Gene expression patterns associated with caste and reproductive status in ants: worker-specific genes are more derived than queen-specific ones

B. FELDMEYER, D. ELSNER and S. FOITZIK
Evolutionary Biology, Institute of Zoology, Johannes Gutenberg University Mainz, Johannes von Müller Weg 6, 55128 Mainz, Germany

Abstract
Variation in gene expression leads to phenotypic diversity and plays a central role in caste differentiation of eusocial insect species. In social Hymenoptera, females with the same genetic background can develop into queens or workers, which are characterized by divergent morphologies, behaviours and lifespan. Moreover, many social insects exhibit behaviourally distinct worker castes, such as brood-tenders and foragers. Researchers have just started to explore which genes are differentially expressed to achieve this remarkable phenotypic plasticity. Although the queen is normally the only reproductive individual in the nest, following her removal, young brood-tending workers often develop ovaries and start to reproduce. Here, we make use of this ability in the ant Temnothorax longispinosus and compare gene expression patterns in the queens and three worker castes along a reproductive gradient. We found the largest expression differences between the queen and the worker castes (~2500 genes) and the smallest differences between infertile brood-tenders and foragers (~300 genes). The expression profile of fertile workers is more worker-like, but to a certain extent intermediate between the queen and the infertile worker castes. In contrast to the queen, a high number of differentially expressed genes in the worker castes are of unknown function, pointing to the derived status of hymenopteran workers within insects.

Keywords: genomics, phenotypic plasticity, social evolution, social insects, transcriptome

Received 24 May 2013; revision received 30 July 2013; accepted 1 August 2013

Introduction
Division of labour is the foundation for the ecological success of insect societies and fundamental to their social organization (Hölldobler & Wilson 1990). In most social insects, behavioural and morphological castes are not genetically determined, but develop through phenotypic plasticity. The queen is the main reproductive caste in the society and does not contribute to other tasks. The queens’ morphology and physiology are adapted specifically for reproduction. Queen lifespan is often greatly elongated compared with the relatively short-lived workers: ant queens can become up to 30 years old (Plateaux 1986; Keller & Genoud 1997).

In insect societies, division of labour occurs not only between queens and workers, but also among the different worker castes. Behavioural specialization in workers is regulated by variation in thresholds to perform certain tasks, which in turn are influenced by age, morphology, experience and genetics of workers (Hölldobler & Wilson 1990; Robinson 1992). Workers specialize in brood care, foraging or nest guarding, and in some species, this behavioural specialization is accompanied by distinct morphologies. For example, ant soldiers may weigh 100 times more than minor workers that perform brood care chores (Hölldobler & Wilson 1990).

In most social Hymenoptera, workers are not reproductively active in the presence of queen-derived chemical signals such as glandular secretions and cuticular hydrocarbons, although behavioural suppression of worker reproduction has been demonstrated in species with small colonies (Monnin 2006; Heinze & D’Ettorre 2009). If the queen dies or is removed, workers might start to develop their ovaries and to engage in
dominance interactions, and become reproductively active within a few weeks (Brunner et al. 2011).

While differentiation between hymenopteran males and females is mediated by the ploidy level (Heimpel & De Boer 2008), the pronounced differences between the diploid female castes are mostly due to external factors, such as the amount or quality of food and temperatures during larval development (Evans & Wheeler 2001a). Queen–worker and worker–worker caste differentiation are thus prime examples of polyphenism, in which the same genome can give rise to divergent phenotypes, characterized by different morphologies, behaviours and life histories (Wheeler 1986). Social insects are therefore ideal models to study the link between gene expression and phenotype, as expression differences in response to environmental signals lead to the development of distinct female castes (Gräff et al. 2007; Hunt et al. 2011; Gadau et al. 2012).

Understanding the genetic basis of caste differences will also give insights into the evolution of phenotypic plasticity in general. Not surprisingly, recent genomic studies have focused on this question by contrasting expression patterns of different castes and developmental stages (Pereboom et al. 2005; Sumner et al. 2006; Gräff et al. 2007; Grozinger et al. 2007; Bonasio et al. 2010; Cardoen et al. 2011; Colgan et al. 2011; Hunt et al. 2011). The emerging picture is that gene expression differences between castes are species specific, with no apparent universal caste-specific expression pattern. However, in two closely related fire ant species, expression differences were larger between developmental stages, sex and caste than between species (Ometto et al. 2011). The tools of genomics not only make it possible to identify caste-specific genes and thus help us to understand the phenotypic plasticity underlying caste differences, but also reveal the evolutionary history of castes (Barchuk et al. 2007; Johnson & Tsutsui 2011; Ferreira et al. 2013).

Although workers of the myrmicine ant genus Temnothorax are monomorphic, they show clear spatial and functional division of labour (Robinson et al. 2009). Worker task specialization is age dependent, with young workers tending the brood and older workers performing foraging duties (Robinson et al. 2009). Queen removal induces reproductive activity in some workers, which engage in dominance interactions and ovary development and start to lay haploid, male-destined eggs (Brunner et al. 2011; Konrad et al. 2012). In this study on the ant Temnothorax longispinosus, we exploit this feature by experimentally inducing the reproductive potential of workers through queen removal. This allows us to elucidate the genetic basis of division of labour along a reproductive gradient and will help us to differentiate between genes associated with fertility per se and those characteristic of queen and worker phenotypes.

Material and methods

Sample collection and behavioural caste determination

Monogynous colonies of Temnothorax longispinosus were collected at the E. N. Huyck Preserve, Rensselaerville, NY, USA (42.516619, -74.138925), in summer 2011. Because queens of this species are singly inseminated, all workers are full-sisters (Foitzik et al. 2004), indicating that genotypic differences are unlikely to influence the assignment of workers to different behavioural castes. Colonies were transferred to small plastic nests upon arrival in Mainz and kept under 15–20 °C 10:14 h night–day conditions. Ants were fed with crickets and honey once per week.

At least twenty workers with half of the brood were isolated from the rest of the colony to induce ovarian development. Workers in T. longispinosus are monomorphic, but can be grouped into distinct behavioural castes that occupy different locations in or outside the nest. After removing the queen, we colour-coded workers (N = 20 per colony, 11 colonies) and observed their behaviour and location twice per day for two weeks (for details, see Konrad et al. 2012) in order to assign them to their behavioural caste, for example brood-carer or forager. Foragers were defined as workers that spent 80% of the observations outside the nest, whereas workers that spent 90% of the observations in contact to the brood were classified as brood-carers. To determine the fertility status of brood-carers, they were cooled down for 20 min. at −20 °C and their ovaries were dissected on ice. Workers with developed ovaries and eggs in development were classified as ‘fertile’, whereas those with short, undeveloped ovaries without eggs were grouped into the ‘infertile’ brood-carer caste. Following reproductive status assessment, individuals were directly transferred to Trizol (for further details, see below).

RNA library preparation and sequencing

A pre-RNA extraction showed that optimally whole bodies of 10 workers and four queens had to be pooled to obtain sufficient RNA for library preparation and sequencing. An additional advantage of pooling several individuals is that it averages out individual gene expression variation, so that we can be more confident that expression differences are due to the different castes and not due to random differences between individuals. We tried to take two workers per caste per colony, where possible. Because not all castes were
available in equal numbers, RNA was extracted from eight fertile brood-carers (two from four colonies each), 10 infertile brood-carers and foragers (two from five colonies each), and four queens (one from four colonies each). Pooled individuals were thoroughly ground in 500 μl Trizol (Invitrogen) and frozen at −80 °C until further processing. For RNA isolation, 200 μl chloroform was added to each sample and the mixture shaken vigorously for 15 s. The mixture was then centrifuged for 15 min at 4 °C and 11 000 g. The upper aqueous phase was transferred to a new 1.5 ml RNase-free tube and precipitated with 200 μl of absolute ethanol. The solution was gently pipetted four times and transferred to an RNeasy mini-spin column (Qiagen). The further procedure followed step 3 onwards of the RNeasy clean-up manual (Qiagen). Illumina library preparation with individually marked (MID) libraries was performed through the sequencing facility affiliated with Mainz University. Three libraries were pooled and paired-end sequenced on one lane of an Illumina HiSeq 2000. The forager library was sequenced later in a similar manner. Original reads have been deposited in the short read archive under the study Accession no. PRJEB4368.

De novo transcriptome assembly and differential expression analyses

Illumina adapter removal was performed using the CLC Genomics Workbench software package (CLC bio). Sequence quality was checked using the program FastQC, version 0.10.1 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Based on this information, eight bases at the 5' end of each sequence were removed, terminal bases with ‘phred’ <20 were cut and sequences <30 bp removed, using an in-house python script. De novo reference transcriptome assembly based on reads of all libraries (CLC word sizes = 15, 25 and 60) resulted in suboptimal contig lengths (average ~400 bp; details not shown). We therefore decided to first assemble each library separately using the De Buijn graph-based CLC assembler (automated word size) followed by a meta-assembly using MIRA (Chevreux et al. 1999) (settings: job = denovo,genome,accurate,sanger).

These contigs were then used as reference for the gene expression analysis, again using the CLC workbench (standard settings). Expression levels were determined as reads per kilobase per mapped reads (RPKM) normalized for gene/contig length and library size (Mortazavi et al. 2008). Statistical significance ($P < 0.05$) of expression-level differences was inferred based on the Kal’s Z-test followed by FDR corrections, as implemented in CLC. Venn diagrams were built using the online available tool Venny (http://bioinfogp.cnb.csic.es/tools/venny).

To estimate dissimilarity of gene expression patterns between the four female castes, we ran a nonmetric multidimensional scaling (NMDS) analysis on the RPKM values of differentially expressed genes using the vegan package in R (www.R-project.org).

Functional annotation

The BLASTx program (Altschul et al. 1990) was used to BLASTx the contigs versus the nonredundant (nr) invertebrate protein database (state December 2012), with cut-off values set to $< e^{-5}$. Functional annotation and enrichment analyses were performed using the Blast2Go online tool with default parameters for the mapping and annotation procedure (Conesa et al. 2005). Assignment of gene ontology (GO) terms to the contigs was performed by importing the above-mentioned BLASTx search results against the nr-Prot database.

Comparison with other studies

Because there was little correspondence between direct BLASTs of *T. longispinosus* contigs and the contigs and transcripts of other studies on caste-specific gene expression patterns, we decided to compare our results on the basis of annotations. If several contigs with the same annotation existed, we chose the contig with the highest read count to infer significant expression differences (as outlined above).

Results

Assembly and BLAST

In total, between 26 and 140 Mio raw reads were obtained after sequencing, of which 20–132 Mio reads remained after trimming (Table S1, Supporting Information). The CLC assembly for each of the libraries resulted in about 55 000–97 000 contigs with an N50 of 824–1140 bp (Table S2, Supporting Information). The subsequent meta-assembly with MIRA resulted in 44 797 contigs with an average contig length of 1437 bp. A BLASTx of these contigs versus the nonredundant protein database gave 19 856 hits with $<e^{-5}$, of which 11 253 were single-gene hits. The tenth highest number of BLAST hits were found in hymenopteran species, with the top four being ants (Fig. S1, Supporting Information), thus giving us confidence in our contig quality.

General gene expression patterns and enrichment analyses

A total of 11 016 significant expression differences were found in pairwise comparisons between castes
(FDR-ᵰ < 0.05; fold change > 2), of which 5346 correspond to single genes. Ordination of the genes with significant expression differences (termed as differentially expressed genes) reveals close resemblance in expression patterns among foragers and infertile workers and substantial differences between fertile workers and queens (Fig 1a). This picture is also reflected in the patterns of differentially expressed genes (DEGs), which are either shared between castes or specifically expressed in only a single caste (private genes). The highest number of private DEGs was found in queens, with more than twice as many private DEGs as the worker castes (Fig 1b). The highest number of shared DEGs was detected between the three worker castes, indicating that worker gene expression to a large extent is independent of their behavioural caste. However, the fertile workers also shared approximately the same number of DEGs with the queens and the two worker castes.

A functional enrichment analysis of queen DEGs versus shared worker genes reflects the reproductive activity of the queens with an overrepresentation of genes belonging to RNA-binding, DNA-binding, ATP-binding and nucleus categories (Fig 2a). In the workers, metabolic processes as well as odorant binding activities are overrepresented. In addition, we contrasted DEGs shared between queens and fertile workers versus the genes shared by all worker castes and found nucleic acid and protein binding, ribonucleoprotein complex, protein folding and gene expression to be overrepresented in the queens and the fertile workers, whereas respiration-related categories are overrepresented in foragers and infertile workers (Fig 2b).

**Annotation rate of caste-specific genes**

A comparison of the annotation rates among private genes between the four different castes revealed differences in the number of genes annotated, with an 86% annotation rate in queens and significantly less in foragers (63%), fertile (55%) and infertile workers (55%) ($\chi^2 = 384.04; P < 0.001$). Annotation in equally expressed genes was even smaller with only 45% homology with described insect sequences, which is comparable with the annotation rate of the complete contig set (48%).

**Caste-specific expression patterns and enrichment analyses**

In contrast to the above general expression pattern analyses, we here investigated pairwise expression differences between castes in more detail (Fig 3a). By determining caste-specific genes, that is, genes that are exclusively regulated differentially in a single caste in comparison with all three other castes, we aimed at identifying genes that are representative for the properties of the focal caste. The highest number of these caste-specific genes was found in the queens (>1000), while the lowest number was found in the infertile workers (41) and foragers (51; Fig 3b). Between the queens and each single worker caste, only 10–20% of the differentially expressed genes were worker-caste dependent, whereas more than 80% were queen specific (Table 1). Worker gene expression, on the other hand, more closely resembles that of other castes, as indicated by the higher number of pair-specific and fewer caste-specific genes (Table 1).

In foragers, the enrichment analysis of the caste-specific genes did not result in any enriched functional
category. In the fertile workers, only two categories were enriched (isoprenoid metabolic process and isoprenoid biosynthetic process) (Fig. S2a+b, Supporting Information). Twenty enriched categories were identified for the infertile workers, of which several are related to metabolism (Fig. S2c, Supporting Information). In queens, 71 categories were enriched, with several involved in ribonucleotide binding, organelles, metabolic processes and protein catabolic processes.

**Vitellogenin**

Ants possess four copies of vitellogenin (Wurm et al. 2010), a multifunctional protein that functions as yolk precursor and is thus indicative for reproductive activity (Amdam et al. 2003). In *T. longispinosus*, expression patterns differ between the different gene copies and the different female castes (Fig 4). While *Vg*2 + 3, as well as the *Vg*-receptor, are overexpressed in the queens, these genes are least expressed in foragers. Conversely, *Vg1* is most highly expressed in the foragers and infertile workers. *Vg6* shows the highest expression level in the fertile workers, followed by the queens and the infertile workers, with lowest expression in the foragers.

**Comparison with other studies**

We found very few similarities between the caste-specific gene expression patterns in *T. longispinosus* and the patterns found in other published studies. Many differentially expressed genes found in other species could not be identified in our annotated contig list, and only few of these conformed to expression patterns (Table 2). Among the genes with similar expression pattern, fatty acid synthase was overexpressed in workers, a gene possibly involved in the synthesis of cuticular hydrocarbons. We also found that transferrin was overexpressed...
in the queens, a gene involved in the transport of iron ions from the hemolymph into the eggs during oogenesis (UniProt annotation). Several of the genes overexpressed in reproductive honeybee workers versus nonreproductive workers (Thompson et al. 2006) were overexpressed in the T. longispinosus queens, whereas the reproductive workers showed similarly low expression values compared with the two sterile worker castes for these genes.

**Discussion**

Our genomic study of the ant T. longispinosus reveals clear differences in gene expression between the four female castes. Queens had many more differentially expressed genes than any of the three worker castes. The fertile workers express mainly worker-caste-specific genes, but they also express several genes belonging to ribosomal-function-related categories, setting this caste apart from the other two worker castes.

A major finding of this study is the large number of worker genes (almost half of the overexpressed genes) without available annotation. This indicates the derived state of hymenopteran workers relative not only to other annotated insects, but also to the queen. The evolutionary history of social hymenopterans with solitary winged wasps as ancestors (Savard et al. 2006; Pilgrim et al. 2008) reflects this pattern on a morphological basis with queens as the ancestral form. In the age of genomics, it is now possible to identify genes and metabolic pathways involved in the development of workers from queens (Sumner et al. 2006; Gräff et al. 2007; Grozinger et al. 2007; Ferreira et al. 2013). In accordance with our results, recent studies of honeybees and primitive eusocial wasps suggest that numerous novel genes evolved in social hymenopterans and that the majority of these genes show worker-specific expression patterns (Barchuk et al. 2007; Ferreira et al. 2013). It appears that the evolution of eusociality lead not only to a decoupling of queen and worker phenotypes, but also to the rapid evolution of genes associated with these (Snell-Rood et al. 2011; Woodard et al. 2011). Given that eusociality evolved independently multiple times in the Hymenoptera (Wilson 1971), it is not surprising that different pathways and sets of genes are involved in caste differentiation in different eusocial lineages of the wasps, bees and ants (Pereboom et al. 2005; Gräff et al.

---

**Table 1** Number of significantly up-regulated genes (FDR- \( P < 0.05 \); fold change \( >2 \)) in the focal caste in comparison with the other castes and the number of pair-specific genes (genes up-regulated in only this specific comparison) with the corresponding percentage in parentheses.

<table>
<thead>
<tr>
<th>Focal caste</th>
<th>Comparison</th>
<th>Total</th>
<th>Pair-specific (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Queen</td>
<td>fer</td>
<td>1657</td>
<td>224 (13.52)</td>
</tr>
<tr>
<td></td>
<td>inf</td>
<td>1944</td>
<td>400 (20.58)</td>
</tr>
<tr>
<td></td>
<td>for</td>
<td>1683</td>
<td>169 (10.04)</td>
</tr>
<tr>
<td>Fertile worker</td>
<td>q</td>
<td>1128</td>
<td>779 (69.06)</td>
</tr>
<tr>
<td></td>
<td>inf</td>
<td>294</td>
<td>84 (28.57)</td>
</tr>
<tr>
<td></td>
<td>for</td>
<td>682</td>
<td>401 (58.80)</td>
</tr>
<tr>
<td>Infertile worker</td>
<td>q</td>
<td>937</td>
<td>734 (78.34)</td>
</tr>
<tr>
<td></td>
<td>fer</td>
<td>277</td>
<td>81 (29.24)</td>
</tr>
<tr>
<td></td>
<td>for</td>
<td>149</td>
<td>41 (27.52)</td>
</tr>
<tr>
<td>Forager</td>
<td>q</td>
<td>1010</td>
<td>723 (71.58)</td>
</tr>
<tr>
<td></td>
<td>fer</td>
<td>1072</td>
<td>824 (76.87)</td>
</tr>
<tr>
<td></td>
<td>inf</td>
<td>184</td>
<td>59 (32.07)</td>
</tr>
</tbody>
</table>

---

**Fig. 3** Overview of differentially expressed genes between castes. (a) Pairwise comparisons between castes with the number of differentially expressed genes (fold change \( >2 \); FDR- \( P < 0.05 \)) and (b) Number of caste-specific genes that are up-regulated in a single caste in comparison with all other castes (q = queen, fer = fertile worker, inf = infertile worker, for = forager).
In contrast, as the transition from solitary to group living presumably evolved only once in ants (Hölldobler & Wilson 1990), the diversity of expression patterns in the ants is unexpected. However, today’s ants exhibit extremely diverse lifestyles, ranging from the basal ponerines with their small colonies and low queen–worker dimorphism, to the socially parasitic inquilines, which have entirely lost the worker caste, to the highly eusocial species, such as the leaf-cutting ants with their morphologically diverse worker castes and pronounced differences between queens and workers in morphology, behaviour and lifespan (Hölldobler & Wilson 1990). This diversity indicates that selection has been acting in different directions with respect to caste differentiation in different ant taxa and may thus explain the variation in gene expression patterns between castes of different ant species. However, discrepancies in gene expression patterns found between different studies (Pereboom et al. 2005; Sumner et al. 2006; Graß et al. 2007; Cardoen et al. 2011) might partly also be due to different experimental set-ups (e.g. different rearing conditions, developmental stages, treatments, single tissue types versus whole bodies for RNA extraction) and sequencing techniques.

**General gene expression patterns**

In total, we identified more than 5000 differentially expressed genes in the four *T. longispinosus* female castes. The queen is the most distinct caste with many genes privately overexpressed and overrepresented on a functional basis. For example, the reproductive status of the queens is reflected in overrepresented categories such as translation, nucleus, ATP-, DNA- and RNA-binding. The gene expression patterns of the fertile workers are largely consistent with other worker castes, and only a few genes and functional categories are shared with the queens. This might be due to the short time period (four weeks) during which these workers could develop their ovaries. Thus, the observed pattern might reflect workers during reorganization, rather than the fully reproductive worker status. However, in honeybees, a similar pattern has been observed, where brain gene expression patterns between reproductive and sterile worker castes differed by only a few hundred genes, in comparison with more than 2000 genes between the queen and the worker castes (Grozinger et al. 2007). This suggests that ovary activation by itself does not lead to large modifications within the workers (Grozinger et al. 2007) and additionally that the large gene expression differences found between the queen and fertile workers may be due mainly to the queen status and associated differences such as longevity, specific pheromone and cuticular hydrocarbon production, and mating state.

In addition to morphological differences, social insect queens display distinct life history traits, such as extended lifespan. Previous gene expression studies in honeybees have provided conflicting results about which genes are involved in oxidative processes related to queen longevity (Evans & Wheeler 2001b; Corona et al. 2005; Grozinger et al. 2007). Whereas two studies found oxidation–reduction genes up-regulated in workers (Evans & Wheeler 2001b; Corona et al. 2005), another study found these genes to be up-regulated in the queen (Grozinger et al. 2007). We found oxidation genes in workers as well as queens. Oxidation–reduction processes and haem binding associated with detoxification are overrepresented in all three worker castes compared with the queen. However, two copies of vitellogenin are overexpressed in the queen, which have
Table 2 Overview of genes with caste-specific differential expression patterns between castes of different hymenopteran species, and the associated *T. longispinosus* expression pattern. The number of genes investigated corresponds to the number of annotated genes available in the corresponding study. (bold, genes with similar expression pattern between castes in *T. longispinosus* compared with the other species; ns, nonsignificant expression differences between *T. longispinosus* castes; q, queen; fer, fertile worker; inf, infertile worker; for, forager; st, sterile; w, worker; ne, newly emerged; non-rep, nonreproductive worker; rep, reproductive worker)

<table>
<thead>
<tr>
<th>Reference (matched genes/genes investigated)</th>
<th>Species/tissue for RNA extraction (comparison between castes)</th>
<th>Genes differentially expressed</th>
<th>Up-regulated in <em>T. longispinosus</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>Thompson et al. 2006 (4/8)</td>
<td><em>Apis mellifera</em>/brain, abdomen (sterile vs. reproductive workers)</td>
<td>Synapsin</td>
<td>st</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ubiquitin</td>
<td>st</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Myosin</td>
<td>st</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Solute carrier</td>
<td>st</td>
</tr>
<tr>
<td>Sumner et al. 2006 (15/29)</td>
<td><em>Polistes canadensis</em>/whole body (queens, newly emerged females, workers)</td>
<td>Fatty acid synthase</td>
<td>w</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Heat shock 70 kDa</td>
<td>w</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Transferrin</td>
<td>q</td>
</tr>
<tr>
<td></td>
<td></td>
<td>alpha-mannosidase</td>
<td>q</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Vitellogenin</td>
<td>q</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tubulin alpha-1 protein</td>
<td>q</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Myosin regulatory</td>
<td>ne</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Troponin C</td>
<td>ne</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Apolipoporphin</td>
<td>w</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Prolyl endopeptidase</td>
<td>q</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Arrestin</td>
<td>w</td>
</tr>
<tr>
<td></td>
<td></td>
<td>60S ribosomal protein</td>
<td>q</td>
</tr>
<tr>
<td>Gräff et al. 2007 (2/6)</td>
<td><em>Lasius niger</em>/whole body (queen vs. worker)</td>
<td>Cytochrome c oxidase</td>
<td>w</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Peroxiredoxin</td>
<td>w</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Vitellogenin</td>
<td>q</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Yellow-g2</td>
<td>q</td>
</tr>
<tr>
<td>Grozinger et al. 2007 (2/10)</td>
<td><em>Apis mellifera</em>/brain (queen, sterile and fertile worker)</td>
<td>Insulin receptor</td>
<td>q</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Transferrin</td>
<td>q</td>
</tr>
<tr>
<td>Bonasio et al. 2010 (0/5)</td>
<td><em>Harpegnathos saltator</em>/whole body (gamergate vs. non-reprod worker)</td>
<td>None</td>
<td></td>
</tr>
<tr>
<td>Toth et al. 2010; (6/28)</td>
<td><em>Polistes metricus</em>/brain (candidate genes associated with foraging/provisioning)</td>
<td>Foraging</td>
<td>ns</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Heat shock factor</td>
<td>ns</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Turtle</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Swiss cheese</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cytochrome p450 reductase</td>
<td></td>
</tr>
<tr>
<td>Cardoen et al. 2011; * (63/1.292)</td>
<td><em>Apis mellifera</em>/whole body (reproductive vs. nonreproductive worker)</td>
<td>Myosin regulatory light chain 2</td>
<td>non-rep</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Actin-related protein 1</td>
<td>non-rep</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ras-like gtp-binding protein rho1</td>
<td>non-rep</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Arginine kinase</td>
<td>non-rep</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Histone rna hairpin-binding protein</td>
<td>rep</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Replication factor c subunit 4</td>
<td>rep</td>
</tr>
<tr>
<td></td>
<td></td>
<td>gtp-binding protein 128up</td>
<td>rep</td>
</tr>
<tr>
<td></td>
<td></td>
<td>alp-dependent rna helicase vasa</td>
<td>rep</td>
</tr>
</tbody>
</table>

© 2013 John Wiley & Sons Ltd
Cuticular hydrocarbons play a crucial role in nestmate recognition, but also in fertility signalling, with queens carrying larger quantities and also longer chained hydrocarbons compared with workers (Heinze & D’Ettorre 2009).

G-protein-coupled receptors are a large family of receptors that include olfactory receptors, which are known to be involved in behavioural responses and chemical communication (Hildebrand & Shepherd 1997). For example, in A. mellifera, the odorant receptor 2 has been identified as a queen pheromone coreceptor (Wanner et al. 2007) and was shown to be up-regulated in the presence of queen mandibular pheromone (Grozinger et al. 2007), suggesting that it plays a role in mediating the regulation of worker reproduction. In our study, two G-protein-related categories are overrepresented in foragers compared with the queens, indicating a queen–forager interaction. However, this could also be caused by the exposure of foragers to additional stimuli outside the nest.

Caste-specific expressed genes

The distinctive status of the queen is demonstrated clearly by the 80% of genes up-regulated specifically in the queen in comparison with all three worker castes. In contrast, only 10–20% of the genes were specifically overexpressed in any of the three worker castes. Among the queen-specific genes, several are involved in reproduction, including oogenesis (maternal protein tudor), oocyte formation (maternal protein exuperantia) or chromosomal segregation in meiosis (claret segregational). In addition, genes associated with the synthesis of long cuticular hydrocarbons such as elongation of very long chain fatty acids protein, fatty acid synthase and acyl-CoA delta(11) desaturase were only expressed in queens. Cuticular hydrocarbons play a crucial role in nestmate recognition, but also in fertility signalling, with queens carrying larger quantities and also longer chained hydrocarbons compared with workers (Heinze & D’Ettorre 2009).

Vitellogenin

Vitellogenin is a multifunctional protein, primarily known for its function as a yolk precursor, produced by all oviparous animals (Amdam et al. 2003). The zinc-binding capacity of its protein product also allows it to act as an antioxidant (Seeuhus et al. 2006). In honeybees, the antioxidant function of vitellogenin has been associated with the regulation of lifespan (Seeuhus et al. 2006; Corona et al. 2007) and the division of labour (Nelson et al. 2007). Higher vitellogenin expression patterns in Lasius niger queens suggest that these functions might also hold for ants (Gräff et al. 2007). Recently, four vitellogenin copies were identified in the fire ant S. invicta (Wurm et al. 2010). These copies presumably evolved neo- or subfunctionalization to acquire caste-specific functions during the divergence from wasps to ants to allow for the complex properties of ant societies. The expression patterns of the different vitellogenin copies that we found for T. longispinosus castes resemble those found for Solenopsis invicta queens and workers (Wurm et al. 2010). In the fourth vitellogenin copy Vg6 (instead of Vg4 as in S. invicta), expression was highest in the fertile workers, followed by the queens and the infertile workers, with the lowest expression in the foragers. In contrast, hardly any expression of Vg4 was detected in S. invicta queens and only low expression in the workers (Wurm et al. 2010). Based on duplication events during the evolution of the different vitellogenin
copies and their new caste-specific roles, it is possible that the fourth vitellogenin copy in *T. longispinosus* serves a different or modified function compared with *S. invicta*, especially taking their different biologies into account.

**Conclusions**

Overall, we find pronounced differences in gene expression between the four female castes along a reproductive gradient in the ant *T. longispinosus*. The queen showed the most distinct expression pattern of all castes. All three worker castes displayed similar expression patterns, but the fertile workers shared more similarities with the queen than the other worker castes.

Our study emphasizes the derived status of the worker castes on a genetic level compared with the queen phenotype as workers showed a large number of novel differentially expressed genes without annotation. This finding supports the emerging picture from other gene expression studies that found a little overlap in expression patterns between taxa, suggesting different pathways during caste evolution in different hymenopteran lineages. Social hymenopterans are thus ideal models for studies on the evolution of phenotypic plasticity, the diversity of the underlying pathways and the functions of associated genes. Annotation of these unknown genes will bring many important insights, because such genes are likely to contain ecologically and evolutionarily interesting information (Tautz & Domazet-Lošo 2011).

**Acknowledgements**

This study was funded by a JGU start-up grant provided by Mainz University and a DFG grant (Fo 298/11-1). We thank three anonymous reviewers, Markus Pfenninger and Mathilde Cordellier for their helpful comments on the manuscript, and Elena Berg for English proofreading. Thanks to Andreas Modlmeier and Tobias Pamminger for help in ant collection and Tobias Haver for his support during the behavioural observations. We are grateful to Benjamin Rieger and Bastian Greshake for kindly providing some of the bioinformatics scripts.

**References**


© 2013 John Wiley & Sons Ltd


B.F. developed the project idea, designed and performed the experiment, analysed the data and drafted the manuscript. D.E. was involved in data analyses. S.F. contributed to ant collection, the development of the study idea, experimental design and data interpretation, and involved in writing of the manuscript.

**Data accessibility**

All sequence data for this study were archived at NCBI’s Short Read Archive (SRA) under study Accession no. PRJEB4368. Assembled contigs are archived on Dryad with accession doi:10.5061/dryad.85pd5. Gene expression data are provided as online supplemental material with this article.

**Supporting information**

Additional supporting information may be found in the online version of this article.

**Table S1** Summary of the sequencing output and the number of reads after trimming.

**Table S2** Summary statistics of caste-specific assemblies conducted in CLC and the following meta-assembly with MIRA.

**Table S3** Results of the gene expression analyses.

**Fig S1** Species distribution of the BLAST hits.

**Fig S2** Enriched functional categories of caste-specific genes.