MOLECULAR ECOLOGY

Molecular Ecology (2016) **25**, 3716–3730 doi: 10.1111/mec.13700

Gene expression and variation in social aggression by queens of the harvester ant *Pogonomyrmex californicus*

MARTIN HELMKAMPF,*[1] ALEXANDER S. MIKHEYEV,†[1] YUN KANG,‡ JENNIFER FEWELL* and JÜRGEN GADAU*

**School of Life Sciences, Arizona State University, 427 East Tyler Mall, Tempe, AZ 85287, USA, †Okinawa Institute of Science and Technology Graduate University, 1919-1 Tancha, Onna-son, Kunigami-gun 904-0495, Japan, ‡College of Letters and Sciences, Arizona State University, 7001 E. Williams Field Road, Mesa, AZ 85212, USA*

Abstract

A key requirement for social cooperation is the mitigation and/or social regulation of aggression towards other group members. Populations of the harvester ant *Pogonomyrmex californicus* show the alternate social phenotypes of queens founding nests alone (haplometrosis) or in groups of unrelated yet cooperative individuals (pleometrosis). Pleometrotic queens display an associated reduction in aggression. To understand the proximate drivers behind this variation, we placed foundresses of the two populations into social environments with queens from the same or the alternate population, and measured their behaviour and head gene expression profiles. A proportion of queens from both populations behaved aggressively, but haplometrotic queens were significantly more likely to perform aggressive acts, and conflict escalated more frequently in pairs of haplometrotic queens. Whole-head RNA sequencing revealed variation in gene expression patterns, with the two populations showing moderate differentiation in overall transcriptional profile, suggesting that genetic differences underlie the two founding strategies. The largest detected difference, however, was associated with aggression, regardless of queen founding type. Several modules of coregulated genes, involved in metabolism, immune system and neuronal function, were found to be upregulated in highly aggressive queens. Conversely, nonaggressive queens exhibited a striking pattern of upregulation in chemosensory genes. Our results highlight that the social phenotypes of cooperative vs. solitary nest founding tap into a set of gene regulatory networks that seem to govern aggression level. We also present a number of highly connected hub genes associated with aggression, providing opportunity to further study the genetic underpinnings of social conflict and tolerance.

Keywords: aggression, cooperation, gene networks, pleometrosis, social evolution, transcriptomics

Received 11 January 2016; revision received 14 April 2016; accepted 2 May 2016

Introduction

The evolution of behavioural strategies occurs at the interface between genetic variation, its expressed behavioural outcomes and selection. Dissecting these interactions for complex social behaviours can be difficult, because the expression of social behaviour by an individual is an interdependent effect of its own genetic architecture, as well as the social genotypes of other individuals in its group (Moore *et al.* 1997, 2002; Wolf *et al.* 1998; Agrawal *et al.* 2001; Fewell 2003; Linksvayer & Wade 2005; Bijma *et al.* 2007; McGlothlin *et al.* 2010; Linksvayer 2015). To understand how genetic architecture mediates social phenotype, we must additionally consider how the social environments normally encountered by that genotype influence gene expression (Page & Amdam 2007; Anholt & Mackay 2012; Linksvayer 2015). Aggression is an integral part of the fabric of

Correspondence: Martin Helmkampf, Fax: +1 (480) 965 6899; E-mail: martin.helmkampf@asu.edu

[1]These authors contributed equally to this work.

© 2016 John Wiley & Sons Ltd

social living, with a strong impact on individual and group fitness (Edwards *et al.* 2006; Alaux *et al.* 2009; Clark & Fewell 2014; Overson *et al.* 2014). In cooperative groups, aggression is often used as a social strategy, by determining differential access to reproduction and other resources, or by modulating and policing cooperative behaviours and reciprocity (reviewed by Collias 1944; Francis 1988; Clutton-Brock & Parker 1995). In primitively eusocial societies, for example, aggression is used in the formation of reproductive dominance hierarchies, for example wasps (O'Donnell 1998; Toth *et al.* 2014), ants (Liebig *et al.* 1999; Monnin & Ratnieks 2001; Hartmann *et al.* 2003) and naked mole rats (Clark & Faulkes 2001; Dengler-Crish & Catania 2007). Similarly, in cooperatively breeding societies, it determines differential mating opportunities and access to other resources (Clutton-Brock & Parker 1995; Creel & MacDonald 1995).

For aggression to be usefully employed as a social strategy in cooperative groups, however, its expression is generally coupled with downregulatory or conflict-resolving behaviours to prevent escalation (de Waal 2000; Silk 2002). In groups or contexts for which these mechanisms are absent, initial aggressive behaviours have the potential to escalate into high-intensity conflict, as the initiator is responded to with aggression by the antagonist. These cases can have severe fitness effects, but also serve as useful contexts in which to dissect the roles of genetic architecture and social environment. The strategy of cooperative nest founding by ant queens provides one such test case for examining the genetic mechanisms underlying aggression, and to explore the interdependence of aggression with social context. In most ant species, new colonies are established individually by recently mated queens (haplometrosis, Hölldobler & Wilson 1977, 1990). However, in several species, queens will form cooperative associations with nonrelatives during colony founding (pleometrosis, Hölldobler & Wilson 1977; Keller 1993; Bernasconi & Strassmann 1999; Johnson 2004; Boomsma *et al.* 2009, 2014). In very few cases, including the California harvester ant *Pogonomyrmex californicus*, these alliances persist through colony maturity, creating polygynous or multifamily units functioning together as a single colony (e.g. Cahan & Julian 1999; Overson *et al.* 2014). Contiguous populations of *P. californicus* show distinct differences in founding strategy. In one population, almost all nests are founded by pleometrosis and mature nests are often polygynous; in the other, most nests are founded by single queens and mature nests are monogynous.

One of the primary behavioural differences between the haplometrotic and pleometrotic queens is in the level of aggression. Some queens from both populations display aggression. However, the likelihood of aggression is much higher when the social group contains a haplometrotic queen, which usually ends with the death of one or both queens. Normally pleometrotic queens, in contrast, show little aggression to other pleometrotic queens (Clark & Fewell 2014; Overson *et al.* 2014). We therefore asked: (i) what are the gene expression correlates of these observed population differences in queen founding strategy (haplometrotic vs. pleometrotic); and (ii) what is the relationship between gene expression and social context (nesting alone vs. in pairs) in the behavioural expression of aggression. To answer these questions, we placed newly mated queens from the two populations into social groups in the laboratory, pairing them either with a queen from the same or alternate social type. We then recorded aggressive behaviour and analysed whole-head gene expression profiles. Finally, we employed a systems biological approach by analysing gene coexpression networks to identify modules of coregulated genes and determine major hub genes that govern aggressive behaviour in *P. californicus*.

Materials and methods

Sample collection, behavioural observations and RNA isolation

Newly mated queens from predominantly solitary founding (haplometrotic) and predominantly cooperatively founding (pleometrotic) populations of *Pogonomyrmex californicus* were collected during the mating flight season of 2013 at Lake Henshaw (33.233°N, 116.762°W) and Pine Valley (32.824°N, 116.528°W) in San Diego County, California, respectively. Specimens were placed as singletons or in pairs in standardized observation nests containing sterilized soil from the collection site and fruit flies and grass seeds as food ad libitum. Only specimens of similar body weight (within 10% of each other, mean = 13.0 mg, total range = 10.0–19.6 mg) were paired to avoid effects of body size on group behaviour. Five experimental groups were established in 48 replicates each: haplometrotic singletons, pleometrotic singletons, purely haplometrotic pairs, purely pleometrotic pairs and mixed pairs of both founding types. Overt aggressive behaviour, defined as grappling, biting and pursuing each other, was recorded for 5 days by semi-continuous scan sampling, that is each specimen was successively observed for <30 s before moving on to the next specimen (see Holbrook *et al.* 2009). On average, ten rounds of observations were completed per day between 8 a.m. and 8 p.m. for a total number of 50 observations. During this period, six specimens were collected immediately

© 2016 John Wiley & Sons Ltd

after they were observed killing their cofoundress following prolonged and intense antagonistic behaviour ('aggressive queens'). This included queens of both populations and occurred between 4 h and 4 days after the start of the experiment. Because aggressive behaviour generally began with one queen making aggressive displays (flared mandibles, standing on top of the other queen) and escalated to overt aggression (i.e. fighting) over many hours, aggressive queens could be clearly identified by scan sampling and monitored specifically for overt aggression. In addition, specimens were collected at the end of the five-day observation period, including 24 specimens from pairs of all association types that did not display aggressive behaviour during the observation period ('nonaggressive queens'), as well as singletons of both populations. The total number of specimens, their population of origin and social environment thus amounted to (see also Table S1, Supporting information):

- Six haplometrotic singletons (Hs)
- Six pleometrotic singletons (Ps)
- Eight specimens in four haplometrotic pairs (Hh)
- Eight specimens in four pleometrotic pairs (Pp)
- Eight specimens of each type in four mixed pairs (Hp and Ph)
- Six aggressive specimens in three haplometrotic (HA) and three pleometrotic (PA) pairs

All 42 specimens were frozen in liquid nitrogen upon collection and stored at −80 °C until total RNA was obtained by TRIzol (Invitrogen, Thermo Fisher Scientific)/chloroform extraction from whole heads.

RNA sequencing and assembly

After RNA extraction, cDNA was synthesized from 100 ng total RNA per sample using polyT oligos and SuperScript II reverse transcriptase (Invitrogen, Thermo Fisher Scientific). To control for variability in RNA yield and assess the experiment's sensitivity and dynamic range, ERCC RNA Spike-In Mixes 1 and 2 (Ambion, Thermo Fisher Scientific) were added alternately to each sample during this step. Second-strand cDNA synthesis was performed using 15 amplification steps, and cDNA libraries were prepared using the Nextera XT DNA Sample Preparation Kit (Illumina, San Diego, CA, USA) to a final concentration of 2 nM. This included the ligation of unique labels, allowing the identification of each of the 42 libraries, corresponding to individual samples, and fragment size selection to 300–700 bp (see (Aird *et al.* 2015) for a more detailed description of cDNA library preparation). After cluster generation, all libraries were sequenced in parallel from both ends on four flow cell lanes using the HiSeq 2000 Sequencing System (Illumina) and v3 chemistry. This yielded ~18.7 million reads or 1.9 gigabases per library, 93% of which achieved a quality score of at least 30 (=base call accuracy of 99.9% or higher), a depth and quality at which RNA sequencing has been shown to be comparable to microarray technology (Wang *et al.* 2011). Reads were quality-trimmed using TRIM GALORE! (v. 0.3.1), and assembled de novo using TRINITY (r20131110) (Grabherr *et al.* 2011).

This assembly encompassed 799 megabases and contained 311 726 transcripts. The assembly was comprehensive, in that 98% of the RNA-seq reads could be mapped back to it as properly paired reads. However, the overwhelming majority of the transcripts occurred at low abundance and had to be filtered prior to differential gene expression analysis. We chose the FPKM cut-off that optimized the average correlation between expected and observed values for each library pair containing different spike-in mixes (FPKM = 0.14, see Fig. S1, Supporting information). Using the lower bound obtained from spike-in data, we could detect RNA concentrations as low as 0.12 attomoles/μL, and over at least 12 orders of magnitude. Filtering in this way, we kept 7890 transcripts (Appendix S1, Supporting information) for differential gene expression and gene coexpression network analyses.

Genetic differentiation between populations

In order to estimate the extent of genetic differentiation between the two populations, we mapped the transcriptome to the *Pogonomyrmex barbatus* reference transcriptome using Stampy (Lunter & Goodson 2011), allowing for a substitution rate of 6%. After marking possible PCR duplicate reads using Picard tools, we called variants using FREEBAYES (Garrison & Marth 2012), keeping variants with a quality score >20. We then used VCFTOOLS (Danecek *et al.* 2011) to compute the Weir and Cockerham F_{ST} between populations (Weir & Cockerham 1984).

Differential gene expression analysis

To quantify gene expression, reads were then remapped to a reference file containing the assembly and ERCC spike-in sequences, using BOWTIE (v1.1.0), as part of the RSEM pipeline (v1.2.11) (Langmead *et al.* 2009; Li & Dewey 2011). The RSEM analysis was performed on TRINITY components, treating TRINITY sequences as isoforms. PCR duplicates were filtered prior to the RSEM expression calculation using Picard Tools. Transcript abundance was measured in expected read counts and FPKM values by RSEM (Appendices S2 and S3,

© 2016 John Wiley & Sons Ltd

Supporting information, respectively), and differences in gene expression levels between samples statistically tested with the R package EDGER (Robinson *et al.* 2009). Differences in gene expression were considered to be significant at a *P*-value level of 0.05 using false discovery rate (FDR) adjustment for multiple testing, and a minimum fold-change of 2 ($\log_2 = 1$) to further increase the specificity of the analysis. Similarity of gene expression profiles across all genes and samples was visualized by the multidimensional scaling function implemented in EDGER, using default parameters. Genes were identified by BLAST against NCBI's nonredundant protein database, the *P. barbatus* Official Gene Set v1.2 (Smith *et al.* 2011), the *Drosophila melanogaster* translated gene set r6.06, and/or the protein database UniProt (*e*-value cut-off of $1 \times 10e^{-5}$). The latter and FlyBase provided functional annotation for genes of particular interest, after verifying gene homology by reciprocal BLAST. Gene sets were tested for enrichment in Gene Ontology (GO) terms with respect to biological processes with the R package GOSTATS based on *D. melanogaster* GO annotation data obtained with BLAST2GO (Conesa *et al.* 2005). Finally, cross-species comparisons were conducted using TBLASTX against transcript data sets provided by the authors of the respective studies (*e*-value cut-off of $1 \times 10e^{-5}$) to determine gene overlap. Gene lists were tested for nonrandom association by Fisher's exact test using the smallest of the following as the total/background: best estimate available in the literature for the number of protein-coding genes in *P. californicus* (17 177 based on *P. barbatus*, Smith *et al.* 2011) or the compared species, or the number genes represented on the microarray used in the compared study.

Gene coexpression network analyses

To describe patterns of interconnectedness between genes that may represent biological pathways, we employed two complementary approaches. First, the R package WGCNA 1.51 (Langfelder & Horvath 2008) was used to construct a network based on how gene expression correlates between genes. We chose an exponent of 9, the lowest value resulting in a scale-free network topology, to transform a correlation matrix of variance-stabilized FPKM values for all 7890 transcripts into an adjacency matrix. After constructing the network, clusters of highly interconnected genes, so called modules, were identified by hierarchical clustering using a minimum module size of 40 and a dendrogram cut height parameter of 0.25. Modules were further characterized by determining highly connected genes within each module and testing each module for enrichment in GO terms. Modules that were significantly associated with population, social context and aggression were identified by multiple linear regression. We considered these traits as dichotomous, independent variables, coding haplometrotic/pleometrotic, single/paired and nonaggressive/aggressive queens as 0/1, respectively. The module eigengenes (each module's gene expression profile for each sample) served as dependent variables. *P*-values were corrected for multiple testing by the Bonferroni method. Additionally, a smaller data set limited to 524 genes that were differentially expressed between aggressive and nonaggressive queens was analysed using WGCNA.

To construct a more detailed coexpression network of genes associated with aggressive behaviour, we designed a second approach based on the same limited data set. For each pair of genes (nodes) *i* and *j*, the correlation coefficient $c(ij)$ was calculated from variance-stabilized FPKM values. From the resulting coexpression matrix, we computed the positive, negative and total degree of each node—$d_p(i)$, $d_n(i)$ and $d_a(i)$—as the sum of all positive, negative and absolute correlation coefficient values, respectively. Due to the size of the network, and because by definition, each node is connected to all other nodes in the network, these values can get quite large. In order to evaluate the 'true' weight of each node, we normalized the weight of each node according to the size of the network. The average positive, negative and total degree $ad_p(i)$, $ad_n(i)$ and $ad_a(i)$ was thus defined as $d_p(i)$, $d_n(i)$ and $d_a(i)$ divided by the number of nodes positively correlated with node *i*, negatively correlated with node *i*, and all nodes in the network, respectively. These values, together with the ratio of positive or negative to total edges of each node, provide a measurement for the connectivity and regulatory importance of a gene, and whether its function is predominantly upregulating or downregulating based on the ratio of its average positive and negative degrees.

Results and discussion

Measuring aggressive behaviour

Of 144 founding associations including 48 purely haplometrotic, purely pleometrotic and mixed pairs, a total of 44 pairs showed at least one instance of overt aggression (i.e. fighting) during the 5-day observation period. Haplometrotic pairs displayed aggressive behaviour significantly more frequently than pleometrotic or mixed pairs, both in terms of the number of nests in which aggressive behaviour was observed (Fig. 1A, chi-square test, $P = 0.02$), and in the absolute number of aggressive acts (Fig. 1B). In 11 of the 21 haplometrotic pairs in which aggression was observed, both queens acted

© 2016 John Wiley & Sons Ltd

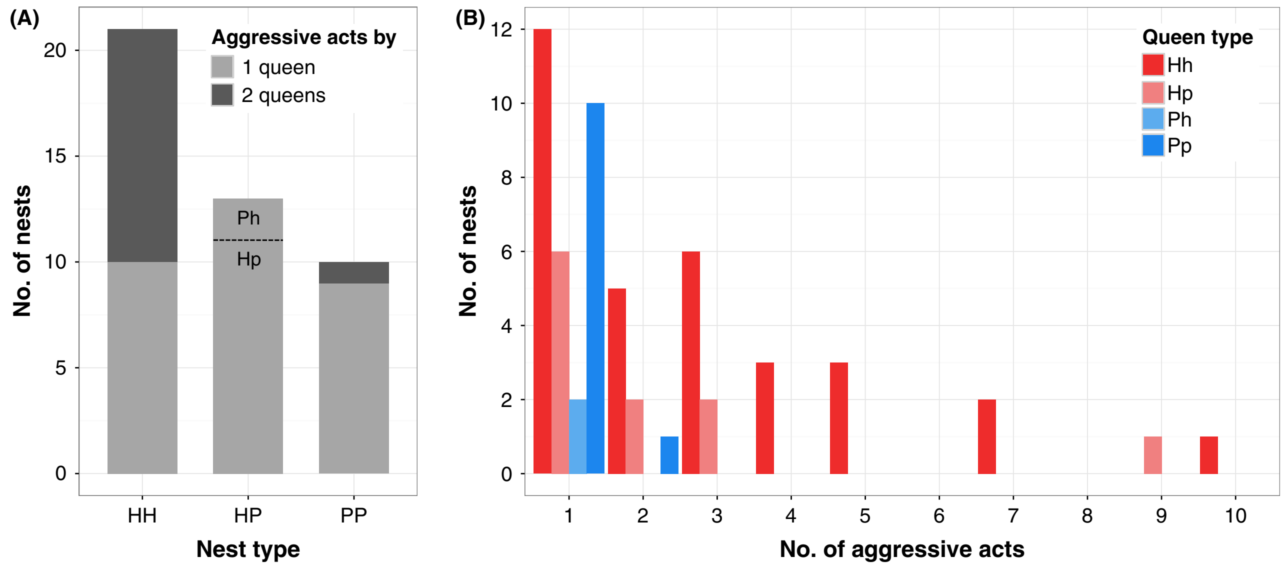

Fig. 1 Frequency of aggressive acts observed in haplometrotic, pleometrotic and mixed associations of *Pogonomyrmex californicus* founding queens. (A) Number of purely haplometrotic (HH), purely pleometrotic (PP) and mixed (HP) associations ('nests') in which aggressive acts were observed. In mixed associations, the dashed line indicates the fraction of nests in which aggressive acts were committed by the haplometrotic (Hp) or pleometrotic (Ph) queen, respectively. (B) Number of aggressive acts committed by haplometrotic queens in purely haplometrotic associations (Hh), haplometrotic queens in mixed associations (Hp), pleometrotic queens in mixed associations (Ph) and pleometrotic queens in purely pleometrotic associations (Pp). This analysis was based on 50 observations of 48 associations of each type. Haplometrotic *P. californicus* foundresses committed aggressive acts significantly more frequently than pleometrotic queens in artificial associations (Chi-square test, $P = 0.02$).

aggressively (Fig. 1A). This phenomenon was almost never observed in other types of associations. These data demonstrate that aggression is amplified when haplometrotic queens interact. The number of nests in which aggressive interactions occurred was generally low and did not differ between pleometrotic pairs and mixed pairs. These data suggest that the presence of a pleometrotic queen can have a dampening effect on aggression, irrespective of the population identity of the other queen.

These findings are consistent with previous studies (Clark & Fewell 2014; Overson *et al.* 2014), but expand on them to explicitly show conflict escalation—both the number of aggressive acts, and the frequency with which both parties engaged in conflict increased dramatically when haplometrotic queens were placed together. The same studies also demonstrated that mortality is highest in haplometrotic associations. Our data support the hypothesis that the transition from solitary to cooperative nest sharing requires some mechanism for reduced aggression and increased tolerance towards other queens. It is likely that reduced aggression is a hallmark of transitioning both to eusociality and to cooperative sociality. This transition has to be accompanied or predated by changes at the genomic and transcriptomic levels affecting developmental pathways. However, we note that the fitness consequences of the behavioural interactions captured by this study are specifically different than for the transition to eusociality itself—these queens are adult nonrelatives who seek each other out to construct nests, rather than adult offspring that remain at the nest (Boomsma *et al.* 2009, 2014).

Overall gene expression profiles

We examined overall gene expression profiles from heads of 42 founding queens differing in population of origin, social environment and level of aggression by RNA sequencing. While the results presented here were obtained from whole heads, we assume that observed gene expression differences can be predominantly attributed to the brain due to the focus on behavioural traits. Dissecting brains, which would have reduced bias from non-neuronal tissue in the head (e.g. glandular and muscular tissue), proved unfeasible due to the number and size of specimens, and appeared to be ineffective in cleanly separating tissue types.

We obtained data on 7890 transcripts that were expressed sufficiently to meet our detection thresholds. Of those, 7530 matched 7017 unique genes in *Pogonomyrmex barbatus*. This suggests the experiment covered slightly less than half the genes expected in the *P. californicus* genome (cf. Smith *et al.* 2011). Multidimensional

© 2016 John Wiley & Sons Ltd

scaling (MDS) analysis across all transcripts revealed that aggressive behaviour had a stronger effect on gene expression than either population or social context (Fig. 2). Although spread out considerably in the multidimensional scaling plot, most of the six aggressive queens were separated from the other queens along the first factor of the MDS analysis. Despite this variance and the small sample size for aggressive queens, this suggests that the physiological state associated with aggression overrides the influence of population on gene expression. As some of the aggressive queens endured prolonged fights before disposing of the rival queen, it is also possible that gene expression was affected by latent injury. Consistent with our observations of the frequency of aggressive behaviour, the difference between the populations with regard to aggression was not absolute (queens from both populations showed aggressive behaviour), but rather manifested in the frequency and propensity to initiate and escalate aggressive behaviour.

Considering the remaining queens (i.e. all but the six chosen for high levels of aggression, including singletons), we found a clear separation between queens of the haplo- and pleometrotic populations along the second MDS factor (Fig. 2), despite the two populations displaying relatively low genetic differentiation (F_{ST} = 0.056 based on 20 587 single nucleotide polymorphisms in our data set). While the haplometrotic queens clustered very tightly, the pleometrotic queens showed more variation along the second MDS axis (Fig. 2).

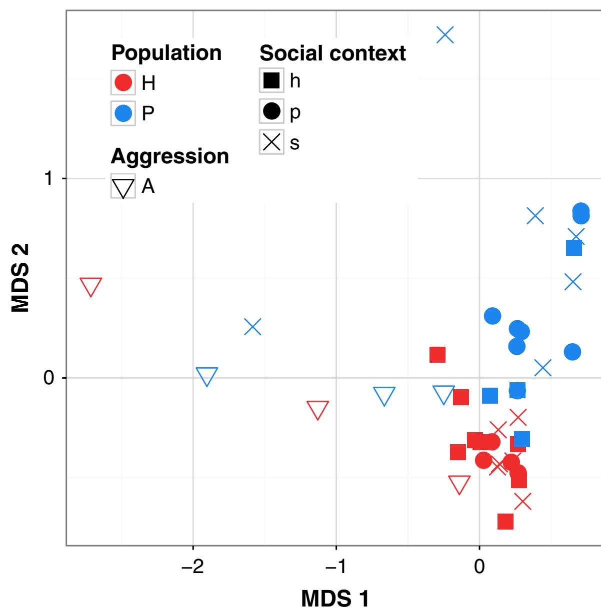

Fig. 2 Similarity of whole-head gene expression profiles of *Pogonomyrmex californicus* foundresses across all 7890 transcripts. Individuals from the haplometrotic population (red) are separated along the second MDS axis from individuals from the pleometrotic population (blue) with a slight overlap. In contrast, social context—whether individuals were nesting alone (crosses), or with a member of the haplometrotic (squares) or pleometrotic (circles) populations—does not lead to a separation of samples. Aggressive individuals (triangles) tend to separate from the other individuals along the first MDS axis but do not form a distinct cluster.

In contrast, the social environment, that is founding nests alone or in association with an individual from the same or other population, did not affect the overall expression profile of abundantly expressed genes. This suggests that underlying, fixed genetic differences between the populations have a measurable effect on gene expression. What makes this result especially noteworthy is the fact that these queens exhibited similarly low levels of aggressive behaviour during observations, and were also matched in their developmental or life history stage. Even in the absence of the most striking behavioural difference, aggression, the population of origin thus exerts its influence on the pattern of gene activity in the heads of *P. californicus* founding queens. This may be a reflection of differences between the two queen types that extend beyond aggression to social effects on behavioural coordination. In addition to aggression, queens of the two populations show distinct differences in their level of task differentiation (with haplometrotic queens showing higher task specialization), propensity to engage in excavation (higher in haplometrotic queens) and brood care activities (higher in pleometrotic queens) (Helms Cahan & Fewell 2004; Jeanson & Fewell 2008; Clark & Fewell 2014).

Differentially expressed genes—population and social environment

To further illuminate the quantitative and qualitative differences in gene expression, we looked at differentially expressed genes across populations and social contexts (Tables 1 and S4–S17, Supporting information). We found 209 unique, differentially expressed genes across all samples excluding aggressive queens. This represents about 1% of the total number of expected genes in the genome of *P. californicus* (cf. Smith *et al.* 2011), using relatively high stringency parameters (FDR = 0.05 and a minimum twofold expression difference). The highest number of differentially expressed genes was found between haplometrotic and pleometrotic queens in pure pairs (Hh–Pp, n = 135 genes; Fig. S2A and Table S5, Supporting information). Genes upregulated in pure-pair haplometrotic queens (Hh) included genes presumably involved in immune response, different metabolic processes and neuronal development and function (e.g. *neuroendocrine protein 7B2*, which encodes a peptidase known to regulate the

© 2016 John Wiley & Sons Ltd

Table 1 Number of differentially expressed genes between *Pogonomyrmex californicus* foundresses from two populations in different social contexts

	Hs (6)	Ps (6)	Hh (8)	Pp (8)	Hp (4)	Ph (4)
Hs (6)	–					
Ps (6)	20	–				
Hh (8)	2	63	–			
Pp (8)	32	0	135	–		
Hp (4)	1	19	0	32	–	
Ph (4)	29	14	19	15	11	–

Capital letters designate population of origin of focal queen: H, haplometrotic; P, pleometrotic. Lower-case letters indicate social context: h, paired with haplometrotic cofoundress; p, paired with pleometrotic cofoundress; s, singleton (unpaired). Numbers in parentheses designate the sample size of each experimental group. Gene expression differences were considered significant at an false discovery rate-adjusted *P*-value cutoff of 0.05 and a minimum fold-difference of 2. See Table S4–S17 (Supporting information) for detailed gene lists of each comparison.

activation and secretion of neuropeptides in vertebrates and insects, Hwang *et al*. 2000). Genes without homologs in other ants or Hymenoptera were also common. Such taxonomically restricted genes (TRGs) have been shown to be widespread in ants and may be behind many lineage-specific adaptations, including social traits (Simola *et al*. 2013; Wissler *et al*. 2013; Jasper *et al*. 2015; Kapheim 2016). For instance, TRGs are overrepresented in differentially regulated genes between social phenotypes in the primitively eusocial wasp *Polistes canadensis* (Ferreira *et al*. 2013), and among genes more strongly expressed in honeybee workers than queens (Johnson & Tsutsui 2011). In line with these findings, our results provide another indication that novel genes play an important role in regulating social behaviour in *P. californicus*.

In pure-pair pleometrotic queens (Pp), the four most strongly expressed genes in comparison with Hh queens were homologous to parts of the deformed wing virus (DWV) genome. This honeybee pathogen is transmitted by *Varroa* mites and related to the Kakugo virus found in the mushroom bodies of aggressive guard bees (Lanzi *et al*. 2006). Contrary to expectations based on the propensity for aggressive behaviour in the haplometrotic queens, these viral transcripts were extremely abundant (transcript group comp62327, FPKM values 10 000–65 000) in seven of our pleometrotic queens, but only one haplometrotic queen. Infection with this virus thus does not seem to be correlated with aggression in *P. californicus*, possibly because the host–parasite relationship is unspecific. Genes particularly strongly expressed in Pp queens also include several taxonomically restricted, functionally unknown genes, as well as various metabolic genes. Of particular note are several genes homologous to *Drosophila melanogaster* genes encoding transcription factors and neuropeptide receptors, including *Rho GTPase-activating protein 190*—which has been shown to play a role in insect olfactory learning and mushroom body development (Billuart *et al*. 2001)—*G-protein-coupled receptor 161*, and *5-hydroxytryptamine receptor 7*, a serotonin receptor which is required for courtship and mating behaviour in *D. melanogaster* (Becnel *et al*. 2011). Based on their function in other organisms, these genes may well contribute to the behavioural differences observed between the two *P. californicus* populations.

The gene expression differences between haplometrotic and pleometrotic queens in pairs (Hh–Pp) were almost an order of magnitude higher than those between haplometrotic and pleometrotic queens that were nesting alone (Hs–Ps, $n = 20$). Whereas the latter represents population differences in the absence of a social environment—the norm for haplometrotic queens—the former compares the worst and best case scenario for haplometrotic and pleometrotic queens, respectively. Because pleometrotic queens predominantly found nests cooperatively, this association (Pp) represents their normal social environment. On the other hand, the normally single-founding, haplometrotic queens do not have an evolutionary history of cofounding and were artificially forced into a social environment in this case (Hh). Thus, the differentiation in gene expression across these social contexts could indicate that the social environment magnifies gene expression differences between queens of the two populations, possibly through latent stress brought about by dominance interactions even in the absence of overt aggression. Similarly, comparing the dominant founding strategies of each population also revealed only a moderate number of differentially expressed genes (Hs–Pp, $n = 32$). These results stand in contrast to the thousands of differentially expressed transcripts found between single haplometrotic and paired pleometrotic foundresses in the monogyne form of the fire ant *Solenopsis invicta* using microarrays (Manfredini *et al*. 2013). While we detected moderate overlap ($n = 40$) between these genes and the differentially expressed genes in haplometrotic and pleometrotic pairs of *P. californicus* (Hh–Pp), it did not prove to be significant (Fisher's exact test, two-tailed $P = 0.19$, using 16 569 genes reported for the *S. invicta* genome as background). However, it is difficult to directly compare the two experiments, which differ in the degree of replication and technology used. With that caveat, there are also biological reasons that differences might exist between the two studies. Manfredini *et al.* examined whole-body gene expression patterns, which in contrast

© 2016 John Wiley & Sons Ltd

to head/brain gene expression patterns are expected to be more broadly affected by physiological differences between haplometrotic and pleometrotic queens. In addition, fire ant queens of the monogyne form only tolerate each other briefly during the colony founding phase (Balas & Adams 1996), whereas pleometrotic associations in *P. californicus* often continue to cooperate throughout colony maturity, so the social phenotypes are not entirely comparable between these two species.

Differentially expressed genes—aggression

Consistent with the results of the MDS analysis, queens showed a strong transcriptional response to performance of aggression: 524 genes (3% of all genes) were differentially expressed between the six aggressive queens and 24 nonaggressive queens in pairs (note that this excludes all singletons) (Fig. S2B and Table S4, Supporting information). Gene ontology (GO) term enrichment analyses revealed an intriguing pattern: of the 243 genes upregulated in nonaggressive queens, significantly more than expected by chance are involved in sensory perception of smell ($P \ll 0.001$). This enrichment was due to an overrepresentation of odorant receptor (OR) genes, which encode receptors initiating the detection of odour molecules in the antennae and other chemosensory organs in insects. Even more strikingly, functional annotation by homology using BLAST revealed that more than half of the 50 most strongly elevated genes in nonaggressive queens are involved in chemoreception, including genes encoding 20 ORs, two gustatory receptors, four odorant binding proteins and a homolog of the *D. melanogaster Sensory neuron membrane protein 1*. In ants, these genes are crucial components of the chemical communication system, allowing them to discern nestmates, caste identity, fertility status and other signals. Queens cooperating during nest founding may express these genes at a higher level as they interact more frequently and generally behave more responsively to the activities of the other queen. This finding could also hint at a potential mechanism to reduce or dampen aggression. Less aggressive queens may invest more heavily in communication 'hardware', which in turn changes their sensitivity, reaction and ultimately also the reaction of the interacting queen.

In contrast, the set of 281 genes upregulated in aggressive queens, with respect to nonaggressive queens, was enriched in genes involved in metabolic processes according to GO term enrichment analyses (52 genes, $P \ll 0.001$). However, manual annotation of the 50 most strongly differentially expressed of these genes revealed a very heterogeneous picture. For instance, the list contains genes related to immune system function, nutrient storage and brain function, including homologs of *Larval serum protein 2* (linked to synapse formation in *D. melanogaster* embryos and possibly adults) (Benes *et al.* 1990; Inaki *et al.* 2007) and *Retinal rod rhodopsin-sensitive cGMP 3′,5′-cyclic phosphodiesterase delta subunit* (Kobayashi *et al.* 2006). Most genes upregulated in aggressive queens relative to nonaggressive queens, however, are implicated in various metabolic processes, including several cytochrome P450s, and mitochondrial function. The latter includes three genes encoded on the mitochondrial genome that are part of the oxidative phosphorylation (OXPHOS) pathway, *cytochrome c oxidase II* and *NADH dehydrogenase subunits 1* and *4*. As this pathway provides most of the energy needs of eukaryotic cells in the form of ATP, this finding correlates with a series of studies showing that changes in metabolic gene expressions is a shared trait for aggressive behaviour across the animal kingdom (Gammie *et al.* 2007; Alaux *et al.* 2009; Ayroles *et al.* 2009; Edwards *et al.* 2009; Anholt & Mackay 2012; Li-Byarlay *et al.* 2014; Toth *et al.* 2014). However, other studies with social insect workers have reported a decrease in energy metabolic gene expression in brains associated with aggression, for instance in honeybees and paper wasps (Alaux *et al.* 2009; Toth *et al.* 2014). In contrast, the expression of many genes related to metabolism was found to be elevated in brains of mice displaying maternal aggression (Gammie *et al.* 2007), and whole bodies of *D. melanogaster* selected for aggressive behaviour. Our results are more in line with the latter two studies, as most genes upregulated in aggressive queens are involved in general metabolic processes and not energy metabolism in particular. Correspondingly, we did not find any functional OXPHOS pathway genes encoded in the nuclear genome to be elevated in expression. While a causal relationship between decreased brain energy metabolism and aggression has been demonstrated in honeybees and *D. melanogaster* (Li-Byarlay *et al.* 2014), we have no measures of the metabolic rate in brains of aggressive queens and are thus are unable to determine whether this upregulation is linked with an increase or decrease of the metabolic rate in founding queens of *P. californicus*.

Finally, we investigated whether the list of differentially expressed genes between aggressive and nonaggressive *P. californicus* queens overlaps with genes found to be associated with aggression in related species. Of the 38 differentially regulated genes shared by honeybees in three aggression-related contexts (hereditary, alarm pheromone induced and age dependent, Alaux *et al.* 2009), three were also found in our list of genes associated with aggression: a gene of unknown function (GB13501), and two encoding a P450 cytochrome and potassium channel, respectively. A similar fraction, six of 63 genes also overlap with brain

© 2016 John Wiley & Sons Ltd

dominance-associated transcripts identified in *Polistes metricus* paper wasps (Toth *et al.* 2014). Finally, 132 genes are shared with the list of genes differentially expressed in *D. melanogaster* lines selected for high and low levels of aggression (out of 1539, Edwards *et al.* 2006). These overlaps, only the last of which is significant (Fisher's exact test, two-tailed: $P \ll 0.001$), are limited in scale, which may be due in part to differences in technology. All three studies above were based on microarrays, and in the case of the honeybee and *Polistes*, did not represent the entirety of genes in the genome. However, the scale of overlap found here is similar to other cross-species comparisons (Toth *et al.* 2014). The fact that we could not identify a common set of genes associated with aggression, as would be indicative of a single, aggression-regulating gene network conserved across insects, may thus also be attributed to different genetic, environmental and social contexts of aggression. Even within the same species, the overlap between aggression-related genes identified in different contexts can be limited (Alaux *et al.* 2009). In addition, gene networks controlling social behaviour are likely to be different between unrelated species, because they require the rewiring and combination of disparate networks inherited from solitary ancestors (Johnson & Linksvayer 2010). This hypothesis has recently found empirical support (Kapheim *et al.* 2015; reviewed in Kapheim 2016).

Gene coexpression network analyses

Using a weighted gene coexpression correlation network approach (WGCNA), we detected eleven modules of transcriptionally correlated genes among all 7890 transcripts (Fig. S3 and Table S18, Supporting information). The size of these modules ranged from 82 to 2079 genes (Table 2), with another 2263 genes (29%) encompassing a group that could not be assigned to any module. Six of these modules were significantly associated with aggressive behaviour (Table 2; modules 1, 3, 5, 8, 9 and 11). In contrast, the social context in nonaggressive queens, that is whether foundresses were nesting alone or in pairs, was not significantly correlated with any module. Similarly, population was only associated with the unassigned group of genes, but none of the modules. To functionally characterize the modules associated with aggressive behaviour, we performed GO term enrichment analyses. The group of unassigned genes, tending to be downregulated in pleometrotic individuals, was enriched in genes related to DNA metabolic (37 genes, $P < 0.001$) and developmental processes (12 genes, $P = 0.009$). Modules positively correlated with aggressive behaviour, thus containing mostly genes upregulated in aggressive individuals, were enriched in genes involved in metabolic processes (module 9, 31 genes, $P < 0.001$ and module 11, various sets of genes), protein localization (module 3, 9 genes, $P = 0.019$) and transport (module 5, 18 genes, $P = 0.008$). In contrast, modules negatively correlated with aggressive behaviour were characterized by genes related to cell communication (module 1, 63 genes, $P = 0.002$) and the perception of chemical stimuli (module 8, 9 genes, $P < 0.001$). This is congruent with the finding that genes involved in metabolism and chemoperception dominate among the differentially expressed genes between aggressive and nonaggressive queens. Such modules of transcriptionally highly interconnected genes, enriched in Gene Ontology categories and correlated with complex phenotypes like aggression, are suggestive of genetic networks that regulate these phenotypes. Such networks are often characterized by common pathways, tissue-specific gene expression and enrichment in transcription factor binding sites (Ayroles *et al.* 2009).

To further elucidate the gene regulatory network of aggression, we repeated the WGCNA analysis with the subset of genes we found to be differentially expressed between aggressive and nonaggressive queens. Of the 524 genes, 377 could be assigned to three modules, containing 181 (module *M*), 127 (module *C*) and 69 (module *U*) genes, respectively. GO analyses revealed these modules to be enriched in genes pertaining to metabolic processes (module *M*, $P < 0.001$, 23 genes) and the perception of chemical stimuli (module *C*, $P < 0.001$, eight

Table 2 Modules of coregulated genes and their association with population, social context and aggression

Module	No. genes	Population	Social context	Aggression
0	2263	**−0.21 (<0.001)**	0.04 (1.000)	−0.03 (1.000)
1	2079	0.02 (1.000)	0.03 (1.000)	**−0.29 (<0.001)**
2	1511	0.02 (1.000)	0.00 (1.000)	−0.12 (1.000)
3	507	−0.04 (1.000)	0.03 (1.000)	**0.30 (<0.001)**
4	479	−0.05 (1.000)	0.06 (1.000)	0.01 (1.000)
5	369	−0.01 (1.000)	0.01 (1.000)	**0.24 (0.016)**
6	142	0.12 (0.200)	−0.11 (0.971)	0.20 (0.053)
7	131	0.06 (1.000)	−0.05 (1.000)	0.20 (0.148)
8	115	−0.02 (1.000)	−0.04 (1.000)	**−0.27 (0.001)**
9	114	0.03 (1.000)	−0.10 (1.000)	**0.30 (<0.001)**
10	98	−0.03 (1.000)	0.02 (1.000)	−0.13 (1.000)
11	82	0.05 (1.000)	−0.06 (1.000)	**0.28 (0.001)**

Among all transcripts, 11 modules were identified by WGCNA, with a fraction remaining unassigned ("module" 0, in grey). Shown are beta-coefficients and Bonferroni-corrected *P*-values (in parentheses) from multiple linear regressions (signs are with respect to pleometrotic, paired and aggressive queens). Coefficients significantly differing from zero (in bold) were only found between modules and aggression. See main text for further characterization of correlated modules.

© 2016 John Wiley & Sons Ltd

genes). Module *U* did not show enrichment in any meaningful categories, most likely due to its small size. As expected, the majority of genes in module *M* and *C* overlap with the modules significantly correlated with aggressive behaviour found in the comprehensive WGCNA analysis.

The second, complementary network approach revealed that on average, nodes (genes) in the coexpression network of genes associated with aggressive behaviour have up- and downregulating interactions with other genes in equal measure (the average ratio of positive edges per node was 0.50). The average total (both positive and negative) degree closely followed a normal distribution with a mean of 0.32. Network characteristics of all examined genes can be found in the Table S19 (Supporting information). Here, we discuss the 20 genes with the highest average total degree ad_a in more detail (Table S2, Supporting information). These highly connected genes, subsequently referred to as 'hub genes', are characterized by having very high average positive (0.42–0.55) and negative degrees (−0.52 to −0.40), as well as a higher ratio of positive edges in comparison with all other genes in the network. While hub genes can be classified by various methods (Sporns *et al.* 2007), we chose to classify them based on their overall connectivity along with average positive and negative degrees, which are readily accessible measurements providing information about the direction of regulation.

Looking at the relationship between the modules and hub genes of aggression determined by both network approaches, we found that ten of the 20 hub genes each are members of WGCNA modules *M* and *C*, respectively. The majority of hub genes were also found among the overall most highly connected genes according to WGCNA (15 are among the top 25), emphasizing a high level of concordance between both approaches. The connectivity in mean average degrees between and within hubs and modules is shown in Fig. 3 (see also Table S3, Supporting information). As expected, the highest connectivity was discovered between the hub genes and modules *M* and *C* with a mean average total degree of 0.57 each, and among the hub genes themselves in particular (mean average total degree = 0.75). On average, there was a slight overrepresentation of positive edges between nodes of these sets of genes, indicating a higher occurrence of upregulating interactions between those genes in the context of aggressive behaviour. Connectivity was found to be slightly lower between modules *M* and *C* as well as within these modules, and with a approximately equal ratio of positive and negative edges. On the other hand, connectivity proved to be less pronounced between module *U* and

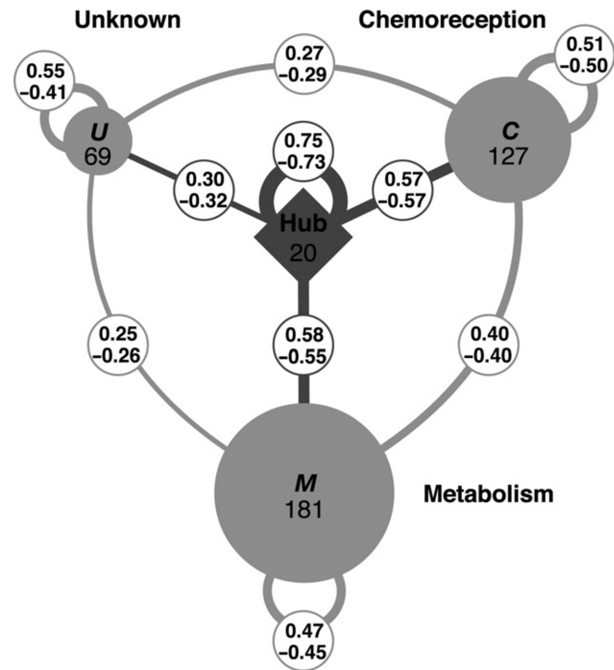

Fig. 3 Mean average positive and negative connectivity between and within hub und module genes identified among 524 differentially expressed genes between aggressive and nonaggressive *Pogonomyrmex californicus* queens. Module *M* is characterized by an overabundance of genes involved in metabolic processes, while module *C* genes are mostly involved in chemoreception (no functional enrichment was found in module *U*). The highest connectivity is found within hub genes, and between hub genes and module *M* and *C* genes.

all other sets, with lower average degrees and a higher ratio of negative than positive edges. This suggests that module *U* genes have mostly downregulating functions.

As expected of master regulatory genes, the orthologs of most hub genes are known to fulfil complex and diverse roles in other organisms (Table S2, Supporting information), and presumably encode transcription factors, regulators of hormone metabolism, and neuronal growth and function. Notably, seven of the 20 hub genes encode proteins that are found in the membranes and matrix of mitochondria, suggesting a role in cell metabolism and growth. For instance, the ortholog of the most highly interconnected hub gene we identified, *walrus*, fulfils a function in the mitochondrial matrix and is involved in the development of various organ systems in *D. melanogaster* (Liu *et al.* 1999). Several hub genes may be directly involved in the regulation of gene expression, either at the level of transcription—like the orthologs of *mustard*, *extra-macrochaetae* and *Histone H2A*—or at the level of translation like *peptidyl-tRNA hydrolase 2*. Other hub genes of note include orthologs of *nervous wreck*, which controls neuromuscular synaptic

© 2016 John Wiley & Sons Ltd

growth as part of the Wnt and BMP signalling pathways in *D. melanogaster* (Rodal *et al.* 2011) and may thus be pivotal for long-term changes in neural connectivity underlying learning and memory in other organisms, and *unkempt*, a morphogene with diverse roles including the regulation of neuron differentiation as part of the insulin receptor/TOR pathway (Avet-Rochex *et al.* 2014). The remaining hub genes encode enzymes in different metabolic pathways, membrane transporters and proteins involved in protein targeting/excretion. Finally, the list of hub genes contains a gene encoding a putative ionotropic glutamate receptor, although the similarity to known genes of this class is relatively low. However, glutamate represents one of the most prominent neurotransmitters in insects and is thus important for many brain functions. While unlikely to directly influence the transcription of many other genes, glutamate receptors may be found downstream in many genetic pathways affecting neural function and behaviour due to the ubiquity and versatility of glutamate. Alternatively, glutamate-mediated neuronal activity may indirectly affect the transcription of many other genes.

Conclusions

We examined the effect of three variables on gene expression in heads of *Pogonomyrmex californicus* founding queens: population of origin, social environment and aggressive behaviour. Overall gene expression profiles (Fig. 2), number of differentially expressed genes (Table 1) and association with gene coexpression modules (Table 2) illustrate that the population of origin more strongly influences gene expression than the social environment. This suggests that underlying, fixed genetic differences measurably affect gene regulation between the two populations, although these differences also appear to be influenced by social context in a complex way. However, we detected the strongest transcriptional changes with respect to aggressive behaviour. Our behavioural analyses indicate a strong gene–environment effect on the expression of aggression. Queens from a social background of solitary founding were more likely to initiate aggression, and aggression was more likely to escalate in the presence of another haplometrotic queen. The high levels of mortality from conflict escalation, especially in haplometrotic pairs, suggest that this may be a barrier to overcome in the evolution of cooperative nest founding; we propose similar conflict escalation could provide an issue for the transition to cooperative sociality more generally. Our transcriptomic analyses suggest that the shift in aggressive behaviour from solitary to cooperative founding may be simply one of degree in gene expression. Aggressive queens, whether from the primarily pleometrotic or haplometrotic populations, showed similar shifts in gene expression profiles away from less aggressive queens of either population. This supports the idea that the molecular mechanism underlying the upregulation of aggression is a shared trait between both populations, which differ not in the ability to be aggressive, but rather the propensity to act aggressively. Indeed, changes in gene regulation can have even more profound effects on the expression of a social phenotype, as was recently shown in ant social parasites (Smith *et al.* 2015).

We further found that genes differentially regulated between haplometrotic and pleometrotic foundresses are predominantly involved in different metabolic processes, immune response, transcription and neuronal function, but also include many putative TRGs. This could reflect hard-wired regulatory differences that have evolved in response to different environmental conditions, for instance with respect to pathogens, microclimate, diet and population density, but also play a role in the profound behavioural differences characterizing the two populations. Similarly, differentially regulated genes between aggressive and nonaggressive foundresses were enriched in genes implicated in metabolism, immune system and neuronal function. Most strikingly, we found that a large fraction of differentially regulated genes between these two groups are involved in chemodetection. This may suggest that nonaggressive foundresses (in which they were almost exclusively upregulated) invest heavily in chemical communication to regulate social interactions—a possible mechanism by which cooperation and reduced aggression may have evolved in this system.

Despite different contexts and manifestations of aggression being studied, the limited overlap of differentially regulated genes in other insect species suggests that gene networks regulating aggressive behaviour are only partially conserved between species. Using a system genetics approach, we found that genes differentially regulated during the transition to an aggressive state in *P. californicus* are highly interconnected and form distinct modules of coregulated genes. The most highly connected genes we identified within these modules are prime candidates for being master control genes of aggressive behaviour in this species, and include genes encoding putative transcription factors, regulators of hormone metabolism, neuronal growth and function, and mitochondrial function. For the future, it will be interesting to further characterize these genes, investigate whether and how they are causally linked to aggression, and what role they may have played during the evolution of tolerance and cooperation in pleometrotic *P. californicus*.

© 2016 John Wiley & Sons Ltd

Acknowledgements

We would like to thank Mandy Man-Ying Tin for assistance with sample processing in the laboratory, the staff at the sequencing centre of the Okinawa Institute of Science and Technology for sequencing the samples, and Miquel Grau Lopez for uploading data to NCBI's Short Reads Archive. We are also grateful to Brian Haney, Ioulia Bespalova, Nathan Smith, Sky Eigen, Marisa Fernandez, Elizabeth Burr and Elizabeth Cash (all Arizona State University) for their help collecting specimens and recording behaviour in the laboratory. This work was supported by Okinawa Institute of Science and Technology subsidy funding (JSPS KAKENHI 25221206, 24770034 to ASM), and the National Science Foundation (DMS- 1313312 to YK and JF).

References

Agrawal AF, Brodie ED, Wade MJ (2001) On indirect genetic effects in structured populations. *The American Naturalist*, **158**, 308–323.

Aird SD, Aggarwal S, Villar-Briones A *et al.* (2015) Snake venoms are integrated systems, but abundant venom proteins evolve more rapidly. *BMC Genomics*, **16**, 647.

Alaux C, Sinha S, Hasadsri L *et al.* (2009) Honey bee aggression supports a link between gene regulation and behavioral evolution. *Proceedings of the National Academy of Sciences of the USA*, **106**, 15400–15405.

Anholt RRH, Mackay TFC (2012) Genetics of aggression. *Annual Review of Genetics*, **46**, 145–164.

Avet-Rochex A, Carvajal N, Christoforou CP *et al.* (2014) Unkempt is negatively regulated by mTOR and uncouples neuronal differentiation from growth control (ed. Desplan C). *PLoS Genetics*, **10**, e1004624.

Ayroles JF, Carbone MA, Stone EA *et al.* (2009) Systems genetics of complex traits in *Drosophila melanogaster*. *Nature Genetics*, **41**, 299–307.

Balas MT, Adams ES (1996) The dissolution of cooperative groups: mechanisms of queen mortality in incipient fire ant colonies. *Behavioral Ecology and Sociobiology*, **38**, 391–399.

Becnel J, Johnson O, Luo J, Nässel DR, Nichols CD (2011) The serotonin 5-HT7Dro receptor is expressed in the brain of *Drosophila*, and is essential for normal courtship and mating (ed. Zars T). *PLoS ONE*, **6**, e20800.

Benes H, Edmondson RG, Fink P *et al.* (1990) Adult expression of the *Drosophila* Lsp-2 gene. *Developmental Biology*, **142**, 138–146.

Bernasconi G, Strassmann J (1999) Cooperation among unrelated individuals: the ant foundress case. *Trends in Ecology and Evolution*, **14**, 477–482.

Bijma P, Muir WM, Ellen ED, Wolf JB, van Arendonk JA (2007) Multilevel selection 2: estimating the genetic parameters determining inheritance and response to selection. *Genetics*, **175**, 289–299.

Billuart P, Winter CG, Maresh A, Zhao X, Luo L (2001) Regulating axon branch stability: the role of p190 RhoGAP in repressing a retraction signaling pathway. *Cell*, **190**, 195–207.

Boomsma JJ, Kronauer DJ, Pedersen JS (2009) The evolution of social insect mating systems. In: *Organization of Insect Societies*, pp. 3–25. Harvard University Press, Cambrigde, Massachusetts.

Boomsma JJ, Huszár DB, Pedersen JS (2014) The evolution of multiqueen breeding in eusocial lineages with permanent physically differentiated castes. *Animal Behaviour*, **92**, 241–252.

Cahan S, Julian G (1999) Fitness consequences of cooperative colony founding in the desert leaf-cutter ant *Acromyrmex versicolor*. *Behavioral Ecology*, **10**, 585–591.

Clark FM, Faulkes CG (2001) Intracolony aggression in the eusocial naked mole-rat, *Hetercephalus glaber*. *Animal Behaviour*, **61**, 311–324.

Clark RM, Fewell JH (2014) Social dynamics drive selection in cooperative associations of ant queens. *Behavioral Ecology*, **25**, 117–123.

Clutton-Brock TH, Parker GA (1995) Punishment in animal societies. *Nature*, **373**, 209–216.

Collias NE (1944) Aggressive behavior among vertebrate animals. *Physiological Zoology*, **17**, 83–123.

Conesa A, Götz S, García-Gómez JM *et al.* (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics*, **21**, 3674–3676.

Creel S, MacDonald DW (1995) Sociality, group size, and reproductive suppression among carnivores. *Advances in the Study of Behavior*, **24**, 203–257.

Danecek P, Auton A, Abecasis G *et al.* (2011) The variant call format and VCFtools. *Bioinformatics*, **27**, 2156–2158.

Dengler-Crish CM, Catania KC (2007) Phenotypic plasticity in female naked mole-rats after removal from reproductive suppression. *Journal of Experimental Biology*, **210**, 4351–4358.

Edwards AC, Rollmann SM, Morgan TJ, Mackay TFC (2006) Quantitative genomics of aggressive behavior in *Drosophila melanogaster*. *PLoS Genetics*, **2**, e154.

Edwards AC, Ayroles JF, Stone EA *et al.* (2009) A transcriptional network associated with natural variation in *Drosophila* aggressive behavior. *Genome Biology*, **10**, R76.

Ferreira PG, Patalano S, Chauhan R *et al.* (2013) Transcriptome analyses of primitively eusocial wasps reveal novel insights into the evolution of sociality and the origin of alternative phenotypes. *Genome Biology*, **14**, R20.

Fewell JH (2003) Social insect networks. *Science*, **301**, 1867–1870.

Francis RC (1988) On the relationship between aggression and social dominance. *Ethology*, **78**, 223–237.

Gammie SC, Auger AP, Jessen HM *et al.* (2007) Altered gene expression in mice selected for high maternal aggression. *Genes, Brain and Behavior*, **6**, 432–443.

Garrison E, Marth G (2012) Haplotype-based variant detection from short-read sequencing. *arXiv*, 1207.3907.

Grabherr MG, Haas BJ, Yassour M *et al.* (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, **29**, 644–652.

Hartmann A, Wantia J, Torres JA, Heinze J (2003) Worker policing without genetic conflicts in a clonal ant. *Proceedings of the National Academy of Sciences of the USA*, **100**, 12836–12840.

Helms Cahan S, Fewell JH (2004) Division of labor and the evolution of task sharing in queen associations of the harvester ant *Pogonomyrmex californicus*. *Behavioral Ecology and Sociobiology*, **56**, 9–17.

Holbrook CT, Clark R, Jeanson R *et al.* (2009) Emergence and consequences of division of labor in forced associations of normally solitary sweat bees. *Ethology*, **115**, 301–310.

© 2016 John Wiley & Sons Ltd

Hölldobler B, Wilson EO (1977) The number of queens: an important trait in ant evolution. *Naturwissenschaften*, **64**, 8–15.

Hölldobler B, Wilson EO (1990) *The Ants*. Harvard University Press, Cambridge, Massachusetts.

Hwang JR, Siekhaus DE, Fuller RS, Taghert PH, Lindberg I (2000) Interaction of *Drosophila melanogaster* prohormone Convertase 2 and 7B2. Insect cell-specific processing and secretion. *The Journal of Biological Chemistry*, **275**, 17886–17893.

Inaki M, Yoshikawa S, Thomas JB, Aburatani H, Nose A (2007) Wnt4 is a local repulsive cue that determines synaptic target specificity. *Current Biology*, **17**, 1574–1579.

Jasper WC, Linksvayer TA, Atallah J, Friedman D, Chiu JC, Johnson BR (2015) Large-scale coding sequence change underlies the evolution of postdevelopmental novelty in honey bees. *Molecular Biology and Evolution*, **32**, 334–346.

Jeanson R, Fewell JH (2008) Influence of the social context on division of labor in ant foundress associations. *Behavioral Ecology*, **19**, 567–574.

Johnson RA (2004) Colony founding by pleometrosis in the semiclaustral seed-harvester ant *Pogonomyrmex californicus* (Hymenoptera: Formicidae). *Animal Behavior*, **68**, 1189–1200.

Johnson BR, Linksvayer TA (2010) Deconstructing the superorganism: social physiology, groundplans, and sociogenomics. *The Quarterly Review of Biology*, **85**, 57–79.

Johnson BR, Tsutsui ND (2011) Taxonomically restricted genes are associated with the evolution of sociality in the honey bee. *BMC Genomics*, **12**, 164.

Kapheim KM (2016) Genomic sources of phenotypic novelty in the evolution of eusociality in insects. *Science*, **13**, 24–32.

Kapheim KM, Pan H, Li C *et al.* (2015) Genomic signatures of evolutionary transitions from solitary to group living. *Science*, **348**, 1139–1143.

Keller L (1993) The assessment of reproductive success of queens in ants and other social insects. *Oikos*, **67**, 177–180.

Kobayashi M, Michaut L, Ino A *et al.* (2006) Differential microarray analysis of *Drosophila* mushroom body transcripts using chemical ablation. *Proceedings of the National Academy of Sciences of the USA*, **103**, 14417–14422.

Langfelder P, Horvath S (2008) WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*, **9**, 559.

Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology*, **10**, R25.

Lanzi G, de Miranda JR, Boniotti MB *et al.* (2006) Molecular and biological characterization of deformed wing virus of honeybees (*Apis mellifera* L.). *Journal of Virology*, **80**, 4998–5009.

Li B, Dewey CN (2011) RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics*, **12**, 323.

Li-Byarlay H, Rittschof CC, Massey JH, Pittendrigh BR, Robinson GE (2014) Socially responsive effects of brain oxidative metabolism on aggression. *Proceedings of the National Academy of Sciences of the USA*, **111**, 12533–12537.

Liebig J, Peeters C, Hölldobler B (1999) Worker policing limits the number of reproductives in a ponerine ant. *Proceedings of the Royal Society of London B: Biological Sciences*, **266**, 1865–1870.

Linksvayer TA (2015) The molecular and evolutionary genetic implications of being truly social for the social insects. *Advances in Insect Physiology*, **48**, 271–292.

Linksvayer TA, Wade MJ (2005) The evolutionary origin and elaboration of sociality in the aculeate Hymenoptera: maternal effects, sib-social effects, and heterochrony. *The Quarterly Review of Biology*, **80**, 317–336.

Liu X, Kiss I, Lengyel JA (1999) Identification of genes controlling Malpighian tubule and other epithelial morphogenesis in *Drosophila melanogaster*. *Genetics*, **151**, 685–695.

Lunter G, Goodson M (2011) Stampy: a statistical algorithm for sensitive and fast mapping of Illumina sequence reads. *Genome Research*, **21**, 936–939.

Manfredini F, Riba-Grognuz O, Wurm Y *et al.* (2013) Sociogenomics of cooperation and conflict during colony founding in the fire ant *Solenopsis invicta* (ed. Tsutsui ND). *PLoS Genetics*, **9**, e1003633.

McGlothlin JW, Moore AJ, Wolf JB, Brodie ED (2010) Interacting phenotypes and the evolutionary process III. Social evolution. *Evolution*, **64**, 2558–2574.

Monnin T, Ratnieks FL (2001) Policing in queenless ponerine ants. *Behavioral Ecology and Sociobiology*, **50**, 97–108.

Moore AJ, Brodie ED, Wolf JB (1997) Interacting phenotypes and the evolutionary process I. Direct and indirect genetic effects of social interactions. *Evolution*, **51**, 1352–1362.

Moore AJ, Haynes KF, Preziosi RF, Moore PJ (2002) The evolution of interacting phenotypes: genetics and evolution of social dominance. *The American Naturalist*, **160**, S186–S197.

O'Donnell S (1998) Dominance and polyethism in eusocial wasp *Mischocyttarus mastigophorus* (Hymenoptera: Vespidae). *Behavioral Ecology and Sociobiology*, **43**, 327–331.

Overson R, Gadau J, Clark RM, Pratt SC, Fewell JH (2014) Behavioral transitions with the evolution of cooperative nest founding by harvester ant queens. *Behavioral Ecology and Sociobiology*, **68**, 21–30.

Page REJ, Amdam GV (2007) The making of a social insect: developmental architectures of social design. *BioEssays*, **29**, 334–343.

Robinson MD, McCarthy DJ, Smyth GK (2009) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, **26**, 139–140.

Rodal AA, Blunk AD, Akbergenova Y *et al.* (2011) A presynaptic endosomal trafficking pathway controls synaptic growth signaling. *The Journal of Cell Biology*, **193**, 201–217.

Silk JB (2002) The form and function of reconciliation in primates. *Annual Review of Anthropology*, **31**, 21–44.

Simola DF, Wissler L, Donahue G *et al.* (2013) Social insect genomes exhibit dramatic evolution in gene composition and regulation while preserving regulatory features linked to sociality. *Genome Research*, **23**, 1235–1247.

Smith CR, Smith CD, Robertson H *et al.* (2011) Draft genome of the red harvester ant *Pogonomyrmex barbatus*. *Proceedings of the National Academy of Sciences of the USA*, **108**, 5667–5672.

Smith C, Helms Cahan S, Kemena C *et al.* (2015) How do genomes create novel phenotypes? Insights from the loss of the worker caste in ant social parasites. *Molecular Biology and Evolution*, **32**, 2919–2931.

Sporns O, Honey CJ, Kötter R (2007) Identification and classification of hubs in brain networks. *PLoS ONE*, **2**, e1049.

Toth AL, Tooker JF, Radhakrishnan S *et al.* (2014) Shared genes related to aggression, rather than chemical communication,

© 2016 John Wiley & Sons Ltd

are associated with reproductive dominance in paper wasps (*Polistes metricus*). *BMC Genomics*, **15**, 75.

de Waal FB (2000) Primates—a natural heritage of conflict resolution. *Science*, **289**, 586–590.

Wang Y, Ghaffari N, Johnson CD *et al.* (2011) Evaluation of the coverage and depth of transcriptome by RNA-Seq in chickens. *BMC Bioinformatics*, **12**, S5.

Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. *Evolution*, **38**, 1358–1370.

Wissler L, Gadau J, Simola DF, Helmkampf M, Bornberg-Bauer E (2013) Mechanisms and dynamics of orphan gene emergence in insect genomes. *Genome Biology and Evolution*, **5**, 439–455.

Wolf JB, Brodie ED, Cheverud JM, Moore AJ, Wade MJ (1998) Evolutionary consequences of indirect genetic effects. *Trends in Ecology and Evolution*, **13**, 64–69.

J.F., J.G. and M.H. conceived of the study and designed the research. M.H. conducted fieldwork and laboratory work, and led the behavioural observations. S.M. performed data assembly and differential gene expression analyses, with additional gene annotation provided by M.H. Y.K. and M.H. carried out the gene coexpression network analyses. M.H. drafted the manuscript, with critical contributions from all authors. All authors approved the final manuscript.

Data accessibility

The raw sequence data for this study have been deposited at NCBI's Short Reads Archive under BioProject number PRJDB4319. Transcripts assembled into contigs are available in Fasta format as Appendix S1 (Supporting information) provided with this article. In addition, gene expression data in counts and FPKM values are also provided as Appendices S2 and S3 (Supporting information), respectively.

Supporting information

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

Fig. S1 Dynamic range of RNA sequencing as determined by ERCC Spike-In Mixes.

Fig. S2 Volcano plots of differentially expressed genes between pure-pair haplometrotic and pleometrotic (A), and aggressive and nonaggressive *Pogonomyrmex californicus* (B) queens.

Fig. S3 Modules of coexpressed genes detected by hierarchical clustering based on a weighted correlation matrix of all 7890 transcripts, using WGCNA.

Table S1 Sample id and experimental groups assigned to *Pogonomyrmex californicus* foundresses used in the present study.

Table S2 Hub genes identified among the 524 differentially expressed genes between aggressive and cooperative foundresses.

Table S3 Mean average positive and negative degree per gene within and between hubs and modules, as well as average ratio of positive and negative edges per gene.

Appendix S1 Sequence data of 7890 assembled, filtered transcripts.

Appendix S2 Matrix of gene expression levels in counts.

Appendix S3 Matrix of gene expression levels in FPKM.

Table S4 Significantly differentially expressed genes between aggressive and nonaggressive queens.

Table S5 Significantly differentially expressed genes between queens in haplometrotic and pleometrotic pairs (Hh–Pp).

Table S6 Significantly differentially expressed genes between haplometrotic and pleometrotic singletons (Hs–Ps).

Table S7 Significantly differentially expressed genes between haplometrotic singletons and pairs (Hs–Hh).

Table S8 Significantly differentially expressed genes between haplometrotic singletons and pleometrotic pairs (Hs–Pp).

Table S9 Significantly differentially expressed genes between haplometrotic singletons and haplometrotic queens in mixed pairs (Hs–Hp).

Table S10 Significantly differentially expressed genes between haplometrotic singletons and pleometrotic queens in mixed pairs (Hs–Ph).

Table S11 Significantly differentially expressed genes between pleometrotic singletons and queens in haplometrotic pairs (Ps–Hh).

Table S12 Significantly differentially expressed genes between pleometrotic singletons and haplometrotic queens in mixed pairs (Ps–Hp).

Table S13 Significantly differentially expressed genes between pleometrotic singletons and pleometrotic queens in mixed pairs (Ps–Ph).

Table S14 Significantly differentially expressed genes between queens in haplometrotic pairs and pleometrotic queens in mixed pairs (Hh–Ph).

Table S15 Significantly differentially expressed genes between queens in pleometrotic pairs and haplometrotic queens in mixed pairs (Pp–Hp).

Table S16 Significantly differentially expressed genes between queens in pleometrotic pairs and pleometrotic queens in mixed pairs (Pp–Ph).

© 2016 John Wiley & Sons Ltd

Table S17 Significantly differentially expressed genes between haplometrotic queens in mixed pairs and pleometrotic queens in mixed pairs (Hp–Ph).

Table S18 Connectivity (k) and module membership (MM) of all 7890 transcripts based on WGCNA gene network approach, sorted by total connectivity (k_{Total}).

Table S19 Connectivity and network characteristics of 524 differentially expressed genes between aggressive and nonaggressive queens based on hub gene approach.

© 2016 John Wiley & Sons Ltd