Testing the kinship theory of intragenomic conflict in honey bees (*Apis mellifera*)

David A. Galbraith[a,1], Sarah D. Kocher[b,c], Tom Glenn[d], Istvan Albert[e], Greg J. Hunt[f], Joan E. Strassmann[g], David C. Queller[g], and Christina M. Grozinger[a,1]

[a]Department of Entomology, Center for Pollinator Research, Pennsylvania State University, University Park, PA 16802; [b]Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138; [c]Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, NJ 08544; [d]Glenn Apiaries, Fallbrook, CA 92088; [e]Department of Biochemistry and Molecular Biology, Pennsylvania State University, University Park, PA 16802; [f]Department of Entomology, Purdue University, West Lafayette, IN 47907; and [g]Department of Biology, Washington University in St. Louis, St. Louis, MO 63130

Edited by Gene E. Robinson, University of Illinois at Urbana–Champaign, Urbana, IL, and approved December 16, 2015 (received for review August 20, 2015)

Sexual reproduction brings genes from two parents (matrigenes and patrigenes) together into one individual. These genes, despite being unrelated, should show nearly perfect cooperation because each gains equally through the production of offspring. However, an individual's matrigenes and patrigenes can have different probabilities of being present in other relatives, so kin selection could act on them differently. Such intragenomic conflict could be implemented by partial or complete silencing (imprinting) of an allele by one of the parents. Evidence supporting this theory is seen in offspring–mother interactions, with patrigenes favoring acquisition of more of the mother's resources if some of the costs fall on half-siblings who do not share the patrigene. The kinship theory of intragenomic conflict is little tested in other contexts, but it predicts that matrigene–patrigene conflict may be rife in social insects. We tested the hypothesis that honey bee worker reproduction is promoted more by patrigenes than matrigenes by comparing across nine reciprocal crosses of two distinct genetic stocks. As predicted, hybrid workers show reproductive trait characteristics of their paternal stock, (indicating enhanced activity of the patrigenes on these traits), greater patrigenic than matrigenic expression, and significantly increased patrigenic-biased expression in reproductive workers. These results support both the general prediction that matrigene–patrigene conflict occurs in social insects and the specific prediction that honey bee worker reproduction is driven more by patrigenes. The success of these predictions suggests that intragenomic conflict may occur in many contexts where matrigenes and patrigenes have different relatednesses to affected kin.

kinship theory | intragenomic conflict | evolutionary biology | social insects | sociogenomics

In honey bees (*Apis mellifera*), intragenomic conflict is hypothesized to arise over worker reproduction (1, 2). Honey bee colonies consist of a single reproductive female (the queen) that lays all the fertilized eggs and thousands of facultatively sterile female workers that perform all the other tasks within the colony (3). Under queenless and broodless conditions, young female workers can activate their ovaries and lay unfertilized eggs that develop into haploid males (drones) (3). Because the queen mates multiply [with an average of 12 drones (4)], most worker competition to produce males will be between half-sisters. These half-sisters do share matrigenes to some degree, but not patrigenes (2) (see Fig. 1 for details). This relatedness asymmetry is predicted to lead to intragenomic conflict over worker reproduction under queenless conditions: patrigenes should more readily favor worker reproduction than matrigenes, and matrigenes should more readily favor the altruistic behavior of remaining sterile and helping to raise the males of others (2). To determine whether intragenomic conflict influences worker reproduction, we uncoupled parent-of-origin effects from lineage-of-origin effects by creating nine reciprocal cross pairs from three original parental pairs (hereafter referred to as three genetic blocks) of Africanized and European honey bee colonies (Fig. 2). This experimental design allowed us to observe parental effects on the hybrid offsprings' reproductive traits and to assess patrigene/matrigene expression levels in both reproductive and sterile workers.

The kinship theory of intragenomic conflict predicts that compared with matrigenes, patrigenes will favor worker reproduction and exhibit enhanced activity on worker reproductive traits. To test this hypothesis, we reared worker offspring generated from the reciprocal cross pairs in queenless, broodless colonies to stimulate ovary activation in a proportion of the workers, as occurs in queenless colonies. Given that Africanized honey bee workers have larger ovaries (5) and are more likely to become reproductively active (6), we predicted that individuals with Africanized fathers will have larger ovaries and be more likely to activate them than individuals with European fathers (and Africanized mothers); equivalently, individuals with Africanized mothers should have smaller ovaries. Indeed, across all nine reciprocal crosses used in this study, individuals who had Africanized fathers had significantly more ovarioles (Fig. 3*A*; Wilcoxon rank-sum, $P < 0.001$) and were more likely to activate their ovaries to produce mature eggs (Fig. 3*B*; Wilcoxon

Significance

Strong support for the theory of kin selection can come from predicting outcomes under circumstances of within-family conflict. Genes inherited from mothers (matrigenes) and fathers (patrigenes) usually work harmoniously in the offspring. However, kin selection theory predicts these genes may be in conflict over interactions among relatives in which they are unequally represented (half-siblings). In honey bees, patrigenes are predicted to favor daughters that lay eggs themselves rather than remaining sterile and rearing their half-sisters' offspring. We tested this prediction, using crosses of distinct genetic stocks. Workers displayed the reproductive characteristics of their paternal genomes, patrigene expression was higher in reproductive tissues, and this patrigene bias increased in reproductive workers. These results provide strong empirical support for kin selection theory.

Author contributions: D.A.G., S.D.K., G.J.H., J.E.S., D.C.Q., and C.M.G. designed research; D.A.G. and T.G. performed research; D.A.G., S.D.K., I.A., and C.M.G. contributed new reagents/analytic tools; D.A.G. analyzed data; and D.A.G., J.E.S., D.C.Q., and C.M.G. wrote the paper with contributions from S.D.K. and G.J.H.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: Transcriptomic data generated from this study have been deposited in the Gene Expression Omnibus (GEO) database, www.ncbi.nlm.nih.gov/geo (accession no. GSE76164). The genomic data have been deposited in the Sequence Read Archive, www.ncbi.nlm.nih.gov/sra (accession no. SRP067574).

[1]To whom correspondence may be addressed. Email: dag5031@gmail.com or cmg25@psu.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1516636113/-/DCSupplemental.

Downloaded by guest on July 24, 2020

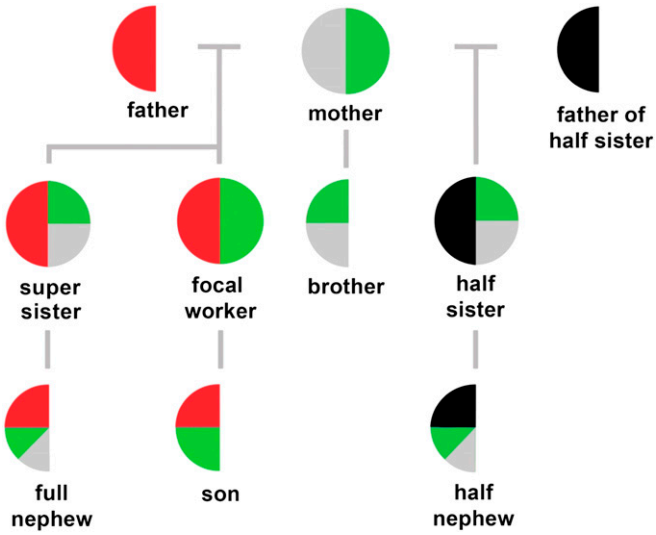

Fig. 1. Genetic basis of the predictions of the intragenomic conflict theory in honey bees. Honey bee colonies are headed by a single reproductive diploid queen that has mated with multiple haploid males. Thus, a focal worker laying eggs under queenless conditions will do so at the cost of rearing fewer sons of half-sisters. The focal worker's patrigenes (red) share no copies with the half-sisters or their sons (half-nephews) because they inherit their patrigenes (black) from another male. Patrigenes will therefore be selected to reproduce at any cost to half-sisters. The focal workers' matrigenes (green; the queen's other gene is gray) do occur in half-sisters and their sons, so matrigenes should be selected to moderate reproduction when the cost to half-sisters is sufficiently high. See ref. 2 for a formal demonstration that the same predictions hold even when the small fraction of full sisters is considered, and also when laying workers replace queen's sons. The schematic for this figure was derived from a design from ref. 51.

rank-sum, $P = 0.002$). Note that two previous studies also found parent-of-origin effects on aggression (7) and ovariole number (8) in honey bees (the latter for a very unusual case in which one race of workers can clonally produce new reproductive females).

To determine whether patrigenes and matrigenes are differentially regulated in sterile and reproductive workers, we collected workers that were sterile (with inactive ovaries) and reproductive (with active ovaries containing mature eggs) from two reciprocal cross pairs (four crosses total), each derived from a different genetic block (see Fig. 2 for details). We predicted that patrigenes would be up-regulated in reproductive workers, regardless of the subspecies of the father.

We sequenced the genomic DNA of the queen mothers and drone fathers of the four crosses and identified 101,372 informative single nucleotide polymorphisms (SNPs) that were present in the genomes of either the mother or father, but not both (allowing us to distinguish matrigenes vs. patrigenes), and were conserved across all crosses (allowing us to identify allele specific expression differences that were replicated in all four crosses). Next, we sequenced the transcriptomes of the reproductive tissues (ovaries and fat bodies) of each of the collected workers from the four crosses (12 sterile and 13 reproductive workers). We identified the matrigenes and patrigenes present in the transcriptomes. Of the informative SNPs identified in the parental genomes, 57,108 were observed within annotated transcripts in the transcriptomes of the worker offspring (OGS3.2) (9).

To provide the most stringent analysis, we generated a list of significantly differentially expressed alleles that were identified using two methods that have been used in previous studies: a generalized linear interactive mixed model (10) using a false discovery rate (FDR) threshold of less than 0.01 and a Storer-Kim test (FDR < 0.01) with additional cutoff thresholds, where the thresholds were previously validated using a meta-analysis (11). There was a significant overlap (Fisher's exact test, $P < 0.0001$) of the lists of transcripts identified in the two analysis methods, and transcripts were only considered significantly biased if they were significant in both analyses. We then compared across the four crosses to identify transcripts in sterile and reproductive workers that exhibited significant expression biases associated with lineage-of-origin and parent-of-origin effects. We identified transcripts in which all SNPs showed the same directional bias in all crosses (the majority of transcripts had multiple SNPs), resulting in 303 transcripts that were consistently biased in all four crosses (Fig. 4). Some transcripts showed biased allelic expression in both sterile and reproductive workers ($n = 77$), whereas others were only biased in one group (sterile only, $n = 87$; reproductive only, $n = 124$). See Dataset S1 for a complete listing of the paternally, maternally, Africanized-, and European-biased transcripts found in sterile and reproductive workers in all four crosses and for a complete list of transcripts that were biased in the two different analyses in individual crosses. We identified several gene ontology categories that were overrepresented in these sets of transcripts (Dataset S1). Finally, we used a previously established method (10, 12) to quantify allele specific expression, using pyrosequencing to validate these results. We evaluated allelic expression levels of 12 genes, using 13 SNPs, in an independent sample of bees. The results were fully consistent with those obtained from the RNAseq analysis (Dataset S1).

Among all paternally biased transcripts in reproductive workers, there were three significant functional categories identified by gene ontology analysis, including anatomical structure morphogenesis, neuron development, and amine catabolic process (see Dataset S2 for a full listing). Maternally biased transcripts in reproductive workers did not produce any significant functional categories. In sterile workers, paternally biased transcripts were significantly associated with six functional categories, including instar larval or pupal development, muscle attachment, or embryonic development ending in birth or egg hatching (Dataset S2), whereas the maternally biased transcripts again did not produce any functional categories.

Because all the workers were reared in queenless and broodless conditions that should begin to stimulate ovary activation, especially by patrigenes, we predicted that more genes would show biased patrigene expression than biased matrigene expression in both classes of workers. This prediction was very strongly supported (Fig. 4; chi square with Yates correction, $P < 0.0001$ for both reproductive and sterile workers). Notably, a previous study of workers reared in the presence of a queen and brood did not show this patrigenic bias (10). Furthermore, we predicted that more transcripts would show a patrigenic bias in reproductive workers compared with sterile workers, and this prediction was also supported (chi square with Yates correction, $P = 0.0259$; Fig. 4).

We also identified the transcripts that were differentially expressed between the sterile and reproductive workers (see Dataset S2 for a listing of these transcripts). Of these, 1,207 had at least one SNP. Comparisons of these significant differentially expressed transcripts with the transcripts that showed significant parent-of-origin biases (303 transcripts) resulted in no significant overlap (Fisher's exact test, $P = 0.289$, see Dataset S3 for a complete list of overlapping transcripts). Thus, there was no tendency for genes that were differentially expressed between sterile and reproductive workers to also show differential allelic expression. This lack of association suggests that full activation of worker ovaries is not the only context driving imprinting.

Several interesting genes exhibited parent-of-origin expression differences, including several paternally biased genes from the ecdysone–vitellogenin pathway, which regulates egg production in insects. Several genes involved in histone modification

GENETICS

Downloaded by guest on July 24, 2020

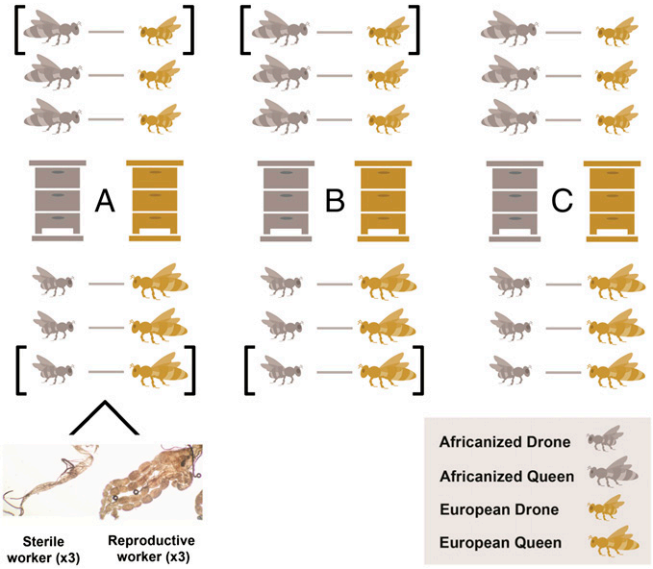

Fig. 2. Schematic of reciprocal cross and experimental design. Graphic representation of reciprocal cross design, where the colors represent the race of the individual (Africanized or European) and the size differences distinguish the queen and the drone. Also shown in the figure are the three separate source colonies or genetic blocks (A, B, and C) that were used for the experimental colonies. Crosses shown in brackets were selected for additional molecular analysis. For these four crosses, we sequenced the genomes of the parents (to identify sequence variations associated with the matrigenes and patrigenes) and the transcriptomes of three reproductive workers and three sterile workers. Reproductive workers had ovaries containing mature eggs, whereas sterile workers had inactive ovaries.

[including homologs of *Drosophila melanogaster, JIL-1* (13), and *Smyd4* (14)] were paternally biased in reproductive workers. Epigenetic processes, such as histone modification, may be involved in mediating differential allelic expression. The reproductive gene *vitellogenin* (*Vg*) (15) showed a paternal expression bias, although it was significantly biased in both reproductive and sterile workers. *Vg* encodes a major yolk precursor protein, although it plays multiple additional roles in regulating honey bee worker behavior (16). The vitellogenin receptor, *yolkless*, was also paternally biased in reproductive workers. Mutations in *yolkless* in *D. melanogaster* results in sterile females with oocytes containing little amounts of yolk compared with flies with a functional *yolkless* gene (17). Expression from the *ecdysone receptor* (*EcR*) and *ecdysone-induced protein 75* (*E75*) genes was also paternally biased in reproductive workers only. The proteins encoded by these genes bind (with other transcription factors; namely, *ultraspiracle*) to upstream regulatory regions in the *Vg* gene and stimulate *Vg* transcription and oogenesis in Dipteran species (18). Thus, several representatives from the ecdysone–vitellogenin pathway are paternally biased and act in the predicted direction of increasing egg production. Future studies are needed to evaluate how variation in allelic expression levels of these candidate genes can result in the observed substantial differences in reproductive physiology.

In summary, these data provide evidence that intragenomic conflict, mediated by differential expression of matrigenes and patrigenes, plays a role in social behavior in honey bee colonies, but the precise mechanisms remain to be determined. In plants and vertebrates, genomic imprinting and allele-specific expression are often mediated by DNA methylation of promoter regions of silenced genes (19). In honey bees and other insects, however, DNA methylation is primarily found in the intragenic regions and is not associated with reducing gene expression levels but, rather, with alternative splicing (20, 21). In other organisms, DNA methylation has been associated with the suppression of transcriptional noise (22–24). In ants, analysis of a single gene found that differential allelic expression was indeed associated with differential DNA methylation levels (25), and thus perhaps in specific gene regions, DNA methylation does affect allelic expression. Methylation has been shown to affect ovarian development in honey bees, consistent with it underlying the patterns we find (26). Alternatively, other epigenetic mechanisms may be operating, including chromatin remodeling via histone modification, which may be mediated by recruiting chromatin remodeling enzymes to specific gene regions via noncoding RNA (27). Chromatin remodeling has been shown to generate parent-of-origin expression differences in *D. melanogaster*, a species that lacks a complete DNA methylation system (28). Finally, our experimental design does not allow us to uncouple the effects of the maternal genome from the mitochondrial genome, and thus it is possible mitochondrial processes play a role (although extant theory on this predicts greater matrigenic expression, and we observe the opposite) (29). Whatever the underlying mechanism may be, our results indicate that parent-of-origin effects on gene expression can be dynamically regulated and change with the physiological state of the individual.

The kinship theory of intragenomic conflict generates numerous detailed predictions of the role of matrigenes and patrigenes in mediating an array of social behaviors in different social contexts (2). Our data provide clear empirical support for a key prediction of this theory in one such novel context: reproduction in honey bee workers. Additional studies are needed to assess the underlying mechanisms and full spectrum of predictions. For example, in species where colonies are headed by one, singly-mated queen (such as bumble bees), our prediction reverses: here maternal alleles should favor worker reproduction (2). But the strong

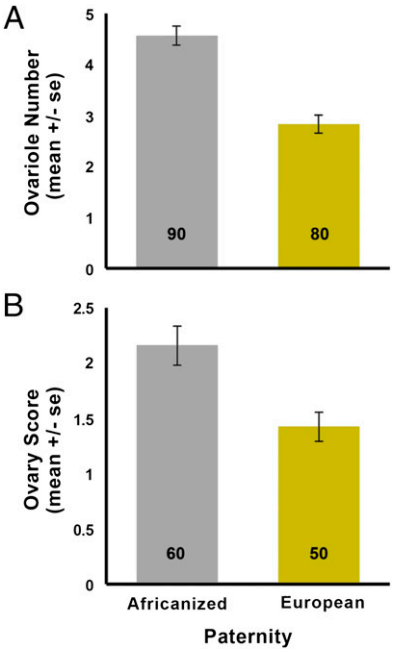

Fig. 3. Patrigenes exert greater influence on ovary size and activation. Honey bee workers with an Africanized father and European mother (*A*) have a greater number of ovarioles ($P < 0.001$) and (*B*) are more likely to become reproductively active ($P = 0.002$) than workers with a European father and an Africanized mother. Because Africanized workers have more ovarioles (5) and are more likely to become reproductively active (6), this suggests patrigenes exhibited enhanced activity on worker reproductive traits relative to matrigenes. Significance was determined using a Wilcoxon rank-sum, and the number of samples is listed in each column.

Downloaded by guest on July 24, 2020

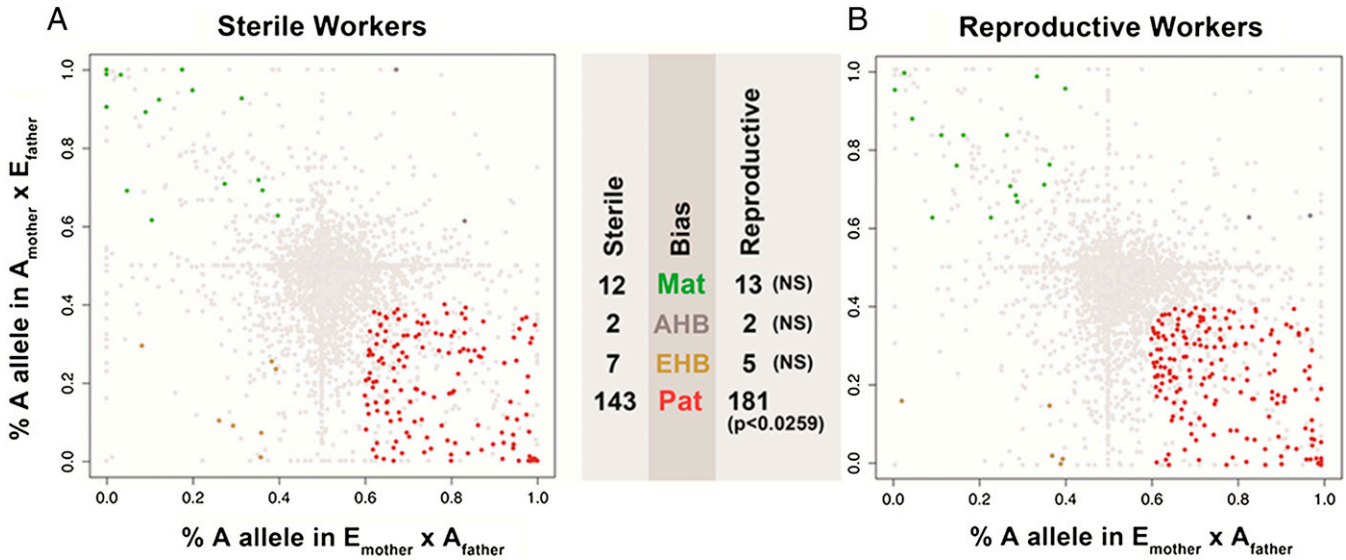

Fig. 4. Worker reproduction is associated with increased expression of patrigenes. Transcript abundance of matrigenes, patrigenes, Africanized alleles, and European alleles were assessed in (*A*) sterile and (*B*) reproductive workers from the four crosses. The *x*-axis represents the proportion of Africanized reads for each transcript in workers with a European mother and Africanized father, whereas the *y*-axis represents the proportion of Africanized reads for each transcript in workers with an Africanized mother and a European father. Each color represents a transcript that is significantly biased in both reciprocal crosses: green is maternal, red is paternal, blue is Africanized, purple is European, and gray is not significant. The significance was determined using an overlap between two statistical tests, a generalized linear interactive mixed model (GLIMMIX) (10), and a Storer-Kim test along with a cutoff threshold of <0.4 and >0.6 expression bias (note that 0.5 is equal expression of both allelic variants) (11).

support we observed for a strange and risky prediction suggests that the theory is correct and that this kind of intragenomic conflict may be a pervasive factor affecting kin-selected behaviors in a wide range of contexts and species.

Methods

Biological Samples. We obtained three Africanized and three European colonies. European colonies were *Apis mellifera ligustica* stock managed by Glenn Apiaries, whereas swarms of Africanized bees were obtained from the Mojave Desert in southern California. The swarms were confirmed to contain Africanized honey bees by using a mitochondrial haplotype analysis (30).

The source colonies were separated into three blocks (A, B, and C), with one Africanized and one European colony assigned to each block. From the two source colonies in each block, we reared three queens and three drones (the parent generation). These queens and drones were crossed by instrumental insemination to create three reciprocal cross pairs per block. This generated a total of nine colonies headed by an Africanized queen inseminated by a European drone, and nine colonies headed by a European queen inseminated by an Africanized drone. Fig. 2 shows a graphical representation of the reciprocal cross pairs. Colonies were generated and maintained at Glenn Apiaries.

Experimental colonies were established using 500 callow workers (F1 generation) from each of the 18 crosses. Workers were paint marked according to their colony of origin and placed into one of three queenless and broodless colonies (A, B, or C; colonies were assigned by block). Each experimental colony was provided with honey frames and a frame of drone comb (to encourage worker egg laying). The colonies were allowed to mature for at least 15 d to promote ovary activation and egg laying. Upon completion of the experiment, each colony was frozen on dry ice and the bees were shipped to Penn State University on dry ice for further processing. Note that honey bee workers recognize their nest mates on the basis of their cuticular hydrocarbon patterns, which are established shortly after emergence as an adult and reflect the colony background (reviewed in ref. 31). Thus, the bees in these experimental colonies should behave as typical for a natural colony.

Parental Effects on Worker Ovary Size and Worker Ovary Activation. We dissected the ovaries of 10 individuals from each reciprocal cross and assessed them for ovary size (based on the number of ovarioles) and activation, using a previously established protocol (32–34). Ovaries were dissected in ice-cold RNAlater (Qiagen). A Wilcoxon rank sum test was used to determine the significance of the differences between individuals with an Africanized father and individuals with a European father.

Eviscerated abdomens with attached fat bodies were stored with the ovaries at −80 °C for transcriptome analysis. Ovaries were chosen for their role in reproduction, and fat bodies were chosen for their role in metabolism, lipid storage, detoxification, and synthesis of vitellogenin, which is required for egg maturation in the ovaries (35, 36).

Sample Preparation for RNA-Seq and DNA-Seq. RNA was isolated from the ovaries and fat bodies of individual worker bees from one reciprocal cross pair (AE and EA) from each of blocks A and B. For each cross, we collected three bees with activated ovaries (containing mature eggs) and three bees with inactivated ovaries, for a total of 12 bees with active ovaries and 13 with inactive ovaries (a submission of 25 samples was required for sequencing, so we selected a 13th sample). RNA was extracted from the tissues of each individual worker, using TRIzol Reagent (Invitrogen). Samples were then treated with TURBO DNase (Ambion) to remove DNA contamination.

Genomic DNA was isolated from the abdomens of the parents (queen and drone) of each cross (for a total of four queens and four drones) used for the RNA-seq analysis, using a Gentra Puregene Tissue Kit (Qiagen).

Quantities of RNA and DNA were measured using a Qubit Fluorometer (Life Technologies) to ensure adequate quantities for sequencing. RNA and DNA samples were sent to the Genome Technology Access Center at Washington University in St. Louis for library construction and sequencing. The samples were sequenced across three and four lanes for the RNA and DNA samples, respectively, on the Illumina HiSEq. 2500 platform. The paired-end sequencing resulted in 2 × 101 bp reads. Transcriptomic data generated from this study have been deposited in NCBI's Gene Expression Omnibus (37) and are accessible through the GEO Series accession no. GSE76164; the genomic data have been deposited in NCBI's Sequence Read Archive (38) and are accessible through the SRA Series accession no. SRP067574.

Generation of Alternate References. The sequences of the genomes of the parents (four queens and four drones) were analyzed to identify SNPs to distinguish between paternal and maternal alleles in the worker offspring. DNA sequencing reads were processed using Trimmomatic (39), removing adaptor sequences and low-quality reads. The trimmed DNA sequencing reads were then aligned to the most recent honey bee genome (Amel_4.5) (9), using Burrows-Wheeler Alignment – Maximal exact match (BWA-MEM) (40). The aligned sequences were used to generate a list of SNPs for each queen and drone, using Freebayes (https://github.com/ekg/freebayes), which can be adjusted to account for ploidy in each individual. Because drones are

GENETICS

Downloaded by guest on July 24, 2020

haploid and queens are diploid, it is necessary to adjust for ploidy to remove biases in SNP detection. Heterozygous SNPs from the queen were then removed, leaving only homozygous SNPs in both the queen and drone. The remaining SNPs were then used to create an alternate reference for each parent by replacing the reference base at the SNP site with the variant detected with Freebayes, using the FastaAlternateReferenceMaker tool, available through the Genome Analysis Tool Kit (41). These alternate references represent the genomes of each parent and will then allow for the tracking of allele specific expression.

Detection of Allelic Expression Differences. RNA sequencing reads were processed using the same method as described for the DNA reads. The trimmed reads were then aligned to the queen- and drone-generated alternate reference files, using Tophat 2.0.11 with the -N 0 option, allowing no mismatches. Reducing the number of mismatches allowed to 0 ensures that reads containing SNPs will only map to the alternate reference containing the SNP and not the other. SNPs that were present in both the queen and the drone were then removed using BEDtools subtractBed (42), as these were not informative for this study. Of the remaining SNPs, only those present in all crosses (coming from either the queen or the drone, but not both) were selected for further analyses to ensure maximum robustness.

Read counts for each of the remaining SNPs were generated and imported into R (www.r-project.org) to test for allelic expression differences, using a Storer-Kim test (43) and previously established cutoff thresholds (11), following a suggested protocol to identify allele-specific expression (11, 44, 45). The Storer-Kim test was used to test the null hypothesis (h_o: $p1 - p2 = 0$) in individuals with active ovaries and inactive ovaries separately, where $p1$ is defined as the proportion of Africanized reads in ExA individuals and $p2$ is defined as the proportion of Africanized reads in AxE individuals. Transcripts with $p1 > 0.5$ and $p2 < 0.5$ were considered as putative paternally biased, and transcripts with $p1 < 0.5$ and $p2 > 0.5$ were considered as putative maternally biased. In addition to the statistical significance of the Storer-Kim test, a threshold based on the degree of allele-specific bias was also used as previously suggested (11), where transcripts with $p1 > 0.6$ and $p2 < 0.4$ were considered to be paternally biased, and $p1 < 0.4$ and $p2 > 0.6$ were considered to be maternally biased. Each SNP was analyzed separately, and for a transcript to be considered as significant, all SNPs were required to show the same directional parent-of-origin bias.

The dataset was also analyzed using a generalized linear interactive mixed model (GLIMMIX), using count data at each SNP to assess parent-specific expression for each transcript. This method has been used previously to identify parent-specific gene expression in honey bees (10). Parent of origin (maternal vs. paternal), race of origin (Africanized honey bee vs. European honey bee), and their interaction (parent × race) were included as fixed terms in the model, whereas SNP and replicate were considered as random factors. A false-discovery threshold of $P < 0.05$ was used to correct for multiple testing. There was a significant overlap (Fisher's exact test, $P < 0.0001$) between the two analysis methods (Storer-Kim and GLIMMIX), and only transcripts that exhibited a significant allelic expression bias in both methods were used for further analyses.

One concern is that despite the high levels of replication and stringency in our analyses, we may have missed a SNP from one of the queen mothers and incorrectly labeled that location as "homozygous" in the queen when it was, in fact, heterozygous. If the "hidden" allele matched the drone's allele, this would have resulted in 0 matrigene reads in half of the offspring and inflated patrigene reads for that transcript. To address this concern, we removed from the "significantly biased transcript list" any genes that showed 0 maternal reads in 1 or more of the 25 individual bees in the analysis. The numbers of significantly biased genes were slightly reduced (see histogram in Dataset S1), but the main results (the strong patrigenic bias and the number of genes exhibiting patrigenic bias increasing significantly from sterile to reproductive workers) remained significant (Dataset S1).

Gene Ontology. For the gene ontology analysis, *D. melanogaster* orthologs of genes containing SNPs exhibiting significant parent-of-origin effects were determined using BLAST (46), with an E-value threshold of 1×10^{-6}. These orthologs were then uploaded to DAVID Bioinformatics Resources 6.7 (47) for gene ontology analysis.

Detection of Differential Gene Expression between Reproductive and Sterile Workers. Trimmed transcriptome sequencing reads were aligned to the most recent honey bee genome assembly (Amel_4.5) (9), using Tophat 2.0.11 (48). Aligned read counts were imported into R statistical software (www.r-project.org). Genes with low read counts (fewer than five reads per gene) were removed. The data were normalized using a trimmed mean of M-values method (49). The EdgeR package in R statistical program was used to identify significantly differentially expressed genes (50).

Independent Validation of the Results Using Pyrosequencing. RNA was extracted from the fat bodies and ovaries of an independent set of bees from one pair of reciprocal crosses, using the same protocol as earlier. Primers were designed to produce a fragment that covers the particular SNP site to be validated, using Pyromark Assay Design 2.0 (Qiagen). The fragments were amplified from the extracted RNA and submitted for pyrosequencing, following a standard protocol for measuring allele-specific mRNA abundance (12).

ACKNOWLEDGMENTS. We thank Harland Patch and Nick Sloff (Penn State) for expert assistance with the graphic design of the figures. This work was supported by the National Science Foundation under Grants MCB 0951200 (to C.M.G.) and IOS 1256416 (to J.E.S. and D.C.Q.) and by the John Templeton Foundation (D.C.Q. and J.E.S.).

1. Haig D (1992) Intragenomic conflict and the evolution of eusociality. *J Theor Biol* 156(3):401–403.
2. Queller DC (2003) Theory of genomic imprinting conflict in social insects. *BMC Evol Biol* 3:15.
3. Winston ML (1991) *The Biology of the Honey Bee* (Harvard University Press, Cambridge, MA).
4. Tarpy DR, Nielsen R, Nielsen DI (2004) A scientific note on the revised estimates of effective paternity frequency in Apis. *Insectes Soc* 51(2):203–204.
5. Linksvayer TA, et al. (2009) The genetic basis of transgressive ovary size in honeybee workers. *Genetics* 183(2):693–707.
6. Makert GR, Paxton RJ, Hartfelder K (2006) Ovariole number—a predictor of differential reproductive success among worker subfamilies in queenless honeybee (Apis mellifera L.) colonies. *Behav Ecol Sociobiol* 60(6):815–825.
7. Shorter JR, Arechavaleta-Velasco M, Robles-Rios C, Hunt GJ (2012) A genetic analysis of the stinging and guarding behaviors of the honey bee. *Behav Genet* 42(4):663–674.
8. Oldroyd BP, et al. (2013) A parent-of-origin effect on honeybee worker ovary size. *Proc Biol Sci* 281(1775):20132388.
9. 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(1):86.
10. Kocher SD, et al. (2015) A Search for Parent-of-Origin Effects on Honey Bee Gene Expression. *G3 (Bethesda)* 5(8):1657–1662.
11. Wang X, Clark AG (2014) Using next-generation RNA sequencing to identify imprinted genes. *Heredity (Edinb)* 113(2):156–166.
12. Wittkopp PJ (2011) Using pyrosequencing to measure allele-specific mRNA abundance and infer the effects of cis- and trans-regulatory differences. *Methods Mol Biol* 772(3):297–317.
13. Jin Y, Wang Y, Johansen J, Johansen KM (2000) JIL-1, a chromosomal kinase implicated in regulation of chromatin structure, associates with the male specific lethal (MSL) dosage compensation complex. *J Cell Biol* 149(5):1005–1010.
14. Thompson EC, Travers AA (2008) A Drosophila Smyd4 homologue is a muscle-specific transcriptional modulator involved in development. *PLoS One* 3(8):e3008.
15. Raikhel AS, et al. (2002) Molecular biology of mosquito vitellogenesis: From basic studies to genetic engineering of antipathogen immunity. *Insect Biochem Mol Biol* 32(10):1275–1286.
16. Nelson CM, Ihle KE, Fondrk MK, Page RE, Amdam GV (2007) The gene vitellogenin has multiple coordinating effects on social organization. *PLoS Biol* 5(3):e62.
17. DiMario PJ, Mahowald AP (1987) Female sterile (1) yolkless: A recessive female sterile mutation in Drosophila melanogaster with depressed numbers of coated pits and coated vesicles within the developing oocytes. *J Cell Biol* 105(1):199–206.
18. Carney GE, Bender M (2000) The Drosophila ecdysone receptor (EcR) gene is required maternally for normal oogenesis. *Genetics* 154(3):1203–1211.
19. Bird A (2002) DNA methylation patterns and epigenetic memory. *Genes Dev* 16(1): 6–21.
20. Foret S, et al. (2012) DNA methylation dynamics, metabolic fluxes, gene splicing, and alternative phenotypes in honey bees. *Proc Natl Acad Sci USA* 109(13):4968–4973.
21. Galbraith DA, Yang X, Niño EL, Yi S, Grozinger C (2015) Parallel epigenomic and transcriptomic responses to viral infection in honey bees (Apis mellifera). *PLoS Pathog* 11(3):e1004713.
22. Huh I, Zeng J, Park T, Yi SV (2013) DNA methylation and transcriptional noise. *Epigenetics Chromatin* 6(1):9.
23. Zilberman D, Gehring M, Tran RK, Ballinger T, Henikoff S (2007) Genome-wide analysis of Arabidopsis thaliana DNA methylation uncovers an interdependence between methylation and transcription. *Nat Genet* 39(1):61–69.
24. Bird AP (1995) Gene number, noise reduction and biological complexity. *Trends Genet* 11(3):94–100.
25. Bonasio R, et al. (2012) Genome-wide and caste-specific DNA methylomes of the ants Camponotus floridanus and Harpegnathos saltator. *Curr Biol* 22(19):1755–1764.
26. Kucharski R, Maleszka J, Foret S, Maleszka R (2008) Nutritional control of reproductive status in honeybees via DNA methylation. *Science* 319(5871):1827–1830.

Downloaded by guest on July 24, 2020

27. Delaval K, Feil R (2004) Epigenetic regulation of mammalian genomic imprinting. *Curr Opin Genet Dev* 14(2):188–195.
28. Joanis V, Lloyd VK (2002) Genomic imprinting in Drosophila is maintained by the products of Suppressor of variegation and trithorax group, but not Polycomb group, genes. *Mol Genet Genomics* 268(1):103–112.
29. Wolf JB (2009) Cytonuclear interactions can favor the evolution of genomic imprinting. *Evolution* 63(5):1364–1371.
30. Garnery L, Solignac M, Celebrano G, Cornuet J-M (1993) A simple test using restricted PCR-amplifled mitochondrial DNA to study the genetic structure of *Apis mellifera* L. *Experientia* 49(11):1016–1021.
31. Grozinger CM (2014) *Honey Bee Pheromones. The Hive and the Honey Bee* (Northern Bee Books, Hebden Bridge, United Kingdom), pp 1–33.
32. Hartfelder K, Bitondi MMG, Santana WC, Simões ZLP (2002) Ecdysteroid titer and reproduction in queens and workers of the honey bee and of a stingless bee: Loss of ecdysteroid function at increasing levels of sociality? *Insect Biochem Mol Biol* 32(2): 211–216.
33. Velthuis HHW (1970) Ovarian development in Apis mellifera worker bees. *Entomol Exp Appl* 13(4):377–394.
34. Amdam GV, Csondes A, Fondrk MK, Page RE, Jr (2006) Complex social behaviour derived from maternal reproductive traits. *Nature* 439(7072):76–78.
35. Amdam GV, et al. (2005) Social reversal of immunosenescence in honey bee workers. *Exp Gerontol* 40(12):939–947.
36. Seehuus SC, Norberg K, Gimsa U, Krekling T, Amdam GV (2006) Reproductive protein protects functionally sterile honey bee workers from oxidative stress. *Proc Natl Acad Sci USA* 103(4):962–967.
37. Edgar R, Domrachev M, Lash AE (2002) Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res* 30(1):207–210.
38. Leinonen R, Sugawara H, Shumway M (2011) The Sequence Read Archive. *Nucleic Acids Res* 39:D19–D21.
39. Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* 30(15):2114–2120.
40. Li H, Durbin R (2010) Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* 26(5):589–595.
41. McKenna A, et al. (2010) The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 20(9):1297–1303.
42. Quinlan AR, Hall IM (2010) BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* 26(6):841–842.
43. Storer BE, Kim C (1990) *Exact Properties of Some Exact Test Statistics for Comparing Two Binomial Proportions. J Am Stat Assoc* 85(409):146–155.
44. Gehring M, Missirian V, Henikoff S (2011) Genomic analysis of parent-of-origin allelic expression in Arabidopsis thaliana seeds. *PLoS One* 6(8):e23687.
45. Wang X, et al. (2008) Transcriptome-wide identification of novel imprinted genes in neonatal mouse brain. *PLoS One* 3(12):e3839.
46. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215(3):403–410.
47. Huang W, Sherman BT, Lempicki RA (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 4(1):44–57.
48. Trapnell C, Pachter L, Salzberg SL (2009) TopHat: Discovering splice junctions with RNA-Seq. *Bioinformatics* 25(9):1105–1111.
49. Robinson MD, Oshlack A (2010) A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol* 11(3):R25.
50. Anders S, Huber W (2010) Differential expression analysis for sequence count data. *Genome Biol* 11(10):R106.
51. Drewell RA, Lo N, Oxley PR, Oldroyd BP (2012) Kin conflict in insect societies: A new epigenetic perspective. *Trends Ecol Evol* 27(7):367–373.

Downloaded by guest on July 24, 2020