

Genomic signatures of evolutionary transitions from solitary to group living

Karen M. Kapheim,^{1,2,3*†} Hailin Pan,^{4*} Cai Li,^{4,5} Steven L. Salzberg,^{6,7} Daniela Puiu,⁷ Tanja Magoc,⁷ Hugh M. Robertson,^{1,2} Matthew E. Hudson,^{1,8} Aarti Venkat,^{1,8,9} Brielle J. Fischman,^{1,10,11} Alvaro Hernandez,¹² Mark Yandell,^{13,14} Daniel Ence,¹³ Carson Holt,^{13,14} George D. Yocum,¹⁵ William P. Kemp,¹⁵ Jordi Bosch,¹⁶ Robert M. Waterhouse,^{17,18,19,20} Evgeny M. Zdobnov,^{17,18} Eckart Stolle,^{21,22} F. Bernhard Kraus,^{21,23} Sophie Helbing,²¹ Robin F. A. Moritz,^{21,24} Karl M. Glastad,²⁵ Brendan G. Hunt,²⁶ Michael A. D. Goodisman,²⁵ Frank Hauser,²⁷ Cornelis J. P. Grimmelihuizen,²⁷ Daniel Guariz Pinheiro,^{28,29} Francis Morais Franco Nunes,³⁰ Michelle Prioli Miranda Soares,²⁸ Érica Donato Tanaka,³¹ Zilá Luz Paulino Simões,²⁸ Klaus Hartfelder,³² Jay D. Evans,³³ Seth M. Barribeau,³⁴ Reed M. Johnson,³⁵ Jonathan H. Massey,^{2,36} Bruce R. Southey,³⁷ Martin Hasselmann,³⁸ Daniel Hamacher,³⁸ Matthias Biewer,³⁸ Clement F. Kent,^{39,40} Amro Zayed,³⁹ Charles Blatti III,^{1,41} Saurabh Sinha,^{1,41} J. Spencer Johnston,⁴² Shawn J. Hanrahan,⁴² Sarah D. Kocher,⁴³ Jun Wang,^{4,44,45,46,47†} Gene E. Robinson,^{1,48†} Guojie Zhang^{4,49†}

¹Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA. ²Department of Entomology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA. ³Department of Biology, Utah State University, Logan, UT 84322, USA. ⁴China National GeneBank, BGI-Shenzhen, Shenzhen, 518083, China. ⁵Centre for GeoGenetics, Natural History Museum of Denmark, University of Copenhagen, Copenhagen, 1350, Denmark. ⁶Departments of Biomedical Engineering, Computer Science, and Biostatistics, Johns Hopkins University, Baltimore, MD 21218, USA. ⁷Center for Computational Biology, McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA. ⁸Department of Crop Sciences, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA. ⁹Department of Human Genetics, University of Chicago, Chicago, IL 60637, USA. ¹⁰Program in Ecology and Evolutionary Biology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA. ¹¹Department of Biology, Hobart and William Smith Colleges, Geneva, NY 14456, USA. ¹²Roy J. Carver Biotechnology Center, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA. ¹³Department of Human Genetics, Eccles Institute of Human Genetics, University of Utah, Salt Lake City, UT 84112, USA. ¹⁴USTAR Center for Genetic Discovery, University of Utah, Salt Lake City, UT 84112, USA. ¹⁵USDA-ARS Red River Valley Agricultural Research Center, Biosciences Research Laboratory, Fargo, ND 58102, USA. ¹⁶CREAF, Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain. ¹⁷Department of Genetic Medicine and Development, University of Geneva Medical School, 1211 Geneva, Switzerland. ¹⁸Swiss Institute of Bioinformatics, 1211 Geneva, Switzerland. ¹⁹Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. ²⁰The Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA. ²¹Institute of Biology, Department Zoology, Martin-Luther-University Halle-Wittenberg, Hoher Weg 4, D-06099 Halle (Saale), Germany. ²²Queen Mary University of London, School of Biological and Chemical Sciences Organismal Biology Research Group, London E1 4NS, UK. ²³Department of Laboratory Medicine, University Hospital Halle, Ernst Grube Str. 40, D-06120 Halle (Saale), Germany. ²⁴German Centre for Integrative Biodiversity Research (iDiv), Halle-Jena-Leipzig, 04103 Leipzig, Germany. ²⁵School of Biology, Georgia Institute of Technology, Atlanta, GA 30332, USA. ²⁶Department of Entomology, University of Georgia, Griffin, GA 30223, USA. ²⁷Center for Functional and Comparative Insect Genomics, Department of Biology, University of Copenhagen, Copenhagen, Denmark. ²⁸Departamento de Biologia, Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, Universidade de São Paulo, 14040-901 Ribeirão Preto, SP, Brazil. ²⁹Departamento de Tecnologia, Faculdade de Ciências Agrárias e Veterinárias, UNESP - Univ Estadual Paulista, 14884-900 Jaboticabal, SP, Brazil. ³⁰Departamento de Genética e Evolução, Centro de Ciências Biológicas e da Saúde, Universidade Federal de São Carlos, 13565-905 São Carlos, SP, Brazil. ³¹Departamento de Genética, Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, 14049-900 Ribeirão Preto, SP, Brazil. ³²Departamento de Biologia Celular e Molecular e Bioagentes Patogênicos, Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, 14049-900 Ribeirão Preto, SP, Brazil. ³³USDA-ARS Bee Research Lab, Beltsville, MD 20705 USA. ³⁴Department of Biology, East Carolina University, Greenville, NC 27858, USA. ³⁵Department of Entomology, Ohio Agricultural Research and Development Center, Ohio State University, Wooster, OH 44691, USA. ³⁶Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, MI 48109, USA. ³⁷Department of Animal Sciences, University of Illinois, Urbana, IL 61801, USA. ³⁸Department of Population Genomics, Institute of Animal Husbandry and Animal Breeding, University of Hohenheim, Germany. ³⁹Department of Biology, York University, Toronto, ON M3J 1P3, Canada. ⁴⁰Janelia Farm Research Campus, Howard Hughes Medical Institute, Ashburn, VA 20147, USA. ⁴¹Department of Computer Science, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA. ⁴²Department of Entomology, Texas A&M University, College Station, TX 77843, USA. ⁴³Department of Organismic and Evolutionary Biology, Museum of Comparative Zoology, Harvard University, Cambridge, MA 02138, USA. ⁴⁴Department of Biology, University of Copenhagen, 2200 Copenhagen, Denmark. ⁴⁵Princess Al Jawhara Center of Excellence in the Research of Hereditary Disorders, King Abdulaziz University, Jeddah 21589, Saudi Arabia. ⁴⁶Macau University of Science and Technology, Avenida Wai long, Taipa, Macau 999078, China. ⁴⁷Department of Medicine, University of Hong Kong, Hong Kong. ⁴⁸Center for Advanced Study Professor in Entomology and Neuroscience, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA. ⁴⁹Centre for Social Evolution, Department of Biology, Universitetsparken 15, University of Copenhagen, DK-2100 Copenhagen, Denmark..

*These authors contributed equally to this work.

†Corresponding author. E-mail: karen.kapheim@usu.edu (K.M.K.); wangj@genomics.org.cn (J.W.); generobi@illinois.edu (G.E.R.); zhanggj@genomics.org.cn (G.Z.)

The evolution of eusociality is one of the major transitions in evolution, but the underlying genomic changes are unknown. We compared the genomes of ten bee species that vary in social complexity, representing multiple independent transitions in social evolution, and report three major findings. First, many important genes show evidence of neutral evolution as a consequence of relaxed selection with increasing social complexity. Second, there is no single road map to eusociality; independent evolutionary transitions in sociality have independent genetic underpinnings. Third, though clearly independent in detail, these transitions do have similar general features, including an increase in constrained protein evolution accompanied by increases in the potential for gene regulation and decreases in diversity and abundance of transposable elements. Eusociality may arise through different mechanisms each time, but would likely always involve an increase in the complexity of gene networks.

The evolution of eusociality involves changes in the unit of natural selection, from the individual to a group (1). Bees evolved eusociality multiple times and are extremely socially diverse (2) (Fig. 1), but all pollinate angiosperms, including many crops essential to the human diet (3). Simple eusociality may be facultative or obligate, and both forms are characterized by small colonies with a reproductive queen and one or more workers that, due to social and nutritional cues, forego reproduction to cooperatively care for their siblings (2). Further evolutionary elaborations have led to complex eusociality, “superorganisms” with colonies of several thousand individuals, sophisticated modes of communication, and morphological specializations for division of labor (4).

Theory predicts that the evolution of simple eusociality involves increased regulatory flexibility of ancestral gene networks to create specialized reproductive and non-reproductive individuals, and the evolution of complex eusociality requires genetic novelty to coordinate emergent properties of group dynamics (5). To test these predictions, we analyzed five de novo and five publicly available draft genome sequences of ten bee species from three families, representing two independent origins of eusociality in Apidae and Halictidae and two independent elaborations of simple to complex eusociality in two apid tribes [Apini (honey bees) and Meliponini (stingless bees); Fig. 1]. The draft genomes were of comparable, high quality (supplementary online text).

We found that the transition from solitary to group life is associated with an increased capacity for gene regulation. We scanned the promoter regions of 5,865 single-copy orthologs among the ten species to calculate a motif score (representing the number and binding strength of experimentally characterized transcription factor binding sites [TFBSs]) for 188 *Drosophila melanogaster* TFs (6) with at least one ortholog in each of the ten bees, and correlated motif score with social complexity, using phylogenetically independent contrasts (7). Of 2,101 significantly correlated motif-gene pairs, 89% were positive and 11% negative, show-

ing that TFs tend to have increased capacity to regulate genes in eusocial species of bees, relative to solitary species (Fig. 2A, supplementary online text).

Further evidence for increased capacity for gene regulation throughout social evolution is a positive ranked correlation between social complexity and the number of genes predicted to be methylated (7) (Spearman’s $\rho = 0.76$, $p = 0.01$; phylogenetically corrected Spearman’s $\rho = 0.64$, $p = 0.06$; Fig. 2B; bioinformatics

predictions validated with bisulfite sequencing data for three invertebrate species; supplementary online text). DNA methylation affects gene expression in a variety of ways (8). Thus, this result suggests an expansion in regulatory capacity with increasingly sophisticated sociality.

The potential for increased regulatory capacity was further revealed at the protein coding level. Increased social complexity also is associated with rapid evolution of genes involved in coordinating gene regulation. A Bayesian phylogenetic covariance analysis (9) of 5,865 single-copy orthologs identified 162 genes with accelerated evolution in species with increased social complexity (7) (Additional Data table S3). These rapidly evolving genes were significantly enriched ($p < 0.05$) for Gene Ontology terms related to regulation of transcription, RNA splicing, ribosomal structure, and regulation of translation (supplementary online text and tables S11 and S12). Similar results have been reported for bee and ant species (10–13); our findings reveal the underlying causes. Approximately two-thirds of these genes are under stronger directional selection in species with increasingly complex eusociality, but we also detected non-adaptive evolution. One-third of the rapidly evolving genes are under relaxed purifying selection in species with complex eusociality, possibly due to reduced effective population sizes (14).

We also found an additional 109 genes, significantly enriched ($p < 0.05$) for functions related to protein transport and neurogenesis, which evolve slower with increased social complexity (supplementary online text, table S13, and Additional Data table S3). This includes orthologs of *derailed 2* and *frizzled*, which function as Wnt signaling receptors in *Drosophila* synaptogenesis (15), and *rigor mortis*, a nuclear receptor involved in hormone signaling (16). A similar pattern of reduced evolutionary rate has been described for genes expressed in human and honey bee brains, potentially due to increasing pleiotropic constraint in complex gene networks (17, 18). Constrained protein evolution of neural and endocrine related genes seems at odds with the evolution of complexity, but this constraint appears to be com-

pensated for, or perhaps driven by, increased capacity for gene regulation.

We next investigated whether these molecular evolution patterns involve similar sets of genes and *cis*-regulatory elements among the early (facultative and obligate simple eusociality) and advanced (complex eusociality) stages of independent social transitions. We identified lineage-specific differences in coding sequences and promoter regions of 1,526 “social genes” for which evolutionary rate (dN/dS) is faster or slower with increased social complexity in two independent origins and two independent elaborations of eusociality (7) (Fig. 1). Among these lineage-specific social genes, we found common patterns of *cis*-regulatory evolution: gains of TFBSs in the promoters of genes that evolve slower with increasing social complexity (Fig. 2C and supplementary online text). This suggests that a shared feature of both independent origins and elaborations of eusociality is increasingly constrained protein evolution with increasing potential for novel gene expression patterns. The TFs responsible for this pattern were different for each social transition, even though our analysis was limited to highly conserved TFs (Table 1). Several function in neurogenesis, neural plasticity, or are prominent regulators of endocrine-mediated brain gene expression in honey bees (19, 20).

We found further lineage-specific differences among the rapidly evolving “social genes” themselves. Genes undergoing accelerated evolution at the origins of eusociality were significantly enriched for GO terms related to signal transduction in both Apidae and Halictidae, but they shared only six genes (6 out of 354 and 167 genes, respectively; hypergeometric test, $p = 0.82$; Fig. 2D; Additional Data tables S5 and S6). Rapid evolution of signal transduction pathways may be a necessary step in all origins of eusociality to mediate intracellular responses to novel social and environmental stimuli (10), but selection appears to have targeted different parts of these pathways in each independent transition. Caste-specific expression and other analyses of these genes are needed to determine their function in eusociality.

Genes showing signatures of rapid evolution with the elaborations of complex eusociality were also highly disparate between honey bees and stingless bees, with only 43 shared genes and no shared enriched GO terms (43 out of 625 and 512 genes, respectively; hypergeometric test, $p = 0.70$; Fig. 2D; Additional Data tables S5 and S6). In addition, only 2 out of 5,865 single-copy orthologs showed a signature of convergent evolution by fitting a dendrogram based on social complexity significantly better than the accepted molecular phylogeny (7) (supplementary online text and fig. S21). Similarly, families of major royal jelly protein genes, sex determining genes, odorant receptors, and genes involved in lipid metabolism expanded in some, but not all, lineages of complex eusocial bees (7) (Table 2; supplementary online text). These results suggest gene family expansion is associated with complex eusociality as predicted (5),

but involves different genes in each case. Despite striking convergence of social traits among the superorganisms (4), the final stages of transformation to this level of biological organization do not necessarily involve common molecular pathways.

The major transitions in evolution involve a reduction in conflict as the level of natural selection rises from the individual to the group (1). Extending this to intragenomic conflict may explain our finding of decreased diversity and abundance of transposable elements (TEs) with increasing social complexity (7) (regression after phylogenetic correction, $F = 8.99$, adjusted $R^2 = 0.47$, $p = 0.017$; Fig. 2E and figs. S42-S44; supplementary online text). This may be a consequence of increased recombination rates among highly eusocial insects (21, 22) or because key features of complex eusociality lead to decreased exposure to parasites and pathogens that horizontally transmit TEs (4, 23). Eusociality in bees may thus provide natural immunity against certain types of intragenomic conflict.

Our results and those in refs. 10-13 support the prediction that changes in gene regulation are key features of evolutionary transitions in biological organization (5). Our results further reveal the convergent adaptive and non-adaptive evolutionary processes common to both the early and advanced stages of multiple independent transitions from solitary to group living. It is now clear that there are lineage-specific genetic changes associated with independent origins of eusociality in bees, and independent elaborations of eusociality in both bees and ants. This includes different sets of genes showing caste-biased expression across species (24–26) and, as we have shown, evolutionary modifications of TEs, gene methylation, and *cis*-regulatory patterns associated with the suite of life history traits that define eusociality. This suggests that if it were possible to “replay life’s tape” (27), eusociality may arise through different mechanisms each time, but would likely always involve an increase in the complexity of gene networks.

REFERENCES AND NOTES

1. J. Maynard Smith, E. Szathmáry, *The major transitions in evolution*. (Oxford University Press, Oxford, England, 1995).
2. C. D. Michener, *The social behavior of the bees*. (Harvard University Press, Cambridge, MA, 1974).
3. A.-M. Klein, B. E. Vaissière, J. H. Cane, I. Steffan-Dewenter, S. A. Cunningham, C. Kremen, T. Tscharntke, Importance of pollinators in changing landscapes for world crops. *Proc. Biol. Sci. B* **274**, 303–313 (2007). [Medline doi:10.1098/rspb.2006.3721](#)
4. H. Hölldobler, E. O. Wilson, *The superorganism: the beauty, elegance and strangeness of insect societies*. (Norton Press, New York, NY, 2009).
5. B. R. Johnson, T. A. Linksvayer, Deconstructing the superorganism: Social physiology, groundplans, and sociogenomics. *Q. Rev. Biol.* **85**, 57–79 (2010). [Medline doi:10.1086/650290](#)
6. L. J. Zhu, R. G. Christensen, M. Kazemian, C. J. Hull, M. S. Enameh, M. D. Basciotta, J. A. Brasefield, C. Zhu, Y. Asriyan, D. S. Lapointe, S. Sinha, S. A. Wolfe, M. H. Brodsky, FlyFactorSurvey: A database of *Drosophila* transcription factor binding specificities determined using the bacterial one-hybrid system. *Nucleic Acids Res.* **39** (Database), D111–D117 (2011). [Medline doi:10.1093/nar/gkq858](#)
7. Materials and methods are available as supplementary material online.
8. H. Yan, R. Bonasio, D. F. Simola, J. Liebig, S. L. Berger, D. Reinberg, DNA methylation in social insects: How epigenetics can control behavior and

- longevity. *Annu. Rev. Entomol.* **60**, 435–452 (2015). [Medline doi:10.1146/annurev-ento-010814-020803](#)
9. N. Lartillot, R. Poujol, A phylogenetic model for investigating correlated evolution of substitution rates and continuous phenotypic characters. *Mol. Biol. Evol.* **28**, 729–744 (2011). [Medline doi:10.1093/molbev/msq244](#)
 10. S. H. Woodard, B. J. Fischman, A. Venkat, M. E. Hudson, K. Varala, S. A. Cameron, A. G. Clark, G. E. Robinson, Genes involved in convergent evolution of eusociality in bees. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 7472–7477 (2011). [Medline doi:10.1073/pnas.1103457108](#)
 11. B. A. Harpur, C. F. Kent, D. Molodtsova, J. M. Lebon, A. S. Alqarni, A. A. Owayss, A. Zayed, Population genomics of the honey bee reveals strong signatures of positive selection on worker traits. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 2614–2619 (2014). [Medline doi:10.1073/pnas.1315506111](#)
 12. J. Roux, E. Privman, S. Moretti, J. T. Daub, M. Robinson-Rechavi, L. Keller, Patterns of positive selection in seven ant genomes. *Mol. Biol. Evol.* **31**, 1661–1685 (2014). [Medline doi:10.1093/molbev/msu141](#)
 13. D. F. Simola, L. Wissler, G. Donahue, R. M. Waterhouse, M. Helmkampf, J. Roux, S. Nygaard, K. M. Glastad, D. E. Hagen, L. Viljakainen, J. T. Reese, B. G. Hunt, D. Graur, E. Elhaik, E. V. Kriventseva, J. Wen, B. J. Parker, E. Cash, E. Privman, C. P. Childers, M. C. Muñoz-Torres, J. J. Boomsma, E. Bornberg-Bauer, C. R. Currie, C. G. Elsik, G. Suen, M. A. Goodisman, L. Keller, J. Liebig, A. Rawls, D. Reinberg, C. D. Smith, C. R. Smith, N. Tsutsui, Y. Wurm, E. M. Zdobnov, S. L. Berger, J. Gadau, Social insect genomes exhibit dramatic evolution in gene composition and regulation while preserving regulatory features linked to sociality. *Genome Res.* **23**, 1235–1247 (2013). [Medline doi:10.1101/gr.155408.113](#)
 14. J. Romiguier, J. Lourenco, P. Gayral, N. Faivre, L. A. Weinert, S. Ravel, M. Ballenghien, V. Cahais, A. Bernard, E. Loire, L. Keller, N. Galtier, Population genomics of eusocial insects: The costs of a vertebrate-like effective population size. *J. Evol. Biol.* **27**, 593–603 (2014). [doi:10.1111/jeb.12331](#)
 15. M. Park, K. Shen, WNTs in synapse formation and neuronal circuitry. *EMBO J.* **31**, 2697–2704 (2012). [Medline doi:10.1038/emboj.2012.145](#)
 16. J. Gates, G. Lam, J. A. Ortiz, R. Losson, C. S. Thummel, *rigor mortis* encodes a novel nuclear receptor interacting protein required for ecdysone signaling during *Drosophila* larval development. *Development* **131**, 25–36 (2004). [Medline doi:10.1242/dev.00920](#)
 17. D. Brawand, M. Soumillon, A. Necsulea, P. Julien, G. Csárdi, P. Harrigan, M. Weier, A. Liechti, A. Aximu-Petri, M. Kircher, F. W. Albert, U. Zeller, P. Khaitovich, F. Grützner, S. Bergmann, R. Nielsen, S. Pääbo, H. Kaessmann, The evolution of gene expression levels in mammalian organs. *Nature* **478**, 343–348 (2011). [Medline doi:10.1038/nature10532](#)
 18. D. Molodtsova, B. A. Harpur, C. F. Kent, K. Seevananthan, A. Zayed, Pleiotropy constrains the evolution of protein but not regulatory sequences in a transcription regulatory network influencing complex social behaviors. *Front. Genet.* **5**, 431 (2014). [Medline doi:10.3389/fgene.2014.00431](#)
 19. D. W. Pfaff, A. P. Arnold, A. M. Etgen, R. T. Rubin, S. E. Fahrbach, Eds., *Hormones, Brain and Behavior* (Elsevier, New York, 2009).
 20. S. Chandrasekaran, S. A. Ament, J. A. Eddy, S. L. Rodriguez-Zas, B. R. Schatz, N. D. Price, G. E. Robinson, Behavior-specific changes in transcriptional modules lead to distinct and predictable neurogenomic states. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 18020–18025 (2011). [doi:10.1073/pnas.1114093108](#)
 21. L. Wilfert, J. Gadau, P. Schmid-Hempel, Variation in genomic recombination rates among animal taxa and the case of social insects. *Heredity* **98**, 189–197 (2007). [Medline doi:10.1038/sj.hdy.6800950](#)
 22. E. S. Dolgin, B. Charlesworth, The effects of recombination rate on the distribution and abundance of transposable elements. *Genetics* **178**, 2169–2177 (2008). [Medline doi:10.1534/genetics.107.082743](#)
 23. S. Schaack, C. Gilbert, C. Feschotte, Promiscuous DNA: Horizontal transfer of transposable elements and why it matters for eukaryotic evolution. *Trends Ecol. Evol.* **25**, 537–546 (2010). [Medline doi:10.1016/j.tree.2010.06.001](#)
 24. B. Feldmeyer, D. Elsner, S. Foitzik, Gene expression patterns associated with caste and reproductive status in ants: Worker-specific genes are more derived than queen-specific ones. *Mol. Ecol.* **23**, 151–161 (2014). [Medline doi:10.1111/mec.12490](#)
 25. P. G. Ferreira, S. Patalano, R. Chauhan, R. French-Constant, T. Gabaldón, R. Guigó, S. Sumner, Transcriptome analyses of primitively eusocial wasps reveal novel insights into the evolution of sociality and the origin of alternative phenotypes. *Genome Biol.* **14**, R20 (2013). [Medline doi:10.1186/gb-2013-14-2-r20](#)
 26. B. G. Hunt, L. Ometto, Y. Wurm, D. Shoemaker, S. V. Yi, L. Keller, M. A. D. Goodisman, Relaxed selection is a precursor to the evolution of phenotypic plasticity. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 15936–15941 (2011). [Medline doi:10.1073/pnas.1104825108](#)
 27. S. J. Gould, *Wonderful Life: The Burgess Shale and the Nature of History* (Norton, New York, 1989).
 28. S. Cardinal, B. N. Danforth, Bees diversified in the age of eudicots. *Proc. Biol. Sci.* **280**, 20122686 (2013). [Medline doi:10.1098/rspb.2012.2686](#)
 29. S. Cardinal, B. N. Danforth, The antiquity and evolutionary history of social behavior in bees. *PLOS ONE* **6**, e21086 (2011). [Medline doi:10.1371/journal.pone.0021086](#)

ACKNOWLEDGMENTS

Data deposition at NCBI: *H. laboriosa*, *D. novaeangliae*, *E. mexicana*, *M. quadrifasciata*, and *M. rotundata* genome assemblies accession numbers PRJNA279436, PRJNA279825, PRJNA279814, PRJNA279820, PRJNA66515. Funding for genome sequencing and analysis: BGI, a U.S. National Institute of Health Pioneer Award (DP1 OD006416) to G.E.R., and EU Marie Curie International Incoming Fellowship (300837) to G.Z. Additional funding: U.S. National Science Foundation grant DEB-0640690 (M.A.D.G.) and DEB-0743154 (G.E.R. and M.E.H.); Danish Council for Independent Research grants 10-081390 (C.L.) and 0602-01170B (C.J.P.G.); Lundbeck Foundation (C.J.P.G.); Georgia Tech-Elizabeth Smithgall Watts endowment (M.A.D.G.); Marie Curie International Outgoing Fellowship PLOF-GA-2011-303312 (R.M.W.); Swiss National Science Foundation award 31003A-125350, Commission Informatique of the University of Geneva, Schmidheiny Foundation, and Swiss Institute of Bioinformatics (E.M.Z.); NSERC Discovery Grant and Early Research Award (A.Z.). This project was conducted under the auspices of the i5K Initiative. We thank the Roy J. Carver Biotechnology Center (sequencing services); N. Lartillot (advice on Coevol); T. Newman (assistance with DNA extractions); and E. Hadley (assistance with figures). Computational support: D. Davidson, N. Band, D. Slater (University of Illinois), J.H. Kidner and H. Scharpenberg (Martin-Luther-University Halle-Wittenberg); Compute Canada; and Center for High Performance Computing at the University of Utah. J. Himes created the illustrations in Fig. 1. We are grateful to T. Pitts-Singer, H. G. Hall, B. N. Danforth, J. Gibbs, and S. Cardinal for providing bee specimens, R. Ayala for identification of *E. mexicana*, J. Vega and the staff at Estación de Biología Chamela, UNAM for support during collecting trips, and S.A. Cameron for insightful discussion.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/cgi/content/full/science.aaa4788/DC1

Materials and Methods

Supplementary Text

Figs. S1 to S44

Tables S1 to S32

Additional Data tables S1 to S12

References (30–151)

Author Contributions

19 December 2014; accepted 6 May 2015

Published online 14 May 2015

10.1126/science.aaa4788

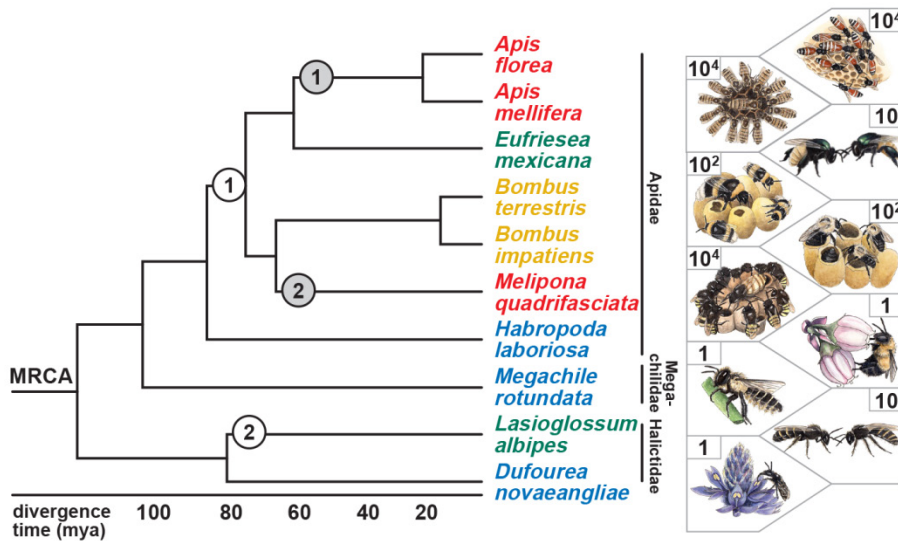


Fig. 1. Phylogeny and divergence times (28) of bees selected for genome analysis. We analyzed two independent origins of simple eusociality from a solitary ancestor, one each in Apidae (white circle 1) and Halictidae (white circle 2), and two independent elaborations of complex eusociality in honey bees (grey circle 1) and stingless bees (grey circle 2). Most bees mate once, but honey bees mate with multiple males. All bees eat pollen and nectar from flowering plants. Species names are colored according to degree of social complexity: blue: ancestrally solitary; green: facultative simple eusociality; orange: obligate simple eusociality; red: obligate complex eusociality. The social biology of *E. mexicana* is unknown, but is representative of the facultative simple eusocial life history (29). Numbers in each box are approximate colony size on a log scale.

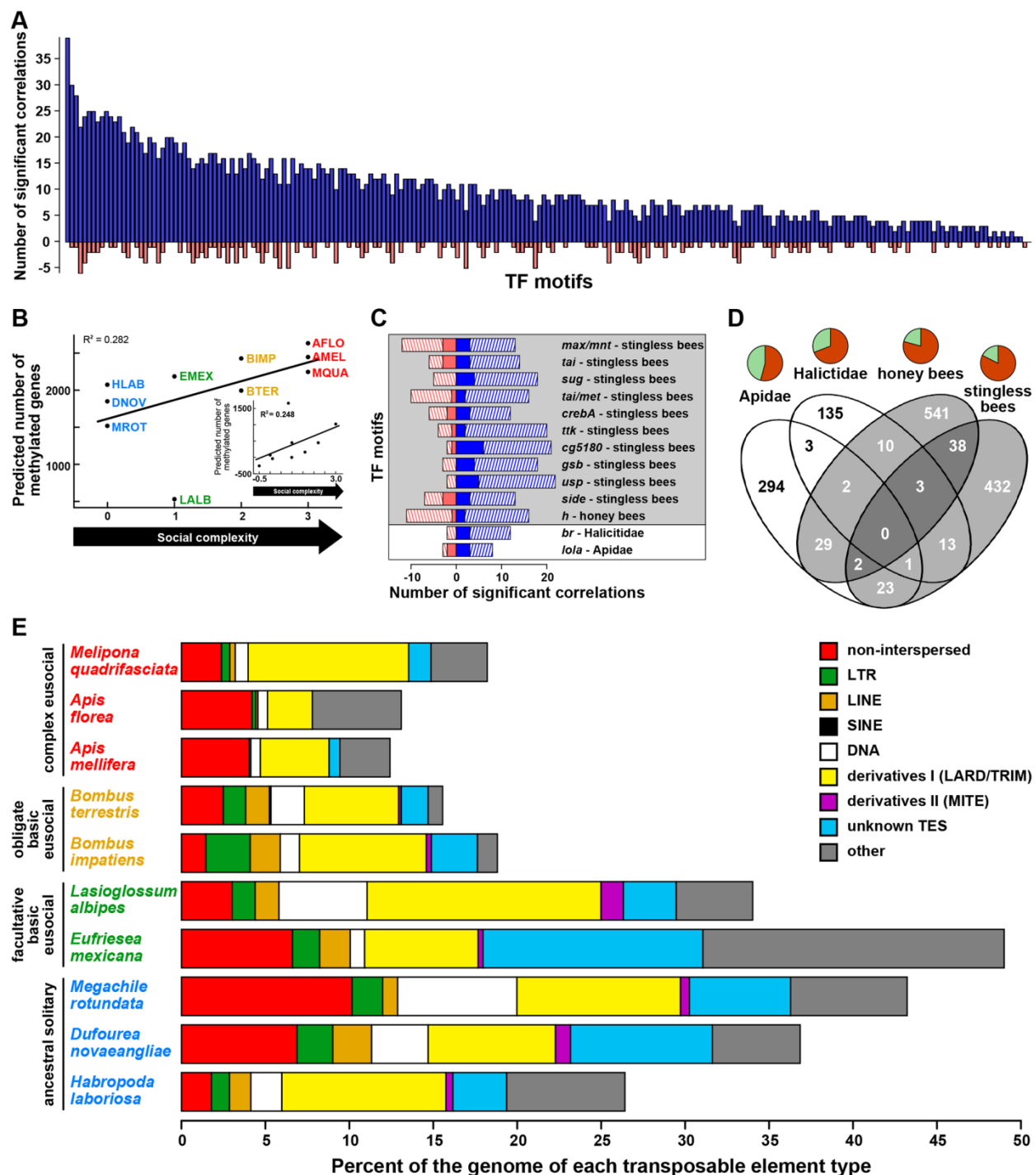


Fig. 2. Genomic signatures of evolutionary transitions from solitary to group life. (A) Increasing social complexity is associated with increasing presence of *cis*-regulatory TFBS in promoter regions. Each bar represents a TFBS for which presence correlates significantly with social complexity (blue: positive; red: negative). (B) Relationship between predicted number of methylated genes and social complexity before and after (inset) phylogenetic correction (see text for statistics). (C) TFBS motifs showing a relationship between social complexity and evolutionary rate of coding and non-coding sequences in different lineages. Bar length indicates the number of significant correlations (blue: positive; red: negative) between each motif score and social complexity (from Table 1) among genes evolving faster (solid) or slower (hatched) in lineages with different levels of social complexity (from 2D). Background shading follows circle shading in Fig. 1. (D) Number of genes for which evolutionary rate is faster or slower in lineages with higher compared to lower social complexity. Pie charts represent the proportion of genes evolving slower (light green) or faster (dark orange) with increased social complexity. Venn diagram shading follows circle shading in Fig. 1. (E) Complex eusocial species have a reduced proportion of repetitive DNA compared to other bees (see text for statistics).

Table 1. Transcription factors (TFs) and corresponding motifs associated with origins and elaborations of eusociality in bees. (Motif names: Fly Factor Survey (6); supplementary online text).

Motif	<i>D. melanogaster</i> TFs	Hypergeometric test p-value
<i>Solitary to simple eusociality – Apidae</i>		
lola_PQ_SOLEXA	<i>Lola</i>	0.0047
<i>Solitary to simple eusociality – Halictidae</i>		
br_PL_SOLEXA_5	<i>Br</i>	0.0016
<i>Simple eusociality to complex eusociality - honey bees</i>		
h_SOLEXA_5	<i>dpn,h</i>	0.0027
<i>Simple eusociality to complex eusociality – stingless bees</i>		
Side_SOLEXA_5	<i>E_spl, HLHm3, HLHm5, HLHm7, HLHmbeta, HLHmdelta, HLHm-gamma, Side</i>	0.0008
usp_SOLEXA	<i>EcR,svp,usp</i>	0.0013
CrebA_SOLEXA	<i>CrebA</i>	0.0040
CG5180_SOLEXA	<i>CG5180</i>	0.0044
tai_Met_SOLEXA_5	<i>Mio_bigmax,tai_Met</i>	0.0045
ttk_PA_SOLEXA_5	<i>Ttk</i>	0.0078
gsb_SOLEXA	<i>gsb,Poxn,prd</i>	0.0083
tai_SOLEXA_5	<i>Tai</i>	0.0100

Table 2. Relative size of select gene families as related to social complexity in bees.

Family	Function	Eusocial bees compared to solitary bees
<i>Differences among bees</i>		
Major royal jelly	Brood feeding	Expanded only in <i>Apis</i>
Sex determination pathway genes	Sex-specific development	Expanded in some eusocial lineages
Odorant receptors	Olfaction	Expanded in complex eusocial lineages
Lipid metabolism genes	Metabolic processing of lipids	Expanded in complex eusocial lineages
<i>Similarities across bees</i>		
Biogenic amines receptors, neuropeptides, GPCRs	Neural plasticity	Similar
Insulin-signaling and ecdysone pathway genes	Insect development, caste determination in honey bees, behavioral plasticity as adults	Similar
Immunity	Infectious disease protection	Similar
Cytochrome P450 monooxygenase genes	Detoxification	Similar