Convergent selection on juvenile hormone signaling is associated with the evolution of eusociality in bees

Beryl M. Jones[1,2,*], Benjamin E.R. Rubin[1,2,*], Olga Dudchenko[3,4,*], Karen M. Kapheim[5], Eli S. Wyman[1,2], Brian G. St. Hilaire[3], Weijie Liu[1,2], Lance R. Parsons[2], S. RaElle Jackson[2], Katharine Goodwin[2], Shawn M. Davidson[2], Callum J. Kingwell[6,7], Andrew E. Webb[1,2], Mauricio Fernández Otárola[8], Melanie Pham[3], Arina D. Omer[3], David Weisz[3], Joshua Schraiber[9,10], Fernando Villanea[9,11], William T. Wcislo[7], Robert J. Paxton[12,13], Brendan G. Hunt[14], Erez Lieberman Aiden[3,4,15,16,ł], Sarah D. Kocher[1,2,ł,§]

*These authors contributed equally
łThese authors contributed equally
§Corresponding Author: skocher@princeton.edu

[1]Department of Ecology and Evolutionary Biology, Princeton University, USA
[2]Lewis-Sigler Institute for Integrative Genomics, Princeton University, USA
[3]The Center for Genome Architecture, Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, USA
[4]Center for Theoretical Biological Physics, Rice University, USA
[5]Department of Biology, Utah State University, USA
[6]Department of Neurobiology and Behavior, Cornell University, USA
[7]Smithsonian Tropical Research Institute, 0843-03092 Panama City, Republic of Panama
[8]Research Center for Biodiversity and Tropical Ecology (CIBET), School of Biology, University of Costa Rica, San José, Costa Rica
[9]Department of Biology, Temple University, USA
[10]Illumina Artificial Intelligence Laboratory, Illumina Inc, San Diego, CA, USA
[11]Department of Anthropology, University of Colorado Boulder, USA
[12]Institute of Biology, Martin-Luther University Halle-Wittenberg, Germany
[13]German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Germany
[14]Department of Entomology, University of Georgia, USA
[15]Shanghai Institute for Advanced Immunochemical Studies, ShanghaiTech, Pudong 201210, China
[16]Faculty of Science, UWA School of Agriculture and Environment, University of Western Australia, Perth WA 6009, Australia

Abstract

Life's most dramatic innovations, from the emergence of self-replicating molecules to highly-integrated societies, often involve increases in biological complexity. Some groups traverse different levels of complexity, providing a framework to identify key factors shaping these evolutionary transitions. Halictid bees span the transition from individual to group reproduction, with repeated gains and losses of eusociality. We generated chromosome-length genome assemblies for 17 species and searched for genes that both experienced positive selection when eusociality arose and relaxed selection when eusociality was secondarily lost. Loci exhibiting these complementary evolutionary signatures are predicted to carry costs outweighed by their importance for traits in eusocial lineages. Strikingly, these loci included two proteins that bind and transport juvenile hormone (JH) – a key regulator of insect development and reproduction. Though changes in JH abundance are frequently associated with polymorphisms, the mechanisms coupling JH to novel phenotypes are not well understood. Our results suggest novel links between JH and eusociality arose in halictids by altering transport and availability of JH in a tissue-specific manner, including in the brain. Through genomic comparisons of species encompassing both the emergence and breakdown of eusociality, we provide insights into the mechanisms targeted by selection to shape a key evolutionary transition.

Keywords: comparative genomics, convergent evolution, eusociality, behavior, bees, juvenile hormone

Introduction

Organisms situated at the inflection point of life's major evolutionary transitions provide a powerful framework to examine the factors shaping the evolution of these traits[1,2]. Halictid bees (Hymenoptera: Halictidae) offer a unique opportunity to study the evolution of eusociality – social colonies with overlapping generations and a non-reproductive worker caste. Within the halictid bees, there have been two independent gains[3] and a dozen losses[4] of eusociality. As a result, closely related species within this group encompass a broad spectrum of social behavior, from solitary individuals that live and reproduce independently to eusocial nests where individuals cooperate to reproduce as a group, and even polymorphic species that produce both solitary and social nests. This evolutionary replication enables a comparative approach to identify the core factors that shape the emergence and breakdown of eusociality and provides insights into the most costly aspects of social life.

We searched for signatures of positive selection associated with the convergent gains of eusociality as well signatures of relaxed selection when eusociality is lost. These complementary patterns

indicate that these loci are associated with costs or trade-offs underlying the maintenance of social traits. We find that some of the targets of selection implicated in the origins and elaborations of eusociality, such as young, taxonomically restricted genes[5] and gene regulatory elements[6], also show relaxation of selective pressures when social behavior is lost, and we uncovered four genes strongly associated with the evolution of eusociality in halictid bees, including the two primary juvenile hormone binding proteins (JHBPs): *apolipoprotein*[7–9] and *hexamerin110*[10,11]. These results, combined with our analysis of JH localization in insects, provide new insights into how JH signaling has likely been modified to shape the evolution of eusociality.

A new comparative genomic resource for studying the evolution of eusociality

We generated comparative genomic resources for halictids consisting of 15 *de novo* genome assemblies and an update of 2 additional assemblies, all with chromosome-length scaffolds (Fig. 1; Fig. S1). We selected species with well-characterized social behaviors that encompass both independent origins and six repeated losses of eusociality[3] (Fig. 1a-b). Sampling closely-related eusocial and solitary species alongside known non-eusocial outgroups provides a powerful framework to examine the molecular mechanisms shaping the evolution of social behavior in this group[1].

Assemblies ranged in size from 267 to 470 Mb, with estimated numbers of chromosomes ranging from 9 to 38 (Table S2). Broadly, we find that, in contrast to mammalian species (Fig. S2), genomic rearrangements among the bee species occur disproportionately within rather than between chromosomes (see Fig. 1c, Fig. S3). Consequently, loci that are on the same chromosome in one bee species also tend to occur on the same chromosome in other bee species. This observation is similar to previous findings in dipteran genomes[12,13] and may indicate a broader trend during the evolution of insect chromosomes.

To increase the quality of genome annotations, we generated

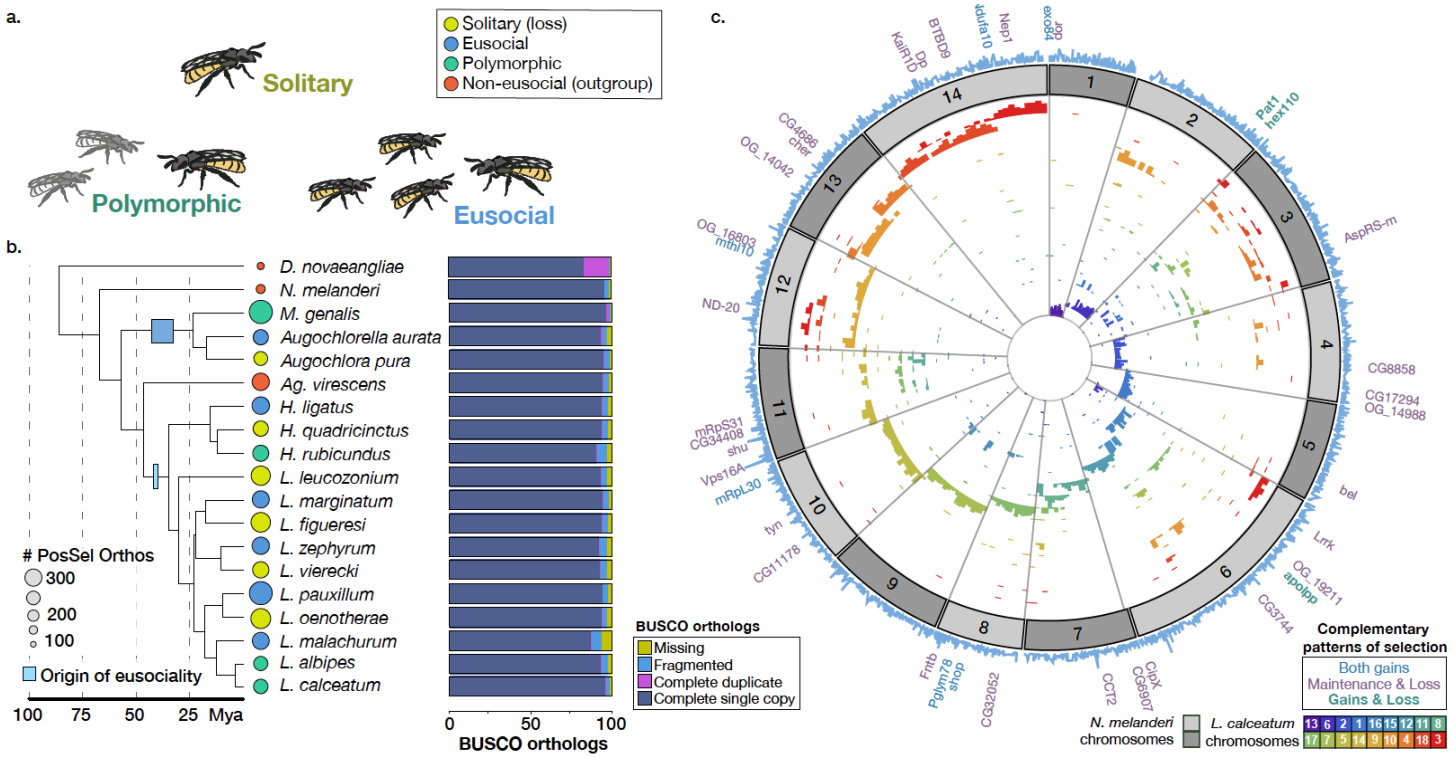

Figure 1. Comparative genomic resources for halictid bees. (a) Halictid bees encompass a wide range of social behaviors, from solitary to eusocial. Solitary species (yellow) nest and reproduce independently. Reproductive queens from eusocial species (blue) initiate colonies with offspring that include non-reproducing workers. Individuals from polymorphic species (green) are capable of both solitary and eusocial reproduction. Non-eusocial outgroups (red) reproduce independently, and include both solitary (*Nomia melanderi*, *Dufourea novaeangliae*) and communal (*Agapostemon virescens*) nesting species. **(b)** Within the Halictinae, there have been at least two independent gains of eusociality and over a dozen independent losses of eusociality. Colors at branch tips indicate the behavior of each species. Circle sizes are proportional to the total number of orthologs with evidence for positive selection on each terminal branch (abSREL tests in HyPhy[16]; FDR<0.1). Light blue rectangles denote the two independent gains of social behavior in this family; their width is proportional to the number of orthologs with evidence for positive selection on each origin branch. Proportions of complete, fragmented, and missing BUSCO orthologs are shown for each genome. **(c)** Genomic data aligned to the outgroup *Nomia melanderi*. The 14 *N. melanderi* chromosomes are represented as a circular ideogram with consecutive chromosomes shown in alternating dark and light gray. The area inside the ideogram comprises 18 color-coded circular tracks, each corresponding to one *L. calceatum* chromosome, and the y-axis of these circular tracks represents the frequency of regions that align to the corresponding region in the *N. melanderi* genome. Usually, most of the alignments fall into a single "wedge", indicating that each *L. calceatum* chromosome corresponds to just one *N. melanderi* chromosome. The pattern is typical for the bees we studied (see Fig. S3). Outer blue line plot indicates the number of branches in the phylogeny where positive selection was detected at each gene (abSREL, p<0.05). The outermost plot highlights genes that show convergent or complementary patterns of selection. Genes experiencing positive selection on both origin branches are labeled in blue, those experiencing an intensification of selection in extant eusocial lineages and complementary relaxation of selection in secondarily solitary species (HyPhy RELAX, FDR<0.1) are in purple, and genes with a complementary pattern of both positive selection on the origin branches and convergent relaxation of selection associated the loss of eusociality are green (note that not all genes in these categories are annotated in *N. melanderi* and are therefore not labeled). See methods section for more details.

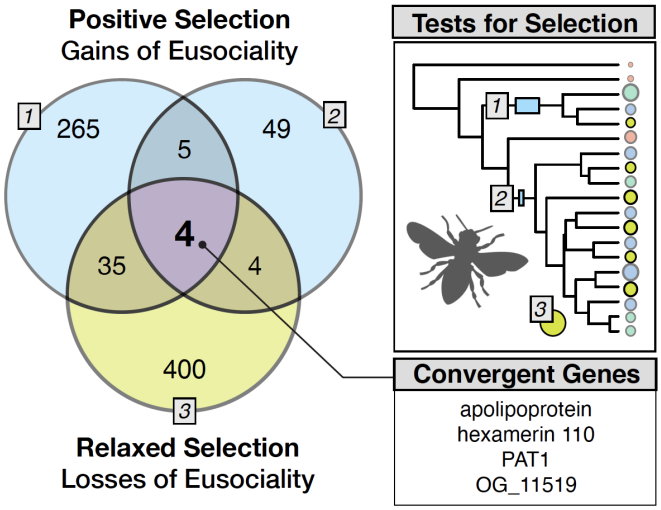

Figure 2. Signatures of selection associated with the gains and losses of eusociality in halictids were identified by testing all orthologs for evidence that dN/dS >1 at a proportion of sites on focal branches (Gains 1 and 2 denoted with gray squares 1-2 on the inset; abSREL[16] tests in HyPhy, FDR<0.05). There were 309 orthologs with evidence of positive selection in the Augochlorini (Gain 1) and 62 in Halictini *s.s.* (Gain 2). Nine of these loci overlapped both origins, which is significantly more than predicted by chance (enrichment ratio=3.97, Fisher's exact p=0.004). Branches representing losses of eusociality (grey square, label 3; yellow circles) were tested for evidence of relaxed selection (HyPhy RELAX[20]); 443 orthologs passed the significance threshold (FDR<0.1), and 4 orthologs overlapped all 3 tests (purple), which is substantially more than predicted by chance (Multi-set Exact Test[21]; fold-enrichment=8.85, p=0.001).

tissue-specific transcriptomes for 11 species, and we also sequenced and characterized 1269 microRNAs (miRs) expressed in the brains of 15 species (Table S12). The number of annotated genes ranged from 11,060 to 14,982, and BUSCO[14] analyses estimated the completeness of our annotations to range from 93.4 to 99.1% (Fig. 1b; Table S2). Whole-genome alignments were generated with progressive Cactus[15] (Fig. 1c), which we then used to identify 52,421 conserved, non-exonic elements present in 10 or more species. All genomes, annotations, and alignments can be viewed and downloaded from the Halictid Geno me Browser (https://beenomes.princeton.edu).

Convergent and complementary signatures of selection are associated with the gains and losses of eusociality

Origins. First, to identify convergent signatures of positive selection associated with transitions from non-eusocial to eusocial life histories, we looked for positive selection on the branches that represent each origin of eusociality: Augochlorini and Halictini *sensu stricto* (Fig. 2). We found 309 genes on the Augochlorini origin branch and 62 genes on the Halictini origin branch with evidence of positive selection (HyPhy abSREL[16], FDR<0.1) (Table S3). On the Augochlorini branch, genes with signatures of positive selection were enriched for cell adhesion (GO:0007155, Fisher's Exact, q=2.01e-4; enrichment=2.92; Table S4). There was no detectable gene ontology (GO) enrichment among positively-selected genes on the Halictini branch (Fisher's Exact, q>0.1). Nine genes showed signatures of positive selection on both branches (Fisher's Exact, p= 0.004, enrichment=3.97), including genes known to regulate lifespan and starvation resistance (*methuselah-like 10*[17]) and neuronal functions (*shopper*[18]), among others (Table S3).

Losses. A unique attribute of halictid bees is that there have been a number of independent losses of eusociality[19] in addition to repeated gains[3]. These reversals provide a powerful lens to identify key genomic factors needed for the maintenance of social living because organisms are expected to minimize investment in traits when social behaviors are lost or unexpressed. This results in the reduction or removal of selective pressures previously maintaining these costly but essential traits[22,23]. Thus, we predicted that genes associated with trade-offs or costly roles in maintaining eusocial societies should show relaxation of constraint in species that have secondarily reverted to solitary nesting. Consistent with this hypothesis, we found 443 genes showing evidence of convergent, relaxed selection on the six branches representing independent losses of eusociality (HyPhy RELAX[20], FDR<0.1; Table S3). These genes are enriched for chromosome condensation (GO:0030261, Fisher's Exact, q=0.067, enrichment=4.09), indicating that they may play a role in chromosome accessibility and gene regulation[24]. They are also enriched for vacuolar transport (GO:0007034, Fisher's Exact, q=0.067, enrichment=3.11; Table S4).

To determine whether or not this pattern is unique to the loss of eusociality, we ran the same tests for relaxed selection using extant eusocial lineages as the focal branches. We found 305 genes with signatures of relaxation in eusocial species (HyPhy RELAX[20], FDR<0.1; Table S3) enriched for four GO terms related to metabolism (Table S4). This is a significantly lower proportion of genes experiencing relaxed selection in eusocial species compared to those experiencing relaxed selection among solitary species (Fisher's exact test $p = 2.42 \times 10^{-7}$, odds-ratio = 1.48), suggesting that the loss of eusociality is more often associated with a release of constraint compared with eusocial maintenance or elaboration.

Convergent selection. Finally, by comparing genes associated with the emergence of eusociality to those associated with its loss, we have the unique ability to identify some of the most consequential molecular mechanisms associated with social evolution. If a shared set of genes is associated with the emergence and maintenance of social behavior in this group, then we would expect to find genes experiencing both positive selection when eusociality emerges and relaxed selection when social behavior is lost. Indeed, we find four genes that meet these criteria: *OG_11519*, a gene with no known homologs outside of Hymenoptera, *Protein interacting with APP tail-1* (*PAT1), apolipoprotein* (*apolpp*), and *hexamerin110* (*hex110*; Fig. 2). This overlap is significantly more than expected by chance (Multi-set Exact Test[21]; fold-enrichment=8.85, p= 0.001). OG_11519 has no identifiable protein domains but is conserved throughout the Hymenoptera. *PAT1* modulates mRNA transport and metabolism[25,26], and *apolpp* and *hex110* are the primary JH binding proteins and collectively bind nearly all the JH in the hemolymph[7–11,27]. Hex110 is associated with caste differentiation in nearly all examined eusocial insects[10,28–30], and ApoLpp has repeatedly been shown to be a major JH binding protein in the hemolymph[8,9,27]. We also find 34 genes that show intensification of selection on extant eusocial lineages and relaxation in secondarily solitary species (HyPhy RELAX, FDR<0.1 for both tests). These genes likely represent trade-offs associated with the maintenance or elaboration of eusociality, and they are enriched for regulation of SNARE complex assembly (GO:0035542, Fisher's Exact, q=0.074, enrichment=80.91; Table S4), which is a key component of synaptic transmission that has also been implicated with variation in social behavior in *L. albipes*[2] and wasps[31].

Young genes and gene regulatory expansions are associated with trade-offs underlying the maintenance of eusociality

Previous studies of eusociality have suggested that, similar to their importance in the evolution of other novel traits[32], younger or taxonomically-restricted genes (TRGs) may play key roles in the

evolution of eusocial behavior[33]. To test this hypothesis, we examined the relationship between gene age and selection associated with eusocial origins, maintenance, and reversions to solitary life histories (Fig. S4). We found a greater proportion of young genes compared with old genes experience relaxed selection when eusociality is subsequently lost (Fig. 3A; Pearson's r=0.869, p=0.002). This relationship does not hold for orthologs showing evidence of relaxed selection on eusocial branches, nor are younger genes more likely than older genes to experience intensified selection associated with either the origins or maintenance of eusociality (Fig. S4) (Table S5). These results suggest that the reversion to solitary life histories is uniquely associated with relaxed selection on young genes, implicating TRGs in traits that are subject to reduced functional constraint when eusociality is lost. Thus, our results partially support other studies in complex eusocial hymenopterans, including honey bees, ants, and wasps, in which young TRGs have been linked to the evolution of the non-reproductive workers[5,34–37], and they further suggest that younger genes are associated with costs or trade-offs linked to eusociality.

Gene regulatory changes have also been implicated in eusocial evolution[37], including the expansion of transcription factor (TF) motifs in the genomes of eusocial species compared with distantly-related solitary species[6]. To assess the degree to which changes in gene regulation may facilitate the evolution of social behavior in halictids, we characterized TF motifs in putative promoter regions in each halictid genome. For each species, we defined these regions as 5kb upstream and 2kb downstream of the transcription start site for each gene[6] and calculated a score for a given TF in each region that reflects the number and/or strength of predicted binding sites[38].

If social species have a greater capacity for gene regulation compared to lineages that have reverted to solitary nesting, then we would expect to find more motifs with scores (reflecting both strength and number of binding sites) that are higher in social taxa compared to secondarily solitary taxa. In support of this hypothesis, we find a greater than 3-fold enrichment of TF motifs that are positively correlated with social lineages compared to secondarily solitary lineages after phylogenetic correction (Fig. 3b), and permutation tests indicate that this difference is highly significant (Fig. 3c). Five of these socially-biased motifs were previously identified as associated with eusocial evolution in bees[6], including the motifs for *lola, hairy, CrebA, CG5180,* and the *met/tai* complex (Table S6).

Interestingly, genes showing relaxed selection associated with the loss of eusociality are enriched for chromosome condensation (Table S4), suggesting there may be trade-offs or costs associated with this increased regulatory potential. To determine whether or not we see a signature of these trade-offs in non-coding sequences, we used phastCons[39] to identify conserved non-exonic elements (CNEEs) in each species. We identified CNEEs that were either slow- or fast- evolving in eusocial and solitary lineages. Using the logic outlined above, we would predict that CNEEs essential to the maintenance of social behavior should be more highly conserved and evolving more slowly in extant eusocial taxa vs. secondarily solitary taxa. After phylogenetic correction, we find 1,876 CNEEs that meet these criteria, with faster rates of evolution in secondarily solitary species (Fig. S5). This is significantly more than expected by chance (Binomial test, $p=5.27e^{-9}$); these regions are proximal to genes enriched for several GO terms, including serotonin receptor signaling (Permutation tests, p=0.007) and adherens junction maintenance (Permutation tests, p=0.009; see Table S7 for complete list). We also find 1,255 CNEEs that are more highly conserved in secondarily solitary lineages (fast in eusocial; slower in solitary). This number is smaller than expected by chance (Binomial test, $p<1e^{-10}$), but these regions are proximal to genes enriched for cytoskeletal organization and processes related to synapse and neurotransmitter functions, among other

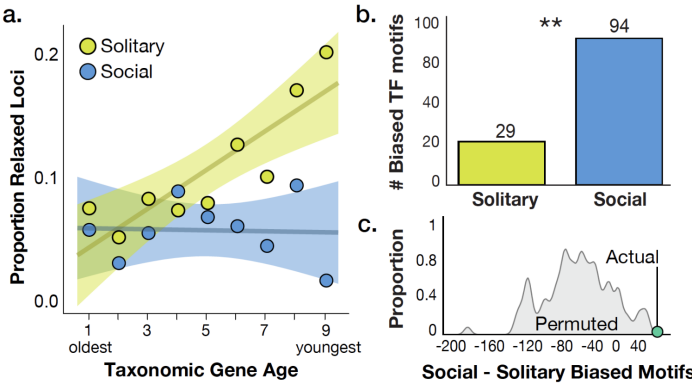

Figure 3. Young genes and increased transcription factor motif presence associated with the maintenance of eusociality. (a) Younger genes are more likely to show signatures of relaxed selection when social behavior is lost. Circles show the proportion of the genes in each age class that show evidence of relaxed selection in solitary (yellow) or social (blue) lineages, from the oldest Bilaterian group (Age=1) to the youngest, halictid-specific taxonomically restricted genes (Age=9; see supplement for complete list). For solitary lineages, we find a significant correlation between gene age and the proportion of those genes with evidence of relaxed selection (Pearson correlation, r=0.869, p=0.002). This relationship does not hold for the social taxa, where genes experiencing relaxed selection in these lineages are found at low frequency in all gene age classes (Pearson correlation, r=-0.0447, p=0.91). Shading represents 95% confidence intervals. (b) STUBB scores[38] were calculated for 223 Drosophila transcription factor binding motifs in each halictid genome, and each motif was tested for overrepresentation in solitary or social genomes; 94 motifs were enriched in social genomes compared to only 29 enriched in solitary genomes. "**" indicates p<0.01. (c) Permutation tests reveal that this approximately 3-fold enrichment in (b) is unlikely to occur by chance (empirical p=<0.01).

GO terms (Permutation tests, p<0.05; Table S7; Fig. S6), potentially providing molecular support to hypotheses that the evolution of eusociality involves altered investments in cognitive function[40,41].

Convergent signatures of selection on Juvenile Hormone Binding Proteins (JHBPs)

We identified a small but robust set of genes with complementary signatures of selection linked to the repeated gains and losses of eusociality (Fig. 2). These complementary patterns highlight the importance of these genes in eusocial lineages: convergent signatures of relaxed selection when eusociality is lost suggest they are also associated with costly trade-offs[23]. Strikingly, 2 of these 4 genes (*apolpp* and *hex110*) are the primary binding proteins for juvenile hormone (JH), an arthropod-specific hormone that regulates many aspects of insect life history including development, reproduction, diapause and polyphenism[42,43]. Together, *apolpp* and *hex110* are thought to bind nearly all JH in insect hemolymph[27].

For both proteins, we find region-specific, faster rates of evolution on eusocial branches compared to non-eusocial outgroups. To identify sites in these genes with evidence of positive selection on the two branches representing independent origins of eusociality, we used a mixed-effects maximum likelihood approach (MEME[44]). For both *apolpp* and *hex110*, we find evidence for positive selection on sites present in functional regions of the protein (Fig. S7), including the receptor binding domain and predicted binding pocket for ApoLpp as well as in all three Hemocyanin domains of Hex110 (associated with storage functions in these proteins[45]) and its predicted binding pocket. Taken together, these results suggest that positive selection shaped JHBP

function as eusociality emerged in two independent lineages of these bees, and that some of these changes may be associated with costs when eusociality is lost.

A molecular model for evolutionary novelty in JH signaling

JH is essential to reproductive maturation in solitary insects, but this signaling system has also been frequently co-opted during major life history transitions[43,46–48], including eusociality[42]. In relatively small eusocial societies like halictids[49,50], paper wasps[51], and bumblebees[52,53], JH has maintained its association with reproductive maturation, but has also gained a new role in mediating aggression and dominance. In the much more elaborate eusocial colonies of honey bees, further modifications to JH signaling have also resulted in a secondary decoupling of JH in workers independent of its ancestral, reproductive role[42]. Although the association between JH titers and division of labor is well established in these taxa, we still do not understand which components of the JH signaling pathways have been targeted or modified by natural selection to evolutionarily link JH and social traits.

Our discovery that JHBPs experience complementary selective pressures when eusociality is gained or lost leads us to propose a model for how this novel association may arise. Specifically, alteration of JHBPs may lead to changes in JH availability in a tissue-specific manner[52], facilitating the decoupling of JH from its ancestral role in reproduction[54] (Fig. 4). Through modification of binding affinity and/or altered cellular uptake of JH, evolutionary changes to *apolpp* and *hex110* could modify the overall and tissue-specific levels of JH[55], with differential responses to JH leading to polyphenisms among individuals[54]. Our model is consistent with theory suggesting that conditional expression of a trait (i.e. eusociality) leads to independent selection pressures and evolutionary divergence[56]; evolution of JHBPs may be one example of such divergence associated with conditional expression of social behavior.

There is a growing body of evidence that Hex110 modulates JH availability via sequestration. Hex110 is associated with caste differentiation in nearly all eusocial insects examined[10,28–30]. In a termite, *Reticulitermes flavipes*, RNAi knockdown of *hex110* leads to an increased proportion of reproductives within the colony, suggesting that Hex110 sequesters JH to prevent it from interacting with downstream targets[28]. Further, gene expression studies in wasps, ants, and honey bees all reveal that abdominal gene expression levels of *hex110* are higher in non-reproducing individuals compared to reproductives, bolstering support for the role of this protein in JH sequestration[29,30,34,57,58]. On the other hand, ApoLpp (a.k.a. lipophorin, apolipoprotein, or lipoprotein) is known to transport its cargo to target tissues via receptor-mediated endocytosis[59] and to cross the blood-brain barrier[60]. Given that ApoLpp has repeatedly been shown to be a major JH binding protein in the hemolymph[8,9], it is likely that ApoLpp also plays a role in JH transport and delivery to tissues, including the brain.

Surprisingly, JH has not previously been characterized in the insect brain, though another major insect hormone, ecdysone, was recently shown to be actively transported into the brain of *Drosophila*[61,62]. We used liquid chromatography coupled mass spectrometry (LC-MS) to measure JH titers in several bee species and found appreciable levels of JH III in all tissues across species, including the brain (Fig. 4a; Fig. S8). In addition, topical application of deuterated JH III (JH III-D3) to the abdomens of *Bombus impatiens* led to incorporation of JH III-D3 into the brain, with an average of 9.5% of total brain JH being of the deuterated form after 24 hours (Fig. 4b). Our results

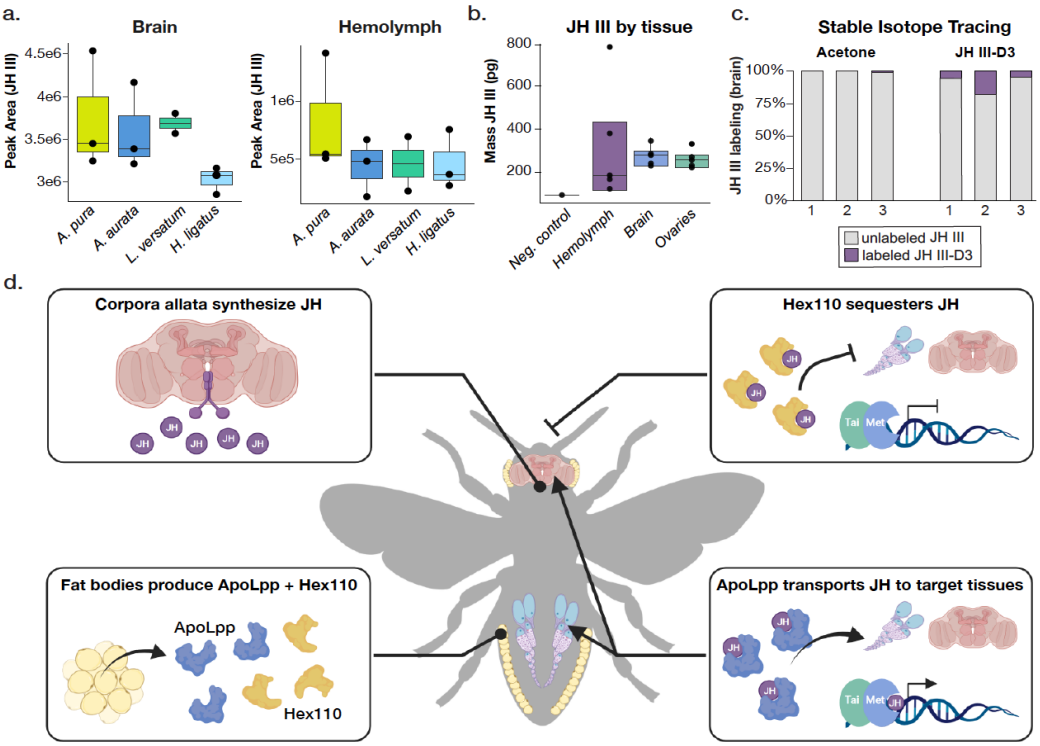

Figure 4. Modifications to JH transport and sequestration could couple behavior with reproduction in the origins of eusociality. (A) Peak areas of JH III in the brain and hemolymph of four halictid species. In all samples, we find detectable levels of JH III in both tissues. n=3 for each tissue of *A. pura, A. aurata,* and *H. ligatus*, and n=2 for each tissue from *L. versatum*. Boxes show median and 25th and 75th percentiles. Whiskers show minimum and maximum values. (B) Tissue samples collected from adult bumblebees confirm there is JH III in the hemolymph, brain, and ovaries (Fig. S8, S10). Fresh dissection buffer (PBS) was used as a negative control. (C) Stable isotope tracing reveals that labelled JH III (JH III-D3) applied onto the abdomens of bumblebees is transported to the brain and detectable by LC-MS 24 hours later. Bees treated with acetone were used as controls. n=3 bee tissues per experimental condition. (D) Hypothetical molecular model. JH is synthesized in the corpora allata (CA) and exerts downstream effects on reproduction and behavior in social insects. We find convergent positive selection on the two primary JH binding proteins[7–9], *apolpp* and *hex110*, when social behavior is gained in halictid bees, and we also find evidence for relaxation of selection on these genes when social behavior is subsequently lost. Hex110 is a storage protein that appears to sequester its cargo rather than deliver it to tissues[10,28]. In contrast, ApoLpp delivers cargo to target tissues, including the brain[59,60]. Together, these results suggest that these two genes could modulate JH availability in a tissue-specific manner[52] – a plausible mechanism by which social behavioral traits (like dominance) can become coupled to ancestrally-reproductive pathways when eusociality arises. Created with BioRender.com.

reveal that JH is indeed present in insect brains and can cross the blood-brain barrier.

Following transport to relevant tissues, JH binds to the JH receptor, Met, to initiate downstream effects[63]. Met is a bHLH-PAS transcription factor that has been well-characterized in flies[63,64], mosquitoes[65], *Tribolium* flour beetles[66], and silk moths[67]. It is coactivated by Taiman (Tai) as the Met/Tai complex, which initiates downstream transcriptional responses. We identified *met/tai* as one of the TF motifs correlated with sociality in halictids (Fig. 3b), and we find evidence for positive selection on *tai* on the Augochlorini origin branch, suggesting that modifications to JH response-elements have also helped to fine-tune JH signaling in this group of bees.

This model provides a testable hypothesis for how JH signaling could be co-opted during evolutionary transitions to and from eusociality. Our findings suggest that the JHBPs, *hex110* and *apolpp*, may function together to alter the availability and uptake of JH by the brain, ovaries, and other tissues, and that modifications to these proteins could generate novel relationships between behavioral and reproductive traits – a key feature of the origins and evolution of eusociality[54].

Conclusions

Sweat bees repeatedly traversed an evolutionary inflection point between a solitary lifestyle and a caste-based eusocial one with multiple gains and losses of this trait. We developed powerful new comparative genomic resources for this group and identified complementary signatures of convergent selection associated with the emergence and breakdown of eusociality. Factors associated with the origins or maintenance of eusociality are also associated with its loss, indicating that there may be trade-offs, constraints, or costs associated with these genomic changes. Strikingly, we find that the functional domains of juvenile hormone binding proteins have been modified by selection as eusociality has been gained and lost in halictids, and that these modifications are likely to affect the transport and availability of JH. Coupled with our finding that JH is present in the insect brain, our results help to explain how the evolution of a major transition is driven by novel linkages between social behaviors and reproduction that have repeatedly evolved alongside eusociality in bees.

Author Contributions

BERR, BMJ, and SDK designed the study. BERR, BMJ, and SDK drafted the initial manuscript. Sample collections were performed by BERR, ESW, BMJ, SDK, MFO, KMK, and RJP. Genomic and transcriptomic libraries were generated by BERR and CK. Hi-C sample processing, library preparation, and assembly: OD, BGS, MP, ADO, and DW. Genome annotation and alignment was performed by BERR and LP. Genomic data was analyzed by BERR, BMJ, SDK, OD, KMK, BGH, JS, and FV. Database generation and curation was performed by AEW. Laboratory experiments were performed by WL and ESW; LC-MS experiments and analyses conducted by SRJ, KEG, and SMD. ELA supervised Hi-C data generation and analysis. SDK supervised the project. All authors contributed to the work presented in this manuscript, discussed the results and implications of this study, and commented on the manuscript.

Acknowledgments

We would like to thank our many colleagues that contributed samples and field support to this dataset, including: Jason Gibbs, Sam Droege, Raphael Jeanson, Joan Milam, Jakub Straka, Marion Podolak, James Cane, and Mallory Hagadorn. We also thank Miriam Richards, Cecile Plateaux-Quenu, Jason Gibbs, and Laurence Packer for discussion and insights on halictid life history and behavior. Tim Sackton, Russ Corbett-Detig, Nathan Clark, Adam Siepel, and Xander Xue provided discussion and advice on data analysis. We also thank Wenfei Tong for the bee drawings and Michael Sheehan for assistance with RNA library preparation. This work was supported by NSF-DEB1754476 awarded to SDK and BGH, NIH 1DP2GM137424-01 to SDK, USDA NIFA postdoctoral fellowship 2018-67012-28085 to BERR, DFG PA632/9 to RJP, a Smithsonian Global Genome Initiative award GGI-Peer-2016-100 to WTW and CK, a Smithsonian Institution Competitive Grants Program for Biogenomics (WTW, KMK, BMJ), a Smithsonian Tropical Research Institute fellowship (to CK), and a gift from Jennifer and Greg Johnson to WTW. MFO was supported by Vicerrectoría de Investigación, UCR, project B7287. ELA was supported by an NSF Physics Frontiers Center Award (PHY1427654), the Welch Foundation (Q-1866), a USDA Agriculture and Food Research Initiative Grant (2017-05741), and an NIH Encyclopedia of DNA Elements Mapping Center Award (UM1HG009375).

Permit Information

Sample permits for collection and import were obtained by SDK, ESW, and MFO (R-055-2017-OT-CONAGEBIO), SDK (P526P-15-04026, and RJP (Belfast City Council, Parks & Leisure Dept).

Data Availability

All sequence data is available in NCBI's SRA database under bioprojects PRJNA613468, PRJNA629833, and PRJNA718331. Accession numbers are given in Table S1. Hi-C data is available on www.dnazoo.org. Code and additional information are available at https://github.com/kocherlab/HalictidCompGen. All genomes, annotations, and alignments can be viewed and downloaded from the Halictid Genome Browser (beenomes.princeton.edu).

References

1. Kocher, S. D. & Paxton, R. J. Comparative methods offer powerful insights into social evolution in bees. *Apidologie* **45**, 289–305 (2014).
2. Kocher, S. D. *et al.* The genetic basis of a social polymorphism in halictid bees. *Nat. Commun.* **9**, 4338 (2018).
3. Gibbs, J., Brady, S. G., Kanda, K. & Danforth, B. N. Phylogeny of halictine bees supports a shared origin of eusociality for *Halictus* and *Lasioglossum* (Apoidea: Anthophila: Halictidae). *Mol. Phylogenet. Evol.* **65**, 926–939 (2012).
4. Danforth, B. N., Conway, L. & Ji, S. Phylogeny of eusocial *Lasioglossum* reveals multiple losses of eusociality within a primitively eusocial clade of bees (Hymenoptera: Halictidae). *Syst. Biol.* **52**, 23–36 (2003).
5. Johnson, B. R. & Tsutsui, N. D. Taxonomically restricted genes are associated with the evolution of sociality in the honey bee. *BMC Genomics* **12**, 164 (2011).
6. Kapheim, K. M. *et al.* Social evolution. Genomic signatures of evolutionary transitions from solitary to group living. *Science* **348**, 1139–1143 (2015).
7. Engelmann, F. & Mala, J. The interactions between juvenile hormone (JH), lipophorin, vitellogenin, and JH esterases in two cockroach species. *Insect Biochem. Mol. Biol.* **30**, 793–803 (2000).
8. de Kort, C. A. D. & Koopmanschap, A. B. Molecular characteristics of lipophorin, the juvenile hormone-binding protein in the hemolymph of the Colorado potato beetle. *Arch. Insect Biochem. Physiol.* **5**, 255–269 (1987).
9. Sevala, V. L., Bachmann, J. A. S. & Schal, C. Lipophorin: A hemolymph juvenile hormone binding protein in the german cockroach, *Blattella germanica*. *Insect Biochem. Mol. Biol.* **27**, 663–670 (1997).
10. Martins, J. R., Nunes, F. M. F., Cristino, A. S., Simões, Z. L. P. & Bitondi, M. M. G. The four hexamerin genes in the honey bee: structure, molecular evolution and function deduced from expression patterns in queens, workers and drones. *BMC Mol. Biol.* **11**, 23 (2010).

11. Ismail, S. M. & Gillott, C. Identification, characterization, and developmental profile of a high molecular weight, juvenile hormone-binding protein in the hemolymph of the migratory grasshopper, *Melanoplus sanguinipes*. *Arch. Insect Biochem. Physiol.* **29**, 415–430 (1995).
12. Nene, V. *et al.* Genome sequence of *Aedes aegypti*, a major arbovirus vector. *Science* **316**, 1718–1723 (2007).
13. Dudchenko, O. *et al.* De novo assembly of the *Aedes aegypti* genome using Hi-C yields chromosome-length scaffolds. *Science* **356**, 92–95 (2017).
14. Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V. & Zdobnov, E. M. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **31**, 3210–3212 (2015).
15. Armstrong, J. *et al.* Progressive alignment with Cactus: a multiple-genome aligner for the thousand-genome era. *Cold Spring Harbor Laboratory* 730531 (2019) doi:10.1101/730531.
16. Smith, M. D. *et al.* Less is more: an adaptive branch-site random effects model for efficient detection of episodic diversifying selection. *Mol. Biol. Evol.* **32**, 1342–1353 (2015).
17. Sung, E. J. *et al.* Cytokine signaling through *Drosophila* Mthl10 ties lifespan to environmental stress. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 13786–13791 (2017).
18. Otto, N. *et al.* The sulfite oxidase Shopper controls neuronal activity by regulating glutamate homeostasis in *Drosophila* ensheathing glia. *Nat. Commun.* **9**, 3514 (2018).
19. Wcislo, W. T. & Danforth, B. N. Secondarily solitary: the evolutionary loss of social behavior. *Trends Ecol. Evol.* **12**, 468–474 (1997).
20. Wertheim, J. O., Murrell, B., Smith, M. D., Kosakovsky Pond, S. L. & Scheffler, K. RELAX: detecting relaxed selection in a phylogenetic framework. *Mol. Biol. Evol.* **32**, 820–832 (2015).
21. Wang, M., Zhao, Y. & Zhang, B. Efficient test and visualization of multi-set intersections. *Sci. Rep.* **5**, 16923 (2015).
22. Wittwer, B. *et al.* Solitary bees reduce investment in communication compared with their social relatives. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 6569–6574 (2017).
23. Lahti, D. C. *et al.* Relaxed selection in the wild. *Trends Ecol. Evol.* **24**, 487–496 (2009).
24. Martin, R. M. & Cardoso, M. C. Chromatin condensation modulates access and binding of nuclear proteins. *FASEB J.* **24**, 1066–1072 (2010).
25. Loiseau, P., Davies, T., Williams, L. S., Mishima, M. & Palacios, I. M. *Drosophila* PAT1 is required for Kinesin-1 to transport cargo and to maximize its motility. *Development* **137**, 2763–2772 (2010).
26. Lobel, J. H., Tibble, R. W. & Gross, J. D. Pat1 activates late steps in mRNA decay by multiple mechanisms. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 23512–23517 (2019).
27. Hidayat, P. & Goodman, W. G. Juvenile hormone and hemolymph juvenile hormone binding protein titers and their interaction in the hemolymph of fourth stadium *Manduca sexta*. *Insect Biochem. Mol. Biol.* **24**, 709–715 (1994).
28. Zhou, X., Tarver, M. R. & Scharf, M. E. Hexamerin-based regulation of juvenile hormone-dependent gene expression underlies phenotypic plasticity in a social insect. *Development* **134**, 601–610 (2007).
29. Hawkings, C., Calkins, T. L., Pietrantonio, P. V. & Tamborindeguy, C. Caste-based differential transcriptional expression of hexamerins in response to a juvenile hormone analog in the red imported fire ant (*Solenopsis invicta*). *PLoS One* **14**, e0216800 (2019).
30. Hunt, J. H. *et al.* A diapause pathway underlies the gyne phenotype in *Polistes* wasps, revealing an evolutionary route to caste-containing insect societies. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 14020–14025 (2007).
31. Wyatt, C. D. R. *et al.* Genetic toolkit for sociality predicts castes across the spectrum of social complexity in wasps. *Cold Spring Harbor Laboratory* 2020.12.08.407056 (2020) doi:10.1101/2020.12.08.407056.
32. Chen, S., Krinsky, B. H. & Long, M. New genes as drivers of phenotypic evolution. *Nat. Rev. Genet.* **14**, 645–660 (2013).
33. Johnson, B. R. Taxonomically restricted genes are fundamental to biology and evolution. *Front. Genet.* **9**, 407 (2018).
34. Warner, M. R., Qiu, L., Holmes, M. J., Mikheyev, A. S. & Linksvayer, T. A. Convergent eusocial evolution is based on a shared reproductive groundplan plus lineage-specific plastic genes. *Nat. Commun.* **10**, 2651 (2019).
35. Harpur, B. A. *et al.* Population genomics of the honey bee reveals strong signatures of positive selection on worker traits. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 2614–2619 (2014).
36. Ferreira, P. G. *et al.* Transcriptome analyses of primitively eusocial wasps reveal novel insights into the evolution of sociality and the origin of alternative phenotypes. *Genome Biol.* **14**, R20 (2013).
37. Simola, D. F. *et al.* Social insect genomes exhibit dramatic evolution in gene composition and regulation while preserving regulatory features linked to sociality. *Genome Res.* **23**, 1235–1247 (2013).
38. Sinha, S., Liang, Y. & Siggia, E. Stubb: a program for discovery and analysis of cis-regulatory modules. *Nucleic Acids Res.* **34**, W555–9 (2006).
39. Siepel, A. *et al.* Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Res.* **15**, 1034–1050 (2005).
40. Dunbar, R. I. M. The social brain hypothesis. *Evol. Anthropol.* **6**, 178–190 (1998).
41. O'Donnell, S. *et al.* Distributed cognition and social brains: reductions in mushroom body investment accompanied the origins of sociality in wasps (Hymenoptera: Vespidae). *Proceedings of the Royal Society B: Biological Sciences* **282**, 20150791 (2015).
42. Hartfelder, K. Insect juvenile hormone: from 'status quo' to high society. *Braz. J. Med. Biol. Res.* **33**, 157–177 (2000).
43. Nijhout, H. F. & Wheeler, D. E. Juvenile hormone and the physiological basis of insect polymorphisms. *Q. Rev. Biol.* **57**, 109–133 (1982).
44. Murrell, B. *et al.* Detecting individual sites subject to episodic diversifying selection. *PLoS Genet.* **8**, e1002764 (2012).
45. Burmester, T. Evolution and function of the insect hexamerins. *Eur. J. Entomol.* **96**, 213–226 (1999).
46. Corbitt, T. S. & Hardie, J. Juvenile hormone effects on polymorphism in the pea aphid, *Acyrthosiphon pisum*. *Entomol. Exp. Appl.* **38**, 131–135 (1985).
47. Zera, A. J. & Denno, R. F. Physiology and ecology of dispersal polymorphism in insects. *Annu. Rev. Entomol.* **42**, 207–230 (1997).
48. Hartfelder, K. & Emlen, D. J. 11 - Endocrine Control of Insect Polyphenism. in *Insect Endocrinology* (ed. Gilbert, L. I.) 464–522 (Academic Press, 2012).
49. Smith, A. R., Kapheim, K. M., Pérez-Ortega, B., Brent, C. S. & Wcislo, W. T. Juvenile hormone levels reflect social opportunities in the facultatively eusocial sweat bee *Megalopta genalis* (Hymenoptera: Halictidae). *Horm. Behav.* **63**, 1–4 (2013).
50. Bell, W. J. Factors controlling initiation of vitellogenesis in a primitively social bee, *Lasioglossum zephyrum* (Hymenoptera: Halictidae). *Insectes Soc.* **20**, 253–260 (1973).
51. Tibbetts, E. A. & Izzo, A. S. Endocrine mediated phenotypic plasticity: condition-dependent effects of juvenile hormone on dominance and fertility of wasp queens. *Horm. Behav.* **56**, 527–531 (2009).
52. Shpigler, H. Y. *et al.* Juvenile hormone regulates brain-reproduction tradeoff in bumble bees but not in honey bees. *Horm. Behav.* **126**, 104844 (2020).
53. Pandey, A., Motro, U. & Bloch, G. Juvenile hormone interacts with multiple factors to modulate aggression and dominance in groups of orphan bumble bee (Bombus terrestris) workers. *Horm Behav* **117**, 104602 (2020).
54. West-Eberhard, M. J. Wasp societies as microcosms for the study of development and evolution. *Natural history and evolution of paper wasps* **290**, 317 (1996).
55. Nijhout, H. F. & Reed, M. C. A mathematical model for the regulation of juvenile hormone titers. *J. Insect Physiol.* **54**, 255–264 (2008).
56. West-Eberhard, M. J. Phenotypic plasticity and the origins of diversity. *Annu. Rev. Ecol. Syst.* **20**, 249–278 (1989).

57. Colgan, T. J. *et al.* Polyphenism in social insects: insights from a transcriptome-wide analysis of gene expression in the life stages of the key pollinator, *Bombus terrestris*. *BMC Genomics* **12**, 623 (2011).
58. Hoffman, E. A. & Goodisman, M. A. D. Gene expression and the evolution of phenotypic diversity in social wasps. *BMC Biol.* **5**, 23 (2007).
59. Van Hoof, D., Rodenburg, K. W. & van der Horst, D. J. Lipophorin receptor-mediated lipoprotein endocytosis in insect fat body cells. *J. Lipid Res.* **44**, 1431–1440 (2003).
60. Brankatschk, M. & Eaton, S. Lipoprotein particles cross the blood-brain barrier in *Drosophila*. *J. Neurosci.* **30**, 10441–10447 (2010).
61. Okamoto, N. *et al.* A Membrane transporter is required for steroid hormone uptake in *Drosophila*. *Dev. Cell* **47**, 294–305.e7 (2018).
62. Okamoto, N. & Yamanaka, N. Steroid hormone entry into the brain requires a membrane transporter in *Drosophila*. *Curr. Biol.* **30**, 359–366.e3 (2020).
63. Charles, J.-P. *et al.* Ligand-binding properties of a juvenile hormone receptor, *Methoprene-tolerant*. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 21128–21133 (2011).
64. Miura, K., Oda, M., Makita, S. & Chinzei, Y. Characterization of the *Drosophila* Methoprene-tolerant gene product: Juvenile hormone binding and ligand-dependent gene regulation. *FEBS J.* **272**, 1169–1178 (2005).
65. Li, M., Mead, E. A. & Zhu, J. Heterodimer of two bHLH-PAS proteins mediates juvenile hormone-induced gene expression. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 638–643 (2011).
66. Zhang, Z., Xu, J., Sheng, Z., Sui, Y. & Palli, S. R. Steroid receptor co-activator is required for juvenile hormone signal transduction through a bHLH-PAS transcription factor, *Methoprene Tolerant*. *J. Biol. Chem.* **286**, 8437–8447 (2011).
67. Kayukawa, T. *et al.* Transcriptional regulation of juvenile hormone-mediated induction of Krüppel homolog 1, a repressor of insect metamorphosis. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 11729–11734 (2012).