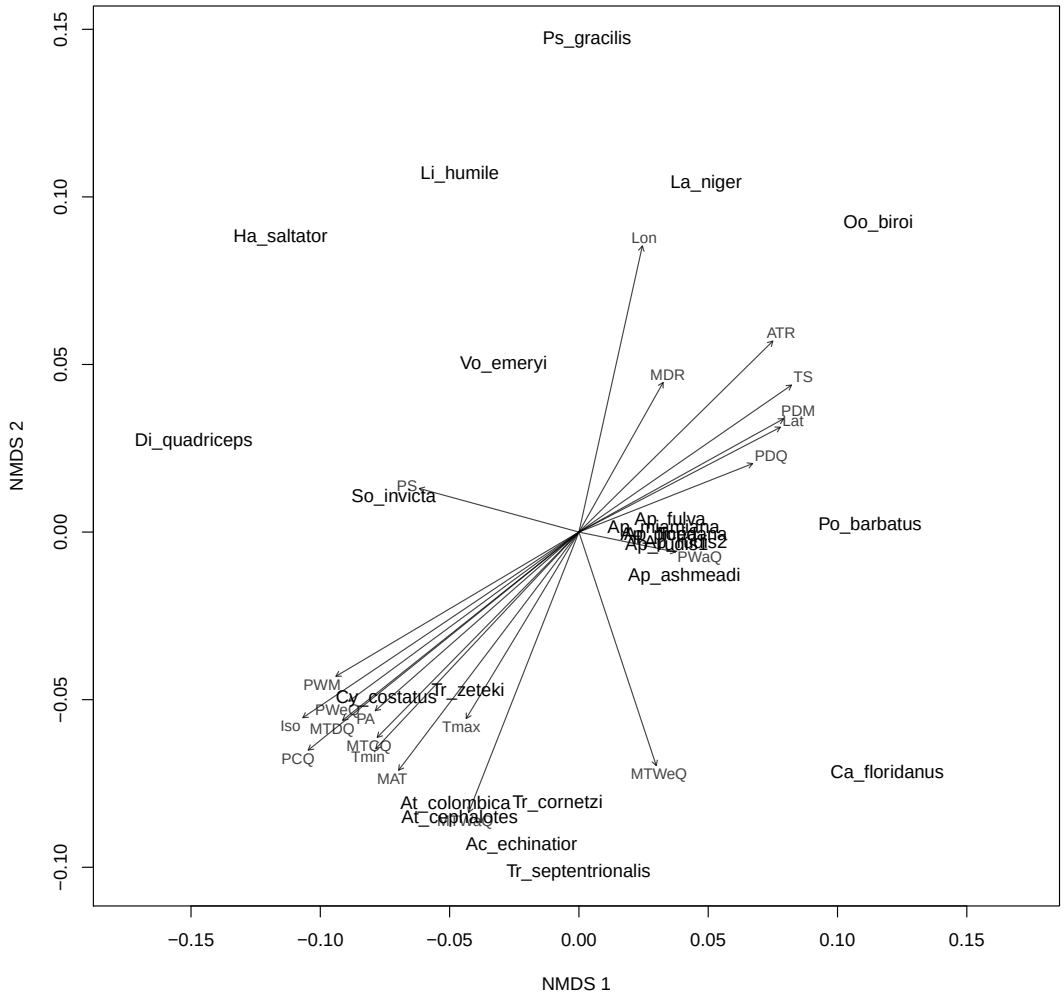

Figure 6. Plot showing an NMDS ordination of MASH genomic distance of all whole-genome ant sequences currently in NCBI and the newly sequenced *Aphaenogaster* spp. from this study. Arrows overlaid on each plot show the correlation vectors (pointing in the direction of and scaled by the correlation) between the full set of climate variables from WorldClim at the sampling locations and the genomic distance of the samples. For variable descriptions see Table 3

DISCUSSION

We have produced seven draft whole-genome sequences of six species of ants in the genus *Aphaenogaster*. These are the first whole-genomes from a previously un-sequenced genus, adding to the sequences of the diverse "formicoid" clade, which contains 90% of all extant ant species (Ward, 2014). Our genomic sequences were comparable in quality to other ant and insect genomes and the patterns of genomic similarity were in line with expectations based on current ant systematics. With the addition of the *Aphaenogaster* sequences, we observed that genome size and similarity display spatial patterns that relate to climate, and the results of our initial biogeographic analysis are consistent with the broad hypothesis that climate has been a force shaping the similarity of ant species.

Although correlative, the climate-genome analysis results are consistent with the hypothesis that ants from regions with more similar climates tend to have similar genomes with cold temperature related variables having the strongest relationships. Previous studies have observed physiological and ecological responses of ants to climate gradients and shifting temperatures (Warren and Chick, 2013; Stanton-Geddes et al., 2016; Diamond et al., 2016; Nguyen et al., 2017; Helms Cahan et al., 2017; Diamond et al., 2017; Penick et al., 2017) that could act as agents of selection or as environmental filters. For example, Warren and Chick (2013) found that cold, but not warm, temperatures limited shifts in the distributions of *A. picea* and *A. rudis*. Diamond et al. (2016) reported that the rate of colonization and occupancy of nests by *Aphaenogaster* species in a five-year experimental warming study (Pelini et al., 2014) declined with temperature in the warm, southern study site (Duke Forest, NC, USA) but not in the cooler, northern study site (Harvard Forest, MA, USA). In addition to the direct impacts of climate, some studies support the importance for the indirect effects of climate via biotic interactions. For example, the distribution of the species *Atta texana* is limited by the cold-tolerance of its fungal symbiont, cultivars of the genus *Attamyces* (Mueller et al., 2011). The evolution of the ant-fungus relationship has lead to reductions in some ant species ranges by cold temperatures. We observed patterns corroborating this in our analysis in the correlation between temperature variables and the clustering of similar genomes of ant species from the tribe Attini (Fig. 6).

It is important to keep in mind that the climate related genomic patterns observed in this study should be considered an initial view of possible biogeographic patterns in ant genomes. As the addition of the these sequences had a marked impact on the statistical results of the climate analysis (see Supplementary Materials Table 2), we expect that further sequencing work will continue to enhance our understanding of the ecological genomics of ants. Also, these findings should be tested with additional sequencing efforts, as we could not control for several potentially important intercorrelated variables, with the most prominent factor being phylogenetics. The biogeographic patterns we observed are likely to be a function

of both phylogeny and responses to ecological variation, as previous research has observed significant climatic variation in insect genomes even after controlling for phylogenetic relatedness (Alfsnes et al., 2017). In addition to our still relatively low power from the small number of ant genomes, we were also unable to control for phylogeny because the most complete phylogeny for ants (Moreau et al., 2006) is primarily resolved to the genus level (Smith et al., 2015) and less than 50% of the current ant genera with whole genome sequences were present in that dataset. Additionally, factors such as sampling bias and sequencing methodology (e.g. 454 versus Illumina) also varied among sequencing efforts, which could have contributed to some of the observed correlations with climate, but we did not attempt to control for these factors statistically. Further analytical and experimental work will be necessary to parse out a clearer understanding of the mechanisms behind these genomic patterns, and future work should methodologically and/or statistically control for such sources of variation in ant genomes as more sequences become available.

CONCLUSION

The addition of the *Aphaenogaster* sequences increases the breadth of global ant genomic sampling. The total number of ant sequences analyzed here, however, is still a relatively small sample (n = 26) of the estimated >16,000 ant species and subspecies (www.antweb.org, accessed 16 April 2018). Efforts such as The Global Ant Genomics Alliance (GAGA)(Boomsma et al., 2017), which aims to greatly increase the number of ant species sequenced from across the world, will provide additional resources for ecological genomics studies. Further work investigating the variation in genomic content and mapping of target coding regions from previous physiological (Nguyen et al., 2017), biochemical (Helms Cahan et al., 2017), and transcriptomic (Stanton-Geddes et al., 2016) studies of *Aphaenogaster* and other ant species will inform predictions of how these species, and the ecosystems that they inhabit, may respond to ongoing climatic change. For instance, determining the genomic factors underlying the temperature response of ant assemblages to climatic gradients (Warren and Chick, 2013; Diamond et al., 2016, 2017) could provide useful insights into the response of these important organisms to non-analog ecosystem states and idiosyncratic community responses (Bewick et al., 2014). In addition, as species distribution models have been significantly improved by the inclusion of genetic information (Ikeda et al., 2016), an ecological genetics approach that couples ant genomic and ecologically relevant data will provide a useful window into the response of many terrestrial ecosystems to a changing climate.

ACKNOWLEDGMENTS

Thank you to the team at the Broad Institute: particularly, James Bochicchio, Sarah Young, Terrance Shay and Caroline Cusick. This work was supported by a US National Science Foundation Dimensions of Biodiversity grant (DEB 11-36646) to NJS, RRD, AME, NJG and SHC.

REFERENCES

Agosti, D., Majer, J. D., Alonso, L. E., and Schultz, T. R. (2000). *Standard methods for measuring and monitoring biodiversity*, volume 233. Smithsonian Institution Press.

Alfsnes, K., Leinaas, H. P., and Hessen, D. O. (2017). Genome size in arthropods; different roles of phylogeny, habitat and life history in insects and crustaceans. *Ecol. Evol.*, 7(15):5939–5947.

Anderson, M. J. (2001). A new method for non-parametric multivariate analysis of variance. *Austral Ecol.*, 26(1):32–46.

Bewick, S., Stuble, K. L., Lessard, J.-P., Dunn, R. R., Adler, F. R., and Sanders, N. J. (2014). Predicting future coexistence in a North American ant community. *Ecol. Evol.*, 4(10):1804–1819.

Bolton, B. (2006). *Bolton's catalogue of ants of the world, 1758-2005*. Harvard University Press.

Boomsma, J. J., Brady, S. G., Dunn, R. R., Gadau, J., Heinze, J., Keller, L., Moreau, C. S., Sanders, N. J., Schrader, L., Schultz, T. R., Sundström, L., Ward, P. S., Wcislo, W. T., and Zhang, G. (2017). The Global Ant Genomics Alliance (GAGA). *Myrmecological News*, 25:61–66.

Brown, P. T. and Caldeira, K. (2017). Greater future global warming inferred from Earth's recent energy budget. *Nature*, 552(7683):45–50.

Burrows, M. T., Schoeman, D. S., Richardson, A. J., Molinos, J. G., Hoffmann, A., Buckley, L. B., Moore, P. J., Brown, C. J., Bruno, J. F., Duarte, C. M., Halpern, B. S., Hoegh-Guldberg, O., Kappel, C. V., Kiessling, W., O'Connor, M. I., Pandolfi, J. M., Parmesan, C., Sydeman, W. J., Ferrier, S., Williams, K. J., and Poloczanska, E. S. (2014). Geographical limits to species-range shifts are suggested by climate velocity. *Nature*, 507(7493):492–495.

Del Toro, I., Ribbons, R. R., and Pelini, S. L. (2012). The little things that run the world revisited: A review of ant-mediated ecosystem services and disservices (Hymenoptera: Formicidae).

DeMarco, B. B. and Cognato, A. I. (2016). A multiple-gene phylogeny reveals polyphyly among eastern North American *Aphaenogaster* species (Hymenoptera: Formicidae). *Zool. Scr.*, 45(5):512–520.

Des Roches, S., Post, D. M., Turley, N. E., Bailey, J. K., Hendry, A. P., Kinnison, M. T., Schweitzer, J. A., and Palkovacs, E. P. (2017). The ecological importance of intraspecific variation. *Nat. Ecol. Evol.*

Diamond, S. E., Chick, L., Penick, C. A., Nichols, L. M., Cahan, S. H., Dunn, R. R., Ellison, A. M., Sanders, N. J., and Gotelli, N. J. (2017). Heat tolerance predicts the importance of species interaction effects as the climate changes. *Integr. Comp. Biol.*, 57(1):112–120.

Diamond, S. E. and Chick, L. D. (2018). Thermal specialist ant species have restricted, equatorial geographic ranges: Implications for climate change vulnerability and risk of extinction. *Ecography (Cop.).*

Diamond, S. E., Nichols, L. M., Pelini, S. L., Penick, C. A., Barber, G. W., Cahan, S. H., Dunn, R. R.,

Ellison, A. M., Sanders, N. J., and Gotelli, N. J. (2016). Climatic warming destabilizes forest ant communities. *Sci. Adv.*, 2(10):e1600842–e1600842.

Diamond, S. E., Sorger, D. M., Hulcr, J., Pelini, S. L., Toro, I. D., Hirsch, C., Oberg, E., and Dunn, R. R. (2012). Who likes it hot? A global analysis of the climatic, ecological, and evolutionary determinants of warming tolerance in ants. *Glob. Chang. Biol.*, 18(2):448–456.

Efron, B. and Tibshirani, R. (2007). On testing the significance of sets of genes. *Ann. Appl. Stat.*, 1(1):107–129.

Ellison, A. M. (2012). *A field guide to the ants of New England*. Yale University Press.

Fick, S. E. and Hijmans, R. J. (2017). WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *Int. J. Climatol.*, 37(12):4302–4315.

Gnerre, S., Maccallum, I., Przybylski, D., Ribeiro, F. J., Burton, J. N., Walker, B. J., Sharpe, T., Hall, G., Shea, T. P., Sykes, S., Berlin, A. M., Aird, D., Costello, M., Daza, R., Williams, L., Nicol, R., Gnirke, A., Nusbaum, C., Lander, E. S., and Jaffe, D. B. (2011). High-quality draft assemblies of mammalian genomes from massively parallel sequence data. *Proc. Natl. Acad. Sci. U. S. A.*, 108(4):1513–8.

Goslee, S. C. and Urban, D. L. (2007). The ecodist Package for Dissimilarity-based Analysis of Ecological Data. *J. Stat. Softw.*, 22(7):1–19.

Hare, E. E. and Johnston, J. S. (2011). Genome Size Determination Using Flow Cytometry of Propidium Iodide-Stained Nuclei. *Mol. Methods Evol. Genet.*, 772(5):3–12.

Helms Cahan, S., Nguyen, A. D., Stanton-Geddes, J., Penick, C. A., Hernáiz-Hernández, Y., DeMarco, B. B., and Gotelli, N. J. (2017). Modulation of the heat shock response is associated with acclimation to novel temperatures but not adaptation to climatic variation in the ants Aphaenogaster picea and A. rudis. *Comp. Biochem. Physiol. Part A Mol. Integr. Physiol.*, 204:113–120.

Hultgren, K. M., Jeffery, N. W., Moran, A., and Gregory, T. R. (2018). Latitudinal variation in genome size in crustaceans. *Biol. J. Linn. Soc.*, 123(2):348–359.

Ikeda, D. H., Max, T. L., Allan, G. J., Lau, M. K., Shuster, S. M., and Whitham, T. G. (2016). Genetically informed ecological niche models improve climate change predictions. *Glob. Chang. Biol.*, 23(1):164–176.

Kaspari, M., Clay, N. A., Lucas, J. A., Revzen, S., Kay, A. D., and Yanoviak, S. P. (2015). Thermal adaptation and phosphorus shape thermal performance in an assemblage of rainforest ants. *Ecology*, 97(4):15–1225.1.

Moreau, C. S. (2006). Phylogeny of the Ants: Diversification in the Age of Angiosperms. *Eur. J. Biochem. Eur. J. Biochem. J. Steroid Biochem. Mol. Cell Nat. Sci. N. Gompel, B. Prud'hom. Nat. J. Piatigorsky, Ann. N.Y. Acad. Sci. Sci.*, 101(281):1249–481.

335 Moreau, C. S., Bell, C. D., Vila, R., Archibald, S. B., and Pierce, N. E. (2006). Phylogeny of the Ants:
336 Diversification in the Age of Angiosperms. *Science (80-.).*, 312(5770).
337 Mousseau, T. A. (1997). Ectotherms Follow the Converse to Bergmann's Rule. *Evolution (N. Y).*,
338 51(2):630.
339 Mueller, U. G., Mikheyev, A. S., Hong, E., Sen, R., Warren, D. L., Solomon, S. E., Ishak, H. D., Cooper,
340 M., Miller, J. L., Shaffer, K. A., and Juenger, T. E. (2011). Evolution of cold-tolerant fungal symbionts
341 permits winter fungiculture by leafcutter ants at the northern frontier of a tropical ant-fungus symbiosis.
342 *Proc. Natl. Acad. Sci.*, 108(10):4053–4056.
343 Neafsey, D. E., Lawniczak, M. K. N., Park, D. J., Redmond, S. N., Coulibaly, M. B., Traoré, S. F., Sagnon,
344 N., Costantini, C., Johnson, C., Wiegand, R. C., Collins, F. H., Lander, E. S., Wirth, D. F., Kafatos,
345 F. C., Besansky, N. J., Christophides, G. K., and Muskavitch, M. A. T. (2010). SNP genotyping defines
346 complex gene-flow boundaries among African malaria vector mosquitoes. *Science*, 330(6003):514–517.
347 Nguyen, A. D., DeNovellis, K., Resendez, S., Pustilnik, J. D., Gotelli, N. J., Parker, J. D., and Cahan,
348 S. H. (2017). Effects of desiccation and starvation on thermal tolerance and the heat-shock response in
349 forest ants. *J. Comp. Physiol. B*, 187(8):1107–1116.
350 Nygaard, S. and Wurm, Y. (2015). Ant genomics (Hymenoptera: Formicidae): Challenges to overcome
351 and opportunities to seize. *Myrmecological News*, 21:59–72.
352 Oksanen, J., Blanchet, F., Kindt, R., Legendre, P., and O'Hara, R. (2016). Vegan: community ecology
353 package.
354 Ondov, B. D., Treangen, T. J., Melsted, P., Mallonee, A. B., Bergman, N. H., Koren, S., and Phillippy,
355 A. M. (2016). Mash: fast genome and metagenome distance estimation using MinHash. *Genome Biol.*,
356 17(1):132.
357 Parmesan, C. (2006). Ecological and Evolutionary Responses to Recent Climate Change. *Annu. Rev.*
358 *Ecol. Evol. Syst.*, 37(1):637–669.
359 Pelini, S. L., Diamond, S. E., Nichols, L. M., Stuble, K. L., Ellison, A. M., Sanders, N. J., Dunn, R. R.,
360 and Gotelli, N. J. (2014). Geographic differences in effects of experimental warming on ant species
361 diversity and community composition. *Ecosphere*, 5(10):art125.
362 Penick, C. A., Diamond, S. E., Sanders, N. J., and Dunn, R. R. (2017). Beyond thermal limits: com-
363 prehensive metrics of performance identify key axes of thermal adaptation in ants. *Funct. Ecol.*,
364 31(5):1091–1100.
365 Petrov, D. A. (2001). Evolution of genome size: new approaches to an old problem. *Trends Genet.*,
366 17(1):23–28.
367 R Core Team (2017). R Core Team (2017). R: A language and environment for statistical computing. *R*

Found. Stat. Comput. Vienna, Austria. URL http//www.R-project.org/., page R Foundation for Statistical Computing.

Rowntree, J. K., Shuker, D. M., and Preziosi, R. F. (2011). Forward from the crossroads of ecology and evolution. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, 366(1569):1322–8.

Siddig, A. A., Ellison, A. M., Ochs, A., Villar-Leeman, C., and Lau, M. K. (2016). How do ecologists select and use indicator species to monitor ecological change? Insights from 14 years of publication in Ecological Indicators. *Ecol. Indic.*, 60:223–230.

Smith, C. R., Helms Cahan, S., Kemena, C., Brady, S. G., Yang, W., Bornberg-Bauer, E., Eriksson, T., Gadau, J., Helmkampf, M., Gotzek, D., Okamoto Miyakawa, M., Suarez, A. V., and 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(11):2919–31.

Spicer, M. E., Stark, A. Y., Adams, B. J., Kneale, R., Kaspari, M., and Yanoviak, S. P. (2017). Thermal constraints on foraging of tropical canopy ants. *Oecologia*, 183(4):1007–1017.

Stanton-Geddes, J., Nguyen, A., Chick, L., Vincent, J., Vangala, M., Dunn, R. R., Ellison, A. M., Sanders, N. J., Gotelli, N. J., and Helms Cahan, S. (2016). Thermal reactionomes reveal divergent responses to thermal extremes in warm and cool-climate ant species. *BMC Genomics*, 17(1):171–186.

Thompson, J. N. (1999). Specific Hypotheses on the Geographic Mosaic of Coevolution. *Am. Nat.*, 153(S5):S1–S14.

Tsutsui, N. D., Suarez, A. V., Spagna, J. C., Johnston, J. S., Gregory, T., Evans, J., Gundersen-Rindal, D., Gardner, T., Gregory, T., Wilson, E., Hölldobler, B., Wilson, E., Li, J., Heinz, K., Johnston, J., Ross, L., Beani, L., Hughes, D., Kathirithamby, J., Geraci, N., Johnston, J., Robinson, J., Wikel, S., Hill, C., Gregory, T., Bennett, M., Leitch, I., SanMiguel, P., Gaut, B., Tikhonov, A., Nakajima, Y., Bennetzen, J., Kazazian, H., Kidwell, M., Comeron, J., Ustinova, J., Achmann, R., Cremer, S., Mayer, F., Hancock, J., Hancock, J., Toth, G., Gaspari, Z., Jurka, J., Redon, R., Ishikawa, S., Fitch, K., Feuk, L., Perry, G., Andrews, T., Fiegler, H., Shapero, M., Carson, A., Chen, W., Cho, E., Dallaire, S., Freeman, J., Gonzalez, J., Gratacos, M., Huang, J., Kalaitzopoulos, D., Komura, D., MacDonald, J., Marshall, C., Mei, R., Montgomery, L., Nishimura, K., Okamura, K., Shen, F., Somerville, M., Tchinda, J., Valsesia, A., Woodwark, C., Yang, F., Zhang, J., Zerjal, T., Zhang, J., Armengol, L., Conrad, D., Estivill, X., Tyler-Smith, C., Carter, N., Aburatani, H., Lee, C., Jones, K., Scherer, S., Hurles, M., Gregory, T., Gregory, T., Petrov, D., Lozovskaya, E., Hartl, D., Devos, K., Brown, J., Bennetzen, J., Bennetzen, J., Ma, J., Devos, K., Ma, J., Bennetzen, J., Oliver, M., Petrov, D., Ackerly, D., Falkowski, P., Schofield, O., Gregory, T., Hebert, P., Kolasa, J., Finston, T., Hebert, P., Foottit, R., Ferrari, J., Rai, K., Ellegren, H., Vandenbussche, R., Longmire, J., Baker, R., Organ, C., Shedlock, A., Meade, A., Pagel, M., Edwards,

S., Hughes, A., Hughes, M., Reinhold, K., Gregory, T., Pittendrigh, B., Clark, J., Johnston, J., Lee, S., Romero-Severson, J., Dasch, G., Gregory, T., Weinstock, G., Robinson, G., Gibbs, R., Worley, K., Evans, J., Maleszka, R., Robertson, H., Weaver, D., Beye, M., Bork, P., Elsik, C., Hartfelder, K., Hunt, G., Zdobnov, E., Amdam, G., Bitondi, M., Collins, A., Cristino, A., Lattorff, H., Lobo, C., Moritz, R., Nunes, F., Page, R., Simoes, Z., Wheeler, D., Carninci, P., Fukuda, S., Hayashizaki, Y., Kai, C., Kawai, J., Sakazume, N., Sasaki, D., Tagami, M., Albert, S., Baggerman, G., Beggs, K., Bloch, G., Cazzamali, G., Cohen, M., Drapeau, M., Eisenhardt, D., Emore, C., Ewing, M., Fahrbach, S., Foret, S., Grimmelikhuijzen, C., Hauser, F., Hummon, A., Huybrechts, J., Jones, A., Kadowaki, T., Kaplan, N., Kucharski, R., Leboulle, G., Linial, M., Littleton, J., Mercer, A., Richmond, T., Rodriguez-Zas, S., Rubin, E., Sattelle, D., Schlipalius, D., Schoofs, L., Shemesh, Y., Sweedler, J., Velarde, R., Verleyen, P., Vierstraete, E., Williamson, M., Ament, S., Brown, S., Corona, M., Dearden, P., Dunn, W., Elekonich, M., Fujiyuki, T., Gattermeier, I., Gempe, T., Hasselmann, M., Kadowaki, T., Kage, E., Kamikouchi, A., Kubo, T., Kucharski, R., Kunieda, T., Lorenzen, M., Milshina, N., Morioka, M., Ohashi, K., Overbeek, R., Ross, C., Schioett, M., Shippy, T., Takeuchi, H., Toth, A., Willis, J., Wilson, M., Gordon, K., Letunic, I., Hackett, K., Peterson, J., Felsenfeld, A., Guyer, M., Solignac, M., Agarwala, R., Cornuet, J., Monnerot, M., Mougel, F., Reese, J., Vautrin, D., Gillespie, J., Cannone, J., Gutell, R., Johnston, J., Eisen, M., Iyer, V., Iyer, V., Kosarev, P., Mackey, A., Solovyev, V., Souvorov, A., Aronstein, K., Bilikova, K., Chen, Y., Clark, A., Decanini, L., Gelbart, W., Hetru, C., Hultmark, D., Imler, J., Jiang, H., Kanost, M., Kimura, K., Lazzaro, B., Lopez, D., Simuth, J., Thompson, G., Zou, Z., Jong, P. D., Sodergren, E., Csuros, M., Milosavljevic, A., Osoegawa, K., Richards, S., Shu, C., Duret, L., Elhaik, E., Graur, D., Anzola, J., Campbell, K., Childs, K., Collinge, D., Crosby, M., Dickens, C., Grametes, L., Grozinger, C., Jones, P., Jorda, M., Ling, X., Matthews, B., Miller, J., Mizzen, C., Peinado, M., Reid, J., Russo, S., Schroeder, A., Pierre, S. S., Wang, Y., Zhou, P., Jiang, H., Kitts, P., Ruef, B., Venkatraman, A., Zhang, L., Aquino-Perez, G., Whitfield, C., Behura, S., Berlocher, S., Sheppard, W., Smith, D., Suarez, A., Tsutsui, N., Wei, X., Wheeler, D., Havlak, P., Li, B., Liu, Y., Sodergren, E., Jolivet, A., Lee, S., Nazareth, L., Pu, L., Thorn, R., Stolc, V., Newman, T., Samanta, M., Tongprasit, W., Claudianos, C., Berenbaum, M., Biswas, S., de Graaf, D., Feyereisen, R., Johnson, R., Oakeshott, J., Ranson, H., Schuler, M., Muzny, D., Chacko, J., Davis, C., Dinh, H., Gill, R., Hernandez, J., Hines, S., Hume, J., Jackson, L., Kovar, C., Lewis, L., Miner, G., Morgan, M., Nguyen, N., Okwuonu, G., Paul, H., Santibanez, J., Savery, G., Svatek, A., Villasana, D., Wright, R., Consort, H., Moreau, C., Bell, C., Vila, R., Archibald, S., Pierce, N., Brady, S., Schultz, T., Fisher, B., Ward, P., Mueller, U., Gerardo, N., Aanen, D., Six, D., Schultz, T., Chapela, I., Rehner, S., Schultz, T., Mueller, U., Wetterer, J., Schultz, T., Meier, R., Gregory, T., Hebert, P., Gregory, T., Shorthouse, D., Wang, J., Jemielity, S., Uva, P.,

Wurm, Y., Graff, J., Keller, L., Bennett, M., Leitch, I., Price, H., Johnston, J., Abouheif, E., Reeve, J., Abouheif, E., Felsenstein, J., Purvis, A., and Rambaut, A. (2008). The evolution of genome size in ants. *BMC Evol. Biol.*, 8(1):64.

Walker, B. J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumar, S., Cuomo, C. A., Zeng, Q., Wortman, J., Young, S. K., and Earl, A. M. (2014). Pilon: An Integrated Tool for Comprehensive Microbial Variant Detection and Genome Assembly Improvement. *PLoS One*, 9(11):e112963.

Ward, P. S. (2014). The Phylogeny and Evolution of Ants. *Annu. Rev. Ecol. Evol. Syst.*, 45(1):23–43.

Ward, P. S., Brady, S. G., Fisher, B. L., and Schultz, T. R. (2015). The evolution of myrmicine ants: phylogeny and biogeography of a hyperdiverse ant clade (Hymenoptera: Formicidae). *Syst. Entomol.*, 40(1):61–81.

Warren, R. J. and Chick, L. (2013). Upward ant distribution shift corresponds with minimum, not maximum, temperature tolerance. *Glob. Chang. Biol.*, 19(7):2082–2088.

SUPPLEMENTARY MATERIALS

	BioProject Accession	BioSample Accession	Lat	Lon
Acromyrmex echinatior	PRJNA62733	SAMN02953789	-79.696513	9.1164638
Atta cephalotes	PRJNA48091	SAMN02953774	-79.696513	9.1164638
Atta colombica	PRJNA343260	SAMN03982875	-79.696513	9.1164638
Camponotus floridanus	PRJNA50201	SAMN02953777	-81.5431872	24.6245746
Cyphomyrmex costatus	PRJNA343963	SAMN03982885	-79.696513	9.1164638
Dinoponera quadriceps	PRJNA301625	SAMN02869781	-79.8697222	9.4008333
Harpegnathos saltator	PRJNA50203	SAMN00016742	75.7138884	15.3172775
Lasius niger	PRJNA269328	SAMN03253098	37.6172999	55.755826
Linepithema humile	PRJNA45799	SAMN02767796	-122.0230146	37.2638324
Monomorium pharaonis	PRJDB3164	SAMD00020277	NA	NA
Ooceraea biroi	PRJNA275884	SAMN02428046	127.6809317	26.2124013
Pogonomyrmex barbatus	PRJNA45797	SAMN02953770	-100.3898876	20.5888184
Pseudomyrmex gracilis	PRJNA377720	SAMN03219222	-70.8119953	-11.7668705
Solenopsis invicta	PRJNA49629	SAMN02953778	-83.357567	33.9519347
Trachymyrmex cornetzi	PRJNA343972	SAMN03982882	-79.696513	9.1164638
Trachymyrmex septentrionalis	PRJNA343973	SAMN03982881	-84.2807329	30.4382559
Trachymyrmex zeteki	PRJNA343251	SAMN03982884	-79.696513	9.1164638
Vollenhovia emeryi	PRJDB3517	SAMD00026325	-100.3898876	20.5888184
Wasmannia auropunctata	PRJDB3443	SAMD00024919	NA	NA

Table 1. NCBI genome database accession information for the previously sequenced ant genomes and coordinates for species those species that could be obtained from the published literature.

	df	*SS*	*MS*	*Pseudo-F*	*R2*	p-*value*
Size Distance						
Lat	1	2707.43	2707.43	2.11	0.07	0.1796
Lon	1	1759.79	1759.79	1.37	0.05	0.2693
MAT	1	118.64	118.64	0.09	0.00	0.7636
Tmin	1	3394.10	3394.10	2.65	0.09	0.1434
Tmax	1	5518.63	5518.63	4.31	0.14	0.0727
PA	1	8349.14	8349.14	6.52	0.21	0.0363
PS	1	5501.51	5501.51	4.29	0.14	0.0679
Residuals	9	11533.39	1281.49		0.30	
Total	16	38882.63			1.00	
MASH Distance						
Lat	1	0.02	0.02	1.66	0.08	0.0683
Lon	1	0.02	0.02	2.07	0.10	0.0295
MAT	1	0.02	0.02	1.95	0.10	0.0332
Tmin	1	0.01	0.01	1.06	0.05	0.3679
Tmax	1	0.01	0.01	1.43	0.07	0.1483
PA	1	0.01	0.01	1.38	0.07	0.1590
PS	1	0.02	0.02	1.56	0.08	0.0871
Residuals	9	0.09	0.01		0.45	
Total	16	0.19			1.00	

Table 2. PerMANOVA pseudo-F table for the analysis of the factors correlated with ant genome size and similarity (MASH distance) only including the previously sequenced NCBI ant specimens.