Worker Reproduction and Caste Polymorphism Impact Genome Evolution and Social Genes Across the Ants

Megan Barkdull [1,*] and Corrie S. Moreau [1,2]

[1]Department of Ecology & Evolutionary Biology, Cornell University
[2]Department of Entomology, Cornell University

*Corresponding author: E-mail: mb2337@cornell.edu.

Accepted: 23 May 2023

Abstract

Eusocial insects are characterized by several traits, including reproductive division of labor and caste polymorphisms, which likely modulate genome evolution. Concomitantly, evolution may act on specific genes and pathways underlying these novel, sociality-associated phenotypes. Reproductive division of labor should increase the magnitude of genetic drift and reduce the efficacy of selection by reducing effective population size. Caste polymorphism has been associated with relaxed selection and may facilitate directional selection on caste-specific genes. Here, we use comparative analyses of 22 ant genomes to test how reproductive division of labor and worker polymorphism influence positive selection and selection intensity across the genome. Our results demonstrate that worker reproductive capacity is associated with a reduction in the degree of relaxed selection but is not associated with any significant change to positive selection. We find decreases in positive selection in species with polymorphic workers, but no increase in the degree of relaxed selection. Finally, we explore evolutionary patterns in specific candidate genes associated with our focal traits in eusocial insects. Two oocyte patterning genes previously implicated in worker sterility evolve under intensified selection in species with reproductive workers. Behavioral caste genes generally experience relaxed selection associated with worker polymorphism, whereas *vestigial* and *spalt*, both associated with soldier development in *Pheidole* ants, experience intensified selection in worker polymorphic species. These findings expand our understanding of the genetic mechanisms underlying elaborations of sociality. The impacts of reproductive division of labor and caste polymorphisms on specific genes illuminate those genes' roles in generating complex eusocial phenotypes.

Key words: comparative genomics, eusocial insect evolution, positive selection, selection intensity.

Significance

Organismal phenotypes mediate patterns of natural selection across the genome, but theoretical predictions about the nature of these effects do not always match empirical findings. In a large comparative genomics study of 22 ant species, we consider how two eusociality-associated traits, worker reproduction and worker polymorphism, impact levels of positive and relaxed selection in ant genomes. We also explore the evolution of specific genes linked to these traits, demonstrating that egg-related patterning genes linked to worker sterility, and behavioral and developmental genes linked to caste polymorphism, evolve in ways consistent with their roles in these traits. Our results add to our understanding of potential mechanisms underlying elaborations of eusociality in ants.

Introduction

Patterns of molecular evolution across taxa are shaped by both neutral and selective mechanisms. These mechanisms include the rate at which mutations occur, the strength of genetic drift, and the efficacy of natural selection in fixing advantageous mutations or purging deleterious ones

(Kimura 1968, 1979; Ohta 1992). Neutral processes in particular can modulate patterns of molecular evolution at a broad, genome-wide scale. For instance, a reduction in population size will increase the strength of genetic drift acting on all genes (Mank et al. 2010; Star and Spencer 2013).

Organismal phenotypes shape evolutionary processes' effects on the genome. For instance, differences in generation time directly impact evolutionary rates by changing the speed at which mutations accumulate (Bromham 2002; Thomas et al. 2010). The number of germline cell divisions that occur prior to reproduction increases the number of de novo mutations passed on to offspring, increasing rates of molecular evolution (Bartosch-Härlid et al. 2003; Jónsson et al. 2017; Wang et al. 2020; Rubin 2022). Different sex-determination systems increase or decrease the strength of genetic drift in sex chromosomes, shifting patterns of evolution in these regions of the genome (Mank et al. 2010).

In social insects, neutral and selective evolutionary mechanisms can be mediated by different sociality-associated phenotypes, leading to variation in patterns of molecular evolution across social taxa. Ants (family: Formicidae) are a particularly useful group in which to examine the effect of these social traits on molecular evolution. In part, this is because eusociality arose only once in the common ancestor of modern ants (Hölldobler and Wilson 1990; Moreau et al. 2006; Gadau et al. 2012). Whereas this is a drawback for comparative work aimed at understanding the origins of eusociality, it is a benefit for comparative studies into elaborations of and variations in eusociality and sociality-associated traits. We can be confident that we are comparing across species with the same evolutionary history leading to the emergence of eusociality.

Shared patterns of molecular evolution have been identified in some ant genomes. Across seven species, ant genomes are characterized by reduced purifying selection compared with the *Drosophila* genome (Simola et al. 2013). In spite of this pattern, there is fairly conserved genome synteny across ant species (Simola et al. 2013). In the same seven ant species, comparative work has identified patterns of positive selection throughout the genome, with 24 functional categories enriched for positive selection in ants (Roux et al. 2014). Compared with bees and flies, ants seem to have increased positive selection on mitochondrial genes, perhaps related to their dramatically increased lifespans (Roux et al. 2014). Recent work also identifies episodes of positive selection during the evolution of the ants, with a burst of positive selection during the evolution of formicoid ants, possibly related to the emergence of extreme division of labor in this clade (Romiguier et al. 2022). In general, ant genomes exhibit specific patterns of molecular evolution, distinct from other hymenopterans and from other insects, and likely related to their eusocial phenotypes.

The single origin of eusociality in ants also facilitates explorations of the effect that different sociality-related traits have on genome evolution (Gadau et al. 2012). One of the defining features of eusociality is reproductive division of labor, with some individuals capable of reproduction and others that are mostly or entirely sterile (Michener 1969; Wilson 1971; Wilson and Hölldobler 2005). Reproductive division of labor is predicted to impact genome evolution. When only one or a few individuals reproduce, effective population size will be reduced, increasing the strength of genetic drift and decreasing the efficacy of natural selection (Kimura 1979). In contrast, if workers are regularly capable of reproduction, effective population size will be relatively increased. Worker attempts to reproduce can also lead to more conflict in the colony, potentially even in clonal species (Liebig et al. 1999; Hartmann et al. 2003). Moreover, because sterile worker phenotypes contribute to fitness only insofar as they impact the fitness of reproductive individuals in the colony, sterile worker genomes should be acted upon by relatively weak indirect kin selection rather than direct selection (Hamilton 1964a, 1964b; Linksvayer and Wade 2009, 2016). The shift toward kin selection should result in a general relaxation of both positive and purifying selection acting on worker-biased genes, and potentially across the genome, in species with sterile workers (Linksvayer and Wade 2009). In contrast, species with reproductive workers should experience a comparatively greater degree of direct selection on worker genes.

Eusociality results in the existence of multiple castes: reproductive queens and non- or less-reproductive workers, which in some groups can be further subdivided into different subcastes. Despite their single origin of eusociality, ants are characterized by wide variation in the degree of worker caste polymorphism across the family. In some species, ant workers are further subdivided into minors and majors (which often behaviorally function as soldiers), or other specialized castes, like repletes. Around 84% of ant genera have monomorphic workers; the rest exhibit various forms of worker polymorphism (Hölldobler and Wilson 1990; Wills et al. 2018). Importantly, these different subcaste systems have multiple independent origins, which facilitates comparative genomic investigations into elaborations of the worker caste (Blanchard and Moreau 2017).

Variations in caste system likely influence molecular evolution (Kapheim 2019). Eusocial insect castes, including ant castes, are typically produced via alternative development of a single genome (developmental plasticity), rather than being genetically determined (with rare exceptions, for instance in *Pogonomyrmex rugosus* and *Pogonomyrmex barbatus*) (Julian et al. 2002; Helms Cahan and Keller 2003; West-Eberhard 2005). Plasticity is associated with increased rates of sequence evolution (Helanterä and Uller 2014; Schrader et al. 2017). There are several potential, nonmutually exclusive reasons for this association (Helanterä

Focal Trait	Predicted effect: Positive selection	Predicted effect: Relaxed selection	Recovered effect: Positive selection	Recovered effect: Relaxed selection
Worker reproduction	⬆	⬇	⬇ N.S.	⬇
Worker polymorphism	⬆	⬆	⬇ N.S.	–

Fig. 1.—Predicted and recovered effects of worker reproduction and worker polymorphism on genome-wide evolutionary patterns. Tan color and upward arrows indicate an increase in the number of genes that experience a selective regime in species with the focal trait relative to those without; blue color and downward arrows indicate a decrease in the number of genes experiencing that selective regime, with a lighter color and N.S. notation to note nonsignificant results based on a bootstrapping test; and gray color and dash indicate no significant change.

and Uller 2014; Schrader et al. 2017). Relaxed selection on a gene may be responsible for initially allowing its expression to become plastic (Leichty et al. 2012; Helanterä and Uller 2014). Conversely, plastic expression can lead to relaxed selection, because the gene is being expressed in only a subset of individuals in the population and is thus partially shielded from selection (Kapheim 2019). A third explanation is that genes expressed plastically between castes are released from caste-antagonistic pleiotropic effects. This frees sites that were ancestrally under purifying selection and allows ancestral alleles or new mutations to be acted upon by directional selection (Schrader et al. 2017). As a result, taxa with more-complex caste systems may be expected to have increased rates of both positive and relaxed selection. However, it is also possible that other phenotypes could obscure this relationship; for instance, species with polymorphic workers also tend to have larger colonies, which could reduce effective population size by reducing the queen-to-worker ratio in these groups (Ferguson-Gow et al. 2014).

Whereas organismal phenotypes modulate neutral and adaptive evolutionary processes, the converse is also true: evolutionary processes will impact particular genes involved in the production of organismal phenotypes. Both reproductive division of labor and worker polymorphism are complex phenotypes, the product of many genes interacting in processes which are not yet fully understood. Nonetheless, a body of research has identified candidate genes implicated in these phenotypes across eusocial insects. Reproductive division of labor has been linked to juvenile hormone signaling, the activity of ecdysteroid hormones, vitellogenesis, and ovarian and oocyte development, among other processes (Ronai et al. 2016; Ghaninia et al. 2017; Jongepier et al. 2018; Araki et al. 2020; Awde et al. 2020). The production of worker subcastes likewise appears to involve diverse genetic mechanisms, including neuropeptide activity, genes implicated in behavioral state, and wing development genes (Gospocic et al. 2017; Rajakumar et al. 2018; Fetter-Pruneda et al. 2021). As these phenotypes have evolved across the ant phylogeny, both selective and neutral

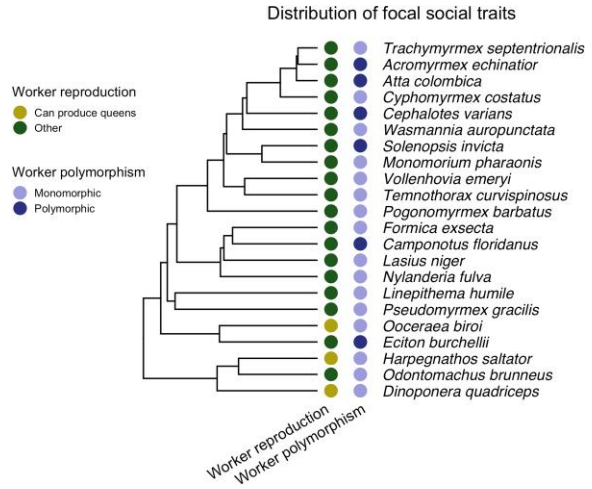

Fig. 2.—Phylogeny of sampled ant species, illustrating character states for focal social traits. Dots at the phylogeny tips are colored to indicate the trait state for that species for worker reproduction and worker polymorphism. Phylogeny adapted from Blanchard and Moreau 2017; sources for character states can be found in supplementary table S1, Supplementary Material online.

evolutionary processes may have shaped the genes responsible for worker reproduction and worker caste polymorphisms.

To understand the impact of eusocial traits on patterns of molecular evolution, we test the following hypotheses, summarized in figure 1. We expect that worker reproduction should be associated with increased rates of positive selection and reduced rates of relaxed selection, as effective population size increases and worker genes become directly exposed to natural selection in these species. Worker polymorphism should be associated with elevated rates of positive selection, along with a reduction in selection intensity, as a result of the relationship between worker polymorphism, gene expression plasticity, and sequence evolution.

Specifically, we test these hypotheses using whole-genome data from 22 ant species that vary in worker reproductivity and worker polymorphism, as summarized in figure 2. We

evaluate the relationship between the two sociality-associated traits and patterns of positive, relaxed, and intensified selection genome wide. We also identify specific candidate genes linked to worker reproduction and worker polymorphism and investigate the selective regimes acting upon these genes in association with each trait.

Results

Orthogroup Identification with OrthoFinder

After filtering coding sequences to retain only the longest isoform for each gene, we inferred orthogroups with OrthoFinder (Emms and Kelly 2015, 2019). This resulted in a set of 32,792 orthogroups across 22 species, including 2,813 single-copy orthogroups and 21,098 non-singleton orthogroups. Subsets of all orthogroups were used for downstream analyses, because BUSTED-PH and RELAX require that an orthogroup contain genes from both species possessing and species lacking the focal trait being tested.

Positive Selection across the ant Phylogeny with BUSTED[S]

We ran BUSTED[S] on all nonsingleton orthogroups in our dataset, on all branches. The BUSTED[S] method assesses whether at least one site in an alignment is under positive selection in at least one branch tested, whereas allowing synonymous substitution rates to vary (Murrell et al. 2015; Wisotsky et al. 2020). We found evidence for positive selection at some point during ant evolution for 53.90% of orthogroups (11,371 of 21,098 orthogroups). These orthogroups are enriched for 37 GO terms, listed in table 1. Of single-copy orthogroups, 83.89% show evidence for positive selection at some point in ant evolution (2,359 of 2,813 single-copy orthogroups).

Positive Selection Associated with Focal Traits with BUSTED-PH

We used BUSTED-PH to assess evidence for each orthogroup having experienced positive selection associated with the presence and absence of each of our focal sociality-associated traits. These analyses were repeated using only single-copy orthogroups, to control for effects of phenomena like gene family expansions; the single-copy orthogroups results reflect the same patterns as observed genome wide (see supplementary table S3, Supplementary Material online). Genome-wide results are summarized in figure 3.

To assess the possible impact of unequal sampling of foreground and background species, we implemented a bootstrapping approach to determine whether our observed results are more extreme than those obtained for sets of foreground species chosen at random (see Methods). Those results are summarized in figure 3 and supplementary table S5, Supplementary Material online.

Worker Reproduction

We tested 11,021 orthogroups for an association between worker reproduction and positive selection. BUSTED-PH identifies 169 orthogroups under positive selection exclusively in species with workers that can produce female reproductive offspring, compared with 1,312 orthogroups under positive selection only in species without such reproductive workers. There are significantly fewer genes under positive selection in species with reproductive workers (chi-squared goodness-of-fit test, $P = 7.49e^{-194}$). However, this result falls in the middle of the distribution of our bootstrapping iterations and so could result from the small number of foreground species in this analysis (figure 3 and supplementary table S5, Supplementary Material online).

Genes under positive selection in species with reproductive workers are enriched for two GO terms. Genes under positive selection in species without reproductive workers are enriched for five GO terms (table 2).

Worker Polymorphism

We examined the effect of worker polymorphism on positive selection in 12,336 orthogroups. We find 349 genes under positive selection exclusively in species with polymorphic workers, compared with 1,324 genes under positive selection solely in monomorphic species. There are significantly fewer genes under positive selection in association with worker polymorphism (chi-squared goodness-of-fit test, $P = 1.37e^{-125}$). Again, this result falls in the center of the distribution of bootstrapping results and so may result from foreground species sampling (figure 3 and supplementary table S5, Supplementary Material online).

Genes under positive selection in species with polymorphic workers are enriched for two GO terms, whereas genes under positive selection in species with monomorphic workers are enriched for five GO terms (table 3).

Shifts in Selection Intensity Associated with Focal Traits with RELAX

To understand how each of our focal traits might be associated with changes in selection intensity, we ran RELAX on the sets of orthogroups also tested with BUSTED-PH (worker reproduction, $N = 10{,}973$ orthogroups; worker polymorphism, $N = 12{,}643$ orthogroups). RELAX assesses selection intensity as a shift in the degree of purifying and/or positive selection in foreground branches relative to background branches (Wertheim et al. 2015). Here, we test whether the number of genes under intensified selection in foreground species is equal to, or different from,

Table 1
Gene Ontology Terms Significantly Enriched for Positive Selection across the ant Phylogeny

Term	GO ID	*P* value	Total orthogroups annotated	Significant orthogroups
Protein binding	GO:0005515	0.00000000000000006	1,836	1,357
Protein phosphorylation	GO:0006468	0.0000000032	277	220
G protein-coupled receptor signaling pathway	GO:0007186	0.00000017	172	140
Signal transduction	GO:0007165	0.000004	403	320
Intracellular signal transduction	GO:0035556	0.0000082	118	76
Regulation of transcription, DNA-templated	GO:0006355	0.00001	309	231
Integral component of membrane	GO:0016021	0.000013	772	552
Proteolysis	GO:0006508	0.000015	411	316
Nucleus	GO:0005634	0.000022	319	241
G protein-coupled receptor activity	GO:0004930	0.00003	158	127
Protein kinase activity	GO:0004672	0.000032	279	221
Transport	GO:0006810	0.000063	744	523
Metalloendopeptidase activity	GO:0004222	0.00019	76	53
Protein serine/threonine kinase activity	GO:0004674	0.00031	36	33
GTPase activator activity	GO:0005096	0.00081	23	22
Transmembrane transporter activity	GO:0022857	0.0009	527	415
Multicellular organism development	GO:0007275	0.0014	25	17
Transmembrane transport	GO:0055085	0.0017	432	357
Phosphatase activity	GO:0016791	0.00172	72	59
Regulation of cellular process	GO:0050794	0.0018	785	609
Extracellular region	GO:0005576	0.0019	136	104
Extracellular matrix	GO:0031012	0.0028	24	16
Guanyl-nucleotide exchange factor activity	GO:0005085	0.00417	49	41
Cell adhesion	GO:0007155	0.0042	42	35
Transcription elongation from RNA polymerase II promoter	GO:0006368	0.0042	8	3
Serine-type endopeptidase activity	GO:0004252	0.00489	145	110
Inorganic anion transmembrane transporter activity	GO:0015103	0.00531	18	17
Cellular component biogenesis	GO:0044085	0.0062	144	106
Ion transport	GO:0006811	0.0064	273	210
Regulation of small GTPase mediated signal transduction	GO:0051056	0.0068	11	11
3′,5′-Cyclic-nucleotide phosphodiesterase activity	GO:0004114	0.00705	9	3
Inorganic anion transport	GO:0015698	0.0072	16	15
Chitin binding	GO:0008061	0.00744	66	52
Gated channel activity	GO:0022836	0.00745	80	63
Phosphate-containing compound metabolic process	GO:0006796	0.0076	503	379
Organonitrogen compound metabolic process	GO:1901564	0.0092	1,245	858
Transferase complex, transferring phosphorus-containing groups	GO:0061695	0.01	25	22

the number of genes under relaxed selection in foreground species. We repeated our analyses using only single-copy orthogroups, to control for the effects of phenomena like gene family expansions (see supplementary table S3, Supplementary Material online). Genome-wide results are summarized in figure 4.

Worker Reproduction

In assessing the relationship between worker reproduction and levels of selection intensity, we identify 685 genes under relaxed selection in species with reproductive workers, compared with 1,927 genes under intensified selection. Thus, there are significantly fewer genes under relaxed selection in species with reproductive workers (chi-squared goodness-of-fit test, $P = 1.89e^{-130}$). The single-copy orthogroup analysis recovers the same pattern of increased selection intensity in species with reproductive workers (supplementary table S3, Supplementary Material online).

Genes under relaxed selection in species with reproductive workers are enriched for three GO terms, whereas those under intensified selection are enriched for 12 GO terms (table 2).

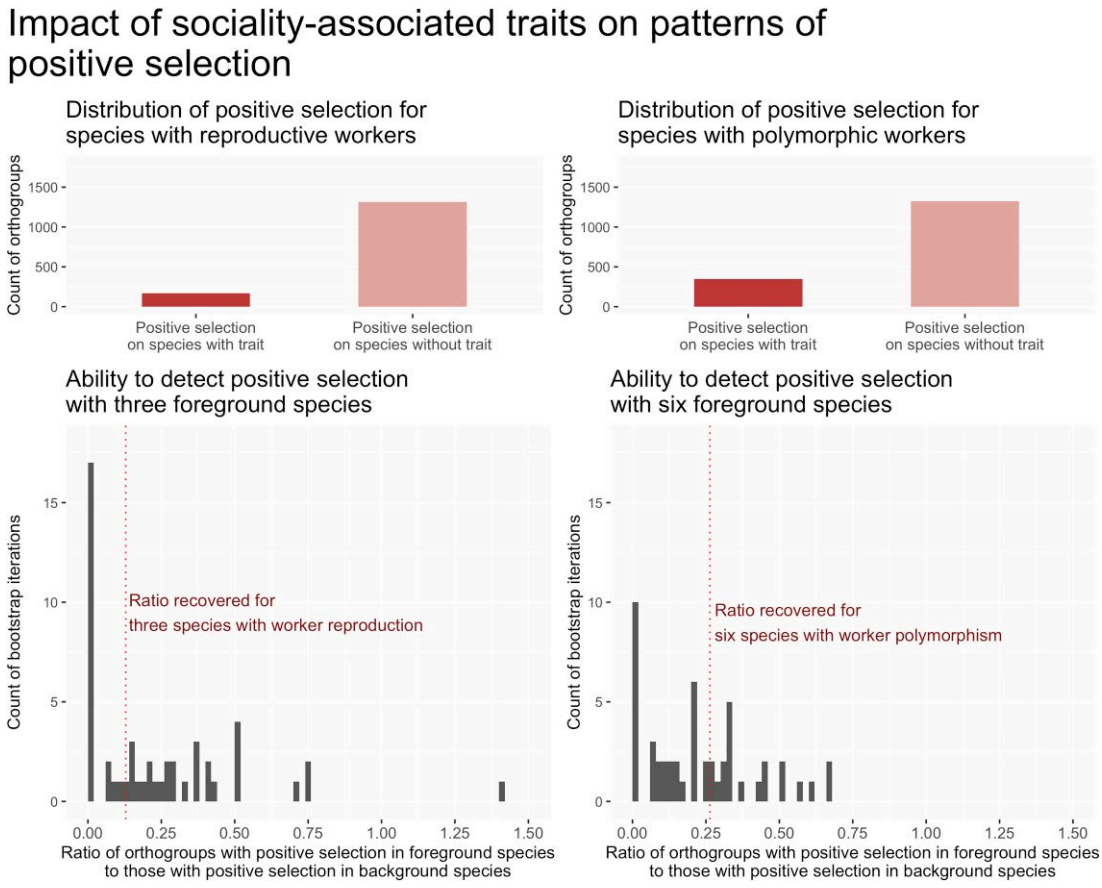

Fig. 3.—Impact of sociality-associated traits on patterns of positive selection across ant genomes. For each trait, the count of orthogroups experiencing positive selection in association with trait presence (dark-red bars) and the count of orthogroups experiencing positive selection in association with trait absence (light-red bars) are shown in the top panel. To assess the potential impact of species sampling on the ability to detect positive selection, we implemented a bootstrapping approach; for each trait, these results are shown in the bottom panel. For both worker reproduction and worker polymorphism, there are fewer genes experiencing positive selection associated with trait presence than with trait absence; however, based on the bootstrapping analysis, we are unable to rule out that this pattern is the result of species sampling, as our recovered ratios fall in the middle of the distribution of bootstrapping results.

Worker Polymorphism

When considering the effects of worker polymorphism on selection intensity, we find that there are very slightly fewer genes under relaxed selection in species with polymorphic workers, although this difference is not significant (chi-squared goodness-of-fit test, $P = 0.3741$). There are 1,557 genes under relaxed selection in species with polymorphic workers, versus 1,607 genes under intensified selection. However, when we assess single-copy orthogroups alone, we do find a shift toward intensified selection in species with polymorphic workers (chi-squared goodness-of-fit test, $P = 5.52e^{-08}$; see supplementary table S3, Supplementary Material online).

Genes under relaxed selection in species with polymorphic workers are enriched for 9 GO terms; genes under intensified selection are enriched for 12 GO terms (table 3).

Evolution of Genes Linked to Focal Phenotypes

Candidate Genes from the Literature

Of the 25 papers we examined per trait, ten papers identified specific candidate genes related to worker reproduction, and six papers identified genes related to worker polymorphism. From these papers, we identified 85 potential candidate genes, 62 linked to worker reproduction and 23 linked to worker polymorphism (supplementary table S6, Supplementary Material online). Of these genes, we were able to find corresponding ant orthogroups for 36 of the worker reproduction candidates and all 23 of the worker polymorphism candidates.

Evolution of Candidate Genes

To reflect different levels of stringency, we report candidate gene results with and without Benjamini–Hochberg correction (i.e., treating each candidate as an independent, a

Table 2
Gene Ontology Terms Enriched for Each Selective Regime Associated with Worker Reproduction

Term	GO ID	*P* value	Total orthogroups annotated	Significant orthogroups
Positive selection in species with the trait				
Transposition, DNA mediated	GO:0006313	0.0062	59	4
Transposase activity	GO:0004803	0.0072	50	4
Positive selection in species lacking the trait				
Intracellular signal transduction	GO:0035556	0.0045	95	26
Microtubule binding	GO:0008017	0.0009	38	15
Phosphatidylinositol phospholipase C activity	GO:0004435	0.0037	5	4
Binding	GO:0005488	0.0090	3,581	648
AP-type membrane coat adaptor complex	GO:0030119	0.0033	3	3
Intensified selection associated with the trait				
Protein phosphorylation	GO:0006468	0.0005	228	64
Intracellular signal transduction	GO:0035556	0.0020	94	30
Regulation of signal transduction	GO:0009966	0.0056	26	11
Protein dephosphorylation	GO:0006470	0.0093	24	10
Zinc ion binding	GO:0008270	0.00072	266	74
Ubiquitin binding	GO:0043130	0.00113	13	8
Microtubule binding	GO:0008017	0.00132	38	16
Protein kinase activity	GO:0004672	0.00211	228	63
Poly(ADP-ribose) glycohydrolase activity	GO:0004649	0.00771	3	3
Phospholipid binding	GO:0005543	0.00963	34	13
COPII vesicle coat	GO:0030127	0.0051	3	3
Integral component of membrane	GO:0016021	0.0092	542	112
Relaxed selection associated with the trait				
Lipid transport	GO:0006869	0.0036	29	7
Metallodipeptidase activity	GO:0070573	0.0031	5	3
DNA-binding transcription factor activity	GO:0003700	0.0076	107	15

priori hypothesis and treating each candidate as one of several related hypotheses). For *P* values, please see supplementary table S6, Supplementary Material online. Here, we focus on the results for each candidate gene treated independently. Most candidate genes are not evolving differently between species with and without the focal trait. However, 13 genes linked to worker reproduction do show differential evolution between species with reproduction workers and those without. Nine of these genes are under intensified selection in species with reproductive workers; two are under relaxed selection in these species; and four genes are under positive selection in species with nonreproductive workers.

Of the worker polymorphism candidate genes, eight are evolving differently in species with monomorphic workers compared with those with polymorphic workers. Four genes are under intensified selection in polymorphic species; four are under relaxed selection; two are under positive selection in polymorphic species; and one is under positive selection in monomorphic species. These results are summarized in table 4.

Discussion

In this study, we make use of published and unpublished ant genomes to test the relationship between social traits, patterns of genome-wide molecular evolution, and the evolution of specific candidate genes linked to social traits. Many of the global patterns we detect are potentially attributable to the impact of sociality-associated traits on effective population size. Eusociality in and of itself is already known to influence genome evolution via an effect on effective population size (Bromham and Leys 2005; Romiguier et al. 2014). Because most eusocial insects, including ants, are characterized by extreme reproductive skew, their effective population sizes are relatively reduced compared with solitary insects. A decrease in effective population size strengthens the impact of genetic drift on molecular evolution and concomitantly reduces the efficacy of natural selection.

Reproductive division of labor in eusocial insects has another evolutionary consequence: selection on worker genes shifts from direct selection because of personal fitness, to indirect or kin selection based on inclusive fitness, that is, benefit provided to the reproductive individual(s) in a colony. Theory predicts that natural selection will act less effectively on worker genes subject to this indirect selection (Linksvayer and Wade 2009). In some, but not all cases, empirical work in social insects has recovered this effect (Warner et al. 2017; Imrit et al. 2020).

Table 3
Gene Ontology Terms Enriched for Each Selective Regime Associated with Worker Polymorphism

Term	GO ID	*P* value	Total orthogroups annotated	Significant orthogroups
Positive selection in species with the trait				
Nucleic acid binding	GO:0003676	0.0012	1,005	43
Receptor ligand activity	GO:0048018	0.0075	29	4
Positive selection in species lacking the trait				
Phospholipid transport	GO:0015914	0.0029	3	3
Protein-containing complex assembly	GO:0065003	0.0069	69	18
Protein binding	GO:0005515	0.0060	1,499	258
Phospholipid transporter activity	GO:0005548	0.0061	6	4
Nucleus	GO:0005634	0.0053	282	52
Intensified selection associated with the trait				
Protein phosphorylation	GO:0006468	0.00086	238	50
Regulation of GTPase activity	GO:0043087	0.00866	7	4
ER to Golgi vesicle-mediated transport	GO:0006888	0.00866	7	4
Multicellular organism development	GO:0007275	0.00872	23	8
Regulation of microtubule-based process	GO:0032886	0.00920	4	3
Regulation of cellular component movement	GO:0051270	0.00920	4	3
Protein binding	GO:0005515	0.0010	1,496	259
Extracellular ligand-gated ion channel activity	GO:0005230	0.0023	49	15
Protein kinase activity	GO:0004672	0.0038	239	49
Enzyme binding	GO:0019899	0.0043	72	19
Signaling receptor binding	GO:0005102	0.0050	43	13
COPII vesicle coat	GO:0030127	0.003	3	3
Relaxed selection associated with the trait				
Carbohydrate metabolic process	GO:0005975	0.00028	112	28
Regulation of DNA-templated transcription, elongation	GO:0032784	0.00211	3	3
Carbohydrate derivative catabolic process	GO:1901136	0.00683	7	4
Coenzyme binding	GO:0050662	0.00082	104	26
Oxidoreductase activity	GO:0016491	0.00113	385	72
Vitamin binding	GO:0019842	0.00345	39	12
DNA-directed 5′-3′ RNA polymerase activity	GO:0003899	0.00622	14	6
Ion channel activity	GO:0005216	0.00770	121	26
Hydrolase activity, hydrolyzing O-glycosyl compounds	GO:0004553	0.00996	44	12

Here, we infer global patterns of positive selection in ant genomes. Positive selection is pervasive across the ant phylogeny, with evidence for positive selection at some point during ant evolution for 53.90% of all orthogroups. Previous comparative research using a smaller set of seven ant genomes also found pervasive positive selection (Roux et al. 2014), though to a lesser degree than found here. The difference in magnitude could be a result of our greater taxon sampling; we also made use of a method with greater power to detect positive selection (Murrell et al. 2015).

We also identify 37 Gene Ontology (GO) terms enriched for positive selection across the ants. These GO terms are linked to a diverse array of biological processes, and unlike previous work, we do not find any evidence for increased positive selection on mitochondrial functions in ants (Roux et al. 2014). The GO terms enriched in our study do have some overlap with the 22 GO terms identified by Roux and colleagues (2014). Both studies find that metalloendopeptidase activity and proteolysis are enriched for positive selection in ants (Roux et al. 2014).

In addition to characterizing broad patterns of positive selection, we examine how two important social traits, worker production of female reproductive offspring and worker polymorphism, shape molecular evolution in ants. These traits appear to modulate positive selection and selection intensity by variously exacerbating or ameliorating the effect of eusociality on effective population size.

In species where female workers can produce female reproductive offspring, we find a reduction in the extent of relaxed selection that is in line with our predictions. Because workers in these species are reproductive, they should experience direct natural selection rather than solely indirect kin selection. This would lead to the observed increase in selection intensity in these species.

In species where workers can produce female reproductive offspring, we see an unexpected reduction in the number of genes under positive selection, though this pattern

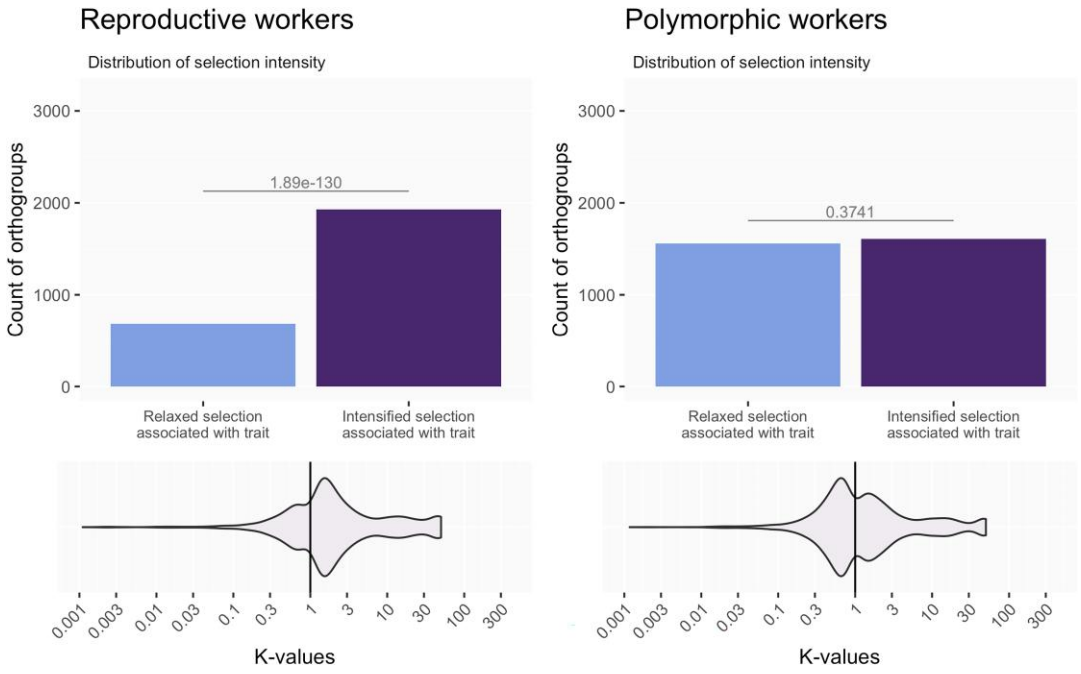

Fig. 4.—Patterns of selection intensity genome wide. In the top panel, for each trait, the count of orthogroups experiencing relaxed selection (in lighter blue) and the count of orthogroups experiencing intensified selection (in darker blue) in association with trait presence are shown. In species with reproductive workers, left, there are significantly fewer orthogroups experiencing relaxed selection versus intensified selection (chi-squared goodness-of-fit test, $P = 1.89e^{-130}$). In species with polymorphic workers, right, there is no significant difference in the number of orthogroups experiencing relaxed versus intensified selection (chi-squared goodness-of-fit test, $P = 0.3741$). In the bottom panel, for each trait, the distribution of significant *k*-values inferred by RELAX is shown. *k*-values measure the degree of selective intensity, with values less than 1 indicating relaxation of selection, and values greater than 1 indicating intensification. Because *k*-values can range from zero to fifty, the *x*-axis is log-transformed for visual clarity.

likely results from the low representation of species with reproductive workers in our dataset. We hypothesized that the pattern could be driven by the inclusion of *Ooceraea biroi*, the clonal raider ant, because its highly unusual reproductive biology may lead to reduced effective population size (Orive 1993; Milgroom 1996). To test this possibility, we separately analyzed the 1,138 orthogroups that did not contain *O. biroi* and which had been tested for positive selection associated with worker reproduction. Whereas the magnitude of the effect was reduced (29 orthogroups under positive selection in association with worker reproduction vs. 95 under positive selection in association with inability of workers to produce reproductive female offspring, compared with 169 and 1,312, respectively, in the entire dataset), the reduction in positive selection remains significant (chi-squared goodness-of-fit test, $P = 1.123e^{-09}$). Thus, clonality in *O. biroi* does not explain why positive selection is reduced in species with reproductive workers.

Worker sterility should primarily impact patterns of natural selection on genes with worker-biased patterns of expression, as these genes are most shielded from direct selection when workers are nonreproductive. This may help explain why we do not detect an increase in positive selection associated with worker reproduction, as worker-biased genes likely comprise a relatively small fraction of the genome. Future research ought to more explicitly interrogate patterns of selection on worker-biased genes to explore this pattern, for instance by precisely identifying genes with worker-biased patterns of expression throughout ontogeny.

A limitation of our approach is that we treat worker reproduction as a binary trait based on the ability of workers to produce reproductive female offspring. This is an imperfect proxy for what is in reality a continuous trait: the amount of reproductive contribution from queens versus workers. For instance, though we treat *Dinoponera quadriceps* and *Harpegnathos saltator* as having worker reproduction because any worker could theoretically produce reproductive female offspring following queen loss, in reality, only one worker per colony achieves dominance and reproduces at a given time (Liebig et al. 1999; Monnin and Peeters 1999). This means that any individual worker is unlikely to pass her genes to the next generation. In contrast, in some of the species we categorize as "nonreproductive" because their workers cannot produce reproductive female offspring, workers may still regularly produce male offspring, thus potentially contributing to the gene pool to a greater degree. We emphasize that future studies aimed at understanding the impact of worker reproduction on

Table 4
Selective Regimes for Focal Trait Candidate Genes; Names of Genes Evolving Differentially with Regard to Trait Presence/Absence Are Bolded

Candidate gene	NCBI gene symbol	Gene function in original study	Positive selection	Selection intensity	Differential evolution
Worker reproduction					
vitellogenin receptor-like isoform 1 XP_003402703	LOC100649042	Encodes a vitellogenin receptor; differentially expressed across castes and life stages in the bumblebee *Bombus terrestris.*	Trait presence or absence is not associated with a shift in selective regime.	Trait presence or absence is not associated with a shift in selective regime.	None
vitellogenin receptor-like isoform 2 XP_003402704	LOC100649042	Encodes a vitellogenin receptor; differentially expressed across castes and life stages in the bumblebee *B. terrestris.*	Trait presence or absence is not associated with a shift in selective regime.	Trait presence or absence is not associated with a shift in selective regime.	None
α-glucosidase	LOC100643608	A carbohydrate metabolism gene; upregulated in queens and reproductive workers of the bumblebee *B. terrestris.*	Trait presence or absence is not associated with a shift in selective regime.	Intensified selection is associated with the trait, prior to FDR correction	Some differential evolution.
glucose dehydrogenase	LOC105288220	A carbohydrate metabolism gene; downregulated in queens and reproductive workers of the bumblebee *B. terrestris.*	Trait presence or absence is not associated with a shift in selective regime.	Relaxed selection is associated with the trait, prior to FDR correction	Some differential evolution.
gemini	LOC105288150	A transcription factor which regulates sterility in honeybee workers; also alternatively spliced between arrhenotokous and thelytokous honeybee strains.	Trait presence or absence is not associated with a shift in selective regime.	Intensified selection is associated with the trait after FDR correction	Some differential evolution.
Ecdysis triggering hormone receptor	LOC105835772	A hormone receptor which regulates ecdysis and juvenile hormone synthesis; linked to a locus associated with thelytoky in honeybees.	Trait presence or absence is not associated with a shift in selective regime.	Trait presence or absence is not associated with a shift in selective regime.	None
Methoprene-tolerant	LOC109861318	Encodes a juvenile hormone receptor; upregulated in reproductive versus nonreproductive adult females in the termine *Cryptotermes secondus.*	Trait presence or absence is not associated with a shift in selective regime.	Trait presence or absence is not associated with a shift in selective regime.	None
Taiman	LOC410637	Dimerizes with ligand-bound juvenile hormone receptor and inhibits metamorphosis; has reproductive female-biased expression in the termite *Cryptotermes secundus* but worker-biased expression in the termite *Macrotermes natalensis.*	Trait presence or absence is not associated with a shift in selective regime.	Intensified selection is associated with the trait after FDR correction	Some differential evolution.

(continued)

Table 4 Continued

Candidate gene	NCBI gene symbol	Gene function in original study	Positive selection	Selection intensity	Differential evolution
Krüppel-homolog 1	LOC112589664	A transcription factor and key regulator mediating the repressive action of juvenile hormone on insect metamorphosis.	Trait presence or absence is not associated with a shift in selective regime.	Intensified selection is associated with the trait, prior to FDR correction	Some differential evolution.
JH epoxidase	LOC105432897	A juvenile hormone biosynthesis gene; upregulated in reproductive females of the termite *Macrotermes natalensis.*	Trait presence or absence is not associated with a shift in selective regime.	Trait presence or absence is not associated with a shift in selective regime.	None
JHAMT	LOC105255785	A juvenile hormone biosynthesis gene; upregulated in reproductive females of the termite *Macrotermes natalensis.*	Trait presence or absence is not associated with a shift in selective regime.	Trait presence or absence is not associated with a shift in selective regime.	None
broad	LOC105194137	A juvenile hormone signalling gene; evolves at a different rate in social termites compared with a nonsocial cockroach.	Trait presence or absence is not associated with a shift in selective regime.	Trait presence or absence is not associated with a shift in selective regime.	None
JH epoxide hydrolase-like 1	LOC108693097	A juvenile hormone degradation gene; differentially expressed between workers and reproductive females across three termite species.	Trait presence or absence is not associated with a shift in selective regime.	Trait presence or absence is not associated with a shift in selective regime.	None
Ecdysone-induced protein 93F	LOC108775748	An adult specifier gene; upregulated in reproductive females of the termite *Cryptotermes secondus* but upregulated in worker females of the termite *Macrotermes natalensis.*	Trait presence or absence is not associated with a shift in selective regime.	Trait presence or absence is not associated with a shift in selective regime.	None
Phosphomevalonate kinase	LOC106751928	Involved in juvenile hormone synthesis; differentially expressed between workers and reproductive females in the termite *Macrotermes natalensis*.	Trait presence or absence is not associated with a shift in selective regime.	Trait presence or absence is not associated with a shift in selective regime.	None
Isopentenyl-diphosphate delta-isomerase	LOC115234887	Involved in juvenile hormone synthesis; differentially expressed between workers and reproductive females in the termite *Macrotermes natalensis*.	Positive selection in the background only, after FDR correction	Trait presence or absence is not associated with a shift in selective regime.	Some differential evolution.
Neofem6/takeout/Cryptotermes secundus protein takeout (LOC111867885), mRNA	LOC112457888	Identified as a putative Takeout homolog in the termite *Cryptotermes secondus*; upregulated in queens of the species.	Trait presence or absence is not associated with a shift in selective regime.	Intensified selection is associated with the trait after FDR correction	Some differential evolution.

(continued)

Table 4 Continued

Candidate gene	NCBI gene symbol	Gene function in original study	Positive selection	Selection intensity	Differential evolution
Neofem9/Histone H2A	EU546151	Identified as a histone 2A homolog in the termite *Cryptotermes secondus*; upregulated in queens of the species.	Trait presence or absence is not associated with a shift in selective regime.	Trait presence or absence is not associated with a shift in selective regime.	None
Neofem 11/XP_974805	EU546153	Upregulated in queens of the termite *Cryptotermes secondus*.	Trait presence or absence is not associated with a shift in selective regime.	Trait presence or absence is not associated with a shift in selective regime.	None
Vasa	LOC105250443	An oocyte patterning gene; Vasa mislocalization in developing worker-laid eggs is associated with inviability in several ant species.	Trait presence or absence is not associated with a shift in selective regime.	Intensified selection is associated with the trait after FDR correction	Some differential evolution.
nanos	LOC115237921	An oocyte patterning gene; nanos mislocalization in developing worker-laid eggs is associated with inviability in several ant species.	Trait presence or absence is not associated with a shift in selective regime.	Trait presence or absence is not associated with a shift in selective regime.	None
lilliputian	LOC105195245	Involved in insulin signalling and morphogenesis; differentially expressed between queen and worker bees in the genus *Mellipona*.	Trait presence or absence is not associated with a shift in selective regime.	Trait presence or absence is not associated with a shift in selective regime.	None
Anarchy	LOC105181529	Mediates programmed cell death; identified as a candidate gene for worker sterility in honeybees.	Trait presence or absence is not associated with a shift in selective regime.	Trait presence or absence is not associated with a shift in selective regime.	None
ark	LOC105282532	A pro-apoptotic peptidase; highly expressed in prepupal worker honeybee ovaries.	Positive selection in the background only, after FDR correction	Trait presence or absence is not associated with a shift in selective regime.	Some differential evolution.
βFTZ-F1	LOC109855278	Integrates juvenile hormone and ecdysone responses in insect metamorphosis.	Trait presence or absence is not associated with a shift in selective regime.	Trait presence or absence is not associated with a shift in selective regime.	None
sdr	LOC105144228	Encodes a a short-chain dehydrogenase/reductase; overexpressed in larval honeybee queen ovaries.	Trait presence or absence is not associated with a shift in selective regime.	Trait presence or absence is not associated with a shift in	None

(continued)

Table 4 Continued

Candidate gene	NCBI gene symbol	Gene function in original study	Positive selection	Selection intensity	Differential evolution
				selective regime.	
oat	LOC105832518	Encodes a putative ornithine-oxo-acid transaminase; overexpressed in larval honeybee worker ovaries.	Positive selection in the background only, after FDR correction	Intensified selection is associated with the trait after FDR correction	Some differential evolution.
fringe/lunatic fringe	LOC112460186	An activator of Notch and insulin signalling; a long-coding RNA comprising one of the fringe/lunatic fringe introns; differentially expressed between developing queen and worker ovaries in honeybees.	Trait presence or absence is not associated with a shift in selective regime.	Trait presence or absence is not associated with a shift in selective regime.	None
vitellogenin receptor	LOC105182956	A vitellogenin receptor; expression levels in honeybee ovaries are correlated with ovary activation.	Trait presence or absence is not associated with a shift in selective regime.	Trait presence or absence is not associated with a shift in selective regime.	None
Myc-B	LOC105195393	Encodes a transcriptional regulator in the Wnt, Hippo, and TGF-beta pathways; differentially expressed between honeybee worker larval transplanted to queen cells at different ages.	Trait presence or absence is not associated with a shift in selective regime.	Trait presence or absence is not associated with a shift in selective regime.	None
casein kinase 1	LOC108772916	Encodes a transcriptional regulator in the Wnt, FoxO, Hedgehog, Hippo, TGF-beta, and longevity regulating pathways; differentially expressed between honeybee worker larval transplanted to queen cells at different ages.	Trait presence or absence is not associated with a shift in selective regime.	Intensified selection is associated with the trait after FDR correction	Some differential evolution.
S-phase kinase-associated protein 2	LOC105143861	Encodes a transcriptional regulator in the FoxO and mTOR pathways; differentially expressed between honeybee worker larval transplanted to queen cells at different ages.	Trait presence or absence is not associated with a shift in selective regime.	Relaxed selection is associated with the trait after FDR correction	Some differential evolution.
neuroparsin-A	LOC105254204	Involved in hormonal regulation of insect reproduction; differentially expressed according to reproductive status in several hymenopteran species.	Trait presence or absence is not associated with a shift in selective regime.	Trait presence or absence is not associated with a shift in selective regime.	None
cGMP-dependent protein kinase 1 (PRKG1)	LOC105189468	A cGMP-dependent protein kinase; downregulated in workers of the bumblebee *B. terrestris* in response to queen presence.	Trait presence or absence is not associated with a shift in selective regime.	Trait presence or absence is not associated with a shift in selective regime.	None

(continued)

Table 4 Continued

Candidate gene	NCBI gene symbol	Gene function in original study	Positive selection	Selection intensity	Differential evolution
SLCO2A1	LOC106748283	Encodes a solute carrier organic anion transporter, acting in prostaglandin transport to regulate reproduction and immunity in insects; upregulated in workers of the bumblebee *B. terrestris* in response to queen presence.	Trait presence or absence is not associated with a shift in selective regime.	Trait presence or absence is not associated with a shift in selective regime.	None
MACF1	LOC106744841	Encodes microtubule-actin cross-linking factor 1; upregulated in workers of the bumblebee *B. terrestris* in response to queen presence.	Positive selection in the background only, after FDR correction	Intensified selection is associated with the trait after FDR correction	Some differential evolution.
mucin-5AC	LOC105200248	Encodes a gel-forming glycosylated protein of unknown function in insects; upregulated in workers of the bumblebee *B. terrestris* in response to brood presence.	Trait presence or absence is not associated with a shift in selective regime.	Trait presence or absence is not associated with a shift in selective regime.	None
Worker polymorphism					
short neuropeptide F receptor	LOC105195500	Regulates foraging behavior in honeybees; more highly expressed in foragers than nurses	Trait presence or absence is not associated with a shift in selective regime.	Trait presence or absence is not associated with a shift in selective regime.	None
vestigial	LOC105251765	Coordinates wing imaginal disc growth; experimental evidences shows this genes is required for development of the soldier subcaste in the ant *Pheidole hyatti.*	Positive selection in the foreground only, after FDR correction	Intensified selection is associated with the trait after FDR correction	Some differential evolution.
vitellogenin (Vg)	LOC115243034	Plays a role in oogensis; has queen-biased expression in the ant *Formica fusca*.	Trait presence or absence is not associated with a shift in selective regime.	Trait presence or absence is not associated with a shift in selective regime.	None
Vg-like A	LOC115239126	A homolog of convential vitellogenin; has forager-biased expression in the ant *Formica fusca*.	Trait presence or absence is not associated with a shift in selective regime.	Trait presence or absence is not associated with a shift in selective regime.	None
Vg-like B	LOC115241951	A homolog of convential vitellogenin; has forager-biased expression in the ant *Formica fusca*.	Trait presence or absence is not associated with a shift in selective regime.	Trait presence or absence is not associated with a shift in selective regime.	None
Vg-like C	LOC115241401	A homolog of convential vitellogenin which lacks most structural features of conventional vitellogenin; has	Trait presence or absence is not associated with a shift in selective regime.	Relaxed selection is associated with the trait, prior to FDR correction	Some differential evolution.

(continued)

Table 4 Continued

Candidate gene	NCBI gene symbol	Gene function in original study	Positive selection	Selection intensity	Differential evolution
		forager-biased expression in the ant *Formica fusca*.			
mAchR	LOC105258758	A muscarinic acetylcholine receptor gene; differentially clock-controlled in foragers and nurses of the ant *Camponotus floridanus*.	Trait presence or absence is not associated with a shift in selective regime.	Trait presence or absence is not associated with a shift in selective regime.	None
DopEcR	Dmel_CG18314	An insect dopamine/ecdysteroid receptor gene; differentially clock-controlled in foragers and nurses of the ant *Camponotus floridanus*.	Trait presence or absence is not associated with a shift in selective regime.	Intensified selection is associated with the trait, prior to FDR correction	Some differential evolution.
Cyclin-dependent kinase 4 (Cdk4)	LOC106744310	A kinase gene whose expression levels cycle with different periodicities in foragers and nurses of the ant *Camponotus floridanus*; also more highly expressed in forager brains than nurse brains.	Trait presence or absence is not associated with a shift in selective regime.	Trait presence or absence is not associated with a shift in selective regime.	None
Alanine—glyoxylate aminotransferase 2-like (Agxt2)	LOC105254428	Expression levels cycle with different periodicities in foragers and nurses of the ant *Camponotus floridanus*; also more highly expressed in nurse brains than forager brains.	Positive selection in the background only, after FDR correction	Relaxed selection is associated with the trait after FDR correction	Some differential evolution.
D-3-phosphoglycerate dehydrogenase(Phgdh)	LOC105252893	Expression levels cycle with different periodicities in foragers and nurses of the ant *Camponotus floridanus*; also more highly expressed in nurse brains than forager brains.	Positive selection in the foreground only, prior to FDR correction	Intensified selection is associated with the trait, prior to FDR correction	Some differential evolution.
Carcinine transporter-like (CarT)	LOC105257575	Expression levels cycle with different periodicities in foragers and nurses of the ant *Camponotus floridanus*; also more highly expressed in nurse brains than forager brains.	Trait presence or absence is not associated with a shift in selective regime.	Trait presence or absence is not associated with a shift in selective regime.	None
Venom carboxylesterase-6	LOC105253143	An orally-transfered protein and juvenile hormone esterase; gene expression levels cycle with different periodicities in foragers and nurses of the ant *Camponotus floridanus*; also more highly expressed in nurse brains than forager brains.	Trait presence or absence is not associated with a shift in selective regime.	Trait presence or absence is not associated with a shift in selective regime.	None
Arylphorin subunit alpha	LOC105250772	Encodes a larval storage protein; upregulated in nurse ants compared with forager ants in *Camponotus floridanus*.	Trait presence or absence is not associated with a shift in selective regime.	Relaxed selection is associated with the trait, prior to FDR correction	Some differential evolution.
Vg1	LOC105205865	A vitellogenin homolog; more highly expressed in major workers than medium or minor	Trait presence or absence is not associated with a	Trait presence or absence is not associated with	None

(continued)

Table 4 Continued

Candidate gene	NCBI gene symbol	Gene function in original study	Positive selection	Selection intensity	Differential evolution
		workers in the ant *Solenopsis invicta*.	shift in selective regime.	a shift in selective regime.	
Vg2	LOC105205782	A vitellogenin homolog; more highly expressed in major workers versus minor workers in the ant *Solenopsis invicta*.	Trait presence or absence is not associated with a shift in selective regime.	Trait presence or absence is not associated with a shift in selective regime.	None
spalt	LOC105208108	Encodes a transcription factor which helps regulate insect wing development; has a soldier-specific pattern of expression in the vestigial wing imaginal discs of *Pheidole morrisi* soldier ants.	Trait presence or absence is not associated with a shift in selective regime.	Intensified selection is associated with the trait after FDR correction	Some differential evolution.
foraging	LOC105202008	Encodes a cGMP-dependent protein kinase linked to foraging behavior and insect polyphenisms; more highly expressed in the heads of foragers than nurses in the ant *Solenopsis invicta*, where RNAi knockdown in foragers induces a transition to nurse behavior.	Trait presence or absence is not associated with a shift in selective regime.	Relaxed selection is associated with the trait, prior to FDR correction	Some differential evolution.
OBP6	LOC108688727	An odorant binding protein differentially expressed between large and tiny workers in the ant *Atta vollenweideri*.	Trait presence or absence is not associated with a shift in selective regime.	Trait presence or absence is not associated with a shift in selective regime.	None
OBP9	LOC105205352	An odorant binding protein differentially expressed between large and tiny workers in the ant *Atta vollenweideri*.	Trait presence or absence is not associated with a shift in selective regime.	Trait presence or absence is not associated with a shift in selective regime.	None
CSP9	LOC105623881	A chemosensory protein differentially expressed between large and tiny workers in the ant *Atta vollenweideri*.	Trait presence or absence is not associated with a shift in selective regime.	Trait presence or absence is not associated with a shift in selective regime.	None
IR93a	LOC105432932	An ionotropic receptor gene more highly expresesed in tiny workers than other workers in the ant *Atta vollenweideri*.	Trait presence or absence is not associated with a shift in selective regime.	Trait presence or absence is not associated with a shift in selective regime.	None
AvOR131	LOC108689490	An odorant binding protein differentially expressed between large and tiny workers in the ant *Atta vollenweideri*.	Trait presence or absence is not associated with a shift in selective regime.	Trait presence or absence is not associated with a shift in selective regime.	None

molecular evolution in social insects should actually quantify the level of reproduction contributed by workers in each species and use methods designed to handle continuous traits to better understand the impact of worker reproduction on molecular evolution. We are also limited by the phylogenetic distribution of sequenced genomes, because many sampled species with sterile workers are in a single clade, and the majority of sampled species do not have reproductive workers. Both factors limit our ability to detect positive selection in the foreground. Improved sampling of ant species with diverse social traits will strengthen future comparative work.

Contrary to our predictions, we find fewer genes under positive selection in species with polymorphic workers than in species with monomorphic workers. Based on a bootstrapping analysis, this pattern may result from the low sampling of these species in our analysis. We also find no difference in the degree of relaxed versus intensified selection genome wide, although single-copy orthogroups show evidence for intensified selection associated with worker polymorphism. We predicted that worker polymorphism would lead to an increase in positive selection because the existence of differential gene expression between castes alleviates caste-antagonistic pleiotropy, freeing differentially expressed genes to be acted upon by positive directional selection (Schrader et al. 2017).

Whereas this may occur for some individual genes, at the level of the genome, that effect may be swamped by another, broader consequence of caste polymorphism. In ants, alternative worker castes are plastic phenotypes which arise via conditional gene expression throughout development. Conditionally expressed genes have reduced exposure to natural selection, including reduced exposure to positive selection (Snell-Rood et al. 2010; Van Dyken and Wade 2010; Kapheim 2019). Plastic development may shield genes from positive selection in polymorphic species.

Species with worker polymorphism also tend to have larger colonies (Ferguson-Gow et al. 2014). In species with large colonies, queen-to-worker ratios are reduced, which may in turn reduce effective population size and concomitantly reduce the strength of natural selection in these species. Together, the shielding effects of plastic development and the relationship between worker polymorphism and colony size may explain the patterns we detect here. We also note that we sample a greater number of monomorphic than polymorphic species in these analyses, possibly limiting the ability to detect positive selection in this group.

We expected worker caste polymorphism would also increase relaxed selection, due to the shielding effect of gene expression plasticity. However, at the level of the genome, we find no difference in the number of genes under relaxed selection compared with intensified selection. Our finding could in part stem from the relatively low representation of polymorphic species in our set of genomes, just five of 22 genomes. Moreover, the dataset includes only one strongly dimorphic species, *Cephalotes varians*; the remainder of the polymorphic species have continuous, size-based worker polymorphisms. Genes may have a more pronounced pattern of conditional expression in species with strong morphological dimorphism than in continuously polymorphic species, leading to a more pronounced effect of caste polymorphism on molecular evolution in the former group. Increased sampling of strongly dimorphic species—for example, by including *Pheidole* ants, *Colobopsis* species with repletes, or some *Carebara* species—would be useful in future studies of caste and molecular evolution.

Whereas we identify significantly enriched GO terms across all selective regimes and for all traits, there are not clear connections between the sociality-associated trait and associated GO terms. This is expected: because we are looking broadly across the entire genome, and because we predict that our sociality-associated traits are mediating genome evolution via neutral processes, our comparative approach is not intended to identify genes or gene categories that are functionally relevant for particular traits. We also note that GO enrichment has some inherent limitations. For instance, processes and pathways relevant to eusociality may be poorly represented, because GO evidence largely comes from model organisms like *Drosophila*; these limitations have been noted in other studies of social insects (Primmer et al. 2013; Korb et al. 2021).

Whereas clearly sociality-related GO terms are not enriched for selection at the level of the genome, we identify individual worker reproduction and worker polymorphism candidate genes with noteworthy patterns of evolution. These candidate genes provide insight into the impact of social elaboration on the genetic basis underpinning sociality in ants.

Of 36 ant orthogroups linked to worker reproduction in eusocial insects, 12 are evolving differently in species with reproductive workers compared with those without. A majority of these genes are under intensified selection in species with reproductive workers, including known hormone signaling and developmental genes. Two genes involved in juvenile hormone signaling are under intensified selection in species with reproductive workers: *Taiman* and *Krüppel-homolog 1*. The Taiman protein forms a complex with ligand-bound methoprene-tolerant (the juvenile hormone receptor); this complex acts to repress metamorphosis (Lozano et al. 2014). Taiman also acts in many other morphogenic developmental pathways (Lozano et al. 2014). In termites, *Taiman* is differentially expressed between reproductive and nonreproductive females (Jongepier et al. 2018).

The protein Krüppel-homolog 1 functions downstream of juvenile hormone signaling. Juvenile hormone induces transcription of *Krüppel-homolog 1*, which impacts

multiple pathways related to both caste polymorphism and reproduction (Zhang et al. 2021). In bumblebees, *Krüppel-homolog 1* is downregulated in subordinate nonreproductive workers compared with dominant reproductive workers (Shpigler et al. 2010). Through interplay between juvenile hormone and insulin signaling, *Krüppel-homolog 1* may help control reproductive status in the queenless ant *Pristomyrmex punctatus* (Araki et al. 2020). It appears that multiple components of the juvenile hormone signaling pathway are important in the evolution of worker reproduction.

It is also noteworthy that two of the worker reproduction candidate genes under intensified selection are known oocyte patterning and development genes. Developmental research has indicated that functional sterility in worker ants results from mislocalization of maternal determinants and other key proteins and RNAs in developing worker oocytes (Khila and Abouheif 2008). Vasa is one such protein. In multiple ant species, Vasa protein mislocalization is associated with inviability of worker-laid eggs (Khila and Abouheif 2008). The *vasa* gene is also expressed in ant queen germline stem cells, where it may be necessary for proper ovariole function (Khila and Abouheif 2010). Our results show that *vasa* is evolving under intensified selection in species with reproductive workers. This pattern indicates that, in these species, *vasa* is reexposed to direct selection in workers, concomitant with a role in proper worker oocyte development.

Whereas selection on oocyte patterning proteins may play a role in the evolution of worker reproduction, developmental research predicts that genes responsible for proper protein localization are even more central to this process (Khila and Abouheif 2008). Developmental proteins are localized in the oocyte by cytoskeletal and microtubule proteins. We identified *microtubule-actin cross-linking factor 1* (*MACF1*) as a worker reproduction candidate gene because it is differentially expressed in bumblebee workers in response to queen and brood presence (Santos et al. 2022). *MACF1* functions in cytoskeletal organization during oocyte development, where it helps properly localize crucial developmental proteins (Gupta et al. 2010; Krishnakumar et al. 2018). Our selection analyses show that *MACF1* is under positive selection only in species without reproductive workers, whereas it evolves under intensified, potentially purifying, selection in species with reproductive workers (Cui et al. 2021). Possibly, *MACF1* is selected towards a new function in species with nonreproductive workers, whereas being constrained in species whose workers produce viable female eggs. Interestingly, the GO term "microtubule binding" follows the same evolutionary pattern as *MACF1*: enriched for positive selection in species without reproductive workers and enriched for intensified selection in species with reproductive workers. This pattern hints that microtubule-related processes are important in worker reproduction more broadly.

Other developmental processes are under positive selection in ant species that lack reproductive workers. In honeybees, differential regulation of apoptosis during development helps establish reproductive differences between queens and workers (Dallacqua and Bitondi 2014). This process involves differential expression of *apoptotic peptidase activating factor (Apaf)-related killer gene (ark)* (Dallacqua and Bitondi 2014). Proapoptotic *ark* is highly expressed in prepupal workers and localizes to developing ovaries (Dallacqua and Bitondi 2014). In our results, *ark* is under positive selection in ant species with nonreproductive workers, suggesting it also helps limit worker reproductive capacity in ants.

Of genes linked to worker polymorphism, we identify eight candidates that are under differential selection between species with monomorphic and polymorphic workers. Four of these genes are evolving under relaxed selection in polymorphic species. All four are genes linked to behavioral caste roles in monomorphic species, which suggests that behavioral plasticity genes become less evolutionarily constrained when morphological subcastes evolve.

Like other hymenopterans, ants have multiple homologs of the gene *vitellogenin* (Gadau et al. 2012; Morandin et al. 2014). These homologs appear to have taken on new roles associated with sociality (Morandin et al. 2014). In particular, the gene *vitellogenin-like-C* lacks most conventional *vitellogenin* structural elements, indicating potential sub- or neofunctionalization (Morandin et al. 2019). In *Formica* ants, *vitellogenin-like-C* is upregulated in workers, particularly foragers, suggesting a role in behavioral division of labor (Morandin et al. 2019; Zhao et al. 2021). It is similarly implicated in a variety of bees (Morandin et al. 2019; Zhao et al. 2021). We find that *vitellogenin-like-c* is under relaxed selection in ant species with polymorphic workers. This is in contrast to patterns found in bees, where *vitellogenin-like-c* evolves under purifying selection (Zhao et al. 2021). Because bees lack morphologically distinct worker subcastes, this pattern suggests that *vitellogenin-like-c* helps generate behavioral division of labor in multiple eusocial insect taxa.

Finally, the gene *foraging* evolves under relaxed selection in worker polymorphic ant species. *Foraging* encodes a cGMP-dependent protein kinase associated not just with foraging behavior in insects, but also with insect polyphenisms ranging from desert locusts to pea aphids to ants (Lucas and Sokolowski 2009; Tarès et al. 2013; Tobback et al. 2013; Chen et al. 2022). RNAi knockdowns and pharmacological manipulations of *foraging* in *Solenopsis invicta* cause behavioral transitions between nurse and forager roles (Chen et al. 2022). Again, our finding supports a general pattern that evolutionary constraint on behavioral

caste genes is reduced when morphological subcastes emerge.

Other candidate genes are under intensified, and in some cases positive, selection in species with worker polymorphisms. Whereas the developmental mechanisms leading to morphologically distinct worker subcastes in ants are not yet fully understood, evidence in *Pheidole* ants indicates that the development of vestigial forewing imaginal discs is necessary to produce soldier morphology (Rajakumar et al. 2012, 2018). Excitingly, we find that two genes implicated in this process are under intensified selection in worker polymorphic ant species.

In insects, the gene *vestigial* encodes a key regulatory protein specifying wing development (Gramates et al. 2022). RNAi knockdown of *vestigial* demonstrates that it is necessary for proper soldier development in *Pheidole* ants (Rajakumar et al. 2018). Thus, *vestigial* is not just a key gene regulating wing polyphenism, but also regulating worker subcaste fate. We find that *vestigial* is evolving under *both* positive and intensified selection in species with polymorphic workers, further implicating *vestigial* in caste polymorphism.

Another wing gene evolves under intensified selection in species with polymorphic workers: *spalt*. *Spalt* is a transcription factor with important regulatory roles in insect development, particularly wing development (Gramates et al. 2022). Like *vestigial*, *spalt* is a key gene in the development of morphologically distinct castes in *Pheidole* (Rajakumar et al. 2012). In species with worker polymorphism, *spalt* is evolving under intensified selection, but not positive selection, suggesting that *spalt* is experiencing purifying selection in these species (Cui et al. 2021). Together, these findings indicate that wing imaginal discs could be involved in generating caste polymorphisms not just in *Pheidole* ants, but in other polymorphic species as well.

Whereas many candidate genes show differential patterns of evolution which expand our understanding of their role in producing complex eusocial traits, a majority of our candidate genes are not evolving differently across species with and without the focal trait. This is interesting in light of the social toolkit hypothesis, which posits that independent origins of sociality are underlaid by evolutionary changes to a common toolkit of genes (Toth and Robinson 2007; Rehan and Toth 2015). A larger scale, comparative investigation of genes linked to social traits would help clarify whether a genetic toolkit also helps produce elaborations of eusociality. Moreover, most of our candidate genes were first identified through differential gene expression studies. This suggests that evolution of cis- and trans-regulatory elements may play a more important role in generating sociality-associated phenotypes than does evolution of protein-coding genes. Comparative work examining the evolution of regulatory elements will clarify their role in the production of social elaborations.

We find that sociality-associated traits impact genome evolution, sometimes in surprising ways. Worker reproduction is associated with reduced positive and relaxed selection, whereas worker polymorphism is associated with reduced positive selection but no change in selection intensity. To more clearly understand the patterns we uncover here, future genomic research should particularly target groups with key sociality-associated traits, including dimorphic workers, and should more precisely quantify traits of interest, for example, by determining the relative reproductive contributions of workers and queens in species where workers can reproduce. We also infer patterns of evolution for many candidate genes linked to both worker reproduction and worker polymorphism and find that a number of notable candidates are evolving differentially in association with these traits. Broader assessments of genes involved in worker reproduction and polymorphism will clarify whether a genetic toolkit produces these social traits. Taken together, our results provide evidence that elaborations to eusociality shape variable patterns of genome evolution and expand our understanding of the genetic mechanisms underlying these elaborations.

Materials and Methods

C. varians Genome Sequencing

Included in our analysis is the newly sequenced genome of the ant *C. varians* (Smith 1876). For genome sequencing, DNA from a single haploid male specimen was extracted using phenol–chloroform DNA extraction (Moreau 2014) and sequenced using the Illumina 2000/2500 platform. A single queen, soldier, worker, callow worker, large larvae, and small larvae were used for transcriptome sequencing. The genome was assembled using ALLPATHS-LG version 52488 and annotated following Rubin (2022). These data are accessioned under NCBI BioProject PRJNA864742.

Taxon Sampling and Data Access

For selection analyses, we included 21 species with publicly available genome sequences, and the newly sequenced *C. varians* genome. For genera with multiple available genomes, and which had identical trait values for all traits being considered in this study (*Atta* and *Trachymyrmex*), we selected only the genome with the highest BUSCO score. For a complete list of species and genome accession numbers, please see supplementary table S1, Supplementary Material online. We filtered coding sequences to retain only the longest isoform per gene using the R package orthologr (Drost et al. 2015) in R version 4.0.5 (R Core Team 2019). To facilitate our phylogenetically informed tests of selection, we made use of the ant phylogeny from Blanchard and Moreau 2017. We translated coding

sequences to amino acid sequences for orthogroup inference using TransDecoder (Haas 2021).

Compiling Species-Level Trait Data

For each species, we compiled data on the following traits: capacity for worker reproduction and degree of worker polymorphism. Data were gathered from the literature, and from consultation with experts (see supplementary table S1, Supplementary Material online).

Identifying Orthologous Sequence Groups

We inferred orthogroups (groups of genes descended from a single gene in the most recent common ancestor of the taxon set) using Orthofinder v2.4.1, using default parameters and specifying a trimmed version of the tree from Blanchard and Moreau (2017) as the species tree (Emms and Kelly 2015, 2019).

Inferring Patterns of Molecular Evolution

Sequence Alignment

Methods for assessing natural selection require aligned nucleotide sequences. To generate these inputs, we ran MAFFT, as implemented by OrthoFinder v2.4.1, to generate multiple sequence alignments for the orthogroup amino acid sequences (Katoh et al. 2002). For orthogroups with fewer than four sequences, for which OrthoFinder does not infer gene trees or multiple sequence alignments, we separately ran FastTree and MAFFT to generate inputs to downstream analyses (Katoh et al. 2002, Price et al. 2010). We then used PAL2NAL with default parameters to produce codon-aware alignments of the corresponding nucleotide sequences (Suyama et al. 2006).

Assessing Positive Selection across the Phylogeny with BUSTED[S]

To explore patterns of positive selection across all genes and all species, we used BUSTED[S] as implemented in the HyPhy package, with default parameters and allowing synonymous rate variation with the flag –srv Yes (Kosakovsky Pond et al. 2020; Wisotsky et al. 2020). This analysis answers the question of whether a gene experienced positive selection for at least one site on at least one branch being tested. Critically, the BUSTED[S] method allows rates of both nonsynonymous and synonymous substitution to vary. We implemented this method because holding synonymous substitution rates fixed, as is common in other methods, is associated with high rates of false positive results (Wisotsky et al. 2020). We implemented Benjamini–Hochberg false discovery rate corrections on the *P* values reported by BUSTED[S].

Assessing Positive Selection in Association with Traits with BUSTED-PH

To test for evidence of positive selection explicitly associated with the focal traits, we used BUSTED-PH, a standalone analysis from the HyPhy package (Kosakovsky Pond et al. 2020). The method works by fitting a step-wise series of four models: First, an unrestricted branch–site random effects likelihood model is fitted, which gives the foreground and background branche–independent dN/dS distributions; next, a constrained model is fitted, where the foreground branches are constrained to $dN/dS \leq 1$. This model and the unrestricted model are compared with a likelihood ratio test, to ascertain if foreground branches are subject to positive selection; third, a constrained model is fitted that constrains the background branches to $dN/dS \leq 1$. As before, this model and the unrestricted model are compared with a likelihood ratio test, to determine if background branches are subject to positive selection; finally, a constrained model is fitted with the same dN/dS distribution for both the foreground and background branches. This model and the unrestricted model are compared with a likelihood ratio test, to test whether selective regimes differ between foreground and background branches.

This method allowed us to identify orthogroups under selection in only species with a trait of interest, and orthogroups under selection in only species lacking that trait of interest. We labeled foreground tips in orthogroup gene trees using a custom R script (available at https://github.com/mbarkdull/FormicidaeMolecularEvolution/blob/main/scripts/LabellingPhylogeniesHYPHY.R). Internal branches were also labelled as foreground if they led exclusively to foreground tips. BUSTED-PH was run with default parameters, including allowing synonymous site variation with the flag –srv Yes.

Because of our unequal sampling of foreground species relative to background species (worker reproduction, $N = 3$ foreground species; worker polymorphism, $N = 6$), we further assessed the impact of foreground species sampling on our ability to detect positive selection. We used two approaches to test for this potential pattern in our results. First, we repeated each positive selection analysis using only the subset of orthogroups with equal representation from foreground and background species for each trait; these results are presented in supplementary table S4, Supplementary Material online. Second, we implemented a bootstrapping approach to measure the effect of foreground species sample number on our ability to detect positive selection. Briefly, we selected a random set of either three (in the case of worker reproduction) or six (in the case of worker polymorphism) species to label as foreground species. We then randomly selected 100 orthogroups to test for positive selection with that set of foreground species. We repeated the test 50 times, selecting a new set of foreground species

for each iteration whereas keeping the set of orthogroups constant. Results were corrected with Benjamini–Hochberg false discovery rate corrections. These results are summarized in figure 3 and supplementary table S5, Supplementary Material online.

Assessing Relaxation of Selection in Association with Traits with RELAX

To test for relaxation of selection associated with the set of focal traits, we used RELAX as implemented in the HyPhy package (Wertheim et al. 2015; Kosakovsky Pond et al. 2020). RELAX assesses changes in the intensity of purifying and/or positive selection concomitantly (Wertheim et al. 2015). The premise of RELAX is that when selection is relaxed, the dN/dS of sites previously under positive (dN/dS > 1) or purifying selection (dN/dS < 1) will shift closer to neutrality (i.e., closer to one). The shift in dN/dS for focal, foreground branches in a phylogeny relative to background branches is described by the selection intensity parameter *k*. If *k* is significantly greater than one, selection is inferred to have been intensified along the foreground branches; if it is significantly less than one, selection is inferred to have been relaxed along the foreground branches. For this analysis, species with the focal trait were assigned as foreground species, and those lacking the trait were assigned as background species, using the same foreground tip labeling approach described above. RELAX was run with default parameters, including allowing synonymous site variation with the flag –srv Yes.

Statistical Analyses

Using the results of BUSTED-PH and RELAX, we performed additional statistical tests to explicitly address our hypotheses. For full statistical details, see supplementary table S2, Supplementary Material online.

BUSTED-PH Downstream Analyses

BUSTED-PH assesses evidence for positive selection in relation to a phenotype. We wanted to understand whether the presence of our focal traits was associated with increased or reduced levels of positive selection. To do this, we implemented Benjamini–Hochberg false discovery rate corrections on the *P* values from BUSTED-PH and subsequently classified orthogroups as under selection only in species with a trait or under selection only in species without a trait. We tested for differences in the proportion of genes in each category for each trait using chi-squared goodness-of-fit tests, fit to the hypothesis of that, of all orthogroups under positive selection, an equal proportion of orthogroups will be under positive selection in foreground and background species (R Core Team 2019).

RELAX Downstream Analyses

RELAX assesses evidence for shifts in selection intensity in association with a phenotype. Here, selection intensity refers to the strength of both positive and purifying selection (Wertheim et al. 2015). We wanted to understand whether our focal traits are associated with particular shifts in selection intensity, so we implemented Benjamini–Hochberg false discovery rate corrections on the *P* values from RELAX and subsequently categorized orthogroups as under intensified or relaxed selection in association with each focal trait. We tested for differences in the proportion of orthogroups in each category using chi-squared goodness-of-fit tests, fit to the hypothesis that, of orthogroups with a shift in selection intensity, the proportion of orthogroups in each category is equal (R Core Team 2019).

Assessing GO Term Enrichment

Annotating Sequences to Assign Gene Function

We searched the sequence of all genes from each genome against the Pfam database using InterProScan with the flags –goterms -appl Pfam -f TSV and provided the output to KinFin in order to generate a putative functional annotation for each orthogroup (Bateman et al. 2004, Jones et al. 2014; Laetsch and Blaxter 2017).

GO Term Enrichment of Genes under Selection in Association with Trait Presence or Absence

Using Fisher's exact test, as implemented in the R package topGO, we tested for GO term enrichment in the gene sets under positive selection in species with our focal traits and in species without the focal traits, and under either relaxed or intensified selection in species with the focal traits (Alexa and Rahnenfuhrer 2010). We identified significantly enriched GO terms ($P < 0.01$) in all three ontologies: biological process, cellular component, and molecular function.

Investigating Evolution of Candidate Genes for Focal Traits

Identifying Candidate Genes in the Literature

We conducted a literature search to identify candidate genes related to worker reproduction and worker polymorphism in eusocial insects. We used the database Web of Science, performing search queries with terms relevant to the genetic basis of worker reproduction and worker polymorphism (see supplementary information, Supplementary Material online, for exact search queries). We used the first 25 results for each search, first skimming each paper to determine if it described any potential candidate genes linked to the focal trait. If it did, we identified homologs for the gene(s) in ants, when possible, and when this was not possible, in the study organism of the original paper. Gene sequences were then

retrieved from NCBI, or in a few cases from the original paper. Finally, we used BLAST to identify the ant orthogroups corresponding to each candidate gene, if possible.

Assessing Relationship between Focal Trait and Evolution of Candidate Genes

Once orthogroups corresponding to each candidate gene were identified, we filtered our BUSTED-PH and RELAX results to identify the specific evolutionary mechanisms acting on each candidate gene in ants, in association with the relevant phenotype. We report both raw *P* values and Benjamini–Hochberg false discovery rate corrected *P* values for these results.

Supplementary Material

Supplementary data are available at *Genome Biology and Evolution* online (http://www.gbe.oxfordjournals.org/).

Acknowledgments

We thank Benjamin E.R. Rubin for preliminary assembly and annotation of the *Cephalotes varians* genome. We thank Michael Sheehan and Leslie Babonis for helpful conversations about this study. This work was supported by the National Science Foundation (NSF DEB 1900357 to C.S.M.). This material is based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant No. DGE – 2139899 to M.B. Any opinion, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Data Availability

The *Cephalotes varians* genome data underlying this article are available via the NCBI GenBank Genome database, under BioProject PRJNA864742. All other data are publicly available, with sources noted in supplementary table S1, Supplementary Material online. The code underlying this article is available via Zenodo, at https://doi.org/10.5281/zenodo.5668329.

Literature Cited

Alexa A, Rahnenfuhrer J. 2010. topGO: enrichment analysis for Gene Ontology. R package version. 2.

Araki M, Miyakawa MO, Suzuki T, Miyakawa H. 2020. Two insulin-like peptides may regulate egg production in opposite directions via juvenile hormone signaling in the queenless ant *Pristomyrmex punctatus*. J Exp Zool B Mol Dev Evol. 334:225–234.

Awde DN, Skandalis A, Richards MH. 2020. Vitellogenin expression corresponds with reproductive status and caste in a primitively eusocial bee. J Insect Physiol. 127:104113.

Bartosch-Härlid A, Berlin S, Smith NGC, Møller AP, Ellegren H. 2003. Life history and the male mutation bias. Evolution 57:2398–2406.

Bateman A, et al. 2004. The Pfam protein families database. Nucleic Acids Res. 32:D138–D141.

Blanchard BD, Moreau CS. 2017. Defensive traits exhibit an evolutionary trade-off and drive diversification in ants. Evolution 71:315–328.

Bromham L. 2002. Molecular clocks in reptiles: life history influences rate of molecular evolution. Mol Biol Evol. 19:302–309.

Bromham L, Leys R. 2005. Sociality and the rate of molecular evolution. Mol Biol Evol. 22:1393–1402.

Chen J, et al. 2022. Role of the *foraging* gene in worker behavioral transition in the red imported fire ant, *Solenopsis invicta* (Hymenoptera: Formicidae). Pest Manag Sci. 78:2964–2975.

Cui R, et al. 2021. Ancestral transoceanic colonization and recent population reduction in a nonannual killifish from the Seychelles archipelago. Mol Ecol. 30:3610–3623.

Dallacqua RP, Bitondi MMG. 2014. Dimorphic ovary differentiation in honeybee (*Apis mellifera*) larvae involves caste-specific expression of homologs of *Ark* and *Buffy* cell death genes. PLoS One. 9:e98088.

Drost H-G, Gabel A, Grosse I, Quint M. 2015. Evidence for active maintenance of phylotranscriptomic hourglass patterns in animal and plant embryogenesis. Mol Biol Evol. 32:1221–1231.

Emms DM, Kelly S. 2015. Orthofinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. Genome Biol. 16:157.

Emms DM, Kelly S. 2019. Orthofinder: phylogenetic orthology inference for comparative genomics. Genome Biol. 20:238.

Ferguson-Gow H, Sumner S, Bourke AFG, Jones KE. 2014. Colony size predicts division of labour in attine ants. Proc Biol Sci. 281:20141411.

Fetter-Pruneda I, et al. 2021. An oxytocin/vasopressin-related neuropeptide modulates social foraging behavior in the clonal raider ant. PLoS Biol. 19:e3001305.

Gadau J, et al. 2012. The genomic impact of 100 million years of social evolution in seven ant species. Trends Genet. 28:14–21.

Ghaninia M, et al. 2017. Chemosensory sensitivity reflects reproductive status in the ant *Harpegnathos saltator*. Sci Rep. 7:3732.

Gospocic J, et al. 2017. The neuropeptide corazonin controls social behavior and caste identity in ants. Cell 170:748–759.e12.

Gramates LS, et al. 2022. Flybase: a guided tour of highlighted features. Genetics 220:iyac035.

Gupta T, et al. 2010. Microtubule actin crosslinking factor 1 regulates the Balbiani body and animal-vegetal polarity of the zebrafish oocyte. PLoS Genet. 6:e1001073.

Haas BJ. 2021. https://github.com/TransDecoder/TransDecoder.

Hamilton WD. 1964a. The genetical evolution of social behaviour. I. J Theor Biol. 7:1–16.

Hamilton WD. 1964b. The genetical evolution of social behaviour. II. J Theor Biol. 7:17–52.

Hartmann A, Wantia J, Torres JA, Heinze J. 2003. Worker policing without genetic conflicts in a clonal ant. Proc Natl Acad Sci U S A. 100:12836–12840.

Helanterä H, Uller T. 2014. Neutral and adaptive explanations for an association between caste-biased gene expression and rate of sequence evolution. Front Genet. 5:297.

Helms Cahan S, Keller L. 2003. Complex hybrid origin of genetic caste determination in harvester ants. Nature 424:306–309.

Hölldobler B, Wilson EO. 1990. *The ants*. London: Harvard University Press.

Imrit MA, Dogantzis KA, Harpur BA, Zayed A. 2020. Eusociality influences the strength of negative selection on insect genomes. Proc Biol Sci. 287:20201512.

Jones P, et al. 2014. Interproscan 5: genome-scale protein function classification. Bioinformatics 30:1236–1240.

Jongepier E, et al. 2018. Remodeling of the juvenile hormone pathway through caste-biased gene expression and positive selection along a gradient of termite eusociality. J Exp Zool B Mol Dev Evol. 330:296–304.
Jónsson H, et al. 2017. Parental influence on human germline de novo mutations in 1,548 trios from Iceland. Nature 549:519–522.
Julian GE, Fewell JH, Gadau J, Johnson RA, Larrabee D. 2002. Genetic determination of the queen caste in an ant hybrid zone. Proc Natl Acad Sci U S A. 99:8157–8160.
Kapheim KM. 2019. Synthesis of Tinbergen's four questions and the future of sociogenomics. Behav Ecol Sociobiol. 73:186.
Katoh K, Misawa K, Kuma K, Miyata T. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. 30:3059–3066.
Khila A, Abouheif E. 2008. Reproductive constraint is a developmental mechanism that maintains social harmony in advanced ant societies. Proc Natl Acad Sci U S A. 105:17884–17889.
Khila A, Abouheif E. 2010. Evaluating the role of reproductive constraints in ant social evolution. Philos Trans R Soc Lond B Biol Sci. 365:617–630.
Kimura M. 1968. Evolutionary rate at the molecular level. Nature 217: 624–626.
Kimura M. 1979. The neutral theory of molecular evolution. Sci Am. 241:98–100, 102, 108 passim.
Korb J, et al. 2021. Comparative transcriptomic analysis of the mechanisms underpinning ageing and fecundity in social insects. Philos Trans R Soc Lond B Biol Sci. 376:20190728.
Kosakovsky Pond SL, et al. 2020. Hyphy 2.5-A customizable platform for evolutionary hypothesis testing using phylogenies. Mol Biol Evol. 37:295–299.
Krishnakumar P, et al. 2018. Functional equivalence of germ plasm organizers. PLoS Genet. 14:e1007696.
Laetsch DR, Blaxter ML. 2017. Kinfin: software for taxon-aware analysis of clustered protein sequences. G3 (Bethesda). 7: 3349–3357.
Leichty AR, Pfennig DW, Jones CD, Pfennig KS. 2012. Relaxed genetic constraint is ancestral to the evolution of phenotypic plasticity. Integr Comp Biol. 52:16–30.
Liebig J, Peeters C, Hölldobler B. 1999. Worker policing limits the number of reproductives in a ponerine ant. Proc R Soc Lond B Biol Sci. 266:1865–1870.
Linksvayer TA, Wade MJ. 2009. Genes with social effects are expected to harbor more sequence variation within and between species. Evolution 63:1685–1696.
Linksvayer TA, Wade MJ. 2016. Theoretical predictions for sociogenomic data: the effects of kin selection and sex-limited expression on the evolution of social insect genomes. Front Ecol Evol. 4:65.
Lozano J, Kayukawa T, Shinoda T, Belles X. 2014. A role for Taiman in insect metamorphosis. PLoS Genet. 10:e1004769.
Lucas C, Sokolowski MB. 2009. Molecular basis for changes in behavioral state in ant social behaviors. Proc Natl Acad Sci U S A. 106: 6351–6356.
Mank JE, Nam K, Ellegren H. 2010. Faster-Z evolution is predominantly due to genetic drift. Mol Biol Evol. 27:661–670.
Michener CD. 1969. Comparative social behavior of bees. Annu Rev Entomol. 14:299–342.
Milgroom MG. 1996. Recombination and the multilocus structure of fungal populations. Annu Rev Phytopathol. 34:457–477.
Monnin T, Peeters C. 1999. Dominance hierarchy and reproductive conflicts among subordinates in a monogynous queenless ant. Behav Ecol. 10:323–332.
Morandin C, et al. 2014. Not only for egg yolk–functional and evolutionary insights from expression, selection, and structural analyses of *Formica* ant vitellogenins. Mol Biol Evol. 31:2181–2193.
Morandin C, Hietala A, Helanterä H. 2019. Vitellogenin and vitellogenin-like gene expression patterns in relation to caste and task in the ant *Formica fusca*. Insectes Soc. 66:519–531.
Moreau CS. 2014. A practical guide to DNA extraction, PCR, and gene-based DNA sequencing in insects. Halteres 5:32–42.
Moreau CS, Bell CD, Vila R, Archibald SB, Pierce NE. 2006. Phylogeny of the ants: diversification in the age of angiosperms. Science 312: 101–104.
Murrell B, et al. 2015. Gene-wide identification of episodic selection. Mol Biol Evol. 32:1365–1371.
Ohta T. 1992. The nearly neutral theory of molecular evolution. Annu Rev Ecol Syst. 23:263–286.
Orive ME. 1993. Effective population size in organisms with complex life-histories. Theor Popul Biol. 44:316–340.
Price MN, Dehal PS, Arkin AP. 2010. Fasttree 2—approximately maximum-likelihood trees for large alignments. PLoS One. 5:e9490.
Primmer CR, Papakostas S, Leder EH, Davis MJ, Ragan MA. 2013. Annotated genes and nonannotated genomes: cross-species use of Gene Ontology in ecology and evolution research. Mol Ecol. 22:3216–3241.
Rajakumar R, et al. 2012. Ancestral developmental potential facilitates parallel evolution in ants. Science 335:79–82.
Rajakumar R, et al. 2018. Social regulation of a rudimentary organ generates complex worker-caste systems in ants. Nature 562: 574–577.
Rehan SM, Toth AL. 2015. Climbing the social ladder: the molecular evolution of sociality. Trends Ecol Evol. 30:426–433.
Romiguier J, et al. 2014. Population genomics of eusocial insects: the costs of a vertebrate-like effective population size. J Evol Biol. 27: 593–603.
Romiguier J, et al. 2022. Ant phylogenomics reveals a natural selection hotspot preceding the origin of complex eusociality. Curr Biol. 32: 2942–2947.e4.
Ronai I, et al. 2016. Anarchy is a molecular signature of worker sterility in the honey bee. Mol Biol Evol. 33:134–142.
Roux J, et al. 2014. Patterns of positive selection in seven ant genomes. Mol Biol Evol. 31:1661–1685.
Rubin BER. 2022. Social insect colony size is correlated with rates of molecular evolution. Insectes Soc. 69:147–157
Santos PKF, Galbraith DA, Starkey J, Amsalem E. 2022. The effect of the brood and the queen on early gene expression in bumble bee workers' brains. Sci Rep. 12:3018.
Schrader L, Helanterä H, Oettler J. 2017. Accelerated evolution of developmentally biased genes in the tetraphenic ant *Cardiocondyla obscurior*. Mol Biol Evol. 34:535–544.
Shpigler H, et al. 2010. The transcription factor Krüppel homolog 1 is linked to hormone mediated social organization in bees. BMC Evol Biol. 10:120.
Simola DF, et al. 2013. Social insect genomes exhibit dramatic evolution in gene composition and regulation while preserving regulatory features linked to sociality. Genome Res. 23:1235–1247.
Smith F. 1876. Descriptions of new species of Cryptoceridae, belonging to the genera *Cryptocerus*, *Meranoplus*, and *Cataulacus*. Trans Entomol Soc Lond. 1876:603–612.
Snell-Rood EC, Van Dyken JD, Cruickshank T, Wade MJ, Moczek AP. 2010. Toward a population genetic framework of developmental evolution: the costs, limits, and consequences of phenotypic plasticity. Bioessays 32:71–81.
Star B, Spencer HG. 2013. Effects of genetic drift and gene flow on the selective maintenance of genetic variation. Genetics 194:235–244.
Suyama M, Torrents D, Bork P. 2006. PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. Nucleic Acids Res. 34:W609–W612.

Tarès S, Arthaud L, Amichot M, Robichon A. 2013. Environment exploration and colonization behavior of the pea aphid associated with the expression of the foraging gene. PLoS One. 8:e65104.
Thomas JA, Welch JJ, Lanfear R, Bromham L. 2010. A generation time effect on the rate of molecular evolution in invertebrates. Mol Biol Evol. 27:1173–1180.
Tobback J, et al. 2013. Developmental- and food-dependent *foraging* transcript levels in the desert locust. Insect Sci. 20:679–688.
Toth AL, Robinson GE. 2007. Evo-devo and the evolution of social behavior. Trends Genet. 23:334–341.
Van Dyken JD, Wade MJ. 2010. The genetic signature of conditional expression. Genetics 184:557–570.
Wang RJ, et al. 2020. Paternal age in rhesus macaques is positively associated with germline mutation accumulation but not with measures of offspring sociability. Genome Res. 30:826–834.
Warner MR, Mikheyev AS, Linksvayer TA. 2017. Genomic signature of kin selection in an ant with obligately sterile workers. Mol Biol Evol. 34:1780–1787.
Wertheim JO, Murrell B, Smith MD, Kosakovsky Pond SL, Scheffler K. 2015. RELAX: detecting relaxed selection in a phylogenetic framework. Mol Biol Evol. 32:820–832.
West-Eberhard MJ. 2005. Phenotypic accommodation: adaptive innovation due to developmental plasticity. J Exp Zool B Mol Dev Evol. 304:610–618.
Wills BD, Powell S, Rivera MD, Suarez AV. 2018. Correlates and consequences of worker polymorphism in ants. Annu Rev Entomol. 63: 575–598.
Wilson EO. 1971. *The insect societies*. Cambridge, MA: Belknap Press of Harvard University Press.
Wilson EO, Hölldobler B. 2005. Eusociality: origin and consequences. Proc Natl Acad Sci U S A. 102:13367–13371.
Wisotsky SR, Kosakovsky Pond SL, Shank SD, Muse SV. 2020. Synonymous site-to-site substitution rate variation dramatically inflates false positive rates of selection analyses: ignore at your own peril. Mol Biol Evol. 37:2430–2439.
Zhang J-J, Xi G-S, Zhao J. 2021. Vitellogenin regulates estrogen-related receptor expression by crosstalk with the JH and IIS-TOR signaling pathway in *Polyrhachis vicina* Roger (Hymenoptera, Formicidae). Gen Comp Endocrinol. 310:113836.
Zhao F, et al. 2021. Molecular evolution of bumble bee vitellogenin and vitellogenin-like genes. Ecol Evol. 11:8983–8992.

Associate editor: Matthew Webster