Genomic dissection of behavioral maturation in the honey bee

Charles W. Whitfield, Yehuda Ben-Shahar, Charles Brillet, Isabelle Leoncini, Didier Crauser, Yves LeConte, Sandra Rodriguez-Zas, and Gene E. Robinson

PNAS published online Oct 25, 2006;
doi:10.1073/pnas.0606909103

Genomic dissection of behavioral maturation in the honey bee

Charles W. Whitfield*†‡, Yehuda Ben-Shahar§¶, Charles Brillet‖, Isabelle Leoncini‖, Didier Crauser‖, Yves LeConte‖, Sandra Rodriguez-Zas*†‡**, and Gene E. Robinson*†‡††

Departments of *Entomology and **Animal Science, †Neuroscience Program, and ‡Institute for Genomic Biology, University of Illinois at Urbana–Champaign, Urbana, IL 61801; §Howard Hughes Medical Institute, ¶University of Iowa College of Medicine, Iowa City, IA 52242; and ‖Laboratoire Biologie et Protection de l'Abeille, Ecologie des Invertébrés, Unité Mixte de Recherche, Institut National de la Recherche Agronomique/Université d'Avignon et des Pays de Vaucluse, Site Agroparc, Domaine Saint-Paul, 84914 Avignon Cedex 9, France

This contribution is part of the special series of Inaugural Articles by members of the National Academy of Sciences elected on May 3, 2005.

Contributed by Gene E. Robinson, August 9, 2006

Honey bees undergo an age-related, socially regulated transition from working in the hive to foraging that has been previously associated with changes in the expression of thousands of genes in the brain. To understand the meaning of these changes, we conducted microarray analyses to examine the following: (*i*) the ontogeny of gene expression preceding the onset of foraging, (*ii*) the effects of physiological and genetic factors that influence this behavioral transition, and (*iii*) the effects of foraging experience. Although >85% of ≈5,500 genes showed brain differences, principal component analysis revealed discrete influences of age, behavior, genotype, environment, and experience. Young bees not yet competent to forage showed extensive, age-related expression changes, essentially complete by 8 days of age, coinciding with previously described structural brain changes. Subsequent changes were not age-related but were largely related to effects of juvenile hormone (JH), suggesting that the increase in JH that influences the hive bee–forager transition may cause many of these changes. Other treatments that also influence the onset age of foraging induced many changes but with little overlap, suggesting that multiple pathways affect behavioral maturation. Subspecies differences in onset age of foraging were correlated with differences in JH and JH-target gene expression, suggesting that this endocrine system mediates the genetic differences. We also used this multifactorial approach to identify candidate genes for behavioral maturation. This successful dissection of gene expression indicates that, for social behavior, gene expression in the brain can provide a robust indicator of the interaction between hereditary and environmental information.

The honey bee, *Apis mellifera*, is one of the model organisms being used to achieve a comprehensive understanding of social life in molecular terms: how social life evolved, how it is governed, and how it influences all aspects of genome structure, genome activity, and organismal function (1). Honey bees offer complex but experimentally accessible social behavior, a compact and well studied brain, and a sequenced genome that provides the foundation for ever-increasing genomic resources.

Honey bees, like many species of social insects, display a division of labor among colony members that is based on behavioral specializations associated with age (2). Adult worker honey bees perform a series of tasks in the hive when they are young (such as brood care or "nursing") and, at ≈2–3 weeks of age, shift to foraging for nectar and pollen outside the hive. The transition to foraging involves changes in endocrine activity, metabolism, circadian clock activity, brain chemistry, brain structure, and brain gene expression (3).

The pace of behavioral maturation in honey bees is not rigid, because the onset age of foraging depends on the needs of the colony. Pheromones and other social cues mediate this behavioral ontogeny and affect foraging onset (4). These cues are thought to act directly or indirectly on physiological factors including juvenile hormone (JH) (5, 6) and molecular pathways associated with the *foraging* and *malvolio* genes, which are among the presumably many genes that play a causal role in honey bee behavioral maturation (7, 8). Variation in the pace of behavioral ontogeny in honey bees also has a genotypic component (9–11).

Microarray analysis is being used to gain a broader appreciation of the genes and molecular pathways involved in age-related division of labor in honey bee colonies (12–15). Nurses and foragers show differences in brain mRNA abundance in approximately one-third of the ≈5,500 genes analyzed (estimated to represent ≈40% of the genes in the bee genome) (12).

To understand the meaning of these changes, we conducted microarray analyses of the bee brain to examine the following: (*i*) the ontogeny of gene expression before the onset of foraging, (*ii*) the effects of genetic and physiological factors that influence the age at onset of foraging, and (*iii*) the effects of foraging experience. First, we show how multiple overlapping influences on brain gene expression can be decomposed into discrete effects, even under naturalistic, free-flying conditions in which bees exhibit typical behavior. Second, we use these results in conjunction with manipulative experiments to test two hypotheses: (*i*) behavior-associated differences in brain gene expression are related to both upstream effectors of behavior (such as JH) and downstream effects of foraging activity; and (*ii*) natural genetic differences in brain gene expression between subspecies are related, at least in part, to differences in upstream effectors of behavior. Third, we use Gene Ontology (GO) analyses to identify biological processes that might be particularly prominent in honey bee behavioral maturation. Fourth, we show how results of these analyses provide a set of candidate genes for socially mediated and genetic differences in behavior.

Results

Brain expression profiles were analyzed by using microarrays derived from honey bee brain ESTs (16); enhanced annotation was provided by results from the honey bee genome project (17). For experiments 1–3, a total of 5,736, 5,559, and 5,637 genes,

Author contributions: C.W.W., Y.B.-S., Y.L., and G.E.R. designed research; C.W.W., Y.B.-S., C.B., I.L., D.C., and Y.L., performed research; C.W.W., Y.B.-S., S.R.-Z., and G.E.R. analyzed data; and C.W.W. and G.E.R. wrote the paper.

The authors declare no conflict of interest.

Freely available online through the PNAS open access option.

Abbreviations: GO, Gene Ontology; JH, juvenile hormone; PC, principal component; PCA, PC analysis; QMP, queen mandibular pheromone.

Data deposition: Gene expression data meet Minimum Information About a Microarray Experiment (MIAME) standards and have been deposited at ArrayExpress (www.ebi.ac.uk/arrayexpress) with accession nos. E-TABM-149, E-TABM-150, and E-TABM-151.

See accompanying Profile on page 16065.

††To whom correspondence should be addressed at: Department of Entomology, University of Illinois at Urbana–Champaign, 320 Morrill Hall, 505 South Goodwin Avenue, Urbana, IL 61801. E-mail: generobi@life.uiuc.edu.

Table 1. Number of genes showing significant expression differences in the honey bee brain as a function of age, behavioral, genetic, environmental, and physiological factors (experiments 1–3)

Factors	$P < 0.05$	$P < 0.001$	$P < 10^{-6}$	5% FDR*
Field colonies				
Age/behavior† (A)	5,275	4,477	3,230	5,258
Subspecies‡ (S)	2,943	1,637	692	2,576
Colony§ (C)	1,501	335	34	679
A × S	708	68	2	39
A × C	1,124	204	24	283
S × C	1,092	198	13	276
Age/behavior contrasts (*t* tests)				
d1 vs. d4	3,280	1,991	992	2,991
d4 vs. d8	2,113	590	44	1,320
d8 vs. d12	937	84	0	50
d12 vs. d17	531	21	0	0
d17 vs. d17F	3,831	2,414	937	3,627
d1 vs. d17F	4,578	3,745	2,749	4,516
Preforaging maturation [d1 vs. (d8, d12, and d17)]	4,093	3,014	1,887	3,957
Hive-bee-to-forager [(d8, d12, and d17) vs. d17F]	4,088	2,965	1,559	3,970
Treatments in laboratory cages¶				
Methoprene	1,587	481	95	894
Manganese	2,539	509	42	1,594
cGMP	1,543	461	84	827
cAMP	999	129	11	151
cGMP vs. cAMP	1,307	327	40	584
Experience deprivation‖				
Hive-restricted vs. F	129	16**	0	2

Gene lists are available at http://stagbeetle.animal.uiuc.edu/papers/PNAS.html. F, forager.

*Gene lists estimated to contain 5% false positives by using the false-discovery rate (FDR) step-up method (19).

†Focal bees were from six age/behavior groups: 0- to 1-h after eclosion (d1), bees from the center of the colony (brood area) at 4, 8, 12, and 17 days of age (d4, d8, d12, and d17), and 16- or 17-day-old foragers (d17F).

‡Focal bees were from two subspecies: *A. m. ligustica* and *A. m. mellifera*.

§Focal bees were cofostered in two host colonies.

¶All tests were contrasts (*t* tests) between treatment and vehicle control, except cGMP vs. cAMP.

‖One-tailed *t* test. Null hypothesis is that hive-restricted vs. F differences in gene expression are not in the same direction as d8, d12, and d17 vs. F differences; only genes significant in the latter contrast at $P < 0.001$ were considered.

**A total of 17 genes were significant using two-tailed test with all genes.

respectively, passed quality criteria and were analyzed (see *Methods*). We used mixed-model ANOVA (18, 19) to determine the number of genes showing differential expression (Table 1). Unless otherwise specified, $P < 0.001$ was used to denote statistical significance when all genes were tested, leading to an expectation of fewer than six false positives per test.

Additional analyses used a set of marker genes, which were shown (12) to be the best 100 genes on the microarray for classifying brain expression profiles of individual bees as nurse or forager. Expression differences for these genes are associated with behavior (either nursing or foraging) and not age (12). We compared the previously determined forager/nurse brain gene expression ratios from this set with ratios for these same genes in the following experiments to determine whether particular comparisons (age, genotype, and treatment) reveal patterns of expression that are more forager-like, more nurse-like, or dissimilar to either.

Experiment 1: Age-Related, Behavior-Related, and Genetic Differences in Brain Gene Expression. We studied 72 individual bees from two subspecies of European honey bees (*A. m. ligustica* and *A. m. mellifera*) that differ in the age at onset of foraging (early and late, respectively; ref. 10 and Fig. 5, which is published as supporting information on the PNAS web site). Bees were cofostered in the same "host" colonies in the field (one *ligustica* and one *mellifera*) and collected at different ages. We generated gene expression profiles for the 72 dissected bee brains using 108 microarrays.

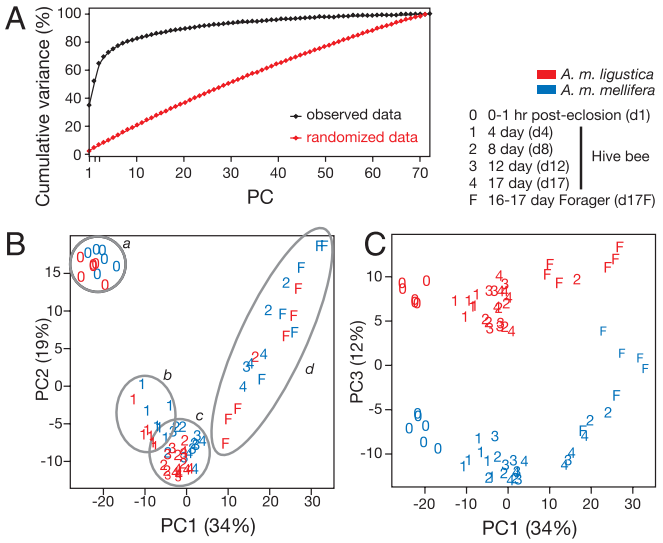

Fig. 1. Division of labor in honey bee colonies and brain gene expression: age-related, behavior-related, and genetic differences. PCA using brain gene expression measurements for all 5,736 genes from all 72 bees in experiment 1 was performed. (*A*) Cumulative variance of PCs. For randomized data, gene expression levels were shuffled among genes (within sample). (*B* and *C*) Individual bees are plotted as a function of PC1 and PC2 (*B*) or PC1 and PC3 (*C*). Age/behavior group and subspecies are indicated in the key. See Table 4 for PCs 1–9.

Expression differences in these brains were extensive. There were significant effects of ontogeny (77% of genes), subspecies (29%), colony (6%), and interactions between these factors (1–4%) (Table 1). Eighty-five percent differed due to at least one of these factors, and 25% differed due to more than one factor.

Although the experiment involved free-flying bees, presumably subject to many influences in the colony and external environment, principal component analysis (PCA) revealed that 65% of variation could be explained by as few as three principal components (PCs) (Fig. 1*A*; >80% was explained by nine PCs; and see Table 4, which is published as supporting information on the PNAS web site). These results indicate that a small number of factors (either controlled or uncontrolled variables) can account for most individual differences in brain gene expression. Analysis of PCs revealed discrete gene expression "axes" that were associated with age, behavior, genotype, and host colony (Fig. 1 *B* and *C* and Table 4).

Differences in brain gene expression reflected in PC1 and -2 (Fig. 1*B*) reveal two distinct axes. The first axis, associated with preforaging maturation, is indicated by three PCA-generated age clusters: newly eclosed bees (cluster *a*), 4-day-old hive bees (cluster *b*), and ≥8-day-old hive bees (cluster *c*). Consistent with the PCA results, *t* tests showed the most extensive gene expression differences between newly eclosed bees and 4-day-olds (1,991 genes), fewer differences between 4- and 8-day-olds (590 genes), and fewer still between subsequent age groups (Table 1; ≥8-day-old age groups showed no significant differences for PC1 and -2; Table 4). These results indicate that early age-related changes in brain gene expression are essentially complete by 8 days of age.

The second axis revealed by PC1 and -2 was associated with differences between hive bees (≥8-day-old) and foragers (Fig. 1*B*, cluster *d*; 2,965 genes in *t* test; Table 1). These differences were highly correlated with differences between age-matched nurses and foragers ($r = 0.91$ for the behavior marker genes; $P = 3.6 \times 10^{-38}$;

Fig. 6*A*, which is published as supporting information on the PNAS web site). Although both the preforaging maturation and hive-to-forager axes involved differential expression of an overlapping set of genes, PCA indicates that they constitute distinct trends, which likely involve different transcriptional regulatory processes.

There were also extensive brain gene expression differences between *A. m. ligustica* and *A. m. mellifera* (PC3, Fig. 1*C*; 1,637 genes, Table 1). They may represent true genotypic differences in brain mRNA abundance or differences in microarray hybridization due to subspecies sequence polymorphisms (the array was derived from bees that were largely *ligustica*; ref. 16). Hybridization differences due to polymorphism, if present, were not extensive because measurements of gene expression were not significantly biased toward *ligustica* (834 and 803 were higher in *ligustica* and *mellifera*, respectively; $P = 0.23$, binomial test). Effects of sequence polymorphism thus should be random with respect to gene expression in other contrasts in this study (which were all genotype-matched) and should not lead to false-positive results in the correlation analyses that follow.

There were no obvious subspecies differences in brain gene expression during preforaging maturation that might explain the differences in age at onset of foraging. Four-day-old *ligustica* and *mellifera* appear to be at the same position in the age-associated axis revealed by PC1 and -2 (Fig. 1*B*), suggesting that their pace of preforaging maturation was similar. However, *ligustica* ≥8-day-old hive bees were significantly more "forager-like" than same-aged *mellifera* hive bees. Expression ratios for *ligustica*/*mellifera* were significantly correlated with reported (12) forager/nurse ratios for the behavior marker genes ($r = 0.23$; $P = 0.026$; Fig. 6*C*). This trend was even stronger ($r = 0.51$; $P = 9.2 \times 10^{-8}$; Fig. 6*D*) when we removed seven outlier hive bees (six *mellifera* and one *ligustica*) that were in cluster *d*, rather than the expected cluster *c*, in Fig. 1*B*. This result led us to hypothesize that the earlier onset age of foraging in *ligustica* may be related to increased activity in some signaling pathway associated with foraging in ≥8-day-old hive bees. We explore this hypothesis in experiment 4 below.

Experiment 2: Effects of Treatments That Influence Onset Age of Foraging on Brain Gene Expression. Three treatments were used: methoprene (a JH analog), manganese [associated with *malvolio*, which encodes a manganese transporter (8)], and cGMP [associated with *foraging*, which encodes a cyclic G-dependent protein kinase (7)]. Bees were genotype-matched (full sisters of primarily *A. m. ligustica* descent) and were housed in small laboratory cages with no possibility for typical nursing or foraging behavior. We examined gene expression in pooled samples of dissected brains ($n = 50$ brains per treatment), using a total of 36 microarrays. Bees not analyzed for gene expression were used to verify treatment efficacy: methoprene, manganese, and cGMP caused precocious foraging, whereas vehicle and cAMP did not (data not shown), as expected (7, 8, 20).

Each treatment significantly affected the expression of >100 of the 5,559 genes tested (Table 1). We tested whether treatments caused significant forager- or nurse-like trends in brain gene expression. We asked this question by using the behavior marker genes and χ^2 (Table 2) and correlation analyses (Table 3, above diagonal). We also explored relationships between treatments, with either the 100 behavior marker genes (Table 3, above diagonal) or all genes on the microarray (Table 3, below diagonal). For comparative purposes, we also analyzed the effects of queen mandibular pheromone (QMP), using data from an independent study (day 3 in ref. 15). QMP delays the onset of foraging (5) and causes nurse-like trends in brain gene expression (15). We also detected nurse-like effects for QMP with our methods of analysis (Tables 2 and 3).

Methoprene and manganese caused significant forager-like changes in brain gene expression (Tables 2 and 3). cGMP did not cause significant effects when compared with vehicle but caused a forager-like trend when compared with cAMP that was marginally significant in a χ^2 test ($P = 0.09$; Table 2) and significant in correlation analysis ($P = 7.8 \times 10^{-5}$; Table 3). cAMP, which does not accelerate foraging, caused a nurse-like trend that was marginally significant in a χ^2 test ($P = 0.057$; Table 2) and significant in correlation analysis ($P = 0.00011$; Table 3). Methoprene and QMP showed highly significant opposing effects with respect to forager/nurse ratios for the 100 behavior marker genes ($r = 0.54$ and -0.54, respectively; Table 3), consistent with their opposing effects on the onset of foraging. These results indicate that treatments that modulate behavior can cause forager- or nurse-like changes in brain gene expression even in the absence of foraging- or nursing-related experience. Methoprene effects were particularly strong; 41 of the 100 behavior marker genes were regulated in the forager-like direction (Table 2).

Table 2. Forager- or nurse-like characteristics of the gene lists from Table 1, using the 100 behavior marker genes: χ^2 analyses

Characteristic	F markers (43)	N markers (57)	Significance
Field colonies			
Preforaging maturation ↑ (2,130)	28	18	$\chi^2 = 9.461$
Preforaging maturation ↓ (1,963)	7	24	$P = 0.0021$
Hive bee-to-forager ↑ (2,073)	37	1	$\chi^2 = 78.10$
Hive bee-to-forager ↓ (2,015)	0	48	**$P = 2.2e{-}16$**
A. m. ligustica ↑ (1,552)	16	13	$\chi^2 = 0.5678$
A. m. mellifera ↑ (1,391)	8	12	$P = 0.45$
Treatments in laboratory cages			
Methoprene ↑ (798)	23	6	$\chi^2 = 20.07$
Methoprene ↓ (789)	2	18	**$P = 7.5e{-}6$**
Manganese ↑ (1,217)	18	3	$\chi^2 = 18.70$
Manganese ↓ (1,322)	4	20	**$P = 1.5e{-}5$**
cGMP ↑ (799)	12	10	$\chi^2 = 0.0058$
cGMP ↓ (744)	7	7	$P = 0.94$
cAMP ↑ (516)	5	9	$\chi^2 = 3.62$
cAMP ↓ (483)	7	1	$P = 0.057$
cGMP > cAMP (663)	15	4	$\chi^2 = 2.83$
cGMP > cAMP (644)	9	10	$P = 0.093$
Inhibitory pheromone treatment in laboratory cages (15)			
QMP ↑ (497)	1	11	$\chi^2 = 16.78$
QMP ↓ (428)	12	0	**$P = 4.2e{-}5$**

The behavior markers were the best 100 genes for classifying individual brain expression profiles as nurse or forager; their expression is associated with behavior and not age (12). We compared the distribution of these genes that were either up-regulated (↑) in foragers (F) or up-regulated in nurses (N) relative to their distribution in each gene list. Numbers in parentheses are the number of genes in each list ($P < 0.05$). Similar results were obtained when gene lists at $P < 0.001$ were used, except trends for manganese and QMP were not significant because of the small number of genes tested. *P* values in bold indicate the strongest effects. Gene expression data for QMP were taken from ref. 15.

Forager-like trends were induced by different treatments that accelerate the onset age of foraging, but the effects of these treatments on brain gene expression were very different. For example, of the hundreds of genes regulated by methoprene and manganese, only 30 genes were up-regulated and 17 were down-regulated by both treatments. Additionally, these treatments were negatively correlated with each other ($r = -0.26$; Table 3, all genes). In contrast, two treatments that differ in their behavioral effect, cGMP and cAMP, showed the strongest positive correlation between any of the treatments ($r = 0.63$; Table 3, all genes), even though cGMP caused a forager-like trend when compared with cAMP for the 100 behavior marker genes. These results suggest that only a small subset of the target genes for each treatment are likely to be related to onset age of foraging.

Experiment 3: Effects of Flight and Foraging on Brain Gene Expression. We used an established manipulation (21) to obtain hive-restricted bees, "presumptive" foragers without foraging experience. We

Table 3. Forager- or nurse-like characteristics of gene lists from Table 1, using 100 behavior marker genes: Correlational analyses

Genes	F/N	Methoprene	Manganese	cGMP	cAMP	cGMP/cAMP	QMP
F/N		**0.54 (1.8e−8)**	**0.50 (2.2e−7)**	0.16 (0.12)	**−0.39 (0.00011)**	**0.39 (7.8e−5)**	**−0.54 (6.5e−9)**
Methoprene	0.17		0.24 (0.019)	**0.35 (0.00043)**	0.22 (0.031)	0.24 (0.021)	**−0.38 (0.00016)**
Manganese	0.22	−0.26		**0.55 (7.5e−9)**	−0.20 (0.051)	**0.69 (9.6e−15)**	**−0.43 (1.5e−5)**
cGMP	−0.08	0.49	0.01		**0.35 (0.00052)**	n.d.	0.11 (0.29)
cAMP	−0.18	0.58	−0.42	0.63		n.d.	0.32 (0.0016)
cGMP/cAMP	0.08	0.05	0.43	n.d.	n.d.		−0.07 (0.47)
QMP	−0.26	0.07	−0.28	0.29	0.28	0.10	

Correlation analysis was performed on forager/nurse brain expression ratios for the behavior marker genes and expression ratios for the same genes from the gene lists generated in experiment 2. Above diagonal: correlation (*r*) and significance (*P*) (in parentheses) for comparisons of the behavior marker genes; regressions significant at $P < 0.001$ are in bold. Below diagonal: correlation (*r*) for all genes; all positive and negative correlations were highly significant ($P < 10^{-6}$) except manganese vs. cGMP. Gene expression data for QMP were taken from ref. 15. n.d., not determined.

examined brain gene expression in nine individual hive-restricted bees and nine individual free-flying foragers from the same colony using 36 microarrays (a third behavioral group was included in the microarray design but was not analyzed in the present study).

Surprisingly, hive-restricted bees were almost indistinguishable from foragers. Only 16 genes showed significant differences between foragers and hive-restricted bees in a direction consistent with differences between foragers and normal hive bees. Only 11 and 3 ($P < 0.05$ and 0.001) of the 100 behavioral marker genes showed significant effects of foraging experience. Our failure to detect more extensive differences between hive-restricted bees and foragers was not because of lack of statistical power, which was comparable with ref. 13 (data not shown). These results indicate that the vast majority of thousands of hive bee–forager differences in brain gene expression observed to date do not depend on flight, light, or other foraging-related stimuli or experience.

The genes affected by foraging experience and the genes affected by JH appear to be more or less distinct. We divided the 100 behavior marker genes into three classes: those regulated by methoprene (47 at $P < 0.05$), those not regulated by methoprene (26 at $P > 0.2$), and the remaining marginally significant set ($0.2 \geq P \geq 0.05$). For the subset regulated by methoprene, forager/hive-restricted ratios were not correlated with forager/nurse ratios ($r = -0.08$; $P = 0.58$; Fig. 2*A*), indicating little or no effect of experience. In contrast, the subset not regulated by methoprene exhibited a strong correlation between these ratios ($r = 0.63$; $P = 0.00062$; Fig. 2*B*), indicating that this subset of genes was strongly influenced by flight, light, or other foraging-related stimuli or experience.

Experiment 4: Subspecies Differences in JH-Target Gene Expression and Circulating JH. The results of experiment 1 led to the hypothesis that the earlier onset age of foraging in *ligustica* is related to increased activity in some forager-associated signaling pathway in ≥8-day-old bees. Additional statistical analyses from experiments 1 and 2 support this hypothesis. We performed rank-correlation analyses between PC3 from experiment 1 (which was associated with subspecies differences; see Fig. 1*C*) and treatment effects from experiment 2. PC3 was significantly correlated with methoprene treatment ($\rho = 0.14$; $P = 2.2 \times 10^{-16}$) but not with cGMP, cAMP, or cGMP vs. cAMP ($P > 0.05$) (Fig. 3*A*). The direction of this correlation indicated a positive relationship between *ligustica* trends (which are positively loaded on PC3 relative to *mellifera*; Fig. 1*C*) and methoprene-regulated gene expression in the brain. These results suggested that differences in brain gene expression between *ligustica* and *mellifera* may be related, at least in part, to differences in some aspect of JH regulation and predict either higher JH titers or higher gene expression response to JH in *ligustica*.

We tested the first of these predictions by comparing circulating titers of JH. The prediction was correct: 14-day-old *A. m. ligustica* cofostered with *mellifera* in either *ligustica* or *mellifera* colonies had significantly ($P = 0.0009$) higher JH titers than *mellifera* (Fig. 3*B*).

Functional Analysis of Honey Bee Brain Gene Expression with GO. We used GO (22) to look for biological processes that might be prominently associated with honey bee behavioral maturation. We examined the gene lists generated in experiments 1–3 (Table 1) for significant associations with specific GO functional categories in the following two ways. We looked for a "representational bias" across GO categories, i.e., a disproportionately high number of genes belonging to a GO category that showed significant regulation (either up or down) relative to the representation of that category on the entire array (Table 5, which is published as supporting information on the PNAS web site). We also looked for a "directional bias" within particular GO categories, i.e., representation of genes in a GO category that were disproportionately up- or down-regulated (Table 5; and see Fig. 7, which is published as supporting information on the PNAS web site). The directional bias tests yielded more extensive results, with significant biases for 87 of 255 GO categories tested (Fig. 7).

Here are a few examples of several GO categories showing

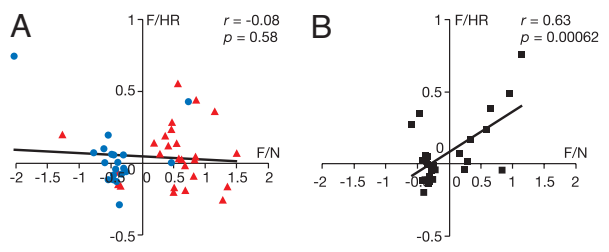

Fig. 2. Foraging experience and JH-target brain gene expression in honey bees. Gene expression ratios for forager/nurse bees from ref. 12 are plotted on the *x* axes; ratios for forager/hive-restricted bees from experiment 3 are plotted on the *y* axes ($\log_2$ values). (*A*) Behavior (nurse/forager) marker genes that were regulated by methoprene ($P < 0.05$; red triangle, up-regulated; blue square, down-regulated). (*B*) Behavior marker genes not regulated by methoprene ($P > 0.2$).

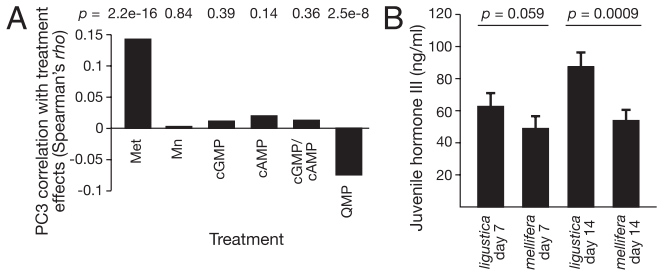

Fig. 3. Physiological basis for subspecies differences in brain gene expression. (*A*) Correlation between PC3 (gene loadings from experiment 1) and physiological treatment effects (experiment 2) (rank-correlation analyses, Spearman's ρ). (*B*) Age-related differences in circulating titers of JH between *ligustica* and *mellifera* bees.

directional bias (Table 5 and Fig. 7). In preforaging maturation, transcription genes ($n = 118$) were disproportionately up-regulated ($P = 8.1e{-}5$), whereas synaptic transmission, ($P = 5.7e{-}6, n = 59$) signal transduction ($P = 9.3e{-}5, n = 198$), and ion transport ($P = 3.4e{-}7$, $n = 86$) were disproportionately down-regulated. Fewer associations were observed for the hive-bee-to-forager transition; these included energy pathways ($P = 0.00080$, $n = 60$) and mitochondria ($P = 1.5e{-}7$, $n = 66$) (both disproportionately down-regulated). Subspecies differences showed no significant biases for any of the 255 GO categories examined.

GO analyses provide additional evidence that the treatments used in experiment 2 act on distinct sets of genes in the brain. For example, for methoprene-regulated genes, down-regulation directional biases were detected for cell communication, signal transduction, cell surface receptor-linked signal transduction, enzyme-linked receptor protein signaling pathway, and receptor activity (Fig. 7). Manganese shared none of these directional biases but had opposing effects on three of these categories and affected 38 other categories. cGMP and cAMP (relative to vehicle) were both commonly associated with only one category (down-regulation of protein folding), but cGMP relative to cAMP was associated with up-regulation of cell communication and regulation of metabolism (both in common with manganese) (Fig. 7).

Candidate Genes for Honey Bee Behavioral Maturation. The results of experiments 1–3 also enabled us to begin to identify specific genes that are candidates for involvement in honey bee behavioral maturation, particularly genes that could play causal roles in the hive-bee-to-forager transition. Although we cannot test causation directly with microarray data, we made three specific predictions that should be true for genes that do play a causal role, and we used these predictions to screen the lists of thousands of genes showing hive bee–forager differences (refs. 12 and 13 and this study). First, mRNA levels in the brain should be correlated robustly with behavior irrespective of age, genotype, or individual differences. Second, regulation of expression should be caused by known effectors of behavior. Third, regulation of expression should not be caused by flight or foraging activity. Subspecies differences that are consistent with the earlier onset of foraging in *ligustica* would provide additional correlative support.

Genes that meet prediction 1 were described in ref. 1; the 100 behavior marker genes used extensively here represent a portion of the genes that meet this prediction. Because rate of behavioral maturation is influenced by social factors (3–6), these behavior marker genes are also socially regulated. Fig. 4 shows a subset of these genes that have functional annotation, divided into sets that were dependent on or independent of foraging experience in experiment 3. Focusing on the experience-independent sets (prediction 3), 15 of these 100 genes were regulated by at least one of the treatments in experiment 2 in a direction consistent with the treatment effect on behavior (prediction 2); 14 of the 15 were regulated by methoprene. The group of genes that showed higher expression in foragers and up-regulation by methoprene included genes with putative roles in signal transduction (CG32703; a MAP kinase that is an ortholog of vertebrate ERK7/8; C.W.W., unpublished results), translation (CG11334), glutamate biosynthesis (CG7470), acid–base homeostasis (*CAH1*), and other functions. The converse class (higher in hive bees and down-regulated by methoprene) included genes with putative roles in cell adhesion (*BM-40-SPARC*), axonogenesis (*fax*), translation (*Ef2b*), and other functions. Four of 15 genes meeting predictions 1–3 also showed subspecies differences consistent with earlier foraging by *ligustica*. For example, CG32703 (ERK7/8) was expressed at higher levels in *ligustica* than *mellifera*, an observation that is consistent with regulation of this gene by methoprene and higher JH titers in *ligustica*. These genes are prime candidates for further analysis.

Gene	Putative function	Hive bee-to-forager	Experience dependent	Pre-foraging maturation	Genotype	Methoprene	Mn	cGMP	cAMP	cGMP vs cAMP	QMP
Forager markers; experience independent											
CG32703	MAP kinase (ERK7)	F***	n.s.	U***	L**	U***	U*	n.s.	D*	U*	n.s.
Inos	inositol-3-phosphate	F***	n.s.	U***	n.s.	n.s.	U*	D**	D*	D*	D*
CG3808 (2.7e-19)	RNA methyltransferase	F***	n.s.	U***	n.s.	U***	D***	U*	U**	n.s.	n.s.
CG30387	receptor signaling protein	F***	n.s.	U***	L*	U*	n.s.	n.s.	n.s.	n.s.	n.s.
CG11334	translation initiation factor	F***	n.s.	U***	L*	U**	n.s.	n.s.	n.s.	n.s.	D*
U2af50	RNA splicing factor	F***	n.s.	U***	M*	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
CG32354	endopeptidase inhibitor	F**	n.s.	U**	L**	U**	n.s.	D**	D*	D*	n.s.
CG6910	oxidoreductase activity	F***	n.s.	U**	M*	U***	D***	n.s.	U*	D*	U*
CG7470	glutamate 5-kinase	F***	n.s.	U***	L*	U***	D***	U*	U**	D**	n.s.
CAH1	carbonate dehydratase	F***	n.s.	U***	L***	U***	U*	n.s.	D*	U**	D*
CG8271	monocarboxylate porter	F***	n.s.	D*	n.s.	U**	n.s.	U*	n.s.	U*	n.s.
Forager markers; experience dependent (significant or marginally significant)											
CG6454 (1.0e-100)	unknown	F***	**	U*	n.s.	n.s.	U**	U*	n.s.	U*	n.s.
(GB10722)	unknown	F***	**	U***	M*	U*	n.s.	n.s.	U*	D*	n.s.
Hsc70Cb	protein folding	F***	p<0.1	D*	n.s.	D**	U**	D***	D**	D*	n.s.
CG18292	kinase regulator	F***	p<0.1	U*	n.s.	n.s.	n.s.	D***	n.s.	D***	n.s.
Nurse markers; experience independent											
Mlc-c	cytoskeletal protein binding	HB***	n.s.	D***	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
BM-40-SPARC	celladhesion; growth factor	HB***	n.s.	D***	n.s.	D**	D*	n.s.	n.s.	D*	n.s.
CG12163	cysteine protease inhibitor	HB***	n.s.	n.s.	M**	D*	D*	n.s.	n.s.	n.s.	n.s.
CG7442	organic cation porter	HB***	n.s.	D***	M**	D**	D*	n.s.	n.s.	n.s.	n.s.
CG11314	mesoderm development	HB***	n.s.	D*	L***	n.s.	n.s.	D*	n.s.	D*	n.s.
zormin	structural constituent of cytoskeleton	HB**	n.s.	n.s.	n.s.	n.s.	n.s.	U*	U*	n.s.	n.s.
SmD3	small nucleolar ribonucleoprotein complex	HB**	n.s.	n.s.	n.s.	n.s.	D*	U**	n.s.	U*	n.s.
Tctp	transposase activity	HB***	n.s.	n.s.	n.s.	D***	n.s.	n.s.	n.s.	n.s.	n.s.
fax	axonogenesis, transmission of nerve impulse	HB***	n.s.	D***	L**	D***	n.s.	D*	n.s.	n.s.	n.s.
Orc1	DNA replication origin binding	HB***	n.s.	U**	M*	n.d.	n.d.	n.d.	n.d.	n.d.	U*
Ef2b	translation elongation factor	HB***	n.s.	U*	n.s.	D**	n.s.	n.s.	n.s.	n.s.	n.s.
eIF-4a (9e-140)	translation initiation factor	HB***	n.s.	U*	L*	n.s.	D*	n.s.	n.s.	n.s.	n.s.
Sh3ß	metal ion binding	HB***	n.s.	D**	M*	D**	D**	n.s.	n.s.	n.s.	n.s.
PebIII	carrier activity	HB***	n.s.	n.s.	L*	U***	D**	n.s.	U***	D***	n.s.
Nurse markers; experience dependent (significant or marginally significant) or not determined											
tun (2e-62)	olfactory learning	HB***	*	D**	L**	D**	n.s.	D*	n.s.	n.s.	n.s.
RfaBp (6e-142)	retinol binding	HB***	p<0.1	D***	n.s.	D*	U*	n.s.	n.s.	n.s.	n.s.
Gasp (7e-26)	chitin binding	HB**	p<0.1	U*	L***	n.s.	n.s.	U**	U*	U*	n.s.
Mmp2 (1e-73)	metalloendopeptidase	HB***	p<0.1	D***	n.s.	U*	n.s.	U*	U*	n.s.	U*
CG10077 (1e-165)	RNA helicase	HB***	p<0.2	n.s.	L**	U*	D*	U**	U*	n.s.	n.s.
CG5418	hydrolase	HB*	n.d.	U***	M***	n.s.	n.s.	n.s.	n.s.	n.s.	U*
CG4572	serine carboxypeptidase	HB*	n.d.	U***	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
CG6688 (5e-39)	ubiquitin-protein ligase	HB*	n.d.	D**	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
CG10962 (2e-44)	oxidoreductase	HB*	n.d.	U***	n.s.	n.s.	D*	n.s.	n.s.	D*	U*
prc (3e-31)	extracellular matrix	HB*	n.d.	D*	n.s.	n.s.	n.s.	n.s.	U*	D*	n.s.

Fig. 4. Candidate genes for honey bee behavioral maturation. Shown are a subset of the 100 behavioral marker genes for which we replicated the findings in ref. 12 in the current study (hive-to-forager differences, $P < 0.05$) and which have functional (GO) annotation. Genes listed without parentheses are putative *D. melanogaster* orthologs based on reciprocal best BLAST match. Genes followed by parentheses are best *Drosophila* match (BLAST *e* value indicated). The gene listed in parentheses is a predicted gene with no *Drosophila* matches at BLAST $e < 10^{-5}$. The color and letter indicate the direction of regulation: red, up-regulated (*U*); blue, down-regulated (*D*); red, higher in forager (*F*); blue, hive bee (*HB*); red, *ligustica* (*L*); blue, *mellifera* (*M*). *, $P < 0.05$; **, $P < 0.001$; ***, $P < 1 \times 10^{-6}$; n.d., not determined; n.s., not significant. Marginally significant *P* values are indicated for experience-dependence. Statistical tests are from Table 1. A total of 15 of the 100 genes listed here met three of three predictions (see *Results*) for genes that could play causal roles in the hive-bee-to-forager transition.

Discussion

Our results demonstrate how a genomic approach can be combined with organismal biology, which, in this case, refers to knowledge about ontogenetic, genetic, physiological, and social components of bee behavior, to help gain insights into the molecular basis of social behavior. This successful dissection of brain gene expression indicates that, for social behavior, gene expression in the brain can provide a robust indicator of the interaction between hereditary and environmental information (23).

Our results, combined with those in refs. 1 and 2, reveal a robust molecular signature for division of labor in honey bee colonies, providing further evidence for a strong connection between brain

gene expression and plasticity in naturally occurring behavior (12). Seeley (24) described four behaviorally distinct "temporal castes" in honey bee colonies that were associated with age, task, and task location. Our PCA revealed trends in brain gene expression that were related to these groups of bees. The first group of newly eclosed bees likely corresponds to Seeley's "cell cleaners" (the first temporal caste, which persists for ≈1 day) and represented the most different and discrete group in PCA (cluster *a* in Fig. 1*B*). Although our four hive-bee age groups (days 4, 8, 12, and 17) overlap in age with two of Seeley's temporal castes, their collection from the brood area at the center of the hive likely places all four age groups (clusters *b* and *c*) in Seeley's "hive center" rather than the subsequent "hive periphery" caste. Seeley's hive periphery caste was not represented in the current study, but results from another study (13) showed that this group (represented by comb builders, guards, and undertakers) also was distinct in brain gene expression profiles from both nurses (hive-center bees) and foragers. The fourth temporal caste described by Seeley consisted of foragers, which did form a distinct group (cluster *d*) in the present study. Thus, four worker groups derived from behavioral observation are distinguished by distinct gene expression profiles in the brain.

Although congruent with Seeley's observations, our findings suggest an interpretation that is only partly temporal. PCA revealed two independent trends in brain gene expression, one associated with age (preforaging maturation) and the other with behavior (hive-bee-to-forager transition). The first trend was essentially complete by 8 days of age and co-occurs temporally with striking structural and molecular changes in the brain (25, 26). Bees can begin foraging as early as 4–5 days of age (27), but most do not initiate foraging this early in life. This trend might reflect changes in brain gene expression associated with development of competence to forage, perform certain hive tasks, or both. Disproportionate up-regulation of genes involved in the frequency, rate, or extent of DNA-dependent transcription early in adulthood suggests that transcriptional mechanisms in the brain might be particularly important for one or more of these behavioral processes.

The second trend in brain gene expression involves bees that have completed the first maturational phase. Behavioral analyses indicate that these ≥8-day-old bees (Fig. 1*B*, cluster *c*) are totipotent; they can transition to the next behavioral stages and become foragers, as is typical, or remain nurses for weeks, if necessary (3). Behavioral transitions after an initial period of maturation may be age-neutral (28), with social and physiological factors influencing the probability of transition on any given day. As a result, the typical onset foraging at 2–3 weeks of age is not "hard wired" but rather an emergent effect of these factors. Consistent with this interpretation, the second trend is primarily associated with behavioral differences, not age, that are clearly related to physiological factors that govern the onset age of foraging.

Studies have shown that several physiological factors are involved in regulating the onset age of foraging in honey bees (7, 8, 20). Our results indicate that methoprene, manganese, and cGMP have very disparate effects on brain gene expression, suggesting multiple pathways. These pathways may be independent, may converge on the relatively small subset of genes that overlapped in response to these treatments, or may act to form a network of interlinked pathways, which might provide robust and flexible regulation in the face of ever-changing environmental and social conditions. Further dissection of treatment effects on brain gene expression into direct and indirect effects may help determine how these physiological factors interact to regulate behavior.

Bees deprived of foraging experience but treated with a JH analog showed forager-like expression profiles, suggesting that the increase in JH that influences the hive-bee–forager transition (20) may cause many of the brain gene expression changes that occur at this time. Extensive studies have demonstrated the role of JH in the regulation of honey bee behavioral maturation (20). JH titers are generally low in nurse bees and high in foragers, and they remain low in "overage" nurses but increase prematurely in "precocious" foragers. Removal of the glands that produce JH delays the onset of foraging, and this delay is eliminated with methoprene treatment (29). JH and vitellogenin are thought to act as mutual repressors in the hive-bee-to-forager transition (28, 30, 31). Although both appear to be key regulators in this process, it seems more probable that JH acts directly on gene expression in the brain.

Our study provides an example of how gene expression analysis can be used to learn about the physiological basis of genotypic differences. These findings strengthen the link (9) between genotypic differences in rate of behavioral maturation and JH titers and responsiveness to JH in honey bees. It was possible to detect this connection even though most *A. mellifera* subspecies differences in brain gene expression are probably unrelated to age at onset of foraging; these subspecies differ in many traits besides division of labor (10).

Experience-dependent changes in brain gene expression are well known, particularly for learning and memory (32). It was thus surprising that gene expression changes in the hive-to-forager transition were primarily experience-independent. Honey bee foraging is cognitively demanding and involves, at a minimum, learning the appearance and location of the hive, learning to navigate in the environment, and learning to extract food from different floral types. Foraging also causes changes in the volume of the neuropil of the mushroom bodies, a region of the brain involved in multimodal sensory integration and learning and memory, and these effects are mimicked by treatment with a cholinergic muscarinic receptor agonist (33). One possibility is that our analysis of whole brains, although sensitive to experience-expectant, hormone-driven, changes, is not sensitive to experience-dependent changes, because they involve more acute, localized changes.

With the possible exception of transcription (discussed above), the meaning of the observed GO directional biases in brain gene expression is unclear, especially those showing extensive down-regulation during preforaging maturation. One possible explanation is that this bias reflects the aftermath of a period of intense brain gene activity during the late pupal period and first few days of adulthood, a time marked by increases in dendritic arborization and presumed synaptogenesis (25). The relative lack of GO representational biases might be related to whole-brain analysis, lack of information on isoforms, or incomplete annotation (≈50% of the spots on this array represent annotated genes). The latter situation will improve substantially with the newly available honey bee genome sequence (17).

These analyses also have identified candidate genes for the regulation of behavioral maturation in honey bees. Especially promising are genes affected by both environmental (social) and hereditary factors. Some of these genes might be pacemakers (23), evolutionarily labile and mechanistically important, and thus of particular importance to an integrative understanding of division of labor in insect societies.

Methods

Animals. Field collections for experiments 1 and 4 were performed at the Laboratory of Bee Biology and Protection, Institut National de la Recherche Agronomique, Avignon, France, and experiments 2 and 3 were performed at the University of Illinois Bee Research Facility, Urbana, IL. In experiment 1, honey bee colonies were derived from two populations whose original sources were the subspecies *A. m. ligustica* and *A. m. mellifera*, based on their area of origin (Italy and Provence, France, respectively). Subspecies determinations were confirmed by allozyme analysis at the *malate dehydrogenase* locus (10) (data not shown). In experiments 2 and 3, colonies were derived from a mixture of European subspecies (predominantly *A. m. ligustica*). To obtain bees of known age, 1-day-old adult bees were obtained by transferring honeycomb frames containing pupae from typical colonies (source colonies) in the field to an incubator (34°C, 95% relative humidity). Bees that

emerged over a 24-h period were marked with a spot of paint (PLA; Testor, Rockford, IL) on the thorax and introduced either into an unrelated host colony or into cages in the laboratory. In experiment 1, each source colony was headed by a naturally mated queen, unrelated to the queens in all other experimental colonies, all approximately the same age. In experiments 2 and 3, each source colony was headed by a queen instrumentally inseminated with semen from a single drone. In experiment 4, we used synthetic queen (Bee Boost; Pherotech, Vancouver, BC, Canada) and brood pheromones [components purchased from Sigma (St. Louis, MO) as in ref. 6] to minimize variation in pheromone availability from live queens and brood, which can affect JH titers (5, 6).

Experiment 1: Age-Related, Behavior-Related, and Genetic Differences in Brain Gene Expression. *A. m. ligustica* and *A. m. mellifera* bees (≈400 each) were marked with a spot of paint on the thorax and cofostered in two typical field host colonies of similar size (≈40,000 adult bees, in two-story hives), one *ligustica* and one *mellifera*. In the absence of replicate host colonies from each subspecies, ANOVA tests (described below) treated host colony variation solely as effects of colony, rather than subspecies. Bees of each subspecies in each colony were sisters from a naturally mated queen (which are polyandrous; ref. 34); different sister groups were used in each colony, unrelated to the host colony. Bees were collected at eclosion (0- to 1-h-olds; newly eclosed bees), at 4, 8, 12, and 17 days of age from the center of the hive irrespective of behavior (hive bees), and as 16- or 17-day-old foragers (easily visible by pollen loads on hind legs or with a distended abdomen that was gently squeezed to test for nectar or water load). Three bees were collected for each combination of subspecies ($n = 2$), host colony ($n = 2$), and age/behavior group ($n = 6$), for a total of 72 bees. Collections were made at the same time of day to minimize circadian effects. Bees were immediately transferred into liquid nitrogen to prevent handling effects on brain gene expression. Marked foragers were destructively sampled at first observation of foraging to determine age at onset of foraging; as expected (10), *ligustica* showed an earlier onset age of foraging than *mellifera* (Fig. 5).

Experiment 2: Effects of Treatments That Influence Onset Age of Foraging on Brain Gene Expression. Groups of 50 1-day-old bees were marked (by treatment) and placed in a wooden cage ($6 \times 12 \times 18$ cm) in an incubator (34°C, 95% relative humidity). Bees were treated orally for 4 days with one of the following substances dissolved in 50% sucrose solution: 40 mg/ml JH analog methoprene (Wellmark International, Schaumburg, IL), 500 mM 8-Br-cGMP (membrane permeable; Sigma), 500 mM 8-Br-cAMP (Sigma), or 20 mM $MnCl_2$ (Sigma); control bees received sucrose alone. Methoprene, cGMP, and $MnCl_2$ administered in this way have been shown to cause precocious foraging in honey bees (7, 8, 35); cAMP treatment does not (7) and was included as a pharmacological control. Bees cannot survive >24 h without ingesting carbohydrates, so all surviving bees must have ingested the treatment. Feeding tubes containing treatment solutions were changed daily (under red light, invisible to bees). There were two cages per treatment. On day 5, all surviving bees from each cage were counted (90–100% survival); some ($n = 50$) were collected for brain gene expression analysis, and some ($n = 40$–50) were placed into a small double-cohort colony (7) of 1,000 bees (the rest were 1 day old) to determine that onset age of foraging was affected by treatment as in the above-referenced studies (data not shown).

Experiment 3: Effects of Flight and Foraging on Brain Gene Expression. Bees were confined to their colony with a previously established technique (21). We glued a plastic bead (1.5–2 mm high) on the dorsal surface of the thorax to increase its height; a screen placed inside the hive prevented these bees from leaving the hive but allowed other bees from the same colony to come and go freely. Hive-restricted bees were exposed to stimuli in the hive (i.e., nectar, pollen, wax, and their nestmates) but could not fly from the hive. We sampled hive-restricted bees that rushed toward the hive entrance when the screen was removed, apparently to attempt to forage. This exhibition of positive phototaxis was taken to mean that hive-restricted bees were presumptive foragers; nurses and other preforagers are typically negatively phototactic, whereas foragers are positively phototactic (36). This assumption was supported by the observation (C.W.W., data not shown) of previously hive-restricted bees returning with pollen loads within 3 h of being allowed to forage. We used a single-cohort colony (initially composed of bees all 1 day of age as in ref. 12), so we were able to collect hive-restricted bees and foragers at 10–11 days of age. [Precocious foraging occurs in single-cohort colonies because of a lack of inhibitory pheromone from older bees (4).] Hive-restricted bees were compared with returning foragers (unrestricted full sisters sampled at the same time). Three different full-sister groups were analyzed.

Experiment 4: Subspecies Differences in JH Titers. *A. m. ligustica* and *A. m. mellifera* bees were cofostered as in experiment 1, except in double-cohort colonies (7). Each colony was established with 1,200 bees: 200 foragers and 200 1-day-old bees of *ligustica*, *mellifera*, and *caucasica*. Bees were collected ($n = 10$) at 7 and 14 days of age from the hive irrespective of behavior. Collections were made at the same time of day (early in the morning, before foraging began) to minimize circadian effects. Bees were collected and immediately placed on ice for hemolymph sampling. Hemolymph samples were obtained and analyzed by using a chiral-specific RIA optimized to detect JH III [the only homolog of JH in honey bees (29)]. Onset age at foraging was determined (6) for other members of these cohorts to confirm that *ligustica* exhibited an earlier onset age at foraging than *mellifera* (Fig. 5), as expected (10).

Microarrays and Initial Data Processing. Methods were as in ref 12. Brains were dissected on dry ice, total RNA was extracted, and mRNA was amplified in a single round of T7 promoter-directed *in vitro* transcription. Filtering included removal of genes abundantly expressed in hypopharyngeal glands (a potential source of tissue contamination in our brain dissections) relative to brain. Intensity signals for cDNAs passing these filters were normalized for microarray position- and intensity-dependent biases by using Lowess smoothing [with the transform.madata function in the R/maanova 0.97–4 software package (19, 37); method = "rlowess"]. After normalization, we collapsed known redundant cDNA values based on gene predictions and annotation from the honey bee genome sequence (38). ESTs corresponding to microarray cDNAs were tested for near-perfect matches (98% identity) to coding (protein) sequence or to genomic sequence within or immediately downstream (500 bp) of predicted genes (using release 1 of the honey bee Official Gene Set; http://racerx00.tamu.edu/bee_resources.html). Redundant cDNA values were averaged (by using untransformed values), and resulting values were assigned to official gene names (which are all prefixed "GB"). Remaining cDNAs not associated with predicted sequences retained their EST identifiers and are presented here by EST accession number (prefixed "BI"). In experiment 1, a total of 6,705 cDNAs passed initial filters; 3,958 of these were collapsed to 2,989 nonredundant genes, and the 2,747 ESTs were unassigned to gene. We refer to the combined set of genes (2,989) and unassigned ESTs (2,747) as genes, although some redundancy is likely to be present in the remaining unassigned ESTs. (Exact numbers differed in the three experiments, but values presented for experiment 1 were typical.)

Microarray Experimental Design and Analysis. All microarray comparisons were direct without use of a common reference sample (Fig. 8, which is published as supporting information on the PNAS web site). This method allowed us to maximize statistical power for particular contrasts while minimizing the number of microarrays.

For example, each pair of sequential age groups in experiment 1 (e.g., 4-day-olds vs. 8-day-olds) was directly compared 12 times, each comparison within subspecies and within host colony. Additionally, 18 microarrays directly compared subspecies (within host colony and age group), and 18 directly compared host colony (within subspecies and age group). A total of 108 microarrays were used to analyze 72 individual bee brains in experiment 1. Individual brains were compared by using an analogous design for experiment 3 (Fig. 8). In experiment 2, brains from treatment groups were pooled and directly compared as indicated (Fig. 8). Resulting data were analyzed by using mixed-model ANOVA (18, 19). All statistical analyses were conducted in R using the R/maanova 0.97–4 software package (19, 37) (see Fig. 8 for statistical models).

PCA. For PCA of individual variation in experiment 1, we first derived gene expression levels for individual brains (I) using the fixed-effects model $y = \mu + A + D + I + \varepsilon$ [this model is blind to biological parameters (ontogeny, subspecies, and colony) tested above]. PCs, PC variances, and PC scores (gene loadings) were calculated from the singular value decomposition (38).

Derivation of 100 Nurse/Forager Behavior Marker Genes; Regression and Rank Correlational Analyses. Unprocessed microarray data from two studies (12, 15) were reanalyzed to provide nonredundant gene expression data comparable with expression data in the present study (experiments 1–3). The 100 behavior marker genes were derived from an independent (12) set of bee brains (30 nurses and 30 foragers) and reanalyzed (by using the same method as in ref. 12 to generate the top 50 set of "predictor" cDNAs); these were the best 100 genes on the microarray for classifying individual brain expression profiles as nurse or forager. Regression analyses across different microarray experiments were performed on $\log_2$-transformed ratios from each experiment (from ANOVA); only genes present in both experiments were analyzed. Spearman's rank-correlation analyses were performed by using PC3 gene loadings (experiment 1) and gene expression ratios for treatments vs. vehicle control (experiment 2).

GO Analyses. Predicted honey bee genes were assigned to orthology groups with *Drosophila melanogaster* genes on the basis of reciprocal best BLASTX match, and GO terms were assigned to bee genes based on annotation of *Drosophila* genes. GO functional terms, parent–child relationships between terms, and *Drosophila* gene GO annotations were downloaded from GO (www.geneontology.org/index.shtml) (all data were downloaded between July and August 2005). Counts of genes in specific categories were performed by using an Access database (Microsoft, Redmond, WA). χ^2 tests (with Yates' continuity correction) were performed in R.

We thank J.-M. Becard and K. Pruiett for technical assistance in the field; A. Stetler and J. Wermeling for brain dissections; T. Newman for assistance with microarrays; P. Kheradpour for microarray data processing scripts; C. Elsik for honey bee–*Drosophila* gene associations; K. A. Hughes for statistical advice; and D. F. Clayton, K. A. Hughes, R. Maleszka, and members of the Robinson and Whitfield laboratories for reviewing the manuscript. This work was supported by a National Science Foundation (NSF) Postdoctoral Fellowship in Bioinformatics (to C.W.W.) and by grants from NSF–National Institutes of Health (to S.R.-Z. and G.E.R.), Burroughs Wellcome (to G.E.R.), and NSF Frontiers in Biological Research (B. Schatz, principal investigator).

1. Robinson GE, Grozinger CM, Whitfield CW (2005) *Nat Rev Genet* 6:257–270.
2. Robinson GE (1992) *Annu Rev Entomol* 37:637–665.
3. Robinson GE (2002) *Am Nat* 160:S160–S172.
4. Leoncini I, Le Conte Y, Costagliola G, Plettner E, Toth AL, Wang M, Huang Z, Becard JM, Crauser D, Slessor KN, Robinson GE (2004) *Proc Natl Acad Sci USA* 101:17559–17564.
5. Pankiw T, Huang ZY, Winston ML, Robinson GE (1998) *J Insect Physiol* 44:685–692.
6. Le Conte Y, Mohammedi A, Robinson GE (2001) *Proc Biol Sci* 268: 163–168.
7. Ben-Shahar Y, Robichon A, Sokolowski MB, Robinson GE (2002) *Science* 296:741–744.
8. Ben-Shahar Y, Dudek NL, Robinson GE (2004) *J Exp Biol* 207:3281–3288.
9. Giray T, Huang Z-Y, Guzman-Novoa E, Robinson GE (1999) *Behav Ecol Sociobiol* 47:17–28.
10. Brillet C, Robinson GE, Bues R, Le Conte Y (2002) *Ethology* 108:115–126.
11. Page RE, Jr, Scheiner R, Erber J, Amdam GV (2006) *Curr Top Dev Biol* 74:253–286.
12. Whitfield CW, Cziko AM, Robinson GE (2003) *Science* 302:296–299.
13. Cash AC, Whitfield CW, Ismail N, Robinson GE (2005) *Genes Brain Behav* 4:267–271.
14. Kucharski R, Maleszka R (2002) *Genome Biol* 3:RESEARCH0007.
15. Grozinger CM, Sharabash NM, Whitfield CW, Robinson GE (2003) *Proc Natl Acad Sci USA* 100 (Suppl 2):14519–14525.
16. Whitfield CW, Band MR, Bonaldo MF, Kumar CG, Liu L, Pardinas JR, Robertson HM, Soares MB, Robinson GE (2002) *Genome Res* 12: 555–566.
17. Honeybee Genome Sequencing Consortium (2006) *Nature* 443:931–949.
18. Wolfinger RD, Gibson G, Wolfinger ED, Bennett L, Hamadeh H, Bushel P, Afshari C, Paules RS (2001) *J Comput Biol* 8:625–637.
19. Cui X, Churchill G (2003) *Genome Biol* 4:210.
20. Bloch G, Wheeler DL, Robinson GE (2002) in *Hormones and Behavior*, eds Pfaff DW, Arnold AP, Etgen AM, Fahrbach SE, Rubin RT (Academic, New York), Vol 3, pp 195–236.
21. Withers GS, Fahrbach SE, Robinson GE (1995) *J Neurobiol* 26:130–144.
22. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, *et al.* (2000) *Nat Genet* 25:25–29.
23. Robinson GE (2004) *Science* 304:397–399.
24. Seeley TD (1982) *Behav Ecol Sociobiol* 11:287–293.
25. Farris SM, Robinson GE, Fahrbach SE (2001) *J Neurosci* 21:6395–6404.
26. Guez D, Belzunces LP, Maleszka R (2003) *Pharmacol Biochem Behav* 75:217–222.
27. Schulz DJ, Huang Z-Y, Robinson GE (1998) *Behav Ecol Sociobiol* 42:295–303.
28. Amdam GV, Omholt SW (2003) *J Theor Biol* 223:451–464.
29. Sullivan JP, Fahrbach SE, Robinson GE (2000) *Horm Behav* 37:1–14.
30. Rutz W, Gerig L, Wille H, Luscher M (1976) *J Insect Physiol* 22:1485–1491.
31. Guidugli KR, Nascimento AM, Amdam GV, Barchuk AR, Omholt S, Simoes ZL, Hartfelder K (2005) *FEBS Lett* 579:4961–4965.
32. Bradley J, Finkbeiner S (2002) *Prog Neurobiol* 67:469.
33. Ismail N, Robinson GE, Fahrbach SE (2006) *Proc Natl Acad Sci USA* 103:207–211.
34. Page RE (1986) *Annu Rev Entomol* 31:297–320.
35. Schulz DJ, Sullivan JP, Robinson GE (2002) *Horm Behav* 42:222–231.
36. Ben-Shahar Y, Leung HT, Pak WL, Sokolowski MB, Robinson GE (2003) *J Exp Biol* 206:2507–2515.
37. Wu H, Kerr MK, Cui X, Churchill GA (2003) in *The Analysis of Gene Expression Data: Methods and Software*, eds Parmigiani G, Garrett ES, Irizarry RA, Zeger SL (Springer, New York).
38. Wall ME, Rechtsteiner A, Rocha LM (2003) in *A Practical Approach to Microarray Data Analysis*, eds Berrar DP, Dubitzky W, Granzow M (Kluwer, Norwell, MA), pp 91–109.