# Thrice Out of Africa: Ancient and Recent Expansions of the Honey Bee, *Apis mellifera*

Charles W. Whitfield,[1,2,3]* Susanta K. Behura,[1] Stewart H. Berlocher,[1,3] Andrew G. Clark,[4] J. Spencer Johnston,[5] Walter S. Sheppard,[6,7] Deborah R. Smith,[8] Andrew V. Suarez,[1,3,9] Daniel Weaver,[10] Neil D. Tsutsui[11,12]

We characterized *Apis mellifera* in both native and introduced ranges using 1136 single-nucleotide polymorphisms genotyped in 341 individuals. Our results indicate that *A. mellifera* originated in Africa and expanded into Eurasia at least twice, resulting in populations in eastern and western Europe that are geographically close but genetically distant. A third expansion in the New World has involved the near-replacement of previously introduced "European" honey bees by descendants of more recently introduced *A. m. scutellata* ("African" or "killer" bees). Our analyses of spatial transects and temporal series in the New World revealed differential replacement of alleles derived from eastern versus western Europe, with admixture evident in all individuals.

The long-standing association between humans and honey bees, *Apis mellifera*, is evidenced by 7000-year-old cave paintings depicting honey collection from wild bee nests (*1*). Managed honey bee colonies satisfy the pollination requirements of both modern agriculture and the demand for products such as honey, wax, and royal jelly (*2*). With the publication of the honey bee genome sequence (*3*) and development of high-density single-nucleotide polymorphism (SNP) markers, tools are now available to study evolutionary processes occurring in both native and introduced populations, including Africanization in the New World, and the genomewide consequences of ancient and recent evolution in this important social insect.

The genus *Apis* comprises 10 species, 9 of which are confined to Asia (*4*). *A. mellifera*, however, is distributed from sub-Saharan Africa to central Asia and northern Europe and is composed of more than two dozen morphologically and geographically distinct subspecies (*5*). It is commonly believed that *A. mellifera* split from its closest relative, *A. cerana*, in western or central Asia (where these species' ranges are closest) and subsequently expanded into Europe and Africa (*5, 6*). Existing populations in Europe and Africa are hypothesized to derive from populations currently found in Asia, where *A. mellifera* occurs as far east as Kazakhstan (*6*). However, an apparent discrepancy between the age of divergence among *A. mellifera* subspecies [0.7 to 1.3 million years (*4, 5, 7*)] and the split 6 to 8 million years ago (*4, 7, 8*) between *A. mellifera* and *A. cerana* suggests alternative scenarios. One hypothesis [suggested by E. O. Wilson, citing a personal correspondence from C. D. Michener (*9*)] is that *A. mellifera* originated in the tropics or subtropics of Africa.

[1]Department of Entomology, University of Illinois at Urbana-Champaign, 505 South Goodwin Avenue, IL 61801, USA. [2]Institute for Genomic Biology, University of Illinois at Urbana-Champaign, IL 61801, USA. [3]Program for Ecology and Evolutionary Biology, University of Illinois at Urbana-Champaign, IL 61801, USA. [4]Department of Molecular Biology and Genetics, Cornell University, Ithaca, NY 14853, USA. [5]Department of Entomology, Texas A & M University, College Station, TX 77843, USA. [6]Department of Entomology, Washington State University, Pullman, WA 99164, USA. [7]Center for Reproductive Biology, Washington State University, Pullman, WA 99164, USA. [8]Entomology program, Department of Ecology and Evolutionary Biology, University of Kansas, Lawrence, KS 66045, USA. [9]Department of Animal Biology, University of Illinois at Urbana-Champaign, IL 61801, USA. [10]Bee Power, LP, and Bee Weaver Apiaries, Inc., Navasota, TX 77868, USA. [11]Department of Ecology and Evolutionary Biology, University of California, Irvine, CA 92697, USA. [12]Institute for Genomics and Bioinformatics, University of California, Irvine, CA 92697, USA.

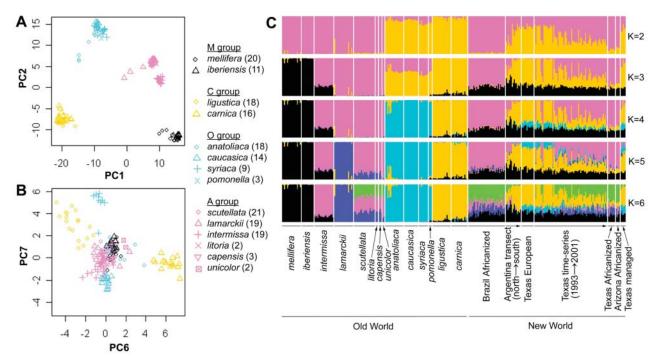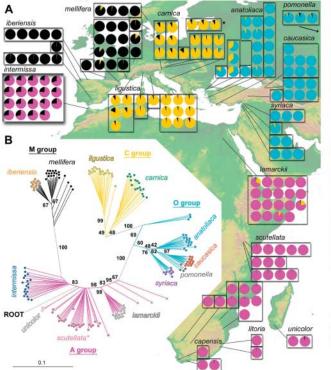*To whom correspondence should be addressed. E-mail: charlie@life.uiuc.edu

**Fig. 1.** Patterns of genetic variation in *A. mellifera*. (**A** and **B**) Principal component plots of Old World individuals based on the 1136 SNP loci (*22*). (A) Bees cluster in four main groupings (M, C, O, and A) in the coordinates of PC1 and PC2. (B) Higher order PCs partition geographical subspecies. See figs. S1 and S2 for additional PCA results. (**C**) Results of *Structure* analysis (*22*). Individuals are represented by vertical lines, grouped by geographical subspecies (Old World) or sampling regime (New World). Division of individuals into colored segments represents the probability of assignment of that individual to each of *K* groups. *K* represents an arbitrary number of hypothetical populations, shown on the right. Colors for *K* = 4 correspond to colors used in PCA (A and B) and Fig. 2, A and C to I.

In North America, introductions of the subspecies *Apis mellifera mellifera* began as early as 1622. Subsequent introductions of *A. m. ligustica* (the "Italian" bee) began in 1859, followed by introductions of at least seven other subspecies from Europe, the Near East, and northern Africa (the descendants of which are collectively called "European") (10). Early introductions in South America are less clear, but probably also involved

**Fig. 2.** Geographical and temporal patterns of diversification. (**A**) Assignment probabilities from *Structure* analysis (Fig. 1C, *K* = 4) of Old World individuals. *A. m. pomonella* (*N* = 3) collected from Kyrgyzstan. Three *A. m. carnica* (indicated by asterisk) were provided by a research facility. (**B**) Neighbor-joining tree based on allele-sharing distance. "ROOT" represents a single derived genotype consisting of 289 SNPs with one common homozygous genotype from *A. cerana* (*N* = 7) and *A. dorsata* (*N* = 4) (32). Branches are colored to correspond to the four groups in (A). The group labeled *scutellata\** also includes *A. m. litoria* (*N* = 2) and *A. m. capensis* (*N* = 3). (**C** to **I**) *Structure* analysis (Fig. 1C, *K* = 4) of New World individuals. (C) Geographic distribu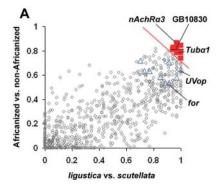tion of Africanization in South America. Hatched region denotes hybrid zone (between latitude 31°04.993′S and 33°11.603′S). (D) Feral bees from Texas before Africanization. (E) Africanized bees (by mitochondria and morphometry) from Texas collected from 1996 to 2000. (F) Africanized bees from Arizona collected from 1996 to 2000. (G) Managed bees from Texas collected in 2005. (H and I) Time series from the Welder Wildlife Refuge (southern Texas) during Africanization (1993 to 2001). (H) Bees with African mitochondria. (I) Bees with European mitochondria.
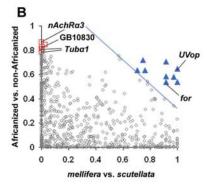
**Table 1.** Candidate loci for selection. Subset of outlier SNPs from Fig. 3A (Set 1) and Fig. 3B (Set 2) (see table S3). The frequency of predominant *A. m. scutellata* allele is indicate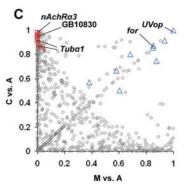d. Syn, synonymous codon change. Groups for South and North America are as indicated in Fig. 2 (arranged by increasing Africanization from left to right). UTR, untranslated region; UV, ultraviolet; GTPase, guanosine triphosphatase; PKG, cGMP-dependent protein kinase.

| | Old World | | | | South America | | | North America | | | | | Near or affected gene | SNP position or effect | Gene product |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| SNP | ligustica | mellifera | scutellata | caucasica | South | Hybrid zone | North | European | 1993–1995 | 1996–1998 | 1999–2001 | Africanized | | | |
| | | | | | | | | | *Set 1* | | | | | | |
| est6550 | 0.00 | 0.95 | 1.00 | 1.00 | 0.06 | 0.83 | 0.99 | 0.21 | 0.38 | 0.64 | 0.88 | 0.91 | GB11704 | 5′ UTR | |
| ahb7495 | 0.00 | 0.94 | 0.97 | 0.00 | 0.13 | 0.42 | 0.97 | 0.08 | 0.29 | 0.50 | 0.46 | 0.91 | GB10583 (nAchRα3) | † | Nicotinic acetylcholine receptor, α3 subunit |
| est8764 | 0.00 | 1.00 | 0.97 | 1.00 | 0.25 | 0.75 | 1.00 | 0.13 | 0.59 | 0.69 | 0.73 | 0.86 | GB10830 | S→A | |
| est9209* | 0.03 | 0.98 | 1.00 | 0.36 | 0.19 | 0.75 | 0.97 | 0.21 | 0.54 | 0.62 | 0.79 | 0.91 | GB10514(Tubα1) | Syn | Tubulin, α1 chain |
| est9211* | 0.03 | 0.98 | 0.97 | 0.32 | 0.19 | 0.75 | 0.97 | 0.21 | 0.55 | 0.67 | 0.79 | 0.95 | GB10514 (Tubα1) | Syn | Tubulin, α1 chain |
| | | | | | | | | | *Set 2* | | | | | | |
| ahb12140 | 0.00 | 0.00 | 1.00 | 0.00 | 0.00 | 0.33 | 0.76 | 0.04 | 0.20 | 0.38 | 0.44 | 0.73 | GB18171 (UVop) | Syn | UV-sensitive opsin |
| ahb11258 | 0.00 | 0.00 | 0.92 | 0.00 | 0.00 | 0.25 | 0.74 | 0.00 | 0.21 | 0.40 | 0.65 | 0.86 | GB15150 | Intron | GTPase activator |
| ahb11774 | 0.08 | 0.03 | 0.94 | 0.00 | 0.00 | 0.50 | 0.76 | 0.04 | 0.23 | 0.55 | 0.54 | 0.50 | GB15050 | ~1500 bp 5′ | |
| est424 | 0.25 | 0.20 | 0.97 | 0.00 | 0.19 | 0.92 | 0.92 | 0.17 | 0.41 | 0.45 | 0.71 | 0.86 | GB19539 | Syn | |
| est10185 | 0.00 | 0.00 | 0.91 | 1.00 | 0.06 | 0.50 | 0.81 | 0.13 | 0.32 | 0.57 | 0.58 | 0.45 | GB18394 (for) | 3′ UTR | *foraging*; PKG |
| ahb9731 | 0.00 | 0.05 | 0.78 | 0.00 | 0.00 | 0.17 | 0.69 | 0.00 | 0.18 | 0.57 | 0.50 | 0.77 | GB15214 | Syn | |

*These two SNPs are 939 base pairs (bp) apart.    †Approximately 3000 bp 5′ of start codon for gene prediction GB10583; this gene has an alternate predicted structure (S.C_Group5.24000030B) (3) that places the SNP 30 bp upstream of the start codon.

**Fig. 3.** $F_{ST}$ of 1136 SNPs in introduced and native populations. (**A**) Differentiation in the New World (Africanized versus non-Africanized; *y* axis) largely corresponds to differentiation between *A. m. scutellata* and *A. m. ligustica* (*x* axis). Filled red squares represent the 1% of SNPs exhibiting the highest differentiation in both axes (Table 1 and table S3, "Set 1"). (**B**) SNPs exhibiting high differentiation between *A. m. scutellata* and *A. m. mellifera* (*x* axis) exhibited little differentiation in the New World (*y* axis). Filled blue triangles represent the 1% of SNPs exhibiting the highest differentiation in both axes (Table 1 and table S3, "Set 2"). (**C**) Few SNPs exhibit high differentiation in both C and M lineages relative to A, consistent with independent derivation of eastern and western European subspecies from Africa. Set 2 loci generally exhibit high differentiation in the New World (B, filled blue triangles) and in the two Old World lineages independently derived from Africa (C, open blue triangles). Several SNPs occurring in or near genes are indicated (abbreviations as in Table 1). "Africanized" refers to all individuals in Fig. 2, C (north of hybrid zone), E, and F. "Non-Africanized" refers to all individuals in Fig. 2, C (south of hybrid zone) and D.

early introduction of western European subspecies (*A. m. mellifera* and *A. m. iberiensis*) and later introduction of eastern European subspecies (*10*). In 1956, a subspecies from the African savannas, *A. m. scutellata*, was intentionally introduced to Brazil. *A. m. scutellata* became established and dispersed northward and southward from Brazil through South and Central America (*11–13*), hybridizing with and displacing previously introduced honey bees. Africanized bees reached their southern limit in Argentina in the 1970s (*14*), but the eventual northward limit in the United States [first invaded in 1990 (*15*)] is unknown. Although ample evidence shows that both European and African alleles occur in Africanized populations (*14, 16–21*), *A. m. scutellata*–derived characteristics, including nesting biology, swarming and absconding behavior, foraging, diet, and mitochondrial DNA (mtDNA), tend to displace European characteristics over time (*13*).

We analyzed 1136 SNP genotypes (*22, 23*) from 328 *A. mellifera*, including 175 individuals from native populations (representing 14 subspecies in Europe, Africa, and Asia) and 153 individuals from introduced populations in North and South America. We inferred the ancestral genotype for a subset of SNPs by analyzing 13 additional individuals from 3 congeneric species, *A. cerana*, *A. dorsata*, and *A. florea*.

In the Old World, there was clear population structure at the level of major groups (each containing multiple subspecies) and geographical subspecies. Principal component analysis (PCA) revealed four major clusters (Fig. 1A) consistent with Ruttner's four morphometrically defined lineages (*5, 24*): M (western and northern Europe), C (eastern Europe), O (Near East and central Asia), and A (Africa). The four groups were broadly consistent with mitochondrial data (*25*) and with assignment by the program *Structure* (Fig. 1C) (*22*). Pairwise $F_{ST}$ between major groups ranged from 0.242 to 0.565 (table S2).

Each of the 10 geographical subspecies with 9 to 20 individuals was genetically distinct, and subspecies could be partitioned in either PCA or *Structure* analysis (Fig. 1B; figs. S2 and S3). $F_{ST}$ among subspecies was 0.501; much lower values were obtained for subspecies within each of the major groups (0.053, 0.129, 0.118, and 0.082 for M, A, O, and C, respectively). Despite considerable differentiation, admixture between geographically proximal populations in the Old World was evident in both PCA (Fig. 1A) and *Structure* analysis (Figs. 1C and 2A). Comparison of linkage disequilibrium (LD) between subspecies (fig. S4) also suggested differences related to admixture, bottlenecks, or both. For example, *A. m. mellifera* and *A. m. ligustica* exhibited relatively high LD for SNP pairs separated by ≤10 kb (mean correlation coefficient $r^2$ of 0.45 and 0.30, respectively), decreasing to 0.18 and 0.06 for SNP pairs separated by 50 to 100 kb). In contrast, *A. m. scutellata* and *A. m. intermissa* exhibited relatively little or no LD (0.08 and 0.13 for ≤10 kb; 0.07 and 0.10 for 50 to 100 kb). LD was low overall and

affected only a small number of SNPs (fig. S4), consistent with extraordinarily high recombination rates in *A. mellifera* [~19 cM/Mb (*26*)].

To explore evolutionary relationships, we generated distance trees using Old World *A. mellifera*, and rooted these trees using an ancestral genotype derived from *A. cerana* and *A. dorsata*. The two groups that comprise European bees (M and C) were the most distantly related, despite their geographic proximity, and M was considerably more similar to A ($F_{ST}$ = 0.242) than to either C or O ($F_{ST}$ = 0.565 and 0.458, respectively) (Fig. 2B, table S2). Unexpectedly, the tree was rooted in Africa (Fig. 2B). Tree rooting and other relationships had strong bootstrap support and were essentially the same when we used parsimony-based methods, alternative SNP subsets, or genotypes from congeneric individuals (rather than a single derived ancestral genotype) (figs. S5 to S7). On the basis of these data, we hypothesize that *A. mellifera* originated in Africa and that there were at least two subsequent expansions into Eurasia—a western expansion into Europe (M) and one or more (independent) eastern expansions into Asia and Europe (O and C).

New World bees collected in Brazil (the original site of *A. m. scutellata* introduction in 1956) were highly Africanized, although all individuals exhibited introgression with the M group (Figs. 1C and 2C; fig. S1). Similar levels of introgression were not evident in any *A. m. scutellata* from Africa (Figs. 1C and 2A; fig. S1.) Two parallel transects in northern Argentina revealed transitions from predominantly African genotypes (north) to predominantly C group genotypes (south) (Fig. 2C and fig. S1B; table S2). *Structure* analysis revealed a consistent minority of M group alleles (Fig. 2C, black) in all individuals north and south of the hybrid zone. Retention of several M group markers has been observed at the population level (*13*); our results indicate that this effect occurs at the individual level. The location of the hybrid zone (Fig. 2C) appears to be stable, matching the general latitude of transition from Africanized to European bees in previous studies published in 1982 and 1991 (*14, 27*).

Feral bees collected in the United States before Africanization were of mixed Eurasian ancestry, dominated by C group but admixed with M and O groups (Fig. 2D and fig. S1A). After Africanization, bees defined as Africanized by morphology exhibited predominantly African ancestry admixed with M, C, and to a lesser extent, O group bees (Fig. 2, E and F; fig. S1A). During Africanization, bees from southern Texas (1993 to 2001) showed a transition from mixed European to substantially, but not exclusively, African ancestry (Fig. 2, H and I; fig. S1C). Individuals in early (1993 to 1995), middle (1996 to 1998), and late collections (1999 to 2001) with African mtDNA were predominantly African at nuclear loci (Fig. 2H, table S2). In contrast, bees with European mtDNA became progressively differentiated from *A. m. ligustica* (C group) and increasingly similar to *A. m. scutellata* (Fig. 2I, table S2). In North America, as in South America,

an unexpectedly consistent minority of M group alleles was evident in all genomes of both non-Africanized and Africanized individuals (Fig. 2, D to I, black). Recently collected bees (1999 to 2001) with African and European mtDNA were virtually indistinguishable at the SNP loci (Fig. 2, H versus I; table S2).

To identify loci with extreme patterns of differentiation, possibly due to natural selection (28), we calculated $F_{ST}$ among Old and New World bees at all 1136 SNP loci (Fig. 3). Pairwise differentiation between non-Africanized and Africanized bees in the New World was highly correlated with differentiation between *A. m. ligustica* and *A. m. scutellata* in the Old World (Fig. 3A) ($r = 0.81$), consistent with the hypothesis that Africanization involves replacement of *A. m. ligustica* alleles by *A. m. scutellata* alleles (SNPs with the highest levels of differentiation in both comparisons are indicated in Fig. 3A by filled red squares and in Table 1 and table S3 as "Set 1"). In contrast, most loci highly differentiated between *A. m. scutellata* and *A. m. mellifera* showed little differentiation between Africanized and non-Africanized bees (Fig. 3B) ($r = -0.37$), indicating that biased replacement of C- but not M-derived alleles [(13) and the present study] may occur throughout the genome. However, a small subset of loci that distinguish *A. m. mellifera* from *A. m. scutellata* also exhibited relatively high differentiation for Africanized versus non-Africanized New World bees (Fig. 3B, filled blue triangles). These few SNPs distinguished both C and M groups from the A group (Fig. 3C and Table 1).

Of 19 SNPs identified as potential sites for selection (Fig. 3), 11 occurred in or near genes (Table 1 and table S3). Six occurred in coding regions, but only one was predicted to cause a nonsynonymous amino acid change (in a gene of unknown function; Table 1). Although some of these SNPs may alter gene function, it is also possible that some or all have hitchhiked with linked polymorphisms that are under selection.

These data provide a global view of the population genetic structure of *A. mellifera*. Future studies can capitalize on the large number of SNPs presented here to further investigate population genetic patterns within and among populations and genes under selection.

### References and Notes

1. E. Crane, *The World History of Beekeeping and Honey Hunting* (Routledge, New York, 1999).
2. K. S. Delaplane, *Crop Pollination by Bees* (CABI, New York, 2000).
3. Honey Bee Genome Sequence Consortium, *Nature*, in press.
4. M. C. Arias, W. S. Sheppard, *Mol. Phylogenet. Evol.* **37**, 25 (2005).
5. F. Ruttner, *Biogeography and Taxonomy of Honeybees* (Springer, New York, 1988).
6. W. S. Sheppard, M. D. Meixner, *Apidologie (Celle)* **34**, 367 (2003).
7. J. M. Cornuet, L. Garnery, *Apidologie (Celle)* **22**, 627 (1991).
8. W. S. Sheppard, S. H. Berlocher, *Apidologie (Celle)* **20**, 419 (1989).
9. E. O. Wilson, in *The Insect Societies* (Belknap, Harvard University, Cambridge, MA, 1971), pp. 94–95.
10. W. S. Sheppard, *Am. Bee J.* **129**, 617 (1989).
11. W. S. Kerr, *S. Afr. Bee J.* **39**, 33 (1967).
12. W. S. Sheppard, D. R. Smith, *Ann. Entomol. Soc. Am.* **93**, 159 (2000).
13. S. S. Schneider, G. De Grandi-Hoffman, D. R. Smith, *Annu. Rev. Entomol.* **49**, 351 (2004).
14. W. E. Kerr, S. D. Delrio, M. D. Barrionuevo, *Am. Bee J.* **122**, 196 (1982).
15. E. A. Sugden, K. R. Williams, *Glean. Bee Cult.* **119**, 18 (1990).
16. T. E. Rinderer, J. A. Stelzer, B. P. Oldroyd, S. M. Buco, W. L. Rubink, *Science* **253**, 309 (1991).
17. T. E. Rinderer, B. P. Oldroyd, W. S. Sheppard, *Sci. Am.* **269**, 52 (December 1993).
18. T. E. Rinderer *et al.*, *Apidologie (Celle)* **24**, 569 (1993).
19. J. J. G. Quezada-Euan, *Apidologie (Celle)* **31**, 443 (2000).
20. J. J. G. Quezada-Euan, L. M. Medina, *Apidologie (Celle)* **29**, 555 (1998).
21. K. E. Clarke, T. E. Rinderer, P. Franck, J. G. Quezada-Euan, B. P. Oldroyd, *Evolution Int. J. Org. Evolution* **56**, 1462 (2002).
22. Materials and methods are available as supporting material on *Science* Online.
23. We identified SNPs from genomic sequence and expressed sequence tags (ESTs) [table S1 and Methods (22)]. Overall, 1536 putative SNPs were examined, of which 1136 were valid polymorphisms. Although genome- and EST-derived SNPs were derived from different populations, both sets (examined separately) provided highly consistent results (e.g., figs. S6 and S7).
24. F. Ruttner, L. Tassencourt, J. Louveaux, *Apidologie (Celle)* **9**, 363 (1978).
25. "C" mitochondrial lineage includes C and part of the O morphological lineages (29, 30). "O" was used to designate a mitochondrial group (31) different from the morphological O group.
26. R. E. Page Jr., J. Gadau, M. Beye, *Genetics* **160**, 375 (2002).
27. W. S. Sheppard, T. E. Rinderer, J. A. Mazzoli, J. A. Stelzer, H. Shimanuki, *Nature* **349**, 782 (1991).
28. P. C. Sabeti *et al.*, *Science* **312**, 1614 (2006).
29. M. R. Palmer, D. R. Smith, O. Kaftanoglu, *J. Hered.* **91**, 42 (2000).
30. D. R. Smith, A. Slaymaker, M. Palmer, O. Kaftanoglu, *Apidologie (Celle)* **28**, 269 (1997).
31. P. Franck, L. Garnery, M. Solignac, J. M. Cornuet, *Apidologie (Celle)* **31**, 167 (2000).
32. M. D. Shriver *et al.*, *Hum. Genomics* **2**, 81 (2005).
33. We thank K. Hartfelder, W. Rubink, C. Tillberg, C. Smith, S. Wittman, M. Sen Sarma, O. Kaftanoglu, O. R. Taylor Jr., R. Crewe, S. Goodman, and M. Sasaki for bee collections; K. Heeres for DNA extraction; S. Edelheit for SNP genotyping; A. Pinto for mitotyping; J. Pinter for genotype analyses and base calling; J. Theirrian for project coordination; and G. E. Robinson, B. S. Gaut, A. D. Long, and P. L. Morrell for advice and manuscript review. We thank M. Adams for facilities and personnel for SNP genotyping; and the Human Genome Sequencing Center, Baylor College of Medicine, and RIKEN for providing sequence data ahead of publication. This work was supported by the Institute for Genomic Biology, School of Integrative Biology, and the Research Board at the University of Illinois (C.W.W.), NSF (A.V.S.), U.S. Department of Agriculture and California Department of Consumer Affairs, Structural Pest Board (N.D.T.).

# Functional CpG Methylation System in a Social Insect

Ying Wang,[1] Mireia Jorda,[2] Peter L. Jones,[1] Ryszard Maleszka,[3] Xu Ling,[4] Hugh M. Robertson,[5,6,7] Craig A. Mizzen,[1,6] Miguel A. Peinado,[2] Gene E. Robinson[5,6,7]*

DNA methylation systems are well characterized in vertebrates, but methylation in *Drosophila melanogaster* and other invertebrates remains controversial. Using the recently sequenced honey bee genome, we present a bioinformatic, molecular, and biochemical characterization of a functional DNA methylation system in an insect. We report on catalytically active orthologs of the vertebrate DNA methyltransferases Dnmt1 and Dnmt3a and b, two isoforms that contain a methyl-DNA binding domain, genomic 5-methyl-deoxycytosine, and CpG-methylated genes. The honey bee provides an opportunity to study the roles of methylation in social contexts.

Among the many important functions of CpG DNA methylation, sex-specific regulation of gene expression (imprinting) in vertebrates stands out because it provides insight into intragenomic conflict (1, 2). Provided social insects have CpG methylation, they would be ideal models to further explore the kin-conflict theory of imprinting, because insect societies are composed of many different types of relatives and they interact with each other in many evolutionarily important contexts (2, 3).

However, although widely conserved from yeast and fungi to plants to vertebrates, DNA methylation in insects is enigmatic. Evidence of CpG-methylated sequences exists for several insect species (4–6), but no bona fide invertebrate deoxycytosine methyltransferases (DNMTs) have been described. Conversely, the model insect *Drosophila melanogaster* shows limited DNA methylation, predominantly in asymmetric CpT and CpA dinucleotides (7), and this is attributed to the only DNMT family member encoded in its genome, dDNMT2, a tRNA^Asp methyltransferase

[1] Department of Cell and Developmental Biology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA. [2] IDIBELL–Cancer Research Institute, L'Hospitalet 08907, Catalonia, Spain. [3] Visual Sciences and Australian Research Council Centre for Molecular Genetics of Development, Research School of Biological Sciences, The Australian National University, Canberra, Australia. [4] Department of Computer Science, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA. [5] Department of Entomology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA. [6] Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA. [7] Neuroscience Program, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA.

*To whom correspondence should be addressed. E-mail: generobi@life.uiuc.edu