

## Accepted Manuscript

Title: A molecular concept of caste in insect societies

Author: Seirian Sumner Emily Bell Daisy Taylor

PII: S2214-5745(17)30110-4  
DOI: <https://doi.org/doi:10.1016/j.cois.2017.11.010>  
Reference: COIS 402



To appear in:

Received date: 26-9-2017  
Revised date: 21-11-2017  
Accepted date: 21-11-2017

Please cite this article as: Seirian SumnerEmily BellDaisy Taylor A molecular concept of caste in insect societies (2017), <https://doi.org/10.1016/j.cois.2017.11.010>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

1 **A molecular concept of caste in insect societies**

2 Seirian Sumner<sup>1</sup>

3 Emily Bell<sup>2</sup>

4 Daisy Taylor<sup>2</sup>

5

6 <sup>1</sup>Centre for Biodiversity and Environmental Research, Medawar Building, University  
7 College London, Gower Street, London WC1E 6BT

8

9 <sup>2</sup>School of Biological Sciences, Bristol Life Sciences Building, University of Bristol,  
10 24 Tyndall Avenue, Bristol BS8 1TQ, UK

11

12 Corresponding author: s.sumner@ucl.ac.uk

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31

32

33

**Abstract**

34 The term 'caste' is used to describe the division of reproductive labour that defines  
35 eusocial insect societies. The definition of 'caste' has been debated over the last 50  
36 years, specifically with respect to the simplest insect societies; this raises the  
37 question of whether a simple categorisation of social behaviour by reproductive state  
38 alone is helpful. Gene-level analyses of behaviours of individuals in hymenopteran

39 social insect societies now provide a new empirical base-line for defining caste and  
40 understanding the evolution and maintenance of a reproductive division of labour.  
41 We review this literature to identify a set of potential molecular signatures that,  
42 combined with behavioural, morphological and physiological data, help define caste  
43 more precisely; these signatures vary with the type of society, and are likely to be  
44 influenced by ecology, life-history, and stage in the colony cycle. We conclude that  
45 genomic approaches provide us with additional ways to help quantify and categorise  
46 caste, and behaviour in general.

47

48 Keywords: social evolution, eusocial insects, genomics, phenotypic plasticity

49

50

51

52

53

54

55

56

57

58

59

60

61

62

63

64

65

66

67

68

69

70

71

## **Introduction**

72

73

74

75

76

77

One of the defining traits of the eusocial insects (ants, some bees and wasps) is a division of reproductive labour achieved through the evolution of dimorphic castes whereby group members specialise in complementary and contrasting behaviours as either queens (reproductives) or workers (non-reproductives). The evolution of queen and worker castes is the key to the ecological success of these insects [1,2]. It is not surprising, therefore, that in the quest to understand how and why eusociality

78 evolves, researchers have focused on revealing the mechanisms, evolution and  
79 function of castes [3–6]. A long-debated issue is the equivalent use of the term  
80 ‘queen’ and ‘worker’ to describe division of labour across all types of eusocial  
81 societies, and specifically: is the concept of caste theoretically relevant [4,7] and  
82 empirically meaningful for the simple eusocial societies?  
83

84 Biologists categorise castes based on behaviour and physiology (Table 1). However,  
85 the last 5 years has seen a burgeoning of data on gene-level differences among  
86 caste in a wide range of eusocial Hymenopteran insects. These data provide an  
87 objective, quantitative measure of caste differentiation. Here [0]we discuss how  
88 recent molecular analyses support the opinion that a dichotomic concept of castes is  
89 too coarse a distinction to properly describe and understand cooperation across the  
90 spectrum of social complexity. We define a set of data-driven hypotheses on putative  
91 molecular signatures of caste for the eusocial Hymenoptera (ants, some bees and  
92 wasps) and advocate an approach that unites these measures of genomic variation  
93 (the ‘molecular phenotype’; e.g. gene expression, regulation and functionality) with  
94 classical measures of phenotypic variation (e.g. behaviour, physiology and  
95 morphology). Due to data availability and potentially confounding contrasting life-  
96 history features, our analyses here are restricted to the eusocial Hymenoptera (ants,  
97 some bees and wasps), although a similar approach could be taken to define  
98 hypotheses for other eusocial groups like the termites [8] or eusocial mammals (e.g.  
99 naked mole rats). We suggest that the integration of quantitative molecular data with  
100 behavioural and physiological traits has much to offer our general understanding of  
101 the evolution of caste differentiation and division of labour in eusocial societies.  
102

### 103 **Defining caste and why it matters**

104 The three key traits of a eusocial species are that it exhibits overlapping generations,  
105 cooperative brood care, and division of labour [9,10]. ‘Caste’ is the mechanism by  
106 which one of these traits - division of labour - is achieved, and it provides phenotypic  
107 variation on which selection can act. The term ‘caste’ was historically used to  
108 describe the fixed, morphologically distinct queens and workers of the complex  
109 eusocial species, like the honeybee, where workers are committed to a lifetime of  
110 functional sterility (e.g. [11,12]). ‘Caste’ has since become a more ubiquitous term,  
111 used to describe division of labour in all eusocial societies, including those that do  
112 not have lifetime, functionally sterile phenotypes. For example, in the space of a  
113 decade, E.O. Wilson’s use of the term ‘caste’ changed from one that required

114 'morphologically distinct' traits [3] to a much looser definition whereby individuals are  
115 'specialised...for prolonged periods of time'...and...'some kind of additional  
116 marker'... which may change with age, and may not involve external morphology  
117 [10]. This latter definition includes facultatively eusocial species, where generations  
118 are overlapping, there is cooperative brood care, but division of labour (and thus  
119 castes roles) can be temporary and distinguished only by behaviour, (e.g.  
120 Stenogastrinae wasps; Halictid bees). Currently, the term 'caste' is used as a catch-  
121 all term to describe the full phenotypic diversity of division of labour in eusocial  
122 insects. However, the applicability of one term (i.e. caste) to describe societal  
123 coordination across all species has been questioned and much of the terminology  
124 surrounding castes remains misleading [1,9,13–15].

125  
126 Determining precisely the traits used to classify individuals into castes has important  
127 implications for the study of eusocial evolution. Indeed, the degree to which castes  
128 are differentiated is a key characteristic used to infer the level of social complexity a  
129 species occupies, and by inference, what stage in the process of social evolution it  
130 may represent [6,16]. The importance of this is illustrated by the recent debate over  
131 whether all eusocial species are in fact 'superorganisms' and truly represent a major  
132 transition in evolution [7]. The key traits used by these authors to determine whether  
133 a species is a superorganism are in fact those which are used to define caste:  
134 specifically, loss of totipotency and permanent commitment (during development) to  
135 a life-time of either reproduction (as queens) or helping (as workers) [7]. This  
136 argument takes the concept of caste beyond semantics and a desire to order the  
137 disorder of biology: if caste is being used as a benchmark for defining evolutionary  
138 processes, we need clearer, more precise ways to categorise and describe it.

139  
140 Ambiguity in the use of the term caste is a long-standing problem (Table 1); we lack a  
141 consistent set of quantifiable traits for defining and categorising caste. Almost all  
142 definitions agree that 'caste' describes **functionally distinct** phenotypes: indeed,  
143 this is necessary for a division of labour. Definitions differ on whether castes are  
144 **irreversible**, such that individuals show lifetime commitment to a specific role, or  
145 flexible, whether a **morphological distinction** is important, or whether **behavioural**  
146 **distinction** alone is sufficient. Complications arise for species which traverse these  
147 criteria; for example, some swarm-founding Polistine wasps (e.g. *Polybia*  
148 *occidentalis*) where caste is apparently irreversible but individuals lack morphological  
149 differentiation. Such confusion can be resolved if species are instead defined as  
150 those in which caste is fixed during development (this includes all species with

151 irreversible castes, but does not require there to be morphological differences), or  
152 whether caste remains plastic throughout life (these are species whose functional  
153 roles are 'reversible', and that lack morphological differences). However, without  
154 detailed knowledge of the plasticity of castes in a species, this definition is of limited  
155 practical use. Moreover, does this mean that species without developmentally  
156 determined roles lack 'real castes' [13]? We require a set of traits that are readily  
157 quantifiable in order to accurately describe castes, and use this to categorise a  
158 species in relation to its level in the evolution of social complexity.

159  
160 The heart of the problem in defining castes is that, like most biological systems, caste  
161 is a complex phenomenon and boundaries between castes are often messy.  
162 Expression of caste traits (as defined in Table 1) varies considerably. Castes can be  
163 so different in appearance that queens and workers can appear to belong to different  
164 species (e.g. *Attine* ants; Figure 1). In contrast, other castes are discernible only from  
165 behavioural observations and ovary dissections (e.g. *Polistes* wasps; Figure 1). The  
166 simpler eusocial societies present the main issue here: they are functionally distinct,  
167 but lack lifetime commitment, morphological distinction and developmental  
168 differentiation. As a result, some authors have suggested that the simplest eusocial  
169 insect societies are better described as 'casteless' and as cooperative breeders  
170 [4,7,12]. Further, traits used to define castes may vary within the lifetime of a colony  
171 (e.g. *Polistes* foundresses can behave like queens until the first workers emerge, and  
172 thereafter behave like workers [18]), and within the lifetime of an individual (e.g.  
173 *Bombus* workers can develop ovaries and become unmated reproductives late in the  
174 colony's life). Finally, ecological conditions can influence expression of caste traits  
175 within and across species: for example, 'worker' traits are only expressed by Halictid  
176 bees living in warm climatic regions, with long summers [19].

177  
178 An appreciation of the sources of 'messiness' when defining castes is essential.  
179 Taking account of how social parameters, ecological variables and molecular  
180 processes influence expression of caste traits is required in order to categorise and  
181 explain the patterns observed between reproductive and non-reproductive individuals  
182 in eusocial insect colonies [13].

183

## 184 **Molecular signatures as quantifiable traits for defining castes**

185 The 21<sup>st</sup> century explosion in molecular techniques allows us to scrutinise the  
186 concept of caste at the genomic level. We propose data-driven hypotheses for how

187 genomic analyses of phenotypes may contribute towards a more quantitative  
188 definition of castes in eusocial insects (summarised in Table 2). These hypotheses  
189 fall into three main facets of genomic variation: gene expression patterns, regulatory  
190 processes and functionality.

### 191 **Differential gene expression patterns as molecular signatures of castes**

192 Levels of differential gene expression between castes differ greatly across species,  
193 and may be indicative of the level of social complexity. For example, castes in  
194 eusocial species with simple societies (e.g. *Polistes*) appear to differ very little in  
195 transcription, with less than 1% of detected genes being differentially expressed  
196 [22,23]. The low levels of transcriptional differentiation that underlie reproductive and  
197 non-reproductive phenotypes in these societies [22–24] reflect the relative lack of  
198 caste specialisation and commitment to specific behavioural and/or physiological  
199 roles. In these species, the majority of differentially expressed genes (DEGs) are  
200 down-regulated in queens compared to workers [22,25,26]. In these societies of  
201 highly plastic phenotypes, queens could be classified as ‘shut-down workers’ – i.e.  
202 reproductive workers with a reduced behavioural repertoire. In bumblebees (*Bombus*  
203 *spp*), reproductive workers show comparable gene expression profiles to queens;  
204 however, non-reproductive workers and queens differ greatly in their patterns of  
205 transcription with a total of 5316 DEGs between castes, 2817 up-regulated in queens  
206 and 2799 up-regulated in non-reproductive workers [27]. The large transcriptional  
207 differences between castes in the more complex eusocial societies (e.g. the  
208 honeybee *Apis mellifera*) reflect developmentally-determined castes, resulting in  
209 individuals which show lifetime commitment to a role and (usually) morphological  
210 differentiation [28,29]; queens typically up-regulate more caste-specific genes than  
211 workers [27,30]. This pattern of caste-specific expression is established during larval  
212 development, where queen-destined larvae up-regulate at least 70% of the  
213 differentially expressed genes [28,29].

214

215 The degree of transcriptional differentiation between castes, along with the putative  
216 contrasting patterns in the direction of caste-biased expression (i.e. whereby there is  
217 a general up-regulation of worker-biased genes in simple societies, but a general up-  
218 regulation in queen-biased genes in more complex societies) are likely to be  
219 important molecular signatures of caste (Table 2). As such, these traits may be  
220 indicative of the level of social complexity, reflecting a hypothetical transcriptional  
221 ‘tipping-point’ in species where caste commitment has evolved.

**222 Regulatory mechanisms as molecular signatures of caste**

223 Epigenetic mechanisms control gene expression by differentially regulating genes in  
224 response to environmental or genetic cues [31]; they can also limit the plasticity of  
225 gene expression, fixing specific transcriptional patterns irreversibly [31–33].

226 Epigenetic canalization is therefore a key hypothesis in the regulation of the inflexible  
227 roles found in eusocial insect species where castes are determined (irreversibly)  
228 during development [34,35]. An important question, however, is whether the same  
229 epigenetic processes regulate caste in the simpler societies, and whether patterns of  
230 epigenetic regulation can be useful signatures of caste evolution and sociality  
231 [34,35].

232  
233 MicroRNAs (miRNAs), 21-23bp RNAs which specifically target mRNAs and control  
234 their translation into proteins, are potential epigenetic candidates in the regulation of  
235 caste determination [29,36]. In the simple societies of *Polistes* wasps, miRNA-  
236 targeted genes show no caste-specific expression between queens and workers [23],  
237 potentially reflecting the caste plasticity of these insects. By contrast, in the more  
238 complex societies, e.g. bumblebees and honeybees, genes targeted by miRNAs  
239 show differential expression between queen and worker-destined larvae [29,36]  
240 (Table 2). Caste-specific targets for miRNAs include the ecdysteroids, involved in  
241 insect development [29,36], as well as genes related to structural differentiation [29],  
242 e.g. *Distal-less* (antennae and proboscis development) and *No extended memory*  
243 (imaginal disc-derived wing morphogenesis) [36]. Although data are currently limited,  
244 the conservation or differential expression of miRNAs among phenotypes may prove  
245 to be useful molecular signatures of caste and play a useful role in identifying the  
246 target developmental pathways for gene-specific silencing experiments.

247  
248 DNA can be chemically modified by the addition of DNA methyltransferases (DNMTs)  
249 which subsequently regulate gene expression; representatives from this group of  
250 enzymes are found in all insect orders, albeit with significant variation [34,37,38].

251 Eusocial Hymenoptera show some of the lowest levels of methylation among insects  
252 [39], and moreover methylation rates and patterns vary considerably across levels of  
253 social complexity [37]. The highly plastic phenotypes of the simple societies, with  
254 putatively reversible castes, show only limited difference in brain methylation  
255 changes between castes, [22,23]. Species with developmentally-determined castes  
256 (which are irreversible) often exhibit high levels of methylation at key genes and  
257 these are associated with caste-biased genes [40] (reviewed in [38]) (Table 2).

258 Levels of brain methylation, therefore, may be an indicator of caste differentiation and



259 social complexity [41]. However, the data are inconclusive. At the species level,  
260 whole body analyses of methylation found little support for a correlation between  
261 methylation and sociality: for example, methylation levels in non-social insects are  
262 not consistently lower than social species, and even within the eusocial Hymenoptera  
263 there is no clear correlation between methylation levels and social complexity [39,42].  
264 One source of variation that may account for the muddy story of the role of  
265 methylation in castes and sociality is the tissue analysed and level of analysis; i.e.  
266 caste-specific and brain tissue [22,23,40] versus species level and/or whole bodies  
267 [8,30,39]

268

269 Chromatin modifications occur via histone post-translational modifications (PTM)  
270 [43,44]. The available data suggest that variation in histone modifications strongly  
271 correlates with caste-biased genes in ants and honeybees [45,46]. Caste  
272 determination during larval development in *Apis* is modulated nutritionally via the  
273 feeding of royal jelly; this substance contains a histone deacetylase inhibitor (HDACi)  
274 which instigates contrasting pathways in queen and worker-destined larvae [46].  
275 HDACi has also been linked to behavioural reprogramming, where it induces an  
276 increase in foraging in ant workers [43]. Queens in both ants and bees exhibit the  
277 highest levels of histone modifications [45,47], with extensive alterations  
278 concentrated in the ovaries; likewise, workers exhibit distinct patterns of histone  
279 PTMs [45]. This work on eusocial species with complex societies suggests that  
280 histone PTMs could provide a molecular signature of caste (Table 2); however, to  
281 date nothing is known about these processes in eusocial species with simple  
282 societies. Determining the capacity of a phenotype to return to developmental  
283 pluripotency via epigenetic reprogramming [48] may help us define castes by the  
284 presence or absence of a molecular marker that imposes commitment on a  
285 phenotype.

### 286 **Functional specialisations as molecular signatures of caste**

287 General patterns of functional enrichment may be useful signatures of caste, and  
288 may be especially useful in determining the level of social complexity exhibited by a  
289 species. Indeed, the degree to which there is caste-specific functional enrichment  
290 appears to depend on the level of social complexity [49–51]. In species with simple  
291 societies, there is little or no functional enrichment of molecular processes between  
292 castes [22,23]; e.g. as few as 6 significantly enriched Gene Ontology (GO) functional  
293 groups were found between castes in the wasp *P. canadensis* [23]. By contrast,  
294 significant levels of functional enrichment are found between castes in species with

295 more complex societies [27], and especially those with developmentally determined  
296 castes, such as the honeybee *A. mellifera*, where caste-biased genes show  
297 significant enrichment for 235 GO functional groups [52].

298

299 Enrichment of specific functional groups, or pathways, could also be a useful  
300 indicator of caste. Across all levels of social complexity, worker-biased genes show  
301 some level of enrichment for metabolic processes [22,24,53,52]; along with  
302 enrichment of cytoskeletal genes (e.g. *actin* and *myosin*) [22,27,52], these patterns  
303 may reflect the augmented energy expenditure associated with worker tasks, rather  
304 than queen (reproductive) tasks. Stress-response/immunological genes and metal-  
305 ion processing genes [24,52,54] which are associated with off-nest behaviours, also  
306 present a possible signature of worker-biased gene pathways. Genes relating to  
307 transcription and translation are up-regulated in queens of eusocial species with  
308 complex societies [52,55]. These pathways are also targets for the epigenetic  
309 regulators discussed above, making a compelling case for considering them as a  
310 potential caste-specific functional group. However queens, at all levels of social  
311 organisation, also express high levels of metabolic genes [52,56], associated with the  
312 energetic costs of reproduction. This overlap in functional enrichment between  
313 phenotypes, particularly in eusocial species with simple societies where functional  
314 speciality is limited [23], makes caste-specific gene pathways difficult to interpret and  
315 highlights the importance of utilising a range of molecular signatures (together with  
316 physiology and behaviour) to interpret caste.

317

## 318 **Summary and Conclusions**

319 The use of the term caste in eusocial insect studies requires further clarification due  
320 to the diverse spectrum of behavioural, morphological and molecular attributes.

321 Based on behavioural and/or morphological data, current literature uses the term  
322 caste as a catch-all to categorise individuals as 'reproductive' or 'non-reproductive',  
323 across the broad spectrum of complexity shown by insect societies. In simple  
324 societies, gene transcription differs very little between queens and workers [22,23]  
325 with the majority of differentially expressed genes being worker-biased [22,26,56],  
326 reflecting the metabolic and immunological costs associated with foraging. At this  
327 stage, workers are fully able to become queens, but 'switch-off' (or do not activate)  
328 their reproductive capabilities, whilst increasing their behavioural repertoire.

329 Epigenetic mechanisms do not appear to be involved at this phase [22,23], though  
330 few studies have looked into this in simple societies.

331  
332 In species where queens and workers are determined developmentally, differential  
333 gene expression between castes is significant [27,30], and there appears to be a  
334 general pattern of an over-abundance of queen-expressed genes [27,30]. Worker-  
335 biased genes continue to reflect the behavioural range of both on and off-nest  
336 activities [52], while queen-biased genes relate to transcription and translation,  
337 reflecting the cellular processes involved in reproduction [30,52,55]. Epigenetic  
338 processes are also important in complex societies; caste-biased genes are  
339 associated with miRNAs and methylation, in particular genes involved in cell  
340 development and differentiation [36,57]. Histone modifications could be the most  
341 important molecular signature of caste; they define queen and worker development  
342 and present caste-specific patterns of expression [45]. However, further work is  
343 required on simple societies and the role histone modifications play at this stage of  
344 eusociality in order to establish whether these are patterns that persist across levels  
345 of social complexity and in different independent lineages of social evolution.

346  
347 The next challenge is to obtain quantitative comparisons of molecular signatures in  
348 species representing the different stages of eusocial complexity, at different stages  
349 across colony cycles and/or experiencing different ecological conditions in order to  
350 fine-tune a holistic set of molecular signatures of caste differentiation that can be  
351 combined with the classical phenotypic traits of behaviour, physiology and  
352 morphology. A holistic approach such as this would also provide an objective way to  
353 determine whether caste is an appropriate term to describe division of labour in the  
354 simplest insect societies of eusocial insects, or whether they are in fact better  
355 described as cooperative breeders. Future work, using comparable methods of  
356 transcriptional profiling (see Kennedy *et al.* [58] on how to overcome methodological  
357 issues of transcriptomic data) alongside classic behavioural studies, may provide  
358 insights into defining the stage of eusocial evolution at which a dichotomy of 'caste',  
359 as opposed to a spectrum of gradual phenotypic variation, is appropriate.

360

361

### 362 **Acknowledgements**

363 We thank M. Bentley for helpful discussions on an earlier version of this manuscript  
364 and two anonymous reviewers for their comments. DT & SS were funded by NERC  
365 grant NE/M012913/2 and EB was funded by a NERC Studentship.

366

367

368

369 **Figure Legend**

370

371 **Figure 1:** Integrating hypotheses on putative molecular signatures for caste  
372 with classical phenotypic traits. Caste traits vary across the spectrum of  
373 eusociality, from facultative/simple eusocial behaviour to obligate/complex  
374 eusocial behaviour. Classical traits for defining Queen (Q) and Worker (W)  
375 castes include Physiology, Morphology and Behaviour. Molecular traits that  
376 may be used additionally to define castes include Gene expression, Gene  
377 Regulation and Functionality. Both morphological differentiation and  
378 behavioural specialisation increase with increasing levels of social complexity,  
379 from facultative and simple societies to obligate and complex societies.  
380 Ovarian physiological traits vary between queens and workers: at all levels of  
381 social complexity queens have fully mature ovaries and are mated (have a full  
382 spermatheca). Workers typically have undeveloped ovaries/are unmated.  
383 However, reproductive physiology in workers can change over time: e.g.  
384 workers in facultative eusocial species can develop their ovaries and mate at  
385 any time. Workers in many obligate eusocial colonies are able to develop their  
386 ovaries but are unable to mate; e.g. in queen-right colonies (QR) of  
387 bumblebees (*Bombus spp*) workers have undeveloped ovaries but in queen-  
388 less colonies (QL), or after a certain stage of colony development, they may  
389 activate their ovaries but remain unmated. In the most complex societies of  
390 obligate eusocial species workers are sterile; e.g. workers in higher Attine  
391 ants lack a spermatheca, and have very regressed ovaries. Molecular  
392 signatures may provide additional traits that help refine the categorisation of  
393 caste roles. To date, there is support for the following hypotheses: **(1) Gene**  
394 **expression:** The degree to which queens and workers differ in gene  
395 expression (measured as the proportion of detected genes that are caste-  
396 biased) appears to increase with the level of social complexity. Queens in  
397 simple eusocial species, like *Polistes*, show a general down-regulation in  
398 gene expression (red arrow) relative to W (green arrow); queens and  
399 reproductive workers in *Bombus* differ only subtly in gene expression (green  
400 arrows, up-regulation of genes), whilst non-reproductive workers are distinct

401 from both (red arrow, down-regulation). Queens in complex societies exhibit  
402 more caste-biased genes (green arrow, up-regulation), while there is little  
403 difference in gene expression across worker castes (blue horizontal arrow).  
404 **(2) Epigenetic gene regulation.** Caste-biased genes may be regulated by  
405 epigenetic processes in the more complex eusocial species, with potentially  
406 higher methylation levels in worker-biased genes, but higher levels of histone  
407 modifications in queens (not shown). In contrast, simple societies appear to  
408 lack caste-specific methylation patterns. **(3) Functional enrichment.** The  
409 degree of functional specialisation of caste-biased genes appears to increase  
410 with social complexity: castes in simple eusocial species show little functional  
411 specialisation whilst in complex eusocial species queens are functionally  
412 distinct and workers lack queen-biased gene pathways. Photo credits:  
413 *Polistes canadensis* (© Emily Bell); *Bombus griseocollis* (WikiMedia Creative  
414 Commons; Source, USGS Bee Inventory and Monitoring Lab); *Atta*  
415 *cephalotes* (WikiMedia Creative Commons; Author, Sarefo). Ovarian  
416 physiology pictures adapted from Mateus [65].

417  
418  
419  
420  
421  
422  
423  
424  
425

425 **References**

- 426  
427 1. Wilson EO: *The Insect Societies*. Harvard University Press; 1971.
- 428 2. Oster GF, Wilson EO: *Caste and Ecology in the Social Insects*. Princeton  
429 University Press; 1978.
- 430 3. Wilson EO: *Sociobiology: The New Synthesis*. Harvard University Press;  
431 1975.
- 432 4. Boomsma JJ: **Lifetime monogamy and the evolution of eusociality**. *Phil*  
433 *Trans Royal Soc B* 2009, **364**:3191–3207.
- 434 5. Bourke AFG: *Principles of Social Evolution*. Oxford University Press; 2011.
- 435 6. Bell E, Sumner S: **Ecology and social organisation of wasps**. *eLS* 2013,  
436 doi:10.1002/9780470015902.a0023597.
- 437 7. Boomsma JJ, Gawne R: **Superorganismality and caste differentiation as**  
438 **points of no return: How the major evolutionary transitions were lost in**  
439 **translation**. *Biol Rev* 2017, doi:10.1111/brv.12330.
- 440 \* **This comprehensive paper puts forward the argument that the term**  
441 **eusociality is used too ambiguously throughout biology. The authors**  
442 **posit that social behaviour should instead be categorised into two**  
443 **distinct groups: cooperative breeders and superorganisms, with the**  
444 **latter representing the ‘point of no return’ to morphologically distinct**  
445 **castes.**
- 446 8. Glastad KM, Gokhale K, Liebig J, Goodisman MAD: **The caste- and sex-**  
447 **specific DNA methylome of the termite *Zootermopsis nevadensis***. *Sci*  
448 *Rep* 2016, **6**:37110.
- 449 9. Crespi BJ, Yanega D: **The definition of eusociality**. *Behav Ecol* 1995,  
450 **6**:109–115.
- 451 10. Wilson EO: **The principles of caste evolution**. In *Experimental Behavioral*  
452 *Ecology and Sociobiology*. Edited by Holldobler B, Lindauer M. Gustav Fisher  
453 Verlag; 1985:307–324.
- 454 11. Darwin C: *On the Origin of Species by Means of Natural Selection, Or, the*  
455 *Preservation of Favoured Races in the Struggle for Life*. John Murray; 1859.
- 456 12. Wheeler WM: **The ant-colony as an organism**. *J Morphol* 1911, **22**:307–  
457 325.
- 458 13. Dew R, Tierney S, Schwarz M: **Social evolution and casteless societies:**  
459 **Need for new terminology and a new evolutionary focus**. *Insectes Soc*  
460 2016, **63**:5–14.

- 461 \*     **The authors outline the issues associated with the current approaches**  
462     **to classifying eusociality, and suggest new terminology to take into**  
463     **account the variability in social organisation, including the use of**  
464     **'casteless' societies for those that lack morphological differentiation.**
- 465 14.     Wheeler WM: *The social Insects: Their Origin and Evolution*. Kegan Paul and  
466     Co. Ltd. London; 1928.
- 467 15.     Costa J, Fitzgerald T: **Developments in social terminology: Semantic**  
468     **battles in a conceptual war.** *Trends Ecol Evol* 1996, **11**:285–289.
- 469 16.     Rehan S, Toth AL: **Climbing the social ladder: The molecular evolution of**  
470     **sociality.** *Trends Ecol Evol* 2015, **30(7)**:426-433.
- 471 \*     **This review provides a framework for investigating the molecular**  
472     **evolution of eusociality, with specific reference to relevant taxa for such**  
473     **studies, and an overview of current hypotheses.**
- 474 17.     Gadagkar R: **The evolution of caste polymorphism in social insects:**  
475     **Genetic release followed by diversifying evolution.** *J Genet* 1997,  
476     **76**:167–179.
- 477 18.     Seppa P, Queller DC, Strassmann JE: **Reproduction in foundress**  
478     **associations of the social wasp, *Polistes carolina*: Conventions,**  
479     **competition, and skew.** *Behav Ecol* 2002, **13**:531–542.
- 480 19.     Schurch R, Accleton C, Field J: **Consequences of a warming climate for**  
481     **social organisation in sweat bees.** *Behav Ecol Sociobiol* 2016, **70**:1131–  
482     1139.
- 483 20.     Michener CD, Denis J: **Were workers of eusocial Hymenoptera initially**  
484     **altruistic or oppressed?** *Proc Natl Acad Sci USA* 1974, **71**:671–674.
- 485 21.     O'Donnell S: **Reproductive caste determination in eusocial wasps**  
486     **(Hymenoptera:Vespidae).** *Annu Rev Entomol* 1998, **43**:323–346.
- 487 22.     Standage DS, Berens AJ, Glastad KM, Severin AJ, Brendel VP, Toth AL:  
488     **Genome, transcriptome and methylome sequencing of a primitively**  
489     **eusocial wasp reveal a greatly reduced DNA methylation system in a**  
490     **social insect.** *Mol Ecol* 2016, **25**:1769–1784.
- 491 23.     Patalano S, Vlasova A, Wyatt C, Ewels P, Camara F, Ferreira PG, Asher CL,  
492     Jurkowski TP, Segonds-Pichon A, Bachman M, et al: **Molecular signatures**  
493     **of plastic phenotypes in two eusocial insect species with simple**  
494     **societies.** *Proc Natl Acad Sci* 2015, **112**:13970–13975.
- 495 24.     Rehan SM, Berens AJ, Toth AL: **At the brink of eusociality:**  
496     **Transcriptomic correlates of worker behaviour in a small carpenter bee.**  
497     *BMC Evol Biol* 2014, **14**:260.

- 498 25. Berens AJ, Hunt JH, Toth AL: **Nourishment level affects caste-related**  
499 **gene expression in *Polistes* wasps.** *BMC Genomics* 2015, **16**:235
- 500 26. Ferreira PG, Patalano S, Chauhan R, Ffrench-Constant R, Gabaldón T,  
501 Guigó R, Sumner S: **Transcriptome analyses of primitively eusocial**  
502 **wasps reveal novel insights into the evolution of sociality and the origin**  
503 **of alternative phenotypes.** *Genome Biol* 2013, **14**:R20.
- 504 27. Harrison MC, Hammond RL, Mallon EB: **Reproductive workers show**  
505 **queenlike gene expression in an intermediately eusocial insect, the buff-**  
506 **tailed bumblebee *Bombus terrestris*.** *Mol Ecol* 2015, **24**:3043–3063.
- 507 28. Chen X, Hu Y, Zheng H, Cao L, Niu D, Yu D, Sun Y, Hu S, Hu F:  
508 **Transcriptome comparison between honeybee queen- and worker-**  
509 **destined larvae.** *Insect Biochem Mol Biol* 2012, **42**:665–673.
- 510 29. Ashby R, Forêt S, Searle I, Maleszka R: **MicroRNAs in honeybee caste**  
511 **determination.** *Sci Rep* 2016, **6**:18794.
- 512 \*\* **By analysing miRNAs, gene expression and DNA methylation together,**  
513 **the authors identified novel genetic pathways involved in caste**  
514 **determination in honeybees. This holistic approach is crucial to further**  
515 **our understanding of caste development.**
- 516 30. Feldmeyer B, Elsner D, Foitzik S: **Gene expression patterns associated**  
517 **with caste and reproductive status in ants: Worker-specific genes are**  
518 **more derived than queen-specific ones.** *Mol Ecol* 2014, **23**:151–161.
- 519 31. Reik W, Dean W, Walter J: **Epigenetic reprogramming in mammalian**  
520 **development.** *Science* 2001, **293**:1089–1094.
- 521 32. Hemberger M, Dean W, Reik W: **Epigenetic dynamics of stem cells and**  
522 **cell lineage commitment: Digging Waddington's canal.** *Nat Rev Mol Cell*  
523 *Biol* 2009, **10**:526–37.
- 524 33. Waddington CH: **Canalization of development and the inheritance of**  
525 **acquired characters.** *Nature* 1942, **150**:563–565.
- 526 34. Yan H, Simola DF, Bonasio R, Liebig J, Berger SL, Reinberg D: **Eusocial**  
527 **insects as emerging models for behavioural epigenetics.** *Nature Rev*  
528 *Genet* 2014, **15**:677–688.
- 529 35. Patalano S, Hore TA, Reik W, Sumner S: **Shifting behaviour: Epigenetic**  
530 **reprogramming in eusocial insects.** *Curr Opin Cell Biol* 2012, **24**:367–373.
- 531 36. Collins DH, Mohorianu I, Beckers M, Moulton V, Dalmay T, Bourke AFG:  
532 **MicroRNAs associated with caste determination and differentiation in a**  
533 **primitively eusocial insect.** *Sci Rep* 2017, **7**:45674.



- 534 37. Yan H, Bonasio R, Simola DF, Liebig J, Berger SL, Reinberg D: **DNA**  
535 **methylation in social insects: How epigenetics can control behavior and**  
536 **longevity**. *Annu Rev Entomol* 2015, **60**:435–452.
- 537 38. Li-Byarlay H: **The function of DNA methylation marks in social insects**.  
538 *Front Ecol Evol* 2016, **4**:1–8.
- 539 39. Bewick AJ, Vogel KJ, Moore AJ, Schmitz RJ: **Evolution of DNA methylation**  
540 **across insects**. *Mol Biol Evol* 2016, **34(3)**:654–665.
- 541 \*\* **This large phylogenetic investigation of DNA methylation identifies the**  
542 **variation in methylation patterns across insect orders. The authors**  
543 **found no correlation between methylation and social evolution**.
- 544 40. Lyko F, Foret S, Kucharski R, Wolf S, Falckenhayn C, Maleszka R: **The**  
545 **honeybee epigenomes: Differential methylation of brain DNA in queens**  
546 **and workers**. *PLoS Biol* 2010, **8**.
- 547 41. Kapheim KM, Pan H, Li C, Salzberg SL, Puiu D, Magoc T, Robertson HM,  
548 Hudson ME, Venkat A, Fischman BJ: **Genomic signatures of evolutionary**  
549 **transitions from solitary to group living**. *Science* 2015, **348**:1139–1143.
- 550 42. Glastad KM, Arsenault SV, Vertacnik KL, Geib SM, Kay S, Danforth BN,  
551 Rehan SM, Linnen CR, Kocher SD, Hunt BG: **Variation in DNA methylation**  
552 **is not consistently reflected by sociality in Hymenoptera**. *Genome Biol*  
553 *Evol* 2017, **9**:1687–1698.
- 554 43. Simola DF, Graham RJ, Brady CM, Enzmann BL, Desplan C, Ray A, Zwiebel  
555 LJ, Bonasio R, Reinberg D, Liebig J, et al.: **Epigenetic (re)programming of**  
556 **caste-specific behavior in the ant *Camponotus floridanus***. *Science* 2016,  
557 **351(6268)** DOI: 10.1126/science.aac6633.
- 558 \*\* **The authors used RNAi and HDACi to modify caste-specific behaviours**  
559 **of worker ants in a comprehensive experiment identifying how single-**  
560 **gene modifications can enable us to understand epigenetic regulators**  
561 **of caste-determination**.
- 562 44. Galbraith DA, Yi S V, Grozinger CM: **Evaluation of possible proximate**  
563 **mechanisms underlying the kinship theory of intragenomic conflict in**  
564 **social insects**. *Integr Comp Biol* 2016, **56**:1206–1214.
- 565 45. Simola DF, Ye C, Mutti NS, Dolezal K, Bonasio R, Liebig J, Reinberg D,  
566 Berger SL: **A chromatin link to caste identity in the carpenter ant**  
567 ***Camponotus floridanus***. *Genome Res* 2013, **23**:486–496.
- 568 46. Spannhoff A, Kim YK, Raynal NJ-M, Gharibyan V, Su M-B, Zhou Y-Y, Li J,  
569 Castellano S, Sbardella G, Issa J-PJ, et al.: **Histone deacetylase inhibitor**

- 570 activity in royal jelly might facilitate caste switching in bees. *EMBO Rep*  
571 2011, **12**:238–243.
- 572 47 Dickman MJ, Kucharski R, Maleszka R, Hurd PJ: **Extensive histone post-**  
573 **translational modification in honeybees.** *Insect Biochem Mol Biol* 2013,  
574 **43**:125–137.
- 575 48. Gowher H, Petell CJ: **An epigenetic switch regulates de novo DNA**  
576 **methylation at pluripotency gene enhancers.** *FASEB J* 2017,  
577 **31(1)**:593.15.
- 578 49. Sobotka JA, Daley M, Chandrasekaran S, Rubin BD, Thompson GJ:  
579 **Structure and function of gene regulatory networks associated with**  
580 **worker sterility in honeybees.** *Ecol Evol* 2016, **6(6)**:1692-1701.
- 581 50. Johnson BR, Jasper WC: **Complex patterns of differential expression in**  
582 **candidate master regulatory genes for social behavior in honeybees.**  
583 *Behav Ecol Sociobiol* 2016, **70(7)**:1033-1043
- 584 51. Friedman DA, Gordon DM: **Ant genetics: Reproductive physiology,**  
585 **worker morphology, and behavior.** *Annu Rev Neurosci* 2016, **39**:41-56.
- 586 52. Berens AJ, Hunt JH, Toth AL: **Comparative transcriptomics of convergent**  
587 **evolution: Different genes but conserved pathways underlie caste**  
588 **phenotypes across lineages of eusocial insects.** *Mol Biol Evol* 2015,  
589 **32**:690–703.
- 590 53. Toth AL, Varala K, Henshaw MT, Rodriguez-Zas SL, Hudson ME, Robinson  
591 GE: **Brain transcriptomic analysis in paper wasps identifies genes**  
592 **associated with behaviour across social insect lineages.** *Proc R Soc B*  
593 *Biol Sci* 2010, **277**:2139–2148.
- 594 54. Toth AL, Tooker JF, Radhakrishnan S, Minard R, Henshaw MT, Grozinger  
595 CM: **Shared genes related to aggression, rather than chemical**  
596 **communication, are associated with reproductive dominance in paper**  
597 **wasps (*Polistes metricus*).** *BMC Genomics* 2014, **15**:75.
- 598 55. Jedlicka P, Ernst UR, Votavov A, Hanus R, Valtero I: **Gene expression**  
599 **dynamics in major endocrine regulatory pathways along the transition**  
600 **from solitary to social life in a bumblebee, *Bombus terrestris*.** *Front*  
601 *Physiol* 2016, **7**:1–20.
- 602 56. Jones BM, Kingwell CJ, Wcislo WT, Robinson GE, Jones BM: **Caste-biased**  
603 **gene expression in a facultatively eusocial bee suggests a role for**  
604 **genetic accommodation in the evolution of eusociality.** *Proc R Soc B Biol*  
605 *Sci* 2017, DOI: 10.1098/rspb.2016.2228.

- 606 57. Bonasio R, Li Q, Lian J, Mutti NS, Jin L, Zhao H, Zhang P, Wen P, Xiang H,  
607 Ding Y, et al.: **Genome-wide and caste-specific DNA methylomes of the**  
608 **ants *Camponotus floridanus* and *Harpegnathos saltator***. *Curr Biol* 2012,  
609 **22**:1755–1764.
- 610 58. Kennedy P, Baron G, Qiu B, Freitak D, Helanterä H, Hunt ER, Manfredini F,  
611 O’Shea-Wheller T, Patalano S, Pull CD, Sasaki T, Taylor D, Wyatt C, Sumner  
612 S: **Deconstructing superorganisms and societies to address big**  
613 **questions in biology**. *Trends Ecol Evol* 2017, **32**:861–872.
- 614 59. Colgan TJ, Carolan JC, Bridgett SJ, Sumner S, Blaxter ML, Brown MJ:  
615 **Polyphenism in social insects: Insights from a transcriptome-wide**  
616 **analysis of gene expression in the life stages of the key pollinator,**  
617 ***Bombus terrestris***. *BMC Genomics* 2011, **12**:623.
- 618 60. Cardoso-Júnior CAM, Fujimura PT, Santos-Júnior CD, Borges NA, Ueira-  
619 Vieira<sup>1</sup>, Hartfelder C, Bonetti K, Goulart LR, Maria A: **Epigenetic**  
620 **modifications and their relation to caste and sex determination and**  
621 **adult division of labour in the stingless bee *Melipona scutellaris***. *Genet*  
622 *Mol Biol* 2017, **40(1)**:61-68
- 623 61. Grozinger CM, Fan Y, Hoover SE, Winston ML: **Genome-wide analysis**  
624 **reveals differences in brain gene expression patterns associated with**  
625 **caste and reproductive status in honeybees (*Apis mellifera*)**. *Mol Ecol*  
626 2007, **16**:4837–4848.
- 627 62. Foret S, Kucharski R, Pittelkow Y, Lockett GA, Maleszka R: **Epigenetic**  
628 **regulation of the honeybee transcriptome: Unravelling the nature of**  
629 **methylated genes**. *BMC Genomics* 2009, **10**:472.
- 630 63. Barchuk AR, Cristino AS, Kucharski R, Costa LF, Simões ZL, Maleszka R:  
631 **Molecular determinants of caste differentiation in the highly eusocial**  
632 **honeybee *Apis mellifera***. *BMC Dev Biol* 2007, **7**:70.
- 633 64. Foret S, Kucharski R, Pellegrini M, Feng S, Jacobsen SE, Robinson GE,  
634 Maleszka R: **DNA methylation dynamics, metabolic fluxes, gene splicing,**  
635 **and alternative phenotypes in honeybees**. *Proc Natl Acad Sci* 2012,  
636 **109**:4968–4973.
- 637 65. Mateus S, Noll FB, Zucchi R: **Caste flexibility and variation according to**  
638 **the colony cycle in the swarm-founding wasp, *Parachartagus fraternus***  
639 **(Gribodo) (Hymenoptera: Vespidae: Epiponini)**. *J Kans Entomol Soc*  
640 2004, **77(4)**:470-483.
- 641  
642

643  
 644 Table 1: Caste and the subsequent division of reproductive labour is a key aspect of  
 645 the definition of 'eusociality'; consequently, a clear and consistent definition of 'caste'  
 646 is essential for the study and categorisation of eusociality. This table provides (a non-  
 647 exhaustive list of) examples of how the definition of the term 'caste' has changed  
 648 over time. Definitions are expanded to include 'caste traits' that comment on whether  
 649 castes differ in function ('functionally distinct'), how permanent caste is  
 650 (irreversible/permanent), and whether their 'caste' is behavioural (B) or morphological  
 651 (M) or both.

Reference	Caste traits			Definition
	Functionally distinct	Irreversible / permanent	Behavioural (B) or morphological (M)	
Michener, 1974 [20]	Y	Y	B & M	Castes are " <b>functionally different</b> groups among the females of a colony. The differences may be only behavioural or physiological, or may also involve structure. The differences are <b>permanent</b> and not due to age".
Wilson, 1975 [3]	Y	N	M	Caste are defined as a "set of individuals of a particular morphological type, or age group or both, that performs a specialized labor in the colony. More narrowly defined, any set of individuals in a given colony that are <b>both morphologically distinct</b> from other individuals and <b>specialized in behavior</b> "
Wilson, 1985 [10]	Y	N	B or M	"a set of colony members...that <b>specialized</b> on particular tasks for prolonged periods of time...distinguished by some kind of additional marker – a larger size, some other kind of <b>anatomical feature</b> , a <b>different age</b> , or even some less apparent <b>physiological trait...</b> ".
Crespi & Yanega, 1995 [9]	N	Y	B	Groups of individuals that become " <b>irreversibly behaviorally distinct</b> at some point prior to reproductive maturity".
O'Donnell, 1998 [21]	Y	Y	M	Castes refer to "female reproductive function, with the requirement that caste membership be <b>developmentally determined and irreversible</b> ".
Boomsma, 2009 [4]	Y	Y	B & M	A "caste is <b>irreversibly determined</b> early in development (before pupation in the Hymenoptera), and to such extent that no individuals or predestined worker cohorts retain the behavioural, and often also physiological, option to disperse and found their own colonies".
Boomsma & Gwane 2017 [7]	Y	Y	M	Lineages of social insects that have not passed the ' <b>point of no return</b> ' to <b>physically distinct</b> castes could be better considered as co-operative breeders.

652  
 653  
 654

655  
656  
657 Table 2: Summary of supporting data for putative molecular signatures of caste,  
658 across levels of eusociality from facultatively eusocial species to complex eusocial  
659 species; signatures are defined in terms of differential gene expression (DE),  
660 regulatory processes (microRNAs (miRNAs), methylation and histone acetylation),  
661 and functional differentiation. Data shown is from the following classifications and  
662 species: Facultatively eusocial species: *Ceratina calcarata* [24] & *Megalopta genalis*  
663 [56], Simple eusocial species: *Polistes canadensis* [23], *Polistes dominula* [22],  
664 *Polistes metricus* [52] and *Dinoponera quadriceps* [23]. Obligate eusocial species  
665 with reproductive workers: *Bombus terrestris* [27,36,55,59], *Melipona scutellaris* [60]  
666 and *Harpegnathus saltator* [57], Obligate eusocial species with sterile workers:  
667 *Camponotus floridanus* [45,57], *Temnothorax longispinosus* [30], *Apis mellifera*  
668 [29,46,47,52,61–64] and *Zootermopsis nevadensis* [8].

669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690

691

692

	<b>Facultatively eusocial/ simple eusocial</b>	<b>Obligate eusocial (reproductive workers)</b>	<b>Obligate eusocial(sterile, specialised workers)</b>
<b>Gene expression</b>	Few genes are DE between queens and workers [22–24]  Most caste biased genes are upregulated in workers [22,26,52,56]	High levels of genes DE between Q & W [27]  Increase in numbers of queen-biased genes compared to workers [27]	High levels of genes DE between Q & W [30,61]  Most caste biased genes are upregulated in queens [30,61]  Low levels of genes DE within worker castes [30]
<b>Regulatory processes: MicroRNAs, methylation, histone acetylation</b>	MiRNAs not targeting caste-biased DEGs [23]  Caste-biased genes are not associated with differential methylation [22,23]	MiRNAs targeting caste-biased genes in larvae [36]  Caste-biased genes associated with differential methylation [57]  Queen-biased genes show lower levels of methylation than worker genes [60]	MiRNAs targeting caste-biased genes in larvae [29]  Caste-biased genes associated with differential methylation [8,57,62]  Queen-biased genes show lower levels of methylation than worker genes [64]  Increased histone activity in Q compared to W [45–47]
<b>Functional specialisation</b>	Functional enrichment of worker-biased genes		
	Metabolic processes, cytoskeleton, stress response [22,24,56]	Metabolic processes, cytoskeleton, stress response [27,59]	Metabolic processes, cytoskeleton, phagocytosis/stress response [30,52,61]
	Functional enrichment of queen-biased genes		
	Glycolysis, metabolic processes, cell development, chromatin organisation, translation [24,52,56]	Glycolysis, metabolic processes, cell development, transcription [27,55]	Glycolysis, metabolic processes, cell development, odorant binding, protein modification, transcription/translation [30,52,63]

693

694

695

696