

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

Matthew K. Lau¹, Aaron M. Ellison¹, Andrew Nguyen^{2,3}, Clint Penick^{4,5}, Bernice DeMarco⁶, Nicholas J. Gotelli², Nathan J. Sanders⁷, Robert R. Dunn⁴, and Sara Helms Cahan²

¹Harvard Forest, Harvard University, Petersham, MA, USA

²Department of Biology, University of Vermont, Burlington, VT, USA

³Department of Entomology and Nematology, University of Florida, Gainesville, FL, USA

⁴Department of Applied Ecology, North Carolina State University, Raleigh, NC, USA

⁵The Biomimicry Center, Arizona State University, Tempe, AZ, USA

⁶Smithsonian Institution, Washington, DC, USA

⁷Environmental 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 found that after controlling for both spatial autocorrelation and phylogenetics ant genome size was marginally correlated with sample site climate similarity. Of all examined climate variables, minimum temperature showed the strongest correlation with genome size. These results suggest that temperature extremes could be a selective force acting on ant genome size and point to the need for more extensive sequencing of ant genomes.

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 is also 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). Studies of ant genomes have shed light on the evolution and social organization of ants (Libbrecht et al., 2013). One promising avenue is the possibility of genome size as an adaptive trait in ants. Recent observational studies have reported biogeographic patterns in genome size in arthropod taxa, e.g. Crustacea (Hultgren et al., 2018), and patterns in insect genomes suggest that climate may constrain genome size with cold temperatures possibly selecting for larger genome sizes (Mousseau, 1997; Petrov, 2001; Alfsnes et al., 2017). Specific to ants, previous research into genome size variation using flow cytometry found that ants have small genomes relative to other insect taxa and that their genomes display large variation across subfamilies with patterns indicative of both gradual and rapid evolution in genome size (Tsutsui et al., 2008).

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*, which are abundant ants that play key roles in the dispersal of understory plant species in North America and temperate Asia, have yet been sequenced. Previous studies have also shown that *Aphaenogaster* species' ecological and physiological responses to climatic change appear to depend both on species identity and on the geographic region in which climatic change occurs (Warren and Chick, 2013; Stanton-Geddes et al., 2016).

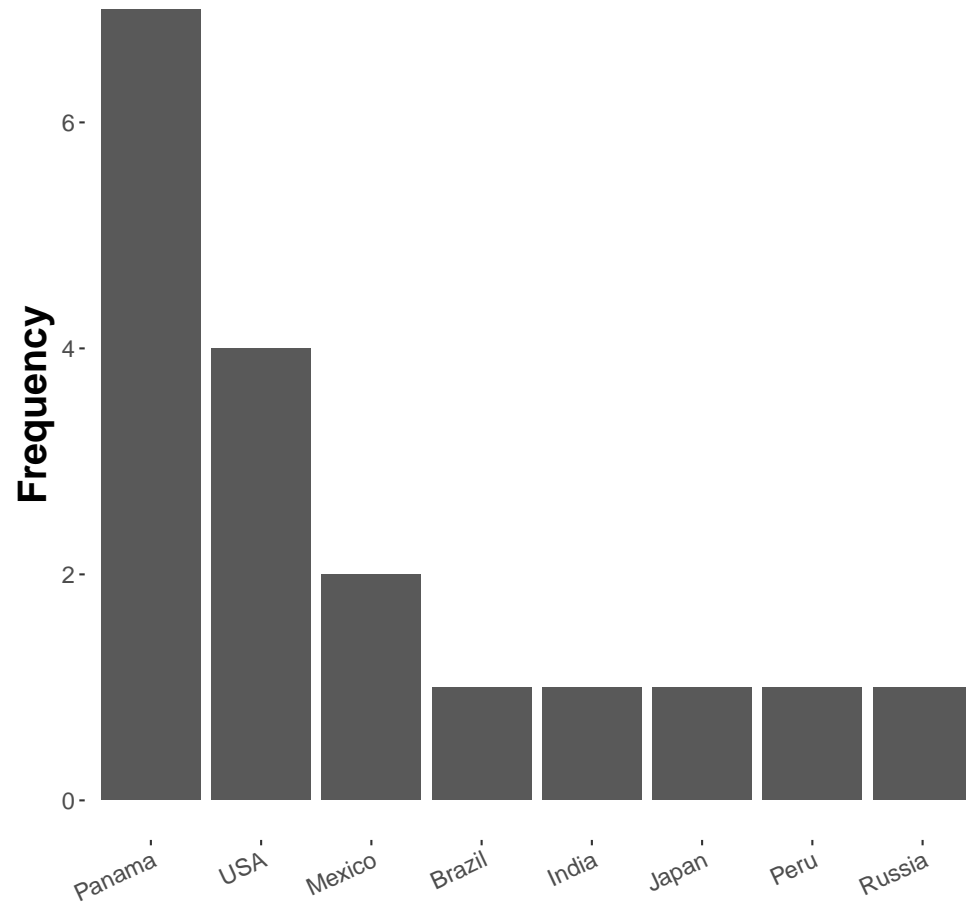


Figure 1. Number of ant (*Formicidae*) whole-genome sequences available in NCBI by country (accessed August 2018).

75 To increase the number of genomes of temperate-zone ant species, we sequenced the genomes
 76 of *Aphaenogaster* species. We conducted whole genome sequencing for six species: *A. ashmeadi*, *A.*
 77 *floridana*, *A. fulva*, *A. miamiana*, *A. picea* and *A. rudis*. These species were collected from across a broad
 78 biogeographic gradient spanning 10 degrees of longitude and 12 degrees of latitude. We also conducted
 79 an initial exploration of biogeographic patterns in ant genome sequences, focusing on genome size. To
 80 do this we analyzed the newly collected *Aphaenogaster* sequences together with all publicly available
 81 ant whole genome sequences. We present the newly sequenced *Aphaenogaster* genomes and investigate
 82 biogeographic (i.e. location and climate) related patterns of ant genomes using all currently sequenced
 83 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

103 final assembled genomes.

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

104 Genome Quality and Composition

105 DNA extractions yielded substantial amounts of high quality DNA with concentrations and quality scores
 106 ranging from 3.45–5.39 ng μ L⁻¹ and 4.05–4.27 ng μ L⁻¹, respectively. All genome assemblies displayed
 107 good coverage, with an average of 70% of fragments mapped (Table 2). Across all species, the length of
 108 the shortest contig at 50% of the genome (i.e. N50) was 18,864 bases; average assembly GC content was
 109 38.18%; and average genome size was 370.45 Mb. Using GAEMR’s BLAST feature to conduct a search
 110 of the contigs against the NCBI’s nucleotide sequence database, we discovered that 38.98% and 22.04%
 111 of the top hits were “ant” and *Aphaenogaster*, respectively. The *Aphaenogaster* genomes compared well
 112 with other ant genome sequences. The sizes of the *Aphaenogaster* genomes were within the range of other
 113 ant genomes based on size from both flow cytometry (Tsutsui et al., 2008) and the previously sequenced
 114 ant genomes available in NCBI (Fig 3).

	<i>A. ashmeadi</i>	<i>A. floridana</i>	<i>A. fulva</i>	<i>A. miamiana</i>	<i>A. picea</i>	<i>A. rudis1</i>	<i>A. rudis2</i>
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*.

115 After detecting and masking repeat regions in the *Aphaenogaster* genomes using *Repeatmasker*
 116 (version 4.0.5 Institute for Systems Biology), we applied MASH distance (Ondov et al., 2016) to measure
 117 pairwise dissimilarity of genomic sequences. The MASH method extends a data compression and
 118 dimensionality-reduction algorithm to generate estimates of sequence similarity with low computational
 119 overhead. Briefly, the pairs of genomic sequences were pre-processed into sets of k-mers of size 21 with

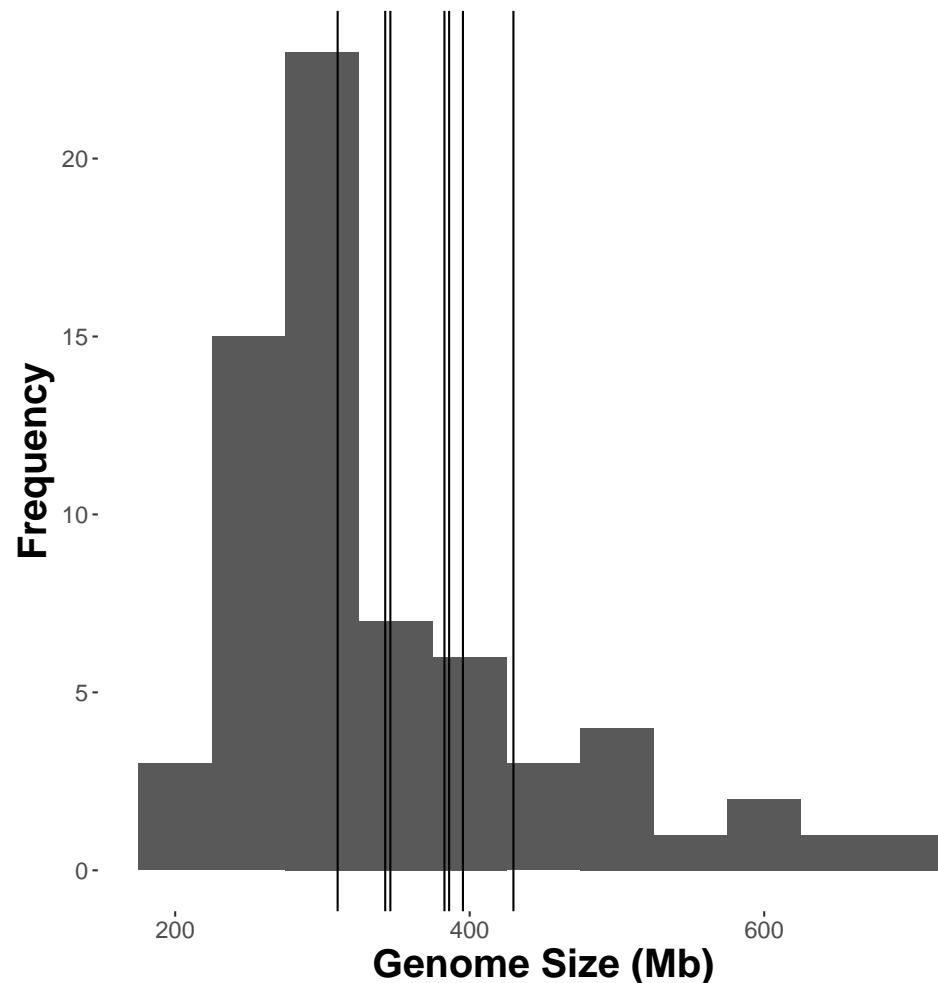


Figure 3. The size of sequenced *Aphaenogaster* genomes were within the size range of previously published observed or estimated genomes of ants. Frequency distribution of previously published genome size estimates using flow cytometry from Tsutsui et al. (2008) and those available via NCBI (accessed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

120 the size of the non-redundant hashes retained set to 1,000. These settings have been demonstrated to
 121 provide good representation of genomic similarity with minimal computational costs (Ondov et al., 2016).
 122 These sets were then used to estimate the Jaccard similarity coefficient (the ratio of shared k-mers to
 123 total k-mers) of subsampled k-mer pairs of genomes. This unbiased estimate of the Jaccard similarity (J)
 124 was then used to calculate the dissimilarity of the two genomes (D) as $D = 1 - J$. All Jaccard similarity
 125 estimates had p -values less than 10^{-14} , which is below the recommended 10^{-3} probability of observing
 126 values of J due to chance.

127 Using the MASH genomic distances, we observed patterns of genomic similarity that were in line
 128 with expectations of ant relatedness. Sequences formed groups that corresponded with subfamilies
 129 (Fig 4). *Aphaenogaster* clustered with other genera from the Myrmicinae and, in general, subfamily
 130 level clustering tended to follow previously observed patterns of subfamily relatedness (Bolton, 2006;

131 Moreau, 2006; Ward, 2014). The *Aphaenogaster* sequences formed a single cluster containing only
132 *Aphaenogaster* species and displayed intra-generic levels of genomic variance comparable to other genera
133 s(e.g., *Trachymyrmex* spp.). The separation of the two *A. rudis* species was initially surprising, as these two
134 samples were collected at the same site (Duke Forest, NC USA) and were identified as the same species
135 based on their morphological characteristics (Ellison, 2012; DeMarco and Cognato, 2016). However, two
136 recent studies of targeted gene regions have demonstrated the polyphyletic nature of *Aphaenogaster rudis*.
137 One study of the evolution of the subfamily Myrmicinae observed that the genus as a whole could be split
138 into at least four different lineages (Ward et al., 2015). Another, more detailed study of the genus in North
139 America found that multiple individuals of *A. rudis* separated out into distinct groupings, each with other
140 species, specifically, individuals of *A. rudis* from North Carolina (USA) were observed to form distinct
141 clusters with individuals of *A. carolinensis*, *A. miamiana*, *A. lamellidens* and *A. texana* (DeMarco and
142 Cognato, 2016).

143 **Biogeographic Patterns of Ant Genomes**

144 We used multivariate correlation analyses to examine biogeographic patterns of ant genomes. Mantel tests
145 of multivariate correlation of distance matrices were used to examine correlations among ant genomes
146 and climate variables. Specifically, we used directional (H_0 : Mantel $r \leq 0$) partial mantel tests, which
147 calculate the correlation between two distance matrices while controlling for the covariance with other
148 matrices (Goslee and Urban, 2007). First, we examined the correlations between genomic similarity
149 (MASH distance), whole-genome size similarity (Euclidean distance of assembly size in total base pairs)
150 and climate variables (also using Euclidean distance). Via partial Mantel tests, we were able to isolate
151 the correlation between genome size and climate by controlling for spatial autocorrelation and potential
152 phylogenetic patterns by including geodesic and MASH distances as terms.

153 Data for climatic variables for each sampling location was obtained from the WorldClim database
154 (version 2.0) at a 2.5 arc minute spatial resolution from the years 1970 to 2000 (Fick and Hijmans, 2017).
155 Although used in the previous analyses of ant genomes, two species, (*W. auropunctata* and *M. pharaonis*),
156 which did not have published location information, were excluded from biogeographic analyses (see
157 Supplementary Materials Table 1).

158 Here, it is important to note that we are using assembly size as an indicator of genome size. As
159 genome size estimates are generally used to set assembly size targets for whole genome sequencing efforts
160 (Hare and Johnston, 2011), we expect there to be a high degree of correlation between assembly size
161 and genome size. Also, as a test of the potential relationship between assembly size and true genome
162 size, we examined the correlation between the average assembly sizes of ant genera that overlapped with
163 flow cytometry estimates of those published in Tsutsui et al. (2008). We found a marginally significant

164 correlation assembly size and flow cytometry at the genus level (Spearman's $\rho = 0.57$, $p\text{-value} = 0.076$),
 165 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 similarity of ant genomes.

166 Using a permutational multivariate analysis of variance (PerMANOVA) procedure, we parsed the
 167 individual variables that were correlated with both genome size and MASH similarity. PerMANOVA is
 168 a flexible multivariate analog of ANOVA that permits the use of a wider set of similarity metrics to be
 169 used for the response matrix (Anderson, 2001), such as the MASH distance. We ran a total of 10,000
 170 permutations of the original distance matrices for each statistical permutation procedure. We chose a
 171 subset of all possible climate variables available via WorldClim for this analysis. A visual inspection
 172 of the sampled climate variable correlations indicated that the primary climate variables, mean annual
 173 temperature (MAT), annual minimum temperature, annual maximum temperature, annual precipitation
 174 and summer precipitation, represented the majority of climate variation (Fig 6). Based on this, we only
 175 included these variables, along with latitude and longitude coordinates, as factors in the PerMANOVAs.

176 After controlling for both spatial autocorrelation and potential phylogenetic patterns, we found a
 177 marginally significant, positive correlation between ant genome size similarity and climate similarity
 178 (Mantel $R = 0.14$, $p\text{-value} = 0.055$). Although genome size similarity and MASH genome similarity were
 179 not significantly correlated (Mantel $R = 0.08$, $p\text{-value} = 0.217$), we included MASH as a covariate in
 180 addition to geodesic distance because previous research indicated that genome size is associated with
 181 phylogenetic relatedness (Alfsnes et al., 2017). We found that different spatial and climatic variables were
 182 associated with the size similarity of ant genomes. Longitude but not latitude was a significant predictor

183 of genome size (Table 4). Temperature of the coldest (Tmin) and hottest (Tmax) month and total annual
 184 precipitation (PA), were significant predictors, but neither mean annual temperature (MAT) nor summer
 185 precipitation (PS) were significant predictors of genome size. Overall, Tmin had the highest correlation
 186 with an R^2 of 0.23. Examining the correlation between genome size and Tmin, we observed a negative
 187 correlation with genome size tending to increase as minimum temperature decreased (Fig 6). When the
 188 newly sequenced *Aphaenogaster* genomes were excluded from the analysis, only annual precipitation
 189 (PA) was a significant predictor of genome size similarity (see Supplementary Materials Table 2).

	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>Pseudo-F</i>	<i>R2</i>	<i>p-value</i>
<i>Assembly Size Similarity</i>						
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	

Table 4. PerMANOVA pseudo-F table for the analysis of the factors correlated with ant assembly size.

190 **Data, Computation, and Statistics**

191 The raw and assembled genome sequences are currently archived at Harvard Forest (Petersham, MA,
 192 USA) and in NCBI's genome database (Genome Accessions NJRK000000000-NJRQ000000000 and
 193 BioSample Accessions SAMN06892346-SAMN06892352). Genomic distance (MASH) computations
 194 were run on the Odyssey cluster supported by the FAS Division of Science, Research Computing Group
 195 at Harvard University. All analyses were conducted in **R** (R Core Team, 2017). Analytical scripts for the
 196 project have been versioned and archived (DOI: 10.5281/zenodo.1341982) and are available online at
 197 <https://zenodo.org/record/1341982>. We used the *vegan* (Oksanen et al., 2016) and *ecodist* (Goslee and
 198 Urban, 2007) packages in R for multivariate analyses.

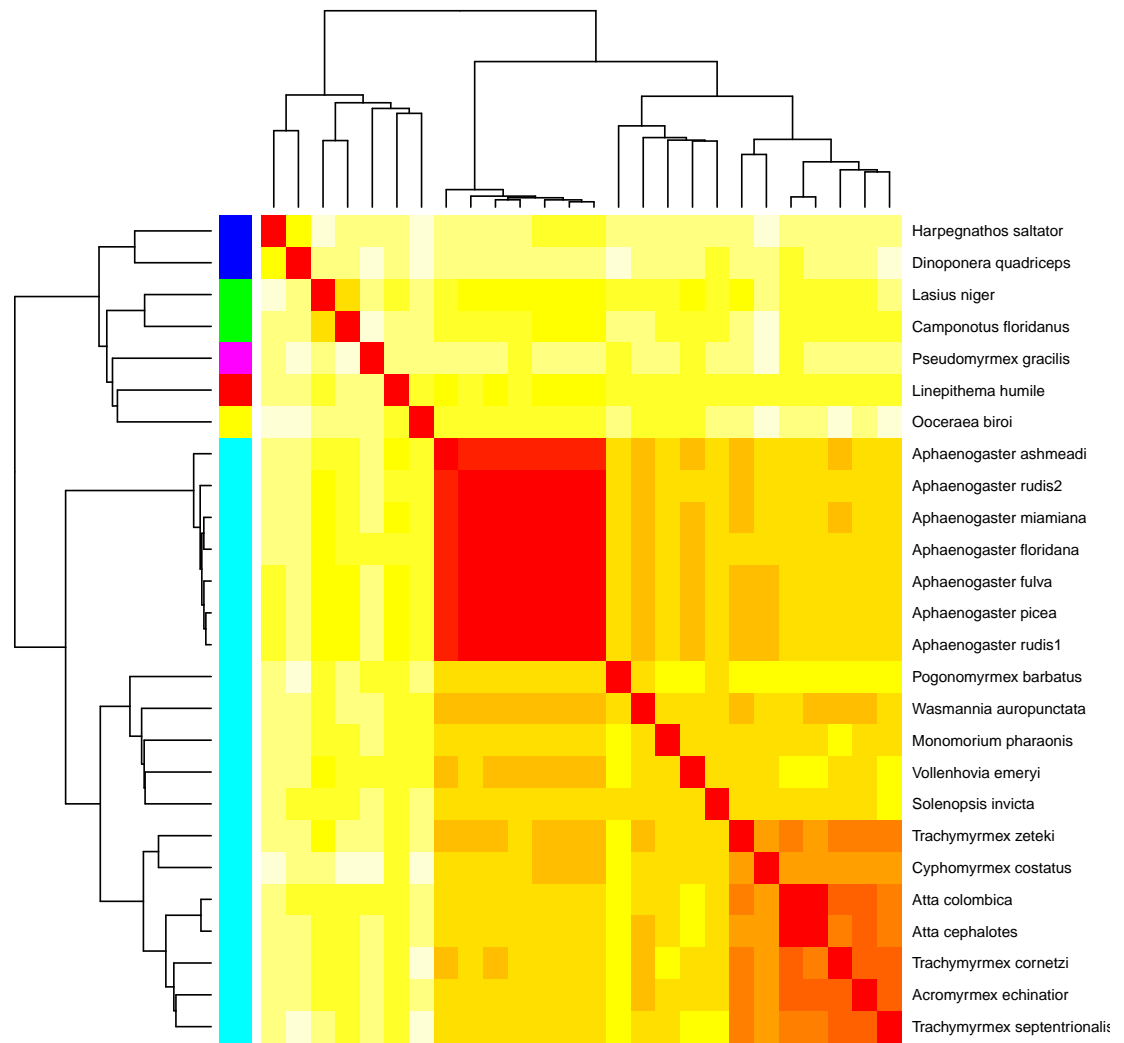


Figure 4. Heatmap of the MASH genomic distances of the *Aphaenogaster* species that we sampled together with other ant species in NCBI. Heat colors shown in the central matrix range from high (white = 1) through moderate (orange = 0.5) to low (red = 0) genomic distance; the diagonal is entirely red because 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).

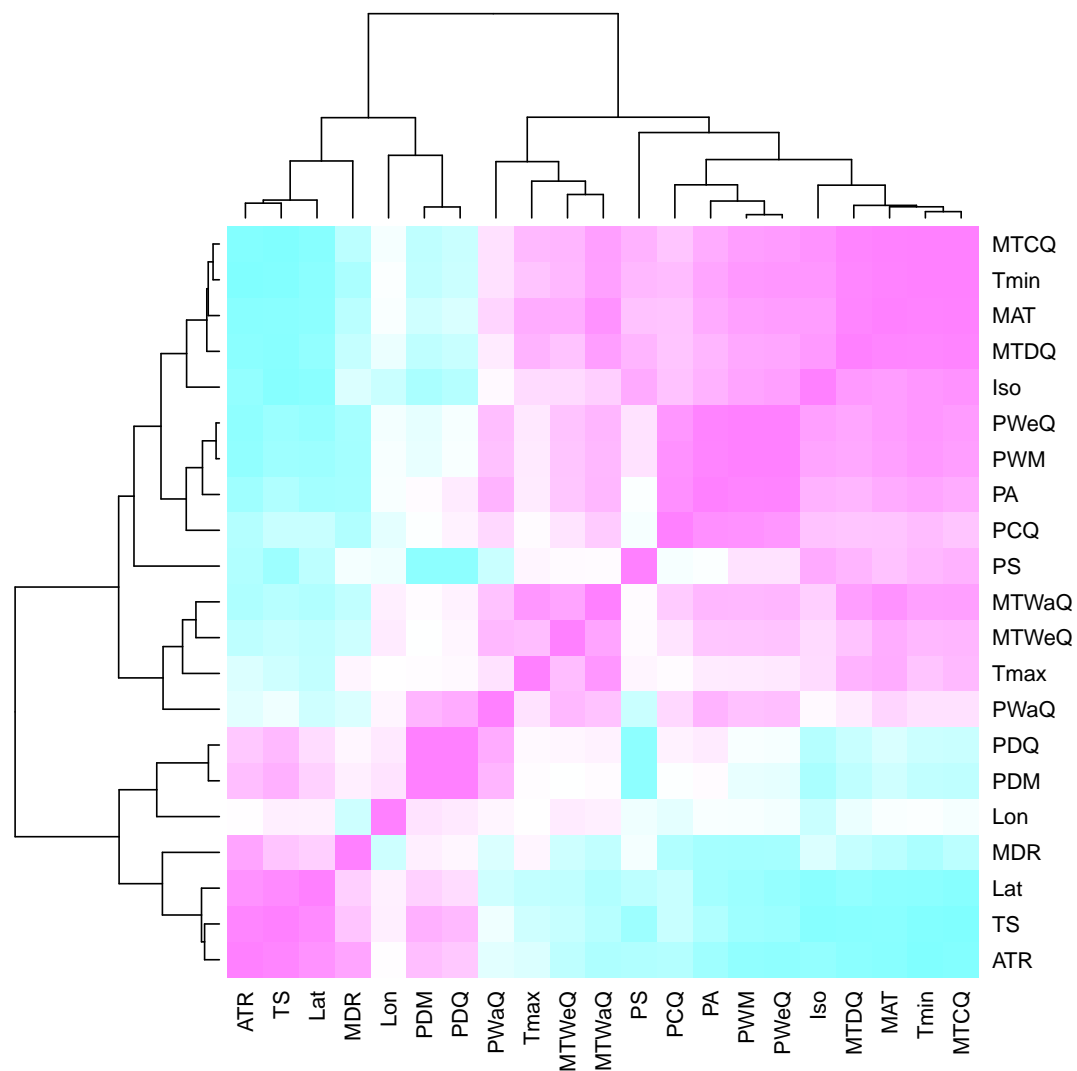


Figure 5. Heatmap of Pearson correlations among climate variables. Cells in the heatmap are colored by the correlation between the two variables that intersect at that location ranging from blue = -1 to white = 0 to pink = 1. The variables are arrayed by hierarchical clustering of the correlations, as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

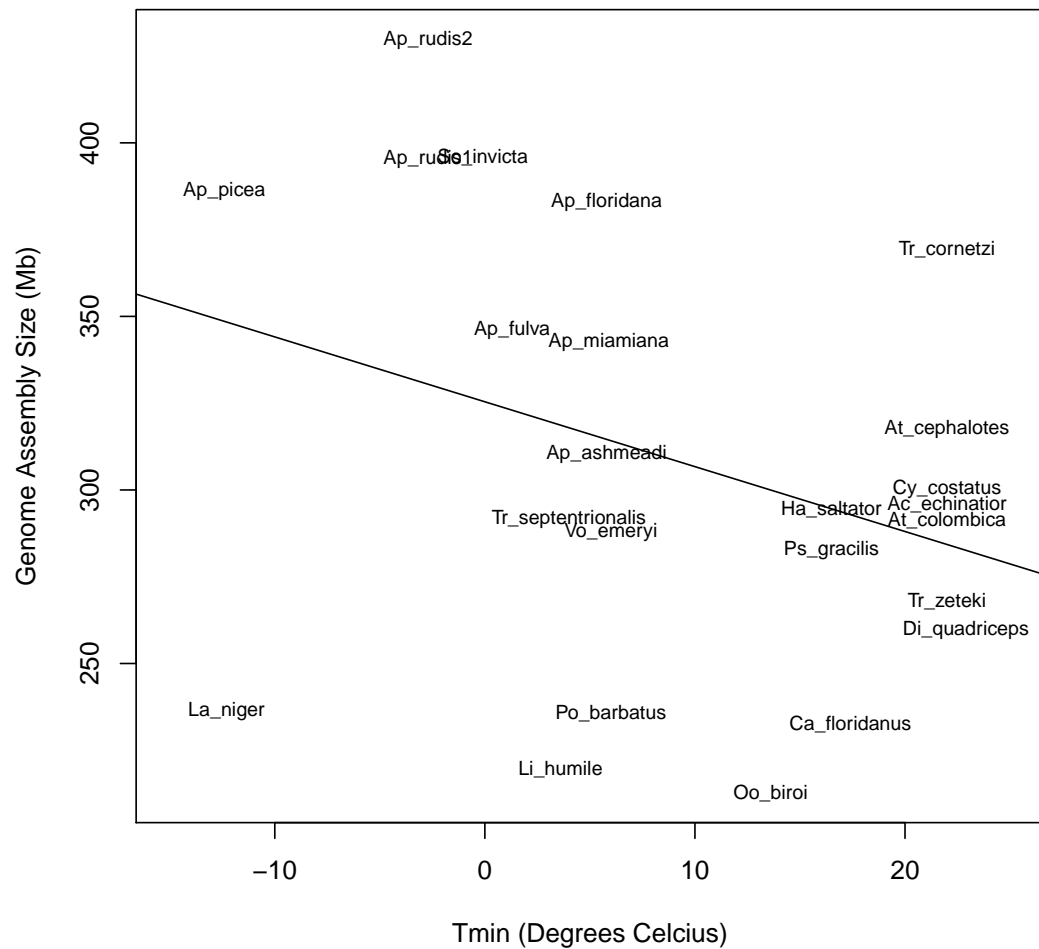


Figure 6. Bivariate plot showing the correlation between ant assembly size and minimum temperature of the coldest month (Tmin). Ants from locations with lower minimum temperatures tended to have larger genomes.

DISCUSSION

We have produced seven draft whole-genome sequences of six species of ants in the genus *Aphaenogaster*. The addition of the *Aphaenogaster* sequences increases the breadth of global ant genomic sampling, as 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 new *Aphaenogaster* sequences, our initial biogeographic analysis revealed that ant genomes from more similar climates have more similarly sized genomes with minimum temperatures having the strongest correlation with genome size, which is consistent with the hypothesis that climate has been a force shaping ant genome size.

Although correlative, our genome analysis results are consistent with the hypothesis that ants from regions with more similar climates tend to have similar sized genomes. 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 led to reductions in some ant species ranges by cold temperatures.

Specific to temperature, we found support for increasing genome size with colder temperatures. We are cautious to offer possible mechanisms for this trend. In the general context of arthropods, genome size appears to be influenced by a complex array of selection pressures, as evidenced by the recent study by Alfsnes et al. (2017), which found that genome size patterns varied greatly among major arthropod taxa with high potential for different mechanisms affecting genome size. For example, insects displayed clear phylogenetic correlations with genome size while genome size patterns in crustaceans were nearly independent of phylogeny but strongly related to biogeographic gradients (e.g. decreasing genome size with increasing maximum observed latitude). In addition, Hultgren et al. (2018) found evidence for

232 increasing genome size with latitude in crustaceans but not decapods, adding another example of the
233 potential complexity of genome size as an adaptive trait.

234 There is the potential for both direct and indirect selection for increased genome size in colder
235 conditions. Hessen et al. (2010) has proposed that there is a complex set of relationships among genome
236 size, developmental rate, cell size and body size and that selection can act at different points in this causal
237 network. One possible direct pathway is that increased expression of gene products that deal with cold
238 stress could lead to larger genomes via whole genome or individual gene duplications Dufresne and Jeffery
239 (2011). Previous work with *Aphaenogaster* supports this hypothesis, as Stanton-Geddes et al. (2016)
240 found that exposure to extreme cold induced expression of genes in the cold-climate *A. picea* more so than
241 in *A. carolinensis* (a more southern, warm climate species). An example of a possible indirect pathway
242 is that cold could select for increased body size to deal with heat-loss in cold conditions (Brown et al.,
243 2004), which could lead to larger genome sizes. There is evidence of cold selecting for greater body size
244 (i.e. Bergmann's Rule) in ants (Heinze et al., 2003; Bernadou et al., 2016), and some have hypothesized
245 that increased body size could lead indirectly to increased genome size via increased cell size (Ryan
246 Gregory, 2005); however, the most recent, broad analysis of genome size in ants (that we are aware
247 of) did not find support for a relationship between ant genome size and body size after controlling for
248 phylogenetic patterns (Tsutsui et al., 2008). Therefore, assuming that our analysis adequately controlled
249 for phylogenetics, indirect selection on genome size via body size is not a likely explanation for our
250 observed relationship between genome size and temperature.

251 It is important to keep in mind that the climate related genomic patterns observed in this study should
252 be considered an initial view of possible biogeographic patterns in ant genomes. As the addition of the
253 these sequences had a marked impact on the statistical results of the climate analysis (see Supplementary
254 Materials Table 2), we expect that further sequencing work will continue to enhance our understanding of
255 the ecological genomics of ants. Also, these findings should be tested with additional sequencing efforts,
256 as we could not control for several potentially important intercorrelated variables, with the most prominent
257 factor being phylogenetics. Additionally, factors such as sampling bias and sequencing methodology
258 (e.g. 454 versus Illumina) also varied among sequencing efforts, which could have contributed to some of
259 the observed correlations with climate. We did not attempt to control for these factors statistically due
260 to the limitations of the current ant genome sample size. Future work should methodologically and/or
261 statistically control for such sources of variation in ant genomes as more sequences become available to
262 elucidate clearer patterns and resolve underlying ecological and evolutionary mechanisms.

CONCLUSION

Although we have increased the total number of sequenced ant genomes by over 30%, the total number of ant sequences analyzed here is still a relatively small sample ($n = 26$) of the estimated $>16,000$ ant species and subspecies (www.antweb.org, accessed 20 Aug 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.
- Bernadou, A., Römermann, C., Gratiashvili, N., and Heinze, J. (2016). Body size but not colony size increases with altitude in the holarctic ant, *Leptothorax acervorum*. *Ecol. Entomol.*, 41(6):733–736.
- 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, J. H., Gillooly, J. F., Allen, A. P., Savage, V. M., and West, G. B. (2004). Toward a metabolic theory of ecology. *Ecology*.
- 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). *Myrmecological News*, 17:133–146.
- 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

317 effects as the climate changes. *Integr. Comp. Biol.*, 57(1):112–120.

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

321 Diamond, S. E., Nichols, L. M., Pelini, S. L., Penick, C. A., Barber, G. W., Cahan, S. H., Dunn, R. R.,
 322 Ellison, A. M., Sanders, N. J., and Gotelli, N. J. (2016). Climatic warming destabilizes forest ant
 323 communities. *Sci. Adv.*, 2(10):e1600842–e1600842.

324 Diamond, S. E., Sorger, D. M., Hulcr, J., Pelini, S. L., Toro, I. D., Hirsch, C., Oberg, E., and Dunn, R. R.
 325 (2012). Who likes it hot? A global analysis of the climatic, ecological, and evolutionary determinants
 326 of warming tolerance in ants. *Glob. Chang. Biol.*, 18(2):448–456.

327 Dufresne, F. and Jeffery, N. (2011). A guided tour of large genome size in animals: what we know and
 328 where we are heading. *Chromosom. Res.*, 19(7):925–938.

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

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

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

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

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

340 Heinze, J., Foitzik, S., Fischer, B., Wanke, T., and Kipyatkov, V. E. (2003). The significance of latitudinal
 341 variation in body size in a holarctic ant, *Leptothorax acervorum*. *Ecography (Cop.)*, 26(3):349–355.

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

346 Hessen, D. O., Jeyasingh, P. D., Neiman, M., and Weider, L. J. (2010). Genome streamlining and the
 347 elemental costs of growth. *Trends Ecol. Evol.*, 25:75–80.

348 Hultgren, K. M., Jeffery, N. W., Moran, A., and Gregory, T. R. (2018). Latitudinal variation in genome
 349 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.

Libbrecht, R., Oxley, P. R., Kronauer, D. J. C., and Keller, L. (2013). Ant genomics sheds light on the molecular regulation of social organization. *Genome Biol.*, 14(7):212.

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.

Mousseau, T. A. (1997). Ectotherms Follow the Converse to Bergmann's Rule. *Evolution (N. Y.)*, 51(2):630.

Mueller, U. G., Mikheyev, A. S., Hong, E., Sen, R., Warren, D. L., Solomon, S. E., Ishak, H. D., Cooper, M., Miller, J. L., Shaffer, K. A., and Juenger, T. E. (2011). Evolution of cold-tolerant fungal symbionts permits winter fungiculture by leafcutter ants at the northern frontier of a tropical ant-fungus symbiosis. *Proc. Natl. Acad. Sci.*, 108(10):4053–4056.

Neafsey, D. E., Lawniczak, M. K. N., Park, D. J., Redmond, S. N., Coulibaly, M. B., Traoré, S. F., Sagnon, N., Costantini, C., Johnson, C., Wiegand, R. C., Collins, F. H., Lander, E. S., Wirth, D. F., Kafatos, F. C., Besansky, N. J., Christophides, G. K., and Muskavitch, M. A. T. (2010). SNP genotyping defines complex gene-flow boundaries among African malaria vector mosquitoes. *Science*, 330(6003):514–517.

Nguyen, A. D., DeNovellis, K., Resendez, S., Pustilnik, J. D., Gotelli, N. J., Parker, J. D., and Cahan, S. H. (2017). Effects of desiccation and starvation on thermal tolerance and the heat-shock response in forest ants. *J. Comp. Physiol. B*, 187(8):1107–1116.

Nygaard, S. and Wurm, Y. (2015). Ant genomics (Hymenoptera: Formicidae): Challenges to overcome and opportunities to seize. *Myrmecological News*, 21:59–72.

Oksanen, J., Blanchet, F., Kindt, R., Legendre, P., and O'Hara, R. (2016). Vegan: community ecology package.

Ondov, B. D., Treangen, T. J., Melsted, P., Mallonee, A. B., Bergman, N. H., Koren, S., and Phillippy, A. M. (2016). Mash: fast genome and metagenome distance estimation using MinHash. *Genome Biol.*, 17(1):132.

Parmesan, C. (2006). Ecological and Evolutionary Responses to Recent Climate Change. *Annu. Rev. Ecol. Evol. Syst.*, 37(1):637–669.

383 Pelini, S. L., Diamond, S. E., Nichols, L. M., Stuble, K. L., Ellison, A. M., Sanders, N. J., Dunn, R. R.,
384 and Gotelli, N. J. (2014). Geographic differences in effects of experimental warming on ant species
385 diversity and community composition. *Ecosphere*, 5(10):art125.

386 Penick, C. A., Diamond, S. E., Sanders, N. J., and Dunn, R. R. (2017). Beyond thermal limits: com-
387 prehensive metrics of performance identify key axes of thermal adaptation in ants. *Funct. Ecol.*,
388 31(5):1091–1100.

389 Petrov, D. A. (2001). Evolution of genome size: new approaches to an old problem. *Trends Genet.*,
390 17(1):23–28.

391 R Core Team (2017). R Core Team (2017). R: A language and environment for statistical computing. *R*
392 *Found. Stat. Comput. Vienna, Austria*.

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

395 Ryan Gregory, T. (2005). Genome Size Evolution in Animals. *Evol. Genome*.

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

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

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

404 Tsutsui, N. D., Suarez, A. V., Spagna, J. C., Johnston, J. S., Gregory, T., Evans, J., Gundersen-Rindal, D.,
405 Gardner, T., Gregory, T., Wilson, E., Hölldobler, B., Wilson, E., Li, J., Heinz, K., Johnston, J., Ross,
406 L., Beani, L., Hughes, D., Kathirithamby, J., Geraci, N., Johnston, J., Robinson, J., Wikel, S., Hill, C.,
407 Gregory, T., Bennett, M., Leitch, I., SanMiguel, P., Gaut, B., Tikhonov, A., Nakajima, Y., Bennetzen, J.,
408 Kazazian, H., Kidwell, M., Comeron, J., Ustinova, J., Achmann, R., Cremer, S., Mayer, F., Hancock,
409 J., Hancock, J., Toth, G., Gaspari, Z., Jurka, J., Redon, R., Ishikawa, S., Fitch, K., Feuk, L., Perry,
410 G., Andrews, T., Fiegler, H., Shapero, M., Carson, A., Chen, W., Cho, E., Dallaire, S., Freeman, J.,
411 Gonzalez, J., Gratacos, M., Huang, J., Kalaitzopoulos, D., Komura, D., MacDonald, J., Marshall, C.,
412 Mei, R., Montgomery, L., Nishimura, K., Okamura, K., Shen, F., Somerville, M., Tchinda, J., Valsesia,
413 A., Woodwark, C., Yang, F., Zhang, J., Zerjal, T., Zhang, J., Armengol, L., Conrad, D., Estivill, X.,
414 Tyler-Smith, C., Carter, N., Aburatani, H., Lee, C., Jones, K., Scherer, S., Hurles, M., Gregory, T.,
415 Gregory, T., Petrov, D., Lozovskaya, E., Hartl, D., Devos, K., Brown, J., Bennetzen, J., Bennetzen, J.,

416 Ma, J., Devos, K., Ma, J., Bennetzen, J., Oliver, M., Petrov, D., Ackerly, D., Falkowski, P., Schofield, O.,
 417 Gregory, T., Hebert, P., Kolasa, J., Finston, T., Hebert, P., Footitt, R., Ferrari, J., Rai, K., Ellegren, H.,
 418 Vandenbussche, R., Longmire, J., Baker, R., Organ, C., Shedlock, A., Meade, A., Pagel, M., Edwards,
 419 S., Hughes, A., Hughes, M., Reinhold, K., Gregory, T., Pittendrigh, B., Clark, J., Johnston, J., Lee,
 420 S., Romero-Severson, J., Dasch, G., Gregory, T., Weinstock, G., Robinson, G., Gibbs, R., Worley, K.,
 421 Evans, J., Maleszka, R., Robertson, H., Weaver, D., Beye, M., Bork, P., Elsik, C., Hartfelder, K., Hunt,
 422 G., Zdobnov, E., Amdam, G., Bitondi, M., Collins, A., Cristino, A., Lattorff, H., Lobo, C., Moritz,
 423 R., Nunes, F., Page, R., Simoes, Z., Wheeler, D., Carninci, P., Fukuda, S., Hayashizaki, Y., Kai, C.,
 424 Kawai, J., Sakazume, N., Sasaki, D., Tagami, M., Albert, S., Baggerman, G., Beggs, K., Bloch, G.,
 425 Cazzamali, G., Cohen, M., Drapeau, M., Eisenhardt, D., Emore, C., Ewing, M., Fahrback, S., Foret,
 426 S., Grimmelikhuijzen, C., Hauser, F., Hummon, A., Huybrechts, J., Jones, A., Kadowaki, T., Kaplan,
 427 N., Kucharski, R., Leboulle, G., Linial, M., Littleton, J., Mercer, A., Richmond, T., Rodriguez-Zas, S.,
 428 Rubin, E., Sattelle, D., Schlipalius, D., Schoofs, L., Shemesh, Y., Sweedler, J., Velarde, R., Verleyen, P.,
 429 Vierstraete, E., Williamson, M., Ament, S., Brown, S., Corona, M., Dearden, P., Dunn, W., Elekonich,
 430 M., Fujiyuki, T., Gattermeier, I., Gempe, T., Hasselmann, M., Kadowaki, T., Kage, E., Kamikouchi, A.,
 431 Kubo, T., Kucharski, R., Kunieda, T., Lorenzen, M., Milshina, N., Morioka, M., Ohashi, K., Overbeek,
 432 R., Ross, C., Schioett, M., Shippy, T., Takeuchi, H., Toth, A., Willis, J., Wilson, M., Gordon, K.,
 433 Letunic, I., Hackett, K., Peterson, J., Felsenfeld, A., Guyer, M., Solignac, M., Agarwala, R., Cornuet,
 434 J., Monnerot, M., Mougél, F., Reese, J., Vautrin, D., Gillespie, J., Cannone, J., Gutell, R., Johnston,
 435 J., Eisen, M., Iyer, V., Iyer, V., Kosarev, P., Mackey, A., Solovyev, V., Souvorov, A., Aronstein, K.,
 436 Bilikova, K., Chen, Y., Clark, A., Decanini, L., Gelbart, W., Hetru, C., Hultmark, D., Imler, J., Jiang,
 437 H., Kanost, M., Kimura, K., Lazzaro, B., Lopez, D., Simuth, J., Thompson, G., Zou, Z., Jong, P. D.,
 438 Sodergren, E., Csuros, M., Milosavljevic, A., Osoegawa, K., Richards, S., Shu, C., Duret, L., Elhaik, E.,
 439 Graur, D., Anzola, J., Campbell, K., Childs, K., Collinge, D., Crosby, M., Dickens, C., Grametes, L.,
 440 Grozinger, C., Jones, P., Jorda, M., Ling, X., Matthews, B., Miller, J., Mizzen, C., Peinado, M., Reid, J.,
 441 Russo, S., Schroeder, A., Pierre, S. S., Wang, Y., Zhou, P., Jiang, H., Kitts, P., Ruef, B., Venkatraman,
 442 A., Zhang, L., Aquino-Perez, G., Whitfield, C., Behura, S., Berlocher, S., Sheppard, W., Smith, D.,
 443 Suarez, A., Tsutsui, N., Wei, X., Wheeler, D., Havlak, P., Li, B., Liu, Y., Sodergren, E., Jolivet, A., Lee,
 444 S., Nazareth, L., Pu, L., Thorn, R., Stolc, V., Newman, T., Samanta, M., Tongprasit, W., Claudianos,
 445 C., Berenbaum, M., Biswas, S., de Graaf, D., Feyereisen, R., Johnson, R., Oakeshott, J., Ranson, H.,
 446 Schuler, M., Muzny, D., Chacko, J., Davis, C., Dinh, H., Gill, R., Hernandez, J., Hines, S., Hume,
 447 J., Jackson, L., Kovar, C., Lewis, L., Miner, G., Morgan, M., Nguyen, N., Okwuonu, G., Paul, H.,
 448 Santibanez, J., Savery, G., Svatek, A., Villasana, D., Wright, R., Consort, H., Moreau, C., Bell, C.,

449 Vila, R., Archibald, S., Pierce, N., Brady, S., Schultz, T., Fisher, B., Ward, P., Mueller, U., Gerardo, N.,
 450 Aanen, D., Six, D., Schultz, T., Chapela, I., Rehner, S., Schultz, T., Mueller, U., Wetterer, J., Schultz,
 451 T., Meier, R., Gregory, T., Hebert, P., Gregory, T., Shorthouse, D., Wang, J., Jemielity, S., Uva, P.,
 452 Wurm, Y., Graff, J., Keller, L., Bennett, M., Leitch, I., Price, H., Johnston, J., Abouheif, E., Reeve, J.,
 453 Abouheif, E., Felsenstein, J., Purvis, A., and Rambaut, A. (2008). The evolution of genome size in ants.
 454 *BMC Evol. Biol.*, 8(1):64.

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

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

459 Ward, P. S., Brady, S. G., Fisher, B. L., and Schultz, T. R. (2015). The evolution of myrmicine ants:
 460 phylogeny and biogeography of a hyperdiverse ant clade (Hymenoptera: Formicidae). *Syst. Entomol.*,
 461 40(1):61–81.

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

SUPPLEMENTARY MATERIALS

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

	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>Pseudo-F</i>	<i>R2</i>	<i>p-value</i>
<i>Assembly Size Similarity</i>						
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	

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.