# Evolution of Gene Expression in Fire Ants: The Effects of Developmental Stage, Caste, and Species

Lino Ometto,*†,[1] DeWayne Shoemaker,[2] Kenneth G. Ross,[3] and Laurent Keller[1]

[1]Department of Ecology and Evolution, University of Lausanne, Biophore, Lausanne, Switzerland

[2]Center for Medical, Agricultural, and Veterinary Entomology, United States Department of Agriculture—Agricultural Research Service

[3]Department of Entomology, University of Georgia

†Present address: Department of Biodiversity and Molecular Ecology, Research and Innovation Centre, Fondazione Edmund Mach, San Michele all'Adige (TN), Italy

*Corresponding author: E-mail: lino.ometto@iasma.it

Associate editor: Douglas Crawford

## Abstract

Ants provide remarkable examples of equivalent genotypes developing into divergent and discrete phenotypes. Diploid eggs can develop either into queens, which specialize in reproduction, or workers, which participate in cooperative tasks such as building the nest, collecting food, and rearing the young. In contrast, the differentiation between males and females generally depends upon whether eggs are fertilized, with fertilized (diploid) eggs giving rise to females and unfertilized (haploid) eggs giving rise to males. To obtain a comprehensive picture of the relative contributions of gender (sex), caste, developmental stage, and species divergence to gene expression evolution, we investigated gene expression patterns in pupal and adult queens, workers, and males of two species of fire ants, *Solenopsis invicta* and *S. richteri*. Microarray hybridizations revealed that variation in gene expression profiles is influenced more by developmental stage than by caste membership, sex, or species identity. The second major contributor to variation in gene expression was the combination of sex and caste. Although workers and queens share equivalent diploid nuclear genomes, they have highly distinctive patterns of gene expression in both the pupal and the adult stages, as might be expected given their extraordinary level of phenotypic differentiation. Overall, the difference in the proportion of differentially expressed genes was greater between workers and males than between workers and queens or queens and males, consistent with the fact that workers and males share neither gender nor reproductive capability. Moreover, between-species comparisons revealed that the greatest difference in gene expression patterns occurred in adult workers, a finding consistent with the fact that adult workers most directly experience the distinct external environments characterizing the different habitats occupied by the two species. Thus, much of the evolution of gene expression in ants may occur in the worker caste, despite the fact that these individuals are largely or completely sterile. Analyses of gene expression evolution revealed a combination of positive selection and relaxation of stabilizing selection as important factors driving the evolution of such genes.

Key words: gene expression, ants, *Solenopsis invicta*, *Solenopsis richteri*, caste, evolution.

## Introduction

Morphological, physiological, and behavioral differences can arise from sequence differences in coding genes or from differences in the level of expression of genes. Examples of the effect of variation in gene expression include tissue differentiation within individuals (e.g., Enard et al. 2002; Khaitovich et al. 2006; Yang et al. 2006), sexual dimorphism in organisms with otherwise almost equivalent genomes (Ellegren and Parsch 2007), and many of the phenotypic differences between closely related species (Enard et al. 2002; Ranz et al. 2003; Rifkin et al. 2003; Cassone et al. 2008). Several studies have demonstrated that such variation in gene expression is an important substrate for adaptive evolution (Ortiz-Barrientos et al. 2006; Whitehead and Crawford 2006a, 2006b; Rebeiz et al. 2009; Babbitt et al. 2010).

An important question is to determine the relative contribution of differences in coding gene sequences and level of gene expression on phenotypic evolution. Both environmental differences and genetic variation (e.g., in transcription factors and regulatory motifs) can translate into gene expression and phenotypic differences. An effective way to investigate the environmental contributions to gene expression variation is to study organisms in which divergent and discrete phenotypes (polyphenism) can develop from an identical genotype. Environmentally induced differences in gene expression associated with polyphenism have been reported in horned beetles (Snell-Rood et al. 2010), termites (Scharf et al. 2003; Weil et al. 2009), honey bees (Evans and Wheeler 2001; Judice et al. 2006; Sen Sarma et al. 2007), bumble bees (Pereboom et al. 2005), wasps (Sumner et al. 2006; Hoffman and Goodisman 2007; Hunt and Goodisman 2010), and ants (Gräff et al. 2007). These studies found evidence that differential gene expression patterns mediate the decoupling between an otherwise identical genotype and the resulting distinct alternative phenotypes, but much remains to be learned about the

mechanisms, causes, and evolution underlying variation in gene expression within- and between-species.

The aim of this study is to investigate gene expression patterns in queens, workers, and males of two closely related ant species to obtain a comprehensive picture of the relative contributions of sex, caste, and species divergence to gene expression evolution. The organizing principle of ant societies is reproductive division of labor, whereby one or a few female individuals (the queens) specialize in reproduction, whereas the others (the workers) participate in cooperative tasks such as building the nest, collecting food, rearing the young, and defending the colony. This reproductive division of labor provides numerous fitness benefits and is one reason for the tremendous ecological success of ants (Hölldobler and Wilson 1990). With the exception of a few atypical lineages (Julian et al. 2002; Volny and Gordon 2002; Helms Cahan and Keller 2003; Fournier et al. 2005), the marked behavioral and morphological differences between queen and worker ants stem from differences in their developmental pathways unrelated to genotype, as any fertilized (diploid) egg has the potential to develop into either caste depending upon environment-induced differentiation in gene expression (reviewed in Schwander et al. 2010). In contrast, the differentiation between males and females generally depends upon whether eggs are fertilized, with fertilized eggs giving rise to females and unfertilized (haploid) eggs giving rise to males.

We studied the relative contribution of sex, caste, and species divergence to gene expression in two species of fire ants, Solenopsis invicta and S. richteri. Our choice of these two species was based on several criteria. First, fire ants are advanced eusocial insects in which the extraordinary degree of caste dimorphism between females is as extreme as found among social animals (including obligate worker sterility enforced by anatomical absence of reproductive organs). Caste differentiation within Solenopsis occurs between the first and second larval instars (Petralia and Vinson 1979), and the observation that the loss of a queen induces many diploid larvae to switch their development from workers to queens is consistent with an environmental caste determination system (Vargo and Fletcher 1986). The evident lack of a genotypic mechanism for caste determination therefore makes Solenopsis an ideal system to investigate gene expression changes associated with radical social evolution. Second, S. invicta and S. richteri are closely related (Pitts et al. 2005), being able to hybridize and even form relatively fit hybrids in nature (Ross et al. 1987; Ross and Robertson 1990; Shoemaker et al. 1996). Third, a microarray containing 21,715 expressed sequence tags (ESTs) representing more than 10,000 different transcripts with diverse molecular functions was recently developed for S. invicta (Wang et al. 2007; Wurm et al. 2009; Wurm et al. 2010).

In addition to studying the contributions of sex, caste, and species divergence to gene expression differences, we analyzed gene expression variation between the pupal and the adult stages. Fire ants and other Hymenoptera are advanced holometabolous insects with extraordinary phenotypic differences among the life stages. Thus, patterns of gene expression are expected to differ substantially from one developmental stage to the next. Finally, we also evaluated the role of selection in driving the evolution of caste-specific and species-specific differences in gene expression.

## Materials and Methods

### Samples

Colonies of S. invicta were collected in Oglethorpe County, Georgia in June 2005 ($n = 3$) and May 2007 ($n = 5$). Colonies of S. richteri were collected in Henderson County, Tennessee in June 2005 ($n = 2$) and May 2007 ($n = 5$). All collected colonies were returned to the laboratory and maintained in a rearing room under standard conditions (Jouvenaz et al. 1977). Allozyme electrophoretic analyses were performed on a subset of ten females from each colony to confirm that all colonies were either pure S. invicta or pure S. richteri (the two species hybridize in their invasive, USA ranges). We also performed allozyme analyses of 30 additional colonies collected from each of these same sites to confirm that each area contained only the respective pure parental species. Five highly informative or fully diagnostic loci were employed in these analyses. The same subset of ten females per study colony was used to confirm by allozyme electrophoresis that all colonies were of the monogyne social form (single queen per colony), a conclusion supported also by the discovery of only a single wingless, reproductive queen per nest (see Ross et al. 2003 for explanation of expected genotype distributions in colonies of the two forms). Three (S. invicta) or two (S. richteri) polymorphic loci were used for these latter analyses. Relevant allozyme methods are detailed in Shoemaker et al. (1996) and Ross et al. (1997).

The samples comprise three castes (workers, queens, and males) collected at two separate developmental stages (pupae and adults) for both species. Only large (major) workers were selected for the worker samples. We made every effort to minimize environmental noise in gene expression by sampling all study individuals from laboratory colonies reared under standard conditions for at least two weeks after field collection. Also, samples were harvested at specific developmental mileposts to minimize temporal variation in gene expression within each of the two developmental stages: all pupae were sampled at the first appearance of "pink eyes," typically within 24 h of pupation, and adults were sampled within 1 h of emergence from the pupal stage. These timepoints were chosen because they make it possible to control most reliably for age and developmental state, variation in which would introduce unwanted noise in our experimental design.

### RNA/DNA Extraction and Preparation

We randomly sampled, and eventually pooled, 1–6 nest mate individuals (average = 4.6 and median = 5) for each

replicate of the 12 specific categories or nodes of interest (two developmental stages $\times$ 3 castes $\times$ 2 species; supplementary table S1, Supplementary Material online). Each pooled sample was promptly snap-frozen in liquid nitrogen in tubes containing 1 g of 1.4-mm ceramic beads (Quackenbush) and then homogenized in Trizol (Invitrogen) using a Retsch bead shaker. Both RNA and DNA were extracted according to the manufacturer's protocol. The quality of the total RNA was assessed visually by gel electrophoresis of a 2 µl aliquot in a 1.5% agarose gel. Only samples with no signs of RNA degradation were retained and used for subsequent microarray analyses.

Total RNA was amplified (Ambion MessageAmp II kit) and then an ~6-µg aliquot of this amplified RNA was mixed with 2 µl of random 9mers (2 µg/µl), 0.5 µl of Alien mRNA Spike mix (Stratagene), and water for a final volume of 17.9 µl. The mix was incubated for 10 min at 70 °C, then held for 5 min on ice. Reverse transcription was performed for 2 h at 50 °C after adding 6 µl of 5 × first-strand buffer, 3 µl of 0.1 M dithiothreitol, 0.6 µl of 50X aminoallyl-deoxynucleotide triphosphate mix (25 mM deoxyadenosine triphosphate, 25 mM deoxycytidine triphosphate, 25 mM deoxyguanosine triphosphate, 15 mM deoxythymidine triphosphate, and 10 mM aminoallyl-deoxyuridine triphosphate), 0.5 µl of RNAse inhibitor (15 U/µl, Invitrogen), and 2 µl of SuperScript III reverse transcriptase (200 U/ml, Invitrogen). The residual RNA was hydrolyzed by adding 15 µl of 0.1 M NaOH and incubating for 10 min at 70 °C. Subsequently, the pH was returned to neutrality by the addition of 15 µl of 0.1 M HCl. The purification of the aminoallyl-labeled cDNA and the coupling to Cy3 or Cy5 dyes (GE Healthcare, Little Chalfont, UK) were done using a modified Qiaquick PCR Purification Kit (QIAGEN) protocol. The combined Cy3- and Cy5-labeled probes were purified using the Qiaquick PCR Purification Kit (QIAGEN) and eluted in 82 µl of elution buffer. We finally added 16.5 µl of 20x standard saline citrate (SSC), 3.0 µl of yeast transfer RNA (10 µg/µl), 1.8 µl of polyA cDNA (2 µg/µl; Sigma), and 1.8 µl of 10% sodium dodecyl sulfate (SDS), then denatured this reaction mix at 95 °C for 45 s and hybridized to the microarray slides at 64 °C overnight (18–22 h). Excess probe was removed by washing the slides at room temperature twice for 5 min in a 2x SSC/0.1% SDS solution, twice for 1 min in a 0.2x SSC solution, 1 min in a 0.1x SSC solution, and twice for 5 min in a 0.1x SSC/0.1% Triton solution.

## Microarray Experiments

Each experimental sample was hybridized against a common reference RNA (which consists of RNA extracted from hundreds of *S. invicta* individuals of different castes [i.e., workers, queens, and males] and of different stages [i.e., adults, pupae, and larvae] to ensure maximum coverage of the transcriptome) labeled with the appropriate alternate dye on our custom-made spotted cDNA microarray (Wang et al. 2007; Wurm et al. 2009). Every sample was equivalent to an independent biological replicate (i.e., each colony was used only once as a source for a given caste and stage; see supplementary table S1, Supplementary Material

online) that was randomly labeled with either Cy3 or Cy5. After hybridization, slides were scanned using an Agilent Technologies Scanner and the signal intensities for each spot were extracted (low-quality spots were flagged and removed prior to analyses) using GenePix software.

## Microarray Analyses

We used the software package limma in R (R Development Core Team 2008) to perform adaptive background correction, within-array print-tip loess normalization and between-array normalization of the intensities (Smyth and Speed 2003; Smyth 2004; Ritchie et al. 2007). To increase the power of the between-array normalization, we included data from (1) 28 additional arrays hybridized with *S. invicta* and *S. richteri* cDNA from worker, queen, or male larvae (excluded from the current study due to insufficient sampling) and (2) 46 arrays hybridized with hybrid *S. invicta* X *richteri* cDNA from various life stages and castes (Ometto L, Ross KG, Shoemaker D and Keller L, in preparation), all of which underwent the same experimental protocols as the samples used in the present study. Stringent quality control criteria were used to filter out spots with unreliable foreground intensity for each microarray slide (modified from Lemos et al. 2008). The criteria were (((F635 Median > CTRL635 + 2 × CTRL635SD) OR (F532 Median > CTRL532 + 2 × CTRL532SD)) AND (((F635 Median-B635) > 2 × B635 SD) OR ((F532 Median-B532) > 2 × B532 SD)) AND ((%>B635 + 2SD > 70) OR (%>B532 + 2SD > 70)) AND ((SNR635 > 2) AND (SNR532 > 2))), where $F_{xxx}$ and $B_{xxx}$ are the foreground and background fluorescence intensities, $SNR_{xxx}$ is the signal to noise ratio, %>$B_{xxx}$+2SD is the percentage of pixels above background mean + 2 × standard deviations (SD), $CTRL_{xxx}$ and $CTRL_{xxx}$SD are the average and SD of the foreground intensities for eight negative control probes spotted onto the microarray (Human beta actin PCR product, Stratagene), and $xxx = 532$ and 635 for Cy3 and Cy5, respectively.

We estimated the relative expression of each clone for each node (caste/species/stage), and the significance of the differential expression of such clones among nodes, using the Bayesian approach implemented in the program Bayesian Analysis of Gene Expression Levels (BAGEL; Meiklejohn and Townsend 2006). To increase the power of the BAGEL estimations, we also included in the analysis the additional 28 and 46 arrays detailed above. A false discovery rate (FDR) was estimated using BAGEL by analyzing a randomized version of the original data set and comparing the number of positives obtained from the randomized and observed data sets. Based on these analyses, a clone was considered to be differentially expressed between two nodes if the associated Bayesian posterior probability was $P < 0.001$, which represents a value with an acceptably low FDR (i.e., FDR $\approx$ 5% when comparing *S. invicta* castes, and FDR $\approx$ 10% when comparing species; supplementary table S2, Supplementary Material online). We define caste-biased genes as those that were significantly over or underexpressed in one caste compared with both of the other castes for a given

developmental stage (supplementary fig. S1, Supplementary Material online). Separate analyses were performed for *S. invicta* and *S. richteri*.

The analysis of gene expression involved multiple samples from two separate species, thus necessitating two main caveats. The first is the possibility of unequal power in detecting differentially expressed genes across comparisons, for example, because of unequal sample sizes. A thorough statistical analysis revealed, however, no substantial difference between the power estimated for our comparisons (Materials and Methods, Supplementary Material online). The second relates to the use of a microarray developed for *S. invicta* to assess gene expression in both *S. invicta* and *S. richteri*. A study by Ranz et al. 2003 showed that the use of a microarray developed for *Drosophila melanogaster* to compare gene expression of *D. melanogaster* and *D simulans* did not lead to biases in the estimation of gene expression levels. Given that the overall nuclear genomic divergence between these two fly species appears greater than that between *S. invicta* and *S. richteri* (Ometto L and Shoemaker D, unpublished data), our microarray analyses likely are not confounded by spurious results due to interspecific sequence divergence. Furthermore, Gilad et al. (2005) showed that single- and multispecies arrays performed equally well when imposing a cutoff in gene expression differences of 1.5-fold, close to the GEL50 (ratio of gene expression level beyond which 50% of differences were significant at $P < 0.001$) and well below the GEL95 ratio thresholds (supplementary table S3, Supplementary Material online). In our analyses, we obtained equivalent patterns of gene expression differences when the GEL95 ratio threshold or the threshold of 1 was used (supplementary table S4, Supplementary Material online). This suggests a negligible effect of the possible sequence mismatches due to interspecific divergence on the levels of microarray hybridization. However, to investigate the possibility that such sequence divergence could result in reduced microarray hybridization, and hence observed expression levels, we calculated the overall level of expression for each caste and species (supplementary table S5, Supplementary Material online). The lack of a consistent difference in expression levels between the species suggests that sequence divergence has negligible effects on levels of microarray hybridization. Finally, because our analyses focus mainly on between-caste, between-sex, and between-stage differences, they should be little affected by interspecific differences in levels of microarray hybridization.

We analyzed our results considering both individual clones and contigs because many contigs are in fact represented by more than one clone spotted onto the microarray (Wang et al. 2007; Wurm et al. 2009; Wurm et al. 2010). In these cases, if one clone for a given contig was differentially expressed (or caste-biased), then that particular contig was also considered differentially expressed or caste-biased. Because the clone-based and contig-based results are very similar (supplementary table S4, Supplementary Material online), we report only results based on contigs. This also makes intuitive sense if we assume that

any one gene is likely represented by a single contig but perhaps by multiple clones. Thus, we use contig synonymously with gene throughout the remainder of the text.

## Analysis of Determinants of Gene Expression Variation

We estimated pairwise gene expression distances between each of the twelve nodes by summing the squared differences between the relative expression values (estimated by BAGEL) across all clones. The resulting pairwise distance matrix was used to construct a distance tree using the neighbor joining method (PHYLIP v3.67; Felsenstein 2005). The resulting tree reflects the similarities among nodes in overall gene expression patterns. We also performed principal components analyses on the gene expression data using R (R Development Core Team 2008) to evaluate the relative importance of caste, developmental stage, and species status on the observed gene expression levels across all genes.

## Analysis of Caste-Biased Gene Expression Evolution

The mode of evolution of gene expression was assessed by computing the ratio $R_X$, which is the variability in gene expression between species over the variability in gene expression within *S. invicta*. The variability in gene expression between species for a contig represented by $n$ clones on the microarray was calculated as $D_X = \left( \sum_{c=1}^{n} \left| \log_2 \frac{X_c^{inv}}{X_c^{ric}} \right| \right) / n$, where $X_c^{inv}$ and $X_c^{ric}$ are the relative expression values estimated for *S. invicta* and *S. richteri*, respectively, by BAGEL for clone $c$ in the developmental stage and caste of interest. Similarly, variability in gene expression within *S. invicta* was estimated as $V_X = \left( \sum_{c=1}^{n} \frac{f * SD_c}{Mean_c} \right) / n$, where the mean and the standard deviation of the $\log_2$ ratios between the normalized dye intensities of the experimental sample and reference are calculated across all *S. invicta* samples for the caste and developmental stage of interest. In order to correct for the unequal sample sizes across nodes, we multiplied $SD_c$ by $f = \sqrt{\frac{mN-1}{mN-m}}$, where $m$ is the sample size and $N = 6$ (for convenience, we normalized to the most common sample size). It follows that $R_X = \frac{D_X}{V_X}$.

Our data set contains a mixture of mitochondrial and nuclear genes, as well as transposable elements. Because genes of the different classes likely experience distinct mutational and selective regimes, we limited our analyses to nuclear genes. To this end, we manually annotated 3,316 of the transcripts with a valid BLASTx hit (Nicolas M, Wang J, Ometto L, unpublished data and personal communication) by verifying the presence of functional domains using InterPro (Hunter et al. 2009). We then assigned these genes to a functional category based on the annotation, distinguishing between nuclear and mitochondrial genes and identifying putative viral and transposable elements.

Values of $V_X$, $D_X$, and $R_X$ were estimated separately for each nuclear gene. All values were then partitioned
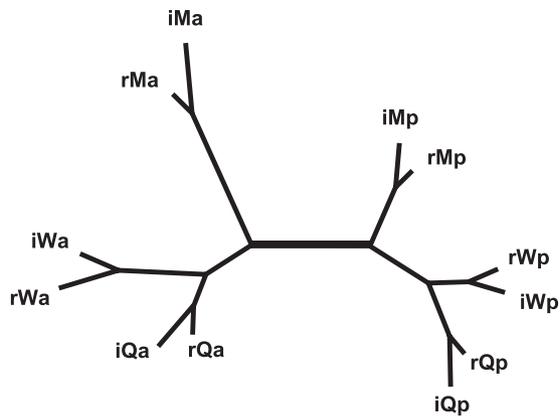
**Fig 1.** Gene expression–based distance tree. Three-letter labels identify species (i = *S. invicta*; r = *S. richteri*), caste (W = workers; Q = queens; M = males), and developmental stage (p = pupae; a = adults).

according to whether the gene was overexpressed in workers, queens, or males, or, alternatively, was not biased in its expression in any caste. We also partitioned the data based on whether or not the gene was differentially expressed between species in a given caste or developmental stage. We then compared average $V_X$, $D_X$, and $R_X$ for these groups of genes using the nonparametric Wilcoxon test to infer possible variation or divergence in the rate of evolution of gene expression.

We also tested for a correlation between the caste specificity of a gene's expression and its variation in expression ($V_X$ and $D_X$). As a proxy for caste specificity, we used the degree of caste overexpression, which, for a gene represented by $n$ clones on the microarray, we estimated as

$$cbx_{caste1} = \left( \sum_{c=1}^{n} \frac{X_c^{caste1}}{X_c^{caste2}} + \sum_{c=1}^{n} \frac{X_c^{caste1}}{X_c^{caste3}} \right)/2$$

where $X_c^{caste}$ is the relative expression value estimated for *S. invicta* by BAGEL for clone $c$ and the superscript refers to the three castes (caste1 represents the caste for which we want to estimate $cbx$). We assume that the larger the value of $cbx$ for a particular gene, the closer that gene is to being wholly caste-specific.

## Results

We performed microarray hybridization experiments on 66 samples representing all three castes and two developmental stages in two fire ant species (supplementary table S1, Supplementary Material online). After discarding microarray spots of poor quality, 22,855 clones remained for subsequent analyses, which correspond roughly to 14,467 putatively unique transcripts or genes.

A gene expression–based distance tree shows that developmental stage was the major determinant of overall gene expression profile, followed by sex, female caste, then species identity (fig. 1). These results were confirmed by principal components analyses: the first component (PC1), on which developmental stage loads heavily, explains more than 40% of the variability in gene expression across samples (fig. 2; supplementary fig. S2, Supplementary

Material online). The second and third components, most strongly related to sex and female caste, respectively, together account for ~34% of the variability in gene expression. Finally, PC4 and PC5, both associated with species, together account for only ~11% of the variability in gene expression.

### Caste-biased genes

In accord with the results of the expression-based distance tree and principal components analyses, Bayesian analyses of gene expression levels revealed that a substantial proportion of genes were significantly overexpressed in one caste compared with the other two at a given developmental stage. The percentage of genes overexpressed in one caste was 22% at the pupal stage in *S. invicta*, 18% at the adult stage in *S. invicta*, 16% at the pupal stage in *S. richteri*, and 14% at the adult stage in *S. richteri* (supplementary table S6 and supplementary fig. S1, Supplementary Material online; fig. 3). In pupae of both species, the percentage of overexpressed genes was highest in males, intermediate in workers, and lowest in queens ($\chi^2$ pairwise goodness of fit tests, all comparisons $P < 10^{-10}$). In adults, the difference was even stronger between queens and the two other castes, whereas there was no statistical difference between males and workers (for both species, $P < 10^{-10}$ for differences between queens and males and between queens and workers, and $P > 0.15$ for the differences between workers and males).

A very similar pattern was found when considering genes significantly underexpressed within a single caste compared with the other two castes within a stage (fig. 3). In pupae of both species, there were many fewer genes underexpressed in both queens and workers compared with males (all $P < 10^{-10}$; supplementary table S6, Supplementary Material online). In contrast, in adults, there were fewer genes underexpressed in queens than in both males and workers (all $P < 10^{-10}$), but more genes underexpressed in workers than males (all $P < 10^{-10}$).

The combined analysis of caste-biased gene expression in both developmental stages provided evidence for a core set of genes important in defining caste identity across pupal and adult ontogeny. Overall, there were significantly more genes than expected by chance that exhibited either caste overexpression or caste underexpression in both developmental stages (supplementary table S6, Supplementary Material online; cumulative hypergeometric distribution, all $P < 10^{-9}$ except for queen underexpressed genes in *S. richteri*, where $P = 0.047$). Although we only analyzed two developmental stages, this result hints at the presence of genes whose characteristic expression marks each caste throughout development.

Two other lines of evidence further suggest a reasonable amount of conservation in caste-biased gene expression patterns during the divergence of the study species. First, 56% of all genes with caste-biased expression in *S. invicta* were also caste-biased in *S. richteri* (cumulative hypergeometric distribution, $P < 10^{-10}$). Second, as in the single-species analyses, the number of queen overexpressed genes
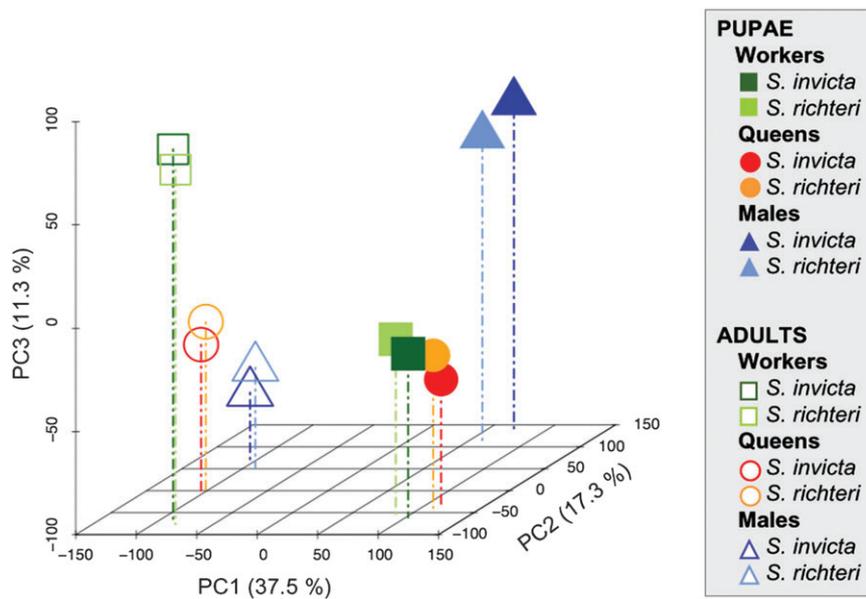
**FIG. 2.** Results of principal components analyses of the gene expression data. Percentages of variability explained by each of the first three components are shown in parentheses.

shared between *S. invicta* and *S. richteri* was significantly lower than the numbers of worker or male overexpressed genes shared between the species (all $P < 10^{-10}$; supplementary table S6, Supplementary Material online).

### Gene Expression Differences Between Species

It is likely that selective pressures acting differently between the two study species vary in intensity among the castes and developmental stages. Therefore, we compared gene expression patterns between *S. invicta* and *S. richteri* separately for each of the three castes and each of the two developmental stages. Overall, 884 genes out of 14,467

(6.1%) were significantly differentially expressed between pupae of the two species in at least one caste (supplementary table S7, Supplementary Material online). The number of such genes was very similar for the queen and worker castes ($\chi^2$ pairwise goodness of fit test, $P = 0.299$), and both values were significantly higher than for males (1.4- and 1.3-fold higher, respectively, $P < 0.001$ in both cases; fig. 4; supplementary table S7, Supplementary Material online). However, a high percentage of such pupa-expressed genes showed a caste-specific pattern of differentiation as they were differentially expressed between the species in only one caste. For instance, of the genes differentially
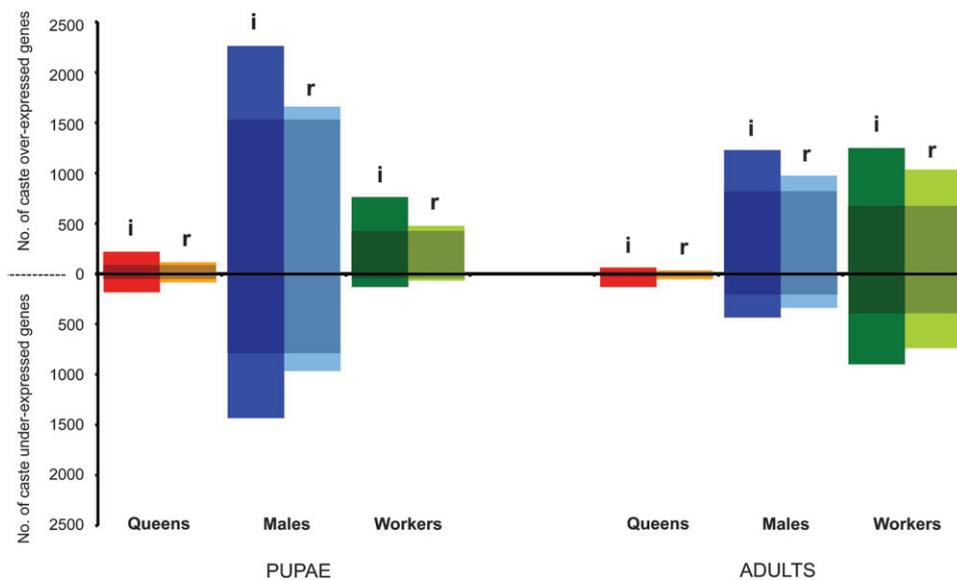


**FIG. 3.** Numbers of genes showing significant overexpression in one caste compared with the other two (upper half of graph) and significant underexpression in one caste compared with the other two (lower half) The darker shaded areas correspond to the number of genes that show the same significant pattern of over or underexpression in *S. invicta* (i) and *S. richteri* (r).
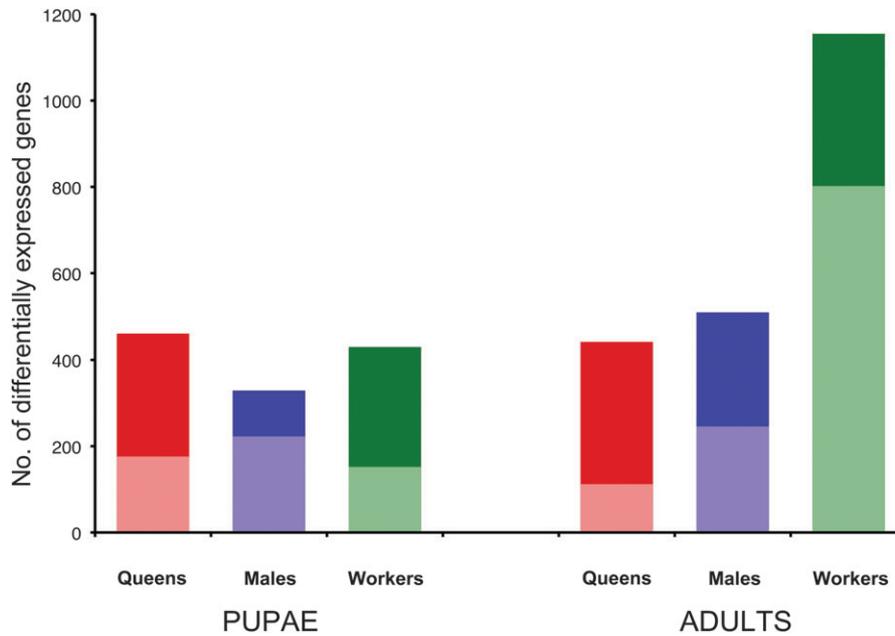
FIG. 4. Genes differentially expressed between *S. invicta* and *S. richteri* in pupae and adults. The light sections of each bar indicate genes that are differentially expressed between species exclusively in one caste. The dark sections indicate genes that are differentially expressed between species in two or all three castes.

expressed between heterospecific pupal workers, 35% were differentially expressed in this caste but not in heterospecific pupal queens or males. The corresponding values were 38% for queen pupae and 68% for male pupae (figs. 4 and 5), which suggests a larger male-specific differentiation in pupal gene expression between species than worker- or queen-specific differentiaton. On the other hand, 43% of the genes differentially expressed between heterospecific pupal workers were also differentially expressed between heterospecific queens (38% for the reciprocal comparison; fig. 5), also suggesting a comparable female-specific differentiation.

The overall number of genes differentially expressed between adults of the two species (1,549, corresponding to 10.7% of the analyzed genes) was higher than for pupae. In contrast to pupae, the difference between species was greatest in workers (2.3- and 2.6-fold for worker-male and worker-queen comparisons, respectively, both $P <$



FIG. 5. Venn diagrams showing the numbers of genes differentially expressed between *S. invicta* and *S. richteri* workers, queens, and males. Overlapping areas indicate differentially expressed genes common to two or three castes. Circle size is proportional to the number of genes.

$10^{-10}$), and there was a marginally significant difference between males and queens ($P = 0.030$). As in pupae, there was a high percentage of genes in adults differentially expressed between the species that showed this pattern in only one caste (fig. 4).

Overall, worker and male overexpressed genes were differentially expressed between species more frequently than expected by chance (supplementary table S7, Supplementary Material online). Notably, about 18% of the worker overexpressed genes in either adult *S. invicta* or *S. richteri* (supplementary tables S6 and S7, Supplementary Material online) were also differentially expressed between adult workers of the two species (cumulative hypergeometric distribution, $P < 10^{-10}$). Similar trends appear for male overexpressed genes (interspecific difference: ~13%, $P < 10^{-10}$) and, to a lesser extent, for queen overexpressed genes (interspecific difference: ~6%, $P = 0.031$) in the adult stage. In pupae, the enrichment in caste overexpressed genes among the species-specific differentially expressed ones is significant only for male genes (~6%, $P < 10^{-10}$). Equivalent patterns are found when analyzing the number of genes that are either caste overexpressed or caste underexpressed but not when analyzing caste underexpressed genes alone. Thus, caste overexpressed genes appear to have differentiated much more than caste underexpressed genes during species divergence.

## Gene Expression Evolution

We estimated both the intraspecific (within *S. invicta*) variation in gene expression, $V_X$, and the interspecific differentiation in gene expression, $D_X$, for all genes. Stabilizing and directional selection are predicted to affect differently the patterns of gene expression variation within and between species. Similar to coding sequences, relatively low $D_X$ and
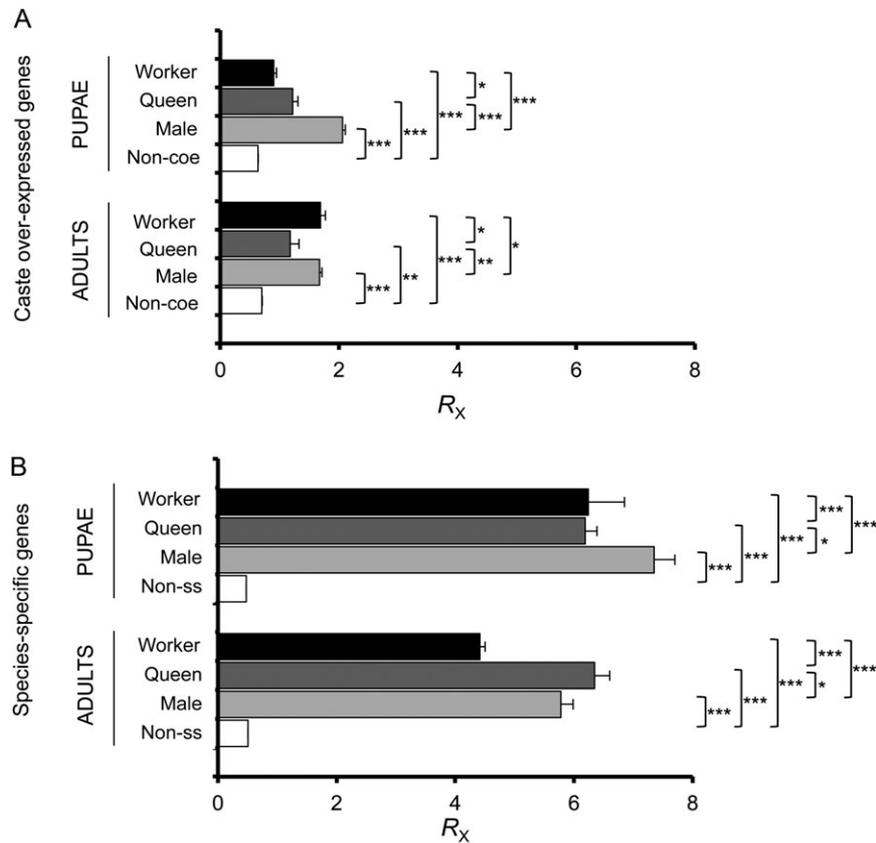
**Fig. 6.** Ratio, $R_X$, of the variability in gene expression between species ($D_X$) to the variability in gene expression within S. invicta ($V_X$) for (A) genes with caste-biased expression and for (B) genes differently expressed between the study species (see supplementary tables S8 and S9, Supplementary Material online and main text for details). Non-coe = noncaste overexpressed; Non-ss = nonspecies-specific. *$P < 0.05$; **$P < 0.001$; ***$P < 10^{-5}$.

$V_X$ values may indicate the presence of strong functional constraints and the action of purifying selection. In contrast, high $D_X$ values may suggest either relaxed selection (when $V_X$ is correspondingly high) or diversifying (positive) selection (when $V_X$ is low). We estimate the ratio $R_X = \frac{D_X}{V_X}$ under the rationale that diversifying selection should lead to a disproportionate increase of $D_X$ compared with $V_X$, as reported in Drosophila for male-biased genes (Meiklejohn et al. 2003). Because neutral expectations for the actual values of $D_X$, $V_X$, and $R_X$ are not necessarily straightforward, we focus instead on the relative differences.

We partitioned all genes into four classes to compare the relative patterns of selection on gene expression in the three castes for each developmental stage: (1) worker overexpressed, (2) queen overexpressed, (3) male overexpressed, and (4) not overexpressed in any caste. In both pupae and adults, the $R_X$ values were significantly higher for the three classes of overexpressed genes than for the genes that were not overexpressed in a given caste ($P < 0.001$ for all castes in both pupae and adults; fig. 6A and supplementary table S8, Supplementary Material online). Furthermore, $V_X$ was lower and $D_X$ was higher for the three classes of genes being overexpressed in one caste than for those that were not (with the exception of $V_X$ in pupae, where worker overexpressed genes had the highest value; supplementary table S8, Supplementary Material online).

This pattern is consistent with caste overexpressed genes being under stronger diversifying (positive) selection between the two lineages and/or being subject to less selective constraints than those genes that are equally expressed across castes (Mank and Ellegren 2009).

In pupae, the average $R_X$ for male overexpressed genes was higher than that for both worker and queen overexpressed genes ($P < 10^{-10}$; fig. 6A). In adults, the average $R_X$ for worker overexpressed genes was slightly larger than that for male overexpressed genes ($P = 0.025$), and both values were greater than for queen overexpressed genes ($P < 0.01$). These results suggest that diversifying selection over both developmental stages plays a more important role in shaping the pattern of expression of worker and, especially, male overexpressed genes than of queen overexpressed genes. The estimates of $V_X$ and $D_X$ further confirm that such positive selection may have been more important for male overexpressed genes than for queen and, perhaps, worker overexpressed genes over both stages (supplementary table S8, Supplementary Material online). For instance, male overexpressed genes had the lowest average $V_X$ values ($P < 0.05$ both in adults and in pupae), yet a relatively high $D_X$ value compared with queen overexpressed genes ($P < 0.05$; worker overexpressed genes have the highest $D_X$ among adults, $P < 0.005$; in pupae, averages are similar across castes, $P > 0.2$). These combined results,

suggesting low variation within species but relatively high differentiation (and correspondingly high $R_X$ values) in gene expression levels between species, are the hallmarks of adaptive evolution associated with lineage divergence. Finally, the findings that worker overexpressed genes have both high $D_X$ and $V_X$, which in pupae result in $R_X$ values lower than for the other caste overexpressed genes ($P <$ 0.005), suggest that worker overexpressed genes likely experienced relaxed selection on gene expression relative to queen and male overexpressed genes in both developmental stages.

We were interested in further evaluating the mode of evolution of gene expression in genes specifically differentially expressed between S. invicta and S. richteri (see e.g., Khaitovich et al. 2004; Jordan et al. 2005). Genes differentially expressed between the species had significantly larger $R_X$ values than genes showing no such expression differences in both pupae and adults (all $P < 10^{-15}$; fig. 6B and supplementary table S9, Supplementary Material online). Moreover, the relatively low $R_X$ values for the latter genes stem from both larger intraspecific variation in gene expression (higher $V_X$) and lower interspecific differentiation in gene expression (lower $D_X$) compared with genes differentially expressed between the species, in all three castes and in both developmental stages (all $P < 10^{-10}$; supplementary table S9, Supplementary Material online). Although low $D_X$ is expected for genes lacking species-specific expression, their high $V_X$ is not. Focusing on genes with species-specific expression in the adult stage, $R_X$ was significantly lower for genes differentially expressed in workers than genes differentially expressed in queens and males ($P < 10^{-10}$), consistent with relaxed diversifying selection on genes exhibiting species-specific expression in the worker caste.

## Discussion

Our results show clear differences in the patterns of gene expression between developmental stages, among castes, and between the two species of fire ant that we studied. Both the principal components analyses and the gene expression–based distance tree revealed that variation in gene expression profiles is influenced more by developmental stage than by caste membership, sex, or species identity. The finding of a very strong effect of major developmental stage is consistent with previous studies of other holometabolous insects with pronounced phenotypic differences in the life stages (Drosophila, vespine wasps), which showed major shifts in gene expression networks among larvae, pupae, and adults (Arbeitman et al. 2002; Hoffman and Goodisman 2007). Our comparisons also reveal that many genes are differentially expressed between castes within just a single developmental stage. These transient differences suggest that phenotypic variation between castes results from the transcriptional deployment of distinct sets of genes at specific developmental stages (Lebo et al. 2009).

The second major contributor to variation in gene expression was the combination of sex and caste. Sex was the stronger determinant of gene expression, as clearly illustrated by the deeper split between females and males than between workers and queens in their gene expression profiles in the expression-based distance tree (fig. 1). Although worker and queen fire ants share equivalent diploid nuclear genomes, they have very distinct patterns of gene expression in both developmental stages that we studied, as might be predicted given the extraordinary level of dimorphism between the two female castes. For instance, in adults, there are more genes differentially expressed between queens and workers than between queens and males (supplementary fig. S1, Supplementary Material online). The percentage of differentially expressed genes is even higher between males and workers, consistent with the fact that individuals of these two classes share neither gender nor reproductive capability. Thus, queens are more similar to the other two castes in terms of gene expression than workers are to males, a pattern evident also from the fact that significantly more genes are overexpressed in workers and males than in queens, both in the pupal and in the adult stages. A similar pattern is observed when considering genes that are underexpressed, with many more genes being underexpressed in males and workers compared with queens. These findings suggest that both gene activation and silencing characterize caste differentiation and function in fire ants. Such caste-specific patterns of gene expression possibly reduce antagonistic pleiotropic effects that can arise in the absence of chromosomal sex determination and genotypic caste determination (Mank et al. 2008).

Our results further indicate that the proportion of genes underexpressed relative to the proportion overexpressed is higher in adult queens than adult workers and males. This finding is consistent with results reported in the honey bee, Apis mellifera (Evans and Wheeler 2001). However, it is important to note that the adult queens used in our experiments were only recently emerged from the pupa, so it is likely that many genes associated with sexual maturation and reproduction become more highly expressed as queens age, translating into a higher number of genes specifically up-regulated in older queens.

Because fire ant workers do not reproduce and thus have only an indirect fitness component, gene expression may be under more relaxed selection in this caste than in queens and males (Linksvayer and Wade 2009). If we assume that regulation of gene expression has a genetic basis and often is under purifying selection in reproductively competent individuals, genes expressed only in workers may be predicted to exhibit both a higher level of within-species polymorphism and a greater between-species divergence than genes also expressed in males and queens. Consistent with this prediction, both intraspecific variability ($V_X$) and interspecific differentiation ($D_X$) in gene expression were higher for worker overexpressed genes than for queen and male overexpressed genes, although the differences were not always statistically

significant (supplementary table S8, Supplementary Material online). To further test these predictions, we can analyze variation in gene expression as a function of the degree of caste-biased gene expression, cbx, where larger values of cbx indicate greater caste specificity in expression (supplementary table S10, Supplementary Material online). An analysis of variance test reveals that both the caste and the degree of caste-biased expression, cbx, significantly affect $V_X$ (all $P < 10^{-10}$) and $D_X$ (all $P < 10^{-4}$), and, more importantly, a significant interaction exists between cbx and caste ($P < 10^{-10}$ for both $V_X$ and $D_X$). Specifically, $V_X$ and $D_X$ increase disproportionately more with cbx for worker overexpressed genes than for queen or male overexpressed genes (supplementary fig. S3, Supplementary Material online). This is consistent with the hypothesis that as genes become more worker overexpressed, variability within species as well as divergence between species in their expression patterns increase in parallel. Faster evolution of genes showing overexpression in one caste has been reported also in wasps (Hunt and Goodisman 2010), suggesting that caste polyphenism generally is an important determinant of regulatory gene evolution.

The gene expression–based distance tree and the principal components analyses suggest that the degree of gene expression divergence between S. invicta and S. richteri varies according to caste. In adults, worker- and male-biased genes are more frequently expressed in a species-specific manner than are queen-biased genes. This pattern is consistent with the fact that adult workers most directly experience the distinct external environments characterizing the different ranges occupied by the two species, namely tropical and subtropical habitats for S. invicta and somewhat more temperate habitats for S. richteri (Trager 1991). This is because fire ant workers forage in the external environment, whereas queens and males are largely confined to the nests, which comprise highly buffered environments that presumably are quite similar between the species. Similarly, fire ant pupae are entirely confined within the nest, where they are cared for by the workers. Consistent with this view, the divergence in gene expression between species ($D_x$) is significantly lower for genes expressed in the pupal stage than in the adult stage ($P < 0.0001$), and the number of such differentially expressed pupal genes is similar among the three castes. The total number of genes exhibiting significant interspecific differentiation (1,549) represents about 10% of the genes we studied. This is much lower than the approximately 50% of genes found to be differentially expressed between the morphologically highly similar D. melanogaster and D. simulans (Ranz et al. 2003), which belong to the same species group and diverged about 2 Ma (Russo et al. 1995). Our finding of relatively low regulatory gene divergence is consistent with S. invicta and S. richteri being close relatives (members of the same species group; Pitts et al. 2005) that are able to produce relatively fit hybrids (Shoemaker et al. 1996). The modest transcriptome divergence we observed evidently is not closely linked to morphological differences between the two species, as they differ in very few diagnostic morphological characters, with their males effectively indistinguishable (Pitts 2002). Instead, traits linked to the physiology and/or behavior may more closely reflect the transcriptional differentiation we detected between S. invicta and S. richteri.

Compared with the relatively low overall differentiation in gene expression between the two fire ant species, the degree of interspecific differentiation in caste-biased gene expression is unexpectedly high. For instance, about 41% of the genes with caste overexpression in adult S. invicta are not similarly overexpressed according to caste in adult S. richteri. In comparison, only 20% of D. melanogaster sex-biased genes were found not to be similarly sex-biased in D. simulans (Ranz et al. 2003). These data suggest a higher evolvability of regulatory elements involved in caste determination than those involved in sex determination during cladogenesis (Mank and Ellegren 2009; Hunt and Goodisman 2010).

An important point to consider in this and other comparative genomic surveys of social Hymenoptera is the different ploidy state of females (diploid) versus males (haploid). Because recessive mutations may be immediately visible to selection in the haploid state, deleterious mutations should be purged more efficiently and beneficial mutations fixed more rapidly in males, barring female sex–limited expression. Several features of our data are consistent with high levels of positive selection on genes expressed in males. In particular, the patterns of expression of genes overexpressed in one caste or differently expressed between species (as measured by $R_X$) suggest stronger positive selection on gene expression in males than in workers or queens. This is particularly evident in the pupal stage, in which spermatogenesis occurs (e.g., Passera and Keller 1992), as there are many more genes overexpressed in male pupae than in worker or queen pupae. Such genes are ideal targets for positive diversifying selection (Ellegren and Parsch 2007) and, indeed, a large fraction of the male overexpressed pupal genes are significantly differentially expressed between the two species. Male overexpressed genes may also evolve at enhanced rates due to sexual selection (e.g., Jagadeeshan and Singh 2005), although sexual selection generally is assumed to be minimal in male ants (Feldhaar et al. 2008).

Our results highlight the role of differential gene expression in species where distinct phenotypes are not merely the result of sexual dimorphism. In fire ants, workers have a considerable number of genes that are specifically up- or down-regulated compared with males and queens. Moreover, workers also have more genes significantly differentially expressed between species than do the other castes. Thus, much of the evolution of gene expression in ants may occur in the worker caste despite the fact that these individuals are largely or completely sterile. This can be explained by a combination of factors, including adult workers experiencing the most diverse environments and exhibiting the broadest behavioral repertoires, and both queens and males having lost ancestral Hymenopteran

somatic-specific traits related to feeding and self-maintenance (see also Toth and Robinson 2007). Workers therefore constitute the extended somatic phenotype of the sexual castes.

## Supplementary Material

Supplementary tables S1–S9 and figs. S1–S3 are available at *Molecular Biology and Evolution* online (http://www.mbe.oxfordjournals.org/).

## Acknowledgments

## References

Arbeitman MN, Furlong EEM, Imam F, Johnson E, Null BH, Baker BS, Krasnow MA, Scott MP, Davis RW, White KP. 2002. Gene expression during the life cycle of *Drosophila melanogaster*. *Science* 297:2270–2275.

Babbitt CC, Fedrigo O, Pfefferle AD, Boyle AP, Horvath JE, Furey TS, Wray GA. 2010. Both noncoding and protein-coding RNAs contribute to gene expression evolution in the primate brain. *Genome Biol Evol.* 2010:67–79.

Cassone B, Mouline K, Hahn MW, White B, Pombi M, Simard F, Costantini C, Besansky N. 2008. Differential gene expression in incipient species of Anopheles gambiae. *Mol Ecol.* 17:2491–2504.

Ellegren H, Parsch J. 2007. The evolution of sex-biased genes and sex-biased gene expression. *Nat Rev Genet.* 8:689–698.

Enard W, Khaitovich P, Klose J, et al. (13 co-authors). 2002. Intra- and interspecific variation in primate gene expression patterns. *Science* 296:340–343.

Evans JD, Wheeler DE. 2001. Expression profiles during honeybee caste determination. *Genome Biol.* 2:RESEARCH0001.

Feldhaar H, Foitzik S, Heinze J. 2008. Review. Lifelong commitment to the wrong partner: hybridization in ants. *Philos Trans Roy Soc B.* 363:2891–2899.

Felsenstein J. 2005. PHYLIP (phylogeny inference package). Seattle (WA): Department of Genome Sciences, University of Washington.

Fournier D, Estoup A, Orivel J, Foucaud J, Jourdan H, Le Breton J, Keller L. 2005. Clonal reproduction by males and females in the little fire ant. *Nature* 435:1230–1234.

Gilad Y, Rifkin SA, Bertone P, Gerstein MB, White KP. 2005. Multi-species microarrays reveal the effect of sequence divergence on gene expression profiles. *Genome Res.* 15:674–80.

Gräff J, Jemielity S, Parker JD, Parker KM, Keller L. 2007. Differential gene expression between adult queens and workers in the ant *Lasius niger*. *Mol Ecol.* 16:675–683.

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

Hoffman EA, Goodisman MAD. 2007. Gene expression and the evolution of phenotypic diversity in social wasps. *BMC Biol.* 5:23.

Hölldobler B, Wilson E. 1990. The ants. Cambridge (MA): The Belknap Press of Harvard University Press.

Hunt BG, Goodisman MA. 2010. Evolutionary variation in gene expression is associated with dimorphism in eusocial vespid wasps. *Insect Mol Biol.* 19:641–652.

Hunter S, Apweiler R, Attwood TK, et al. (38 co-authors). 2009. InterPro: the integrative protein signature database. *Nucleic Acids Res.* 37:D211–D215.

Jagadeeshan S, Singh RS. 2005. Rapidly evolving genes of *Drosophila*: differing levels of selective pressure in testis, ovary, and head tissues between sibling species. *Mol Biol Evol.* 22:1793–1801.

Jordan IK, Mariño-Ramírez L, Koonin EV. 2005. Evolutionary significance of gene expression divergence. *Gene.* 345:119–126.

Jouvenaz D, Allen G, Banks W, Wojcik D. 1977. A survey for pathogens of fire ants, *Solenopsis* spp., in the southeastern United States. *Flor Entomol.* 60:275–279.

Judice CC, Carazzole MF, Festa F, Sogayar MC, Hartfelder K, Pereira GAG. 2006. Gene expression profiles underlying alternative caste phenotypes in a highly eusocial bee, Melipona quadrifasciata. *Insect Mol Biol.* 15:33–44.

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.

Khaitovich P, Weiss G, Lachmann M, Hellmann I, Enard W, Muetzel B, Wirkner U, Ansorge W, Pääbo S. 2004. A neutral model of transcriptome evolution. *PLoS Biol.* 2:E132.

Khaitovich P, Enard W, Lachmann M, Pääbo S. 2006. Evolution of primate gene expression. *Nat Rev Genet.* 7:693–702.

Lebo MS, Sanders LE, Sun F, Arbeitman MN. 2009. Somatic, germline and sex hierarchy regulated gene expression during *Drosophila* metamorphosis. *BMC Genomics* 10:80.

Lemos B, Araripe LO, Hartl DL. 2008. Polymorphic y chromosomes harbor cryptic variation with manifold functional consequences. *Science* 319:91–93.

Linksvayer TA, Wade MJ. 2009. Genes with social effects are expected to harbor more sequence variation within and between species. *Evolution* 63:1685–1696.

Mank JE, Ellegren H. 2009. Are sex-biased genes more dispensable? *Biol Lett.* 5:409–412.

Mank JE, Hultin-Rosenberg L, Zwahlen M, Ellegren H. 2008. Pleiotropic constraint hampers the resolution of sexual antagonism in vertebrate gene expression. *Am Nat.* 171:35–43.

Meiklejohn CD, Parsch J, Ranz JM, Hartl DL. 2003. Rapid evolution of male-biased gene expression in *Drosophila*. *Proc Natl Acad Sci U S A.* 100:9894–9899.

Meiklejohn CD, Townsend JP. 2006. A Bayesian method for analysing spotted microarray data. *Brief Bioinform.* 6:318–330.

Ortiz-Barrientos D, Counterman BA, Noor MAF. 2006. Gene expression divergence and the origin of hybrid dysfunctions. *Genetica* 129:71–81.

Passera L, Keller L. 1992. The period of sexual maturation and the age at mating in *Iridomyrmex humilis*, an ant with intranidal mating. *J Zool.* 228:141–153.

Pereboom JJM, Jordan WC, Sumner S, Hammond RL, Bourke AFG. 2005. Differential gene expression in queen-worker caste determination in bumble-bees. *Proc R Soc B.* 272:1145–1152.

Petralia R, Vinson S. 1979. Developmental morphology of larvae and eggs of the imported fire ant, *Solenopsis invicta*. *Ann Entomol Soc Am.* 72:472–484.

Pitts J. 2002. A cladistic analysis of the Solenopsis saevissima species-group (Hymenoptera: Formicidae). Athens (GA): University of Georgia.

Pitts J, McHugh J, Ross KG. 2005. Cladistic analysis of the fire ants of the *Solenopsis saevissima* species-group (Hymenoptera: formicidae). *Zool Scr.* 34:493–505.

R Development Core Team. 2008. R: a language and environment for statistical computing. Vienna (Austria): R Foundation for Statistical Computing.

Ranz JM, Castillo-Davis CI, Meiklejohn CD, Hartl DL. 2003. Sex-dependent gene expression and evolution of the *Drosophila* transcriptome. *Science* 300:1742–1745.

Rebeiz M, Pool JE, Kassner VA, Aquadro CF, Carroll SB. 2009. Stepwise modification of a modular enhancer underlies adaptation in a *Drosophila* population. *Science* 326:1663–1667.

Rifkin SA, Kim J, White KP. 2003. Evolution of gene expression in the *Drosophila melanogaster* subgroup. *Nat Genet.* 33:138–144.

Ritchie M, Silver J, Oshlack A, Holmes M, Diyagama D, Holloway A, Smyth G. 2007. A comparison of background correction methods for two-colour microarrays. *Bioinformatics* 23:2700–2707.

Ross K, Robertson J. 1990. Developmental stability, heterozygosity, and fitness in two introduced fire ants (*Solenopsis invicta* and *S. richteri*) and their hybrid. *Heredity* 64:93–103.

Ross K, Vander Meer R, Fletcher D, Vargo E. 1987. Biochemical phenotypic and genetic studies of two introduced fire ants and their hybrid (Hymenoptera: formicidae). *Evolution* 41:280–293.

Ross KG, Krieger MJB, Shoemaker DD. 2003. Alternative genetic foundations for a key social polymorphism in fire ants. *Genetics* 165:1853–1867.

Ross KG, Krieger MJB, Shoemaker DD, Vargo EL, Keller L. 1997. Hierarchical analysis of genetic structure in native fire ant populations: results from three classes of molecular markers. *Genetics* 147:643–655.

Russo CA, Takezaki N, Nei M. 1995. Molecular phylogeny and divergence times of drosophilid species. *Mol Biol Evol.* 12:391–404.

Scharf ME, Wu-Scharf D, Pittendrigh BR, Bennett GW. 2003. Caste- and development-associated gene expression in a lower termite. *Genome Biol.* 4:R62.

Schwander T, Lo N, Beekman M, Oldroyd BP, Keller L. 2010. Nature versus nurture in social insect caste differentiation. *Trends Ecol Evol.* 25:275–282.

Sen Sarma M, Whitfield CW, Robinson GE. 2007. Species differences in brain gene expression profiles associated with adult behavioral maturation in honey bees. *BMC Genomics.* 8:202.

Shoemaker DD, Ross KG, Arnold ML. 1996. Genetic structure and evolution of a fire ant hybrid zone. *Evolution* 50:1958–1976.

Smyth G. 2004. Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol.* 3:3.

Smyth G, Speed T. 2003. Normalization of cDNA microarray data. *Methods* 31:265–273.

Snell-Rood EC, Cash A, Han MV, Kijimoto T, Andrews J, Moczek AP. 2010. Developmental decoupling of alternative phenotypes: insights from the transcriptomes of horn-polyphenic beetles. *Evolution* doi: 10.1111/j.1558-5646.2010.01106.x.

Sumner S, Pereboom JJM, Jordan WC. 2006. Differential gene expression and phenotypic plasticity in behavioural castes of the primitively eusocial wasp, Polistes canadensis. *Proc R Soc B.* 273:19–26.

Toth AL, Robinson GE. 2007. Evo-devo and the evolution of social behavior. *Trends Genet.* 23:334–341.

Trager J. 1991. A revision of the fire ants, *Solenopsis geminata* group (Hymenoptera: formicidae, Myrmicinae). *J New York Entomol Soc.* 99:141–198.

Vargo E, Fletcher D. 1986. Evidence of pheromonal queen control over the production of male and female sexuals in the fire ant, *Solenopsis invicta*. *J Comp Physiol A.* 159:741–749.

Volny VP, Gordon DM. 2002. Genetic basis for queen-worker dimorphism in a social insect. *Proc Natl Acad Sci U S A.* 99:6108–6111.

Wang J, Jemielity S, Uva P, Wurm Y, Gräff J, Keller L. 2007. An annotated cDNA library and microarray for large-scale gene-expression studies in the ant *Solenopsis invicta*. *Genome Biol.* 8:R9.

Weil T, Korb J, Rehli M. 2009. Comparison of queen-specific gene expression in related lower termite species. *Mol Biol Evol.* 26:1841–1850.

Whitehead A, Crawford DL. 2006a. Neutral and adaptive variation in gene expression. *Proc Natl Acad Sci U S A.* 103:5425–5430.

Whitehead A, Crawford DL. 2006b. Variation within and among species in gene expression: raw material for evolution. *Mol Ecol.* 15:1197–1211.

Wurm Y, Wang J, Keller L. 2010. Changes in reproductive roles are associated with changes in gene expression in fire ant queens. *Mol Ecol.* 19:1200–1211.

Wurm Y, Uva P, Ricci F, Wang J, Jemielity S, Iseli C, Falquet L, Keller L. 2009. Fourmidable: a database for ant genomics. *BMC Genomics.* 10:5.

Yang X, Schadt EE, Wang S, Wang H, Arnold AP, Ingram-Drake L, Drake TA, Lusis AJ. 2006. Tissue-specific expression and regulation of sexually dimorphic genes in mice. *Genome Res.* 16:995–1004.