Supporting Information

Alaux et al 10.1073/pnas.0907043106

SI Text

Soldier Trap. To obtain soldier bees, a net trap fitting the dimensions of the beehive with 4 triangular pieces of white wood (two 55.5 × 90 cm and two 45.5 × 90 cm) was placed on top of the opened hive. After stimulation, the trap was removed and the hive-top closed to prevent the soldiers inside it from escaping. The trap was then placed inside a freezer at −18 °C for 5 min.

cis*-Regulatory Analysis.** ***Determining genes targeted by a transcription factor motif. We masked exons (of adjacent genes) and short tandem repeats of period ≤5 bp [Tandem Repeat Finder (1), command-line arguments 2 3 5 80 10 25 5 –m –d] present in the promoter, and removed all masked positions, thus obtaining promoters of variable lengths (≤5 Kbp). We "scanned" each resulting promoter with a sliding window of 500-bp length, in shifts of 250 bp. Each window was scored for presence of the motif by using the Hidden Markov Model-based approach used in ref. 2. This raw score was converted to an empirical P value based on raw scores of randomly generated windows of the same length and guanine-cytosine content. Each promoter was thus associated with a list of "score P values" of varying sizes because the masked promoters are of varying lengths and are covered by varying number of 500-bp windows. Let k be the number of window P values associated with a gene's promoter, and let β be a significance threshold (e.g., 0.05) at which we wish to associate the gene with the motif. We computed $\alpha = 1 - (1 - \beta)^{1/k}$ and used this as a threshold on the window-score P value. If any window has score P value below this threshold α, then the corresponding promoter (and gene) is designated as a target of the motif.

Associating motifs with gene sets. Each motif was tested for association with differentially regulated gene sets in each experiment, as follows: Let U_T and D_T be the sets of genes up- and down-regulated in experiment T (i.e., alarm pheromone, old, or soldier). Each set was partitioned into genes that are targeted by the motif (as determined above, with $\beta = 0.01$) and genes that are not. A Fisher exact test was performed on the 2 × 2 contingency table (up vs. down, motif target vs. nontarget), and the 1-tailed P value was the statistical measure of association between the gene set U_T and the motif. By repeating the same test for the 2 × 2 contingency table: (down vs. up, motif target vs. nontarget), we obtained a P value for the association between the gene set D_T and the motif.

Motifs associated consistently with genes influenced by heredity and environment. A motif was tested for consistent association with up-regulated gene sets from all three experiments as follows. (The same procedure was repeated with down-regulated gene sets.) Let P_1, P_2, and P_3 be the 1-tailed P values of association of the motif with the up-regulated gene sets in the three experiments, respectively. We considered the combined measure

$$P_c = 1 - (1 - P_1)\,(1 - P_2)\,(1 - P_3).$$

We denote this random variable as P_c and its observed value as $P_c^{(o)}$ Under the null hypothesis that each of P_1, P_2, P_3 is uniformly distributed, we computed the probability π_c that the combined measure P_c has a value less than or equal to the observed value $P_c^{(o)}$, i.e., a 1-tailed P value for the combined measure P_c, as follows:

$$\pi_c(x) = \Pr(P_c \le x) = \Pr\left(\prod_j (1 - P_j) \ge 1 - x\right)$$

$$= 1 - \Pr\left(\prod_j (1 - P_j) \le 1 - x\right)$$

$$= 1 - (1 - x)\sum_{i=0}^{2} (-1)^i \frac{(\ln(1 - x))^i}{i!},$$

where the last step is due to the fact that $1 - P_j$ is uniformly distributed in [0,1] under the null hypothesis.

The random variable P_c has the desirable property that it is low only if each of P_1, P_2, P_3 is low, and thus captures consistent motif association (low P value) in all three experiments. (Contrast this with the product of the three P_j's, which may be low even if one or two of the P_j's is close to 1.) However, the P value π_c computed above corresponds to the strong null hypothesis that every individual P_j is uniformly distributed. We therefore empirically estimated the appropriate thresholds to use for this P value, as described below.

Multiple hypothesis correction. We simulated random gene sets of the same size as the original gene sets in each experiment, obtained the π_c for each motif, calculated how many motifs have a π_c below a specific threshold τ, and computed an average of this number over 1,000 independent simulations. This is the empirical expectation of the number of tests with π_c below τ, and was used to estimate a false discovery rate (FDR) corresponding to each $\pi_c(P_c^{(o)})$ in the tests with the original gene sets. Only motifs with estimated FDR < 0.001 are reported in Table 5.

Extreme value distribution. We considered the minimum $\pi_c(P_c^{(o)})$ over all motifs, representing the extent of the strongest consistent motif association. We then computed the corresponding minimum value from each of the 1,000 random simulations described above, and the histogram thus obtained provides us with an estimate of the Extreme Value Distribution of $\pi_c(P_c^{(o)})$. This allowed us to estimate an empirical P value for the minimum $\pi_c(P_c^{(o)})$.

mRNA Quantification by Real-Time Quantitative RT-PCR (qRT-PCR). We determined whether alarm pheromone exposure induces an up-regulation of the transcription factor *c-Jun* in antennal lobes (ALs) (Fig. S1). Seven control and pheromone-exposed workers per colony were analyzed. We also verified the expression levels of 4 genes (*Uvop*, *Gβe*, *trp*, and *TpnCI*) in whole brains of bees from two Africanized honey bee (AHB) colonies compared with bees from two European honey bee (EHB) colonies (Fig. S3). Five to seven workers (mainly AHB workers) per colony were analyzed. Because the expression levels of the tested genes did not differ significantly between the two AHB colonies and between the two EHB colonies (1-tailed permutation test, $P = 0.18$ for each test), results were pooled for analysis.

Whole brains and ALs were dissected, frozen, and then homogenized in TRIzol (Invitrogen Life Technologies) before extracting RNA with the Qiagen RNeasy kit for total RNA with on-column DNase I treatment (Qiagen). Forty nanograms from the two ALs and 200 ng from whole brains were reverse-transcribed with random hexamers by using the RETROscript kit (Ambion). mRNA quantification was performed using an ABI Prism 7900 sequence detector and the SYBR green detection method (Applied Biosystems). *c-Jun* mRNA levels were normalized to the *rp49* "housekeeping" gene as control (expression level did not vary: $P = 0.64$ for each test) and those of *Uvop*,

$G\beta_e$, *trp*, and *TpnCI* were normalized to the external control *rcp1* (expression level did not vary; $P = 0.11$ for each test). A list of primers is given in Table S5.

Measurements of Mitochondrial Enzyme Activities. Activity of NADH:ubiquinone oxidoreductase was assayed as previously described (3, 4). Briefly, brain mitochondria were lysed by freeze-thawing 3 times in hypotonic buffer (25 mM KH_2PO_4, pH 7.2, 5 mM $MgCl_2$) and added to an assay buffer consisting of 65 μM ubiquinone$_1$, 130 μM NADH, 2 μg/mL antimycin A, and 2.5 mg/mL BSA in the presence or absence of rotenone. Oxidation of NADH was measured at 340 nm for 5 min at 30 °C by using a SpectraMax M2 microplate reader (Molecular Devices), after which 5 μg/mL rotenone was added to the reaction. Activity was then measured for another 5 min and the difference in rate before and after treatment with rotenone was used to determine complex I activity. The molar absorbance coefficient used for NADH was 6.22 $mM^{-1}cm^{-1}$.

Complex IV activity was measured by using a Cytochrome *c* Oxidase Assay kit (Sigma) as per manufacturer instructions. The decrease in absorption at 550 nm, representing oxidation of ferrocytochrome *c*, was monitored over time on a SpectroMax M2 microplate reader (Molecular Devices) at room temperature. Activity was then calculated by using the following formula: $(\Delta absorbance_{550\ nm}/min \times 28.8)/40 \times 21.84$, where 21.84 represents the $\Delta\varepsilon^{mM}$ between ferrocytochrome *c* and ferricytochrome *c* at 550 nm, and 28.8 represents the dilution factor multiplied by the total reaction volume.

Complex V (ATP synthase) activity was determined by measuring the decrease in NADH absorption at 340 nm in the presence or absence of 2 μg/mL oligomycin, a specific inhibitor of the F_1F_0 complex of ATPase, at 30 °C as previously described (5). Mitochondria were freeze-thawed as before and added to an assay medium consisting of 4 mM ATP, 25 mM Tris-Cl (pH 7.4), 25 mM KCl, 5 mM $MgCl_2$, 200 μM NADH, 1.5 mM phosphoenolpyruvate, 5 units lactate dehydrogenase, 3 units pyruvate kinase, and 2.5 μM rotenone.

In addition, assays of ATP production in the presence or absence of 5 μg/mL rotenone (complex I inhibitor), 0.25 mM KCl (complex IV inhibitor), or 2 μg/mL oligomycin (complex V inhibitor) were also performed by using an ATPlite Luminescence Assay System (Perkin-Elmer). Freeze-thawed mitochondria were prepared as above and the lysates incubated at 30 °C in a reaction buffer containing 65 μM ubiquinone$_1$, 130 μM NADH, and 2.5 mg/mL BSA. Aliquots were removed every minute and processed according to manufacturer instructions, and luminescence was detected by using a SpectroMax M2 microplate reader.

All measurements were corrected for background and crude protein concentration and analyzed by using 1-way ANOVA with Tukey posthoc testing.

Functional analysis. The list of GO categories from Fig. 2 is shown below:

1. Multicellular organismal process.
2. Response to stimulus.
3. Metabolic process.
4. Biological regulation.
5. Binding.
6. Electron carrier activity.
7. Catalytic activity.
8. Structural molecule activity.
9. Transmembrane transporter activity.
10. System process.
11. Detection of stimulus.
12. Cellular metabolic process.
13. Oxidation reduction.
14. Regulation of biological process.
15. Protein binding.
16. Isomerase activity.
17. Oxidoreductase activity.
18. Structural constituent of ribosome.
19. Substrate-specific transmembrane transporter activity.
20. Neurological system process.
21. Detection of external stimulus.
22. Alcohol metabolic process.
23. Regulation of cellular process.
24. Intramolecular lyase activity.
25. Ion transmembrane transporter activity.
26. Sensory perception.
27. Alcohol biosynthetic process.
28. Regulation of cell cycle.
29. Inositol-3-phosphate synthase activity.
30. Cation transmembrane transporter activity.
31. Sensory perception of light stimulus.
32. Polyol biosynthetic process.
33. Regulation of S phase.
34. Inorganic cation transmembrane transporter activity.
35. Visual perception.
36. Inositol biosynthetic process.
37. Regulation of S phase of mitotic cell cycle.
38. Monovalent inorganic cation transmembrane transporter activity.
39. Positive regulation of S phase of mitotic cell cycle.
40. Hydrogen ion transmembrane transporter activity.

1. Benson G (1999) Tandem repeats finder: A program to analyze DNA sequences. *Nucleic Acids Res* 27:573–580.
2. Alaux C, et al. (2009) Regulation of brain gene expression in honey bees by brood pheromone. *Genes Brain Behav* 8:309–319.
3. Birch-Machin MA, Turnbull DM (2001) Assaying mitochondrial respiratory complex activity in mitochondria isolated from human cells and tissues. *Methods Cell Biol* 65:97–117.
4. Tieu K, et al. (2003) D-beta-hydroxybutyrate rescues mitochondrial respiration and mitigates features of Parkinson disease. *J Clin Invest* 112:892–901.
5. Bosetti F, Yu G, Zucchi R, Ronca-Testoni S, Solaini G (2000) Myocardial ischemic preconditioning and mitochondrial F1F0-ATPase activity. *Mol Cell Biochem* 215:31–37.
6. Hunt GJ, et al. (2007) Behavioral genomics of honeybee foraging and nest defense. *Naturwissenschaften* 94:247–267.
7. Kanehisa M, Goto S (2000) KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* 28:27–30.

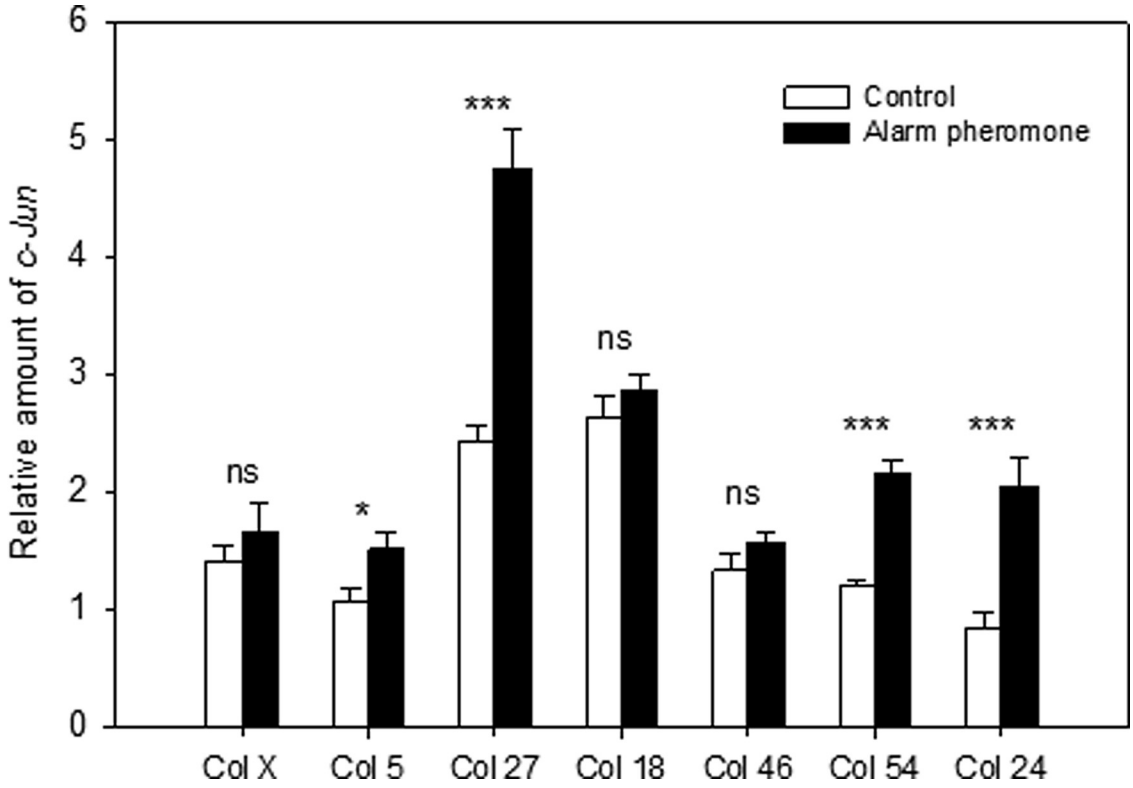


Fig. S1. Effect of alarm pheromone exposure on *c-Jun* responses. Alarm pheromone exposure caused up-regulation of *c-Jun* in the antennal lobes of bees in 4 of 7 colonies. Data normalized to expression levels of *rp49*. Means ± SE shown. Significant differences were determined by using a 1-tailed permutation test (*, $P < 0.05$; ***, $P < 0.001$).

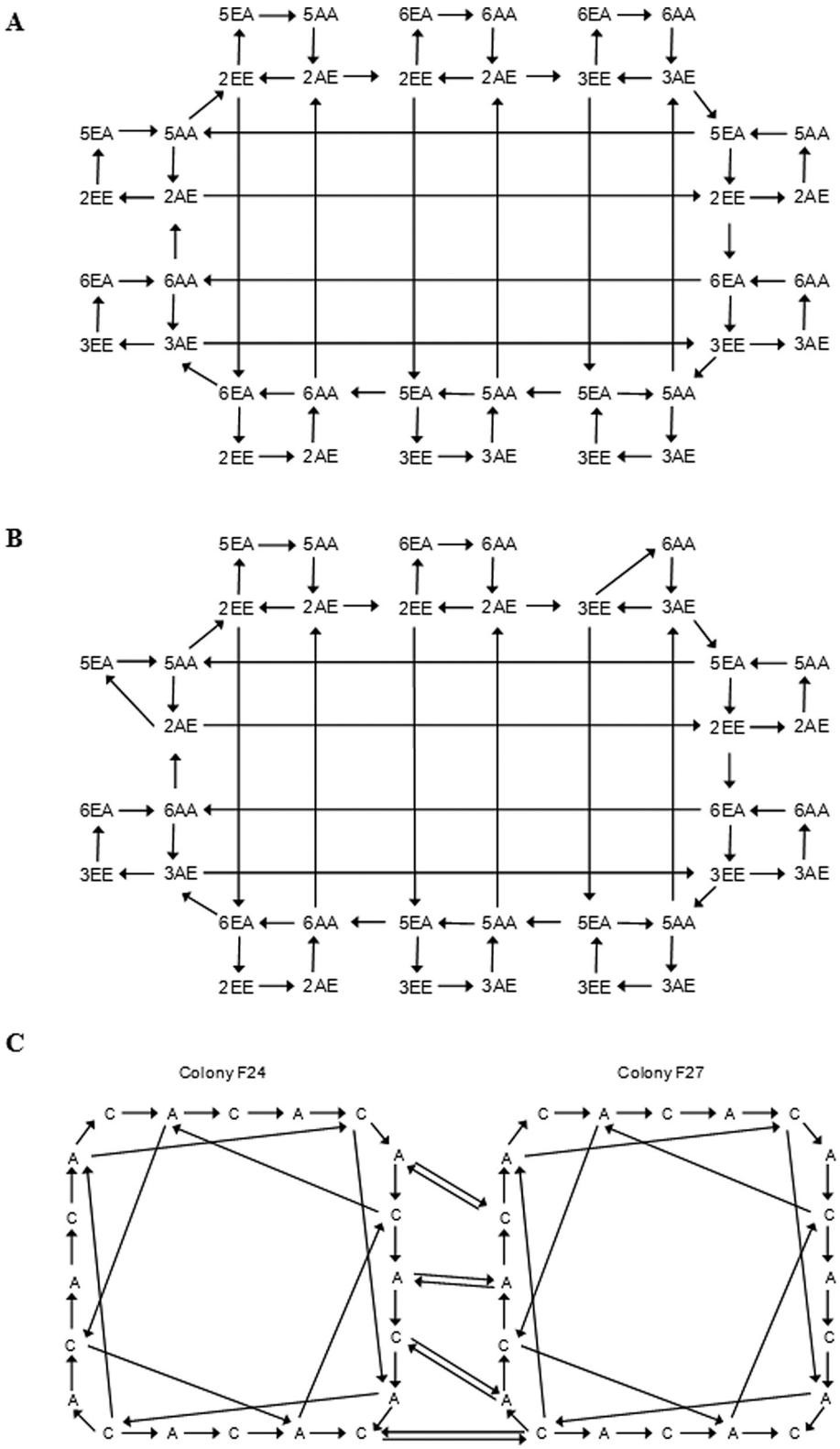
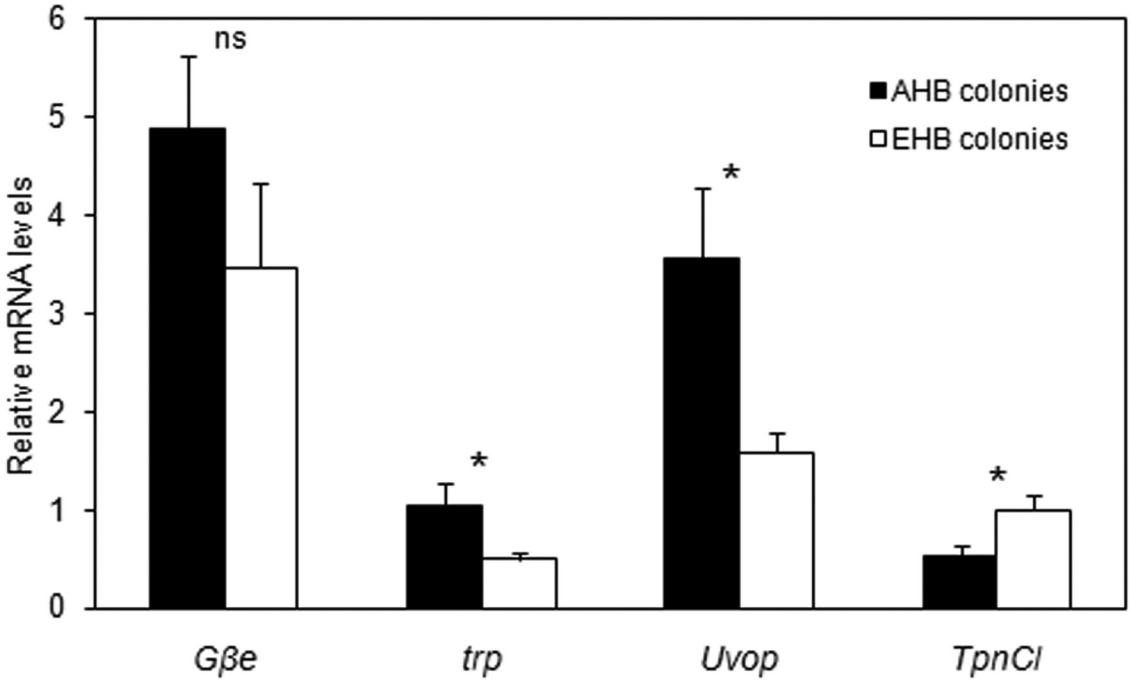


Fig. S2. Microarray experimental designs. Arrow tail indicates Cy3-labeled sample and arrow head indicates Cy5-labeled sample. Each sample was labeled an equal number of times with Cy5 and Cy3. (*A*) AHB/EHB experimental design for guards and foragers. 2E and 3E, EHB colonies; 5A and 6A, AHB colonies; last letter indicates worker genotype; 60 arrays for each behavioral group. (*B*) AHB/EHB experimental design for soldiers. Same as in *A* except that in colonies 2E and 6A, four EHB soldiers were analyzed, 58 arrays. (*C*) Alarm pheromone experimental design. C, control bees; A, bees exposed to alarm pheromone; bees from 2 colonies, 64 arrays.

A



B

Honeybee annotations	Gene name	Microarrays (AHB/EHB)	qRT-(AHB/EHB)
GB10545	*Gβe*	1.5	1.4
GB16264	*trp*	1.4	2.1
GB18171	*Uvop*	1.4	2.2
GB10545	*TpnCl*	0.8	0.5

Fig. S3. Validation of microarray results with real-time qRT-PCR. (*A*) Brain expression levels of four genes identified from microarray analysis as showing different expression levels between AHB vs. EHB. Data normalized to expression levels of *rcp1*. (*B*) Ratio of mean expression levels (AHB/EHB) from array and qRT-PCR results are shown. qRT-PCR samples were from independent biological replicates, not used for microarray analysis. Means ± SE shown. Significant differences were determined by using a 1-tailed permutation test (∗, $P < 0.05$).

Table S1. Aggression-related brain gene regulation in honey bees: Number of genes overlapping between gene sets regardless of direction

	Old bee				AHB guard				AHB soldier				AHB forager			
Experiment	Expected	Observed no.	*RF*	*P* value	Expected	Observed no.	*RF*	*P* value	Expected	Observed no.	*RF*	*P* value	Expected	Observed no.	*RF*	*P* value
Alarm pheromone	55.3	151	2.7	<0.0001	10	24	2.4	<0.0001	21.8	55	2.5	<0.0001	2.3	7	3	<0.008
Old bee					28.8	54	1.9	<0.0001	63.7	147	2.3	<0.0001	6.8	21	8.1	<0.0001
AHB guard (I)									11.6	81	7	<0.0001	1.2	24	19.5	<0.0001
AHB soldier (I)													2.7	29	10.7	<0.0001
Alarm pheromone	55.3	151	2.7	<0.0001	19.8	63	3.2	<0.0001	33.6	64	1.9	<0.0001	13.9	27	1.9	<0.0001
Old bee					57.1	128	2.2	<0.0001	98.3	195	2	<0.0001	40.3	85	2.1	<0.0001
AHB guard (C)									35.4	155	4.4	<0.0001	14.6	91	6.2	<0.0001
AHB soldier (C)													24.9	106	4.3	<0.0001

Data are from the following: Experiment 1, AHB vs. EHB (guards, soldiers, and foragers); Experiment 2, alarm pheromone induced; and Experiment 3, old vs. young bees. Expected, the number of genes expected to overlap in all three experiments by chance alone; *RF*, representation factor; I, individual genotype; C, colony genotype. The statistical significance of the overlap between two lists was determined using an exact hypergeometric test and an extended version of the hypergeometric test for the overlap between three lists (see *SI Text*).

Table S2. Correlation analysis for genes differentially expressed as a function of genotype, alarm pheromone exposure, and age

	Alarm pheromone		Old bee	
Experiment	*r*	*P* value	*r*	*P* value
AHB Guard (I)	0.139	0.004	0.062	0.02
AHB Soldier (I)	0.229	<0.001	0.277	<0.001
AHB Forager (I)	0.161	<0.001	0.302	<0.001
AHB Guard (C)	0.207	<0.001	0.059	0.026
AHB Soldier (C)	0.189	<0.001	0.349	<0.001
AHB Forager (C)	0.357	<0.001	0.15	<0.001
Alarm pheromone	—	—	0.429	<0.001
Old bee	0.65	<0.001	—	—

Correlations were performed on expression values for genes regulated by alarm pheromone or differentially expressed between old and young individuals and the expression values of the same genes from the other lists that are listed. Log_2-transformed ratios were used. I, individual genotype; C, colony genotype.

Table S3. Genes regulated by alarm pheromone and/or differentially expressed between AHB/EHB, and found in independently derived aggression quantitative trait loci (6)

Bee genes	Fly orthologs	Molecular function	Biological process
Alarm pheromone			
GB10390 ↓	*Pros25*	Endopeptidase activity	ATP-dependent proteolysis; cell proliferation
GB18971 ↑	*l(1)G0232*	Rhodopsin-like receptor activity	G-protein-coupled receptor protein signaling pathway
GB12020 ↓	*ubl*	Unknown	Protein modification process
GB12504 ↑	*CkIIβ*	Casein kinase activity	Mushroom body development; circadian rhythm; signal transduction
GB15016 ↑	*Hsc70-3*	ATPase activity	Sleep; response to heat; defense response; protein folding; RNA interference
GB14502 ↑	*MED22*	RNA polymerase II transcription mediator activity	Regulation of transcription from RNA polymerase II promoter
GB13698 ↓	*Psf2*	Transcription elongation regulator activity	Positive regulation of S phase of mitotic cell cycle
AHB/EHB, individual genotype			
Guard			
GB19009 ↓	—	—	—
GB16258 ↑	—	—	—
Soldier			
GB19009 ↓	—	—	—
GB19676 ↑	*CG8165*	Regulation of transcription from RNA polymerase II promoter	Transcription from RNA polymerase II promoter
GB15582 ↑	*14-3-3ε*	Regulation of transcription, protein complexes, protein trafficking	Response to external stimulus; nonassociative learning
AHB/EHB, colony genotype			
Guard			
GB14502 ↓	*MED22*	RNA polymerase II transcription mediator activity	Regulation of transcription from RNA polymerase II promoter
GB11198 ↓	*CG34347*	Actin binding; structural constituent of cytoskeleton	Cytoskeleton organization and biogenesis
Soldier			
GB14824 ↑	*CG7675*	Oxidoreductase activity, acting on CH-OH group of donors	Metabolic process
GB18618 ↑	*htt*	Microtubule binding; protein binding	Axon cargo transport
GB19676 ↑	*CG8165*	Regulation of transcription from RNA polymerase II promoter	Transcription from RNA polymerase II promoter
GB19319 ↓	*RpI12*	Zinc ion binding; transcription regulator activity	Transcription from RNA polymerase I promoter
GB17741 ↓	*CG7849*	Unknown	Unknown
GB19681 ↓	*Ctr1A*	Copper ion transmembrane transporter activity	Copper ion transport
Forager			
GB14824 ↓	*CG7675*	Oxidoreductase activity, acting on CH-OH group of donors	Metabolic process
GB18557 ↑	*α-Spec*	Cytoskeletal protein binding	Synaptic transmission; regulation of cellular component organization and biogenesis

Arrows indicates up- or down-regulated by alarm pheromone and in AHB vs. EHB bees. Known functions of orthologs of these genes are shown. Quantitative trait loci were derived in a separate study based on an assay that measured propensity to sting in response to a disturbance to the colony (in the field), and involved different bees derived from crosses between AHB and EHB (6). —, unknown.

Table S4. Genes that were up- and down-regulated in all three experiments, i.e., responsive to heredity, alarm pheromone, and age, are shown with their fly and human orthologs

Bee oligo	Bee gene	Fly orthologs	Human orthologs
Up-regulated			
AM00762	—	—	—
AM01806	—	—	—
AM02192	—	—	—
AM02336	—		—
AM02847	GB10220	—	—
AM02986	GB10356	—	—
AM03461	GB10836	*Hsc70Cb*	*HSPA4* (*hsp70*)
AM03846	GB11223	*His3*:*CG33803*	*HIST2H3*
AM04192	GB11572	*Inos*	*AC008397.7* (*MIP synthase*)
AM04834	GB12215	*Droj2*	*DNAJA4*
AM05318	GB30506	*Gp93*	*HSP90B1*
AM05381	GB12766	*Arr2*	—
AM06103	GB13501	—	—
AM06766	GB14166	—	—
AM08435	GB15853	*Kap-α1*	*KPNA*
AM08741	GB16163	—	—
AM09086	GB16507	*CG1600*	—
AM10198	GB17624	*CHORD*	*CHORDC1*
AM10664	GB18094	*tj*	*MAF*
AM12432	GB19885	*CG1516*	*PC*
AM12790*	GB11487	*moody*	*MTNR1a*
Down-regulated			
AM02193	—	—	—
AM02582	—	—	—
AM03140	GB10513	*CG14286*	*C8orf33*
AM03851	GB11228	*His3.3A*	*H3F3B*
AM06007	GB13404	*Ts*	*TYMS*
AM06580	GB13982	*Faa*	*FAH*
AM07997*	GB15409	*Cyp6 g1*	—
AM08023	GB15436	*CG10674*	*C19orf56*
AM08856	GB16276	*SmB*	*SNRPB*
AM10511	GB17938	*CG1756*	—
AM10649	GB18078	*CG9646*	*C22orf9*
AM10750	GB18183	—	—
AM11126	GB18566	—	—
AM11255	GB18694	—	—
AM11556	GB19000	*Tim9a*	*TIMM9*
AM12010	GB19459	*Psf1*	*GINS1*
AM12536	GB19989	*CG4875*	*AC068533.6* (*RPC9*)

Asterisks indicate genes that were regulated at a slightly less stringent threshold in one experiment (*moody* was up-regulated in AHB foragers at FDR = 0.097 and *Cyp6 g1* was down-regulated by alarm pheromone at FDR = 0.057). —, unknown.

Table S5. Oxidative phosphorylation pathway downregulated in aggressive bees

	Bee gene	Fly orthologs	Alarm Pheromone	Old Bee	AHB Soldier (I)	AHB Soldier (C)
Complex I NADH-coenzyme Q oxidoreductase	GB15948	CG10320				
	GB18920	Pdsw				
	GB10916	CG9306				
	GB10474	CG12400				
	GB13526	CG3621				
	GB10859	CG2014				
	GB15438	CG15434				
	GB17095	CG9140				
	GB16917	CG12203				
	GB10406	CG6020				
	GB15102	mitochondrial acyl carrier protein 1				
Complex II Succinate-Q oxidoreductase	GB12875	CG6666				
Complex III Q-cytochrome c oxidoreductase	GB12164	CG30354				
	GB14417	CG14482				
	GB10344	Oxen				
	GB12510	CG4769				
Complex IV Cytochrome c oxidase	GB18028	CG14235				
	GB15238	CG9065				
	GB30388	cyclope				
	GB20012	CG10664				
	GB15816	Cytochrome c oxidase subunit Va				
	GB17614	CG2249				
Complex V ATP synthase	GB16494	ATPase coupling factor 6				
	GB18417	CG4692				
	GB15629	Oligomycin sensitivity-conferring protein				
	GB10989*	Vha68-2				
	GB16751	lethal (2) 06225				
	GB14791	bellwether				
	GB15291	ATP synthase-γ chain				

This was performed by using the DAVID annotation tool (http://david.abcc.ncifcrf.gov/) that presents information on molecular pathways from the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways (http://www.genome.jp/kegg/) (7). Asterisk indicates up-regulation. I, individual genotype; C, colony genotype. Colors indicate the three different experiments.

Table S6. Primer sequences

Primer	Sequence (5′ to 3′)
c-Jun, forward	CGTGGCGGCATCCAAA
c-Jun, reverse	CCCTTCAGCAATTTAACCTTATCTTC
Uvop, forward	CTGCCATCACAATTTGCTTCTTAT
Uvop, reverse	TGTAACACCAGGTGTTAAAAGTGCTT
$G\beta_e$, forward	TGCGCTCAGCATGGGTAAT
$G\beta_e$, reverse	TGTCCATACCACCACAAGCAA
trp, forward	GAAGTGTCACCGACTACGAGGAT
trp, reverse	CTGCCGCCGGTTTGG
TpnCl, forward	TTCTTCGCAAAGCGTTCGA
TpnCl, reverse	CAATCTGAGAATATCAGCCACCAT
rcp1, forward	TCAATTAACTCGGAATCGGA
rcp1, reverse	CCTGGATTTCCCTGCTGAT
rp49, forward	GGAACTGGAAGTTTTAATGATGCA
rp49, reverse	CAACAATGGATTTACGTTTTTTACTG

Other Supporting Information Files

Dataset S1 (XLS)
Dataset S2 (XLS)