Forum

The Global Ant Genomics Alliance (GAGA)

Jacobus J. Boomsmma*, Centre for Social Evolution, Department of Biology, University of Copenhagen, 2100 Copenhagen, Denmark. E-mail: jboomsmma@bio.ku.dk
Seán G. Brady, Department of Entomology, National Museum of Natural History, Smithsonian Institution, Washington, DC 20013-7012, USA.
Robert R. Dunn, Department of Applied Ecology, North Carolina State University, Raleigh, NC 27695-7613, USA;
Center for Macroecology, Evolution, and Climate, Natural History Museum of Denmark, University of Copenhagen, 2100 Copenhagen, Denmark.
Jürgen Gadau, Institute for Evolution and Biodiversity, University of Münster, D-48149 Münster, Germany.
Jürgen Heinze, LS Zoologie / Evolutionsbiologie, University of Regensburg, D-93040 Regensburg, Germany.
Laurent Keller, Department of Ecology and Evolution, University of Lausanne, CH-1015 Lausanne, Switzerland.
Lukas S. Moreau, Integrative Research Center, Department of Science and Education, Field Museum of Natural History, Chicago, Illinois, USA.
Nathan J. Sanders, Center for Macroecology, Evolution, and Climate, Natural History Museum of Denmark, University of Copenhagen, 2100 Copenhagen, Denmark.
Lukas Schrader, Centre for Social Evolution, Department of Biology, University of Copenhagen, 2100 Copenhagen, Denmark; Institute for Evolution and Biodiversity, University of Münster, D-48149 Münster, Germany. E-mail: Lukas.Schrader@bio.ku.dk
Ted R. Schultz, Department of Entomology, National Museum of Natural History, Smithsonian Institution, Washington, DC 20013-7012, USA.
Lotta Sundström, Centre of Excellence in Biological Interactions, Department of Biosciences, University of Helsinki, Helsinki, Finland.
Philip S. Ward, Department of Entomology & Nematology, University of California at Davis, One Shields Avenue, Davis, CA 95616, USA.
William T. Wcislo, Smithsonian Tropical Research Institute, Apartado 0843-03092, Balboa, República de Panamá.
Guojie Zhang (contact author), Centre for Social Evolution, Department of Biology, University of Copenhagen, 2100 Copenhagen, Denmark; China National GeneBank, BGI-Shenzhen, 518083 Shenzhen, China; State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, 650223 Kunming, China. E-mail: guojie.zhang@bio.ku.dk
The GAGA Consortium; for details, see Appendix S1, as digital supplementary material to this article, at the journal's web pages.
* Authors listed in alphabetical order.

Key words: Phylogenomics, caste polymorphism, division of labor, ecology, symbiosis, life history, evolution, reproductive conflict.

For centuries, scholars, the lay public, and children alike have been enthralled by ants because of their fascinating behaviors and social organization. Decades of biological research have documented that the ants have evolved a stunning global diversity (WARD 2014) (Fig. 1). More than 15,000 species belonging to over 330 genera have been named to date (ANTWEB 2017) and new species are discovered almost every day. While some of these species are extraordinarily rare, known from a single collection or site, others have spread around the world. Collectively, ants are among the most abundant insects in most terrestrial ecosystems on every continent except Antarctica.

Over more than 100 million years of history, the ants have been impressive social innovators. They evolved advanced division of labor; assembly line processing of food resources; mass-raiding predatory columns; seed-harvesting and storage; sophisticated chemical, acoustic and visual communication systems; fungus-farming for food; herding aphids as livestock for meat and honeydew; practicing biological and chemical pest control; and social immune systems akin to efficient public health systems (HOLLDÖBLER & WILSON 1990, BOURKE & FRANKS 1995, CREMER & SIXT 2009).

What sets ants apart from all but a handful of other animal lineages is their advanced reproductive division of labor with specialized, morphologically distinct castes for breeding (queens), nursing and food-gathering (workers), and sometimes defense (soldiers), all of which are unconditionally devoted to a single colony (CROZIER & PAMILO 1996). Dozens of ant species have been intensively studied by ecologists, geneticists, biogeographers, and evolutionary biologists, both in the field and in the laboratory. Some ant species have also become model organisms in studies of epidemiology, aging, development, and neurobiology.

During the last seven years, 20 completely sequenced ant genomes have been published, so the ants are among the genetically best-known insect families (Tab. 1). These studies have provided remarkable information about the genetic underpinnings of key adaptations such as those that accompanied fungus farming and allowed communication and recognition systems to become highly sophisticated (SIMOLA et al. 2013), but the big comparative picture of genomic diversity across the entire family is still lacking. The fact that ant genomes tend to be modest in size has greatly facilitated genome sequencing efforts so far, and now allows us to pursue much more ambitious objectives.

We have recently launched the Global Ant Genomics Alliance (GAGA) to fundamentally advance the multidisciplinary genomics-based study of ant biology. Through this program, we will obtain full high-quality genome se-
Fig. 1: The evolutionary and biological diversity of the ants (Formicidae) captured with their tree of life and representative images. The phylogenetic tree is based on over 1000 species from 287 genera and shows the relationship between the seventeen currently recognized subfamilies (BRANSTETTER & al. 2017). (A) Threat display of a Solenopsis fire ant worker (Myrmicinae); (B) plant ants of the genus Pseudomyrmex (Pseudomyrmicinae) on their mutualistic Acacia host plant in Cerro Zuela (Panama); (C) a minor and a medium worker of the fungus-farming ant Acromyrmex echinatior (Myrmicinae) caring for their fungal symbiont; (D) a raiding column of the army ant Eciton burchellii (Dorylinae); (E) a Gigantiops destructor worker (Formicinae); (F) a lab colony of Pheidole californica (Myrmicinae); (G) cooperation in nest construction between workers of Oecophylla smaragdina (Formicinae) in Queensland, Australia; (H) foraging worker of Dinoponera australis (Ponerinae); (I) broodcare behavior in a lab colony of Cardiocondyla obscurior (Myrmicinae). Photo credits: Lukas Schrader (A, B, C, G, I); Alex Wild (D, E, F, H).

sequences of at least one representative of each ant genus that can be feasibly collected. GAGA will provide the overarching framework necessary for a series of large-scale comparative research programs to explain the genomic underpinnings of key innovations that globally characterize the ants as a group, as well as those that characterize the distinct subgroups.

At the same time, GAGA will actively encourage sequencing efforts that focus on more detailed species-level genomic comparisons in lineages (often tribes) of ants with unique life histories that are likely to have retained more specific signatures of molecular adaptation in their genomes. We expect such projects to be coordinated by PI groups with their own funding independent of the GAGA
Tab. 1: The 20 ant species for which sequenced genomes are available, in alphabetical order. Scaffold N50 and contig N50 are standard parameters for evaluating and comparing the quality of genome assemblies. Different sequencing strategies and assembly methods produce draft genomes of varying quality as illustrated by the figures in the table. The GAGA consortium aims to produce a consensus on the genome assembly standard for ant genomics research. With access to the best possible sequencing technologies, we expect to be able to produce top-quality assemblies with only a few thousand gaps, similar to the latest reference genomes produced for humans. We anticipate that GAGA will thus be able to address many biological questions that could not be resolved with existing fragmented and less well assembled genomes. Examples of such questions include chromosome evolution, synteny between genes far apart on chromosomes, reconstruction of gene-family evolutionary histories, and the evolution of gene regulatory regions.

<table>
<thead>
<tr>
<th>Species name</th>
<th>Assembly size Mb</th>
<th>Scaffold N50 Mb</th>
<th>Contig N50 Mb</th>
<th>Biological questions addressed</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acromyrmex echinatior</td>
<td>300</td>
<td>1,094.27</td>
<td>62.71</td>
<td>Mutualism with fungi, gene loss, arginin-loss, neuropeptides, sex determination</td>
<td>NYGAARD &amp; al. (2011)</td>
</tr>
<tr>
<td>Atta cephalotes</td>
<td>318</td>
<td>5,154.50</td>
<td>14.24</td>
<td>Mutualism with fungi, arginin-loss, serine proteases, hexamerin</td>
<td>SUEI &amp; al. (2011)</td>
</tr>
<tr>
<td>Atta colombica</td>
<td>293</td>
<td>2,037.15</td>
<td>15.00</td>
<td>Multigenome analysis of mutualistic performance in ants and fungal crops</td>
<td>NYGAARD &amp; al. (2016)</td>
</tr>
<tr>
<td>Camponotus floridanus</td>
<td>238</td>
<td>602.92</td>
<td>24.13</td>
<td>Genomic comparison, caste, epigenetics of aging and behavior</td>
<td>BONASIO &amp; al. (2010)</td>
</tr>
<tr>
<td>Cardiocondyla obscurior</td>
<td>182</td>
<td>2,570.86</td>
<td>14.94</td>
<td>Genome reduction, transposable elements, differential methylation</td>
<td>SCHRAER &amp; al. (2014)</td>
</tr>
<tr>
<td>Cyphomyrmex costatus</td>
<td>323</td>
<td>1,016.47</td>
<td>27.20</td>
<td>Multigenome analysis of mutualistic performance in ants and fungal crops</td>
<td>NYGAARD &amp; al. (2016)</td>
</tr>
<tr>
<td>Dinoponera quadriiceps</td>
<td>260</td>
<td>1,361.23</td>
<td>29.91</td>
<td>Phenotypic plasticity, comparison with wasps</td>
<td>PATALANO &amp; al. (2015)</td>
</tr>
<tr>
<td>Formica selysi</td>
<td>224</td>
<td>16.10</td>
<td>No data</td>
<td>Non-recombining social chromosome comparison with Solenopsis invicta</td>
<td>PURCELL &amp; al. (2014)</td>
</tr>
<tr>
<td>Harpegnathos saltator</td>
<td>297</td>
<td>598.19</td>
<td>38.03</td>
<td>Genomic comparison, caste, epigenetics of aging and behavior</td>
<td>BONASIO &amp; al. (2010)</td>
</tr>
<tr>
<td>Lasius niger</td>
<td>236</td>
<td>17.06</td>
<td>16.38</td>
<td>Transposable elements, detoxification</td>
<td>KONOROV &amp; al. (2017)</td>
</tr>
<tr>
<td>Linepithema humile</td>
<td>216</td>
<td>1,386.36</td>
<td>35.86</td>
<td>Gene family expansions, odorant receptors, royal jelly, desaturases</td>
<td>SMITH &amp; al. (2011a)</td>
</tr>
<tr>
<td>Oooceraea biroi</td>
<td>214</td>
<td>1,291.49</td>
<td>31.93</td>
<td>Queen loss, aomicitic parthenogenesis</td>
<td>OXLEY &amp; al. (2014)</td>
</tr>
<tr>
<td>Pogonomyrmyx barbatus</td>
<td>235</td>
<td>793.75</td>
<td>11.61</td>
<td>Caste-specific gene networks, gene family expansions, comparisons with bees</td>
<td>SMITH &amp; al. (2011b)</td>
</tr>
<tr>
<td>Pseudomyrmex gracilis</td>
<td>282</td>
<td>317.68</td>
<td>38.83</td>
<td>Rates of genome-wide evolution in mutualisms with plants</td>
<td>RUBIN &amp; MOREAU (2016)</td>
</tr>
<tr>
<td>Solenopsis invicta</td>
<td>353</td>
<td>720.58</td>
<td>14.67</td>
<td>Vitellogenin, sex differentiation, telomerase, odorant receptors</td>
<td>WURM &amp; al. (2011)</td>
</tr>
<tr>
<td>Trachymyrmex septentrionalis</td>
<td>295</td>
<td>2,447.26</td>
<td>14.51</td>
<td>Multigenome analysis of mutualistic performance in ants and fungal crops</td>
<td>NYGAARD &amp; al. (2016)</td>
</tr>
<tr>
<td>Trachymyrmex cornetti</td>
<td>402</td>
<td>632.81</td>
<td>12.44</td>
<td>Multigenome analysis of mutualistic performance in ants and fungal crops</td>
<td>NYGAARD &amp; al. (2016)</td>
</tr>
<tr>
<td>Trachymyrmex zeteki</td>
<td>269</td>
<td>1,333.95</td>
<td>18.99</td>
<td>Multigenome analysis of mutualistic performance in ants and fungal crops</td>
<td>NYGAARD &amp; al. (2016)</td>
</tr>
<tr>
<td>Vollenhovia emeryi</td>
<td>288</td>
<td>1,346.09</td>
<td>32.42</td>
<td>Genomic changes in inquiline social parasites</td>
<td>SMITH &amp; al. (2015)</td>
</tr>
</tbody>
</table>

objectives, but we hope they will interact with GAGA to share data and facilitate comparisons of genomic data sets with sister lineages that will be sequenced under the auspices of GAGA.

GAGA is among the most ambitious new global initiatives for comparative genomics (PENNI H 2017), following in the footsteps of the recently completed avian comparative genomics project (ZHANG & al. 2014), which
demonstrated the power of large-scale collaborations for uncovering global patterns of phenotypic diversification, evolutionary convergence, and genome-level evolutionary constraints. By applying this comprehensive approach to elucidate the global genomic diversity of the ants, we aim to address a series of fundamental questions in ecology, evolution, and developmental biology, and we expect to obtain significant novel insights.

Because the ants are already one of the best-studied insect lineages, they offer many opportunities to combine genomic insights (NYGAARD & WURM 2015) with comparative data on natural history, adaptive radiation, ecological functionality, species distributions, and complex behavior coordinated by advanced communication. Other opportunities for significant improvements in biological understanding will emerge from the comparative analyses of genomic characteristics that we hope to connect with, for example, degrees of division of labor between queen and worker castes, and ancestral haplodiploid or evolutionarily derived novel sex determination systems.

In addition to the genomic data, we will coordinate the compilation of macroecological datasets on global ant biodiversity and biogeography to be supplemented by as much core life-history data as the global community can provide. Life histories of ant species and genera vary dramatically in terms of colony structure, reproduction, diet, habitat, range size, interactions with other ants, and ecological dominance. This extraordinary diversity offers unique opportunities for examining the molecular and genomic correlates of ecological abundance and life-history traits (DUNN & al. 2009, GIBB & al. 2015) and for assessing putative explanations for such correlations from first principles.

We expect that GAGA will significantly advance our knowledge of phylogenomics, providing unprecedented resolution of the ant tree of life that will allow novel understanding of the forces that drove the adaptive radiations of the ant subfamilies. Explaining the evolution and elaboration of physically differentiated castes will be of paramount importance, both in terms of phenotypic functionality and developmental biology. Independent of phylogeny, GAGA will also offer key insights in the convergent evolution of adaptations to extreme environments and the recurrent origins of symbioses with microbes and macroscopic partners serving nutritional or protective purposes.

Finally, the ants stand out as the most diverse lineage for the expression of reproductively altruistic behavior, mediated by an enormous variation in primarily chemical communication systems. Although haplodiploidy has been retained throughout the ants, some lineages evolved parthenogenesis or experimented with mixtures of sexual and asexual propagation of either worker or queen castes (SCHWANDER & al. 2010). Haplodiploidy also endowed the ants with a unique burden of potential reproductive conflicts over sex-ratio investments, the regulation of male-production by workers, and the emergence of multi-queen colonies in most if not all major lineages (BOURKE & FRANKS 1995, CROZIER & PAMILO 1996, BOOMSMA & al. 2014). The investigation of genomic rearrangements that induce changes in social organization and control of potential reproductive conflicts has only just begun and stands to gain significantly from the large-scale availabil-


Funding for sequencing the first ca. 200 genomes has been secured via a Chinese biodiversity genomics grant embedded in the Kunming Institute of Zoology of the Chinese Academy of Sciences, and we anticipate that funding for 100 - 200 more genomes will become available as GAGA develops and reaches its full potential in the coming one to two years. With regard to technology, GAGA aims to produce high-quality genome assemblies via complementary approaches that can accommodate both common ant species, for which sufficient biomass is relatively easy to obtain, and rare species, for which genomes will have to be generated from minute quantities of high quality DNA.

GAGA will be supported by a worldwide bioinformatics analysis group with rich experience in comparative ant genomics. The consortium will further provide transparent guidelines for accessing the most cutting-edge sequencing technologies, so the best value for the money can be ensured throughout the GAGA program. We aim to establish a common standard for data collection and analysis, which will allow GAGA to serve the global community of ant biologists in the best possible way, and to make sure that GAGA-generated genome sequences will comply with the ambitious assembly and annotation standards that soon will be universally required for published reference genomes.

We are reaching out to the international community of myrmecologists and other biological specialists to participate – with collecting efforts, unique expertise in the annotation of specific gene families, and with grant applications to fund complementary genome-sequencing efforts focusing on specific ant groups for which researchers aim to obtain multiple species-level genomes and for which GAGA can provide the outgroup genomes. We hope that the global community of ant biologists will spearhead new research programs on wide-ranging research questions that can be facilitated under the GAGA umbrella. GAGA will ensure fair credit to all contributing researchers independent of their type of employment and career stage. Collection trips to obtain the first ca. 100 ant species are planned for 2017 and 2018 and we hope to hear from other myrmecologists who are interested in joining the global sampling endeavor.

Interested researchers can contact GAGA via the project website (http://antgenomics.dk) to learn more about the initiative, to inspect and develop the list of ant species targeted for collection and sequencing, to propose new or alternative species within genera, or to offer expertise to contribute to ongoing projects. We are also eager to hear from PIs who would like to propose and fund larger comparative projects so they can spearhead analyses for complementary papers, similar to the process that generated a number of flagship papers in the bird program on which the coordinators were neither first nor last authors. We aim to publish the overarching results of the GAGA efforts as a coordinated set of papers in 2021. We anticipate that not all global ant genome-sequencing efforts during the coming five years can be embedded into GAGA, but we will be eager to discuss pre-publication data-sharing with other groups to encourage sharing of credits in the timeliest way possible.
In sum, we consider GAGA to be a platform for numerous existing research programs and a key instrument for facilitating novel collaborations that address fundamental questions about one of the most interesting and most amazing animal lineages on the planet – the ants.

References


