PROCEEDINGS OF THE ROYAL SOCIETY B

Proc. R. Soc. B (2006) **273**, 19–26
doi:10.1098/rspb.2005.3291
Published online 18 October 2005

Differential gene expression and phenotypic plasticity in behavioural castes of the primitively eusocial wasp, *Polistes canadensis*

Seirian Sumner[1,2,*], Jeffrey J. M. Pereboom[2,3] and William C. Jordan[2]

[1]*Smithsonian Tropical Research Institute, Apartado 2072, Balboa, Ancon, Republic of Panama*
[2]*Institute of Zoology, Zoological Society of London, Regent's Park, London NW1 4RY, UK*
[3]*Centre for Research and Conservation, Royal Zoological Society of Antwerp, Koningin Astridplein 26, 2018 Antwerp, Belgium*

Understanding how a single genome can produce a variety of different phenotypes is of fundamental importance in evolutionary and developmental biology. One of the most striking examples of phenotypic plasticity is the female caste system found in eusocial insects, where variation in reproductive (queens) and non-reproductive (workers) phenotypes results in a broad spectrum of caste types, ranging from behavioural through to morphological castes. Recent advances in genomic techniques allow novel comparisons on the nature of caste phenotypes to be made at the level of the genes in organisms for which there is little genome information, facilitating new approaches in studying social evolution and behaviour. Using the paper wasp *Polistes canadensis* as a model system, we investigated for the first time how behavioural castes in primitively eusocial insect societies are associated with differential expression of shared genes. We found that queens and newly emerged females express gene expression patterns that are distinct from each other whilst workers generally expressed intermediate patterns, as predicted by *Polistes* biology. We compared caste-associated genes in *P. canadensis* with those expressed in adult queens and workers of more advanced eusocial societies, which represent four independent origins of eusociality. Nine genes were conserved across the four taxa, although their patterns of expression and putative functions varied. Thus, we identify several genes that are putatively of evolutionary importance in the molecular biology that underlies a number of caste systems of independent evolutionary origin.

Keywords: paper wasp; social insect; caste determination; suppression subtractive hybridization

1. INTRODUCTION

An individual's phenotype is defined by its morphological, physiological and behavioural characteristics. Plasticity in these traits enables individuals to maximize their fitness in a variable environment, and thus plays an important role in adaptation, speciation and evolution (West-Eberhard 2002). Phenotypic adaptation is implemented at the level of the genes in response to fluctuating biotic and abiotic cues (Evans & Wheeler 2001; West-Eberhard 2002; Hofmann 2003). The environment of a social animal is particularly variable because of the additional influences of interactions with other group members. It is not surprising therefore that the eusocial insects (bees, wasps, ants and termites) exhibit some of the most extreme examples of phenotypic plasticity in the form of their female castes (West-Eberhard 1989; Hölldobler & Wilson 1990).

The eusocial insects exhibit a vast array of social complexity, encompassing varying degrees of caste differentiation. The most complex societies (e.g. honeybees and most ants) have morphologically distinct queen and worker castes that are determined pre-imaginally, and have sterile workers (cannot mate and produce diploid eggs). Castes in more simple societies (primitively eusocial insects, e.g. hover wasps and allodapine bees) are temporary, determined during adulthood and differ only in behaviour. Behavioural queen/worker castes are likely to represent the ancestral state of morphological castes, and thus provide insights into the evolution and origin of queen/worker castes that may not be obtained from studying advanced societies, where castes are more derived. As sociality has evolved multiple times in the social Hymenoptera (Wilson 1971), a key question is whether the genes associated with different caste systems are conserved across taxa and across different levels and origins of social complexity (Robinson 2002; Robinson & Ben-Shahar 2002). Differential gene expression amongst queen/worker castes has been studied previously in two species of advanced eusocial insects (the honeybee (*Apis mellifera*) and bumble-bee (*Bombus terrestris*) Corona *et al.* 1999; Evans & Wheeler 1999; Hepperle & Hartfelder 2001; Pereboom *et al.* 2005). Related studies have looked at the gene expression associated with worker behaviour in honeybee adults (Ben-Shahar *et al.* 2002; Kucharski & Maleszka 2002; Robinson 2002; Whitfield *et al.* 2003), wing polymorphism in ant queens and workers (Abouheif & Wray 2002; Tian *et al.* 2004), and soldier and worker caste differentiation in termites (Miura *et al.* 1999; Miura *et al.* 2003; Scharf *et al.* 2003). However, to date, there have been no studies on the relationship between genes and queen/worker castes in primitively eusocial insect societies.

Polistes paper wasps are primitively eusocial, exhibiting highly plastic behavioural phenotypes that are morphologically similar and determined during adulthood

* Author for correspondence (seirian.sumner@ioz.ac.uk).

Received 27 May 2005
Accepted 7 August 2005

(Reeve 1991). *Polistes* females may vary in size, but on emergence all adults are capable of mating and laying haploid (male) and diploid (female) eggs. Although a female can remain as a worker all her life, she may become a queen if the existing queen dies or is over thrown, or if she founds a new colony (West-Eberhard 1969; Hughes *et al.* 1987; Peters *et al.* 1995). Nest-mates compete to realise their reproductive potential, forming a hierarchy of social dominance. The lowest ranked females are newly emerged from their pupal cells. Workers are intermediately ranked; they forage and care for brood. Queens are the highest ranked, and lay eggs (West-Eberhard 1986; Giray *et al.* 2005). The social hierarchy is maintained by physical aggression from higher ranked nestmates (Hughes *et al.* 1987) or by age/order of arrival at the nest (Hughes & Strassman 1988). Caste membership therefore changes gradually during a female's lifetime. Delaying caste determination until adulthood and maintaining high caste plasticity throughout life enables individuals to exploit their highly dynamic social environment (Jeanne 1972).

We present the first data on the molecular basis of behavioural queen/worker caste polyphenisms in a primitively eusocial society. *Polistes canadensis* is an ideal species in which to study castes because it is restricted to the relatively aseasonal tropics where, in contrast to temperate species of *Polistes*, both workers and queens are long-lived (Pickering 1980), caste is not influenced by the time of year that a female emerges (Mead & Gabouriaut 1993), and size/nutrition is unlikely to influence caste as queens do not hibernate (Strassmann *et al.* 1984; Sullivan & Strassmann 1984). We first identified genes that were differentially expressed between pairs of the three behavioural castes within *P. canadensis* (young females, *Y*; workers, *W*; and queens, *Q*). Next, we tested the hypothesis that the developmental gradient of young females to workers and workers to queens involves the gradual change in regulation of the same set of genes by comparing expression patterns of those genes identified as differentially expressed between one caste pair in other caste pairs. Lastly, we tested the hypothesis that the genes underlying behavioural castes (found in *Polistes*) are conserved in adults of advanced eusocial taxa (e.g. *Apis* and *Bombus*).

2. MATERIAL AND METHODS

(a) *Sample collection*

In March 2002, *P. canadensis* colony members from 12 nests in Panama's Canal Zone were individually marked. Nests were post-emergence (i.e. had already produced adult offspring) with nine females ± 0.24 (s.e.). Censuses of nestmates were taken twice a day, every 3–5 days for one month, and behaviour recorded during 1 h observation periods on most census days. Queens were behaviourally dominant and the only egg-layer on their nests, and workers were low ranked females who foraged (West-Eberhard 1986). There was no queen turnover during the monitoring period. After one month queens and workers were removed directly from their nests using forceps and stored immediately either in liquid nitrogen (N_2) or RNAlater reagent (Ambion), and subsequently at $-80\,^{\circ}C$ (for N_2 samples) or $-20\,^{\circ}C$ (for RNAlater samples). Young females were preserved in the same way within 12 h of emergence on collected nest combs.

(b) *Finding differentially expressed genes*

(i) *Construction of enriched libraries*

Suppression subtractive hybridization (SSH) was used to identify differentially expressed genes (Diatchenko *et al.* 1996). Total RNA was extracted from five individuals of each phenotype (Trizol, Invitrogen) and purified into Messenger RNA (mRNA) (Nucleotrap kit, BD Clontech). Complementary DNA (cDNA) was synthesized from 2 μg mRNA, pooled from all five individuals per caste. Forward and reverse subtractions for each caste pair were performed following the PCR-Select cDNA subtraction protocol (BD Clontech). PCR-amplified products were cloned using pCR2.1 vector and grown in microtitre plates under standard conditions (Invitrogen).

(ii) *Differential screening and selection of candidate genes*

Differential screening eliminates false positives (Jin *et al.* 1997). Duplicate macroarrays, consisting of 960 clones per paired library, were constructed on nylon membranes using standard techniques (Hybond NX, Amersham). Subtracted cDNA probes derived from the SSH, were radio labelled (alpha-P^{32} dATPs) and hybridized to the macroarray membranes at 65 °C in Church buffer overnight (Sambrook & Russell 2001). Membranes were subsequently washed in SSC/SDS solution following the manufacturer's instructions (Amersham), and exposed to X-ray film (Kodak BioMax MS) for 3–72 h. Each macroarray was stripped and hybridized in turn with both members of its library's paired subtraction. Clones were chosen for sequencing if the product that hybridized with the opposite caste was at least twofold weaker than the hybridization product resulting with their own sample. Homologous matches with sequences in online databases (GenBank, Swiss-Prot, EMBL) were identified using BLAST-X and BLAST-N algorithms.

(iii) *Quantification of differential expression*

Total RNA was extracted from a further 10 individuals of each phenotype. Northern blots containing 15 μg total RNA from pooled samples (five individuals per phenotype) were constructed. RNA was quantified prior to loading using a spectrophotometer, and again after running the gel by recording the fluorescence of the mRNA smear for each sample using a CCD camera and Syngene GENEGENIUS software. Probes were prepared from the chosen clones by PCR amplification and isolation of the amplified band (QIAquick gel extraction kit, Qiagen), and then radio labelled and hybridized to blots overnight at 65 °C. Blots were washed and exposed to X-ray film as described above. Peak heights were quantified from computer images of the scanned autoradiographs, corrected for the amount of RNA loaded on the gel (see above), and reported as ratios normalized across the three phenotypes to the phenotype showing strongest expression.

3. RESULTS

Clones were chosen at random from those showing the strongest differential expression and 173 were successfully sequenced. Nine genes were present in more than one clone, with a maximum of eight clones containing the same sequence. Homologues with sequences on web-based databases were obtained for 105 sequences. Forty-three sequences were chosen for northern blot

Table 1. Normalized relative expression levels for 43 genes screened in young females, workers and queens of *Polistes canadensis*, ranging from 0 (no expression detected), through intermediate values that indicate expression level as a proportion of the caste with the highest expression (i.e. 1). (Gene identity was inferred from the closest homologous match with sequences in other organisms (given in 'related species' column, with the probability of finding homologues by chance indicated by the '*E*-value' column). Functional group for each gene was determined using Gene Ontology and literature specific to the Genbank homologues. The letters in the last column indicate the relevant graph in figure 2, equal expression (less than twofold expression difference) across castes (=), or gave no detectable signal (n). Asterisk indicates clones which were subjected to Northern analysis in duplicate using separate blots, all of which gave comparable results. The 43 sequences have been deposited in Genbank (dbEST) under accession numbers DT319104 (dbEST Id 30947136) to DT319146 (dbEST Id 30947178).)

EST number	putative identity	related species; accession number (*E*-value)	functional group	NE	W	Q	figure 2 panel
1	egg-derived tyrosine-like	*Anopheles gambiae*; EAA43538 (7×10^{-4})	binding protein	0	0	1	(*a*)
2	poly-ADP-ribose polymerase	*A. gambiae*; Q7QBC7 (5×10^{-52})	metabolic enzyme	0	0	1	(*a*)
3	CG7291*	*Drosophila melanogaster*; Q9VO62 (6×10^{-9})	hypothetical protein	0	0	1	(*a*)
4	unknown (probe 34)	*none*	unknown	0	0	1	(*a*)
5	ENSANGP07937	*A. gambiae*; XP_321166.21 (2×10^{-7})	metabolic enzyme	0.16	1	0.68	(*a*)
6	fatty acid synthase/P270*	*Bombyx mori*; Q01678 (6×10^{-38})	metabolic enzyme	0.16	1	0.69	(*a*)
7	heat shock 70 kDa protein	*Manduca sexta*; Q9U639 (7×10^{-68})	RNA processing	0.46	1	0.78	(*a*)
8	hypothetical protein	*Dictyostelium discoideum*; Q54E26 (0.002)	unknown	0.35	1	0.81	(*a*)
9	imaginal disc growth factor	*Diaprepes abbreviatus*; Q5PXZ1 (5×10^{-29})	growth factor	0.19	1	0.91	(*a*)
10	transferrin	*Apis mellifera*; Q86PH6 (6×10^{-29})	binding protein	0.15	0.12	1	(*a*)
11	phospholipase C-like	*Plasmodium falciparum*; AEO14831 (0.008)	metabolic enzyme	0	0.2	1	(*a*)
12	porin*	*A. gambiae*; Q8T4K0 (2×10^{-19})	membrane protein	0.43	0.36	1	(*a*)
13	alpha-mannosidase	*Mus musculus*; AK075650 (2×10^{-11})	protein metabolism	0	0.21	1	(*a*)
14	vitellogenin	*A. mellifera*; AJ517411 (2×10^{-13})	storage protein	0	0.13	1	(*a*)
15	tubulin alpha-1 chain*	*D. melanogaster*; PO663 (4×10^{-53})	structural protein	0	0.37	1	(*a*)
16	CG30051	*D. melanogaster*; Q4V5Q0 (2×10^{-15})	unknown	0	0.88	1	(*a*)
17	pol protein	*Solenopsis invicta*; Q8ITE3 (2×10^{-35})	transposable element	1	0	0	(*b*)
18	unknown (probe 30)	*none*	unknown	1	0	0	(*b*)
19	unknown (probe 45)	*none*	unknown	1	0	0	(*b*)
20	myosin regulatory light chain	*D. melanogaster*; P18432 (2×10^{-43})	structural protein	1	0.28	0	(*b*)
21	penelope transposable element	*Drosophila virilis*; U49102 (1×10^{-6})	transposable element	1	0.19	0	(*b*)
22	y-box protein	*B. mori*; Q6F6B1 (7×10^{-11})	RNA processing	1	0	0.1	(*b*)
23	ENSANG12317	*A. gambiae*; Q7PVG9 (8×10^{-4})	binding protein	1	0.88	0.23	(*b*)
24	troponin C*	*D. melanogaster*; P47948 (1×10^{-7})	binding protein	1	0.59	0.29	(*b*)
25	protease inhibitor	*B. mori*; Q8T7L9 (4×10^{-9})	inhibitory enzyme	1	0.51	0.31	(*b*)
26	apolipophorin	*S. invicta*; Q6B972 (6×10^{-2})	binding protein	0.8	1	0.35	(*b*)
27	prolyl endopeptidase	*D. melanogaster*; Q9VKW5 (7×10^{-32})	metabolic enzyme	0.6	0	1	(*c*)
28	MRJP	*A. mellifera*; O18330 (5×10^{-9})	pigmentation protein	0.34	0	1	(*c*)

(*Continued.*)

Table 1. (*Continued.*)

EST number	putative identity	related species; accession number (*E*-value)	functional group	NE	W	Q	figure 2 panel
29	hexamerin 2	*Camponotus festinates*; Q9U5Y8 (3×10^{-26})	storage protein	0.25	0	1	(*c*)
30	CG31605 - PA	*D. melanogaster*; Q96114 (7×10^{-9})	unknown	1	0	0.63	(*c*)
31	CG8057 (AMP-activated protein kinase)*	*D. melanogaster*; Q9V541 (3×10^{-48})	metabolic enzyme	0.1	1	0.19	(*d*)
32	arrestin	*Locusta migratoria*; P32122 (3×10^{-75})	binding protein	0.23	1	0.45	(*d*)
33	p24 delta1 cargo receptor*	*Xenopus laevis*; Q9PT59 (5×10^{-21})	protein transport	0.59	0.57	1	=
34	ENSANGP18540	*A. gambiae*; Q7Q1W2 (6×10^{-9})	unknown	0.89	0.92	1	=
35	60S ribosomal protein	*Spodoptera frugiperda*; Q962T2 (2×10^{-31})	RNA processing	0.95	0.83	1	=
36	cytochrome c oxidase	*Stenoponia americana*; Q8M3BO (4×10^{-48})	metabolic enzyme	0.73	1	0.9	=
37	ATPase subunit 6	*Ornithodoros moubata*; Q8HQK0 (5×10^{-4})	metabolic enzyme	0.74	1	0.75	=
38	peroxiredoxin	*D. melanogaster*; Q9V5K7 (3×10^{-39})	metabolic enzyme	0.96	1	0.64	=
39	unknown (probe 40)	*none*	unknown	1	0.67	0.64	=
40	hymenoptaecin precursor*	*A. mellifera*; Q10416 (4×10^{-10})	antibacterial poly-peptide	*	*	*	n
41	nebulin	*Homo sapiens*; Q9Y5Z1 (3×10^{-39})	binding protein	*	*	*	n
42	P450 cytochrome precursor	*D. melanogaster*; Q9V6D6 (1×10^{-35})	metabolic enzyme	*	*	*	n
43	bacteriophage S13	*Bacteriophage S13*; PO7931 (2×10^{-20})	viral component	*	*	*	n

analysis (see table 1). Choice of sequences was based on three criteria: (i) all sequences showing significant homology to genes known to be differentially expressed in castes of other social Hymenoptera ($n=9$); (ii) a set of sequences showing significant homology to genes not known to be associated with caste ($n=30$); (iii) a set of sequences with no known homologues ($n=4$). Thirty-nine genes gave a detectable signal and over half of these were detected in all three phenotypes to varying degrees. Seven genes were equally expressed across all phenotypes (genes 33–39; i.e. have less than twofold differences in expression). Seven were detected in only one of the phenotypes: four were detected only in queens (genes 1–4), three in young females (genes 17–19), but none were detected only in workers. Nine genes were subjected to northern blot analysis for all three phenotypes on replicate blots containing independently isolated mRNAs (see table 1). All replicates gave highly comparable results (paired *t*-test $t_{(0.05(2)27)}$ $p=0.102$).

To determine how distinct the patterns of differential gene expression were between different phenotypes we analysed the correlation of expression levels between pairs of phenotypes of each of the 39 genes using Kendall partial rank–order correlations, which eliminates the possibility that a correlation is due to the association between any two phenotypes with the third (Seigel & Castellan 1988). The overall expression levels of genes in queens and young females were not significantly correlated ($\tau=-0.18$, $p>0.1$; figure 1*a*), indicating that these two phenotypes differ substantially in gene expression patterns. In contrast, there were significant correlations in gene expression between queens and workers ($\tau=0.31$, $p<0.005$; figure 1*b*) and between workers and young females ($\tau=0.38$, $p<0.001$; figure 1*c*), indicating that workers do not differ substantially from queens and young females in gene expression.

We further assessed how gene expression patterns changed with the phenotypic transitions that commonly occur in a *P. canadensis* colony by looking for trends in the direction of gene expression between young females, workers and queens. The majority of the genes (81% (26/32)) were either up- or down-regulated (more than twofold expression difference) with increasing social status (figure 2*a* ($n=16$), *b* ($n=10$), respectively), indicating that workers generally express caste-associated genes at levels that are intermediate between those of young females and queens. Only six genes deviated from this pattern, either being not detected (figure 2*c*, $n=4$) or up-regulated in workers (figure 2*d*, $n=2$). Some of the expression differences might be explained in terms of the genes' functional groups (table 2). All of the genes associated with metabolism in our sample had expression levels that were twofold (or more) higher in queens than in young females ($n=7$ genes, Fisher's exact $p=0.00058$),

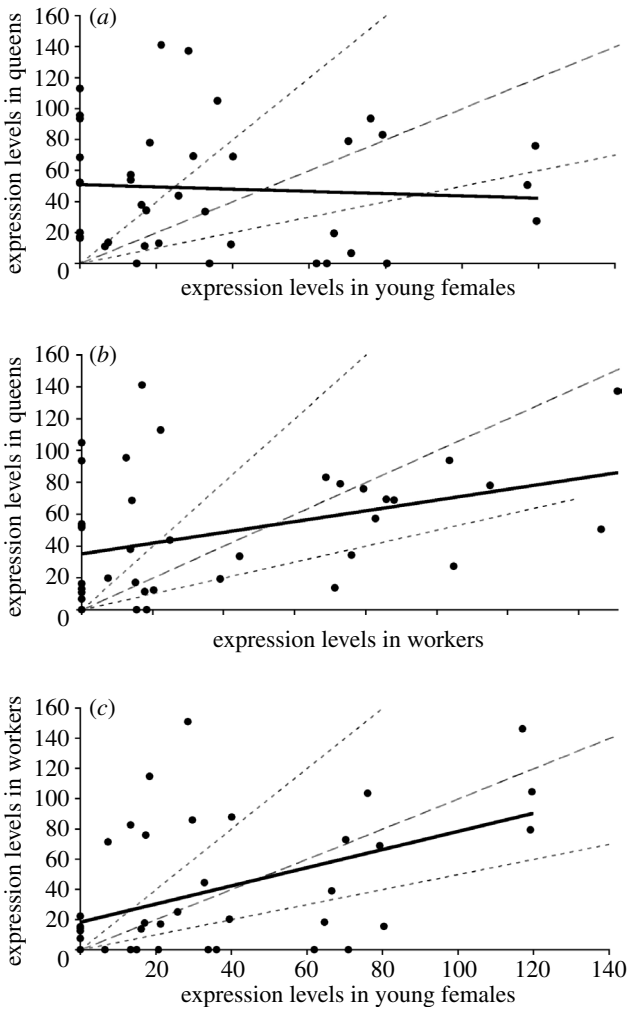

Figure 1. Scatter plots and regression lines (solid) for raw expression levels (corrected for amount loaded on the gel) of genes between the different caste pairs. The dashed line shows equal expression between the two phenotypes. The dotted lines show the upper and lower limits of twofold expression levels.

indicating that queens over express metabolic genes relative to young females.

We investigated whether any of the genes we have identified in our study were also differentially expressed in adult castes of eusocial insects in more complex (socially advanced) societies where queen and worker castes differ in morphology. We excluded larval gene expression studies because gene functions in larvae and adults stages are not comparable (Pereboom *et al.* 2005). Gene expression data for adult castes were available for three species: two bees (*A. mellifera* and *B. terrestris*) and one ant (the fire ant, *Solenopsis invicta*). Eight of the genes we found in *P. canadensis* were differentially expressed between castes of one or more of these species (see table 3). Another gene, hexamerin 2, was expressed equally in alate and dealate queens of *S. invicta*. There were some consistent patterns of expression for particular genes by queens and workers of the different species. Worker specific genes included those associated with heat shock and imaginal disk development, which were up-regulated in foraging workers for both *P. canadensis* and *A. mellifera*, and peroxiredoxin which was worker expressed in both *P. canadensis* and *B. terrestris*. Cytochrome oxidase was associated with the absence of reproduction in *B. terrestris* workers, but up-regulated in both queens and workers of *P. canadensis*. Three genes associated with mature (dealate) queens in *S. invicta* were up-regulated by *P. canadensis* queens (vitellogenin, major royal jelly protein (MRJP) and hexamerin 2). Transferrin was up-regulated in virgin queens and foragers of *A. mellifera* but up-regulated by only queens in *P. canadensis*.

4. DISCUSSION

We have isolated 32 genes whose differential expression is associated with young female, worker and queen phenotypes in adult females of the paper wasp, *P. canadensis*. The social biology of a *Polistes* colony is governed by the ontogenetic transitions from young adult female to queen (see §1). Our first hypothesis was that the gene expression patterns associated with the three behavioural phenotypes would reflect these important biological changes. Our data support this. First, overall gene expression levels in workers correlated significantly with both queens and young females, whilst there was no significant correlation in gene expression levels between queens and young females (figure 1). Second, 81% of the caste-associated genes were either up- or down-regulated along the social gradient, such that workers usually express genes at levels that are intermediate between those of young females and queens (figure 2). Thirdly, only two genes were up-regulated in workers (figure 2*d*). These findings demonstrate that the worker expression patterns are intermediate to those of queens and young females, reflecting the flexible, temporary nature of the behavioural castes in *Polistes*. The workers used in our study had specialized in foraging. It remains to be seen whether individual *P. canadensis* workers who show a tendency to forage, nest guard or nurse brood show greater differentiation at the transcriptome level than is apparent from our study (Giray *et al.* 2005).

Young, newly emerged females in simple societies are likely to be analogous to the totipotent, early instar larvae in more complex societies, which have yet to embark on a queen or worker developmental pathway. The gene expression patterns of young females were more similar to workers than queens. This is consistent with the biology of *P. canadensis*, since in nature young females invariably become workers after emergence, rather than queens (Pickering 1980). However, there were some genes that were specifically up-regulated by young females, and these may be associated with either totipotency or the final stages of maturation. These genes include homologues of transposable elements, which are thought to be involved in mutagenesis in *Drosophila* (Kaminker *et al.* 2002).

Genes that putatively code for metabolic enzymes were up-regulated in queens, which is consistent with egg production being metabolically costly. Queens expressed nine genes that were not detected in workers, and up-regulated a further six genes relative to workers (table 1). Not all these genes are obviously associated with reproduction, suggesting that there is more to being a queen than simply producing eggs. For example, transferrin is up-regulated eightfold in queens relative to workers and has anti oxidant and antibacterial functions (Kucharski & Maleszka 2003). Caste determination mechanisms in the eusocial Hymenoptera are closely linked with changes in levels of ecdysteroids and

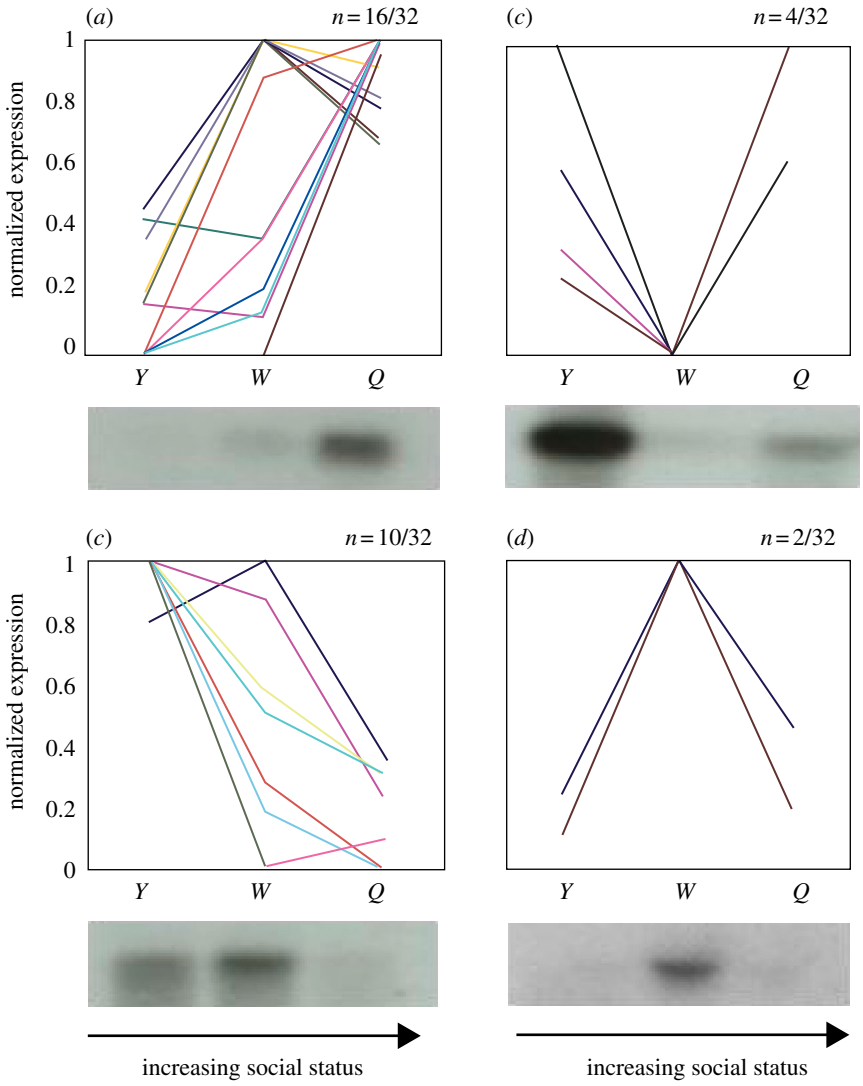

Figure 2. Normalized expression of differentially expressed genes in young females (*Y*), workers (*W*) and queens (*Q*). Patterns are determined using greater than or equal to twofold difference in expression between castes. Zero indicates no detectable expression, 1 indicates the phenotype expressing that gene at the highest level, and intermediate values show expression as a proportion of the caste with the highest expression. Each line represents a different gene.

juvenile hormone (JH) (Bloch *et al.* 2002). We found no genes that were obviously associated with production of these hormones in our study, although the uptake of vitellogenin is stimulated by the presence of JH (Bloch *et al.* 2002), and was up-regulated in *P. canadensis* queens.

Our second hypothesis was that genes associated with queen and worker behaviours in adult *P. canadensis* wasps would also be differentially expressed in other eusocial taxa. There is currently little gene expression data available for adult eusocial insects. We found data on two bees (*A. mellifera* and *B. terrestris*) and one ant (*S. invicta*) to compare with *P. canadensis*, although the latter examined only dealate and alate queens. Heat shock proteins are associated with stress, particularly exposure to high temperatures, as well as being molecular chaperones. These were up-regulated in foragers of *A. mellifera* relative to nursing workers (Whitfield *et al.* 2003), and in foragers of *P. canadensis* relative to queens, highlighting the potential importance of this family of genes in foraging behaviour. Vitellogenin was up-regulated in mature queens of *P. canadensis* and *S. invicta*, but in *A. mellifera* it was expressed by both queens and nurse (hive) workers.

Table 2. Mean normalized gene expression (mean±s.e.) for each caste within main functional groups.

functional group	number of genes	young	workers	queens
binding proteins	6	0.53±0.19	0.6±0.18	0.55±0.14
metabolic enzymes	10	0.41±0.11	0.62±0.13	0.78±0.08
protein processing	3	0.31±0.17	0.26±0.17	1±0
RNA processing	3	0.8±0.17	0.61±0.31	0.63±0.27
storage proteins	2	0.12±0.12	0.07±0.07	1±0
transposable elements	2	1±0	0.1±0.1	0

In *A. mellifera* vitellogenin is involved with brood-food production (jelly) (Amdam *et al.* 2003). Vitellogenin in *Polistes*, and probably in other primitively eusocial species, has not evolved such a complex role and is instead a queen-associated gene involved mainly in egg production

Table 3. Comparison of genes expressed by behavioural castes in *P. canadensis* with genes associated with morphological castes amongst adults of more complex eusocial insects. (In each case the phenotype with the highest expression level is indicated. 'Not expressed' indicates gene ortholog has been analysed but no signal was detected; 'n', no expression data available for adults.)

putative gene	*P. canadensis* (queen/forager)	*B. terrestris* (Pereboom *et al.* 2005; queen/worker; non-reproductive/ reproductive worker)	*A. mellifera* (forager/nurse)	*S. invicta* (Tian *et al.* 2004; dealate/alate queen)
heat shock proteins	forager[a]	n	forager (Whitfield *et al.* 2003)	n
imaginal disc development	forager	not expressed	forager (Whitfield *et al.* 2003)	n
vitellogenin	queen[a]	n	nurse/queen (Amdam *et al.* 2003)	dealate
MRJP	queen[a]	n	forager (Kucharski & Maleszka 2002)	dealate
hexamerin 2	queen[a]	not expressed	n	dealate/alate
60S ribosomal protein	queen	worker	n	n
cytochrome oxidase	forager	worker	n	dealate
peroxiredoxin	forager	queen/non-reproductive worker	n	n
transferrin	queen[a]	n	forager, virgin queen (Kucharski & Maleszka 2003)	n

[a] Twofold or more difference in expression between queens and workers.

(Giray *et al.* 2005). MRJP is involved in secretion of royal jelly in *A. mellifera* workers (Whitfield *et al.* 2003). In *P. candensis*, it was not detected in foragers but was up-regulated strongly in queens. This is comparable with *S. invicta* where MRJP-like genes are up-regulated in dealate queens, suggesting that they are associated with mature queen behaviour (Tian *et al.* 2004). Both MRJP and vitellogenin appear to be caste-associated genes in adults across taxa, but their functions in *Apis* have apparently diverged.

Polistes, *Solenopsis*, *Apis* and *Bombus* represent four independent origins of eusociality (Carpenter 1991; Cameron & Mardulyn 2001). Given the relatively small number of genes examined in any of these studies (with the exception of those on *Apis*) it was surprising to find 9 of the 32 genes identified in this study were also differentially expressed during caste determination across 2–4 of these taxa. These genes are putatively of evolutionary importance in queen and worker behaviours, since they have arisen independently in several eusocial taxa through either conservation or convergence. However, few genes have been examined in more than two species, there is limited evidence for consistency in the direction of caste-associated expression across taxa, and specific gene functions in the different taxa remain undetermined.

Our study provides the first base line data for sociogenomic studies on the proximate mechanisms underlying behavioural phenotypes in primitively eusocial insects. Within-family comparative studies of species sharing a common evolutionary ancestor, but showing a range of social complexity, promise to further our understanding of the molecular basis of caste evolution and phenotypic plasticity.

We thank M. Dijkstra and N. J. B. Isaac for assistance in the field, A. F. G. Bourke and M. J. West-Eberhard for comments, and W. T. Wcislo for logistical support. S.S. was supported by a Postdoctoral Fellowship and the research was funded by a Scholarly Studies Grant, both awarded by the Smithsonian Institute.

REFERENCES

Abouheif, E. & Wray, G. A. 2002 Evolution of the gene network underlying wing polyphenism in ants. *Science* **297**, 249–252. (doi:10.1126/science.1071468)

Amdam, G. V., Norberg, K., Hagen, A. & Omholt, S. W. 2003 Social exploitation of vitellogenin. *Proc. Natl Acad. Sci. USA* **100**, 1799–1802. (doi:10.1073/pnas.0333979100)

Ben-Shahar, Y., Robichon, A., Sokolowski, M. B. & Robinson, G. E. 2002 Influence of gene action across different time scales on behaviour. *Science* **296**, 741–744. (doi:10.1126/science.1069911)

Bloch, G., Wheeler, D. E. & Robinson, G. E. 2002 Endocrine influences on the organisation of insect societies. In *Hormones, brain behaviour* (ed. D. Pfaff, A. Arnold, A. Etgen, S. Fahrbach, R. Moss & R. Rubin), vol. 3, pp. 195–235. New York: Academic.

Cameron, S. A. & Mardulyn, P. 2001 Multiple molecular data sets suggest independent origins of highly eusocial behaviour in bees (Hymenoptera: Apinae). *Syst. Biol.* **50**, 194–214. (doi:10.1080/10635150151125851)

Carpenter, J. M. 1991 Phylogenetic relationships and the origin of social behaviour in the Vespidae. In *The social biology of wasps* (ed. K. C. Ross & R. W. Matthews), pp. 7–32. Ithaca, NY: Cornell University Press.

Corona, M., Estrada, E. & Zurita, M. 1999 Differential expression of mitochondrial genes between queens and workers during caste determination in the honeybee, *Apis mellifera*. *J. Exp. Biol.* **202**, 929–938.

Diatchenko, L. *et al.* 1996 Suppression subtractive hybridization: a method for generating differentially regulated or tissue-specific cDNA probes and libraries. *Proc. Natl Acad. Sci. USA* **93**, 6025–6030. (doi:10.1073/pnas.93.12.6025)

Evans, J. D. & Wheeler, D. E. 1999 Differential gene expression between developing queens and workers in the honey bee, *Apis mellifera*. *Proc. Natl Acad. Sci. USA* **96**, 5575–5580. (doi:10.1073/pnas.96.10.5575)

Evans, J. D. & Wheeler, D. E. 2001 Gene expression and the evolution of insect polyphenisms. *BioEssays* **23**, 62–68. (doi:10.1002/1521-1878(200101)23:1<62::AID-BIES1008>3.0.CO;2-7)

Giray, T., Giovanetti, M. & West-Eberhard, M. J. 2005 Juvenile hormone, reproduction, and worker behaviour in the neotropical social wasp *Polistes canadensis*. *Proc. Natl Acad. Sci. USA* **102**, 3330–3335. (doi:10.1073/pnas.0409560102)

Hepperle, C. & Hartfelder, K. 2001 Differentially expressed regulatory genes in honey bee caste development. *Naturwissenschaften* **88**, 113–116. (doi:10.1007/s001140000196)

Hofmann, H. 2003 Functional genomics of neural and behavioural plasticity. *J Neurobiol.* **54**, 272–282. (doi:10.1002/neu.10172)

Hölldobler, B. & Wilson, E. O. 1990 *The ants*. Berlin: Springer.

Hughes, C. R. & Strassman, J. E. 1988 Age is more important than size in determining dominance among workers in the primitively eusocial wasp, *Polistes instabilis*. *Behaviour* **107**, 1–15.

Hughes, C. R., Beck, M. O. & Strassmann, J. E. 1987 Queen succession in the social wasp, *Polistes annularis*. *Ethology* **76**, 124–132.

Jeanne, R. L. 1972 Social biology of the Neotropical wasp, *Mischocyttarus drewseni*. *Bull. Mus. Comp. Zool. Harv. Univ.* **144**, 63–150.

Jin, H., Cheng, X., Diatchenko, L., Siebert, P. D. & C, H. 1997 Differential screening of a subtracted cDNA library: a method to search for genes preferentially expressed in multiple tissues. *BioTech.* **23**, 1084–1087.

Kaminker, J. S., Bergman, C. M., Kronmiller, B., Carlson, J., Svirskas, R., Patel, S., Frise, E., Wheeler, D. A., Lewis, S. E., Rubin, G. M., Ashburner, M. & Celniker, S. E. 2002 The transposable elements of the *Drosophila melanogaster* euchromatin: a genomics perspective. *Genome Biol.* **3**, 0084.1–0084.20. (See http://www.fruitfly.org/about/pubs/genome_biology/gb-2002-3-12-research0084.pdf.)

Kucharski, R. & Maleszka, R. 2002 Evaluation of differential gene expression during behavioural development in the honeybee using microarrays and northern blots. *Genome Biol.* **3**. (doi:10.1186/gb-2002-3-2-research0007)

Kucharski, R. & Maleszka, R. 2003 Transcriptional profiling reveals mutilfunctional roles for transferrin in the honeybee, *Apis mellifera*. *J. Insect Sci.* **3**, 1–8.

Mead, F. & Gabouriaut, D. 1993 Post-eclosion sensitivity to social context in *Polistes dominulus* Christ females (Hymenoptera: Vespidae). *Insect. Soc.* **40**, 11–20. (doi:10.1007/BF01338829)

Miura, T., Kamikouchi, A., Takeuchi, H., Natori, S., Kubo, T. & Matsumoto, T. 1999 Soldier caste-specific gene expression in the mandibular glands of *Hodotermopsis japonica* (Isoptera: Termopsidae). *Proc. Natl Acad. Sci. USA* **96**, 13 874–13 879. (doi:10.1073/pnas.96.24.13874)

Miura, T., Koshikawa, S. & Matsumoto, T. 2003 Winged presoldiers induced by a juvenile hormone analog in *Zootermopsis nevadensis:* implications for plasticity and evolution of caste differentiation in termites. *J. Morphol.* **257**, 22–32. (doi:10.1002/jmor.10100)

Pereboom, J. J. M., Jordan, W. C., Sumner, S., Hammond, R. L. & Bourke, A. F. G. 2005 Differential gene expression in queen–worker caste determination in bumblebees. *Proc. R. Soc. B* **272**, 1145–1152. (doi:10.1098/rspb.2005.3060)

Peters, J. M., Queller, D. C., Strassmann, J. E. & Solis, C. R. 1995 Maternity assignment and queen replacement in a social wasp. *Proc. R. Soc. B* **260**, 7–12.

Pickering, J. 1980 Sex ratio, social behaviour and ecology in *Polistes*, *Pachysomoides* and *Plasmodium*. Ph.D. thesis, Harvard University, Cambridge, Massachusetts.

Reeve, H. 1991 In *Social biology of wasps* (ed. K. G. Ross & R. W. Matthews), pp. 99–148. Ithaca, NY: Cornell University Press.

Robinson, G. E. 2002 Genomics and intergrative analysis of division of labour in honeybee colonies. *Am. Nat.* **160**(Suppl.), S160–S172. (doi:10.1086/342901)

Robinson, G. E. & Ben-Shahar, Y. 2002 Social behaviour and comparative genomics: new genes or new gene regulation? *Genes Brain Behav.* **1**, 197–203. (doi:10.1034/j.1601-183X.2002.10401.x)

Sambrook, J. & Russell, D. W. 2001 *Molecular cloning: a laboratory manual*, 3rd edn. New York: Cold Spring Harbour Laboratory Press.

Scharf, M. E., Wu-Scharf, D., Pittendrigh, B. R. & Bennett, G. W. 2003 Caste- and development-associated gene expression in a lower termite. *Genome Biol.* **R62**, 1–11.

Seigel, S. & Castellan, N. J. 1988 *Non-parametric statistics for the behavioural sciences*. New York: McGraw-Hill.

Strassmann, J. E., Lee Jr, R. E., Rojas, R. R. & Baust, J. G. 1984 Caste and sex differences in cold-hardiness in the social wasps, *Polistes annularis* and *P. exclamans* (Hymenoptera: Vespidae). *Insect. Soc.* **31**, 291–301.

Sullivan, J. D. & Strassmann, J. E. 1984 Physical variability among nest foundresses in the polygynous social wasp, *Polistes annularis*. *Behav. Ecol. Sociobiol.* **15**, 249–256. (doi:10.1007/BF00292986)

Tian, H., Vinson, B. & Coates, C. J. 2004 Differential gene expression between alate and dealate queens in the red imported fire ant, *Solenopsis invicta* Buren (Hymenoptera: Formicidae). *Insect Biochem. Mol. Biol.* **34**, 937–949. (doi:10.1016/j.ibmb.2004.06.004)

West-Eberhard, M. J. 1969 The social biology of polistine wasps. *Misc. Publ. Mus. Zool. Univ. Mich.* **140**, 1–101.

West-Eberhard, M. J. 1986 Dominance relations in *Polistes canadensis* (L.), a tropical social wasp. *Monit. Zool. Ital. (NS)* **20**, 263–281.

West-Eberhard, M. J. 1989 Phenotypic plasticity and the origins of diversity. *Ann. Rev. Ecol. Syst.* **20**, 249–278. (doi:10.1146/annurev.es.20.110189.001341)

West-Eberhard, M. J. 2002 *Developmental plasticity and evolution*. New York: Oxford University Press.

Whitfield, C. W., Cziko, A. M. & Robinson, G. E. 2003 Gene expression profiles in the brain predict behaviour in individual honey bees. *Science* **302**, 296–299. (doi:10.1126/science.1086807)

Wilson, E. O. 1971 *The insect societies*. Cambridge, MA: Harvard University Press.

As this paper exceeds the maximum length normally permitted, the authors have agreed to contribute to production costs.