A molecular concept of caste in insect societies

Seirian Sumner¹
Emily Bell²
Daisy Taylor²

¹Centre for Biodiversity and Environmental Research, Medawar Building, University College London, Gower Street, London WC1E 6BT
²School of Biological Sciences, Bristol Life Sciences Building, University of Bristol, 24 Tyndall Avenue, Bristol BS8 1TQ, UK

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

Highlights

• The categorisation of social insect individuals into two distinct castes - reproductive and non-reproductive forms - may be too simplistic: ecology, life-history and level of social complexity generate variation within and among caste

• The molecular technological advances of the 21st century provide us with an additional empirical base-line for defining and understanding the caste system in eusocial insects.

• Molecular signatures, combined with behavioural, physiological and morphological traits, may provide a quantitative and objective method for defining and categorising caste.
Abstract

The term ‘caste’ is used to describe the division of reproductive labour that defines eusocial insect societies. The definition of ‘caste’ has been debated over the last 50 years, specifically with respect to the simplest insect societies; this raises the question of whether a simple categorisation of social behaviour by reproductive state alone is helpful. Gene-level analyses of behaviours of individuals in hymenopteran social insect societies now provide a new empirical base-line for defining caste and understanding the evolution and maintenance of a reproductive division of labour. We review this literature to identify a set of potential molecular signatures that, combined with behavioural, morphological and physiological data, help define caste more precisely; these signatures vary with the type of society, and are likely to be influenced by ecology, life-history, and stage in the colony cycle. We conclude that genomic approaches provide us with additional ways to help quantify and categorise caste, and behaviour in general.

Keywords: social evolution, eusocial insects, genomics, phenotypic plasticity
**Introduction**

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

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

**Defining caste and why it matters**

The three key traits of a eusocial species are that it exhibits overlapping generations, cooperative brood care, and division of labour [9,10]. ‘Caste’ is the mechanism by which one of these traits - division of labour - is achieved, and it provides phenotypic
variation on which selection can act. The term 'caste' was historically used to
describe the fixed, morphologically distinct queens and workers of the complex
eusocial species, like the honeybee, where workers are committed to a lifetime of
functional sterility (e.g. [11,12]). 'Caste' has since become a more ubiquitous term,
used to describe division of labour in all eusocial societies, including those that do
not have lifetime, functionally sterile phenotypes. For example, in the space of a
decade, E.O. Wilson's use of the term 'caste' changed from one that required
'morphologically distinct' traits [3] to a much looser definition whereby individuals are
'specialised... for prolonged periods of time'... and... 'some kind of additional
marker'... which may change with age, and may not involve external morphology
[10]. This latter definition includes facultatively eusocial species, where generations
are overlapping, there is cooperative brood care, but division of labour (and thus
castes roles) can be temporary and distinguished only by behaviour, (e.g.
Stenogastrinae wasps; Halictid bees). Currently, the term ‘caste’ is used as a catch-
all term to describe the full phenotypic diversity of division of labour in eusocial
insects. However, the applicability of one term (i.e. caste) to describe societal
coordination across all species has been questioned and much of the terminology
surrounding castes remains misleading [1,9,13–15].

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

Ambiguity in the use of the term caste is a long-standing problem (Table 1); we lack a
consistent set of quantifiable traits for defining and categorising caste. Almost all
definitions agree that ‘caste’ describes functionally distinct phenotypes: indeed,
this is necessary for a division of labour. Definitions differ on whether castes are
irreversible, such that individuals show lifetime commitment to a specific role, or flexible, whether a morphological distinction is important, or whether behavioural distinction alone is sufficient. Complications arise for species which traverse these criteria; for example, some swarm-founding Polistine wasps (e.g. *Polybia occidentalis*) where caste is apparently irreversible but individuals lack morphological differentiation. Such confusion can be resolved if species are instead defined as those in which caste is fixed during development (this includes all species with irreversible castes, but does not require there to be morphological differences), or whether caste remains plastic throughout life (these are species whose functional roles are ‘reversible’, and that lack morphological differences). However, without detailed knowledge of the plasticity of castes in a species, this definition is of limited practical use. Moreover, does this mean that species without developmentally determined roles lack ‘real castes’ [13]? We require a set of traits that are readily quantifiable in order to accurately describe castes, and use this to categorise a species in relation to its level in the evolution of social complexity.

The heart of the problem in defining castes is that, like most biological systems, caste is a complex phenomenon and boundaries between castes are often messy.

Expression of caste traits (as defined in Table 1) varies considerably. Castes can be so different in appearance that queens and workers can appear to belong to different species (e.g. Attine ants; Figure 1). In contrast, other castes are discernible only from behavioural observations and ovary dissections (e.g. *Polistes* wasps; Figure 1). The simpler eusocial societies present the main issue here: they are functionally distinct, but lack lifetime commitment, morphological distinction and developmental differentiation. As a result, some authors have suggested that the simplest eusocial insect societies are better described as ‘casteless’ and as cooperative breeders [4,7,12]. Further, traits used to define castes may vary within the lifetime of a colony (e.g. *Polistes* foudndresses can behave like queens until the first workers emerge, and thereafter behave like workers [18]), and within the lifetime of an individual (e.g. *Bombus* workers can develop ovaries and become unmated reproductives late in the colony’s life). Finally, ecological conditions can influence expression of caste traits within and across species: for example, ‘worker’ traits are only expressed by Halictid bees living in warm climatic regions, with long summers [19].

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

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

The 21st century explosion in molecular techniques allows us to scrutinise the concept of caste at the genomic level. We propose data-driven hypotheses for how genomic analyses of phenotypes may contribute towards a more quantitative definition of castes in eusocial insects (summarised in Table 2). These hypotheses fall into three main facets of genomic variation: gene expression patterns, regulatory processes and functionality.

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

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

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

Regulatory mechanisms as molecular signatures of caste

Epigenetic mechanisms control gene expression by differentially regulating genes in
response to environmental or genetic cues [31]; they can also limit the plasticity of
gene expression, fixing specific transcriptional patterns irreversibly [31–33].
Epigenetic canalization is therefore a key hypothesis in the regulation of the inflexible
roles found in eusocial insect species where castes are determined (irreversibly)
during development [34,35]. An important question, however, is whether the same
epigenetic processes regulate caste in the simpler societies, and whether patterns of
epigenetic regulation can be useful signatures of caste evolution and sociality
[34,35].

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

DNA can be chemically modified by the addition of DNA methyltransferases (DNMTs)
which subsequently regulate gene expression; representatives from this group of
enzymes are found in all insect orders, albeit with significant variation [34,37,38].
Eusocial Hymenoptera show some of the lowest levels of methylation among insects
[39], and moreover methylation rates and patterns vary considerably across levels of
The highly plastic phenotypes of the simple societies, with putatively reversible castes, show only limited difference in brain methylation changes between castes, [22,23]. Species with developmentally-determined castes (which are irreversible) often exhibit high levels of methylation at key genes and these are associated with caste-biased genes [40] (reviewed in [38]) (Table 2). Levels of brain methylation, therefore, may be an indicator of caste differentiation and social complexity [41]. However, the data are inconclusive. At the species level, whole body analyses of methylation found little support for a correlation between methylation and sociality: for example, methylation levels in non-social insects are not consistently lower than social species, and even within the eusocial Hymenoptera there is no clear correlation between methylation levels and social complexity [39,42]. One source of variation that may account for the muddy story of the role of methylation in castes and sociality is the tissue analysed and level of analysis; i.e. caste-specific and brain tissue [22,23,40] versus species level and/or whole bodies [8,30,39].

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

**Functional specialisations as molecular signatures of caste**

General patterns of functional enrichment may be useful signatures of caste, and may be especially useful in determining the level of social complexity exhibited by a
species. Indeed, the degree to which there is caste-specific functional enrichment appears to depend on the level of social complexity [49–51]. In species with simple societies, there is little or no functional enrichment of molecular processes between castes [22,23]; e.g. as few as 6 significantly enriched Gene Ontology (GO) functional groups were found between castes in the wasp *P. canadensis* [23]. By contrast, significant levels of functional enrichment are found between castes in species with more complex societies [27], and especially those with developmentally determined castes, such as the honeybee *A. mellifera*, where caste-biased genes show significant enrichment for 235 GO functional groups [52].

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

**Summary and Conclusions**

The use of the term caste in eusocial insect studies requires further clarification due to the diverse spectrum of behavioural, morphological and molecular attributes. Based on behavioural and/or morphological data, current literature uses the term caste as a catch-all to categorise individuals as ‘reproductive’ or ‘non-reproductive’, across the broad spectrum of complexity shown by insect societies. In simple societies, gene transcription differs very little between queens and workers [22,23]
with the majority of differentially expressed genes being worker-biased [22,26,56],
reflecting the metabolic and immunological costs associated with foraging. At this
stage, workers are fully able to become queens, but ‘switch-off’ (or do not activate)
their reproductive capabilities, whilst increasing their behavioural repertoire.
Epigenetic mechanisms do not appear to be involved at this phase [22,23], though
few studies have looked into this in simple societies.
In species where queens and workers are determined developmentally, differential
gene expression between castes is significant [27,30], and there appears to be a
general pattern of an over-abundance of queen-expressed genes [27,30]. Worker-
baised genes continue to reflect the behavioural range of both on and off-nest
activities [52], while queen-biased genes relate to transcription and translation,
reflecting the cellular processes involved in reproduction [30,52,55]. Epigenetic
processes are also important in complex societies; caste-biased genes are
associated with miRNAs and methylation, in particular genes involved in cell
development and differentiation [36,57]. Histone modifications could be the most
important molecular signature of caste; they define queen and worker development
and present caste-specific patterns of expression [45]. However, further work is
required on simple societies and the role histone modifications play at this stage of
eusociality in order to establish whether these are patterns that persist across levels
of social complexity and in different independent lineages of social evolution.
The next challenge is to obtain quantitative comparisons of molecular signatures in
species representing the different stages of eusocial complexity, at different stages
across colony cycles and/or experiencing different ecological conditions in order to
fine-tune a holistic set of molecular signatures of caste differentiation that can be
combined with the classical phenotypic traits of behaviour, physiology and
morphology. A holistic approach such as this would also provide an objective way to
determine whether caste is an appropriate term to describe division of labour in the
simplest insect societies of eusocial insects, or whether they are in fact better
described as cooperative breeders. Future work, using comparable methods of
transcriptional profiling (see Kennedy et al. [58] on how to overcome methodological
issues of transcriptomic data) alongside classic behavioural studies, may provide
insights into defining the stage of eusocial evolution at which a dichotomy of ‘caste’,
as opposed to a spectrum of gradual phenotypic variation, is appropriate.
Acknowledgements
We thank M. Bentley for helpful discussions on an earlier version of this manuscript and two anonymous reviewers for their comments. DT & SS were funded by NERC grant NE/M012913/2 and EB was funded by a NERC Studentship.

Figure Legend

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

(2) **Epigenetic gene regulation.** Caste-biased genes may be regulated by
epigenetic processes in the more complex eusocial species, with potentially
higher methylation levels in worker-biased genes, but higher levels of histone
modifications in queens (not shown). In contrast, simple societies appear to
lack caste-specific methylation patterns. (3) **Functional enrichment.** The
degree of functional specialisation of caste-biased genes appears to increase
with social complexity: castes in simple eusocial species show little functional
specialisation whilst in complex eusocial species queens are functionally
distinct and workers lack queen-biased gene pathways. Photo credits:
*Polistes canadensis* (© Emily Bell); *Bombus griseocollis* (WikiMedia Creative
Commons; Source, USGS Bee Inventory and Monitoring Lab); *Atta
ccephalotes* (WikiMedia Creative Commons; Author, Sarefo). Ovarian
physiology pictures adapted from Mateus [65].


* This comprehensive paper puts forward the argument that the term eusociality is used too ambiguously throughout biology. The authors posit that social behaviour should instead be categorised into two distinct groups: cooperative breeders and superorganisms, with the latter representing the ‘point of no return’ to morphologically distinct castes.


The authors outline the issues associated with the current approaches to classifying eusociality, and suggest new terminology to take into account the variability in social organisation, including the use of ‘casteless’ societies for those that lack morphological differentiation.


This review provides a framework for investigating the molecular evolution of eusociality, with specific reference to relevant taxa for such studies, and an overview of current hypotheses.


** By analysing miRNAs, gene expression and DNA methylation together, the authors identified novel genetic pathways involved in caste determination in honeybees. This holistic approach is crucial to further our understanding of caste development.


This large phylogenetic investigation of DNA methylation identifies the variation in methylation patterns across insect orders. The authors found no correlation between methylation and social evolution.


The authors used RNAi and HDACi to modify caste-specific behaviours of worker ants in a comprehensive experiment identifying how single-gene modifications can enable us to understand epigenetic regulators of caste-determination.


