Social divergence: molecular pathways underlying castes and longevity in a facultatively eusocial small carpenter bee

Wyatt A. Shell and Sandra M. Rehan

Department of Biology, York University, 4700 Keele Street, Toronto, Ontario, Canada M3J 1P3

Unravelling the evolutionary origins of eusocial life is a longstanding endeavour in the field of evolutionary-developmental biology. Descended from solitary ancestors, eusocial insects such as honeybees have evolved ontogenetic division of labour in which short-lived workers perform age-associated tasks, while a long-lived queen produces brood. It is hypothesized that (i) eusocial caste systems evolved through the co-option of deeply conserved genes and (ii) longevity may be tied to oxidative damage mitigation capacity. To date, however, these hypotheses have been examined primarily among only obligately eusocial corbiculate bees. We present brain transcriptomic data from a Japanese small carpenter bee, *Ceratina japonica* (Apidae: Xylocopinae), which demonstrates both solitary and eusocial nesting in sympatry and lives 2 or more years in the wild. Our dataset captures gene expression patterns underlying first- and second-year solitary females, queens and workers, providing an unprecedented opportunity to explore the molecular mechanisms underlying caste-antecedent phenotypes in a long-lived and facultatively eusocial bee. We find that *C. japonica*'s queens and workers are underpinned by divergent gene regulatory pathways, involving many differentially expressed genes well-conserved among other primitively eusocial bee lineages. We also find support for oxidative damage reduction as a proximate mechanism of longevity in *C. japonica*.

1. Introduction

The emergence of obligate eusocial life, in which a reproductive individual is supported by many non-reproductive relatives [1,2], is considered one of the major evolutionary transitions in biological complexity [3]. A defining feature of eusocial organization is reproductive division of labour, a pronounced form of phenotypic plasticity highlighting the capacity for a single genotype to produce multiple phenotypes under different environmental conditions [4]. For example, across various social insects including honeybees, a female develops either into a reproductive queen or a non-reproductive worker depending on the dietary and environmental conditions of her infancy [5]. This ontogenetic split is highly consequential, as queens and workers go on to exhibit substantially different physiologies and behavioural states. While a queen spends most of her life inside the nest, producing brood and interacting with her colony, her worker offspring perform a variety of complex, age-related tasks to support and protect the nest [6]. This phenotypic variation by caste extends to life expectancy: workers live for two to six weeks [7], but queens may spend up to 5 years continuously producing brood [8].

Research exploring caste-associated variations in honeybees and other highly social insects have provided a wealth of insights into the molecular mechanisms underpinning differences in behaviour and longevity [9–11]. For example, we now appreciate that both the developmental and behavioural plasticity of honeybees emerges from a highly directional regulatory network of modular suites of co-expressed genes [12,13]. Further, studies have shown that variations in longevity seen among social insects may be tied to individual
reproductive capacity [8,14,15], metabolic requirements and oxidation-reduction activity [16–19]. The relatively abbreviated honeybee worker lifespan, for instance, may partly be explained by the high metabolic costs of frequent flight, which causes increasing harm as oxidation-reduction activity fades with worker age [7,16]. By contrast, not only do long-lived queens rarely fly, there is evidence that reproduction-associated genes, such as vitellogenin, may also mitigate oxidative damage in this species [20,21].

Advanced eusocial bee lineages are known to have evolved from solitary ancestors [22]. Research among bee lineages of less derived sociality thus presents an exciting opportunity to empirically test major evolutionary hypotheses regarding the emergence of social organizational systems [22–24]. For instance, recent works have indicated that similar behavioural states and reproductive roles among independent social bee lineages may be underpinned by consistent differential expression of deeply conserved genes [25–27]. Research has also revealed that worker-like roles often feature expanded regulatory elements and evidence of positive protein evolution [28,29], as predicted by genetic release (in which otherwise pleiotropically constrained genes may be available for co-option into novel roles) [30]. Gene co-option and positive selection are expected to lead to the development of increasingly novel (taxonomically restricted) genes during further evolutionary derivations of social traits (e.g. emergence of a true worker caste) [22,31]. Additionally, comparative research is finding growing support for the role of sociality itself as exerting considerable influence on gene expression pathways associated with behavioural states [32–34] and lifespan among bee lineages [35]. There remains, however, a paucity of datasets which capture the transcriptomics of emergent social traits and ageing among facultatively eusocial bee taxa, which would allow for empirical testing of hypotheses that seek to bridge the effects of ageing and social environment on phenotypic plasticity [4,11,23,36].

*Ceratina japonica* is a long-lived species of small carpenter bee, which forms either solitary or eusocial nests in sympatry across its native range in Japan [37–39]. As is common to all social species of small carpenter bees around the globe [40–42], *C. japonica* forms or re-uses a linear burrow within the stems of pithy plants in which it rears around eight offspring a year [37,38]. In solitary nests, brood cell provisioning, guarding and rearing are accomplished by a single reproductive female. By contrast, social nests typically contain two adult females: the larger mother is reproducitively dominant and guards the nest, while the smaller daughter forages [38,39]. *Ceratina japonica* also lives and reproduces for 2 years [38]. In this way, *C. japonica* presents a novel opportunity to simultaneously test hypotheses regarding the evolution of social phenotypic plasticity and longevity in a species demonstrating facultative eusociality. Here, we investigate *C. japonica* brain transcriptomic data to (i) identify patterns of differential gene expression and cis-regulatory enrichment among this species’ three naturally co-occurring female classes (i.e. queens, workers and solitary females); (ii) examine molecular signatures of ageing within and among these classes and (iii) determine the degree to which molecular elements associated with sociality and longevity in *C. japonica* may be conserved among other social insect lineages. We predict that (i) *C. japonica’s* facultative primitive eusociality will be largely underpinned by deeply conserved and differentially expressed genes (DEGs), with only limited evidence of novel or taxonomically restricted genes; (ii) *C. japonica’s* queen and worker behavioural classes will be underpinned by distinct regulatory pathways and (iii) some measure of oxidation reduction will play a role in older individuals of all three behavioural classes regardless of other regulatory differentiation.

2. Material and methods

(a) Sample collection, sequencing and molecular analysis

*Ceratina japonica* queens, workers and solitary females were collected at dawn from nests found in broken stems of *Hydrangea* sp. around Sapporo, Japan in July 2015 following [43]. Individuals were flash frozen in liquid nitrogen upon collection and remained at or below −196°C until dissection. Nest type (solitary versus social) was determined through nest dissection, female identification and scoring individual wing wear and ovarian development. Social nests are established in pre-formed burrows (dirty walls) [38,39]. In social nests, workers develop from the first laid egg and are primarily foragers [38,39]. However, workers are capable of reproduction most worker-laid eggs are consumed by queens [38,39]. Queens, by contrast, are both the dominant egg layers and primary guards; both females regularly engage in trophallaxis [38,39]. Solitary nests are established in newly dug burrows (clean walls), in which the reproductive female lays eggs, forages for her brood and guards her offspring as she is able [38,39]. Nest types are largely stable from 1 year to the next: once established, nests typically remain solitary or social [38]. Social and solitary nest statuses were determined during nest dissection: those containing eggs and/or larvae along with two adult females were deemed social and those with only one adult female solitary. Individual ages were determined using nest condition (clean nest walls year 1; soiled, reused nests year 2; electronic supplementary material, table S1) and adult female wing wear (little to no wing wear, scored between 1 and 3, indicated young, first-year bees; moderate to severe wing wear, scored between 4 and 5, indicated old, second-year bees) following well-established methods (electronic supplementary material, table S1) [38,44]. In social nests, the wing wear of both females was taken into consideration to aid in assessing age. To assign classes within social nests, individuals were assessed for combined metrics of both body size (intersegmental distance and wing length) and ovarian development (sum of lengths of three largest ovarioles). In accordance with known biology [38,39,45], the larger and more reproducitively established of the two individuals was assigned to the queen class (electronic supplementary material, table S1) [38]. Following class determination, a total of 18 individuals (three young queens and three old; three young workers and three old; and three young solitary females and three old) were selected for brain RNA extraction using the QIAGEN RNaseasy Kit and protocol (cat. no. 73404; electronic supplementary material, table S2). Brain tissue was selected for comparison with other studies, relevance to behaviour, and as one of the most transcriptionally active tissues [46]. Acceptability of RNA sample quality was confirmed on an Agilent Tape Station 2200 before being submitted to Génome Québec (Montréal, Canada) for library preparation and 150 base pair (bp) paired-end Illumina HiSeq 2500 sequencing. Illumina sequencing produced an average of 39.9 MB of raw sequence data for each of our 18 brain samples (718.4 MB in total; electronic supplementary material, table S3). On average, 93% of raw sequence data for each sample passed initial quality checks for further alignment via FastQC, enabling the mapping of an average of 9231 genes at 35× read coverage across all samples after adapter removal via Trimmomatic (representing
3. Results

(a) Differential gene expression by behavioural class and age

Principal components analysis of total gene expression data revealed the majority of variation to be among workers and queens or solitary females. A total of 89% of this variation explainable by the first (5183 genes, 58%), second (1876 genes, 21%) and third components (893 genes, 10%; electronic supplementary material). The majority of this variation was significantly upregulated among old workers included cytochrome P450 6A1, and 6A1). Functional enrichment for old worker DEGs indicated a shift with age towards flight, locomotion and both metabolic (e.g. lipid metabolism) and immune-associated processes (e.g. haemocyte proliferation). Three cytochrome P450 s, an orthologue of major royal jelly protein 2, and both heat shock 70 kDa protein and oxidoreductase YrbE were upregulated in old queens over young. GO term enrichment revealed a shift towards both oxidation reduction and integral-mediated cell adhesion as queens’ age. Despite few DEGs between young and old solitary females, GO term enrichment revealed that solitary females also underwent a detectable shift towards increased immune responsiveness, protein methylation and oxidoreductase activity with age (electronic supplementary material, table S8).

Queens and workers shared more significantly upregulated genes in common than either did with solitary females, both among young (N = 5, 10%) and old individuals (N = 10, 11%). DEGs upregulated in both young queens and workers included putative gustatory receptor 23a, while old queens and workers shared both cytochrome p450 6A1 and major royal jelly protein 2 (electronic supplementary material, table S8). Young queen and worker DEG sets were enriched for 11 GO terms in common, but this count dropped with age to just three, representing a significant reduction in functional similarity (χ2-test, χ2 = 9.159, d.f. = 4, p = 0.0419; electronic supplementary material, table S6). Notably, all classes were enriched for oxidoreductase when old (electronic supplementary material, table S8).

(b) Network analysis

A total of 8936 genes and 16 of our 18 samples passed filtering checks for use in weighted gene co-expression network analysis (WG-CNA; electronic supplementary material, figure S4 [49]). Topographic rendering revealed that the most strongly trait-associated genes underlying queens (N = 447; electronic supplementary material, figures S9–S11 and tables S9, S14), workers (N = 224; electronic supplementary material, figures S15–S17 and tables S10, S14) and solitary females (N = 231; electronic supplementary material, figures S12–S14 and tables S11, S14) clustered mainly by class (figure 3; electronic supplementary material, figure S8). Worker-associated modules featured hub genes (i.e. genes with module membership (a measure of intramodular connectivity) and gene significance values greater than 0.9; p-values < 0.05) that were among the most isolated from the core network (figure 3). Worker hub genes included cytochrome b561, catalase and esterase FE4. Functional enrichment among genes from worker-associated modules suggested a role for immune processes (e.g. Toll signalling pathway), metabolic activity (e.g. oligosaccharide metabolism) and both muscular and neuronal pathways (electronic supplementary material, table S14).

(c) Cis-regulatory enrichment

Transcription factor (TF) binding site enrichment was determined using Stubb [50] and cis-Metalysis [51] referencing the JASPAR insect and vertebrate TF database [52]. This analysis identified significantly enriched and functionally unique up- or downregulating TFs associated both with behavioural (N = 429) and age classes (N = 568; electronic supplementary material, tables S15 and S16). Workers featured many more upregulating TFs than queens or solitary
were detected between Table S18). By far, the strongest correlations in gene expression social Hymenoptera (electronic supplementary material, among the reproductive dominants or subordinates of other queens or workers were similarly upregulated C. japonica material, tables S18 OrthoFinder v 2.3.2 [57] (figure 6; electronic supplementary expressed in both this study and others via BLASTn and included a handful found to be significantly differentially categories: non-reproductive females (workers) and reproductive females (i.e. queens and solitary females). (Online version in colour.)

Figure 1. (a) Summary illustrations of Ceratina japonica nesting biology. In social nests ((ii), reused nest), the queen acts as the primary reproductive and nest guard (in grey) while worker forages, in solitary nests (bottom, newly formed nest), the lone reproductive female produces, guards and forages for her brood (in grey). (b) Heat map of significantly DEGs (FDR corrected p values < 0.05; N = 471) identified among young (first year) and old (second year) queens, workers and solitary females (three samples per group; relative expression values are in log2fold change). Hierarchical cluster analysis reveals strong support for at least two focal categories: non-reproductive females (workers) and reproductive females (i.e. queens and solitary females). (Online version in colour.)

females, both by class (NWorkers = 223, versus NQueens = 30, NSolitary = 90) and age (NOldWorkers = 308, versus NOldQueens = 35, versus NOldSolitary = 31). Here, we consider enrichment only for those TFs with known neural, immune and reproductive regulatory associations. Overall, workers featured significantly greater neural-associated TFs (χ² = 6.85, d.f. = 2, p = 0.0325), but significantly fewer reproduction-associated TFs than expected given total TF counts (χ² = 11.12, d.f. = 2, p = 0.00385). By contrast, solitary females were enriched for significantly more reproduction-associated TFs than expected (electronic supplementary material, table S6). Counts of immune-associated TFs were not significantly different than expected for any role given respective total TF counts (χ² = 0.157, d.f. = 2, p = 0.924). Enriched binding site motifs accommodated TFs that were themselves differentially expressed in our dataset (i.e. BH1, eve, RREB1; figure 4) suggesting active regulation at time of sampling.

(d) Comparative analyses of gene expression
Rank-rank hypergeometric overlap (RRHO) analyses of gene expression variation [53] between queens and workers of C. japonica and comparable roles in eight additional hymenopteran species (C. australensis [25]; M. genalis [28]; C. calcara [29]; E. robusta; E. tridentata [34]; A. mellifera [54]; T. longispinosa [55]; and Polistes metricus [56]) collectively identified a total of 3328 genes significantly correlated in all comparisons to bees and a paper wasp (p < 0.0001; figure 5; electronic supplementary material, table S17). Notably, this list of genes included a handful found to be significantly differentially expressed in both this study and others via BLASTn and OrthoFinder v 2.3.2 [57] (figure 6; electronic supplementary material, tables S18–S20). Specifically, DEGs upregulated in C. japonica queens or workers were similarly upregulated among the reproductive dominants or subordinates of other social Hymenoptera (electronic supplementary material, table S18). By far, the strongest correlations in gene expression were detected between C. japonica and other Ceratina species (i.e. C. calcara [29] and C. australensis [25]) and between C. japonica and other primitive eusocial bees (M. genalis [28] and E. robusta [34]).

RRHO comparisons of gene expression by age within C. japonica and both C. calcara [29] and A. mellifera [58] identified an additional 2402 genes significantly correlated in all but one comparison among classes (p < 0.0001; electronic supplementary material, table S23). Overall, this suite of analyses further corroborated the trends of molecular divergence between C. japonica’s worker and queen classes (electronic supplementary material, figure S18). Orthologues consistently upregulated in older foraging individuals in all three species included probable cytochrome P450 305a1, glucose dehydrogenase and alpha-glucosidase.

DEGs uniquely upregulated in workers also accounted for the majority of C. japonica DEGs identified as under positive selection in Shell et al. [34] (NWorkers = 21, NQueens = 6, NSolitary = 4; electronic supplementary material, table S21) and the majority of C. japonica DEGs orthologous with genes identified under positive selection in studies of honeybees (59], N = 17/19, 89%) and other small carpenter bees [60] (N = 14/14, 100%; electronic supplementary material, table S22). GO term enrichment for this set included processes associated with positive chemotaxis and both carbohydrate and amino sugar metabolism (electronic supplementary material, table S22). DEGs uniquely upregulated in queens accounted for three and solitary females no orthologues under positive selection in other taxa (electronic supplementary material, table S18). Notably, both queen and worker DEGs under positive selection were significantly enriched for oxidoreductase activity.

4. Discussion
Here, we provide the first insights into the molecular dynamics underlying both caste-antecedent behavioural classes and ageing in Ceratina japonica, a long-lived species of small carpenter bee capable of eusocial nesting. We uncover considerable transcriptomic variation among queens, workers and solitary females, and highlight how
regulatory enrichment and a modular co-expression network may play important roles in *C. japonica*’s social phenotypes. We also show how correlations between DEGs underlying *C. japonica*’s form of eusociality and those observed across other social taxa provide empirical support for the role of deeply conserved genes in the evolution of social traits [22,61]. Finally, we reveal how gene expression patterns shift by class over *C. japonica*’s 2-year lifespan and identify the putative role of oxidation reduction as a critical proximate mechanism for longevity in a long-lived and facultatively social bee.

(a) Workers are highly distinct from queens and solitary females

Our analyses highlight reproductively subordinate workers as especially distinct from queens and solitary females. Workers are underpinned by almost 70% of all DEGs and the majority of genes under positive selection in this and other social bees [33,59,60]. Workers also feature extended enrichment for neural-associated TFs, some of which are themselves differentially expressed (e.g. homeobox protein...
**Figure 4.** Heat map highlighting TFBS motifs with known neural, immune or reproductive regulatory roles, significantly enriched in the promoter regions of genes associated with queens, workers or solitary females (for full list, see electronic supplementary material, table S15). Motif names are presented in order of phenotypic association with queens, workers or solitary females (for full list, see electronic supplementary material, table S15). The rightmost column contains phenotype-association and module information for all motifs from WGCNA. Queens are generally downregulated and workers generally upregulated, and solitary females fall somewhere in between. (Online version in colour.)

Barth1). Other worker enriched TFs, such as apterus, CTCF and Lim3, have been previously associated with worker-like phenotypes, including foraging, across other incipiently social (e.g. *C. australensis* [25]; *C. calcatrix* [29]) and eusocial bees (e.g. *A. mellifera* [38]; *E. robusta, Exoneurella tridentata* [34]). Network analysis further indicates that genes expressed in workers are not upregulated in solitary female or queen-associated pathways [49,62]. Our study thus corroborates previous observations that caste-associated phenotypes may be tied to highly modular sets of co-expressed genes [12,29]. Evidence of positive selection almost exclusively in worker-associated genes suggests support for the hypothesis of genetic release [30], which predicts that pleiotropic constraints are likely to loosen in the non-reproductive role of a social lineage.

Three genes—odorant receptor coreceptor and two copies of general odorant binding protein 69a (Obp69a)—were identified as both highly expressed worker-associated DEGs and functionally important worker hub genes. Odorant-binding proteins (OBPs) and receptors (ORs) underpin insect olfactory behaviour and pheromone binding [63]. At least one OBP, Obp69a, has also been shown to help moderate social interactions among fruit flies [64]. The role of ORs and OBPs in chemical communication is relatively well-appreciated among honeybees (*Apis mellifera* [65,66]) and other advanced eusocial Hymenoptera (i.e. ants [67] and wasps [68]), as they likely contribute to nest-mate recognition and the direction of caste behaviour and physiology [69,70]. OBPs have also been found to play an important role in other eusocial bees (e.g. *Bembix terestris* [71]) and still-flexible reproductive dominance hierarchies of incipiently social small carpenter bees [25,46]. While their exact physiological function within *C. japonica* remains a question for future research, the evident involvement of ORs and OBPs in *C. japonica*’s worker class highlights chemical communication as of potential importance for the behavioural dynamics of this species [72,73].

(b) Social phenotypes appear to be undergoing molecular divergence

Overall, the molecular pathways underlying *C. japonica*’s queen and worker phenotypes are highly class specific and appear to have opposing directional regulation. For example, a set of 128 enriched TFBS motifs upregulated worker-associated and downregulated queen-associated expression profiles. Extensive gene downregulation in queens is perhaps not surprising given behavioural observations of this species: queens are at least partly characterized by two negative key tasks—continuous resting and abandonment of foraging—[74]. By comparison, we find that solitary female regulatory enrichment and directionality partly overlaps with both queens and workers. Network analysis corroborates this phenotypic directionality: modules are generally clustered by class, but many genes remain partially co-expressed among classes. In sum, rather than the decoupled queen and worker pathways of advanced eusocial Hymenoptera (e.g. *Temnothorax longispinosus* [55]; *A. mellifera* [54]), *C.

<table>
<thead>
<tr>
<th>motif</th>
<th>regulatory role</th>
<th>queen</th>
<th>worker</th>
<th>solitary</th>
<th>DEG</th>
<th>WGCNA</th>
</tr>
</thead>
<tbody>
<tr>
<td>ap</td>
<td>muscle development, juvenile hormone production, neuronal path finding</td>
<td>solitary</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>br_Z1</td>
<td>oogenesis; reproduction; dendrite morphogenesis; salivary gland morph and function</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>br_Z2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>br_Z3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eip74EF</td>
<td>development, reproduction, immunity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RIREB1</td>
<td>transcriptional regulator</td>
<td>W &gt; Q, W &gt; S</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CREB1</td>
<td>neuronal differentiation, circadian rhythm</td>
<td>queens, workers</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FEV</td>
<td>serotonergic neurons</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deaf1</td>
<td>peripheral tissue antigens; involved in visual learning</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B.H1</td>
<td>eye and external sensory organ development</td>
<td>W &gt; S</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>eve</td>
<td>central nervous system development</td>
<td>W &gt; Q, S &gt; Q</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amt</td>
<td>master regulator of hypoxia, xenobiotic, and angiogenetic pathways</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Proc. R. Soc. B* 289, 20212663
**Figure 5.** RRHO plots portraying correlations between genes ranked by association with queens through workers in *C. japonica* (all x-axes) to homologous genes associated with comparable roles in eight additional hymenopteran species (y-axes). Overall gene list comparisons and overlaps are presented in the electronic supplementary material, table S17. The most positively and significantly correlated comparisons were between *C. japonica* and (i) other ceratine species, despite variations in complexity of social phenotype and (ii) other primitively eusocial bees, *Megalopta genalis* and *Eunera robusta*, despite phylogenetic distance between lineages. Although correlations are also significant among advanced eusocial bees and wasps, they are also more localized by queen or worker phenotype at this level. (Online version in colour.)

*C. japonica*’s queens and workers may still be undergoing molecular divergence, both from solitary females and each other. Overall, *C. japonica*’s brain gene expression patterns more closely resemble those of other facultatively eusocial bees, such as *Euglossa dilemma* [26] or *M. genalis* [28], in which observably caste-antecedent phenotypes remain flexible at both the molecular and phenotypic levels.

**(c) Molecular signatures of longevity in *Ceratina japonica***

Class-specific shifts in gene expression and regulatory enrichment with age suggest queens, workers and solitary females undergo separate changes according to life history. Although workers may experience the most dramatic changes of any single phenotype, both workers and queens upregulate some of the same immune-associated genes with age. These results suggest that *C. japonica* social nest females may experience age-associated changes that are more typical of species with larger and better-established social nest environments (e.g. *A. mellifera* [75]; *Bombus terrestris* [76]). Extensive regulatory increases and expanded flight, odorant binding and metabolic activity in *C. japonica* workers mirror the age-associated changes observed in workers of more advanced eusocial bees (e.g. *A. mellifera* [6]). Further, elevated immune activity among queens and workers suggests even a small social nest may invite similar pathogenic and oxidative challenges to those of larger eusocial colonies [77,78].

Queens, workers and solitary females of *C. japonica* all live up to 2 years or more [45], suggesting longevity may be underpinned by some shared molecular mechanisms despite distinct life histories. Accordingly, we identified significant enrichment for oxidation-reduction processes across old individuals regardless of phenotypic class (electronic supplementary material, table S5). An ability to mitigate oxidative damage has been consistently and independently implicated as critical to the longevity of many aerobic taxa, from bivalves [79] to humans, mammals and birds [80–82], and among social insects [17,83,84]. Among advanced eusocial bees, older foragers may offset oxidative damage through increased expression of redox-related proteins, such as catalase [83]. Catalase and four copies of esterase fe4 were found to be strongly and significantly associated with *C. japonica* workers during multiple analyses. Like catalase, esterase fe4 plays an essential role in mitigating damage by oxidation or environmental stress (such as insecticides) across many insects, including social bees (*A. mellifera* [85]; *A. cerana* [86]). Longevity in *C. japonica*’s worker class may also be explained by the fact that workers remain reproductively viable [32,33], a trait that is positively correlated with longevity across social insects [14,18,21,87]. Overall, the transcriptomics of ageing in *C. japonica* supports the propositions that (i) individual longevity in social insect lineages may be tied to ancestrally reproductive pathways and activity [21,87,88] and (ii) oxidation–reduction activity is a key predictor of aerobic organisinal lifespan [17,82,89].

(d) Deeply conserved and differentially expressed genes underlie primitive eusociality

Comparative analyses reveal that genes and gene expression patterns associated with queen and worker classes in *C. japonica* are both well-conserved and highly correlated across other *Ceratina* and among primitively eusocial bees. Both *C. calcara* and *C. austroleonis* demonstrate incipiently social phenotypes [90,91] and, as relatively close congers of *C. japonica*, share many of the life-history traits consistent across the small carpenter bee genus (e.g. stem nesting and alloparental care [37]). Molecular consistencies among the small carpenter bees may therefore be explained as much by broad similarities in biology as they are by shared ancestry within a genus [40]. By contrast, *C. japonica* is a distant cousin of primitively eusocial *E. robusta* (*Xylocopinae*) [92] and of a separate family from *M. genalis* (*Halictinae*) [93]. Strongly positive and significant correlations in gene expression patterns detected among these species are therefore likely better explained by consistencies in phenotypic traits [94–96] than by shared phylogenetic ancestry. Our study thus joins recent works among Hymenoptera (e.g. bees [34]) and other social animals (e.g. snapping shrimps [33]) in offering empirical support for the theory that sociality itself may exert a broadly consistent influence on otherwise separate evolutionary trajectories [32]. Our results also offer support for the theory that non-obligate forms of eusocial behaviour are underpinned primarily by deeply conserved genes and DEGs. Comparisons of *C. japonica* to species that have passed the evolutionary
### References


---

### Figure 6

A representative subset of all genes associated with queen or worker phenotypes in *C. japonica* identified during both orthofinder and RRHO analyses and matched with strong statistical support to genes in *Ceratina* (i.e. *C. calcarata* and/or *C. australensis*), primitively eusocial bees (e.g. *E. robusta*, *M. genalis* or *E. dilemna*), advanced eusocial bees (e.g. *A. mellifera*) and/or other eusocial Hymenoptera (e.g. *Polistes metricus* and *Tennmothorax longispinosus*). Gene accession numbers are located in the electronic supplementary material, table S7. Blue boxes indicate orthologous genes and similar regulatory contexts between *C. japonica* and at least one species from the indicated group; white boxes indicate a lack of contextual or regulatory overlap. Genes associated with queen and worker phenotypes in *C. japonica* are very well conserved among similar roles in other ceratinine taxa, both primitively and advanced eusocial bee lineages, and to a lesser extent among other eusocial Hymenoptera (electronic supplementary material, tables S17, S19 and S20). (Online version in colour.)

<table>
<thead>
<tr>
<th>gene family</th>
<th>queen</th>
<th>worker</th>
</tr>
</thead>
<tbody>
<tr>
<td>TBC1 domain family member 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>muscle LIM protein Mlp84B</td>
<td></td>
<td></td>
</tr>
<tr>
<td>purine nucleoside phosphorylase</td>
<td></td>
<td></td>
</tr>
<tr>
<td>transcription factor kayak</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cuticulin-1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cadherin-89D</td>
<td></td>
<td></td>
</tr>
<tr>
<td>neurogenic locus notch homologue protein 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>trehalase</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

‘point of no return’ into advanced eusociality [97] also indicate that taxonomically and phenotypically restricted genes may play an increasingly focal role in these lineages [22].

**Data accessibility.** Read data are accessible via NCBI PRJNA 413373.

**Authors’ contributions.** W.A.S.: formal analysis, investigation, visualization and writing—original draft; S.M.R.: conceptualization, data curation, formal analysis, funding acquisition, investigation, project administration, supervision and writing—review and editing.

All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

**Competing interests.** We declare we have no competing interests.

**Funding.** This work was supported by funding from an NSERC Discovery Grant to S.M.R. and NSF Graduate Research Fellowship 1450271 to W.A.S.

**Acknowledgements.** We thank Scott Groom for assistance with field collection of bees and Génome Québec for RNA library preparation and Illumina sequencing.


75. Iovinella I et al. 2011 Differential expression of odorant-binding proteins in the mandibular glands of the honey bee according to caste and age. J. Proteome Res. 10, 3439–3449. (doi:10.1021/pr1000754)