Developmental plasticity shapes social traits and selection in a facultatively eusocial bee

Karen M. Kapheim[a,b,1], Beryl M. Jones[c], Hailin Pan[d,e,f], Cai Li[g], Brock A. Harpur[h], Clement F. Kent[i], Amro Zayed[i], Panagiotis Ioannidis[j], Robert M. Waterhouse[k,l], Callum Kingwell[b,m], Eckart Stolle[n], Arián Avalos[o], Guojie Zhang[d,e,f], W. Owen McMillan[b], and William T. Wcislo[b]

[a]Department of Biology, Utah State University, Logan, UT 84322; [b]Smithsonian Tropical Research Institute, 0843-03092 Panama City, Republic of Panama; [c]Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ 08544; [d]State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, 650223 Kunming, China; [e]China National Genebank, BGI-Shenzhen, Shenzhen, 518083 Guangdong, China; [f]Centre for Social Evolution, Department of Biology, University of Copenhagen, DK-2100 Copenhagen, Denmark; [g]Bioinformatics and Computational Biology Laboratory, The Francis Crick Institute, NW1 1AT London, United Kingdom; [h]Department of Entomology, Purdue University, West Lafayette, IN 47907; [i]Department of Biology, York University, Toronto, ON M3J 1P3, Canada; [j]Foundation for Research and Technology Hellas, Institute of Molecular Biology and Biotechnology, Vassilika Vouton, 70013 Heraklion, Greece; [k]Department of Ecology and Evolution, University of Lausanne, 1015 Lausanne, Switzerland; [l]Swiss Institute of Bioinformatics, University of Lausanne, 1015 Lausanne, Switzerland; [m]Department of Neurobiology and Behavior, Cornell University, Ithaca, NY 14853; [n]Institute of Biology, Martin Luther University Halle-Wittenberg, 06120 Halle (Saale), Germany; and [o]Honey Bee Breeding, Genetics, and Physiology Laboratory, US Department of Agriculture-Agricultural Research Service, Baton Rouge, LA 70820

Edited by Joan E. Strassmann, Washington University in St. Louis, St. Louis, MO, and approved April 29, 2020 (received for review January 9, 2020)

Developmental plasticity generates phenotypic variation, but how it contributes to evolutionary change is unclear. Phenotypes of individuals in caste-based (eusocial) societies are particularly sensitive to developmental processes, and the evolutionary origins of eusociality may be rooted in developmental plasticity of ancestral forms. We used an integrative genomics approach to evaluate the relationships among developmental plasticity, molecular evolution, and social behavior in a bee species (*Megalopta genalis*) that expresses flexible sociality, and thus provides a window into the factors that may have been important at the evolutionary origins of eusociality. We find that differences in social behavior are derived from genes that also regulate sex differentiation and metamorphosis. Positive selection on social traits is influenced by the function of these genes in development. We further identify evidence that social polyphenisms may become encoded in the genome via genetic changes in regulatory regions, specifically in transcription factor binding sites. Taken together, our results provide evidence that developmental plasticity provides the substrate for evolutionary novelty and shapes the selective landscape for molecular evolution in a major evolutionary innovation: Eusociality.

social evolution | *Megalopta genalis* | transcription factor binding | genetic accommodation | gene regulation

Changing conditions during development can yield new phenotypes from a single genotype, without an underlying mutation or other sources of genetic variation (1–3). These novel phenotypes are exposed to selection and may confer fitness benefits or costs. This suggests that developmental variation can play a role in evolution, but in order to do so, mechanisms producing novel phenotypes must be written into the genetic code in a process called genetic accommodation (2). This evolutionary process, where novel phenotypes arise first through developmental plasticity, complements the more well-known pattern where novel phenotypes first arise through genetic variants (4). In both cases, beneficial and harmful novelties are sorted by selection, helping to shape patterns of adaptive evolution, but the key difference lies in the initial source of phenotypic variation. An outstanding question in evolutionary biology is thus how often phenotypic selection arising from environmental variance is translated into genetic change, and thus the extent to which plasticity in preexisting developmental pathways helps shape the evolution of novel phenotypes (2, 4–12). Answering this question requires a better understanding of the mechanisms and fitness consequences of development as they relate to the expression of novel traits (8, 12–15).

Developmental plasticity is likely important in the evolution of animal behavior, because behavioral traits are sensitive to environmental cues (15–18), and changes in behavior promote selection on associated physiological and morphological characters (16, 17, 19, 20). Eusociality is a complex phenotype that represents a major evolutionary innovation in animal behavior. It involves coordinated changes in a suite of traits, resulting in reproductive queens and workers that forego reproduction and cooperate to help raise their siblings. A hallmark of eusociality is that extreme differences in physiology, morphology, and behavior of queens and workers develop from a shared genome in response to varying environmental conditions (e.g., nutrition,

EVOLUTION

Significance

Developmental processes are an important source of phenotypic variation, but the extent to which this variation contributes to evolutionary change is unknown. We used integrative genomic analyses to explore the relationship between developmental and social plasticity in a bee species that can adopt either a social or solitary lifestyle. We find genes regulating this social flexibility also regulate development, and positive selection on these genes is influenced by their function during development. This suggests that developmental plasticity may influence the evolution of sociality. Our additional finding of genetic variants linked to differences in social behavior sheds light on how phenotypic variation derived from development may become encoded into the genome, and thus contribute to evolutionary change.

Author contributions: K.M.K., B.M.J., B.A.H., C.F.K., A.Z., E.S., G.Z., W.O.M., and W.T.W. designed research; K.M.K., B.M.J., H.P., C.L., B.A.H., C.F.K., A.Z., P.I., R.M.W., C.K., E.S., A.A., and G.Z. performed research; K.M.K., B.M.J., B.A.H., C.F.K., P.I., R.M.W., E.S., and A.A. analyzed data; and K.M.K., B.M.J., H.P., B.A.H., C.F.K., R.M.W., E.S., and W.T.W. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

This open access article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

Data deposition: The genomic data are available at the National Center for Biotechnology Information Bioproject, https://www.ncbi.nlm.nih.gov/bioproject/ (*Megalopta genalis* genome assembly accession no. PRJNA494872; raw sequences from the population genetic study accession no. PRJNA494983). Genome annotation files are available at Zenodo (DOI: 10.5281/zenodo.3735554). Other results are available in Datasets S1–S5. Code is available at https://github.com/kapheimlab/mgen_genome_plasticity.

[1]To whom correspondence may be addressed. Email: karen.kapheim@usu.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2000344117/-/DCSupplemental.

First published May 29, 2020.

Downloaded by guest on April 7, 2021

social crowding). Developmental plasticity is thus central to evolutionary elaborations of eusociality, and plays a critical role in generating eusocial phenotypes. However, preexisting variation in developmental processes may also have facilitated the origins of eusociality (21). Like other complex behaviors (22, 23), the evolution of eusociality involves coordinated shifts in the timing and tissue-specificity of expression among conserved sets of interacting genes (24). Research with social ants, bees, and wasps has shown that social behavior is associated with differential expression of hundreds of interacting genes, many of which are ancient and pleiotropic (25–28). This suggests that the genes regulating social behavior in eusocial species likely descended from genes with different functions in the solitary ancestors that gave rise to eusociality.

One way that existing genes may become repurposed for new functions is through regulatory change, so they are expressed in new spatiotemporal and epistatic contexts (2). Comparative studies have shown that genes with regulatory functions are under selection in eusocial species (29, 30). Furthermore, evolutionary transitions from solitary to eusocial lifestyles are associated with changes in how gene-expression networks are organized (31), specifically with increased binding potential of transcription factors in the promoter regions of genes in conserved hormone pathways (30). This suggests that social traits may emerge from evolutionary changes in the way ancient genes are expressed.

Gene networks that regulate developmental plasticity may be a nexus of such regulatory changes that enable the evolution of social traits from ancient genes. Many of the genes underlying social behavior in insects respond to variable developmental conditions (e.g., diet, aggression, hormone levels) to produce adults with different behavioral traits (e.g., workers, queens) (32–36). Accordingly, the leading hypotheses proposed to explain how eusociality evolves from a solitary ancestor evoke developmental processes as a source of initial phenotypic variation for traits related to social behavior. These developmental processes include diapause (37), developmental heterochrony (38), and reproductive or endocrine ground plans (39–41). In support of these hypotheses, genes that are differentially expressed between ant and honey bee castes are also differentially expressed during reproductive development in flies (42). Furthermore, genes with a conserved role in insect sex differentiation (i.e., *feminizer*, *doublesex*) are differentially expressed between castes in ants (43), stingless bees (44), and honey bees (36, 45). These observations suggest genes that regulate development have been co-opted for eusociality, but the evolutionary processes by which this may have occurred is unknown.

Alternatively, if social evolution is initiated by novel genetic variants (e.g., mutation), then genes underlying eusociality should initially be limited in number and taxonomic breadth (46). Evidence for this comes from the finding that taxonomically restricted honey bee genes are enriched among those with a signature of positive selection (47) and are disproportionately expressed in derived tissues with unique functions in sociality (48). Furthermore, genes that are differentially expressed between social castes, and thus likely play a role in social behavior, are enriched for genes without known orthologs in a variety of social insect species (49–51), and have reduced pleiotropy among ants (52). Deciphering the relative contributions of novel genetic variants and plasticity-led evolution in the origins of eusociality thus requires explicit analysis of the relationship between the genes that regulate developmental plasticity and the genes underpinning eusociality, and how these are shaped by natural selection.

We explored the relationship between development and eusociality in a facultatively eusocial bee, *Megalopta genalis*. Females of this species deploy either a solitary or social strategy, and solitary and eusocial nests co-occur within populations (Fig. 1*A*). Investigating the mechanisms underpinning these alternative strategies provides a window into factors that may have been important at the evolutionary origins of eusociality. We generated a draft genome assembly and annotation for *M. genalis* comparable in quality to other bee genomes (N50 scaffold > 3.6 Mbp, 98% complete BUSCOs) (*SI Appendix*, *Supplemental Results*, Figs. S1–S4, and Tables S1–S4). We aligned RNA sequences from two previous studies of *M. genalis* to this genome assembly to investigate the mechanistic links between development and eusociality (53, 54). Specifically, we tested the hypothesis that eusociality evolves from existing genes that regulate development by comparing gene-expression changes among life stages and sexes (53), with gene-expression differences among social phenotypes within a single life stage and sex (54) (Fig. 1*B*). We approached this hypothesis from a transcriptomic perspective because co-option of developmental pathways for novel functions is likely to involve shifts in gene expression (2, 10).

We also used population genomics to investigate how development influences molecular evolution associated with eusociality. We sequenced whole genomes of nine solitary and nine social nest foundresses from a single population and analyzed over 4.5 M single nucleotide polymorphisms (SNPs) to identify genomic signatures of selection and genetic variants associated with social phenotype. We used these data to test the hypothesis that variation in preexisting developmental processes help shape selection on traits related to eusociality.

Finally, we identified genetic variants associated with social strategy to explore ways in which mechanisms that generate plastic phenotypes can become genetically encoded. Our results provide evidence that the mechanisms underlying social plasticity are shared with ancient developmental processes, such as sex differentiation and metamorphosis, that molecular evolution of genes related to eusociality is influenced by their role in regulating these developmental processes, and that eusociality emerging from developmental plasticity may be written into the genome via genetic variants that alter gene regulatory networks.

Results

Gene-Expression Changes during Development Are Correlated with Gene-Expression Differences among Social Phenotypes. If developmental plasticity facilitates the origins of eusociality, then genes that are differentially expressed among social phenotypes should also be differentially expressed during sex differentiation or metamorphosis. Correlation analysis showed that differences in gene expression, measured as $\log_2$ fold-change (logFC), between sexes and life stages were significantly correlated with gene-expression differences among adult females with different social phenotypes (see Fig. 1*B* for contrasts). As expected, the strongest correlations in gene-expression differences were among social types within similar tissues (e.g., brain, abdomen). All 15 correlation tests of logFC for gene-expression differences between social types were significant for genes expressed in abdominal tissue (e.g., logFC queen vs. worker correlated with logFC queen vs. solitary) and 14 (93%) were significant for genes expressed in brain tissue (Spearman's ρ, Benjamini–Hochberg adjusted $P < 0.001$) (Fig. 2 and Dataset S5). However, there was also a strong correlation of expression differences among social types with expression differences between sexes (i.e., adult males and females). This pattern was tissue-specific, with sex differences in the abdomen correlating most strongly with social differences in the abdomen (six of six significant correlations) (Fig. 2) and, similarly, for brain tissue (five of six significant correlations) (Fig. 2). Contrasts between reproductive social phenotypes (e.g., queen vs. replacement queen) were among the most strongly correlated with expression differences between sexes. This finding suggests that the correlations we observe are not based exclusively on genes expressed in activated ovaries or related to differences in reproduction. We also observed strong

Downloaded by guest on April 7, 2021

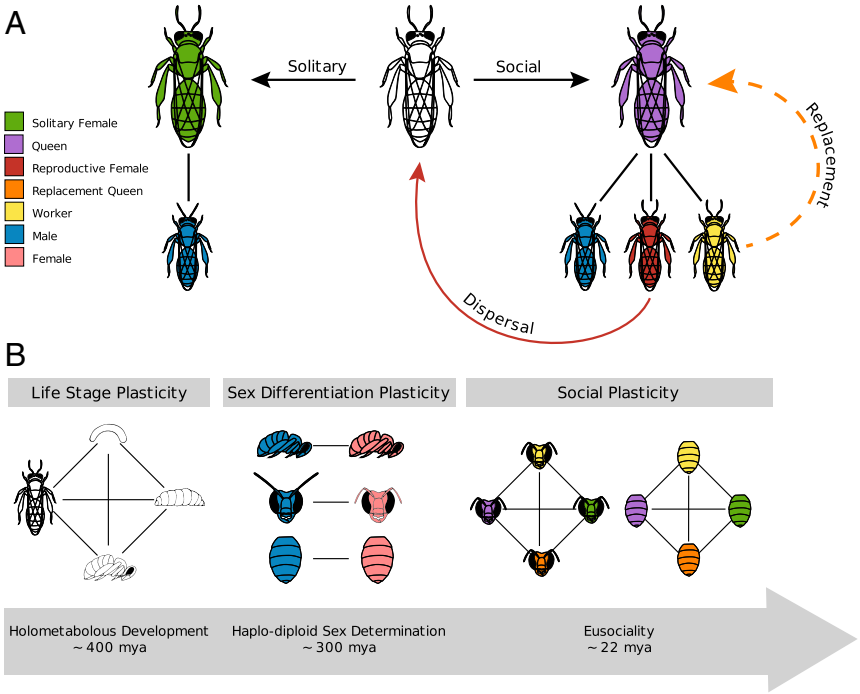

Fig. 1. Study organism and experimental design. (*A*) Schematic of the social biology of *M. genalis*. Foundresses (white) may remain solitary (green) by producing all male (blue) offspring in the first brood or become queens (purple) of eusocial nests. Among eusocial nests, females of the first brood of offspring are typically workers. Reproductive daughters (red) are produced in later broods. If a queen is usurped, dies, or is experimentally removed from the nest, a worker commences egg-laying, becoming a "replacement queen" (orange dashed line) (78, 109). (*B*) Contrasts used to evaluate three types of phenotypic plasticity in *M. genalis*, ranging from least to most recent in origin (110–112). Lines indicate gene-expression contrasts that evaluate life stage plasticity (among eggs, larvae, pupae, and adult abdomens in males and females), sex differentiation plasticity (between males and females in pupae, brains, and abdomens of adults), and social plasticity (among adult females characterized as workers [yellow], queens [purple], solitary females [green], and replacement queens [orange] in brain and abdomens).

correlations among gene-expression differences across life stages and social types, although this pattern was more evident in the abdominal tissue of social types than in brain tissue (Fig. 2). Of 90 correlation tests between gene-expression changes during development (i.e., sex or life-stage contrasts) and gene-expression differences in abdominal tissue among social types, 73 (81%) were significant (Spearman's ρ, Benjamini–Hochberg adjusted $P < 0.001$) (Dataset S5). A majority of correlation tests (64 of 90 [71%]) were also significant when comparing gene-expression differences during development and in the brain tissue of social types. Together, these results demonstrate that genes that are dynamically expressed during development are also dynamically expressed across social types. Conversely, genes for which expression is stable during development are not generally differentially expressed among social types, and thus are not likely to have a role in social behavior.

Plasticity of Developmental Gene Expression Is Predictive of Gene-Expression Plasticity among Social Types. If plasticity in gene networks underlying developmental processes led to novel phenotypes important in the evolution of eusociality, then gene-expression plasticity during development should be predictive of gene-expression plasticity related to social behavior. We tested this prediction by calculating a plasticity index for every gene among developmental stages, sexes, and adult social types in our two independent datasets. This index measures departure from uniform expression for each gene by summarizing the magnitude of expression changes across multiple contexts (i.e., life stages, sexes, social types) (Fig. 1*B*). It is derived from the Euclidean distance (i.e., square root of the sum of squares of logFC), and is thus similar to the mean absolute logFC across conditions, but summing the squares gives large expression differences more weight. Thus, genes with higher expression in a single group have a higher plasticity index than mean absolute logFC, and this provides a better characterization of expression plasticity (55). The plasticity index has been previously used to describe the evolutionary properties of genes that function in social behavior (42, 55), and here we used it to determine how well developmental plasticity predicts social plasticity in terms of gene expression.

We used model selection methods to identify the best predictors of expression plasticity among social types (i.e., social plasticity index for genes expressed in the abdomen or brain). Potential predictor variables included other types of gene-expression plasticity (i.e., plasticity index for developmental stages, sex, and social types in other tissues), evolutionary history of each gene (i.e., orthogroup age, evolutionary rate, duplicability, and copy number variation), and other properties of gene expression (i.e., average expression level, coefficient of expression variation within social types [C.V.]) (Dataset S1).

We found that developmental expression plasticity is highly predictive of social-expression plasticity. Gene-expression plasticity across both sex and life stage were among the strongest significant predictors of expression plasticity between social types in two separate models describing brain and abdominal gene expression ($F_{abdomen} = 723.25$, $P_{abdomen} < 2.2e\text{-}16$, $n = 9{,}284$ genes; $F_{brain} = 1407.60$, $P_{brain} < 2.2e\text{-}16$, $n = 9{,}251$ genes) (Fig. 3 *A*–*D*). Estimated coefficients for stage and sex plasticity were 3 to 19× as high as the next strongest predictor (orthogroup age), even though these data were derived from a completely different set of individuals, sexes, tissues, and developmental stages (Fig. 3 *E* and *F* and *SI Appendix*, Tables S5–S8). A quadratic model best described the relationship between developmental and

Downloaded by guest on April 7, 2021

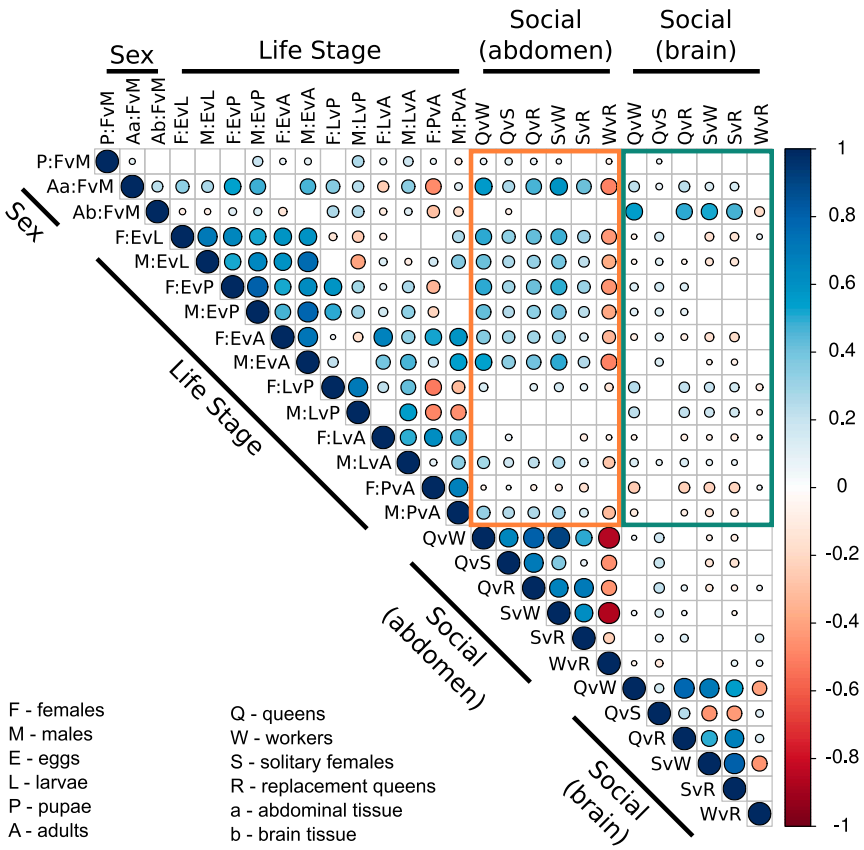

Fig. 2. Correlation of logFC across life stages, sexes, and social phenotypes. Differential expression of genes from each pair of conditions was compared to all other pairs to identify contrasts with significantly similar gene-expression changes. Circle size and color represent correlation strength (Spearman's ρ); positive or negative correlation indicates concordance or discordance in direction of differential expression, depending on order in which the dyad is listed (e.g., QvW vs. WvR in red indicates a similar set of genes is up-regulated in Q and R compared with W); boxes highlight expression changes during development correlated with expression differences in abdominal (orange) or brain (green) tissue among social types; empty cells indicate correlations were not statistically significant (Benjamini–Hochberg adjusted $P < 0.001$); correlation and P values are in Dataset S5.

social plasticity, indicating that genes with moderate levels of developmental plasticity tended to have the highest levels of social plasticity and genes with high levels of developmental plasticity had similar or slightly lower social plasticity (Fig. 3 *A–D*). The linear component of the developmental plasticity indices consistently had the largest model coefficients (Fig. 3 *E* and *F*), indicating that developmental plasticity is a strong positive predictor of social plasticity. Permutation analysis revealed that the observed coefficients derived from model selection were within the range of those obtained when gene-expression data are permuted across individuals (*SI Appendix*, Tables S5 and S6), further suggesting that genes, which are coexpressed in one context (e.g., social types), tend to also be coexpressed in other contexts (e.g., development).

Model results also suggested that the molecular signatures of sociality operate independently across tissues. Another type of plasticity, expression plasticity within tissues and social types (C.V.), was a strong predictor of the social plasticity index in brain tissue (coefficient = 0.42, $t = 49.23$, $P < 2e\text{-}16$), but not abdominal tissue. The plasticity index between social types in brain and abdominal tissues were significant predictors of one another, but the effect size relative to other variables was small, considering these samples came from the same set of individuals (coefficient brain social plasticity on abdomen social plasticity = −0.006; coefficient abdomen social plasticity on brain social plasticity = −0.031) (*SI Appendix*, Tables S7 and S8).

Shared Components of Developmental and Social Gene-Expression Networks Are Enriched for Ancient Genes with Evidence of Positive Selection in Social Species. Development is largely governed by regulatory relationships in gene-expression networks (2, 56). We therefore investigated the relationship between developmental and social gene-expression networks using independent components analysis (ICA). ICA models the expression of each gene in each sample as a linear-weighted sum of independent components (ICs). Each IC represents the effect of independent processes influencing the overall expression profile within each dataset (57) (*SI Appendix*, Figs. S7–S9). We identified 5 ICs for the set of genes differentially expressed in abdominal tissue between two or more social phenotypes (false-discovery rate [FDR] adjusted $P < 0.05$, $|\text{logFC}| \geq 1.2$, $n = 1{,}469$ genes, 55 bees), and 47 ICs for the set of genes differentially expressed between two or more life stages or sexes ($n = 6{,}859$ genes, 50 bees). We next used graph correlational analysis to identify shared processes influencing social and developmental plasticity. This revealed sets of social ICs and developmental ICs that are significantly correlated with one another (reciprocal ICs) (Fig. 3*G*).

If ancestral developmental variation facilitates social evolution, then highly correlated social ICs and developmental ICs should be enriched for conserved genes related to social behavior. In support of this prediction, we find that most reciprocally correlated ICs were significantly enriched for genes that belong to ancient orthogroups that include genes shared by vertebrates (hypergeometric tests: $ICs_{1–47}$, $P = 8.78e^{-07}$; $ICs_{2–14}$, $P = 0.04$; $ICs_{3–9}$, not significant [n.s.]; $ICs_{4–16}$, $P = 8.78e^{-07}$;

Downloaded by guest on April 7, 2021

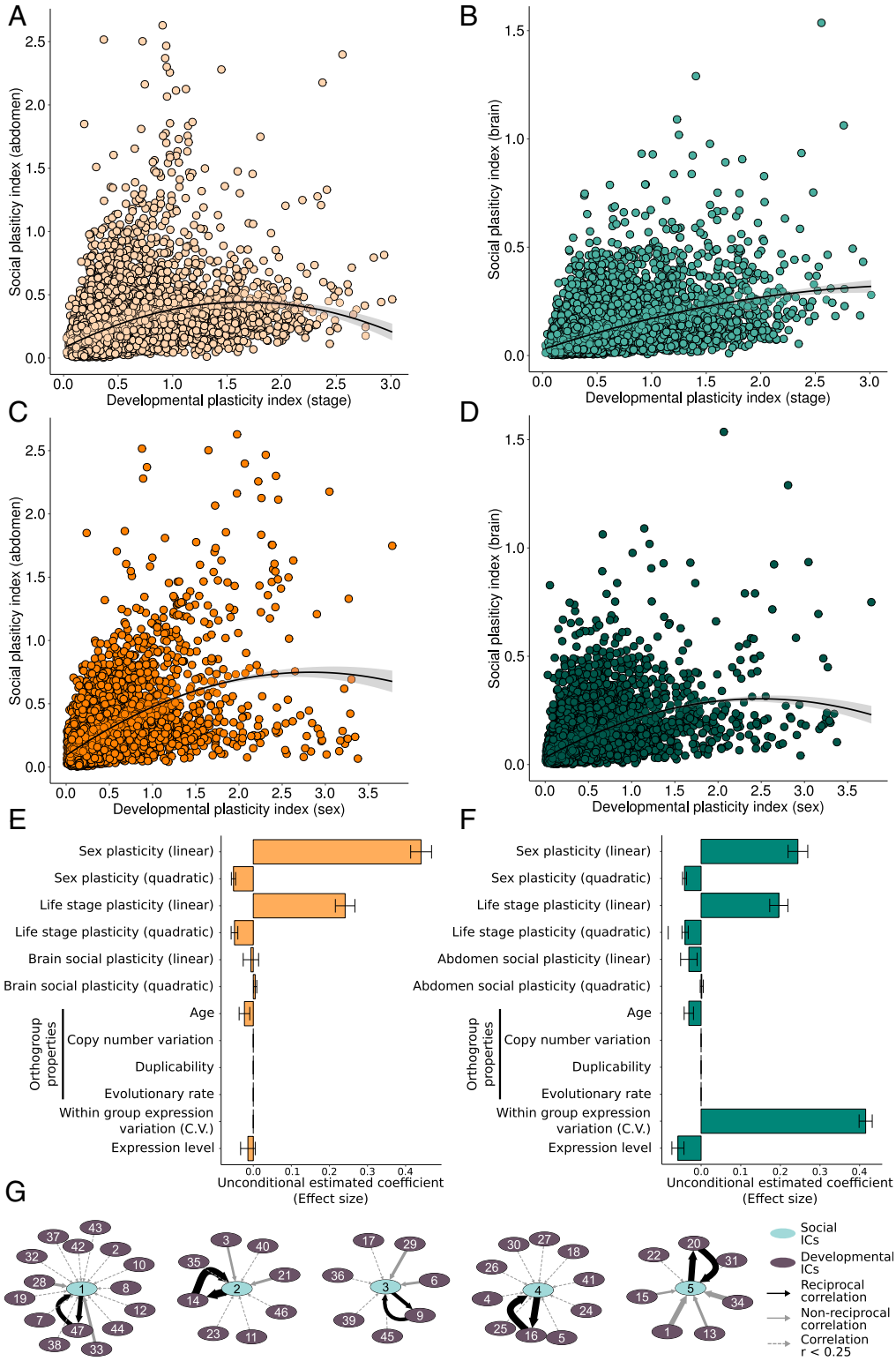

Fig. 3. Gene-expression patterns associated with eusociality are associated with developmental plasticity. (*A*–*F*) Expression plasticity index among life stages and sexes were strong significant predictors in statistical models predicting expression plasticity among adult female social phenotypes for genes expressed in both abdominal (*A*, *C*, and *E*) and brain (*B*, *D*, and *F*) tissues. Social plasticity indices were fit to a quadratic model based on goodness-of-fit tests and the nature of the relationship between social and developmental plasticity indices (*SI Appendix*, Fig. S5). (*A*–*D*) Line is fitted to a quadratic model, shading indicates 95% confidence interval. (*E* and *F*) Unconditional coefficients are estimated from multimodel inference (intercept coefficients not shown), Coefficients for both linear and quadratic terms are shown. (*G*) ICA to infer regulatory module networks in gene-expression data identify ICs associated with social plasticity and developmental plasticity. Social ICs are significantly correlated with ICs activated across developmental life stages, suggesting social gene expression networks are embedded within developmental gene networks. Arrow width indicates correlation strength. Numbers label each IC.

Downloaded by guest on April 7, 2021

$ICs_{5–20}$, $P = 0.003$). Most reciprocally correlated ICs were also enriched for genes with evidence of positive selection in *M. genalis* (see McDonald–Kreitman test results below; hypergeometric tests: $ICs_{1–47}$, $P = 0.003$; $ICs_{2–14}$, $P = 0.0005$; $ICs_{3–9}$, $P = 0.003$; $ICs_{4–16}$, $P = 0.003$; $ICs_{5–20}$, n.s.). Most highly and reciprocally correlated ICs were enriched for genes under positive selection in highly eusocial species compared to solitary species (58) (hypergeometric tests: $ICs_{1–47}$, $P = 3.74e^{-05}$; $ICs_{2–14}$, $P = 0.04$; $ICs_{3–9}$, $P = 0.004$; $ICs_{4–16}$, n.s.; $ICs_{5–20}$, n.s.). These results support the hypothesis that components of gene regulatory networks that function in social behavior are derived from conserved developmental gene networks.

Genes Involved in Sociality Are Enriched for Genes under Positive Selection. We next investigated the evolutionary consequences of developmental processes by exploring the relationship between molecular evolution and developmental and social plasticity. Previous research in ants and bees showed that genes differentially expressed among castes (i.e., caste-biased genes) evolve faster and undergo stronger adaptive evolution than noncaste-biased genes (27, 47, 59–63). However, these studies focused on derived, obligately eusocial species that likely experience different regimes of natural selection than those relevant to the origins of eusociality. In contrast, there was no evidence for rapid molecular evolution in caste-biased genes in *Polistes canadensis* wasps, which have simpler forms of eusociality, although no females nest alone as a viable behavioral alternative (50). Both social and solitary behavior are viable lifestyles for *M. genalis*, which provides the opportunity to evaluate the hypothesis that developmental plasticity influences adaptive evolution relating to the origins of eusociality.

We tested this hypothesis with whole-genome resequencing data from 18 *M. genalis* females and 1 outgroup *M. amoena* female. First, we used model selection methods to identify the best predictors of the neutrality index (NI) for each gene. NI measures the direction and degree of departure from neutral evolution, with NI close to zero indicating strong positive selection and NI > 1 indicating negative selection (64). We found that genes with high levels of social plasticity in both the brain and abdomen had NI close to zero, indicating that genes with the greatest magnitude of expression differences among social phenotypes tend to be under the strongest positive selection (Fig. 4 *A* and *B*). The plasticity index of expression across social phenotypes was the strongest significant predictor of NI (compound Poisson linear model: $t_{\text{abdomen-social-plasticity}} = 2.43$, $P_{\text{abdomen-social-plasticity}} = 0.01$; $t_{\text{brain-social-plasticity}} = 3.79$, $P_{\text{brain-social-plasticity}} = 0.0002$; $n = 8{,}418$ genes) (*SI Appendix*, Table S9). The estimated model coefficient for social plasticity was 1.8× (brain) or 1.1× (abdomen) as high as the next strongest predictor (orthogroup age) (*SI Appendix*, Table S9). All of the genes predicted to be under positive selection with NI also have direction of selection > 0, indicating that our estimates of positive selection were not conflated with weak purifying selection (Dataset S1) (65).

Second, we found significant enrichment of caste-biased genes among those under positive selection in a McDonald–Kreitman test (25% of genes under positive selection are caste-biased, hypergeometric test: $P = 0.03$). Most of these genes (65%) were up-regulated in queen versus worker abdomens, and include genes that function in reproduction in honey bees (e.g., orthologs of the *Drosophila melanogaster* genes *notch* and *yolkless*), indicating that reproductive function is likely under strong selection, as shown for other eusocial bees with a solitary phase prior to the emergence of workers (60, 66).

Developmental Plasticity Helps Shape Selection on Social Traits. We next investigated how the relationship between adaptive evolution and caste-biased gene expression is influenced by developmental processes. If ancestral patterns of developmental plasticity influence selection on social behavior, then signatures of selection on caste-biased genes may stem from their role in nonsocial processes, such as sex determination or metamorphosis. Consistent with this prediction, we found that genes under positive selection (as identified with a McDonald–Kreitman test) were significantly enriched for those exhibiting expression bias among males and females (32%, hypergeometric test: $P = 0.009$) or stages of development (87%, hypergeometric test: $P = 0.005$). All but one of the caste-biased genes under positive selection were also differentially expressed as a function of development (Fig. 4*C*). Together, these results support the hypothesis that selection on genes involved in social behavior is likely influenced by their expression patterns in developmental contexts.

Genetic Variants Associated with Social Phenotype May Influence Gene Regulation. While developmental plasticity can produce new phenotypes that are exposed to selection, the mechanisms that generate these new variants must be encoded in the genome

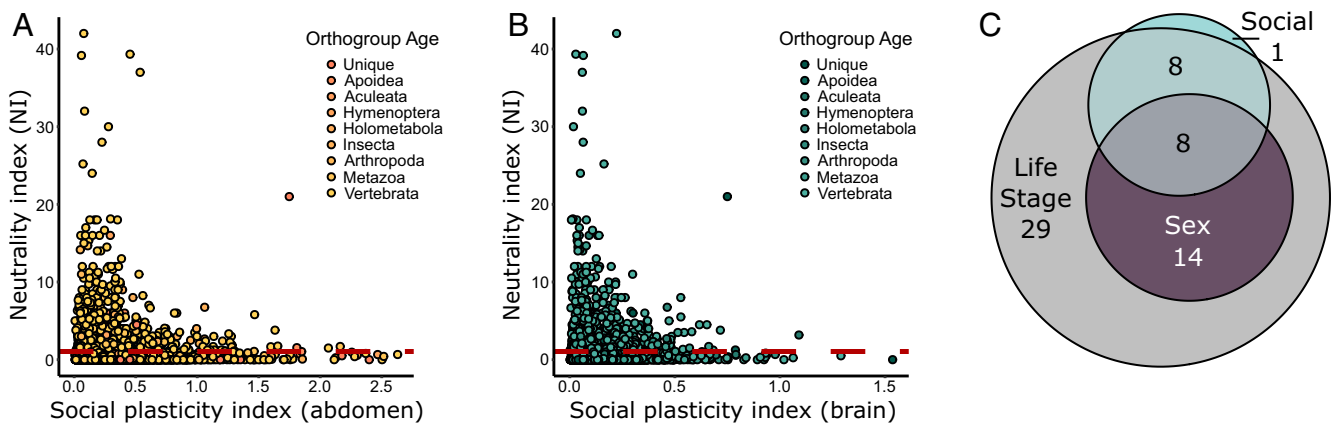

Fig. 4. The relationship between eusociality and molecular evolution is influenced by developmental plasticity. Social plasticity of gene expression in the abdomen (*A*) and brain (*B*) are significant predictors of adaptive evolution, as inferred from analysis of the NI (~0, positive selection; NI > 1, negative selection; red line, NI = 1). The orthogroup age of each *M. genalis* gene reflects the taxonomic span of species with orthologs in each orthogroup. (*C*) Genes under positive directional selection are enriched for genes that are differentially expressed across developmental stages (Life Stage), between the sexes (Sex), and among social types (Social). Numbers in circles indicate the number of positively selected genes that are also differentially expressed in each category of comparison.

Downloaded by guest on April 7, 2021

to contribute to evolution (2). We therefore identified tentative genetic variants associated with *M. genalis* foundress strategy (e.g., social or solitary) (Fig. 1*A*). We identified 26 significantly differentiated regions in the genomes of solitary and social foundresses (45-SNP windows, average $F_{st} > 0.15$, Storey's Q < 0.01) (*SI Appendix*, Figs. S10 and S11 and Dataset S2). Individual SNPs with high F_{st} (>0.2) in these significantly differentiated regions were not enriched for any particular genomic feature (e.g., exons, introns), relative to the total number of SNPs in each type of feature ($\chi^2 = 8.97$, $P = 0.78$).

A likely means by which plasticity is encoded into the genome is through genetic variants that affect regulatory relationships of existing gene networks (2, 67–70). Previous research suggests that the transition from a solitary to eusocial lifestyle likely involved adaptive changes in gene regulation, specifically in transcription factor (TF) binding activity (30). We therefore investigated how genetic variants associated with *M. genalis* foundress strategy within the 26 significantly differentiated regions are predicted to alter TF binding probability. Among SNPs with significantly high F_{st} (>0.2) in promoter regions or introns within these differentiated regions, 32.1% (42 of 131) significantly affected the binding probability of one or more TFs between social forms. These included nearly half of 13 TFs involved in neurogenesis or endocrine-mediated gene expression that have been previously identified as important in social evolution (*tai*/*met*, *CG5180*, *side*, *h*, *br*, and *lola*) (30) (Fig. 5, *SI Appendix*, Table S10, and Datasets S3 and S4). These results suggest genetic variants that confer regulatory effects on gene expression may be an important mode by which plastic traits can become integrated into the genome.

Discussion

Developmental plasticity is an important source of phenotypic novelty, and developmental gene networks that generate novel variants may therefore contribute to the evolution of new traits, even if genetic variation is initially absent (2). This plasticity-led pattern of evolution has been documented less frequently than evolutionary change that is initiated when new phenotypes arise from mutation or other sources of genetic variation. This is particularly true for complex behavioral phenotypes, such as eusociality, despite theoretical and increasing empirical support that plasticity in ancestral developmental mechanisms shapes subsequent evolution (2, 4, 7, 11, 12, 15, 16, 21, 71, 72).

Our findings provide multiple lines of evidence that developmental plasticity may have facilitated the evolutionary origins of eusociality. Contrary to expectations under evolution initiated by mutation or other sources of genetic variants, we do not find evidence that genes underpinning eusociality are taxonomically restricted or limited in number. Instead, we find that the transcriptional underpinnings of social plasticity in a facultatively eusocial sweat bee are highly related to transcriptional plasticity in development. Changes in gene expression during metamorphosis and sex differentiation are highly correlated with differences in gene expression among females that differ in social behavior (Fig. 2), and expression plasticity during development is the strongest predictor of expression plasticity related to social phenotype (Fig. 3 *A*–*F*). Furthermore, gene-expression networks related to social plasticity are embedded within developmental gene networks (Fig. 3*G*), and the most correlated components of these networks are enriched for ancient genes under positive selection in social species that represent multiple independent origins of eusociality. Overall, these results suggest that mechanisms underpinning eusociality are derived from tinkering of developmental gene regulatory networks important for other traits.

Complex behavioral phenotypes, like eusociality, are particularly likely to be influenced by selection acting upon preexisting plasticity (9, 15, 21). However, direct tests of this hypothesis have been limited by the fact that there are no known eusocial populations recently derived from—and contemporaneous with—solitary ones. Many bee species exhibit facultative sociality across latitudinal and altitudinal clines, but numerous biotic and abiotic factors differ for populations in different environments (73). In contrast, species like *M. genalis*, in which solitary and eusocial phenotypes co-occur within a population, are relatively rare, but permit direct comparisons of individuals that express solitary and social behavior under roughly constant ecological conditions. This kind of natural experiment provides a clear comparative picture of the intrinsic mechanisms underlying sociality, without the confounding effects of environmental variation, and thus serves as a proxy for a transitional state that is likely to have occurred in the evolutionary transition from a solitary ancestor (74).

An additional advantage to focusing on a facultatively eusocial species like *M. genalis* is that our analyses are unbiased with regard to gene orthology. Highly conserved genes are more likely to function in ancient processes like metamorphosis or sex differentiation. Thus, cross-species comparisons limited to conserved genes could lead to an inflated estimate for the role of development in eusocial evolution. By focusing on differences in social phenotypes within a single population, we are able to include all genes, even those with no identifiable orthologs, and without a priori knowledge of their function in core processes. Our finding that developmental gene-expression networks provide the transcriptional underpinnings for the expression of social plasticity in *M. genalis* is thus unbiased with regard to gene orthology, and is based on the full complement of genes in this species.

Our results complement existing support for plasticity-led evolution of sociality in this species, which involves both phenotypic and genetic accommodation (2, 21). Phenotypic accommodation is adaptive adjustment of a developmentally induced phenotype (2, 75), and evidence for this comes from detecting preexisting plasticity and cryptic genetic variation in new developmental contexts (4). Evidence that there has been phenotypic accommodation of sociality in *M. genalis* comes from the fact that the facultative expression of sociality is induced by the social environment. Nest-founding females actively adjust the quality and quantity of pollen and nectar provided to their developing daughters (76), and this yields ecologically relevant physiological variation among adults that is reflective of differences between queens and workers (77). Additionally, the presence of queens in the nest suppresses reproductive maturation in her worker daughters (78), likely in part through aggressive behavioral interactions (79–83). Our present results also add to the evidence for phenotypic accommodation by revealing cryptic genetic differentiation among social and solitary females against an otherwise undifferentiated genomic background.

We also found evidence that there has been subsequent modification of sociality through genetic accommodation (in the sense of ref. 11). We identified significant genetic differences in TF binding propensity between social- and solitary-biased alleles, suggesting that the regulation of genes that function in social behavior has been refined through molecular evolution. We additionally found that socially expressed genes are under positive selection, indicating that the regulation of genes related to social behavior is adaptively shaped by evolution in this ancestral proxy population. Moreover, caste-biased genes in *M. genalis* have been previously shown to be under strong positive selection in obligately eusocial species, suggesting that there has been adaptive refinement of these genes in derived lineages (54). Our study thus provides the missing pieces of an emerging picture for plasticity-led social evolution in a facultatively eusocial bee.

Conclusion

Accumulating evidence for a role of developmental plasticity in evolution has led some to criticize the proximate-ultimate

Downloaded by guest on April 7, 2021

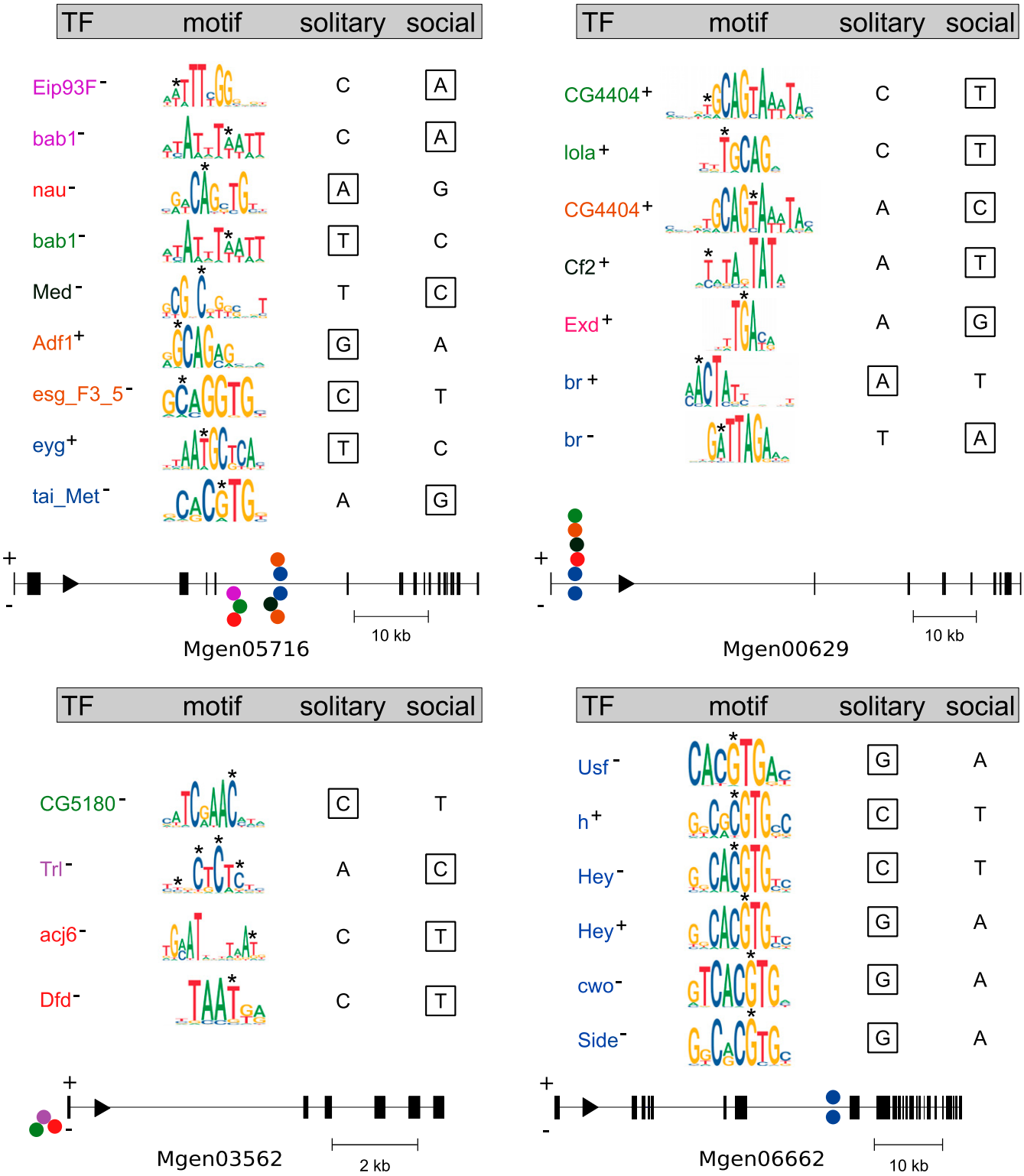

Fig. 5. Examples of altered TF binding probabilities among SNPs in high F_{st} outlier regions near four genes. Filled circles indicate position of SNPs within or near gene models, which affect predicted binding of similarly colored TF(s). Positions within the motif affected by the SNP are denoted with asterisks, and allele version (social or solitary), which is predicted to have motif binding, is marked via squares. Filled boxes in gene models indicate exons, and the filled triangle denotes direction of transcription. Strand with predicted TF binding is denoted with a ± next to the TF name.

framework by which the study of novel traits has advanced for the last 50 y (84). This criticism is based on the assertion that this partition hinders thinking about reciprocal processes that influence both proximate and ultimate explanations, such as developmental plasticity. Perhaps no study of novel phenotypes has been so successful at integrating proximate and ultimate analyses than those in the field of sociogenomics, which aims to study the evolution of social life in molecular terms within a Darwinian inclusive fitness

Downloaded by guest on April 7, 2021

framework (85). Despite the unabated success of this field, some key findings suggest that additional factors need to be addressed for a complete understanding of social evolution. This includes the finding that many pathways regulating social behavior are ancient and highly pleiotropic, and caste-biased gene expression only partially explains observed patterns of molecular evolution (86, 87). Our study shows that existing variation in developmental processes helps fill this gap, providing support for a central role of developmental plasticity as a source of functional variation that can both generate novel phenotypes in major evolutionary transitions, and shape the evolutionary trajectories of these phenotypes.

Materials and Methods

Detailed materials and methods can be found in *SI Appendix*.

Genome Assembly and Annotation. We collected *M. genalis* bees from Barro Colorado Island, Panama, in 2014 and 2015. Genomic DNA from four full sibling males was isolated and sequenced for whole-genome assembly from four Illumina shotgun libraries and three Illumina mate-pair libraries. Over 639 million reads were used for the assembly with SOAPdenovo2 (88). We predicted gene models using both homology-based methods and de novo methods, integrated by GLEAN (89). The *M. genalis* gene set, comprising 12,865 genes, was mapped to the OrthoDB v9 (90) orthology database at the Hymenoptera, Insecta, Arthropoda, and Metazoa nodes in order to delineate shared orthologs between *M. genalis* and other insects and other animals. We detected and quantified repetitive elements in the *M. genalis* genome de novo from short reads and annotated the genome sequence assemblies (contigs and scaffolds ≥ 500 bp; n = 6,057, L = 405.985 Mbp).

Population Genomic Analysis. For population genetic analyses, genomic DNA was isolated from the thoraces of 18 females (9 solitary, 9 social), 1 additional adult male, and 1 female *Megalopta amoena*. *M. amoena* is also facultatively eusocial and lives a similar lifestyle to *M. genalis* on Barro Colorado Island, where this specimen was collected (91). We sequenced shotgun libraries for each sample on a HiSEq. 4000 to an average depth of >39 million reads per individual. Variant identification was conducted following GATK best practices. After filtering variants, we tested for admixture, relatedness, and departure from Hardy–Weinberg equilibrium. We ran individual McDonald–Kreitman tests and estimated the NI for each gene and corrected for FDR using the Simes/Benjamini/Hochberg method (92). We then estimated the selection coefficient for all genes in the genome using SNIPRE (93). We estimated F_{st} between social and solitary nests using a Bayesian implementation to estimate the significance of each F_{st} estimate (94–96). We extracted 26 45-SNP windows that were significant outliers after a permutation test with a stringent FDR and were highly differentiated relative to all SNPs in the genome (Storey's Q < 0.01; F_{st} > 0.15), and extracted all SNPs within and around (±2,000 bp) these windows.

Differential Gene-Expression Analysis. RNA sequences from refs. 53 and 54 were aligned to the *M. genalis* genome assembly with STAR (97). This included the sequences from the whole brains of 7 queens, 7 solitary females, 9 workers, 7 replacement queens and from the abdomens (without the gut) of 7 queens, 7 solitary females, 6 workers, and 5 replacement queens. We also had sequences from 3 female eggs, 1 male egg, 6 female larvae, 2 male larvae, 4 female pupae, 4 male pupae, abdominal and brain tissue of 5 adult females and 10 adult males. These bees were collected directly from natural nests in the field, reared under ambient conditions in outdoor enclosures, or collected from observation nests (also in the field). They were matched for reproductive status within each social category (workers 10 to 24 d old, replacement queens 10 to 31 d old, queens 18 to 137 d old, solitary females 68 to 121 d old). Counts of reads per gene were calculated with featureCounts in SubRead (98). Analysis of differential expression was completed with DESeq2 (99) in R v3.4.4 (100). We calculated a plasticity index for each gene in our expression set (n = 9,284) following Schrader et al. (55). More detail about the plasticity index is available in *SI Appendix*.

Model Selection to Predict Expression Plasticity. We used model selection methods implemented through glmulti (101) in R v3.4.4 (100) to investigate the factors that best predict gene-expression plasticity among social types in both brain and abdominal tissue. We included the following parameters in the full model of social plasticity index in abdominal tissues: Average expression, sex plasticity index (linear and quadratic terms), stage plasticity index (linear and quadratic terms), social plasticity index for the brain (linear and quadratic terms), orthogroup age, coefficient of expression variation in the abdomen across social types (to account for variance in expression unrelated to social type), and metrics of general evolvability, including orthogroup evolutionary rate, orthogroup duplicability, orthogroup copy number variation. The model for social plasticity index in the brain included the same set of parameters, except social plasticity in the abdomen was swapped for social plasticity in the brain, and coefficient of expression variation was specific to the brain. Predictor variables were checked for multicollinearity using the VIF function built by Zuur et al. (102) and a threshold of VIF = 3 (103). We then scaled all predictor variables to have mean of 0 and SD of 1 to allow direct comparison of estimated coefficients. We fit these models with a generalized linear models (GLM; family = γ, link = log), and used Akaike Information Criterion (AIC) to identify the best fitting models. GLM family and link parameters were determined by comparing models fit to different distributions using the plot function in R v3.4.4. We obtained unconditional estimates of model coefficients and 95% confidence intervals with multimodel inference.

Correlation Analysis. We used the function rcorr with the Spearman method in the R package Hmisc v4.1.1 (104). We used p.adjust to apply a Benjamini–Hochberg correction for multiple testing to the P values.

Independent Components Analysis. We ran ICA for each dataset using the clusterFastICARuns function in MineICA (105), with 500 iterations and average linkage with hclust clustering. We compared ICs between the development dataset and the social datasets using a correlation graph approach, implemented with the function runCompareIcaSets. We extracted the top 1,000 contributing genes to reciprocal ICs with the function compareGenes (type = "union", cutoff = 3), and used phyper in R v3.5.1 to test for significant overlap with genes identified as under positive selection in the McDonald–Kreitman test, genes identified as under selection associated with sociality in previous comparative genomics studies (30, 58), and genes belonging to orthogroups that share ancestry with vertebrates.

Model Selection for Predictors of NI. We used compound Poisson GLMs with package cplm (106) in R v3.4.4 (100) to model the NI for each gene. We selected among models with different combinations of parameters based on AIC, and found strong support (ΔAIC >10) for the full model that included average expression, tau index, orthogroup age, and plasticity indices for stage, sex, and social type (both abdomen and brain) with both linear and quadratic terms.

Enrichment Tests of Genes under Positive Selection. We used hypergeometric tests, implemented with phyper in R v3.4.4 (100), to test for significant enrichment among genes that are conditionally expressed (|log FC| > 1.2 and Benjamini–Hochberg adjusted P < 0.05) in each context (sex, stage, social type) with those that were identified as under positive directional selection in the McDonald–Kreitman test.

TF Binding Probability Analysis. We scanned 201-bp regions surrounding intronic SNPs or SNPs upstream of 22 genes in significantly differentiated regions (131 SNPs from Dataset S2 with F_{st} > 0.2 between solitary and social forms) for TF binding sites using FIMO (107). We used F_{st} > 0.2 as a cutoff to restrict our analysis of differences in TF binding probability to those SNPs that are most differentiated between solitary and social foundresses. (F_{st} > 0.2 is approximately the 95% quantile for F_{st} of individual SNPs across the genome and corresponds to P < 0.02.) TF position weight matrices for 223 motifs were obtained from FlyFactorSurvey (108) for representative *Drosophila* motifs, as in ref. 30 (Dataset S3). The SNP of interest was deemed to have an effect on TF binding if only one allele had a significant (P < 0.0001) match to a given TF with FIMO, and/or if the ratio of FIMO scores was greater than |1.5|.

Data Availability. The *M. genalis* genome assembly is available at the National Center for Biotechnology Information (BioProject PRJNA494872). Raw sequences from the population genetic study are available at National Center for Biotechnology Information (BioProject PRJNA494983). Genome annotation files are available at Zenodo (DOI: 10.5281/zenodo.3735554). Other results are available as Datasets S1–S5. Code is available at https://github.com/kapheimlab/mgen_genome_plasticity.

ACKNOWLEDGMENTS. We thank G. Trujillo and E. Velasquez for assistance in the field; L. Schrader and J. Oettler for providing advice on use of the plasticity index; C. A. Blatti III and L. Sloofman for providing curated transcription factor motif weight matrices; and M. J. West-Eberhard, G. E. Robinson, A. P. Moczek, and three anonymous reviewers for providing

Downloaded by guest on April 7, 2021

insightful feedback on earlier versions of this manuscript. Sequencing was performed at the University of Illinois at Urbana–Champaign (UIUC) Roy J. Carver Biotechnology Center. Computational support was provided by the University of Utah Center for High Performance Computing and the UIUC Computer Network Resource Group/Biocluster. J. Johnson (Life Sciences Studios) created the illustrations in Fig. 1. This work was supported by the United States Department of Agriculture National Institute of Food and Agriculture Grant 2018-67014-27542 (to K.M.K.); Utah Agricultural Experiment Station, Utah State University and approved as journal paper number 9240 (Project 1297 to K.M.K.); the National Sciences and Engineering Research Council (Discovery to A.Z.); the York Research Chair in Genomics (A.Z.); European Commission Horizon 2020 Research Infrastructure Program INFRAVEC2-731060 (to P.I.); Swiss National Science Foundation Grant PP00P3_170664 (to R.M.W.); the Smithsonian Institution Competitive Grants Program for Biogenomics (W.T.W., K.M.K., B.M.J.); the Smithsonian Institution Global Genome Initiative (W.T.W., C.K.); a Smithsonian Tropical Research Institute short-term fellowship (to C.K.); a Cornell University fellowship (to C.K.); and general research funds from the Smithsonian Tropical Research Institute (W.T.W.).

1. H. F. Nijhout, Development and evolution of adaptive polyphenisms. *Evol. Dev.* **5**, 9–18 (2003).
2. M. J. West-Eberhard, *Developmental Plasticity and Evolution*, (Oxford University Press, 2003).
3. T. Uller, A. P. Moczek, R. A. Watson, P. M. Brakefield, K. N. Laland, Developmental bias and evolution: A regulatory network perspective. *Genetics* **209**, 949–966 (2018).
4. N. A. Levis, D. W. Pfennig, Evaluating "plasticity-first" evolution in nature: Key criteria and empirical approaches. *Trends Ecol. Evol.* **31**, 563–574 (2016).
5. R. Lande, Adaptation to an extraordinary environment by evolution of phenotypic plasticity and genetic assimilation. *J. Evol. Biol.* **22**, 1435–1446 (2009).
6. J. M. Baldwin, A new factor in evolution. *Am. Nat.* **30**, 441–451 (1896).
7. A. P. Moczek *et al*., The role of developmental plasticity in evolutionary innovation. *Proc. Biol. Sci.* **278**, 2705–2713 (2011).
8. K. N. Laland *et al*., The extended evolutionary synthesis: Its structure, assumptions and predictions. *Proc. Biol. Sci.* **282**, 20151019 (2015).
9. T. D. Price, A. Qvarnström, D. E. Irwin, The role of phenotypic plasticity in driving genetic evolution. *Proc. Biol. Sci.* **270**, 1433–1440 (2003).
10. J. Gerhart, M. Kirschner, The theory of facilitated variation. *Proc. Natl. Acad. Sci. U.S.A.* **104** (suppl. 1), 8582–8589 (2007).
11. T. Schwander, O. Leimar, Genes as leaders and followers in evolution. *Trends Ecol. Evol.* **26**, 143–151 (2011).
12. S. F. Gilbert, T. C. G. Bosch, C. Ledón-Rettig, Eco-Evo-Devo: Developmental symbiosis and developmental plasticity as evolutionary agents. *Nat. Rev. Genet.* **16**, 611–622 (2015).
13. N. Aubin-Horth, S. C. P. Renn, Genomic reaction norms: Using integrative biology to understand molecular mechanisms of phenotypic plasticity. *Mol. Ecol.* **18**, 3763–3780 (2009).
14. E. K. Fischer, C. K. Ghalambor, K. L. Hoke, Can a network approach resolve how adaptive vs nonadaptive plasticity impacts evolutionary trajectories? *Integr. Comp. Biol.* **56**, 877–888 (2016).
15. N. A. Levis, D. W. Pfennig, Plasticity-led evolution: A survey of developmental mechanisms and empirical tests. *Evol. Dev.* **22**, 71–87 (2019).
16. W. T. Wcislo, Behavioral environments and evolutionary change. *Annu. Rev. Ecol. Syst.* **20**, 137–169 (1989).
17. S. C. P. P. Renn, M. E. Schumer, Genetic accommodation and behavioural evolution: Insights from genomic studies. *Anim. Behav.* **85**, 1012–1022 (2013).
18. R. A. Duckworth, The role of behavior in evolution: A search for mechanism. *Evol. Ecol.* **23**, 513–531 (2009).
19. C. D. Schlichting, M. A. Wund, C. D. Schlichting, M. A. Wund, Phenotypic plasticity and epigenetic marking: An assessment of evidence for genetic accommodation. *Evolution* **68**, 656–672 (2014).
20. M. J. West-Eberhard, Phenotypic plasticity and the origins of diversity. *Annu. Rev. Ecol. Syst.* **20**, 249–278 (1989).
21. B. M. Jones, G. E. Robinson, Genetic accommodation and the role of ancestral plasticity in the evolution of insect eusociality. *J. Exp. Biol.* **221**, jeb153163 (2018).
22. S. B. Carroll, Evolution at two levels: On genes and form. *PLoS Biol.* **3**, e245 (2005).
23. M. C. King, A. C. Wilson, Evolution at two levels in humans and chimpanzees. *Science* **188**, 107–116 (1975).
24. G. E. Robinson, Y. Ben-Shahar, Social behavior and comparative genomics: New genes or new gene regulation? *Genes Brain Behav.* **1**, 197–203 (2002).
25. A. L. Toth, S. M. Rehan, Molecular evolution of insect sociality: An eco-evo-devo perspective. *Annu. Rev. Entomol.* **62**, 419–442 (2017).
26. W. A. Shell, S. M. Rehan, Behavioral and genetic mechanisms of social evolution: Insights from incipiently and facultatively social bees. *Apidologie* **49**, 13–30 (2018).
27. C. Morandin *et al*., Comparative transcriptomics reveals the conserved building blocks involved in parallel evolution of diverse phenotypic traits in ants. *Genome Biol.* **17**, 43 (2016).
28. S. M. Rehan *et al*., Conserved genes underlie phenotypic plasticity in an incipiently social bee. *Genome Biol. Evol.* **10**, 2749–2758 (2018).
29. D. F. Simola *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).
30. K. M. Kapheim *et al*., Social evolution. Genomic signatures of evolutionary transitions from solitary to group living. *Science* **348**, 1139–1143 (2015).
31. S. Patalano *et al*., Molecular signatures of plastic phenotypes in two eusocial insect species with simple societies. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 13970–13975 (2015).
32. S. V. Arsenault, B. G. Hunt, S. M. Rehan, The effect of maternal care on gene expression and DNA methylation in a subsocial bee. *Nat. Commun.* **9**, 3468 (2018).
33. W. Mao, M. A. Schuler, M. R. Berenbaum, A dietary phytochemical alters caste-associated gene expression in honey bees. *Sci. Adv.* **1**, e1500795 (2015).
34. J. R. Withee, S. M. Rehan, Social aggression, experience, and brain gene expression in a subsocial bee. *Integr. Comp. Biol.* **57**, 640–648 (2017).
35. K. Zhu *et al*., Plant microRNAs in larval food regulate honeybee caste development. *PLoS Genet.* **13**, e1006946 (2017).
36. A. Roth *et al*., A genetic switch for worker nutrition-mediated traits in honeybees. *PLoS Biol.* **17**, e3000171 (2019).
37. J. H. Hunt, G. V. Amdam, Bivoltinism as an antecedent to eusociality in the paper wasp genus *Polistes*. *Science* **308**, 264–267 (2005).
38. T. A. Linksvayer, M. J. Wade, The evolutionary origin and elaboration of sociality in the aculeate Hymenoptera: Maternal effects, sib-social effects, and heterochrony. *Q. Rev. Biol.* **80**, 317–336 (2005).
39. G. V. Amdam, K. Norberg, M. K. Fondrk, R. E. Page Jr., Reproductive ground plan may mediate colony-level selection effects on individual foraging behavior in honey bees. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 11350–11355 (2004).
40. M. J. West Eberhard, "Flexible strategy and social evolution" in *Animal Societies: Theories and Facts*, Y. Ito, J. L. Brown, J. Kikkawa, Eds. (Japan Scientific Societies Press, 1987), pp. 35–51.
41. M. J. West Eberhard, "Wasp societies as microcosms for the study of development and evolution" in *Natural History and Evolution of Paper Wasps*, S. Turillazzi, M. J. West Eberhard, Eds. (Oxford University Press, 1996), pp. 290–317.
42. M. R. Warner, L. Qiu, M. J. Holmes, A. S. Mikheyev, T. A. Linksvayer, Convergent eusocial evolution is based on a shared reproductive groundplan plus lineage-specific plastic genes. *Nat. Commun.* **10**, 2651 (2019).
43. A. Klein *et al*., Evolution of social insect polyphenism facilitated by the sex differentiation cascade. *PLoS Genet.* **12**, e1005952 (2016).
44. D. V. Brito *et al*., Molecular characterization of the gene *feminizer* in the stingless bee *Melipona interrupta* (Hymenoptera: Apidae) reveals association to sex and caste development. *Insect Biochem. Mol. Biol.* **66**, 24–30 (2015).
45. B. R. Johnson, W. C. Jasper, Complex patterns of differential expression in candidate master regulatory genes for social behavior in honey bees. *Behav. Ecol. Sociobiol.* **70**, 1033–1043 (2016).
46. G. J. Thompson, P. L. Hurd, B. J. Crespi, Genes underlying altruism. *Biol. Lett.* **9**, 20130395 (2013).
47. B. A. Harpur *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).
48. W. C. Jasper *et al*., Large-scale coding sequence change underlies the evolution of postdevelopmental novelty in honey bees. *Mol. Biol. Evol.* **32**, 334–346 (2015).
49. B. Feldmeyer, D. Elsner, S. Foitzik, Gene expression patterns associated with caste and reproductive status in ants: Worker-specific genes are more derived than queen-specific ones. *Mol. Ecol.* **23**, 151–161 (2014).
50. P. G. Ferreira *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).
51. F. Manfredini *et al*., Sociogenomics of cooperation and conflict during colony founding in the fire ant *Solenopsis invicta*. *PLoS Genet.* **9**, e1003633 (2013).
52. C. Morandin, A. S. Mikheyev, J. S. Pedersen, H. Helanterä, Evolutionary constraints shape caste-specific gene expression across 15 ant species. *Evolution* **71**, 1273–1284 (2017).
53. B. M. Jones, W. T. Wcislo, G. E. Robinson, Developmental transcriptome for a facultatively eusocial bee, *Megalopta genalis*. *G3 (Bethesda)* **5**, 2127–2135 (2015).
54. B. M. Jones, C. J. Kingwell, W. T. Wcislo, G. E. Robinson, Caste-biased gene expression in a facultatively eusocial bee suggests a role for genetic accommodation in the evolution of eusociality. *Proc. Biol. Sci.* **284**, 20162228 (2017).
55. L. Schrader, H. Helanterä, J. Oettler, Accelerated evolution of developmentally biased genes in the tetraphenic ant *Cardiocondyla obscurior*. *Mol. Biol. Evol.* **34**, 535–544 (2017).
56. D. W. Pfennig, I. M. Ehrenreich, Towards a gene regulatory network perspective on phenotypic plasticity, genetic accommodation and genetic assimilation. *Mol. Ecol.* **23**, 4438–4440 (2014).
57. W. Saelens, R. Cannoodt, Y. Saeys, A comprehensive evaluation of module detection methods for gene expression data. *Nat. Commun.* **9**, 1090 (2018).
58. S. H. Woodard *et al*., Genes involved in convergent evolution of eusociality in bees. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 7472–7477 (2011).
59. M. R. Warner, A. S. Mikheyev, T. A. Linksvayer, Genomic signature of kin selection in an ant with obligately sterile workers. *Mol. Biol. Evol.* **34**, 1780–1787 (2017).
60. B. A. Harpur *et al*., Queens and workers contribute differently to adaptive evolution in bumble bees and honey bees. *Genome Biol. Evol.* **9**, 2395–2402 (2017).
61. E. Privman *et al*., Positive selection on sociobiological traits in invasive fire ants. *Mol. Ecol.* **27**, 3116–3130 (2018).
62. B. G. Hunt *et al*., Sociality is linked to rates of protein evolution in a highly social insect. *Mol. Biol. Evol.* **27**, 497–500 (2010).
63. B. G. Hunt *et al*., Relaxed selection is a precursor to the evolution of phenotypic plasticity. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 15936–15941 (2011).

Downloaded by guest on April 7, 2021

64. D. M. Rand, L. M. Kann, Excess amino acid polymorphism in mitochondrial DNA: Contrasts among genes from *Drosophila*, mice, and humans. *Mol. Biol. Evol.* **13**, 735–748 (1996).
65. N. Stoletzki, A. Eyre-Walker, Estimation of the neutrality index. *Mol. Biol. Evol.* **28**, 63–70 (2011).
66. K. A. Dogantzis *et al*., Insects with similar social complexity show convergent patterns of adaptive molecular evolution. *Sci. Rep.* **8**, 10388 (2018).
67. D. Thompson, A. Regev, S. Roy, Comparative analysis of gene regulatory networks: From network reconstruction to evolution. *Annu. Rev. Cell Dev. Biol.* **31**, 399–428 (2015).
68. A. M. Hulse, J. J. Cai, Genetic variants contribute to gene expression variability in humans. *Genetics* **193**, 95–108 (2013).
69. M. Ackermann, W. Sikora-Wohlfeld, A. Beyer, Impact of natural genetic variation on gene expression dynamics. *PLoS Genet.* **9**, e1003514 (2013).
70. E. Cannavò *et al*., Genetic variants regulating expression levels and isoform diversity during embryogenesis. *Nature* **541**, 402–406 (2017).
71. D. W. Pfennig *et al*., Phenotypic plasticity's impacts on diversification and speciation. *Trends Ecol. Evol.* **25**, 459–467 (2010).
72. M. A. Wund, Assessing the impacts of phenotypic plasticity on evolution. *Integr. Comp. Biol.* **52**, 5–15 (2012).
73. W. T. Wcislo, J. H. Fewell, "Sociality in bees" in *Comparative Social Evolution*, D. R. Rubenstein, P. Abbot, Eds. (Cambridge University Press, 2017), pp. 50–83.
74. C. D. Michener, From solitary to eusocial: Need there be a series of intervening species? *Fortschr. Zool.* **31**, 293–306 (1985).
75. M. J. West-Eberhard, Phenotypic accommodation: Adaptive innovation due to developmental plasticity. *J. Exp. Zoolog. B Mol. Dev. Evol.* **304**, 610–618 (2005).
76. K. M. Kapheim, S. P. Bernal, A. R. Smith, P. Nonacs, W. T. Wcislo, Support for maternal manipulation of developmental nutrition in a facultatively eusocial bee, Megalopta genalis (Halictidae). *Behav. Ecol. Sociobiol. (Print)* **65**, 1179–1190 (2011).
77. K. M. Kapheim *et al*., Physiological variation as a mechanism for developmental caste-biasing in a facultatively eusocial sweat bee. *Proc. Biol. Sci.* **279**, 1437–1446 (2012).
78. A. R. Smith, K. M. Kapheim, S. O'Donnell, W. T. Wcislo, Social competition but not subfertility leads to a division of labour in the facultatively social sweat bee *Megalopta genalis* (Hymenoptera: Halictidae). *Anim. Behav.* **78**, 1043–1050 (2009).
79. K. M. Kapheim, T. Y. Chan, A. R. Smith, W. T. Wcislo, P. Nonacs, Ontogeny of division of labor in a facultatively eusocial sweat bee Megalopta genalis. *Insectes Soc.* **63**, 185–191 (2016).
80. A. R. Smith, M. Simons, V. Bazarko, J. Harach, M. A. Seid, Queen–worker aggression in the facultatively eusocial bee *Megalopta genalis*. *Insectes Soc.* **66**, 479–490 (2019).
81. A. R. Smith, M. Simons, V. Bazarko, M. Seid, The influence of sociality, caste, and size on behavior in a facultatively eusocial bee. *Insectes Soc.* **66**, 153–163 (2019).
82. L. Arneson, W. T. Wcislo, Dominant-subordinate relationships in a facultatively social, nocturnal bee, *Megalopta genalis* (Hymenoptera: Halictidae). *J. Kans. Entomol. Soc.* **76**, 183–193 (2003).
83. W. T. Wcislo, V. H. Gonzalez, Social and ecological contexts of trophallaxis in facultatively social sweat bees, *Megalopta genalis* and *M. ecuadoria* (Hymenoptera, Halictidae). *Insectes Soc.* **53**, 220–225 (2006).
84. K. N. Laland, K. Sterelny, J. Odling-Smee, W. Hoppitt, T. Uller, Cause and effect in biology revisited: Is Mayr's proximate-ultimate dichotomy still useful? *Science* **334**, 1512–1516 (2011).
85. G. E. Robinson, C. M. Grozinger, C. W. Whitfield, Sociogenomics: Social life in molecular terms. *Nat. Rev. Genet.* **6**, 257–270 (2005).
86. K. M. Kapheim, Synthesis of Tinbergen's four questions and the future of sociogenomics. *Behav. Ecol. Sociobiol.* **73**, 186 (2019).
87. K. M. Kapheim, Genomic sources of phenotypic novelty in the evolution of eusociality in insects. *Curr. Opin. Insect Sci.* **13**, 24–32 (2016).
88. R. Luo *et al*., SOAPdenovo2: An empirically improved memory-efficient short-read de novo assembler. *Gigascience* **1**, 18 (2012).
89. C. G. Elsik *et al*., Creating a honey bee consensus gene set. *Genome Biol.* **8**, R13 (2007).
90. E. M. Zdobnov *et al*., OrthoDB v9.1: Cataloging evolutionary and functional annotations for animal, fungal, plant, archaeal, bacterial and viral orthologs. *Nucleic Acids Res.* **45**, D744–D749 (2017).
91. W. T. Wcislo *et al*., The evolution of nocturnal behaviour in sweat bees, *Megalopta genalis* and *M. ecuadoria* (Hymenoptera : Halictidae): An escape from competitors and enemies? *Biol. J. Linn. Soc. Lond.* **83**, 377–387 (2004).
92. Y. Benjamini, Y. Hochberg, Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc. B* **57**, 289–300 (1995).
93. K. E. Eilertson, J. G. Booth, C. D. Bustamante, SnIPRE: Selection inference using a Poisson random effects model. *PLOS Comput. Biol.* **8**, e1002806 (2012).
94. H. Li, A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* **27**, 2987–2993 (2011).
95. S. Y. Kim *et al*., Design of association studies with pooled or un-pooled next-generation sequencing data. *Genet. Epidemiol.* **34**, 479–491 (2010).
96. M. D. Shapiro *et al*., Genomic diversity and evolution of the head crest in the rock pigeon. *Science* **339**, 1063–1067 (2013).
97. A. Dobin *et al*., STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
98. Y. Liao, G. K. Smyth, W. Shi, The R package *Rsubread* is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. *Nucleic Acids Res.* **47**, e47 (2019).
99. M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
100. R Core Team, *R: A Language and Environment for Statistical Computing*, (R Foundation for Statistical Computing, Vienna, Austria, 2016).
101. V. Calcagno, C. de Mazancourt, glmulti: An R package for easy automated model selection with (generalized) linear models. *J. Stat. Softw.* **34**, 636–645 (2010).
102. A. F. Zuur, E. N. Ieno, N. Walker, A. A. Saveliev, G. M. Smith, *Mixed Effects Models and Extensions in Ecology with R*, (Springer-Verlag, 2009).
103. A. F. Zuur, E. N. Ieno, C. S. Elphick, A protocol for data exploration to avoid common statistical problems. *Methods Ecol. Evol.* **1**, 3–14 (2010).
104. F. E. Harrell Jr., Hmisc: Harrell Miscellaneous (2018). Version 4. https://cran.r-project.org/package=Hmisc. Accessed 26 May 2020.
105. A. Biton, MineICA: Analysis of an ICA decomposition obtained on genomics data (2012). Version 1.4.0. https://bioc.ism.ac.jp/packages/2.14/bioc/html/MineICA.html. Accessed 26 May 2020.
106. Y. Zhang, Likelihood-based and Bayesian methods for Tweedie compound Poisson linear mixed models. *Stat. Comput.* **23**, 743–757 (2013).
107. C. E. Grant, T. L. Bailey, W. S. Noble, FIMO: Scanning for occurrences of a given motif. *Bioinformatics* **27**, 1017–1018 (2011).
108. L. J. Zhu *et al*., FlyFactorSurvey: A database of *Drosophila* transcription factor binding specificities determined using the bacterial one-hybrid system. *Nucleic Acids Res.* **39**, D111–D117 (2011).
109. K. M. Kapheim, A. R. Smith, P. Nonacs, W. T. Wcislo, R. K. Wayne, Foundress polyphenism and the origins of eusociality in a facultatively eusocial sweat bee, *Megalopta genalis* (Halictidae). *Behav. Ecol. Sociobiol.* **67**, 331–340 (2013).
110. J. L. Rainford, M. Hofreiter, D. B. Nicholson, P. J. Mayhew, Phylogenetic distribution of extant richness suggests metamorphosis is a key innovation driving diversification in insects. *PLoS One* **9**, e109085 (2014).
111. T. Gempe, M. Beye, Function and evolution of sex determination mechanisms, genes and pathways in insects. *BioEssays* **33**, 52–60 (2011).
112. S. G. Brady, S. Sipes, A. Pearson, B. N. Danforth, Recent and simultaneous origins of eusociality in halictid bees. *Proc. Biol. Sci.* **273**, 1643–1649 (2006).

Downloaded by guest on April 7, 2021