

## NEWS AND VIEWS

## MEETING REVIEW

**Ant genomics: strength and diversity in numbers**

CHRISTOPHER D. SMITH,\* CHRISTOPHER R. SMITH,† ULRICH MUELLER‡ and JÜRGEN GADAU§

\*SFSU Dept. Biology, 1600 Holloway Ave, San Francisco, CA 94132, USA, †Department of Biology, Earlham College, 801 National Road West, Richmond, IN 47374, USA, ‡Integrative Biology, University of Texas at Austin, Austin, TX 78712, USA, §School of Life Sciences, Arizona State University, Tempe, AZ Tempe, AZ 85287-4501, USA

**Abstract**

A recent workshop held at the Arizona State University Center for Social Dynamics and Complexity gathered over 50 prominent researchers from around the globe to discuss the development of genomic resources for several ant species. Ants play crucial roles in many ecological niches and the sequencing of several ant genomes promises to elucidate topics ranging from the genetic basis for social complexity, longevity and behaviour to systems biology and the identification of novel antimicrobial compounds. Unlike other species, most ant genomes are being generated by individual labs and small collaborations without the annotation and computational resources that support prominent model organism genome databases such as those for the fruitfly and roundworm. Attendees summarized their current progress and future plans for several ant genomes and discussed how best to coordinate the analysis and annotation of ant sequences to benefit the broad research interests of the social insect community.

**Keywords:** behaviour, bioinformatics, ecological genetics, genomics, insects, transcriptomics

Received 6 August 2009; revision received 1 October 2009; accepted 5 October 2009

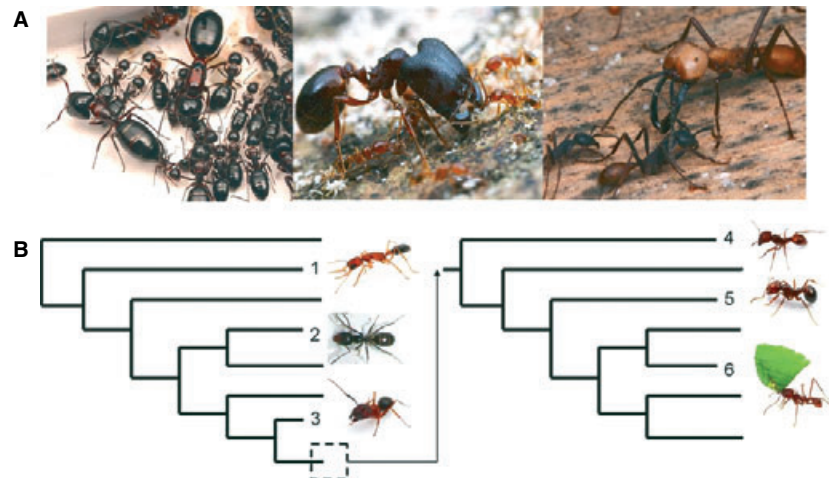
The genomics revolution has had a major impact on understanding the molecular basis of development and disease, but has focused on only a few model organisms and humans. There are approximately 1.6 million described plants, animals and fungi on Earth, but fewer

than 10 genomes of these species are well annotated and described. Reduced sequencing cost and improvement in comparative analysis now allow genomic tools to be developed for the many non-model species that impact fundamental aspects of the biosphere and our daily lives. For instance, the budding field of sociogenomics (Robinson *et al.* 2005) aims to explain complex interactions and emergent social behaviours by understanding processes from the bottom-up (from genes to societies). The molecular mechanisms underlying complex societies are only beginning to be understood, largely as a result of genome resources that were developed first for the honeybee (Honeybee Genome Sequencing Consortium 2007). Ants are another ecologically and functionally diverse group of species whose genomes promise to answer key biological questions ranging from behaviour to biomedicine to ecosystem health. The diversity in social organization, behavioural and phenotypic plasticity, morphology, communication and ecology in the ants is unrivaled and make ants one of several emerging model taxa (Smith *et al.* 2009). To develop a coordinated research agenda, a workshop on ant genomics was held on 16–17 January 2009 at the Center for Social Dynamics and Complexity at Arizona State University. More than 60 scientists from the leading ant research laboratories in both North America and Europe met to coordinate the financing, labour, sequencing and annotation of a phylogenetically representative (Fig. 1, Table 1) set of ant genomes within the next 3–5 years. A list of workshop participants and/or laboratories involved in ant genomics discussions can be found at the ant genomics portal (<http://www.antgenomics.org>).

Genomic information on ants promises to benefit diverse research fields well beyond those focusing on social insects. For example, large-scale transcriptome microarray (Wang *et al.* 2008) and candidate gene 'macroarray' (Goodisman *et al.* 2005; Graff *et al.* 2007) approaches with fewer markers have already shed new light on processes ranging from caste development to ageing. These studies have built on molecular studies of single genes and gene networks implicated in ant behaviour and caste differentiation (Tian *et al.* 2004), evolutionary development (Abouheif & Wray 2002), aggression (Lucas & Sokolowski 2009), reproduction (Khila & Abouheif 2008, Lu *et al.* 2009) and foraging behaviours (Ingram *et al.* 2005). Indeed, recent honeybee studies have shown that DNA methylation differences are associated with the developmental switch that

Correspondence: Christopher Smith;  
E-mail: smithcd@sfsu.edu

¶The first two authors contributed equally to this manuscript.



**Fig. 1** Ant forms and phylogeny. (A) Morphological diversity in the ants, from right to left, a queen (centre) and different sized workers of a carpenter ant; a minor and major worker of marauder ants (the major is 500× as massive as the minor); and a major and minor of army ants. (B) Ants funded for large-scale genomic or transcriptomic sequencing put into a phylogenetic framework (phylogeny modified from Brady *et al.* (2006)); the left tree is of the major ant lineages whereas the right tree is a more detailed version of the myrmicine branch which includes the most intensively studied taxa. The ants pictured are those funded, they are: 1 = *Harpegnathos saltator* (jumping ant), 2 = *Linepithema humile* (Argentine ant), 3 = *Camponotus floridanus* (carpenter ant), 4 = *Pogonomyrmex barbatus* (harvester ant), 5 = *Solenopsis invicta* (fire ant) and 6 = three species of attine fungus gardening ants from the genera *Atta*, *Acromyrmex* and *Apterostigma* (the fungi and other symbionts are also scheduled to be sequenced). All photos © Alex Wild, used with permission.

**Table 1** Summary of ant genome and transcriptome projects

Species	Project	Research group(s)	Sequencing technology	Funding source(s)	Progress
<i>Solenopsis invicta</i> (5)	EST Genome	Keller L Shoemaker D Fire Ant Genome Consortium†	Sanger Illumina 454	Multiple Sources‡	>10× genome coverage Annotated EST library
<i>Linepithema humile</i> (2)	Transcriptome Genome	Smith CD Tsutsui ND	454 Titanium (some paired-end)	Multiple Sources§	7× genome coverage Mixed life stage transcriptomes
<i>Harpegnathos saltator</i> (1)	Transcriptome EST	Berger S Liebig J	Illumina, Sanger	Howard Hughes Medical Institute	>90× genome coverage Annotated EST library
<i>Camponotus floridanus</i> (3)	Genome	Reinberg D			
<i>Pogonomyrmex barbatus</i> (4)	Transcriptome Genome	Gadau J Smith CR	454 Titanium (some paired-end) and SOLiD	National Science Foundation	>10× genome coverage
<i>Atta cephalotes</i> <i>Apterostigma dentigerum</i> 14 fungal & bacterial species (6)	Transcriptome Genome Meta-genome	Clifton S Currie C Gerardo N Slater S Suen G Taylor J Weinstock G	454 Paired-end Titanium	Roche Pharmaceuticals	Unknown

Numbers in parentheses correspond to numbers in Fig. 1.

†Members include M. Goodisman, J. Heinze, K. Ingram, K. McCubbin, P. Pamilo, C.-J. Shih, W.-J. Wu, S. Y.

‡AR and J Leenards Foundation, Swiss National Science Foundation, Rub Foundation, Agassiz Foundation, Herbette Foundation, Chuard-Schmid Foundation, University of Lausanne, USDA-CSREES-AFRI.

§National Institute of Mental Health, San Francisco State University, University of California Berkeley.

drives caste development (Elango *et al.* 2009). Below we explore a short list of the most exciting questions that could be immediately addressed with multiple ant genomes in hand.

### Why ants?

Ants are dominant components of all ecosystems, exceeding in biomass most other animal groups in tropical and temperate environments (Hölldobler & Wilson 1990). In addition to providing essential ecosystem services, ants are some of the most ecologically and economically costly invasive species. Ants are also an incredibly diverse group with 14 602 described species (<http://www.antweb.org>) representing diverse morphologies (Fig. 1), life histories and ecologies. Their evolutionary and ecological success is chiefly because of their social organization and complex division of labour. Thus, the development of genomic tools for ants will inform a diverse range of fields and major research areas (several are described below) ranging from medicine, to development, to behaviour and evolutionary ecology.

The predominant strength and weakness of ant genomics are one-in-the-same, diversity. While comparative genomics in ants can address many fundamental questions, developing genomic resources for many species is not only costly, but consensus over the most appropriate species is difficult to reach. With few funding agencies willing to back a single large project (10–20 genomes), a noncentralized strategy emerged with different groups independently seeking funding (Table 1). A common goal of sharing resources and facilitating research within the ant genomics community was agreed upon. Furthermore, meeting attendees agreed on the need for database and annotation standards similar to those adopted for the fruitfly (*Drosophila* 12 Genomes Consortium 2007), honeybee and wasp genome projects. As described below, progress is rapid and funding for several ant genomes and transcriptomes has already been acquired and is currently being put to use (Table 1).

### What genomic resources are becoming available for ants?

As a nongenetic model system, ants lack many of the basic genomic tools available for other model organisms. Two reviews published in 2008 (Goodisman *et al.* 2008; Smith *et al.* 2008) discussed a relative paucity of ant genomic tools and active research. While many resources are not yet publicly available, sequence data are rapidly being generated and the number of genomic studies on ants has multiplied several fold (Table 1) (Wurm *et al.* 2009).

Fortunately, most ants have moderately sized genomes (200–600 Mb), making it feasible to sequence many species at relatively high coverage (i.e. at least 6×) (Tsutsui *et al.* 2008). EST libraries and genome sequencing is underway for seven species (Table 1, Fig. 1), with sequences more tentatively planned for other species. It is expected that by 2011 these resources will be made public, assembled and will have preliminary computational gene annotations. Most of the new genome and transcriptome sequences will be generated using 'next generation' sequencing, especially the platforms available from Roche (454), Illumina (Solexa) and ABI (SOLiD). While the obvious advantage to this approach is enough raw sequence to represent a high percentage of the genome, such assemblies suffer from fragmentation into numerous short contigs that are difficult to map to full length chromosomes. Another issue is that techniques to merge data sets from different sequencing platforms into a single assembly are just being developed. Furthermore, as nongenetic model systems, most ants lack markers that would aid the physical mapping of contigs onto chromosomes. Nonetheless, the planned assemblies will yield sequence of high quality that is sufficient for gene finding, promoter analyses, SNP detection and development of microarray platforms. Improvements to the assembly of ant genomes will be facilitated by the sequencing of haploid males and the generation of linkage maps from SNPs of brothers.

### Current and potential insights from ant genomics and transcriptomics

#### *Integration of behavioural genomics, developmental genomics and sociocomplexity*

Like the honeybee, ants have a complex division of labour with many behavioural castes (e.g. nurse ants, guards, foragers). Unlike the honeybee, however, many ant lineages have independently evolved worker subcastes that differ remarkably in morphology (Fig. 1A). Despite nearly identical genome sequences, the worker castes in such polymorphic societies differ in adult morphology as much as highly diverged species. Studying the differentiation of behavioural and morphological castes at the transcriptome level will provide insights into the mechanisms of diverse types of plasticity, from developmental to neurological systems. Moreover, ants have a functional DNA methylation system that, like the honeybee (Elango *et al.* 2009), is probably involved in caste differentiation (Kronforst *et al.* 2008; Kucharski *et al.* 2008) making them exciting systems to explore how epigenetic modifications affect the evolution of developmental and social processes.

##### *Communication and systems biology*

The ecological dominance of ants is based on efficient communication (largely chemical) between colony members. Understanding the genomics underlying kin recognition and territoriality should help devise control strategies of the most-successful ant invaders, whose success often depends on greater tolerance of conspecifics. Ant genomics will provide unique insights into the mechanistic basis of information flow, behavioural response thresholds and systems optimization of many interacting individuals. Furthermore, ant genomics will provide insights into the evolution of chemosensory and neuronal genes regulating social processes operating at multiple levels of organization (e.g. individual and colony).

##### *Disease management in social environments*

As social organisms living in dense colonies of related individuals, ants provide an opportunity to study epidemiological processes within a designed interaction network. The interactions of individuals (castes, work chains, etc.) as well as the architecture of the interaction space are presumably under selection to maximize colony efficiency while also minimizing the spread of diseases. Both of these can be studied at phenotypic and genomic levels. Furthermore, many ants are engaged in complex symbiotic interactions. The genomics of these interactions will give novel insights into the co-evolutionary dynamics of disease resistance, including the potential discovery of novel antifungal or antimicrobial peptides. Planned genome, transcriptome and metagenome sequences of leaf-cutter ants and their symbionts promised to shed considerable light on these interactions (Table 1).

##### *Genetics of longevity, ageing and behaviour*

Ant queens are among the longest-lived insects known (up to 30 years), often exceeding the average age of their workers or males by one or more orders of magnitude (Hölldobler & Wilson 1990). In contrast to most other species, ants show a positive correlation between reproduction and longevity (Schrempf *et al.* 2005). The adaptations permitting this astounding prolongation of life are not understood, but genomic comparison of the insulin signalling pathway and free radical metabolism genes between independent origins of sociality (e.g. honeybees and wasps) will shed light on the genetics and evolution of ageing and fertility.

##### *Quantitative and reverse genetics*

Although laboratory maintenance of many species has been achieved, controlled laboratory crosses have only

been described for a few species. Genera such as *Cardiocondyla*, *Harpegnathos* and *Monomorium*, can be mated and propagated indefinitely in the laboratory, facilitating heritability and quantitative-genetics analyses through inbreeding, cross-breeding and backcrossing. Haplodiploid genetics will facilitate development of molecular maps and mapping polygenic traits, but will be limited until improved husbandry techniques are developed for most species. RNA interference has been successfully used to decrease gene expression in an ant (Lu *et al.* 2009) and multiple groups are working to develop transposable element (TE) or bacterial systems to overexpress or knock down gene targets. The creation of transgenic lines in other insect species via TEs (Pavlopoulos *et al.* 2004) suggests that these approaches should be well suited to ants and will help ameliorate the lack of traditional genetic tools.

##### *Comparative genomics*

Because of their ancient origin, well over 100 million years ago (Brady *et al.* 2006) and their remarkable subsequent diversification into more than 14 000 described species, genome-scale comparisons between ants offer a unique opportunity to observe what differences have driven the evolution of this hyperdiverse taxon. Comparison with the fruit fly, honeybee, pea aphid, flour beetle, silkworm moth and the jewel wasp genomes may aid in identifying DNA sequences associated with eusociality. Another outstanding question is whether the simple repeat and TE landscapes of ant genomes resemble eusocial honeybees (AT-rich sequence, few TEs) or solitary parasitoid wasps (AT-rich, diverse TEs). Such repetitive regions will further complicate genome assembly, but may also give insight into the life history of ants and suggest whether eusociality creates conditions that prune selfish DNA from the genome. Similarly, comparisons between judiciously chosen sister taxa may be particularly instructive, such as between an inquiline parasite (species existing exclusively as workerless queens that parasitize other ant species) and its unmodified sister lineage. Other ant species have lost the queen morph and form societies headed by reproducing workers. Such transitions to worker-less or queen-less ant species have evolved multiple times among ants, permitting identification of the common genomic footprints associated with these key changes in social organization.

##### **Looking forward**

Myrmecology has entered the genomic era, particularly comparative genomics, because of a long tradition of interpreting diversity within a comparative context. The scale of integrating genome knowledge across so many species is daunting and will require coordinated and stan-

standardized databases, new software tools, data-sharing policies and unprecedented cooperation. Undaunted, the ant genomics community has decided to follow the ancient wisdom, 'look to the ant thou sluggard and consider her ways and be wise' (Proverbs 6:6).

### Acknowledgements

The authors are grateful to all workshop participants and hope to have captured the direction of the workshop and the collective wisdom of its participants through the text of this document. On behalf of all participants, the authors thank the Center for Social Dynamics and Complexity as well as the School of Life Science at ASU for hosting and financially supporting the workshop. Christopher D. Smith was supported by NIMH grant 5SC2MH086071. The authors would also like to thank the two anonymous referees for their comments on this manuscript. We also thank Shelley Berger, Nicole Gerardo, Laurent Keller, Juergen Liebig, Danny Reinberg, DeWayne Shoemaker, and Neil Tsutsui for sharing unpublished results.

### References

- Abouheif E, Wray GA (2002) Evolution of the gene network underlying wing polyphenism in ants. *Science*, **297**, 249–252.
- Brady SG, Schultz TR, Fisher BL, Ward PS (2003) Evaluating alternative hypotheses for the early evolution and diversification of ants. *Proceedings of the National Academy of Science of the United States of America*, **103**, 18172–18177.
- Drosophila 12 Genomes Consortium (2007) Evolution of genes and genomes on the Drosophila phylogeny. *Nature* **450**, 203–218.
- Elango N, Hunt BG, Goodisman MAD, Yi SV (2009) DNA methylation is widespread and associated with differential gene expression in castes of the honeybee, *Apis mellifera*. *Proceedings of the National Academy of Science of the United States of America*, **106**, 11206–11211.
- Goodisman MAD, Isoe J, Wheeler DE, Wells MA (2005) Evolution of insect metamorphosis: a microarray-based study of larval and adult gene expression in the ant *Camponotus festinatus*. *Evolution*, **59**, 858–870.
- Goodisman MAD, Kovacs JL, Hunt BG (2008) Functional genetics and genomics in ants (Hymenoptera: Formicidae): the interplay of genes and social life. *Myrmecological News*, **11**, 107–117.
- Graff J, Jemielity S, Parker JD, Parker KM, Keller L (2007) Differential gene expression between adult queens and workers in the ant *Lasius niger*. *Molecular Ecology*, **16**, 675–683.
- Hölldobler B, Wilson EO (1990) *The Ants*. Belknap Press of Harvard University Press, Cambridge, MA.
- Honeybee Genome Sequencing Consortium (2007) Insights into social insects from the genome of the honeybee *Apis mellifera*. *Nature*, **443**, 931–949.
- Ingram KK, Oefner P, Gordon DM (2005) Task-specific expression of the foraging gene in harvester ants. *Molecular Ecology*, **14**, 813–818.
- Khila A, Abouheif E (2008) Reproductive constraint is a developmental mechanism that maintains social harmony in advanced ant societies. *Proceedings of the National Academy of Science of the United States of America*, **105**, 17884–17889.
- Kronforst MR, Gilley DC, Strassmann JE, Queller DC (2008) DNA methylation is widespread across social hymenoptera. *Current Biology*, **18**, R287–R288.
- Kucharski R, Maleszka J, Foret S, Maleszka R (2008) Nutritional control of reproductive status in honeybees via DNA methylation. *Science*, **319**, 1827–1830.
- Lu H, Vinson SB, Pietrantonio PV (2009) Oocyte membrane localization of vitellogenin receptor coincides with queen flying age, and receptor silencing by RNAi disrupts egg formation in fire ant virgin queens. *FEBS Journal*, **276**, 3110–3123.
- Lucas C, Sokolowski MB (2009) Molecular basis for changes in behavioral state in ant social behaviors. *Proceedings of the National Academy of Science of the United States of America*, **106**, 6351–6356.
- Pavlopoulos A, Berghammer AJ, Averof M, Klingler M (2004) Efficient transformation of the beetle *Tribolium castaneum* using the Minos transposable element: quantitative and qualitative analysis of genomic integration events. *Genetics*, **167**(2), 737–746.
- Robinson GE, Grozinger CM, Whitfield CW (2005) Sociogenomics: social life in molecular terms. *Nature Reviews Genetics*, **6**, 257–270.
- Schrempf A, Heinze J, Cremer S (2005) Sexual cooperation: mating increases longevity in ant queens. *Current Biology*, **15**, 267–270.
- Smith CR, Toth AL, Suarez AV, Robinson GE (2008) Genetic and genomic analyses of the division of labour in insect societies. *Nature Reviews Genetics*, **9**, 735–748.
- Smith CR, Dolezal A, Eliyahu D, Holbrook CT, Gadau J (2009) Ants (Formicidae): models for sociocomplexity. *Cold Spring Harbor Protocols*, **4**, doi: 10.1101/pdb.emo125
- Tian HS, Vinson SB, Coates CJ (2004) Differential gene expression between alate and dealate queens in the red imported fire ant, *Solenopsis invicta buren* (hymenoptera : Formicidae). *Insect Biochemistry and Molecular biology*, **34**, 937–949.
- Tsutsui ND, Suarez AV, Spagna JC, Johnston S (2008) The evolution of genome size in ants. *BMC Evolutionary Biology*, **8**, 64.
- Wang J, Ross KG, Keller L (2008) Genome-wide expression patterns and the genetic architecture of a fundamental social trait. *Proceedings of the Public Library of Science Genetics*, **4**, e1000127.
- Wurm Y, Uva P, Ricci F *et al.* (2009) Fourmidable: a database for ant genomics. *BMC Genomics*, **10**, 5.

---

The authors share a passion for the genetics of caste determination and social complexity, behavioural ecology, and evolutionary history of ants. They are also interested in the development of molecular and genomic tools to further the understanding of ants and their roles in Earth's ecosystem. C.D.S. has worked on several insect genome projects and is interested in using ants to study the genetic and epigenetic components of behavioural disorders. C.R.S. is interested in the proximate and ultimate factors underlying division of labour in social insects. U.M. is interested in the behaviour, evolution, ecology, microbiology, and natural history of social insects. J.G. is an evolutionary biologist interested in the genetic basis of species differences and speciation in Hymenoptera and the proximate and ultimate mechanisms of genetic variability in social insects.

---