Supporting Information

Rittschof et al. 10.1073/pnas.1420369111

SI Materials and Methods

Study Subjects and Behavioral Paradigms.

Mouse. Before the experiment, we housed male "resident" mice (C57BL/6J strain, received at 3 wk of age) in cages (one male per cage) along with one sexually viable female per resident (C3H/HeJ strain, received at 5 wk of age) for at least 1 wk. Males were 6 wk of age when the females were first introduced. We did not change the bedding in these cages during this housing period to retain the olfactory cues that contribute to territoriality in the resident mice. Two hours before testing, we removed female companion mice from resident cages. We randomly assigned males to either the experimental or the control treatment group ($n = 3$ males per treatment).

During the resident–intruder trial, we lowered the resident cage into a blank-walled experimental chamber designed to minimize external stimuli. This chamber was equipped with a top–down camera to record behavioral response. For the experimental animals, we then introduced an unfamiliar male "intruder" mouse (BALB/cJ strain), contained within a stainless steel wire mesh cage, into the cage of the resident. The wire mesh cage prevented males from making physical contact, thus preventing injury to the intruder mouse. For the control animals, a small paper cup housed within the wire mesh container was placed into the resident cage.

After 10 min, we removed the intruder (or the cup) and kept the resident mouse in a dark and quiet place for an additional 15 min. Following the holding period, we euthanized the resident mice, using cervical dislocation. Mouse work was performed with oversight from the University of Illinois Institutional Animal Care and Use committee (IACUC), under IACUC protocol 13358.

Stickleback. We collected males from the Navarro River, a freshwater population, and maintained them in the laboratory on a 16:8 (light:dark) photoperiod at 18 °C. Males were housed separately in 26.50-L tanks [36 cm (length) × 33 cm (width) × 24 cm (height)] and provided with nesting material including algae, sand, and gravel. Opaque dividers were inserted between males' tanks the night before the experiment to prevent visual interactions between neighbors.

At the time of the behavioral experiment, all males were in the "territorial" phase of the nesting cycle; i.e., they were defending a territory but had incomplete nests (1). We randomly assigned males to experimental and control treatment groups ($n = 3$ males per treatment). Males in the experimental treatment were exposed to a live, reproductively mature male intruder confined to a glass flask. Intruders were always smaller than the territorial holder. We presented males in the control treatment with an empty flask. After introducing the treatment objects into the males' tanks, we monitored males until they first oriented to the treatment object. We then observed male behaviors for 5 min following this first orientation to confirm aggression toward the intruding male in the experimental group. We removed the objects 5 min after first orientation.

We performed treatments in pairs (one experimental and one control animal). Thirty minutes after the first orientation, pairs were netted and quickly killed by decapitation within seconds following an IACUC-approved protocol (no. 06178) of the University of Illinois at Urbana–Champaign.

Honey bee. All experimental bees were female worker bees that were full sisters (offspring of a queen inseminated by a single drone), thus minimizing variation in genetic background. Honeycomb frames containing honey bee pupae were collected from a source colony and housed in a 34 °C incubator until adult emergence. One-day-old adult bees were collected, individually marked, and divided into 12 groups of 10. We arbitrarily assigned groups to experimental and control treatments ($n = 6$ groups of bees per treatment). The 10 bees were housed together in a 7.0 × 8.0 × 9.0-cm Plexiglas box with a small piece of honey comb. Bees were supplied honey and water ad libitum.

The intruder assay was modified from refs. 2 and 3. Assays were performed on groups of bees (10 bees per group, see above) in a ventilated room kept between 25 °C and 28 °C when bees were 7 d old. We performed treatments in pairs. For experimental groups, we introduced an unrelated intruder bee (a forager collected from a natural colony) to the container and monitored and recorded each individual bee's response to the intruder for 5 min following the first orientation toward the intruder. We measured aggressive behaviors, including lunging, biting, dragging the intruder, and attempting to sting. For the control treatment, we introduced a small piece of brass, slightly larger than the size of a bee, for 5 min. Although bees occasionally inspected the object, its presence did not otherwise alter their behavior. After 5 min we removed the intruder bee (or object) and then left the bees undisturbed for an additional 25 min to allow any event-related brain transcriptional changes to occur. We then flash froze all bees in liquid N_2. For RNA-seq analysis, we selected the experimental group member that showed the highest number of aggressive behaviors. A bee from each control group was selected at random.

RNA Sample Preparation.

Mouse. The whole brain was immediately removed from the skull, and a coronal section was used to separate the rhombencephalon from the rest of the brain. The coronal section began at the level of the optic chiasm, which delimits the anterior part of the hypothalamus, and passed through the anterior commissure. This divided the remaining tissue into a rostral portion (section A) and a caudal portion (section B). Section B was then split by coronal section approximately equally into a rostral (B1) and a caudal (B2) section. Sections B1 and B2 were laid rostral-side down and two 2.5-mm punches of the ventral hypothalamus (VH) were taken from the ventral surface of sections B1 and B2, one punch from each hemisphere. Two 2.5-mm punches were then taken immediately dorsal to the hypothalamus for bed nucleus of the stria terminalis (BNST). Finally, four 3-mm punches each were taken from the medial cortical areas of B1 and B2 for somatosensory cortex. Total RNA was isolated from dissected regions, using TRIzol Reagent (Life Technologies) according to the manufacturer's protocol. RNA was subsequently purified on columns with the RNA Clean and Concentrator-25 kit (Zymo Research).

Stickleback. Immediately following decapitation, we dissected out the diencephalon on dry ice by cutting the two lobes away from the cerebellum and removing the entire structure from the skull. The diencephalon was placed individually in Eppendorf tubes containing 500 μL of TRIzol Reagent (Life Technologies). Total RNA was isolated immediately using TRIzol Reagent according to the manufacturer's recommendation and subsequently purified on columns with the RNeasy kit (QIAGEN). RNA was eluted in a total volume of 30 μL in RNase-free water.

Honey bee. We dissected whole brains using RNA-later ICE (Life Technologies). We homogenized the brains and extracted nucleic acids using RNeasy kits (QIAGEN).

Samples for all three species were treated with DNase (QIAGEN) to remove genomic DNA during the extraction procedure. RNA quantity was assessed using a Nanodrop spectrophotometer (Thermoscientific), and RNA quality was assessed using the Agilent

Bioanalyzer 2100. RNA was immediately stored at −80 °C until used in sequencing library preparation.

RNA-seq Library Preparation. Poly-A RNA was enriched from 1–2 μg of total RNA by using Dynabeads Oligo$(dT)_{25}$ (Life Technologies), following the manufacturer's protocol. Two rounds of poly(A) enrichment were performed with a final elution in 14 μL of water. The poly-A–enriched RNA was used to prepare RNA-seq libraries, using the NEXTflex Directional RNA-seq Kit (dUTP based) with Illumina compatible adaptors (Bioo Scientific). Manufacturer's instructions were followed and 13–15 cycles of PCR amplification were performed depending on the starting input of total RNA. Libraries were quantified on a Qubit, using the dsDNA High Sensitivity Assay Kit (Life Technologies), and library size was assessed on a Bioanalyzer High Sensitivity DNA chip (Agilent). Libraries were pooled and diluted to a final concentration of 10 nM. All 12 bee libraries were pooled together. The three mouse experimental and control samples were pooled together, resulting in two total pools, and the three stickleback experimental and control samples were pooled together, resulting in two total pools. Final library pools were quantified using real-time PCR, using the Illumina compatible kit and standards (KAPA) by the W. M. Keck Center for Comparative and Functional Genomics at the Roy J. Carver Biotechnology Center (University of Illinois). Single-end sequencing was performed on an Illumina HiSEq 2500 instrument by the W. M. Keck Center for Comparative and Functional Genomics at the Roy J. Carver Biotechnology Center (University of Illinois). The one pool of 12 bee samples was sequenced on two lanes and each pool of mouse and stickleback samples was sequenced on one lane.

RNA-seq Data Processing.

Mouse. We aligned reads to the mouse reference genome (NCBI build 37.2 genome file), using TopHat2 (2.0.6) (4) and Bowtie2 (2.0.5) (5). Reads were assigned to features following the NCBI build 37.2 annotation file. RNA-seq produced an average of about 73 million reads per sample.

Stickleback. FastQC was used to assess read quality and the FastX toolkit was used to filter low-quality reads and residual adaptor sequences (hannonlab.cshl.edu/fastx_toolkit/). RNA-seq produced an average of 72 million reads per sample. We aligned reads to the *Gasterosteus aculeatus* reference genome (the repeat masked reference genome, Ensembl release 70), using TopHat (1.4.1) (6) and Bowtie (0.12.8) (7). Reads were assigned to features according to the Ensembl release 70 gene annotation file (ftp.ensembl.org/pub/release-70/gtf/gasterosteus_aculeatus/).

Honey bee. All honey bee samples were sequenced across two different lanes of the Illumina flow cell. RNA-seq produced an average of 18.5 million reads per sample per lane. We aligned reads to the current version of the *Apis mellifera* reference genome (v. 4.5) (8), using TopHat2 (2.0.6) (4) and Bowtie2 (2.0.5) (5). Reads were assigned to features according to the honey bee Official Gene Set (v. 3.2) (8).

For the TopHat alignments across all species, we designated library type as "first strand" and otherwise used default parameters. The outputs from TopHat were converted using SAMtools 0.1.18 (9) to calculate the number of reads per gene. We calculated reads per gene using the htseq-count function in the Python package HTSeq (0.5.4), using union mode, with strandedness specified (10). Because each honey bee sample was sequenced on two different lanes, counts per gene were pooled before the differential expression analysis.

Identifying differential expression. For each species, we assessed differential expression between experimental and control animals, using the "exactTest" function in the R software package EdgeR (11). This approach models the distribution of counts (per gene) across samples as a negative binomial distribution. Count data were normalized by library size and library composition (using the "calcNormFactors" function).

A gene was included in the differential expression analysis if the number of counts was greater than or equal to 1 count per million in any sample. We estimated dispersion across samples, using the "estimateCommonDisp" function. We retained all genes from the output of the EdgeR analysis instead of using EdgeR to designate a cutoff for significant differential expression. By retaining all genes we could perform a rank-based gene set enrichment analysis (below) and apply this approach consistently across all analyses (single-species and homologous triplet analyses, see below). We ranked all genes in the output on the basis of raw P value.

Homologous Triplet Analysis. Gene lists for each species were filtered to contain only homologous genes (*Materials and Methods*). The within-species gene lists were then assigned new rank-based P values before calculating the three-species combined significance score for each triplet of homologous genes. We reassigned P values because using raw P values could bias our results if (i) homologous genes have disproportionately low P values compared with other genes in the within-species analyses and (ii) species showed large differences in P values, such that very small P values in one species contribute disproportionately to the combined P-value calculation. Assigning new P values accounted for the first problem and minimized the effects of the second problem.

We created a homologous triplet for each EOG, which contained one representative gene from that EOG group per species. We selected as the representative gene for each species the gene that showed the lowest P value among all of the paralogs in an EOG. In cases where there was more than one gene in an EOG for a particular species, we adjusted the P value for the representative gene before calculating the combined significance score for each homologous triplet, using the formula

$$p_{\text{corr}} = 1 - (1 - p)^x,$$

where x is the number of paralogs in the EOG for that species. This procedure allowed us to account for the fact that EOGs with a large number of paralogs have a higher probability of a lower P value due to chance.

Once representative P values were corrected, we used Fisher's method in the R package MADAM (12) to calculate a combined P value for each triplet group of evolutionarily conserved genes, following the formula

$$S = -2\sum_{i=1}^{k} \ln p,$$

where S is the combined significance score and $k = 3$ (number of P values to be combined, one per species). We then ranked the groups of triplets on the basis of this combined significance score.

Gene Functional Analyses.

Gene annotations. We used mouse Gene Ontology (GO) annotations curated by the database Babelomics 4.3 (13). For stickleback and bee, we derived GO assignments, using protein family annotations from the database PANTHER (14). Bee and stickleback protein sequences were blasted against all genomes in the database (~82 genomes). This procedure assigns proteins to PANTHER families on the basis of structural information as well as phylogenetic information. Genes were then annotated using GO information derived from the ~82 sequenced genomes in the PANTHER database.

Gene set enrichment analyses. Ranked gene lists were assessed for GO enrichment, using the logistic regression function in Babelomics 4.3. This is a threshold-free enrichment analysis that determines which GO categories are represented by genes near the top of a ranked list

(15). We performed these analyses on both single species and homologous triplet ranked gene lists. For the homologous triplet analysis, where there was a mouse, stickleback, and bee gene representative per triplet group of genes (each with its own set of GO terms), we assigned functional enrichment on the basis of the mouse gene representative.

Gene set enrichment analysis post hoc analyses. Gene set enrichment analysis (GSEA) is a rank-based analysis that identifies GO terms that are strongly represented by genes showing relatively high levels of differential expression across the experimental and control groups. This approach was necessary to allow for the same statistical approach across both single species and homologous triplet analyses. However, because the GSEA is threshold free, it is not adequate to identify specific genes that account for significant enrichment of a particular functional term. For a subset of significantly enriched GO terms of interest, we performed a post hoc analysis to determine which genes (or triplet groups of genes) within that category accounted most strongly for the enrichment of the GO term. To do this, we performed iterative hypergeometric tests to determine the threshold cutoff point within the ranked list of genes that corresponded to the strongest significance for enrichment of the GO category of interest. Honey bee genes identified in this way were checked against a list of genes known to be prone to contamination by the hypopharyngeal gland tissue to verify the likelihood the signal originated from the brain (below and Dataset S9).

Honey bee hypopharyngeal gland gene list. Nurse and forager worker bees were identified and collected following published procedures (16). We dissected honey bee hypopharyngeal glands (HPGs) out of heads, using RNA-Later ICE. We generated three samples of nurse HPGs and three samples of forager HPGs (we combined HPGs from two individuals for each sample). For a different set of individuals collected at the same time, we lyophilized heads and dissected brains, generating three total samples, each composed of one nurse and one forager brain. We extracted RNA, generated RNA-seq libraries, and sequenced samples as above.

Raw FASTQ data were trimmed for sequencing adapters and quality from the 3′ end, using the program Trimmomatic v0.22 (17), using a minimal phred33 quality score of 20 and a minimal length of 30. Sequences were then aligned using Tophat2 v2.0.8, using the default parameters for single-end reads (–library-type fr-firststrand) and the *A. mellifera* v4.5 reference genome and OGS 3.2 gene models. Raw read counts were generated using htseq-count from HTSeq v0.5.4 with the OGS 3.2 gene models. The htseq-count parameters used were -s reverse -m union -a 0 -t gene -i ID.

The raw read counts were input into R v3.0.2 for data preprocessing and statistical analysis. A total of 4,657/15,319 genes did not have at least 10 counts in at least three samples and were filtered out. The remaining 10,662 genes were analyzed for differential expression, using EdgeR v3.4.0. The raw count values were used in a negative binomial generalized log-linear model (18) that accounted for the total library size for each sample and an extra trimmed mean of M values normalization factor (19) for any biases due to changes in total RNA composition of the samples. Pairwise comparisons were pulled as contrasts from the model and a false discovery rate correction (20) was done separately for each comparison.

Our goal was to identify genes whose expression level in the brain may be biased by HPG contamination. To do this, we generated lists of genes that were both highly expressed in the HPG and highly enriched in the HPG compared with the brain. For each gene in nurses and foragers separately, we calculated log counts per million in the HPG and log fold change, comparing the HPG to the brain samples. We selected genes that had a log counts per million greater than 9 and were found to be in the top 3% of genes enriched in HPG compared with brain (log fold change >4). This resulted in a list of 45 genes for nurses and 31 genes for foragers. There was a high level of overlap between the nurse and forager lists, and both lists were combined to generate a list of 54 genes whose brain signal is prone to HPG contamination in the honey bee.

cis-Motif Analysis.

Whole-genome scanning for motif occurrences, using Stubb. The Stubb algorithm (21) was used to score each 500-bp window of the genome (with 250-bp shifts) for presence of TF binding sites, defined as matches to the position weight matrix (PWM) or "motif" corresponding to that TF. The Stubb score of a DNA sequence window reflects both the number and strengths of putative binding sites in that sequence. The Stubb score uses a fixed "background model" of genomic composition (separate for each genome). To account for the significant heterogeneity of local G/C composition in each genome [especially in the honey bee genome (22)], the Stubb score of each window was converted into an empirical *P* value based on its rank among all genomic windows of similar G/C content. Motif scanning was performed separately for each motif in our collection, which included PWMs in the JASPAR database (23) (129 motifs) and from the Taipale laboratory (24) (239 motifs). A tandem-repeat masker (25) was used before scoring to prevent Stubb from misinterpreting short tandem repeats as weak binding sites. The result of this step was a genome-wide motif score profile for each of the three genomes, for each motif in our collection.

Motif enrichment tests. We considered the differentially expressed genes (up- or down-regulated in experiment vs. control) in each of the three species. The numbers of up-/down-regulated genes at FDR ≤ 0.1 were $n = 518$ and 291 (mouse), 372 and 127 (fish), and 46 and 107 (bee). Each of these six sets of genes (hereafter called "DEG" sets) was separately tested for enrichment of (*i*) individual motifs and (*ii*) Boolean combinations of motifs. Each test involved defining a "motif target set" comprising genes that have high scoring windows for that motif (or motif pair) in their control regions. The enrichment was quantified by a hypergeometric test of overlap between a DEG set and a motif target set.

Defining motif target sets. For tests involving individual motifs, a gene's control region was defined as the intergenic region between the gene and the nearest gene on either side. To correct for the variability in lengths of control regions, we adjusted the empirical *P* value of each window as

$$p = 1 - (1 - p_{mk})^{\text{win}},$$

where p_{mk} is the empirical *P* value of the window and win is the number of scored windows in the control region to which this window belongs. The lowest adjusted *P* value (p in equation above) in a control region was assigned as the motif score of the gene, and 500 genes with the lowest scores (strongest motif presence) were designated as the motif target set.

For tests involving pairs of motifs, motif target sets were defined based on high-scoring windows (lowest adjusted *P* values) in the 1-kbp sequence immediately upstream of the gene, to ensure that the presence or absence of binding sites of a pair of TFs was considered in proximity to each other. Note that a valid Boolean combination of "motif 1" and "motif 2" may be "motif 1 and not motif 2" and the motif target genes in such cases must be in the motif target set of motif 1 and outside the motif target set of motif 2. Also, in defining motif target sets for (Boolean combinations of) pairs of motifs, we required that motif targets of each motif have windows with adjusted *P* value ≤ 0.01, in addition to being in the top 500 highest-scoring genes for that motif.

Meta-associations. *cis*-Metalysis (26) was used to identify enriched motifs or motif pairs, in the promoters of up- or down-regulated DEGs across species. Briefly, *cis*-Metalysis considers the

hypergeometric test *P* values of enrichment tests from each of the three species and combines these *P* values into a "meta *P* value" representing the statistical significance of a "meta-association" spanning multiple species. A meta-association reported by *cis*-Metalysis may be significant due to significant *P* values in all three species, in two of the three species, and even due to a suitably strong *P* value in any one species. (We ignored meta-associations of the third kind here, because we were specifically interested in a motif being significantly associated with DEG sets in two or three species.) For tests involving individual motifs, *cis*-Metalysis was run in the "flexible" mode, which allowed for meta-associations where the motif was significantly enriched in up-regulated genes in one species and significantly enriched in down-regulated genes of another species. For tests involving pairs of motifs, *cis*-Metalysis was run in the "identical" mode, where the individual associations forming the meta-association must involve differentially regulated genes of the same "directionality" (up- or down-regulated; details in ref. 26).

Empirical false discovery rate. Each *cis*-Metalysis run tests a large number of motifs for meta-associations. For tests involving individual motifs, the flexible mode of *cis*-Metalysis results in testing of different combinations of up- and down-regulated gene sets from each species. For tests involving pairs of motifs, the identical mode of *cis*-Metalysis results in testing of different Boolean combinations of each pair of motifs. To account for the multiple-hypothesis testing problem thus introduced, we estimated empirical FDR values corresponding to each meta *P* value reported by *cis*-Metalysis. To this end, we generated 100 randomized versions of the dataset analyzed by *cis*-Metalysis. In each randomized version the motif target sets were randomly redefined (while maintaining their sizes and mutual overlap sizes). We collected all meta-associations reported by *cis*-Metalysis on a randomized dataset, and for any $\tau \in [0,1]$ we counted the number of meta-associations below (stronger than) τ. This count of "false discoveries" was then averaged over the 100 randomized datasets, to give us an estimate F for the number of false discoveries at significance level τ. Then, for each meta *P* value τ on the real dataset, letting D denote the number of meta-associations with meta *P* value $\leq \tau$, we estimated a corresponding empirical FDR as the fraction F/D.

qPCR. M-MuLV reverse transcriptase (New England Biolabs; M0253L) was used to generate cDNA from pools of RNA dissected from VH, BNST, and cortex of experimental or control animals, respectively. qPCR reactions of 10 μL were conducted in triplicate, using Power SYBR Green PCR master mix (Applied Biosystems). Expression levels were normalized relative to the average expression of the mouse 18S rRNA gene. Primer sets are reported in Table S3. Samples were analyzed in an Applied Biosciences 7900HT thermocycler, and the delta-delta-Ct method (27) was used to determine normalized quantities for comparisons.

Immunohistochemistry. For tissue collection, all animals were deeply anesthetized using isoflurane and transcardially perfused with 4% paraformaldehyde. The brains were quickly dissected and immersed in 30% sucrose (in cold PBS, pH 7.4) until the tissues sank. They were then frozen in Tissue-Tek O.C.T. Compound (VWR; no. 25608-930). Brains were sagitally cut into 14-μm sections using a Leica RM2155 microtome, loaded onto Super Plus charged slides, and allowed to air dry for 20 min before storing at −80 °C.

For IHC, slides were allowed to reach room temperature, fixed in acetone for 10 min at room temperature, and washed in PBS. Samples were then treated with Proteinase K (5 μg/mL) for 5 min before staining. Sections were blocked with BSA and 0.1% NaN3 in PBS at pH 7.4 (Antibody Diluent Reagent Solution; Invitrogen, no. 003218) for 30 min at room temperature and then incubated overnight with primary antibody at 4 °C.

The following primary antibodies were used: Emx1 (1:100; Santa Cruz Biotechnology, no. sc-28220), Foxg1 (1:100; Abcam, no. ab23470), and Nr5a1/SF-1 (1:100; Millipore, no. 07-618).

Biotin–Goat Anti-Rabbit IgG (H+L) DS Grd (Invitrogen; no. 656140) or Biotin-Goat Anti-Mouse IgG (H+L) (Invitrogen; no. 626540) was used as a secondary antibody at a dilution of 1:200, followed by Streptavidin Alexa Fluor 594 Conjugate (1:400; Molecular Probes, Life Technologies; no. S32356) for 20 min. Tissue was counterstained using 10 μg/mL Hoechst 33342 nuclear staining (Invitrogen; no. H3570). Sections were mounted using Prolong Gold Antifade Reagent (Molecular Probes, Life Technologies; no. P36934). Fluorescent images were captured using a NanoZoomer high-resolution scanner (Hamamatsu) and an ApoTome Structured Illumination Optical Sectioning System (Zeiss) incorporated into an Axiovert 200M microscope (Zeiss).

1. Wootton RJ (1976) *The Biology of Sticklebacks* (Academic, London).
2. Breed MD (1983) Nestmate recognition in honey bees. *Anim Behav* 31(1):86–91.
3. Breed MD, et al. (1995) The role of wax comb in honey bee nestmate recognition. *Anim Behav* 50:489–496.
4. Kim D, et al. (2013) TopHat2: Accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol* 14(4):R36.
5. Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9(4):357–359.
6. Trapnell C, Pachter L, Salzberg SL (2009) TopHat: Discovering splice junctions with RNA-Seq. *Bioinformatics* 25(9):1105–1111.
7. Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 10(3):R25.
8. Elsik CG, et al.; HGSC production teams; Honey Bee Genome Sequencing Consortium (2014) Finding the missing honey bee genes: Lessons learned from a genome upgrade. *BMC Genomics* 15:86.
9. Li H, et al.; 1000 Genome Project Data Processing Subgroup (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25(16):2078–2079.
10. Anders S, Pyl PT, Huber W (2014) HTSeq - a Python framework to work with high-throughput sequencing data. *Bioinformatics*, 10.1093/bioinformatics/btu638.
11. Robinson MD, McCarthy DJ, Smyth GK (2010) edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26(1):139–140.
12. Kugler KG, Mueller LA, Graber A (2010) MADAM - An open source meta-analysis toolbox for R and Bioconductor. *Source Code Biol Med* 5:3.
13. Medina I, et al. (2010) Babelomics: An integrative platform for the analysis of transcriptomics, proteomics and genomic data with advanced functional profiling. *Nucleic Acids Res* 38(web server issue):W210–W213.
14. Mi H, Muruganujan A, Thomas PD (2013) PANTHER in 2013: Modeling the evolution of gene function, and other gene attributes, in the context of phylogenetic trees. *Nucleic Acids Res* 41(database issue):D377–D386.
15. Sartor MA, Leikauf GD, Medvedovic M (2009) LRpath: A logistic regression approach for identifying enriched biological groups in gene expression data. *Bioinformatics* 25(2):211–217.
16. Whitfield CW, Cziko AM, Robinson GE (2003) Gene expression profiles in the brain predict behavior in individual honey bees. *Science* 302(5643):296–299.
17. Lohse M, et al. (2012) RobiNA: A user-friendly, integrated software solution for RNA-Seq-based transcriptomics. *Nucleic Acids Res* 40(web server issue):W622–W627.
18. McCarthy DJ, Chen Y, Smyth GK (2012) Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res* 40(10):4288–4297.
19. Robinson MD, Oshlack A (2010) A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol* 11(3):R25.
20. Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J R Stat Soc B* 57(1):289–300.
21. Sinha S, van Nimwegen E, Siggia ED (2003) A probabilistic method to detect regulatory modules. *Bioinformatics* 19(suppl 1):i292–i301.
22. Sinha S, Ling X, Whitfield CW, Zhai C, Robinson GE (2006) Genome scan for cis-regulatory DNA motifs associated with social behavior in honey bees. *Proc Natl Acad Sci USA* 103(44):16352–16357.
23. Portales-Casamar E, et al. (2010) JASPAR 2010: The greatly expanded open-access database of transcription factor binding profiles. *Nucleic Acids Res* 38(database issue):D105–D110.
24. Jolma A, et al. (2013) DNA-binding specificities of human transcription factors. *Cell* 152(1-2):327–339.
25. Benson G (1999) Tandem repeats finder: A program to analyze DNA sequences. *Nucleic Acids Res* 27(2):573–580.
26. Ament SA, et al. (2012) New meta-analysis tools reveal common transcriptional regulatory basis for multiple determinants of behavior. *Proc Natl Acad Sci USA* 109(26): E1801–E1810.
27. Schmittgen TD, Livak KJ (2008) Analyzing real-time PCR data by the comparative C(T) method. *Nat Protoc* 3(6):1101–1108.

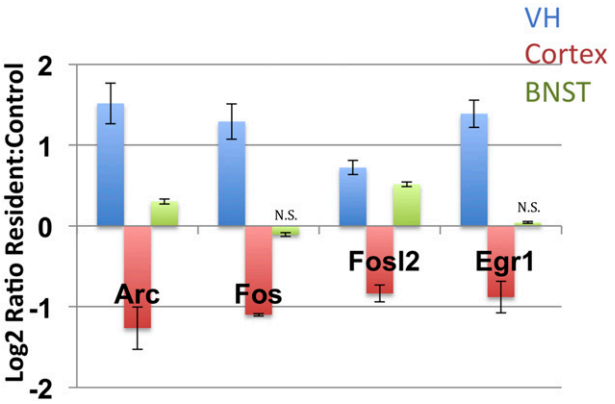

Fig. S1. qPCR results for a selection of IEGs measured in the mouse ventral hypothalamus (VH), bed nucleus of the stria terminalis (BNST), and cortex. Results are expressed as the $\log_2$ of the ratio of experimental to control expression levels (+SE) following normalization of each to 18S as an endogenous control. All bars except the two noted with "N.S." are significantly different from zero ($P < 0.05$).

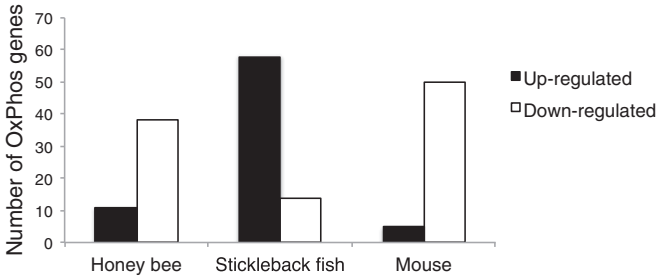

Fig. S2. Genes annotated to "oxidative phosphorylation" (GO: 0006119) showed a significant pattern of regulation for all species (sign test, $P < 0.0001$) when comparing the direction of expression between the experimental and control conditions.

Table S1. Genes responsible for the enrichment of the GO term G-protein–coupled receptor activity (identified by the GSEA post hoc analysis)

Gene symbol	Gene name
Mouse	
Adora2a	Adenosine Receptor A2a receptor
Gpr88	G-protein–coupled receptor 88
Drd1a	Dopamine receptor D1
Stickleback	
ENSGACG00000010585	Bradykinin receptor
ENSGACG00000010586	Bradykinin receptor
ENSGACG00000010329	None
ENSGACG00000012535	Somatostatin receptor
ENSGACG00000007555	Taste receptor type 1
ENSGACG00000002297	Cadherin-3
ENSGACG00000001434	Rhodopsin
ENSGACG00000000716	Opsin 1
ENSGACG00000015110	None
ENSGACG00000000189	None
ENSGACG00000001398	None
ENSGACG00000019960	Prostaglandin D2 receptor 2
ENSGACG00000015114	Cadherin 26
ENSGACG00000010287	Opsin 1
Honey bee	
GB47118	Cadherin
GB50034	Rhodopsin 2
GB41643	Rhodopsin 5
GB46500	Ecdysis triggering hormone receptor activity
GB54316	Crustacean cardioactive peptide receptor
GB44824	None
GB53589	Frizzled 2
GB54361	None

Table S2. Results of the post hoc analysis in mouse that identified significant transcription factors belonging to the GO term "Sequence-specific DNA binding transcription factor activity"

Mouse	Direction of change	Found in homology analysis
Egr3	Up	No
Pitx1	Down	No
Shox2	Down	No
Nr4a3	Up	Yes
Foxg1	Up	Yes
Egr2	Up	No
Emx1	Up	Yes
Egr4	Up	No
Pitx2	Down	No
Tbr1	Up	No

Table S3. Primers for mouse qPCR

Gene	FWD_name	Forward primer	RVS_name	Reverse primer
Arc	Arc_Q1_F	CACCAAAACCCAGGGGACAT	Arc_Q1_R	TATGAATCACTGCTGGGGGC
Fos	Fos_Q1_F	GTTCGTGAAACACACCAGGC	Fos_Q1_R	GGCCTTGACTCACATGCTCT
Fosl2	Fosl2_Q1_F	TGTACACGTGTGTCCTCTGC	Fosl2_Q1_R	GCCAGAGAAGGACTGTTCCC
Egr1	Egr1_Q1_F	AGTGATGAACGCAAGAGGCA	Egr1_Q1_R	TAGCCACTGGGGATGGGTAA

Other Supporting Information Files

Dataset S1 (XLSX)
Dataset S2 (XLSX)
Dataset S3 (XLSX)
Dataset S4 (XLSX)
Dataset S5 (XLSX)
Dataset S6 (XLSX)
Dataset S7 (XLSX)
Dataset S8 (XLSX)
Dataset S9 (XLSX)