**Genes involved in convergent evolution of eusociality in bees**

S. Hollis Woodard, Brielle J. Fischman, Aarti Venkat, Matt E. Hudson, Kranthi Varala, Sydney A. Cameron, Andrew G. Clark, and Gene E. Robinson

***Eusociality****:* A highly organized form of animal society. A species of animal is considered eusocial if its individuals live in groups that meet three criteria: 1. Reproductive division of labor; only a few members of society get to have offspring. 2. Cooperative care of offspring; members of the society help care for offspring that are not their own. 3. Multiple generations (for example, parents and offspring) live together.

***Evolution***: Change in inherited characteristics of populations over generations. Multiple factors, including natural selection, contribute to evolution.

***Convergent, divergent***: In convergent evolution, two species that are not closely related evolve to have similar traits; for example, both some birds and some butterflies use plant nectar for food. In divergent evolution, two species that are closely related evolve to be more different; for example, the shape of beaks in different species of finches in the Galapagos have become very different over time, as species adapt to different food sources. These terms can be used to describe molecular evolution, as well as evolution on the level of phenotypes.

 ***Accelerated rate of evolution***: A quicker accumulation of evolutionary changes over time, often detected on the molecular level, in one species relative to another. Accelerated evolution can indicate an increase in the influence of natural selection on the evolution of a species.

Proceedings of the National Academy of Sciences (2011). 18: 7472-7477.

**Abstract**

**Eusociality** has **evolved** at least 11 different times in insects. There are many types of eusocial lifestyles, ranging from species living in small colonies with open conflict over reproduction (**primitively eusocial**), to species in which colonies contain hundreds of thousands of highly specialized sterile workers produced by one or a few queens (high or **advanced eusociality**). Although the evolution of eusociality has been intensively studied, the genetic changes involved in the evolution of eusociality are relatively unknown. We examined patterns of genetic changes across three independent origins of eusociality. We did this by sequencing mRNA of nine socially diverse bee species, and comparing the **sequence** from each species with each other, and with **genome** sequence from the honey bee *Apis mellifera*. We found a group of 212 genes with changes in amino acid sequence indicating **accelerated** **evolution** across all types of eusociality studied. We also found unique groups of 173 and 218 genes with accelerated evolution specific to either highly or primitively eusocial lineages, respectively. These results demonstrate that **convergent evolution** can involve a complicated pattern of genetic changes in both shared and lineage-specific groups of genes. Genes involved in signal transduction, gland development, and carbohydrate metabolism are among the most notable rapidly evolving genes in eusocial lineages. These findings provide a starting point for linking specific genetic changes to the evolution of eusociality.

 ***Advanced eusocialit****y*: In species displaying advanced eusociality, social groups are larger; physical differences between individuals that reproduce and those that do not are more extreme; and the colonies can live through multiple years.

 ***Primitive eusociality***: In species displaying primitive eusociality, social groups are smaller; physical differences between individuals are smaller; and colonies are established each spring and end each fall. Offspring that will become reproductive hibernate in the winter.

 ***Solitary***: In solitary species, individuals may live near each other, but they do not rely on interactions with each other for survival or reproduction.

***Genome***: All the inherited information of a particular species—in organisms other than viruses, most of this information is physically encoded in DNA.

***Molecular evolution***: Evolution viewed on the level of changes in DNA, RNA, or protein sequence.

***Sequence***: The order of nucleotide bases (in DNA or RNA) or amino acids (in protein) for a given molecule. All three types of sequence (DNA, RNA and amino acid) are related to each other, and determine the shape a molecule of RNA or protein will be, the interactions it will have with other proteins or molecules and the functions it can perform.

 ***bp***: This stands for base pair, and refers to a pair of hydrogen-bonded nucleotide bases in a DNA molecule. Base pairs are used as a unit to describe how long a piece of DNA or RNA is.

***kbp***: This stands for kilo base pairs, or 1,000 base pairs. Bacterial genes are often around 1 kbp long, while human genes are often 20-50 kbp long.

***Gbp***: This stands for Giga base pairs, or 1,000,000,000 base pairs. The human genome is 3.2 Gbp long.

**Introduction**

Eusociality, the phenomenon in which some female offspring give up their own chance to reproduce to care cooperatively for their siblings, is an unusual and extreme class of social behavior (1). This type of social behavior has evolved multiple times, but only in a small number of lineages, primarily in the insects (11 or more times; ref. 2). The evolution of eusociality has long fascinated biologists because it requires that individuals help others at a reproductive cost to themselves, even though natural selection favors individuals with greater reproductive success (3). Although eusociality has been studied for many years, and theories for its evolution have been carefully developed and debated (4, 5), relatively little is known about genetic changes associated with eusocial evolution (6). Changes in DNA sequence, and where they occur in the genome, have the potential to inform us about the evolutionary processes involved in the evolution of eusociality (7). Comparing genetic changes that occurred during the evolution of eusociality in different species can help us understand whether independent occurrences of eusociality involved similar or different genetic changes.

We explored the genetic basis of eusocial evolution in bees. Several characteristics make bees a good choice for this study. There is a wide diversity of social lifestyles among bee species, from **solitary** to primitive eusociality to advanced eusociality (8). Additionally, eusociality has evolved independently at least six times (9-12) in the bees, more than in any other group. These features make it possible to compare multiple, independent origins of different social lifestyles among relatively closely related species. Furthermore, the extensive knowledge of bee natural history (8, 13, 14) makes it easier to develop hypotheses about how genetic changes detected in eusocial bees influence their behavior.

To study patterns of **molecular evolution** associated with eusociality in bees, we extracted mRNA from a set of nine bee species, and sequenced ~1 Gbp of short stretches of this mRNA. The set of bee species we chose reflects the remarkable social diversity in bees by including eusocial and non-eusocial species. Among these bee species there have been three origins of eusociality (9, 10). The eusocial bee species we studied exhibit two different forms of eusocial lifestyle, “highly eusocial” and “primitively eusocial” (ref. 8; Fig. 1). We compared the short stretches of mRNA sequence with the genome sequence from the highly eusocial honey bee *Apis mellifera* (15). Using these comparisons, we assembled the short stretches into partial or nearly-complete gene sequences, or **gene alignments**, for many genes in the nine other species. We searched among the alignments for genes with more mutations resulting in amino acid substitution in eusocial relative to non-eusocial lineages. Faster (“accelerated”) rates of protein changes (due to amino acid substitution) can indicate positive natural selection that has spread specific changes in protein function throughout populations (16). If genes contributing to certain traits show accelerated rates of protein changes in several **lineages**, this can suggest an association between changes in the genome and the evolution of those traits.

***Gene or sequence alignment:*** a method of lining up the nucleotide sequences of DNA or RNA, or the amino acid sequence of protein, to find parts of the sequence that are similar: for example, comparing the sequence of the same gene in multiple species to discover how similar they are. The authors have used the word “alignments” here to refer to the regions of the sequenced DNA that were found to be very similar across several bee species.

***Mutation*:** An accidental change in DNA sequence.

 ***Synonymous mutation*:** A change in the DNA that results in a new sequence of nucleotides in the DNA and RNA, but not in a change in the amino acid sequence of the encoded protein. This can happen because more than one RNA base triplet (codon) encodes for one amino acid: for example, AGU and AGC both code for serine, so a mutation that changes AGU to AGC would be synonymous.

 ***Non-synonymous mutation*:** A change in the DNA that results in a new sequence of nucleotides in the DNA and RNA, and a corresponding change in the amino acid sequence of the encoded protein.

***Phylogenetic tree*:** A diagram that shows evolutionary relationships between related species or taxa (group of related species). Taxa that branch directly off from each other are more closely related, while taxa separated by more branch points are more distant. In some phylogenetic trees, the length of each branch is used to represent how much time has gone by since taxa became distinct from each other.

 ***Lineage*:** A lineage is one branch of a phylogenetic tree, representing a population that has never split into different taxa, even though it may have still have changed over time.



Figure 1. Bee species used to identify genes evolving rapidly in eusocial lineages. This phylogenetic tree of species used in this study based on previously published trees (9–11) and the sequence in this study. Diamonds represent independent origins of eusociality.

1. What was the researchers’ main goal in this study?

2. Describe the lifestyle differences you would expect between *Apis mellifera*, *Bombus impatiens*, and *Megachile rotundata*.

**Methods**

*Bee Collection and Sequencing.* Bees from each species were caught while foraging, or directly from their nests. To make sure that the RNA in the tissues was preserved, each bee was rapidly frozen immediately after its collection, and stored frozen until use. To make sure that as many genes as possible were represented in the RNA obtained, bees of several ages, behavioral groups and castes were collected for each species. RNA was extracted (removed and purified) from the brains and bodies of the bees, and the RNA of bees from each species was pooled together. The RNA for each species was then broken into small fragments, and the sequence of each fragment was determined.

***EST*:** This stands for “expressed sequence tag.” Expressed sequence tags are short (much shorter than a full gene) sections of sequence that are obtained from RNA converted into DNA, rather than from genome DNA. This means that they most likely to be part of a gene, not part of non-expressed DNA sequence.

***Orthologous*:** If two genes found in two related species that both evolved from a single gene present in their most recent common ancestor, those two genes are orthologous to each other; each gene is the other gene’s ortholog. Orthologous genes will have their own unique sequence changes in each species, but overall their sequences will be similar because they originated from the same gene. Because they have similar sequences, the proteins they code for are expected to have similar functions.

*EST and Alignment Assembly.* EST reads (sequences of RNA fragments) were compared to the sequence of the honey bee genome. Using the honey bee genome as a reference, a software program looked for overlaps in the fragments and assembled them into longer pieces of RNA sequence. After the sequence fragments were connected as much as possible, probable genes were identified by comparing sequences with known sequences of individual honey bee genes to identify **orthologs**. *Apis mellifera* gene sequences can be found at BeeBase (Official Honey Bee Gene Set; www.beebase.org).

*Phylogeny.* Sequences that could be identified as genes or gene fragments were aligned using a software package called MrBayes. That is, while accounting for mutations that had changed individual base pairs in some species but not others, gene sequences were lined up so that corresponding base pairs were matched with each other. By examining the different sequence changes that occurred in different species, the software then calculated the most likely evolutionary relationships between the different species. These relationships were displayed as a phylogenetic tree. Because many genes were used to produce this result, it is highly likely that our phylogenetic tree accurately reflects the evolutionary history of these species.

*Evolutionary Tests.* To address the central question of our study, we wanted to find genes that genes that might be contributing in some way to the evolution of social behavior. To do this, we used a software package (called PAML) that examined the sequence changes that happened in each gene, in each species. This is similar to the phylogeny analysis described above, but with two important differences. First, instead of using the data from all the genes together, each gene was examined individually. Second, the software used in this analysis was comparing the rates of synonymous mutations and non-synonymous mutations. If a gene has a higher ratio of non-synonymous to synonymous mutation rates in certain lineages, it is likely to be undergoing positive selection in those lineages. We generated lists of genes that undergoing positive selection in eusocial species compared with solitary species, reasoning that these genes are more likely to be contributing in some way to biological processes that promote eusociality.

*Gene Ontology Enrichment Analysis.* To quantify and describe what types of genes we found to be potentially related to the evolution of eusociality, we performed a **Gene Ontology** (GO) enrichment analysis. To do this, we used a previously created list that matches honey bee genes to *Drosophila melanogaster* (fruit fly) **orthologs** (15). Orthologous fly genes with some known information about their functions were available for most (n = 3,451) genes in our dataset. We then used software called DAVID to group genes from our study according to their possible function, and determine whether any of these groups were larger than would have been expected to occur by chance (40).

***P-value:*** The p-value is a way of reporting the results of a statistical test. The p-value gives the likelihood that the result observed in an experiment would have occurred by chance alone, rather than being explained by the factor being tested in the experiment. A p-value of 0.07, for example, means that if the experimental factor being tested has no affect at all, there is a 7% chance this or an even more extreme result would be obtained.

 *p<0.05*: This is a criterion many scientists use as evidence that their experimental hypothesis is likely to be correct. It means that there is at most a 5% chance that their results are explained by chance alone.

**Results**

**Characterization of Alignments**. The RNA we sequenced from multiple bee species corresponded to ~33% of the genes (3,638 out of 11,062) in the *A. mellifera* Official Gene Set.

**Phylogenetic Tree Inference from EST Data**. We used our set of 3,638 alignments to estimate the phylogenetic relationships among bee species. The phylogenetic tree we generated was identical in structure to trees in other studies (9–11).

**Heterogeneous Patterns of Molecular Evolution Among Bee Lineages**. We searched among our sequences for genes with more frequent non-synonymous mutations (resulting in amino acid sequence changes) in eusocial relative to noneusocial lineages. We performed two tests that compare models of **neutral and non-neutral sequence evolution**. These tests search for genes in which the ratio of non-synonymous to synonymous nucleotide substitutions is higher in certain species. Test 1 identified genes in which the ratio is higher in all eusocial lineages as a group, compared with non-eusocial lineages. Test 2 identified genes in which the ratio is highest in either all primitively or all highly eusocial lineages as a group, relative to all other lineages. These tests can overlap. That is, a gene may be evolving more rapidly in all eusocial relative to non-eusocial lineages, as well as evolving most rapidly in either the highly or primitively eusocial lineages.

Our tests of protein evolution discovered a number of genes evolving differently between eusocial and non-eusocial lineages, and between different eusocial lineages. For test 1, we found 212 out of 3,638 genes (6%) evolving significantly more rapidly in all eusocial lineages relative to non-eusocial lineages (“All Eusocial” gene list). For test 2, we found 173 genes (5%) evolving most rapidly in highly eusocial lineages (“Highly Eusocial” gene list) and 218 genes (6%) in primitively eusocial lineages (“Primitively Eusocial” gene list), relative to other lineages (P < 0.05 in all three cases). Table 1 shows the most significant genes (based on **P value**) on each list. These results show that changes in some genes may support all types of eusocial behavior, while changes in other genes may support specific types of eusociality.

***Signal transduction*:** The molecular process that communicates a stimulus received by the cell through a receptor in the surface membrane (ie detecting the presence of a hormone or a neurotransmitter, heat or cold, too much or too little sugar in the blood, etc.) to components inside the cell that create a physiological response (ie gene expression, change in cell structure, etc.) Genes in this category could encode proteins that form the receptors, or the molecules inside the cell that respond to the activated receptors.

***Gland development*:** Formation of an organ that creates and excretes a biological substance, either into the body (for example, a hormone) or outside the body (for example, sweat). Genes in this category encode proteins that help glands form properly during growth.

***Carbohydrate metabolism*:** Formation or breakdown of carbohydrates (such as sugars) in the body. Genes in this category include those that break down sugar for energy, or change one type of sugar to another for storage or use.

**Biological Processes Evolving More Rapidly in Eusocial Relative to Non-Eusocial Bees**. It is hard for biologists to look at a list of genes that are evolving, and figure out what those changes are actually doing inside the cell. One way to begin to figure this out is by grouping genes according to what the function of each genes’s protein product (for example, does it help digest sugar? Is it part of the cell membrane?). These groups are called functional categories. A list of genes produced in an experiment can then be tested, to see if some functional categories have more genes from the list in them, than would likely happen by chance. We performed this test on our All Eusocial, Highly Eusocial, and Primitively Eusocial gene lists. “**Gland development**” and “cell surface receptor-linked **signal transduction**” were among the functions overrepresented only in the All Eusocial gene list (P < 0.05).

**Carbohydrate metabolism**-related categories were enriched in both the All Eusocial and Highly Eusocial gene lists, suggesting that these genes with this function are evolving both more rapidly in eusocial relative to non-eusocial lineages and most rapidly of all in highly eusocial lineages (Fig. 2). Fifteen of the 26 genes encoding glycolytic enzymes (proteins involved in carbohydrate metabolism) showed evidence of accelerated evolution in one or both of these lists Two of the most rapidly evolving genes on the Highly Eusocial gene list encode glycolytic enzymes.



Figure 2. Biological functions with evidence of accelerated evolution in eusocial lineages. Terms listed in each oval represent categories of genes evolving more rapidly in those lineages.

Transcription-related categories were enriched in both the All Eusocial and Primitively Eusocial gene lists, but not in the Highly Eusocial gene list. This enrichment exclusively in the All Eusocial and Primitively Eusocial gene lists suggests a similar pattern to that seen with carbohydrate-metabolism related genes in the All Eusocial and Highly Eusocial gene lists, only here with an emphasis in primitively eusocial lineages.

**Lifestyle- and Lineage-Specific Patterns of Molecular Evolution**. Some biological processes were enriched exclusively in either the Highly Eusocial or Primitively Eusocial gene lists and were not enriched in the All Eusocial gene list. For example, we detected evidence of accelerated evolution in brain-related functional categories in primitively eusocial bees, but not in highly eusocial bees.

3. What are three categories of genes that were found to be targets of selection during eusocial bee evolution?

**Discussion**

We identified several hundred genes showing signs of accelerated evolution. This includes some with signs of accelerated evolution in all eusocial lineages in our study, and some that were specific to a certain type of eusocial lifestyle. These results demonstrate that convergent evolution, in this case evolution of eusocial behavior, can involve several types of molecular changes. Genes involved in gland development, signal transduction, and carbohydrate metabolism were among the most rapidly evolving genes identified in this study. These findings provide a starting point for linking specific genetic changes to the evolution of eusociality in bees, which will be an important challenge for the future.

Several major steps will be needed to achieve the goal of discovering how specific changes in DNA sequence led to the evolution of eusociality. First, we need to improve our ability to predict how changes in amino acid sequence affect protein function. We also need to know how changes in protein function affect a particular biological process. Finally, we must discover how evolutionary changes in a particular biological process might affect traits associated with eusociality (20). For the present, we provide some speculation for how changes in the biological processes highlighted in our findings might affect social behavior.

***Pheromone:*** A chemical produced by an organism and excreted from the body, to communicate to other individuals of the same species. The particular odor of a pheromone will have a specific meaning and encourage certain behaviors on the part of the organisms that detect it.

Genes associated with gland development appear to have been a strong target of selection during eusocial bee evolution. This is not surprising: communication through release of chemicals and odors (pheromones) is an important way that social insects influence the behavior and physiology of colony members. Compared to solitary insects, eusocial insects have remarkably diverse gland functions and produce many unique gland secretions, including pheromones, food for young insects, and antimicrobial compounds (8, 13, 14). It is possible that at least some of the protein-coding sequence changes identified here are related to the evolution of complicated systems of chemical communication found in social bees.

Another category that appears to have been a strong target of selection during eusocial bee evolution is genes involved in signal transduction. A previously studied gene, *Amfor,* which is involved in food-searching behavior in both solitary and social insects, also belongs in this functional category (21). Several genes on the All Eusocial gene list in this category have known roles in behavior and neuronal function (22). Our results provide further evidence that signal transduction may be a general target of selection during behavioral evolution.

Genes associated with carbohydrate metabolism appear to have been a particularly strong target of selection during eusocial bee evolution. Our finding of a shared pattern of accelerated evolution across all eusocial lineages in our study may reflect the fact that many eusocial bees rely more heavily on honeys with high sugar concentration in their diet than do non-eusocial species (8), although all bees use nectar as their carbohydrate source. In addition, several characteristics shared by all eusocial insects, including worker–queen caste determination and worker–worker division of labor, are influenced by nutrition (6). Our results are consistent with these findings and also suggest that sequence changes in carbohydrate metabolism-related genes may have been involved in the evolution of these novel eusocial traits in bees.

Additional changes in carbohydrate metabolism-related genes were also detected in the highly eusocial bee lineages, but not in primitively eusocial bee lineages. This result may be due to the evolution of unique metabolic demands in the highly eusocial lifestyle, such as year-round nest thermoregulation (8), extended lifespan in queens (10-fold longer than workers; refs. 13, 27), and greatly increased foraging activity (14).

We were initially surprised to detect a signature of accelerated evolution in brain-related functional categories in primitively eusocial bee lineages, but not in highly eusocial bee lineages. Scientists who study human and primate brains have hypothesized that social life requires more learning, memory and reasoning abilities, and therefore is a strong selective force in brain evolution (28). This hypothesis tries to explain why humans have larger brains relative to their body size than other primates. It might be assumed that these demands are greater in the larger and more complex colonies of the highly eusocial bees, and we might have expected to find evidence of rapid evolution in brain-related genes in highly eusocial relative to primitively eusocial lineages (29). However, perhaps it is the primitively eusocial society members that face greater requirements for complex learning and memory abilities. In primitively but not advanced eusocial species, the social role of each bee changes frequently, and the balance between cooperation and competition is more dynamic (8, 13, 29). This finding of accelerated evolution in brain-related genes exclusively in primitively eusocial bees might eventually help us understand more about the evolution of behavioral differences that exist between primitively and highly eusocial species.

Our finding of shared sets of rapidly evolving genes across three independent lineages that gave rise to eusociality in bees suggests that there might also be some common molecular roots for eusocial evolution, despite the incredible social diversity found among bees. Among the biological processes that appear to have been under selection across all eusocial lineages in our study, carbohydrate metabolism stands out. It has been suggested that there is, a set of highly conserved genes and molecular pathways that were co-opted for novel, social functions during eusocial evolution (26). Our results provide additional support for the possibility that genes related to carbohydrate metabolism are key components in these conserved genes (6, 26). The existence of conserved genes for eusociality can be tested further, because there are at least another eight independent gains of eusociality in the bees, ants, wasps, and termites (2). The insect societies provide rich material to explore how changes in DNA sequence are associated with the evolution of social life.

4. What reasoning did the researchers provide for the importance of each of the three categories in social behavior?

5. Why would brain-related gene categories show accelerated rates of evolution in primitively eusocial species?

**References**

1. Maynard Smith J, Szathmary E (1995) *The Major Transitions in Evolution* (Oxford Univ. Press, New York).

2. Wilson EO, Hölldobler B (2005) Eusociality: Origin and consequences. *Proc Natl Acad Sci* USA 205:13367–13371.

3. Darwin C (1859) *On the Origin of Species* (Murray, London).

4. Queller DC (1992) A general model for kin selection. *Evolution* 46:376–380.

5. Nowak MA, Tarnita CE, Wilson EO (2010) The evolution of eusociality. *Nature* 466:1057–1062.

6. Smith CR, Toth AL, Suarez AV, Robinson GE (2008) Genetic and genomic analyses of the division of labour in insect societies. *Nat Rev Genet* 9:735–748.

7. Wong A (2010) Testing the effects of mating system variation on rates of molecular evolution in primates. *Evolution* 64:2779–2785.

8. Michener CD (1974) *The Social Behavior of the Bees* (Harvard Univ. Press, Cambridge, MA).

9. Cameron SA, Mardulyn P (2001) Multiple molecular data sets suggest independent origins of highly eusocial behavior in bees (Hymenoptera:Apinae). *Syst Biol* 50:194–214.

10. Schwarz MP, Richards MH, Danforth BN (2007) Changing paradigms in insect social evolution: Insights from halictine and allodapine bees. *Annu Rev Entomol* 52:127–150.

11. Cardinal S, Straka J, Danforth BN (2010) Comprehensive phylogeny of apid bees reveals the evolutionary origins and antiquity of cleptoparasitism. *Proc Natl Acad Sci USA* 107:16207–16211.

12. Danforth BN, Sipes S, Fang J, Brady SG (2006) Recent and simultaneous origins of eusociality in halictid bees. *Proc Natl Acad Sci USA* 103:15118–15123.

13. Wilson EO (1971*) The Insect Societies* (Harvard Univ. Press, Cambridge, MA).

14. Roubik DW (1992*) Ecology and Natural History of Tropical Bees* (Cambridge Univ. Press, Cambridge, UK).

15. HBGSC (2006) Insights into social insects from the genome of the honeybee *Apis mellifera*. *Nature* 443:931–949.

16. Yang ZH, Bielawski JP (2000) Statistical methods for detecting molecular adaptation. *Trends Ecol Evol* 15:496–503.

17. Schneider A, et al. (2010) Estimates of positive Darwinian selection are inflated by errors in sequencing, annotation, and alignment. *Genome Biol Evol* 1:114–118.

18. Wong KM, Suchard MA, Huelsenbeck JP (2008) Alignment uncertainty and genomic analysis. *Science* 319:473–476.

19. Wong WSW, Yang Z, Goldman N, Nielsen R (2004) Accuracy and power of statistical methods for detecting adaptive evolution in protein coding sequences and for identifying positively selected sites. *Genetics* 168:1041–1051.

20. Dean AM, Thornton JW (2007) Mechanistic approaches to the study of evolution: the functional synthesis. *Nat Rev Genet* 8:675–688.

21. Ben-Shahar Y, Robichon A, Sokolowski MB, Robinson GE (2002) Influence of gene action across different time scales on behavior. *Science* 296:741–744.

22. Root CM, et al. (2008) A presynaptic gain control mechanism fine-tunes olfactory behavior. *Neuron* 59:311–321.

23. Ament SA, Corona M, Pollock HS, Robinson GE (2008) Insulin signaling is involved in the regulation of worker division of labor in honey bee colonies. *Proc Natl Acad Sci USA* 105:4226–4231.

24. Corona M, et al. (2007) Vitellogenin, juvenile hormone, insulin signaling, and queen honey bee longevity. *Proc Natl Acad Sci USA* 104:7128–7133.

25. Patel A, et al. (2007) The making of a queen: TOR pathway is a key player in diphenic caste development. *PLoS ONE* 2:e509.

26. Toth AL, Robinson GE (2009) Evo-devo and the evolution of social behavior: Brain gene expression analyses in social insects. *Cold Spring Harb Symp Quant Biol* 74:419–426.

27. Finkel T, Holbrook N (2000) Oxidants, oxidative stress and the biology of ageing. *Nature* 405:239–247.

28. Dunbar RIM (2009) The social brain hypothesis and its implications for social evolution. *Ann Hum Biol* 36:562–572.

29. Gronenberg W, Riveros AJ (2009) *Organization of Insect Societies: From Genome to Sociocomplexity*, eds Gadau J, Fewell J (Harvard Univ. Press, Cambridge, MA), pp 377–401.

30. Silva A, et al. (1998) CREB and memory. *Annu Rev Neurosci* 21:127–148.

31. Sokolowski MB (2010) Social interactions in “simple” model systems. *Neuron* 65:780–794.

32. Nozawa M, Suzuki Y, Nei M (2009) Reliabilities of identifying positive selection by the branch-site and the site-prediction methods. *Proc Natl Acad Sci USA* 106:6700–6705.

33. Suzuki Y, Nei M (2002) Simulation study of the reliability and robustness of the statistical methods for detecting positive selection at single amino acid sites. *Mol Biol Evol* 19:1865–1869.

34. Wolf JBW, Kunstner A, Nam K, Jakobsson M, Ellegren H (2009) Nonlinear dynamics of nonsynonymous (d(N)) and synonymous (d(S)) substitution rates affects inference of selection. *Genome Biol Evol* 1:308–319.

35. Lein ES, et al. (2007) Genome-wide atlas of gene expression in the adult mouse brain. *Nature* 445:168–176.

36. Katoh K, Misawa K, Kuma K, Miyata T (2002) MAFFT: A novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res* 30:3059–3066.

37. Drummond AJ, et al. (2010) Geneious v5.1. Available at http://www.geneious.com.

38. Ronquist F, Huelsenbeck JP (2001) MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* 19:1572–1574.

39. Yang ZH (2007) PAML 4: Phylogenetic analysis by maximum likelihood. *Mol Biol Evol* 24:1586–1591.

40. Huang DW, Sherman BT, Lempicki RA (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 4:44–57.

**Acknowledgements**

We thank G. Bloch, C. Chanchao, K. Kapheim, M. de Lama, K. Hartfelder, M. P. Schwarz, T. Pitts-Singer, and S. Wongsiri for bee samples; E. Hadley for help with figures; T. Newman for laboratory assistance; C. Rassmussen for field assistance and bee identification; T. O’Conner for sequence editing; B. Smith for bioinformatic assistance; J. Huelsenbeck, A. Meade, and L. Jermiin for substitution model discussion; T. D. Seeley for helpful discussion; members of the Clark laboratory for helpful suggestions on the analyses; S. A. Ament, A. L. Toth, J. B. Whitfield, the Social Insect Reading Group (UIUC), and members of the Robinson and Cameron laboratories for comments on the manuscript; and P. Bouffard, T. Harkins, T. Jarvie, R. Martone, and J. Osterberger (Roche/454) for expert technical support. This work was supported by 454 Life Sciences (Roche Diagnostics Corporation) via the Roche 1GB contest (G.E.R., S.A.C., A.G.C., M.E.H., and Saurabh Sinha), National Science Foundation (NSF) Grant DEB07-43154 (to G.E.R. and M.E.H.), an NSF Predoctoral Fellowship (to B.J.F.), and National Institutes of Health Training Grant PHS2T32DC006612 (to S.H.W.; A. Feng, Principal Investigator).