Insect Molecular Biology

Insect Molecular Biology (2010) **19**(5), 641–652 doi: 10.1111/j.1365-2583.2010.01021.x

Evolutionary variation in gene expression is associated with dimorphism in eusocial vespid wasps

B. G. Hunt and M. A. D. Goodisman

School of Biology, Georgia Institute of Technology, Atlanta, GA, USA

Abstract

Phenotypic diversity is frequently generated by differences in gene expression. In this study, we addressed the relationship between homology in gene expression and phenotype among four species of eusocial wasps. Specifically, we investigated the evolution of caste-specific and sex-specific gene expression patterns associated with caste polyphenisms and sexual dimorphisms. We also identified several genes with functions relevant to their phenotype-specific roles. Our results suggest that gene expression profiles associated with caste polyphenisms may evolve rapidly relative to those associated with sexes. Thus, caste-biased genes may undergo less regulatory constraint or be subject to greater neutral variation in expression than sex-biased genes.

Keywords: caste, gene regulation, homology, sexual dimorphism, social insect.

Introduction

Understanding the molecular basis of phenotypic diversity is one of the fundamental challenges of evolutionary biology. In recent years, a surge of attention and debate has focused on the relative evolutionary roles played by protein-coding sequence, *cis*-regulatory elements, epigenetic inheritance, and developmental processes (West-Eberhard, 2003; Pigliucci *et al*., 2006; Hoekstra & Coyne, 2007; Moczek, 2007; Carroll, 2008; Jablonka & Raz, 2009). This attention has not been in vain. Among other findings, a large body of research now demonstrates that phenotypic diversification frequently arises through changes in the spatial and temporal expression of functionally conserved proteins (Carroll, 2008).

Polyphenisms, which arise when environmental differences give rise to discrete phenotypic classes from a single genotype, provide an opportunity to study phenotypic divergence in the context of a shared genetic background. Insects, in particular, provide numerous examples of polyphenisms marked by the canalization of highly elaborate, plastic phenotypes (Evans & Wheeler, 2001; Nijhout, 2003). Examples include beetle horns (Moczek *et al*., 2007), female aphid dispersal and parthenogenetic morphs (Brisson & Stern, 2006) and eusocial insect castes (Wilson, 1971; Wheeler, 1986). In each of these cases differences in gene expression, rather than differences in genotype, are associated with the presence of alternate phenotypes (Sumner, 2006; Brisson *et al*., 2007; Goodisman *et al*., 2008; Smith *et al*., 2008; Kijimoto *et al*., 2009). In addition, sex-specific differences in phenotype, although often triggered by genetic differences, are also associated with widespread differences in gene expression (Jin *et al*., 2001; Ellegren & Parsch, 2007; Ayroles *et al*., 2009). Thus gene regulation has been widely linked to the developmental fate and adult function of distinct phenotypic classes.

Many comparative evolutionary studies have investigated the degree of similarity in the molecular basis of independently arising, convergent phenotypes. For example, discrete queen and worker castes have arisen multiple times independently in eusocial insects and exhibit limited, but notable, convergence in their underlying gene expression patterns (Toth *et al*., 2007; Smith *et al*., 2008). Interestingly, many studies have also observed surprising flexibility in the molecular basis of similar phenotypes with a shared evolutionary origin in different species, such as gene expression patterns associated with sexual dimorphisms (Zhang *et al*., 2007) and caste polyphenisms (Abouheif & Wray, 2002; Weil *et al*., 2009). An important set of evolutionary questions thus concerns the levels of functional drift and neutral variation influencing gene expression profiles associated with homologous phenotypes (Wray & Abouheif, 1998; True & Haag, 2001; Khaitovich *et al*., 2004).

First published online 6 June 2010.

Correspondence: Brendan G. Hunt, School of Biology, Georgia Institute of Technology, 310 Ferst Drive, Atlanta, GA 30332, USA. Tel.: +404 8311810; fax: +404-894-0519; e-mail: brendan.hunt@gatech.edu

In this study, we sought to understand how gene expression patterns evolve in homologous phenotypic classes of eusocial insects. Hymenopteran eusocial insects, which include ants, some bees, and some wasps, are attractive candidates for studying the association between variation in phenotype and gene expression for several reasons. First, the caste system that defines eusociality has arisen at least nine independent times within the Hymenoptera (Wilson, 1971; Hughes *et al*., 2008). Specifically, in all hymenopteran eusocial insects, labour is divided within a colony such that queens and males reproduce and female workers engage in tasks related to brood rearing and colony defence (Wilson, 1971). Thus, queen and worker castes, which often exhibit extreme polyphenism, are prime targets for studying the evolutionary maintenance of alternate phenotypes arising in a single sex. Second, hymenopterans do not exhibit chromosomal sex determination. Instead, males are generally produced by unfertilized, haploid eggs (arrhenotokous parthenogenesis) and females are produced by fertilized, diploid eggs (Heimpel & de Boer, 2008). Thus, males and females possess the same genes, and sex differences, like caste differences, develop in large part through variation in gene expression associated with differential gene activation and differences in ploidy (Aron *et al*., 2005; Hoffman & Goodisman, 2007). Finally, although castes and sexes are both examples of dramatic dimorphisms, castes have arisen far more recently than sexes during the course of evolution. Our study takes advantage of the unique characteristics of eusocial hymenopterans to reveal the potential evolutionary consequences of differences between caste polyphenisms and sexual dimorphisms on the evolution of caste-specific and sex-specific gene expression patterns.

We focused our investigation on wasps of the family Vespidae because they exhibit wide variation in social complexity. For example, wasps in the subfamily Vespinae exhibit dramatic polyphenisms and have large colony sizes, whereas wasps in the subfamily Polistinae tend to exhibit relatively small colony sizes and behaviorally-defined castes (Hunt, 2007). In this study, we examined relative gene expression levels for 21 genes in castes and sexes of four vespid wasp species with a common eusocial ancestor (Hines *et al*., 2007; Hunt, 2007) using quantitative real-time PCR. Two of the four species were informative of differences arising between female castes and sexes, whereas the other two species provided information on sex differences alone, due to partial sampling of female castes. We predicted that gene expression profiles specific to a given female caste would evolve as fast or faster than gene expression profiles specific to a given sex (Jin *et al*., 2001; Khaitovich *et al*., 2004; Cutter & Ward, 2005; Zhang *et al*., 2007; Ayroles *et al*., 2009). Gene expression associated with castes may be particularly fast-evolving because castes are relatively recent evolutionary innovations (Hughes *et al*., 2008), in contrast with sexes, and recent experimental evidence suggests that castes exhibit flexible underlying patterns of gene regulation across taxa (Weil *et al*., 2009).

Results

We analysed gene expression in three vespine wasps, *Vespula maculifrons*, *Vespula squamosa* (commonly known as yellowjackets; Macdonald & Matthews, 1981, 1984) and *Dolichovespula maculata* (hornets; Archer, 2006), as well as one polistine wasp, *Polistes exclamans* (paper wasps; Turillazzi & West-Eberhard, 1996; Hunt, 2007). Relative levels of gene expression were determined for adult queens, workers and males in *V. maculifrons* and *V. squamosa*, in adult workers and males in *D. maculata*, and in adult females (caste undetermined) and males in *P. exclamans* (Table 1, Figs 1, 2). Data were collected for up to 21 genes in each species (Fig. 1, see Experimental procedures).

We first investigated whether sexes or castes showed more conserved patterns of gene regulation across taxa. Comparisons of relative gene expression levels for the nine genes with data from sexes and castes of all four species revealed an overall clustering of male and female profiles across taxa (Fig. 2A). Comparisons of relative gene expression levels for the 14 genes with data from sexes and castes of both *Vespula* species revealed that seven genes (contigs 61, 155, 350, 446, 504, 506, and 677) were consistently expressed more highly in females than males (with statistically higher expression in female castes of at least one taxon; Fig. 2B, C). Meanwhile, two of 14 genes (contigs 339, and 626) were consistently more highly expressed in males than females in both species (with statistically higher expression in males of at least one taxon; Fig. 1 and 2B, C).

We next determined whether genes showed conserved expression relationships between queen and worker castes in both *Vespula* species. Comparisons of relative gene expression levels in queens and workers revealed that some genes were consistently differentially expressed between castes in both species. Specifically, four of 14 genes (contigs 350, 504, 506, and 583) were more highly expressed in queens than workers in *V. maculifrons* and *V. squamosa* (with statistically significant queen upregulation in one species in all four cases; Fig. 1 and 2D). In contrast, seven of 14 genes had divergent relationships with respect to queen and worker gene expression differences in *V. maculifrons* and *V. squamosa*, demonstrating the labile nature of caste-biased gene expression over evolutionary time (Figs 1 and 2D). Five of these genes (contigs 61, 339, 446, 626, and 677) exhibited statistically different expression levels between

Table 1. Gene annotation and relative expression patterns among males (M), queens (Q), and workers (W) in *Vespula squamosa* and *Vespula maculifrons*, between workers and males in *Dolichovespula maculata*, and between late-season females (F) and males in *Polistes exclamans*

	Gene ontology similarity*	Relative expression pattern†				Gene ontology annotation‡
Contig ID	Putative orthologue	*V. squamosa*	*V. maculifrons*	*D. maculata*	*P. exclamans*	Biological process terms
Contig52	regucalcin	W ≈ M > Q	–	W > M	–	Cellular calcium ion homeostasis, positive regulation of ATPase activity, regulation of calcium-mediated signalling
Contig61	Ribosomal protein L13	Q > W ≈ M	NS	W > M	–	Mitotic spindle elongation, mitotic spindle organization, translation
Contig88	Vitellogenin-2	–	Q > W ≈ M	W > M	–	Determination of adult life span
Contig155	Ribosomal protein L32	NS	Q ≈ W > M	W > M	F > M	Mitotic spindle elongation, mitotic spindle organization, translation
Contig170	lipoma HMGIC fusion partner-like 3	NS	M > Q ≈ W	–	F > M	biological process
Contig272	Ribosomal protein L8	NS	–	W > M	–	Mitotic spindle elongation, mitotic spindle organization, translation
Contig339	Tropomyosin 2	M > W > Q	M > Q ≈ W	M > W	–	Heart development
Contig350	Stubarista	Q > M	Q ≈ W > M	W > M	F > M	Mitotic spindle elongation, mitotic spindle organization, translation
Contig422	Ribosomal protein S18	–	Q > W ≈ M	W > M	F > M	Mitotic spindle organization, translation, translational initiation
Contig430	Twinstar	W > Q ≈ M	W ≈ M > Q	W > M	F > M	Actin filament organization, axonogenesis, leg segmentation
Contig446	Elongation factor 1alpha100E	W > M	W > Q > M	W > M	F > M	Translation, translational elongation
Contig462	Ribosomal protein L15	–	Q ≈ W > M	W > M	F > M	Translation
Contig498	string of pearls	NS	–	W > M	–	Translation
Contig504	CG4692	NS	Q > W > M	NS	NS	Proton transport
Contig506	Ribosomal protein L23	NS	Q > M	–	F > M	Mitotic spindle elongation, mitotic spindle organization, translation
Contig531	Ribosomal protein LP0	–	–	W > M	F > M	DNA repair, translation
Contig583	Ribosomal protein L17	W > M	Q ≈ W > M	NS	F > M	Mitotic spindle elongation, mitotic spindle organization, translation
Contig626	wings up A	W ≈ M > Q	M > W ≈ Q	NS	F > M	Muscle maintenance, muscle organ development, skeletal muscle tissue development
Contig644	Superoxide dismutase	NS	NS	W > M	F > M	Aging, determination of adult life span, removal of superoxide radicals
Contig664	40S ribosomal protein S27	NS	NS	NS	F > M	Cell proliferation, translational elongation
Contig677	60S ribosomal protein L23a	W > Q	Q ≈ W > M	W > M	–	Translational elongation

*BLAST similarity scores and hit species information given in Table S3.
†A caste's level of expression was designated as less than or greater than another if $P < 0.05$. When a caste had intermediate but nonsignificant expression differences from two significantly different groups, it was not included in the Table. – = no data, NS = nonsignificant differences (see Fig. 1).
‡Where more than three gene ontology(GO) biological process terms existed for a gene, only three terms were chosen. A full list of matching GO terms and accessions are presented in Table S3.

queens and workers within at least one taxon. Furthermore, queen and worker gene expression profiles had qualitatively different relationships in *V. maculifrons* and *V. squamosa* – queen and worker gene expression profiles clustered together more closely in *V. maculifrons* than in *V. squamosa* (Fig. 2A, B). Together these results suggest that relative gene expression levels between castes vary widely among taxa, and gene expression profiles associated with sexes are less labile than those associated with castes.

We further tested whether *Vespula* genes in our study were convergent, with respect to caste-specific regulation,

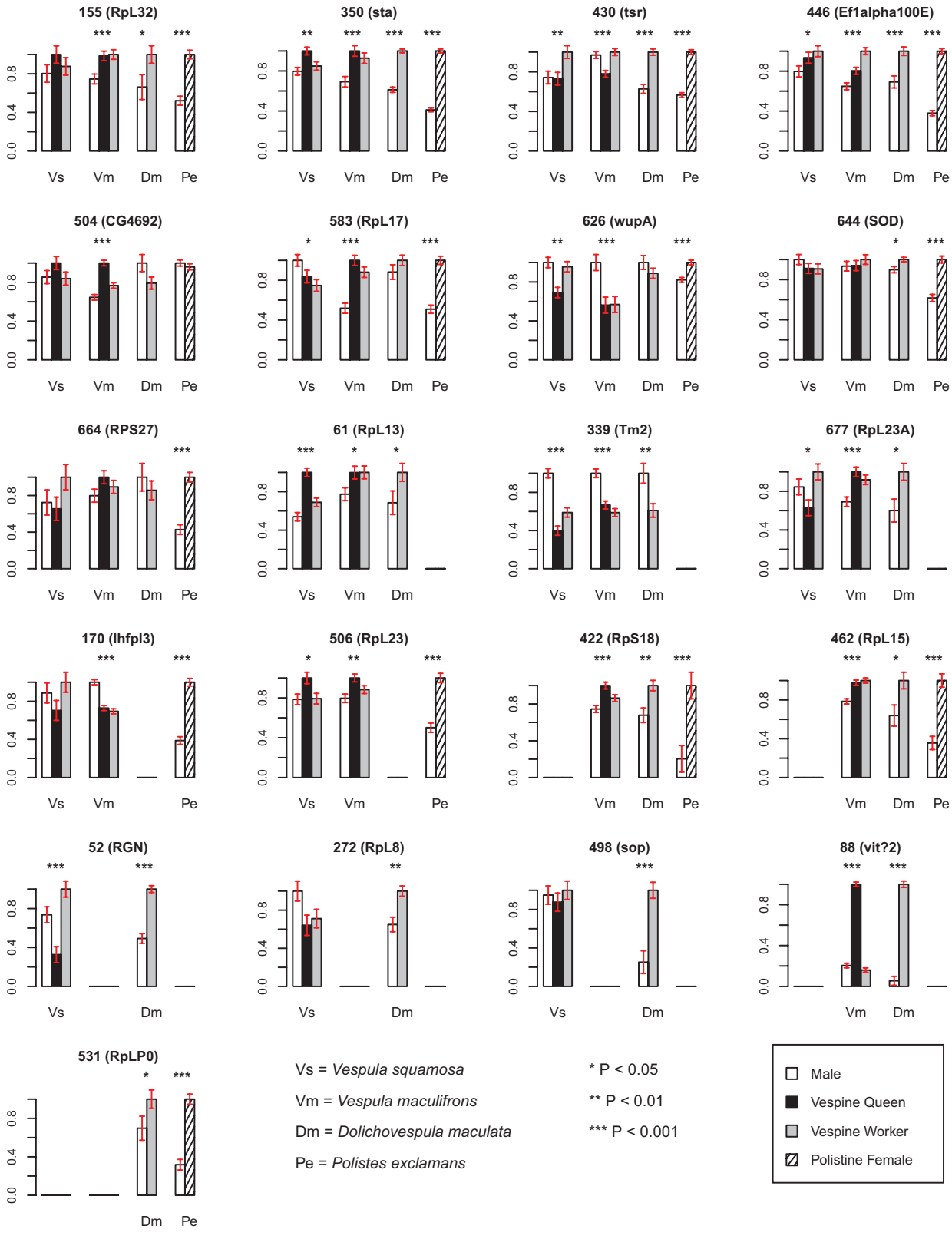

Figure 1. Relative expression differences between castes and sexes in vespid wasps. Normalized least square means of relative expression levels are plotted with 95% confidence intervals for four vespid wasp species. *P*-values denote the outcome of an ANOVA test of differences between castes.

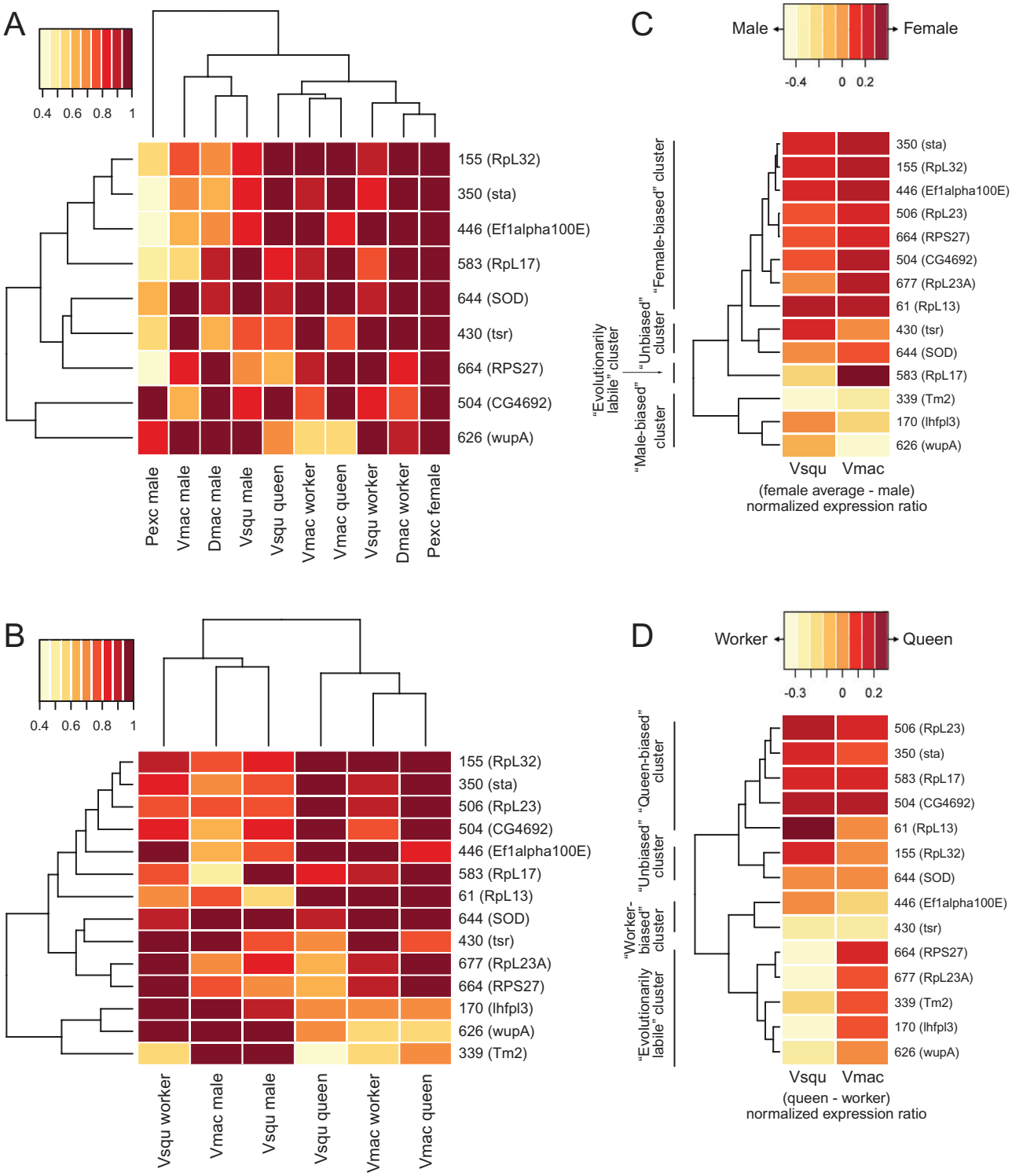

Figure 2. Gene expression profiles associated with castes and sexes in vespid wasps. Increasing or decreasing block saturation indicates higher or lower relative levels of expression for a particular gene in a given caste. (A) Normalized gene expression intensities in four species illustrate that male and female gene expression profiles cluster across taxa; all genes with data from every phenotypic class in all four species are shown. (B) Normalized gene expression intensities in *Vespula squamosa* (Vsqu) and *Vespula maculifrons* (Vmac) illustrate that female queen and worker gene expression profiles cluster in *V. maculifrons* but not *V. squamosa*. (C) Differences in normalized gene expression intensities between females and males indicate that male-biased and female-biased genes cluster among *Vespula* species. (D) Differences in normalized gene expression intensities between queens and workers indicate that a limited proportion of genes exhibit conserved differences with respect to caste. All genes with data for queens, workers and males in both *Vespula* species are shown in panels B–D.

Table 2. Comparison of gene expression between *Apis mellifera* queens (Q) and workers (W) and two species of *Vespula* wasps

Contig ID	*A. mellifera* gene name (ID)	BLAST E-value (score)	*A. mellifera* expression	*Vespula squamosa* expression‡	*Vespula maculifrons* expression‡
Contig88	Vitellogenin (GB13999-PA)	6e-63 (236)	Q > W*	–	Q > W
Contig350	Similar to stubarista CG14792-PA, isoform A (GB19082-PA)	1e-149 (524)	W > Q*	Q ≈ W	W ≥ Q
Contig430	Similar to Cofilin/actin-depolymerizing factor homologue (Protein D61) (Protein twinstar) (GB18917-PA)	3e-82 (299)	Q ≈ W*	W > Q	W > Q
Contig446	Elongation factor 1-alpha (GB16844-PA)	6e-74 (274)	Q ≈ W*, Q > W†	W ≥ Q	W > Q
Contig498	Ribosomal protein S2e (GB10861-PA)	1e-151 (530)	Q > W*	Q ≈ W	–
Contig504	Similar to CG4692-PB, isoform B (GB18417-PA)	4e-30 (126)	Q > W*	Q ≈ W	Q > W
Contig644	CuZn superoxide dismutase (GB10133-PA)	6e-43 (170)	Q > W*	Q ≈ W	Q ≈ W

>Indicates significantly higher expression, ≈ indicates nonsignificant expression difference, – indicates no data.
*Data comparing expression levels in adult queen and worker brains (Grozinger *et al*., 2007).
†Data comparing 4th instar queen-destined and worker-destined larvae; other genes in Table 2 are unreported (either due to nonsignificant expression differences or technical issues) in comparisons of queen- and worker-destined 3rd, 4th, and 5th instar larvae (Barchuk *et al*., 2007).
‡According to a post-hoc Tukey's HSD analysis for the effect of caste in our whole-body expression analysis (Table S1; see experimental procedures).

with the independently eusocial honeybee *Apis mellifera* (Table 2; Barchuk *et al*., 2007; Grozinger *et al*., 2007). Two of the seven genes we compared had convergent relationships between expression pattern and caste, both of which were upregulated in queens. These were orthologues of *vitellogenin* and CG4692 (see Discussion). These results are consistent with the presence of few genes showing convergent caste-specific expression patterns in distantly related, independently eusocial taxa.

Finally, we examined the variability of gene regulation among castes and sexes to see if relative gene expression levels differed among these phenotypic classes within each taxon. We found that all 21 genes exhibited significant differences in expression among some phenotypic classes for at least one species, suggesting that gene expression differences are widespread among dimorphic adult phenotypes (Fig. 1). We also observed widespread differences in gene expression between individuals of the same caste, even for *V. squamosa* samples, which were all collected at a single time point, providing additional support to the finding that gene expression is highly variable between individuals (Table S1; Oleksiak *et al*., 2002).

Discussion

Evolution of gene expression in castes and sexes

One of the most striking patterns that emerged from our analysis is that caste-biased gene expression profiles appear to have undergone greater rates of change than sex-biased gene expression profiles in the *Vespula* genus (Fig. 2B, C, D). Previous studies have demonstrated that sex-biased genes undergo elevated gene expression divergence between taxa relative to non-sex biased genes (Meiklejohn *et al*., 2003; Ranz *et al*., 2003). However, our data suggest that caste-biased gene expression may evolve at even faster rates than sex-biased gene expression. The observation that caste-biased gene expression patterns are largely species-specific agrees with recent work comparing caste-biased gene expression patterns among closely related termites (Fig. 2D; Weil *et al*., 2009). Considering the flexibility in caste-specific gene expression patterns we observed within a single genus (Fig. 2B, D), the fact that we observed little convergence between vespid wasps and *A. mellifera* is not surprising (Table 2). Together, these results suggest that only a limited subset of caste-biased genes may contribute to a common 'genetic toolkit' underlying the evolution of eusocial castes (Toth & Robinson, 2007). However, we also caution that differences observed between *Vespula* and *A. mellifera* gene expression patterns likely arise, in part, because the honeybee data were generated from adult brains (Grozinger *et al*., 2007) and whole-body larvae (Barchuk *et al*., 2007), while our *Vespula* data were generated from whole-body adults (Table 2).

Because female caste dimorphisms are relatively recent innovations, whereas sexual dimorphisms are widespread and ancient among animals, it is possible that greater resolution has been achieved in intralocus conflict over phenotypic optima arising between sexes than between castes (Bonduriansky & Chenoweth, 2009). This may help to explain the labile gene expression patterns observed among castes of different species. Alternatively, genes that exhibit caste-biased expression patterns may simply be subject to fewer pleiotropic constraints or exhibit higher dispensability, and a lower impact on fitness, than genes associated with sex-bias (Mank *et al*., 2008; Mank & Ellegren, 2009). Accordingly, although prior transcriptomic studies have revealed that variation in gene expression is not inherently neutral (Denver *et al*., 2005; Rifkin *et al*., 2005), the flexibility in caste-biased gene regulation could represent a larger role for neutral variation in shaping caste-specific patterns of gene expression relative to sex-specific gene expression (Khaitovich *et al*., 2004).

We also observed that queen and worker gene expression profiles clustered together more closely in *V. maculifrons* than in *V. squamosa* (Fig. 2A, B). This may reflect species-level differences in the degree of dimorphism between queen and worker castes because *V. squamosa* castes are among the most dimorphic of vespid wasps (Greene, 1991). Alternatively, the clustering of worker and male phenotypes in *V. squamosa* may be an effect of the limited number of loci used in our study (Hoffman & Goodisman, 2007). Thus, the relationship between levels of caste-biased gene expression and the degree of caste dimorphism should be more fully addressed on a transcriptome-wide scale in the future.

There were readily apparent differences in the gene expression profiles of males and females in our data (Fig. 2A), as has been observed previously in taxa such as *Drosophila melanogaster* (Ayroles *et al*., 2009). In hymenopteran insects, where males are haploid and females are diploid, differences in gene expression between the sexes arise both as a function of differential gene activation and differential gene dosage related to ploidy levels (Birchler *et al*., 2005; Mileyko *et al*., 2008; Mank, 2009). Based purely on ploidy levels, one would expect many genes to be more highly expressed in females than males. However, hymenopterans may undergo dosage compensation, sometimes only in limited tissues (Aron *et al*., 2005) or genes (Mank, 2009). Furthermore, recent theory suggests that changes in gene copy number may result in disproportionally large changes in gene expression levels in the context of biological gene interaction networks (Mileyko *et al*., 2008). Thus, it remains unclear precisely how gene dosage in haplodiploid organisms alters gene regulation in males and females.

It is notable that we observed widespread female-biased gene expression at the loci we assayed (Fig. 2A, C). This is particularly evident for *P. exclamans,* which exhibited increased female bias relative to the three vespine species in our study (Fig. 2A). This is consistent with prior research demonstrating that sex-specific gene expression patterns change in line with increasing evolutionary divergence (Zhang *et al*., 2007). However, we do not expect that female bias in our data could arise from differences in ploidy levels because the total numbers of gene transcripts were held constant for each of our samples using global cDNA normalization. Thus, this female bias potentially reflects differences in the variance in gene expression levels among genes within sexes, where relatively few male-biased loci contribute disproportionately to the total pool of male gene transcripts. Alternatively, this could reflect our limited sampling of loci.

As with ploidy, effects of experimental design must be considered in order to interpret our results appropriately. In particular, given the limited number of genes in our study our conclusions must be vetted against the possible effects of ascertainment bias. First, our methods preferentially favoured genes with high levels of sequence conservation because we selected only those loci with enough conservation to facilitate the design of cross-species primers (see Experimental procedures). For example, 12 of the 21 genes were chosen as candidates for primer development based on sequence similarity to genes in an ant with developmental gene expression data (Table S2; Goodisman *et al*., 2005). Accordingly, we cannot say whether the observed differences in regulatory lability between caste-biased genes and sex-biased genes would hold true for rapidly evolving genes. Second, expressed sequence tag (EST) sequences used for primer development in this study were chosen based on preliminary analysis according to diverse biological interests, including differential expression between queen and worker castes (Table S2). Specifically, nine genes in our dataset had preliminary gene expression data for *V. squamosa* (Table S2; Hoffman & Goodisman, 2007). Of these, four genes were previously found to exhibit caste-biased gene expression patterns and five had no a *priori* predictions with respect to caste (Table S2). None of the genes in our study had a *priori* predictions with respect to sex (Table S2). Regardless, we do not believe that the selection of genes previously shown to display caste-biased expression has spuriously given rise to more labile regulation for caste-biased genes than sex-biased genes because two of the three predicted caste-biased genes analysed in both *Vespula* species exhibited conserved expression patterns among castes (contigs 504 and 583; Figs 1 and 2D).

Functional significance of gene expression

Several genes in our study displayed particularly informative patterns of expression in the context of functional annotation and prior studies of the molecular differences underlying eusocial insect castes. For instance, orthologues of a few genes have previously been linked to the division of labour in eusocial insects. Namely, *vitellogenin* is a yolk protein believed to play an important role in egg production and antioxidant function related to queen longevity and is strikingly upregulated in queens of diverse eusocial insect taxa (Table 2; Scharf *et al*., 2005; Sumner *et al*., 2006; Corona *et al*., 2007; Graff *et al*., 2007). Here we find that *vitellogenin* is dramatically upregulated in queens relative to workers and males in *V. maculifrons* and upregulated in *D. maculata* workers relative to males (Fig. 2). Although *V. squamosa* data is not presented for *vitellogenin* because its dilution curve did not meet our standards (see Experimental procedures), *V. squamosa vitellogenin* expression patterns were qualitatively similar to *V. maculifrons*, with dramatically higher expression in

queens than males or workers. The upregulation of *vitellogenin* in *D. maculata* workers relative to males may be explained by high levels of worker reproduction in this species (Foster *et al*., 2001). In contrast, *V. maculifrons* and *V. squamosa* workers rarely reproduce (Kovacs & Goodisman, 2007) and workers and males in these species show limited *vitellogenin* expression (Fig. 1).

Superoxide dismutase (SOD) is another well-studied gene that has been hypothesized to play a key role in aging because of its association with longevity in *Drosophila* (Orr & Sohal, 1994). Although SOD is upregulated in brains of queens in the honeybee *A. mellifera* (Table 2; Grozinger *et al*., 2007), a SOD orthologue was found to be down-regulated in queens relative to workers and males in the ant *Lasius niger* (Parker *et al*., 2004). In *V. maculifrons* and *V. squamosa*, SOD is not significantly differently expressed among queens, workers and males (Fig. 1). However, SOD expression likely varies among tissues and over time. We did not investigate variation in expression in different tissues or developmental stages. Thus, we can only state that our results generally support the view that antioxidant activity by SOD does not contribute to longevity in social insects (Parker *et al*., 2004), but further work is needed to address this question fully in vespid wasps.

Finally, CG4692, which has been found to be upregulated in *A. mellifera* queens, is also upregulated in queens of *V. maculifrons* (Table 2; Grozinger *et al*., 2007). CG4692 is associated with proton transport, which may bear functional significance to caste differences in line with the widely observed queen upregulation of metabolic genes (Cristino *et al*., 2006; Grozinger *et al*., 2007; Smith *et al*., 2008).

A few male-biased genes in our study also had noteworthy functions. Two genes associated with musculature appear to be consistently upregulated in males of vespine wasps: orthologues of *wings up A* and, most strikingly, *tropomyosin 2* (Table 1). Remarkably, haploid males in hymenopteran social insects have doubled the nuclear DNA content of muscles to effectively restore diploidy in these tissues (Aron *et al*., 2005; see above). This important evolutionary innovation, known as endoreduplication, may have occurred in response to the demands of flight for dispersal and mating. Our results are consistent with a scenario in which endoreduplication results in higher levels of transcription (Edgar & Orr-Weaver, 2001).

Conclusion

To our knowledge, this is among the first studies to investigate evolutionary variation in gene expression among both castes and sexes. Our results suggest that caste-biased gene expression may be more labile than sex-biased gene expression and demonstrate the utility of social insects in addressing questions of the molecular evolution of phenotypic dimorphisms. Several directions for future investigations are suggested by our study. Firstly, differences in gene regulation that give rise to sexual dimorphism in hymenopteran eusocial insects should be more intensively studied (e.g. Hoffman & Goodisman, 2007). Secondly, more detailed comparisons of caste-biased gene expression in taxa with different levels of morphological differentiation between castes (such as vespines and polistines) will provide valuable insight into the evolutionary elaboration of eusociality (Hunt, 2007; Toth *et al*., 2007, 2009). Finally, we recognize that our study is limited in the number of loci examined. With the rapid advancement of comparative genomic resources for eusocial hymenopterans, future studies can test our hypotheses on a transcriptome-wide scale (Smith *et al*., 2008). Indeed, our hope is that this study will act as an impetus for others to further examine patterns of gene expression in the castes and sexes of vespid wasps and other eusocial insects.

Experimental procedures

Wasp samples

We sampled adult wasps from colonies collected in the metro Atlanta, Georgia (USA) area from October to November 2007, with the exception of one *D. maculata* colony collected in August 2007. We sampled three males, three pre-reproductive queens (gynes, referred to as 'queens' throughout the manuscript), and three workers from one *V. squamosa* and two *V. maculifrons* colonies. We also sampled individuals from two *D. maculata* colonies. Three workers were sampled from one of these colonies, and three males and three workers were sampled from the other. Finally, we collected two *P. exclamans* colonies and sampled three late-season females and three males from each colony. Live wasps were flash frozen using liquid nitrogen and stored at −80°C.

Expressed sequence tag sequencing and primer development

Gene-specific PCR primers were designed to assess gene expression levels. ESTs for putative target genes were sequenced from *V. squamosa* cDNA libraries following Hoffman & Goodisman (2007). Chromatogram files were trimmed for quality and vector contamination using Sequencher (Gene Codes, Ann Arbor, MI, USA). Contigs were assembled from a combination of ESTs described in Hoffman & Goodisman (2007; GenBank accession numbers EG325041–EG327184) and ESTs sequenced thereafter (GenBank accession numbers GW787760–GW792031) using CAP3 (Huang & Madan, 1999) with default settings other than percent identity (set to P = 90) and overlap (set to O = 40). EST sequences used for primer development in this study were chosen based on preliminary analysis according to diverse biological interests, including differential expression between gyne and worker castes, differential expression between larvae and adults, genes upregulated in multiple

developmental stages and castes (Hoffman & Goodisman, 2007), and those with orthologues previously studied in ants (Goodisman *et al*., 2005) (Table S2).

Gene-specific primers were designed with the goal of universal amplification among orthologues in vespid wasp taxa. BLAST alignments (Altschul *et al*., 1997) between *V. squamosa* sequences and the honeybee *A. mellifera* were used to identify putatively conserved regions for primer placement. Primers were designed using Primer3 (Rozen & Skaletsky, 2000) after masking repeats with RepeatMasker (Smit *et al*. 1996–2004). We used primers that amplified in all four taxa during non-quantitative PCR tests.

We annotated EST sequences from *V. squamosa* based on sequence similarity using gene ontology (GO) terms. Each sequence served as a query to search the GO database (release 2009-06-23) using the AmiGO search tool (Carbon *et al*., 2009) with a similarity threshold of 1e-5 and BLAST filtering enabled. If a hit with *D. melanogaster* was found, its annotation was used. Otherwise, the best hit with GO biological process annotation in another, more distantly related species was used. GO biological process terms and gene names were assigned to each *V. squamosa* EST.

Quantitative real-time reverse transcription PCR

We measured the level of gene expression in each sample using quantitative real-time reverse transcription PCR (QRT-PCR; Heid *et al*., 1996). First, RNA extraction was performed using whole bodies. Briefly, tissue was homogenized using a liquid nitrogen-chilled mortar and pestle immediately following removal from freezer. RNA was extracted using TRI Reagent (Molecular Research Center, Inc., Cincinnati, OH, USA) following standard RNA isolation protocol. For large individuals, a fraction of the homogenate was used so as not to exceed ~10% of TRI Reagent volume. Total RNA concentration was determined by UV spectrophotometry and normalized across samples. RNA was then treated with DNase to remove genomic DNA contamination using the TURBO DNA-*free* kit (Ambion, Austin, TX, USA). Removal of genomic DNA was verified by PCR and DNase treatment was repeated if necessary.

We synthesized first strand cDNA using SuperScript III reverse transcriptase (Invitrogen, Carlsbad, CA, USA) and a poly $dT_{14–18}$ primer mix in a 50 μl reaction volume. cDNA concentration was normalized across samples using agarose gel electrophoresis (Goodisman *et al*., 2005) because no empirically verified exogenous controls exist for vespid wasps. Briefly, we loaded three different volumes of each sample (6 μl, 4 μl, and 2 μl) into a 192-well agarose gel. cDNA intensity was measured from gel images using the ImageJ program (Rasband 1997–2009). A linear regression was generated based on the dilutions of each sample (Mean $R^2 = 0.99$), and the intensity predicted from the linear regression model at a 6 μl volume was used to normalize cDNA concentrations across samples. Nine samples with low cDNA concentrations had to undergo a second cDNA synthesis and subsequent normalization to a subsample of previously normalized cDNA. We used the intensity of size-standard ladders distributed throughout the gel to ensure uniform background intensity. A 1:10 dilution of normalized cDNA was aliquoted, stored at –20°C, and used for quantitative PCR within 2–3 days of thawing.

We performed QRT-PCR on StepOnePlus and Prism 7000 sequence detection systems from Applied Biosystems (Foster City, CA, USA). All samples of *V. maculifrons* and *V. squamosa* for a given gene were run in triplicate on the same 96-well plate. Samples on every plate were run alongside one species-specific dilution curve consisting of five to six dilutions (1:2, 1:4, 1:8, 1:16, 1:32, 1:64) of cDNA with a higher starting concentration than those used for expression analysis. This curve was used to control for differences in PCR efficiency between species and to generate a standard curve for relative quantification between samples within a species (Applied Biosystems, 2001). We did not make direct comparisons in intensity across species or genes. For a given primer set, all samples of *D. maculata* and *P. exclamans* were run in the same manner on a second 96-well plate.

Statistical analysis

The starting transcript abundance for each sample, relative to other samples from the same species for the same gene, was estimated from cycle threshold (Ct) values (calculated by StepOnePlus and Prism 7000 sequence detection system software) and the applicable standard curve (Applied Biosystems, 2001). In order to ensure that standard curves were of acceptable technical quality and that cDNA amplification was of acceptable efficiency, only data calibrated with standard curves exhibiting R^2 values > 0.90 are presented.

A nested ANOVA was used to test for significant differences in gene expression between castes and between individuals within castes for each species, treating males as a 'caste' to incorporate sex differences. Relative gene expression levels (the relative Ct values for samples of a given species after making intensity adjustments according to the species- and plate-specific standard curve) were used to calculate least square means for each caste using a standard least squares model. For visualization and subsequent analysis, least square means of each caste within a species were normalized such that the highest relative expression level for a given gene was assigned a value of 1.0 and expression levels of other castes became a fraction thereof. For *Vespula* species, which had three sampled castes, a post-hoc Tukey HSD test on least square means was used to determine which castes differed significantly in expression level for a given gene. ANOVA, standard least square calculations, and post-hoc tests were performed using the JMP statistical software package (SAS Institute Inc, Cary, NC).

Heatmaps with dendrograms were generated from normalized relative expression levels using R (R Development Core Team, 2008). Euclidean distances between rows and columns were used to generate heatmaps, and dendrograms were computed by performing a hierarchical cluster analysis using a set of dissimilarities for the *n* objects being clustered, according to the 'complete' method.

Apis mellifera *orthologue analysis*

We compared gene expression differences observed between *Vespula* queens and workers to gene expression differences observed between *A. mellifera* queens and workers (Barchuk *et al*., 2007; Grozinger *et al*., 2007). *A. mellifera* orthology was determined by sequence similarity; BLASTX was used to determine sequence similarity of *V. squamosa* sequences to those in a database composed of proteins representing the complete *A. mellifera* official gene set (version 1; Honeybee Genome Sequencing Consortium, 2006). The best hit for each *V. squa-*

mosa sequence was taken and the corresponding *A. mellifera* gene identifier was used to assign a gene name. We then used *A. mellifera* gene IDs to assign relative gene expression in brains of *A. mellifera* queens and sterile workers (Grozinger *et al*., 2007) and in *A. mellifera* whole-bodies of queen-destined and worker-destined larvae (Barchuk *et al*., 2007).

Acknowledgements

This research was supported by the U. S. National Science Foundation (grant DEB 0640690). We thank J.L. Kovacs and S.V. Yi for helpful discussion and two anonymous reviewers for comments that helped to improve the manuscript.

References

Abouheif, E. and Wray, G.A. (2002) Evolution of the gene network underlying wing polyphenism in ants. *Science* **297**: 249–252.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J.H., Zhang, Z., Miller, W. *et al*. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**: 3389–3402.

Applied Biosystems (2001) Relative quantitation of gene expression. *User Bulletin #2: ABI Prism 7700 Sequence Detection System*.

Archer, M.E. (2006) Taxonomy, distribution and nesting biology of species of the genus *Dolichovespula* (Hymenoptera, Vespidae). *Entomol Sci* **9**: 281–293.

Aron, S., de Menten, L., Van Bockstaele, D.R., Blank, S.M. and Roisin, Y. (2005) When hymenopteran males reinvented diploidy. *Curr Biol* **15**: 824–827.

Ayroles, J.F., Carbone, M.A., Stone, E.A., Jordan, K.W., Lyman, R.F., Magwire, M.M. *et al*. (2009) Systems genetics of complex traits in *Drosophila melanogaster*. *Nat Genet* **41**: 299–307.

Barchuk, A.R., Cristino, A.S., Kucharski, R., Costa, L.F., Simoes, Z.L.P. and Maleszka, R. (2007) Molecular determinants of caste differentiation in the highly eusocial honeybee *Apis mellifera*. *BMC Dev Biol* **7**: 70.

Birchler, J.A., Riddle, N.C., Auger, D.L. and Veitia, R.A. (2005) Dosage balance in gene regulation: biological implications. *Trends Genet* **21**: 219–226.

Bonduriansky, R. and Chenoweth, S.F. (2009) Intralocus sexual conflict. *Trends Ecol Evol* **24**: 280–288.

Brisson, J.A. and Stern, D.L. (2006) The pea aphid, *Acyrthosiphon pisum*: an emerging genomic model system for ecological, developmental and evolutionary studies. *BioEssays* **28**: 747–755.

Brisson, J.A., Davis, G.K. and Stern, D.L. (2007) Common genome-wide patterns of transcript accumulation underlying the wing polyphenism and polymorphism in the pea aphid (*Acyrthosiphon pisum*). *Evol Dev* **9**: 338–346.

Carbon, S., Ireland, A., Mungall, C.J., Shu, S., Marshall, B., Lewis, S. and the AmiGO Hub and the Web Presence Working Group (2009) AmiGO: online access to ontology and annotation data. *Bioinformatics* **25**: 288–289.

Carroll, S.B. (2008) Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. *Cell* **134**: 25–36.

Corona, M., Velarde, R.A., Remolina, S., Moran-Lauter, A., Wang, Y., Hughes, K.A. *et al*. (2007) Vitellogenin, juvenile hormone, insulin signaling, and queen honey bee longevity. *Proc Natl Acad Sci U S A* **104**: 7128–7133.

Cristino, A.S., Nunes, F.M.F., Lobo, C.H., Bitondi, M.M.G., Simoes, Z.L.P., Costa, L.D. *et al*. (2006) Caste development and reproduction: a genome-wide analysis of hallmarks of insect eusociality. *Insect Mol Biol* **15**: 703–714.

Cutter, A.D. and Ward, S. (2005) Sexual and temporal dynamics of molecular evolution in *C. elegans* development. *Mol Biol Evol* **22**: 178–188.

Denver, D.R., Morris, K., Streelman, J.T., Kim, S.K., Lynch, M. and Thomas, W.K. (2005) The transcriptional consequences of mutation and natural selection in *Caenorhabditis elegans*. *Nat Genet* **37**: 544–548.

Edgar, B.A. and Orr-Weaver, T.L. (2001) Endoreplication cell cycles: more for less. *Cell* **105**: 297–306.

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

Evans, J.D. and Wheeler, D.E. (2001) Gene expression and the evolution of insect polyphenisms. *BioEssays* **23**: 62–68.

Foster, K.R., Ratnieks, F.L.W., Gyllenstrand, N. and Thoren, P.A. (2001) Colony kin structure and male production in *Dolichovespula* wasps. *Mol Ecol* **10**: 1003–1010.

Goodisman, M.A.D., Isoe, J., Wheeler, D.E. and Wells, M.A. (2005) Evolution of insect metamorphosis: a microarray-based study of larval and adult gene expression in the ant *Camponotus festinatus*. *Evolution* **59**: 858–870.

Goodisman, M.A.D., Kovacs, J.L. and Hunt, B.G. (2008) Functional genetics and genomics in ants (Hymenoptera: Formicidae): the interplay of genes and social life. *Myrmecol News* **11**: 107–117.

Graff, J., Jemielity, S., Parker, J.D., Parker, K.M. and Keller, L. (2007) Differential gene expression between adult queens and workers in the ant *Lasius niger*. *Mol Ecol* **16**: 675–683.

Greene, A. (1991) *Dolichovespula* and *Vespula*. In *The Social Biology of Wasps* (Ross, K.G. and Matthews, R.W., eds), pp. 263–305. Cornell University Press, Ithaca.

Grozinger, C.M., Fan, Y.L., Hoover, S.E.R. and Winston, M.L. (2007) Genome-wide analysis reveals differences in brain gene expression patterns associated with caste and reproductive status in honey bees (*Apis mellifera*). *Mol Ecol* **16**: 4837–4848.

Heid, C.A., Stevens, J., Livak, K.J. and Williams, P.M. (1996) Real time quantitative PCR. *Genome Res* **6**: 986–994.

Heimpel, G.E. and de Boer, J.G. (2008) Sex determination in the Hymenoptera. *Annu Rev Entomol* **53**: 209–230.

Hines, H.M., Hunt, J.H., O'Connor, T.K., Gillespie, J.J. and Cameron, S.A. (2007) Multigene phylogeny reveals eusociality evolved twice in vespid wasps. *Proc Natl Acad Sci U S A* **104**: 3295–3299.

Hoekstra, H.E. and Coyne, J.A. (2007) The locus of evolution: evo devo and the genetics of adaptation. *Evolution* **61**: 995–1016.

Hoffman, E.A. and Goodisman, M.A.D. (2007) Gene expression and the evolution of phenotypic diversity in social wasps. *BMC Biology* **5**: 23.

Honeybee Genome Sequencing Consortium (2006) Insights into social insects from the genome of the honeybee *Apis mellifera*. *Nature* **443**: 931–949.

Huang, X.Q. and Madan, A. (1999) CAP3: a DNA sequence assembly program. *Genome Res* **9**: 868–877.

Hughes, W.O.H., Oldroyd, B.P., Beekman, M. and Ratnieks, F.L.W. (2008) Ancestral monogamy shows kin selection is key to the evolution of eusociality. *Science* **320**: 1213–1216.

Hunt, J.H. (2007) *The Evolution of Social Wasps*. Oxford University Press, New York, NY.

Jablonka, E. and Raz, G. (2009) Transgenerational epigenetic inheritance: prevalence, mechanisms, and implications for the study of heredity and evolution. *Q Rev Biol* **84**: 131–176.

Jin, W., Riley, R.M., Wolfinger, R.D., White, K.P., Passador-Gurgel, G. and Gibson, G. (2001) The contributions of sex, genotype and age to transcriptional variance in *Drosophila melanogaster*. *Nat Genet* **29**: 389–395.

Khaitovich, P., Weiss, G., Lachmann, M., Hellmann, I., Enard, W., Muetzel, B. *et al*. (2004) A Neutral Model of Transcriptome Evolution. *Plos Biol* **2**: e132.

Kijimoto, T., Costello, J., Tang, Z.J., Moczek, A.P. and Andrews, J. (2009) EST and microarray analysis of horn development in *Onthophagus* beetles. *BMC Genomics* **10**: 504.

Kovacs, J.L. and Goodisman, M.A.D. (2007) Irregular brood patterns and worker reproduction in social wasps. *Naturwissenschaften* **94**: 1011–1014.

Macdonald, J.F. and Matthews, R.W. (1981) Nesting biology of the eastern yellowjacket, *Vespula maculifrons* (Hymenoptera, Vespidae). *J Kans Entomol Soc* **54**: 433–457.

Macdonald, J.F. and Matthews, R.W. (1984) Nesting biology of the southern yellowjacket, *Vespula squamosa* (Hymenoptera, Vespidae): social parasitism and independent founding. *J Kans Entomol Soc* **57**: 134–151.

Mank, J.E. (2009) The W, X, Y and Z of sex-chromosome dosage compensation. *Trends Genet* **25**: 226–233.

Mank, J.E. and Ellegren, H. (2009) Are sex-biased genes more dispensable? *Biol Lett* **5**: 409–412.

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

Meiklejohn, C.D., Parsch, J., Ranz, J.M. and Hartl, D.L. (2003) Rapid evolution of male-biased gene expression in *Drosophila*. *Proc Natl Acad Sci U S A* **100**: 9894–9899.

Mileyko, Y., Joh, R.I. and Weitz, J.S. (2008) Small-scale copy number variation and large-scale changes in gene expression. *Proc Natl Acad Sci U S A* **105**: 16659–16664.

Moczek, A.P. (2007) Developmental capacitance, genetic accommodation, and adaptive evolution. *Evol Dev* **9**: 299–305.

Moczek, A.P., Andrews, J., Kijimoto, T., Yerushalmi, Y. and Rose, D.J. (2007) Emerging model systems in evo-devo: horned beetles and the origins of diversity. *Evol Dev* **9**: 323–328.

Nijhout, H.F. (2003) Development and evolution of adaptive polyphenisms. *Evol Dev* **5**: 9–18.

Oleksiak, M.F., Churchill, G.A. and Crawford, D.L. (2002) Variation in gene expression within and among natural populations. *Nat Genet* **32**: 261–266.

Orr, W.C. and Sohal, R.S. (1994) Extension of life-span by overexpression of superoxide dismutase and catalase in *Drosophila melanogaster*. *Science* **263**: 1128–1130.

Parker, J.D., Parker, K.M., Sohal, B.H., Sohal, R.S. and Keller, L. (2004) Decreased expression of Cu-Zn superoxide dismutase 1 in ants with extreme lifespan. *Proc Natl Acad Sci U S A* **101**: 3486–3489.

Pigliucci, M., Murren, C.J. and Schlichting, C.D. (2006) Phenotypic plasticity and evolution by genetic assimilation. *J Exp Biol* **209**: 2362–2367.

R Development Core Team (2008) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org. Accessed in 2008.

Ranz, J.M., Castillo-Davis, C.I., Meiklejohn, C.D. and Hartl, D.L. (2003) Sex-dependent gene expression and evolution of the *Drosophila* transcriptome. *Science* **300**: 1742–1745.

Rasband, W.S. (1997–2009) *ImageJ*. U. S. National Institutes of Health, Bethesda, Maryland, USA. http://rsb.info.nih.gov/ij/. Accessed in 2008.

Rifkin, S.A., Houle, D., Kim, J. and White, K.P. (2005) A mutation accumulation assay reveals a broad capacity for rapid evolution of gene expression. *Nature* **438**: 220–223.

Rozen, S. and Skaletsky, H. (2000) Primer3 on the WWW for general users and for biologist programmers. *Methods Mol Biol* **132**: 365–386.

Scharf, M.E., Wu-Scharf, D., Zhou, X., Pittendrigh, B.R. and Bennett, G.W. (2005) Gene expression profiles among immature and adult reproductive castes of the termite *Reticulitermes flavipes*. *Insect Mol Biol* **14**: 31–44.

Smit, A., Hubley, R. and Green, P. (1996–2004) *RepeatMasker Open-3.0*. http://www.repeatmasker.org. Accessed in 2008.

Smith, C.R., Toth, A.L., Suarez, A.V. and Robinson, G.E. (2008) Genetic and genomic analyses of the division of labour in insect societies. *Nat Rev Genet* **9**: 735–748.

Sumner, S. (2006) Determining the molecular basis of sociality in insects: progress, prospects and potential in sociogenomics. *Ann Zool Fenn* **43**: 423–442.

Sumner, S., Pereboom, J.J.M. and Jordan, W.C. (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, A.L. and Robinson, G.E. (2007) Evo-devo and the evolution of social behavior. *Trends Genet* **23**: 334–341.

Toth, A.L., Varala, K., Newman, T.C., Miguez, F.E., Hutchison, S.K., Willoughby, D.A. *et al*. (2007) Wasp gene expression supports an evolutionary link between maternal behavior and eusociality. *Science* **318**: 441–444.

Toth, A., Bilof, K., Henshaw, M., Hunt, J. and Robinson, G. (2009) Lipid stores, ovary development, and brain gene expression in *Polistes metricus* females. *Insectes Soc* **56**: 77–84.

True, J.R. and Haag, E.S. (2001) Developmental system drift and flexibility in evolutionary trajectories. *Evol Dev* **3**: 109–119.

Turillazzi, S. and West-Eberhard, M.J. (1996) *Natural History and Evolution of Paper-Wasps*. Oxford University Press, New York, NY.

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

West-Eberhard, M.J. (2003) *Developmental Plasticity and Evolution*. Oxford University Press, New York, NY.

Wheeler, D.E. (1986) Developmental and physiological determinants of caste in social Hymenoptera: evolutionary implications. *Am Nat* **128**: 13–34.

Wilson, E.O. (1971) *The Insect Societies*. Harvard University Press, Cambridge.

Wray, G.A. and Abouheif, E. (1998) When is homology not homology? *Curr Opin Genet Dev* **8**: 675–680.

Zhang, Y., Sturgill, D., Parisi, M., Kumar, S. and Oliver, B. (2007) Constraint and turnover in sex-biased gene expression in the genus *Drosophila*. *Nature* **450**: 233–238.

Supporting Information

Additional Supporting Information may be found in the online version of this article under the DOI reference: DOI 10.1111/j.1365-2583.2010.1021.x

Table S1. Results of statistical analysis testing if gene expression levels differ among castes or among individuals within castes of each species.

Table S2. Prior data informing selection of loci for primer development in the present study.

Table S3. Gene ontology biological process annotation.

Table S4. Primers used for amplifying gene regions in four vespid wasp species.

Please note: Neither the Editors nor Wiley-Blackwell are responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.