THE EVOLUTION AND MECHANISMS
OF CASTE PLASTICITY IN VESPID WASPS

A thesis submitted in fulfilment of

the requirements for the degree of

Doctor of Philosophy

Benjamin Aaron Taylor

Department of Genes, Environment & Evolution

University College London

2021

Declaration

I, Benjamin Aaron Taylor, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Benjamin Aaron Taylor

Abstract

Social insects are ecologically dominant predators, pollinators, herbivores and detritivores across many terrestrial ecosystems. Key to the ecological success of these species is a uniquely strong division of labour between reproductives ('queens') and non-reproductives ('workers'). In some social insect species, reproductive division of labour is obligate and developmentally determined, but many other taxa possess full reproductive plasticity, which is the basal state for social insect evolution. Answering the question of how division of reproductive labour is maintained in the presence of reproductive plasticity is an important prerequisite to understanding how and why this plasticity has been lost in the most derived social insect taxa.

In this thesis, I address this question using two species of social wasp which exhibit strong division of reproductive labour but full reproductive plasticity. Two chapters of the thesis examine responses to queen loss in the European paper wasp *P. dominula,* in order to understand the mechanisms by which groups accommodate the loss of a reproductive. In Chapter 2 I show that in this species, groups generate replacement reproductives rapidly and with little conflict by relying on an age-based succession criterion. In Chapter 3 I analyse the transcriptomic mechanisms that underlie this succession process, and show that variation in individuals' phenotypes only partially explains their transcriptomic responses, a result that suggests hidden costs of queen loss.

In Chapter 4, I analyse individual-level transcriptomic data from a facultatively social tropical hover wasp, *Liostenogaster flavolineata*, which forms linearly age-based dominance hierarchies in which individuals exhibit progressively reduced foraging effort as they move up in rank. I show that despite differences in social structure, variation in

gene expression in colonies of this species is surprisingly similar to that of obligately social species such as *P. dominula*. I also find that genes that are associated with indirect fitness in *L. flavolineata* are more strongly evolutionarily conserved than genes associated with direct fitness, a surprising result that runs counter to results obtained for other social insect species. Additionally, in Chapter 5 I argue for a reconceptualization of the loss of reproductive plasticity that has occurred in more complex insect societies.

Taken as a whole, this thesis sheds light on the behavioural and transcriptomic mechanisms by which distinct fitness strategies are maintained in reproductively skewed societies as well as revealing potential limitations of these mechanisms, emphasising the value of reproductively plastic social insects as models for the evolution of sociality.

Impact statement

This thesis advances our understanding of the behavioural and molecular mechanisms that facilitate the division of reproductive labour that characterises social insect societies. Social insect species have substantial anthropological value: these taxa perform key roles in virtually every terrestrial ecosystem, and that impact is felt especially keenly in agricultural ecosystems. The threats to human agriculture posed by ongoing declines in insect pollinators are now widely publicised, but emerging evidence suggests that social wasps—which are prolific predators of agricultural pests—may be experiencing equally strong declines, the effects of which upon crop yields are little known. These species are ecologically and anthropologically important precisely because their social structures permit an outstanding degree of cooperative efficiency, and thus the question of how such societies are maintained impacts directly upon human social, ecological and financial concerns.

Beyond this specific intersection with human interests, however, the topics addressed in this manuscript reflect one of biology's central themes. Fundamental to this work is a ubiquitous question: why and how do individuals engage in cooperation? The coming together of distinct subunits to form new levels of evolutionary individuality has been one of the most important drivers of life's diversity, and yet our understanding of this process remains incomplete. The paradoxical nature of evolution, a conflict-driven process that nonetheless repeatedly produces cooperators, has been and remains a source of fascination for me. It would be my greatest pleasure to believe that the work I present here brings us an increment closer to understanding that process.

Acknowledgements

Table of contents

List of figures

List of tables

List of abbreviations

AUC – Area under curve

BLAST – Basic local alignment search tool

DNA – Deoxyribonucleic acid

GO – Gene ontology

RNA – Ribonucleic acid

ROC – Receiver operating characteristic

SVM – Support vector machine

WGCNA – Weighted gene co-expression network analysis

Chapter 1 – Introduction

1.1 Social evolution

Evolution by natural selection is a fundamentally and inescapably competitive process, yet the story of life on Earth is also a story of cooperation. Chromosomes are societies of genes, multicellular organisms are societies of cells, and social insect groups are societies of multicellular individuals. Despite the fundamental role that the formation of cooperative groups has played in evolution, our understanding of the reasons that individuals do or do not form such groups is limited, as is our knowledge of the mechanisms by which cooperative groups maintain their cohesion once formed. This is especially true for those cooperative groups in which reproduction is monopolised by a minority of individuals, because conflicts of interest between reproducers and non-reproducers should undermine the unity of such groups. In this thesis, I use social insect societies as a model with which to explore the question of how cooperative groups maintain social cohesion in the face of high potential within-group conflict.

1.1.1 Fitness strategies and their payoffs

Fitness is a measure of the genetic contribution that an individual makes towards the next generation of the population to which it belongs (Hamilton 1964a; Barker 2009). The simplest way in which an individual can increase its fitness is by producing its own offspring; fitness gained this way is referred to as *direct fitness*. Direct fitness is only one component of an individual's potential lifetime fitness, however. Because an individual organism by definition shares a larger proportion of genes with its relatives than with the 'average' individual in the population, organisms can also increase the share of their own genes in the next population by increasing the direct fitness of their relatives. For example, an animal might gain fitness by helping raise the offspring of its parents or siblings rather than producing and raising its own offspring. Fitness that an individual gains by helping this way is known as *indirect fitness*, and the sum of that individual's direct and indirect fitness is its *inclusive fitness* (Hamilton 1964a).

Organisms typically cannot maximise both direct and indirect fitness simultaneously: an individual that spends time and energy helping relatives is not spending that time and energy rearing its own young, so helping behaviour is only expected to evolve if the fitness payoff for helping outweighs the fitness payoff for pursuing direct fitness (Hamilton 1964a,b). The choice of whether or not to adopt an indirect fitness strategy is

16

formalised in Hamilton's rule, which states that it is evolutionarily favourable to help a relative when

$$rB - C > 0$$

where r represents the relatedness of the altruistic individual to the individual it is helping, B represents the fitness benefit experienced by the beneficiary, and C represents the fitness cost experienced by the altruist as a result of their helping behaviour (Hamilton 1964a,b; West et al. 2007). Relatedness (r) is expressed as a measure of genetic similarity relative to the population, reaching a maximum value of 1 in genetically identical individuals. The benefit to the individual that receives help (B), and the cost to the altruist (C), can be expressed in terms of number of offspring. Thus C is the number of extra offspring that the beneficiary of altruism produced as a result of the helping behaviour it received, and B is the number of offspring that the altruist could have successfully produced using the time and energy it spent on the helping behaviour. We can express Hamilton's rule more generally in order to make it applicable to any choice an individual makes between two reproductive strategies. Strategy 1 should be pursued over strategy 2 only when

$$r_1B_1 - C_1 > r_2B_2 - C_2$$

with each strategy having its own costs and benefits.

A wide variety of factors can influence the benefits and costs of different fitness strategies, and therefore an individuals' decision of whether to help relatives or to focus on direct fitness. Briefly, organisms are more likely to invest in indirect fitness when within-group relatedness is high and can be easily assured (Griffin et al. 2004; Hughes et al. 2008); when opportunities for direct fitness are constrained by a lack of access to resources or mates, or when helpers are subfertile (Carrete et al. 2006; Holman 2014a); when the benefits of helping behaviour are high, for example in species that are reliant on resources that cannot be efficiently exploited by solitary individuals (Cortes-Avizanda et al. 2011; Faulkes & Bennett 2013); and possibly also when environmental stochasticity is high (Kennedy et al. 2018; but see Cornwallis et al. 2017 for an alternative view).

In most species that exhibit some degree of social behaviour, the distinction between direct and indirect fitness strategies is not absolute. For example, in many mammal and

bird species, helping behaviour directed towards relatives is performed primarily by individuals that are either too young or too old to reproduce themselves (Nattrass et al. 2019), is expressed facultatively in response to transient ecological conditions (Holman 2014b), or is performed communally such that individuals can be considered to simultaneously perform and receive helping behaviour. As a result, individuals in such species can expect to accrue a combination of direct and indirect fitness gains throughout their lives, rather than specialising in one or another fitness strategy.

Social insects, which arguably represent the most extreme examples of social behaviour in the Metazoa, provide a contrast to the flexible fitness roles found in most social animals. Within social insect groups, individuals are divided into more or less distinct reproductive roles, with the majority of individuals acting as non-reproductive 'workers' that are fully (and in some cases irreversibly) committed to an indirect fitness strategy. The compartmentalization of fitness into direct and indirect components in social insect colonies has resulted in these taxa becoming models for social evolution, as well as foci of evolutionary and ecological interest in their own right.

1.1.2 Insect sociality

Social insect species as commonly understood are those that either obligately or facultatively form multi-member groups in which there is a significant partitioning of reproductive and non-reproductive labour between individuals. This definition excludes some lesser-known species that exhibit remarkably interesting and complex behaviours that surely fall under a broader definition of 'social', such as the wonderfully derived patterns of biparental care exhibited by many species in the burying beetle genus *Nicrophorus* (Suzuki 2000; Benowitz & Moore 2016). It nonetheless remains the case that when biologists (and indeed laymen) refer to social insects, they almost invariably have in mind the socially complex colonies found in many species of bees, wasps, ants and termites (Brian 2012). What these species ultimately have in common is a high degree of reproductive skew, with the majority of individuals forgoing direct reproduction in favour of indirect fitness gains accrued by aiding a relative. This thesis deals with the question of why and how this distinctive reproductive skew is maintained.

Beyond the one definitive trait of strong reproductive skew, social insect species also share a great number of features that are to a greater or lesser extent related to this skew, such as large colony sizes and the construction of static communal nest structures. Within this category, however, is an enormous degree of variability. In some species group sizes

rarely exceed 10, while in others they may number in the millions (Bourke 1999); in some species reproductives and non-reproductives are distinguished only behaviourally, while in others they exhibit extreme dimorphism (Gadagkar 1997). The question of how to best describe and categorise variation in insect sociality remains unsettled, with several frameworks proposed.

Most frameworks of insect sociality either explicitly or implicitly support the notion that social insect species fall into two broad categories. One category of species ('primitively eusocial', 'simple' or 'cooperatively breeding' social insects) is that in which some proportion of individuals are reliant upon indirect fitness gains from altruistic behaviour directed towards relatives, but commitment to reproductive roles is not complete. Individuals in such species retain the plasticity to switch from non-reproductive to reproductive roles. Model examples in this category include halictid bees, paper wasps and many 'lower' termites. The other category ('complex', 'highly eusocial' or 'superorganismal' social insects) consists of species in which reproductive roles are irreversibly determined during development, resulting in the majority of individuals existing as sterile or subfertile workers. Species in this category include honeybees, vespine wasps, most 'higher' termites, and the large majority of ant species.

Additionally, many frameworks make a distinction between species that are facultatively social, i.e. those in which individuals can reproduce solitarily but may also form reproductively skewed groups, and those in which reproduction only occurs as part of a social group. The tropical hover wasps that I describe in Chapter 4 of this thesis represent an example of facultative sociality, and in that chapter I use this fact to explore the molecular signatures that distinguish individuals that are invested in indirect fitness strategies from those that are invested in direct fitness.

It is possible to conceptualise variation in insect social systems as representing a form of 'social ladder' along which species can be placed depending upon factors such as colony size, strength of reproductive skew, and the presence and absence of evolutionary novelties such as morphologically specialised subcategories among non-reproductives (Evans 1958; Rehan & Toth 2015). Within this framework, some species exhibit more or less 'complex', 'advanced' or 'derived' forms of social system, but the differences between each 'rung' on this social ladder are relatively small and evenly-spaced, so that variation in insect sociality is best understood as a smooth spectrum.

An alternative view that has received increased attention in the recent literature is the 'superorganismal' hypothesis that the two broad categories of insect society—those with reproductive plasticity and those without—represent a uniquely meaningful discontinuity that should be given primacy in our understanding of insect social evolution (Boomsma 2009, 2013; Boomsma & Gawne 2018). As long as they have a chance of transitioning to a reproductive role, non-reproductives may be selected to 'freeload', withholding their foraging or nest-building labour in order to save energy for future direct fitness opportunities. Once non-reproductives are irreversibly reliant upon reproductives for reproduction, however, the interests of the two groups will be fully aligned, allowing the evolution of highly specialised and derived traits that maximise colony efficiency (Hölldobler & Wilson 2009; Boomsma & Gawne 2018). An extreme example is the suicidal nest-defence behaviour of some bees and ants (Shorter & Rueppell 2012), a behaviour that eliminates the actor's future direct fitness prospects and therefore can probably only have evolved once those prospects are already nullified. Within this framework, species with irreversible reproductive division of labour are 'superorganisms' that recapitulate many familiar features of multicellular organisms: obligate division of germline and soma in the form of queens and workers (Helanterä 2016); homeostasis (Modlmeier et al. 2019); physiology (Johnson & Linksvayer 2010); and even 'cancers' such as those represented by Cape honeybee social parasites (Oldroyd 2002). This framework posits that superorganisms are meaningfully distinct from species with plastic division of reproductive labour and that other differences, such as gradations in colony size or the presence of specialised non-reproductive subtypes, are secondary (Boomsma & Gawne 2018).

For the purposes of this thesis, I use the term 'superorganismal' as shorthand for the type of insect society that totally lacks reproductive plasticity, and likewise 'non-superorganismal' for societies in which individuals retain the ability to fully transition between non-reproductive and reproductive states. In Chapter 5, I discuss the evolutionary significance of this distinction in detail. More broadly, part of my rationale for focusing on the mechanisms that govern reproductive plasticity in non-superorganismal species in this thesis is the idea that the loss of such plasticity is a necessary precursor for the evolution of at least some of the derived traits that characterise the most complex social insect societies. No part of this thesis is incompatible with the social ladder framework, however. Whether the loss of reproductive plasticity represents a unique evolutionary discontinuity or just one in a series of gradual steps, it remains the

case that the causes and consequences of reproductive plasticity in less derived species form a necessary part of our understanding of the evolution of insect sociality as a whole.

1.2 Reproductive plasticity and alternative fitness strategies

1.2.1 The evolutionary significance of reproductive plasticity

Solitary breeding is the likely basal state (or 'ground-plan') for all social insect lineages. Solitary insects perform reproductive tasks sequentially, switching frequently between them in response to environmental cues such as seasonal fluctuations in the availability of resources (Koštál 2006) or expected mortality risk (Javoiš & Tammaru 2004; Cotter et al. 2011), as well as in response to their physiological state. The sequential execution of reproductive and non-reproductive tasks by the same individual, and the capacity to modify their expression in response to external and internal conditions, implies high reproductive plasticity at the level of the individual. This, however, comes at the cost of trade-offs where the optimisation of the phenotype to the opposing selective demands of each task is limited by the time and energetic costs of continuously remodelling traits (Flatt & Heyland 2011; Murren et al. 2015). These costs will be particularly large in unpredictable environments and for traits that take a long time to modify (Gabriel 2006).

Notably, progressively provisioning eumenine (such as *Synagris cornuta* and *Zethus miniatus*) and sphecid wasps (such as *Ammophila pubescens*), exhibit a particularly pronounced sequential division of reproductive labour. After laying eggs, females in these species go through extended periods of brood care and provisioning during which time their ovaries diminish, in effect cycling between distinct periods of reproductive and non-reproductive labour (Field 1992; Kelstrup et al. 2018). Indeed, it has been proposed that the uncoupling of reproductive and non-reproductive phenotypes in solitary insects such as progressively provisioning wasps represents the evolutionary antecedent of the division of labour that occurs in social insect societies (the 'ovarian ground-plan hypothesis'; West-Eberhard 1996; Amdam et al. 2006; Kelstrup et al. 2018).

The reproductive division of labour exhibited by social insect societies effectively represents a parallelisation of reproductive and non-reproductive activities. Rather than a single individual switching sequentially between phenotypes, in insect societies these phenotypes are partitioned between individuals, freeing reproductives to dedicate a larger proportion of their time and energy towards egg-laying. This parallelisation of roles represents an efficient solution to the trade-offs between reproductive and non-

reproductive investment (Hölldobler & Wilson 2009; Peeters & Molet 2010), which likely contributes to the ecologically dominant role that many such species play in terrestrial ecosystems. A particularly stark outcome of the efficiency of reproductive partitioning in social insect societies is the reversal of the longevity-fecundity trade-off: whereas in most species individuals that expend more energy on reproduction also die younger (all other things being equal; Djawdan et al. 2004), in social insects reproductive individuals may live far longer than non-reproductives (Von Wyschetzki et al. 2015; Blacher et al. 2017).

The primary beneficiaries of insect societies' division of reproductive labour are queens. Because social insects are usually not clonally related, the indirect fitness benefits that workers gain by enhancing the reproductive output of the queen are only partial—*prima facie*, it would be selectively beneficial for any given worker to usurp the queen if she could. In superorganismal species, this conflict is primarily resolved by developmentally-imposed subfertility or sterility of workers, which renders most individuals incapable of reproducing efficiently and thereby aligns queens' and workers' interests. But in those taxa that retain reproductive plasticity, including both species that are the foci of this thesis, the question of how workers are incentivised and/or coerced to continue working rather than reproducing remains unresolved.

1.2.2 Direct fitness incentives

The simplest possible explanation for the persistence of helping among social insect workers is that helpers gain access to direct fitness opportunities that would otherwise be unavailable. This is particularly the case for societies in which between-individual relatedness is low, as is the case in some groups of *P. dominula* co-foundresses (Queller et al. 2000; Field & Leadbeater 2016): in such groups, there are no indirect fitness benefits to be gained, and so subordinates must instead somehow accrue direct fitness gains via their helping activities. One form that such gains might take is in the form of 'cheating', where subordinates lay their own eggs on the nest, but the evidence that this actually occurs is poor. Besides the fact that such cheating is unlikely to actually result in significant direct fitness gains (since the large majority of early-season eggs develop as workers that do not themselves reproduce), dominant foundresses are very efficient at maintaining a reproductive monopoly either by preventing their subordinates from laying eggs or by engaging in oophagy of cheaters' eggs (Monnin et al. 2009; Dapporto et al. 2010).

A second potential source of direct fitness for non-reproductives is reproductive inheritance. If subordinate individuals have a chance of inheriting the reproductive role (for example, in instances of predation), then accepting a non-reproductive role may be selectively advantageous even in the absence of indirect fitness incentives. Indeed, nest inheritance is thought to be an important incentive for subordinate foundresses in *P. dominula* (Leadbeater et al. 2011; Field & Leadbeater 2016) and may also be an important incentive for workers in *L. flavolineata* (Field et al. 2000). In such systems, working may be conceptualised as a form of 'payment' that individuals must give in return for being allowed to remain on the nest and thereby remain in contention to inherit the reproductive position (Ragsdale 1999). In larger groups, however, inheritance becomes increasingly unlikely, especially past the early stages of colony founding: once the first generation of workers has emerged, the chances of queen loss are expected to diminish considerably (Strassmann et al. 2004). Even where nest inheritance occurs, the question of how a replacement queen is chosen from among the group members remains substantially unresolved. I return to this question in Chapter 2.

1.2.3 Indirect fitness incentives

Where opportunities to gain direct fitness via either cheating or nest inheritance are minimal, workers must be incentivised by indirect fitness benefits, which implies that within-group relatedness must be relatively high. One of the strongest lines of evidence favouring a key role for kin selection in favouring the persistence of reproductive division of labour in social insects is the fact that obligate monogamy appears to be the ancestral state for all independent origins of insect sociality (Hughes et al. 2008). Since monandry ensures that a queen's daughters will be strongly related to their siblings ($r = 0.5$ assuming a Fisherian sex ratio), the association between this trait and insect sociality implies that high relatedness is a prerequisite for reproductive division of labour (Hughes et al. 2008; Quiñones & Pen 2017). Superorganisms are the exception that proves this rule: many extant species exhibit strong polyandry, drastically reducing relatedness between workers and their reproductive siblings, but this appears to be a derived trait that has only appeared after the loss of reproductive plasticity has rendered workers obligately subfertile (Boomsma 2009; Boomsma & Gawne 2018).

1.2.4 Coercion

Queens may further incentivise non-reproductives to work via coercion. Coercion may be direct, as in the case of aggression targeted towards underperforming workers (Reeve & Gamboa 1987; Kikuta & Tsuji 1999). Alternatively, it may be indirect: the 'maternal manipulation hypothesis' states that in the earliest insect societies, mothers might have induced subfertility in their daughters by under-nourishing them during development, reducing their potential to accrue direct fitness and thereby incentivising them to remain on the nest as workers (Kapheim et al. 2011, 2015). The social organisation of superorganisms, in which the majority of individuals are induced to develop as subfertile workers, can likewise be conceptualised as a form of coercion (Ratnieks & Wenseleers 2008). Coercion can nonetheless only provide a partial explanation for the maintenance of reproductive division of labour, because even sterile workers must receive some fitness benefit. If workers received zero fitness, then working would not persist even in the absence of costs, and so the ultimate causes of working must still be understood in terms of either direct or indirect fitness benefits, with coercion modulating rather than generating fitness incentives (Linksvayer 2010; Kapheim et al. 2015).

1.3 Mechanisms of caste determination

Multiple mechanistic hypotheses have emerged to describe the molecular evolution of insect sociality. These hypotheses, which are largely mutually compatible, have generated a number of predictions about the nature of the molecular mechanisms that may have underlain the evolution of reproductively differentiated castes.

1.3.1 Solitary ground plan hypotheses

Two separate but related hypotheses posit that social insect castes may have their origins in environmentally responsive genes (West-Eberhard 1996, 2003; Toth & Rehan 2017). The *ovarian ground plan hypothesis* proposes that the ancestral state for social insects is a progressively-provisioning ancestor that went through periods of ovarian activation and quiescence, and that the decoupling of these two phases resulted in the early evolution of castes (Toth & Rehan 2017). This hypothesis predicts that gene networks associated with worker or queen behaviour in the early stages of eusociality may have their origins in gene networks associated with foraging or reproductive behaviour in solitary species (West-Eberhard 1996; Toth & Rehan 2017). Most genetic studies of this hypothesis to

date have relied on data from species with obligately differentiated castes, however (Amdam et al. 2006).

The *maternal heterochrony hypothesis* proposes that worker behaviour in social insect societies has its origins in maternal care behaviour expressed in solitary ancestors (Linksvayer & Wade 2005). Since queens in simple eusocial societies must perform the full range of maternal behaviours during the founding stages of the colony, this hypothesis predicts that the gene networks involved in worker caste expression in mature colonies will be related to the gene networks involved in queen behaviour during the founding stage, as well as those networks involved in maternal behaviour in solitary ancestors (Linksvayer & Wade 2005). Evidence for this hypothesis is mixed, with its predictions supported in a species of paper wasp but not in bumblebees (Toth et al. 2007; Woodard et al. 2014).

1.3.2 The genetic toolkit hypothesis

Following from solitary ground-plan hypotheses, and drawing inspiration from findings in evolutionary developmental biology that convergent morphological innovations in different lineages have often involved conserved sets of genes (True & Carroll 2002), the genetic toolkit hypothesis proposes that a conserved set of genes and gene networks has been involved in multiple independent evolutionary origins of insect sociality (Toth & Robinson 2007; Toth et al. 2010). This hypothesis thus predicts that the same sets of genes may be involved in caste expression even when comparing phylogenetically distinct social insect taxa.

Comparisons of caste gene expression across the Hymenoptera have produced mixed results with regards to the predictions of the genetic toolkit hypothesis: while some genes involved in foraging and aggression do appear to be conserved across multiple distinct social lineages, the numbers of shared genes between lineages is small and in some cases inconsistent with regards to the direction of caste bias (Toth et al. 2010, 2014; Wyatt et al. 2020). It is possible that particular pathways or networks of genes are more likely than individual genes to be conserved across the domains of social insect life (Berens et al. 2015), but a paucity of cross-taxon comparisons means that this hypothesis remains unconfirmed.

1.3.3 The 'novel genes' hypothesis

The 'novel genes' hypothesis proposes that novel protein-coding gene sequences have been important in the evolution of social insect innovations such as caste (Sumner 2014). This hypothesis predicts that taxonomically-restricted genes may comprise a large proportion of the genome in social insect lineages, and this prediction has received significant support from empirical studies. Novel genes have been identified in each new social insect genome that has been sequenced and such sequences appear to overrepresented among caste-biased genes (Feldmeyer et al. 2014). Multiple studies have found that worker-biased genes are more likely to be taxonomically restricted than queen-biased genes, suggesting that the evolution of worker phenotypes has involved more functional innovation than that of queen phenotypes (Ferreira et al. 2013; Feldmeyer et al. 2014; Jones et al. 2017; Warner et al. 2017), although the strongest evidence has come from more derived species with morphologically distinct castes. In Chapter 4 of this thesis, I test this hypothesis in the facultatively social hover wasp *L. flavolineata*, with surprising results.

1.3.4 Epigenetic mechanisms of caste determination

Several molecular mechanisms other than gene expression *per se* are thought to play a role in maintaining social insect caste expression. These mechanisms should be considered complementary to, rather than distinct from, gene expression itself, since such mechanisms will both be the products of gene expression and likely to produce their effects partly through further downstream effects on the transcriptome. While the bioinformatic portions of this thesis focus on transcriptomic data, here I briefly review the evidence for the role of other molecular mechanisms for context and completeness.

Epigenetic mechanisms offer a stable means of altering gene expression levels without changes in gene sequence, and as such ought to be prime candidates for a role in caste expression, especially in species with incomplete caste plasticity. Drawing comparisons with mammalian development (in which levels of epigenetic modification increase with increasing cell differentiation), Patalano et al. (2012) have proposed that reduced reproductive plasticity in the transition to insect sociality ought to be associated with increasing evidence of methylation, while Maleszka et al. (2014) have emphasised the potential role of methylomic degeneracy in producing functionally divergent castes from a single genome.

In the last decade the role of methylation in social insect differentiation has been the topic of extensive research (Kronforst et al. 2008; Weiner et al. 2012; Yan et al. 2014), but it has proven difficult to establish a consistent pattern of association between methylation and social complexity (Weiner et al. 2013; Bewick et al. 2017). Methylation appears to be associated with caste in some species of wasps (Weiner et al. 2013) and bees (Li et al. 2018) with reproductive plastic castes, but two separate studies have found no significant role for methylation in caste regulation in *Polistes canadensis* (Patalano et al. 2015) and *Polistes dominula* (Standage et al. 2016), and indeed both these species appear to have lost a key methyltransferase gene (DNMT2). In superorganismal ants, several papers have linked caste to levels of methylation (Kucharski et al. 2008; Bonasio et al. 2012; Foret et al. 2012), but the methodology of these and other methylation studies has been called into question (Libbrecht et al. 2016).

Histone acetylation alters the accessibility of chromatin, and thus joins DNA methylation as a possible stable source of reproductive differentiation. To my knowledge there is no evidence of a role for histone acetylation in reproductive differentiation in insect species that retain full reproductive plasticity. The discovery of histone deacetylase inhibitor (HDACi) as a component in the royal jelly of honey bees suggests a role for this mechanism in superorganismal reproductive caste determination (Spannhoff et al. 2011), as does the recent finding that caste-specific chromatin modification patterns are present early in honey bee development (Wojciechowski et al. 2018). Royal jelly also contains noncoding microRNAs (miRNAs), another possible reproductive regulator (Guo et al. 2013).

Hormonal control, which represents a form of within-organism signalling that is less transient than gene expression but less enduring than epigenetic mechanisms, appears to be an important correlate of caste at all levels of insect sociality. Juvenile hormone (JH) is the primary regulator of reproduction in solitary insects (Nijhout 1998; Roy et al. 2018), including in progressively provisioning wasps that may represent the likely basal state for wasp sociality (Giray et al. 2005; Tibbetts et al. 2013), and this role appears to be maintained in social species. JH is associated with reproductive differentiation in non-superorganismal bees (Hartfelder et al. 2006), termites (Korb 2015) and wasps (Kelstrup et al. 2015); and likewise in ants (Penick et al. 2012; Libbrecht et al. 2013) and honeybees (Capella & Hartfelder 2002; Bomtorin et al. 2014). It thus appears that JH is a mediator of reproductive differentiation across insect lineages of all degrees of sociality.

1.3.5 Mechanisms of caste plasticity

A large number of studies have investigated the behavioural and molecular mechanisms by which castes are maintained in insect societies and by which caste differences emerge during development in superorganismal species (e.g. Korb 2015; Vojvodic et al. 2015; Ashby et al. 2016; Klein et al. 2016; Smith et al. 2018; Wojciechowski et al. 2018; Sasaki & Harada 2020), but relatively little attention has been given to the mechanisms of caste plasticity itself in species that retain the ability to transition between reproductive roles (Chandra et al. 2018; Libbrecht et al. 2018). This is surprising given that the loss of this plasticity is a critical precursor for the extreme form of sociality represented by superorganisms (Boomsma & Gawne 2018): understanding how such plasticity is achieved in species that have distinct but plastic castes must, therefore, be a prerequisite to understanding insect sociality as a whole. Several questions in this area remain substantially unanswered. To what degree is the apparently high caste plasticity in less derived social insect species reflected at the molecular level? To what extent are workers that transition to become queens able to fully match the phenotypic and molecular roles of foundress queens? By what mechanisms are replacement queens selected, and how is queen loss reflected at the level of e.g. worker gene expression? And, how is the persistence of potential reproductive plasticity reflected in the phenotypic and molecular profiles of workers? In Chapters 3 and 4 of this thesis, I aim to provide answers to some of these questions.

1.4   Thesis overview

1.4.1 Study species

This thesis concentrates on two species of social wasp within the family Vespidae (**Table 1.1**). Here I present only a summary overview of these two species' ecology, but relevant aspects are discussed in greater depth in the relevant chapters of this thesis.

Polistes dominula (commonly referred to as *Polistes dominulus* or *Polistes gallicus* in older literature), the Eureopean paper wasp, is a temperate species whose native range covers much of southern Europe, North Africa, and temperate parts of Asia (Carpenter 1996). It is also an important invasive species in North America (Johnson & Starks 2004), where it appears to be displacing native paper wasps (Gamboa et al. 2004). *P. dominula* gynes mate in late Summer or early Autumn before overwintering (Reeve 1991). In early Spring of the following year, gynes emerge and found small colonies either singly or in

multi-foundress groups, which may consist of unrelated individuals (Reeve 1991; Field & Leadbeater 2016). As in other social Hymenopterans, fertilised eggs always develop as females while unfertilised eggs always develop as males, allowing *P. dominula* queens to determine the sex of their offspring. Workers are always female and males are usually absent until later in the season when gynes emerge, although a small number of early males may be produced (possibly to take advantage of opportunities presented by unmated replacement queens; Turillazzi 1980; Strassmann et al. 2004).

Ethologist Leo Pardi described several aspects of social organization in *P. dominula* (which he identified at that time as *P. gallicus*) in a number of pioneering papers in the 1940s (Pardi 1942, 1946, 1948), and since this time the species has become a model for investigations into the behavioural mechanisms that govern small insect societies (Starks et al. 2006). Features of *P. dominula*'s ecology that make it well-suited for such investigations include the fact that nests are not enclosed, facilitating easy observation and manipulation, and that group sizes are small (usually well under 100), making social dynamics relatively easy to track. Nests are frequently found in relatively dense aggregations around man-made structures such as roof tiles, which again facilitates collection and observation. *P. dominula* workers retain full reproductive plasticity, and the loss of a queen is known to result in the inheritance of the nest by a former worker (Strassmann et al. 2004).

	Polistes dominula	*Liostenogaster flavolineata*
Sociality	Obligate sociality	Facultative sociality
Caste plasticity	Yes	Yes
Native range	Eurasia (Temperate)	South Asia (Tropical)
Group size	Usually <50	Usually <10
Dominance hierarchy	Linear, age-based	Linear, age-based

Table 1.1. Focal species studied in this thesis. Images: Eugene Zelenko & David Baracchi.

My other focal species is the tropical hover wasp *Liostenogaster flavolineata*, which is native to South-East Asia (Cronin et al. 2011). Like *P. dominula*, *L. flavolineata* groups occupy open-faced nests that are easy to observe and manipulate, and workers are fully capable of inheriting the reproductive position in a group (Bridge & Field 2007). Unlike *P. dominula*, however, this species occupies a tropical environment, meaning that nests can continue to operate uninterrupted as long as they are not destroyed by predators or weather. Although within-group relatedness is high in *L. flavolineata*, it has been speculated that small group sizes (<10 wasps/group) and lack of seasonal constraints may result in direct fitness forming a large component of lifetime fitness for workers in this species (Field et al. 1998, 2000), since these factors increase the likelihood that individuals will be able to inherit the reproductive position within a colony. While several papers have explored the social structure of this species at a behavioural level (Field et al. 2000; Bridge & Field 2007; Baracchi et al. 2015), molecular data pertaining to hover wasp sociality are almost entirely lacking.

1.4.2 Terminology

As stated above, for the purposes of this thesis I have preferred to remain agnostic regarding the truth or otherwise of specific ontological frameworks that attempt to describe differences between different insect societies. I have also attempted where possible to avoid certain terms that are common in the social insect literature but that lack either specificity or consistency in their usages. In particular, I do not refer to any taxon as 'eusocial' because this term appears to hold different meanings for different researchers, to the extent that some reserve the term exclusively for the very most derived social insect species (Wilson 1971, 1975; Boomsma & Gawne 2018), while others extend the category to cover such diverse groups as killer whales (McAuliffe & Whitehead 2005) or humans (Nowak et al. 2010; Joiner et al. 2016). I instead simply refer to insects as social or solitary dependent upon whether or not they form reproductively skewed groups. Except where I use 'superorganismal' as shorthand for species with obligately irreversible loss of reproductive plasticity, I have attempted to restrict myself to referring to variation between species in relative terms, e.g. 'more derived' or 'more complex'.

In lieu of the rather cumbersome 'reproductive' and 'non-reproductive', hereafter I will use the standard parlance of social insect research: dominant reproductive females are 'queens', related non-reproductives are 'workers', and the state of being reproductive or non-reproductive is an individual's 'caste'. Where the non-reproductive individuals in a

group are not necessarily genetically related to the queen, as in *Polistes* co-foundress associations (Queller et al. 2000; Field & Leadbeater 2016), I instead use the term 'subordinates'.

1.4.3 A note on vespid phylogeny

The vespid family comprises over 5000 species, most of them solitary but including a number of social taxa spanning a large range of social complexity. These include such well-studied taxa as the yellowjacket wasps *Vespula* spp. and *Dolichovespula* spp., the hornets *Vespa* spp., and the *Polistes* paper wasps, which together form a monophyletic group (Vespinae). The phylogenetic position of the hover wasps (subfamily Stenogastrinae), which also exhibit sociality, has remained a point of contention within the literature (reviewed in Bank et al. 2017). The question of whether the Stenogastrinae represent a sister group to the Vespinae is important for understanding the evolution of wasp sociality. If these two groups form a monophyletic clade, then the most parsimonious explanation is that sociality has evolved just once among vespids (i.e. vespid sociality is a synapomorphy), but if the Stenogastrinae are phylogenetically separated from the Vespinae then sociality may have evolved twice (i.e. vespid sociality is homoplastic). Though several conflicting phylogenies have been claimed (Schmitz & Moritz 1998; Carpenter 2003; Hines et al. 2007; Pickett & Carpenter 2010), the most recent and most complete arrangement is that generated by Bank et al. (2017). This phylogeny places the Stenogastrinae as basal to the remainder of the Vespidae, strongly implying that sociality evolved twice independently in this family. This finding provides important context for results that I present in Chapter 4.

1.4.4 Thesis structure

The remainder of this thesis is divided into 3 data chapters, a short review chapter, and a final discussion chapter that serves as a synthesis of and conclusion to the preceding work. The structure of each of the data chapters broadly reflects that of a journal publication, with each chapter given its own methods and discussion section.

In **Chapter 2**, I examine individual- and group-level responses to queen loss in the European paper wasp *Polistes dominula* in order to understand the mechanisms by which groups accommodate the loss of a reproductive. Specifically, I test whether queen succession in this species is regulated by convention- or conflict-based mechanisms, and

I investigate which traits predict the identity of successor queens. I find that groups are able to generate replacement reproductives rapidly and with little conflict by relying on an age-based succession criterion in which the oldest individual(s) in a group take on the newly vacant reproductive role. I furthermore show that older individuals do not appear to be physiologically advantaged to become queens, which leads me to conclude that age acts as a 'conventional' signal to ease reproductive succession rather than providing inherent benefits. I also contrast *P. dominula* to tropical species in the same genus in order to argue that the convention-based succession mechanism possessed by this species is a response to ecological constraints upon colony lifespan. The results I present in this chapter shed light on the mechanisms by which caste stability can be achieved in the absence of obligate worker subfertility, an important step toward a full understanding of the earlier stages of insect social evolution.

In **Chapter 3**, I analyse the transcriptomic mechanisms that underlie the rapid and low-conflict succession process exhibited by *P. dominula*. I test the hypotheses that individuals' gene expression profiles will closely track their phenotypes, and that as a consequence replacement queens will strongly resemble their predecessors at the transcriptomic level, while individuals that remain as workers will not exhibit significant transcriptomic responses to queen loss. Combining a novel machine learning approach with standard methods, I show that individual-level transcriptomic responses to queen loss in *P. dominula* in fact consist of two components. One component correlates with caste expression as expected, with individuals that transition phenotypically from workers to queens acquiring a more queenlike gene expression profile than those that remain as workers. The other component, however, is a general de-differentiation in which all individuals become transcriptomically intermediate between queens and workers from control colonies, regardless of behaviour or ovarian development. This second effect may simply reflect a stress response to the loss of the queen, but might instead represent a form of 'priming' in which the loss of one queen causes individuals to anticipate subsequent additional instances of queen turnover. As well as indicating an overlooked disconnect between caste-biased transcriptomes and phenotypes, and demonstrating a potentially powerful new tool with which to explore the molecular correlates of caste, this chapter's results reveal potential limitations of the seemingly-efficient succession process investigated in Chapter 2. Understanding such limitations of caste plasticity is important if we are to understand why obligately subfertile castes were selected for in other social insect lineages.

In **Chapter 4**, I examine the gene expression consequences of investment into alternative fitness strategies in a facultatively social species of tropical hover wasp, *Liostenogaster flavolineata*, that represents an independent origin of sociality among the Vespidae. I hypothesise that the shifts in fitness interests that exist within groups of this species will be tracked strongly at the level of gene expression, and that as a result the strongest determinant of within-group variation in transcriptomes will be dominance rank rather than caste. I also take advantage of the fact that this species represents an independent origin of insect sociality to test the hypothesis that direct and indirect fitness biased genes should show divergent patterns of evolutionary conservation. At the phenotypic level, I confirm that *L. flavolineata* workers strongly vary their foraging effort based on their rank within a group's social hierarchy, indicating that individuals modulate their investment into indirect fitness gains based on their likelihood of attaining the reproductive role in the future. Taking advantage of this system to identify genes whose expression is associated with shifts in fitness strategies I find that, contrary to my predictions, caste is a much stronger predictor of gene expression than rank or foraging rate. These results indicate that despite being fully capable of nesting independently, members of *L. flavolineata* groups exhibit patterns of gene expression that closely resemble those of more derived species that are obligately social. Additionally, I compare my results to those obtained for *P. dominula* in Chapter 3; doing so, I find evidence that genes associated with indirect fitness in *L. flavolineata* are more strongly conserved than those associated with direct fitness, an unexpected result that demonstrates the potential value of this little-studied taxon as a focus for bioinformatic investigations of the evolution of insect sociality.

Chapter 5 consists of a short review of patterns of reproductive plasticity across the spectrum of insect sociality. In the context of multi-level selection theory and the concept of major evolutionary transitions in individuality, in this chapter I argue that the transition to 'superorganismal' societies can be reconceptualised as involving a shift in reproductive plasticity between levels of biological complexity, rather than as a loss of reproductive plasticity *per se*.

Chapter 6 serves as a synthesis of the preceding chapters, in which I review the key results generated by the thesis.

In **Appendix A**, I present the results of a simple analysis in which I used published social insect transcriptomic data to generate simulated gene expression datasets and then perform power analysis based on different sequencing approaches; the results of this analysis informed the experimental design that I used in Chapter 3. **Appendix B** gives details of the *L. flavolineata* genome assembly against which the transcriptomes analysed in Chapter 4 were aligned. Supplementary figures for each chapter are provided in **Appendix C**. In-text references to supplementary figures and tables are prepended by the letter S (e.g. "Figure S1.1"). As some of the supplementary tables for this thesis are very large and would take up excessive space in print, supplementary tables are provided externally as .xlsx files.

1.4.5 Notes on contributions and publications

Unfortunately, time- and pandemic-related restrictions necessitated that I take advantage of pre-existing data for my final data chapter. The fieldwork described in Chapter 4 was performed by Daisy Taylor and the *L. flavolineata* genome assembly was completed by Heinz Himmelbauer (University of Natural Resources and Life Sciences, Vienna) and Roderic Guigo (Centre for Genomic Regulation, Barcelona) as a collaboration with my supervisor, Seirian Sumner. All other work is my own, with contributions from Alessandro Cini and Rita Cervo, both of whom advised on experimental design and fieldwork for Chapter 2; Michael Bentley, who advised on the SVM methodology employed in Chapter 3; and Chris Wyatt, who troubleshot various bioinformatic questions that arose in the course of the thesis. Seirian Sumner and Max Reuter, my supervisors, contributed to all aspects of the thesis.

To date, the work herein has generated three journal articles (one currently in press). Taylor et al. (2020) forms the bulk of Chapter 2 of this thesis, Taylor et al. (Nature Communications, in press) forms Chapter 3, and Taylor et al. (2019) forms Chapter 5.

Chapter 2 – Phenotypic and group-level correlates of queen succession in *Polistes dominula*

The results presented in this chapter have been published as: Taylor, B. A., Cini, A., Cervo, R., Reuter, M., & Sumner, S. (2020). Queen succession conflict in the paper wasp *Polistes dominula* is mitigated by age-based convention. *Behavioral Ecology, 31*(4), 992-1002.

2.1   Introduction

Reproduction in insect social groups is dominated by one or a few individuals, with the remaining group members relegated to a non-reproductive worker role. The question of how such societies deal with potential conflict over the partitioning of reproductive opportunities remains unresolved. As in other social insect species in which individuals retain a high degree of reproductive plasticity, inheritance of the breeding position within *Polistes dominula* colonies represents a potentially important fitness incentive for workers. The value of inheriting the reproductive role in a group following queen loss is also a potential source of conflict between workers, however—if replacement queens are chosen by aggressive contests between workers, queen loss could result in significantly deleterious disruption to colony functioning. Reproductive succession is thought to be determined via an age-based convention in some *Polistes* species, but there is also evidence for contest-based succession systems in which the replacement queen uses physical aggression to overpower and thereby subordinate her nestmates.

In this chapter, I provide evidence that queen succession in *P. dominula* is determined via convention rather than contest, with little disruption to the colony's social functioning. I use queen removal experiments and fine-scale behavioural analyses to confirm that age is a strong predictor of succession, and that behavioural responses to queen removal are restricted to the oldest individuals rather than being experienced equally across the group. The results presented here provide one of the most comprehensive and detailed experimental analyses of the dynamics of reproductive succession in a social insect to date, thereby shedding light on the mechanisms by which insect societies are able to maintain cohesion in the face of within-group conflict.

2.1.1 Mechanisms of reproductive succession

In cooperatively-breeding groups, such as those of non-superorganismal social insects, the loss of a reproductive dominant opens up opportunities for conflict within otherwise stable societies. Subordinate individuals may compete for the newly vacant reproductive role, but this competition is likely to come at a cost to the group as a whole (Gobin et al. 2003; Strassmann et al. 2004). It is thus predicted that selection will favour the evolution of mechanisms that facilitate the replacement of lost dominant individuals without outright conflict (Aureli & de Waal 2000; Aureli et al. 2002). Subordinate individuals may instead compete for a vacant reproductive role via 'conventional' traits that serve to differentiate between candidates reproductives without direct conflict. For example, while dominance and reproductive succession are determined by aggressive interactions in many vertebrate societies (e.g. Creel et al. 1992; Clarke & Faulkes 1997; Nichols et al. 2012), in other societies wholly or partially conventional phenotypic traits such as song complexity, age or length of tenure may serve as cues to dictate dominance without the need for outright conflict (e.g. East & Hofer 1991; Spencer et al. 2004; Duncan et al. 2018). Likewise, evidence exists for both conflict-based and convention-based reproductive succession mechanisms in insect societies. Injurious fights over queen succession are particularly likely to occur when candidate reproductives are unrelated, as in the case of pleometrosis in ants (e.g. Bernasconi & Strassmann 1999) or multiple-foundress nesting in social wasps (e.g. West-Eberhard 1969).

Reproductive succession in invertebrate groups with higher relatedness appears to more often involve traits such as age (e.g. Bridge and Field 2007). Evidence for the role of age in predicting dominance is particularly strong in *Polistes* paper wasps (Pardi 1948; Strassmann & Meyer 1983; Miyano 1986; Hughes & Strassmann 1988; Tsuji & Tsuji 2005) but it remains unclear whether the role of this trait in fact reflects differences in e.g. fighting ability or ovarian development, or whether it represents a genuinely arbitrary convention. While several studies have involved observations of *Polistes* colonies from which queens have either been experimentally removed or naturally lost, few of these have paid close attention to the social dynamics that give rise to replacement reproductives. In particular, only a very small number of studies have investigated the mechanisms of queen succession in nests following the eclosion of workers (**Table 2.1**). Reproductive conflicts during this worker phase are expected to be radically different from those that occur during the founding phase because workers are usually closely

related to one another and to the queen, which subordinate foundresses may not be (Zanette & Field 2008; Leadbeater et al. 2011; Field & Leadbeater 2016).

The close relatedness among siblings within post-eclosion nests means that fitness interests of individuals in such groups are more strongly aligned than in foundress groups, where relatedness can be low (Zanette & Field 2008; Leadbeater et al. 2011), and this may have favoured the evolution of more robust conflict-resolution mechanisms in post-eclosion nests. Unlike clonal social insects (e.g. social aphids Uematsu et al. 2013), however, groups of *Polistes* workers are not genetically identical and so some degree of selfish behaviour may persist.

2.1.2 Chapter overview

In this chapter I conduct a detailed analysis of the social dynamics of queen succession in *P. dominula*. By removing queens from monogynous colonies, I generate detailed data on behavioural and ovarian development with which to test several hypotheses pertaining to the social mechanisms of queen replacement in this species. I begin by determining the individual-level traits that dictate queen succession (Section 2.3.1). I predict that queen succession is explained either by age, which is known to influence dominance in many *Polistes* species (Pardi 1948; Strassmann & Meyer 1983; Miyano 1986; Hughes and Strassmann 1988; Tsuji and Tsuji 2005), suggesting convention; or alternatively by size, a trait thought to influence fighting ability in *Polistes* (Turillazzi & Pardi 1977; Cervo et al. 2008; Tibbetts & Shorter 2009; but see Reeve et al. 2000; Cant et al. 2006), suggesting contest. I then test the hypothesis that queen succession reflects pre-existing physiological constraints on ovarian development (Section 2.3.2). Next, I examine group-level responses to queen removal in order to test the prediction that queen removal will be followed by significant social disruption if succession is dictated by conflict, but not if it is dictated by convention (Section 2.3.3). I additionally investigate the distribution of behavioural responses to queen removal (Section 2.3.4): these should be evenly distributed among multiple individuals within each colony under a contest-based succession mechanism, whereas in the case of convention-based succession only one or a few individuals should undergo significant behavioural changes. Finally, I test the hypothesis that group-level responses to queen removal will be minimized when there is reduced ambiguity regarding the identity of the replacement queen (Section 2.3.5).

Table 2.1. Review of published literature examining the effects of queen loss in post-emergence *Polistes* colonies. *Fighting*: evidence of fighting between potential successors. *Colony disruption*: evidence of significant disruption to colony functioning or growth following queen loss. *Conventional cues*: evidence that traits other than size or fighting ability predict queen succession. *Lack of colony disruption*: evidence that colony functioning or growth are unaffected by queen loss. Parentheses indicate partial evidence. Fuller details are provided in supplementary table **S2.1**.

Paper	Species	Data type: Behavioural (B) / Genetic (G) / Ovarian (O) / Chemical (C)	Subordinate foundresses present?	Evidence for conflict		Evidence for convention	
				Fighting	Colony disruption	Conventional cues	Lack of colony disruption
Hughes et al 1987	Polistes annularis	B	Both considered	Y			Y
West-Eberhard 1969	Polistes canadensis	O	N		Y		
Miyano 1986	Polistes chinensis antennalis	B	N	Y	Y	Y	
Dapporto et al 2005	Polistes dominula [1]	CO					
Strassman et al 2004	Polistes dominula [1]	BO	N	(Y)	Y		
Monnin et al 2009	Polistes dominula [1]	BO	Y				Y
Tibbetts & Huang 2010	Polistes dominula [1]	BC	N	(Y)			
Strassman & Meyer 1983	Polistes exclamans	B	N	(Y)	Y	Y	
West-Eberhard 1969	Polistes fuscatus	O	N				
Reeve & Gamboa 1983	Polistes fuscatus	B	N		(Y)		
Page et al 2002	Polistes fuscatus	BO	Not reported				
Pardi 1948	Polistes gallicus [2]	BO	Both considered			Y	
Hughes & Strassman 1988	Polistes instabilis	BO	N			Y	
Miyano 1991	Polistes jadwigae	BO	N	Y	Y	Y	
Metcalf & Whitt 1977	Polistes metricus	G	Both considered			Y	(Y)
Metcalf 1980	Polistes metricus	BG	N				
Hagiwara & Kojima 2002	Polistes nipponensis	B	N				
Metcalf 1980	Polistes variatus	BG	N				

1 *P. dominula* is frequently referred to as *P. dominulus*

2 *P. gallicus* and *P. dominula* are difficult to distinguish and were not widely recognised as separate until ~1980; it is thus possible that Pardi's (1948) data actually derive from *P. dominula*

2.2 Methodology

2.2.1 Colony collection and ageing of wasps

Polistes dominula colonies ($N = 76$) were collected from rural areas near Florence, Italy, and transferred to a laboratory in the University of Florence during mid-May 2017, before the emergence of the first brood. Only nests with at least 20 cells and at least one capped cell were collected. *P. dominula* wasps typically do not initiate flight when the ambient temperature is below ~22°C (Weiner et al. 2011) and colonies were therefore collected in the morning when temperatures were still low (<20°C) to ensure that all nestmates were collected together. Colonies were maintained in glass boxes (15 cm x 15 cm x 15 cm) under natural light conditions with ad libitum access to water, sugar, dipteran larvae and nest materials (cardboard and paper). On cool days (midday temperatures <25 °C) heaters were used to maintain high daytime temperatures within the laboratory; otherwise, temperatures were ambient. Immediately after colony collection, individuals were given unique markings by applying spots of coloured paint (Testor Corporation) to the thorax and/or to the tips or upper portions of the wings.

Both single foundress (monogynous) and multi-foundress (polygynous) colonies were collected (mean 1.9 foundresses/colony; range 1-5 foundresses/colony). In order to ensure that all adult offspring emerged into an equivalent social environment on each colony, all subordinate foundresses were removed from the multiple-foundress colonies. This loss of subordinate foundresses is not an unnatural state for *Polistes* colonies, which experience high rates of foundress loss in the wild (Strassmann 1981; Miyano 1986; Strassmann et al. 2004). Colonies were observed for five minutes hourly for two days following relocation to the lab in order to identify the dominant foundress. Dominant foundresses were identified based on two criteria: firstly, the dominant spends the large majority of her time on the nest and occupies the central portion of the nest carton, whilst her subordinate co-foundresses forage (Baracchi 2017); secondly, dominant individuals could be observed antennating subordinate individuals (see Methods; Pardi 1948; Jandt et al. 2014), while subordinates were never observed antennating dominants. In all cases, the dominant foundress of a colony was identified within two days of collection and before the eclosion of adult workers. Once the dominant foundress of a colony had been identified, all subordinates were removed from the nest box. Colonies were then checked daily and any newly emerged individuals were given unique identification markings. This

allowed me to assign ages to each emerging individual in each colony, to an accuracy of 24 hours.

2.2.2 Queen removal experiments

Mature colonies were randomly allocated either to a queen removal (QR), or to a control (C) treatment. Control colonies were subjected to a sham removal in which an individual was removed from and then immediately placed back onto the nest. To capture the process of queen succession, I further randomly separated colonies into those in which I sampled wasps and assessed their reproductive dominance early during the succession period and those where this was done late in succession. A previous queen removal study in *P. dominula* (Strassmann et al. 2004) found that the process of queen replacement is ongoing at three days following queen loss but largely complete after twelve days. In line with this, I chose three and twelve days following queen removal as focal timepoints for my analyses. The four treatments are henceforth referred to as QR3 ($n=20$), QR12 ($n=20$), C3 ($n=20$) and C12 ($n=15$).

A colony was considered 'mature' once it had produced a minimum of four adult offspring (Dapporto et al. 2005). Each colony was manipulated according to its treatment group on the morning of the fourth day following maturity. All manipulations were performed between 07:00 and 08:00, during which time period the temperatures in the lab were still well below 22°C. For colonies belonging to QR3 or QR12 treatments, manipulation consisted of removing the queen and immediately preserving her body in a 1.5 mL Eppendorf tube containing 80% EtOH, which was stored at -20°C for later dissection. For C3 and C12 treatments, manipulation consisted of removal of a randomly pre-selected non-foundress individual. Following this manipulation, each colony was maintained under standard lab conditions for an additional period of time dictated by treatment group: three days for QR3 and C3 colonies, and twelve days for QR12 and C12 colonies (**Figure 2.1A**).

2.2.3 Ovarian and body size measurements

All individuals were dissected after completion of the experiment. For each individual, an index of ovarian development was obtained by removing and photographing the ovarioles, and subsequently measuring the mean length of the six largest oocytes/ova present. This is a standard measure of ovarian development and has been shown to strongly correlate with more complex ovarian indices in *P. dominula* (Cini et al. 2013).

Individuals with extremely underdeveloped oocytes (all oocytes <15 μm in length) were assigned an ovarian index of 15 μm as measurements were imprecise below this value.

I obtained measures of body size by removing the left forewing of each individual and measuring the maximum distance from the tip of the wing to the thoracic wing joint, which is a reliable proxy for body size in *Polistes* wasps (Haggard & Gamboa 1980; Cant et al. 2006). Wing length was used in place of head width because head tissues had been consumed in a separate analysis. All measurements were performed using *Fiji*, a distribution package of the image processing program *ImageJ* 1.52 (Schindelin et al. 2012; Schneider et al. 2012). All measurements were performed by a single observer, blind with respect to each individual's identity and treatment group.
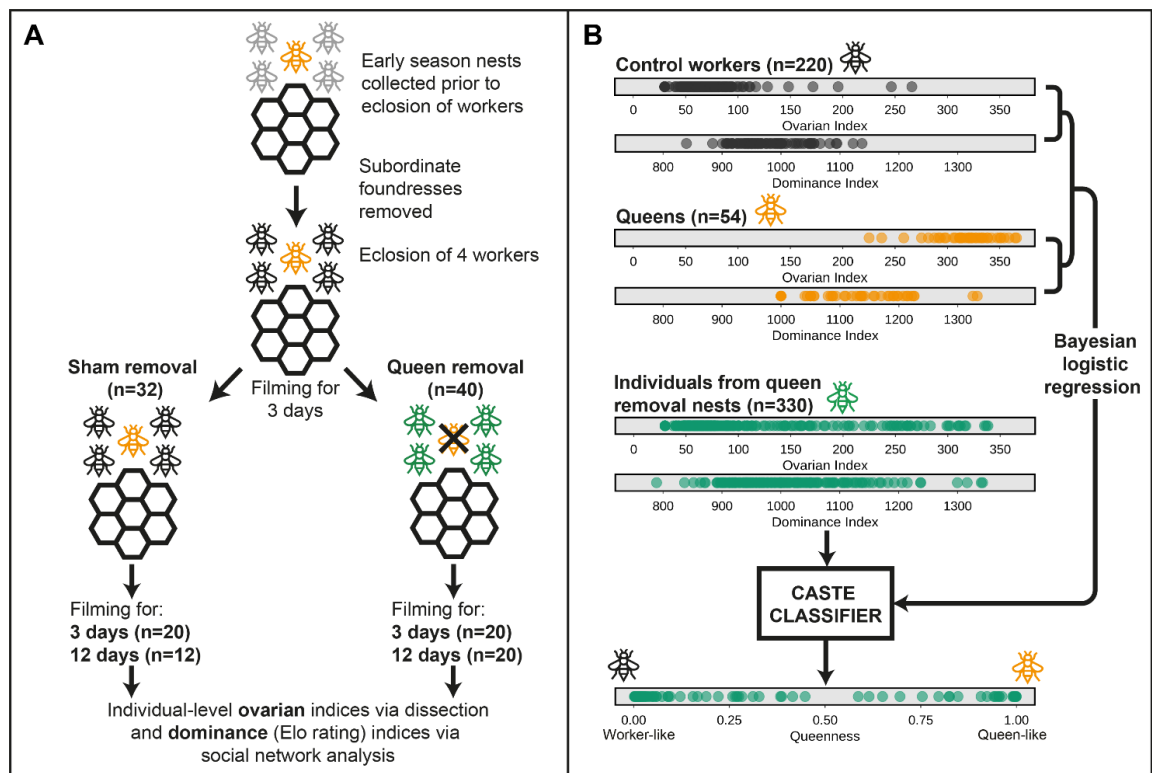
Figure 2.1. Summarized experimental design describing the generation of phenotypic data in this chapter. (A) Early-season nests were transferred to the lab prior to the eclosion of workers and subordinate foundresses were removed. After at least 4 workers had eclosed, nests were filmed for three days and then assigned to either control or queen removal treatments. Following treatment (sham or queen removal) nests were filmed for a further three or twelve days and then all individuals were dissected to generate ovarian development indices following Cini et al. (2013). Footage of nests was used to generate dominance indices in the form of Elo ratings (Elo 1978; Neumann et al. 2011). (B) Ovarian and dominance indices from queens (orange) and control workers (black) were used to produce a logistic regression model for caste classification (0=worker, 1=queen). Data from individuals on queenless post-removal nests (green) were then passed through this model to fit caste estimates and thereby identify individuals with high 'queenness', i.e. those that exhibited strongly queen-like phenotypes following queen removal.

2.2.4 Behavioural recording of nestmate interactions

To obtain data regarding the individual- and group-level mechanisms of queen replacement, I recorded the individuals' behaviour before and after manipulation. Colonies were filmed before and after experimental manipulation in order to assess individual-level and group-level effects of queen loss. Recordings were made for 30 minutes daily for three days before and three days after experimental manipulation, using Sony HDR-CX405 HD video cameras mounted on tripods. Additionally, colonies in the QR12 and C12 treatments were filmed for 30 minutes daily on days 10 through 12 following experimental manipulation. All filming occurred between the hours of 10:30 and 16:30 when *P. dominula* activity levels are at their highest (Cini et al. 2013). The time of filming for each colony and day was randomized. Between 07:00 and 08:00 on the morning following the end of the treatment-specific time period, the entire colony was terminated by removing all individuals and preserving their bodies in alcohol as described above.

Digitally recorded behaviours for each colony were annotated using BORIS observation software (Friard & Gamba 2016). To permit efficient analysis of the >200 hours of videos recorded, behaviours not directly related to dominance interactions (e.g. nest building) were not recorded. An observer recorded the time and duration of each instance of dominance behaviour that occurred within each video, and also tracked the proportion of time spent on and off the nest carton by each individual within each 30-minute video. In line with previous behavioural observations of *P. dominula* colonies, mounting followed by antennation was by far the most common behaviour observed at all treatment stages (Pardi 1948; Tibbetts & Huang 2010; Jandt et al. 2014). More aggressive dominance interactions, such as wing chewing, were observed extremely rarely (<20 times across the entire experiment) and so all subsequent analyses focus solely on mounting followed by antennation (hereafter referred to simply as 'antennation').

To reduce the likelihood of false positives, antennation bouts were only recorded if they were longer than 1 second in duration. Antennation bouts that did not occur in conjunction with an unambiguous instance of mounting were not recorded, as these may simply reflect affiliative or communicative functions rather than dominance behaviour. In each instance of antennation, the dominant and subordinate individuals' identities were recorded. Dominant and subordinate actors in a bout of antennation are usually easily identified: the dominant strikes the subordinate with her antennae while the subordinate remains still

and lowers her head and body to the nest surface (Pardi 1948; Jandt et al. 2014). In a small proportion (~5%) of antennation bouts, antennation was resisted by the targeted individual, resulting in an inconclusive struggle between the two individuals without either establishing a clear dominant role. For the purposes of downstream social network analyses, these bouts were recorded as draws.

In order to test whether queen removal results in a group-level change in dominance behaviour, I calculated the mean and variance of antennation rate within each colony for each three-day time period. I additionally calculated the average proportion of time individuals spent off the nest carton (a proxy measure of participation in off-nest activities such as foraging or inter-nest drifting) for each time period, to test whether queen removal exacts a cost in terms of the group's ability to continue normal colony functions. As interactions were only recorded when they occurred on the nest carton itself, antennation rate was measured relative to the amount of time that an individual was present on the nest.

2.2.5 Assignment of Elo ratings to estimate within-group dominance

I used the Elo rating system (Elo 1978; Albers & De Vries 2001; Neumann et al. 2011) to determine a dominance score for each individual within each observation period. The Elo rating system assigns each individual an equal, arbitrary starting value and then uses the results of sequential pairwise interactions, each with a winner and loser, to adjust the individuals' values based on the discrepancy between the actual outcome of the interaction and that expected from previous interactions. Thus, an individual experiences a large gain in Elo rating if she wins unexpectedly (i.e. if she defeats an individual of higher prior Elo value than herself), but only a small gain if she was already expected to win the contest (i.e. if she already has a higher Elo value than the individual she defeats). Elo ratings are well-suited to my data, as they can accommodate repeated interactions between specific pairs of individuals.

I collated all observed behaviours for a given colony within each of three set time periods: the three days prior to manipulation, the three days immediately following manipulation (for all colonies), and days 10 through 12 following manipulation (for QR12 and C12 colonies only). I then generated Elo rankings for each individual during each period using the AniDom package in R (Farine & Sanchez-Tojar 2018) with an initial Elo value of 1000 and scaling constant $K = 100$. Elo rankings are sensitive to the order of interaction,

but each video represents only a small portion of the total number of interactions that may have occurred within a 3-day period and the order of in which interactions were observed may not have been meaningful. To remove potential bias, I therefore randomly re-ordered the list of contests collated within each 3-day period 1000 times and obtained an Elo ranking for each individual for each permutation; final Elo scores were then calculated as the mean score across all permutations for the given time period. Each time period (three days pre-removal, and days 1-3 and 10-12 post-removal) was treated as independent. In addition, I discarded dominance values for any individual that was observed on the nest carton for fewer than 30 minutes within the focal three-day time period rather than arbitrarily assign 'neutral' Elo scores to individuals that were under-observed.

2.2.6 Social network analyses

As a complement to individual-level measures of dominance, I used social network analyses to capture group-level consequences of queen removal. Network characteristics were calculated using the behavioural interactions for each colony for each three-day period, again treating each period as independent. Collated lists of behavioural interactions were converted to social networks in R using the *igraph* package (Csardi and Nepusz 2006) with each individual representing a node and each interaction representing an edge (connection) between nodes. For each network, I then generated two measures of social network structure. The *transitivity coefficient* (or *clustering coefficient*) of a network measures the global density of closed node triads, i.e. the proportion of instances in which, when an individual A has interacted with two other individuals B and C, those two individuals have also interacted with one another. Transitivity may be interpreted as a measure of the cohesiveness of the group (Croft et al. 2008). *Degree centrality* is the extent to which a particular node (individual) occupies a central location within a network (Croft et al. 2008). Variance in degree centrality indicates the level of social monopoly within a group, and high degree centrality variance indicates that one or a few individuals dominate the network relative to a larger number of poorly-connected individuals. If new queens establish themselves by directing frequent dominance behaviour towards their nestmates, then we should expect to observe an increase in the variance of degree centrality in colonies following queen removal.

2.2.7 Estimation of individual-level transition from worker to queen roles

In order to robustly assign queen identity, I fit a Bayesian logistic regression model with the R package *arm* (Gelman & Su 2018) using ovarian development indices and Elo ratings from queens and workers from queenright control colonies as the independent variables and caste as a binary response variable (**Figure 2.1B**). This allowed me to subsequently estimate the degree to which individuals on post-removal colonies phenotypically resembled workers or queens. Workers were coded with a value of 0 and queens were coded with a value of 1. I applied a flat Bayesian prior to the model output to account for the fact that phenotypic scores for workers and queens exhibit perfect separation on queenright colonies (Gelman et al. 2008), whereas my expectation is that workers transitioning to queens must necessarily pass through some intermediate stage in which their phenotype is wholly or partially intermediate between that of a 'normal' queen and 'normal' worker. For each QR3 and QR12 colony, I then estimated the 'queenness' of each individual by supplying the model with ovarian scores and dominance scores pertaining to the three days prior to colony termination. Using this method, each individual from each QR3 and QR12 colony was assigned a value from 0 to 1 indicating the degree of phenotypic identity to workers or queens from queenright colonies (0 = 100% similarity to worker phenotype; 1 = 100% similarity to queen phenotype), except where Elo ratings could not be established due to insufficient observational data.

2.2.8 Hypothesis testing

General linear mixed models (GLMMs) except those with queenness as the response were constructed in R using the *lmer* function, part of the package *lme4* (Bates et al. 2015). For individual-level analyses, response variables were queenness (Section 2.4.3), Elo rating, or ovarian development. Models with queenness as the response were constructed using the *metafor* R package (Viechtbauer 2010), which allowed me to weight individuals' queenness estimates by the inverse of the standard error of those estimates. For all individual-level models, fixed effects were wing length, age and pre-manipulation Elo rating and their pairwise interactions. The possibility that caste-biased traits such as ovarian development might be negatively correlated within nests (if, for example, replacement queens suppress their nestmates' phenotypic plasticity) led me to question the validity of including colony as a random effect in the individual-level models. To

account for this, I ran each individual-level model both with and without the random effect. In most cases the inclusion or exclusion of the random effect did not qualitatively affect the results—in the one case where it did (the result described in Supplementary Table S2.4), I report the results of the latter fixed-effect model alongside the mixed-effect model.

For group-level analyses, response variables were within-group mean antennation rate, antennation rate variance, network transitivity, degree centrality or mean time off-nest. Fixed effects were treatment group (Control vs Queen removal), and experimental stage (Pre-manipulation vs Post-manipulation days 1-3 vs Post-manipulation days 10-12), number of individuals in the colony, and the interaction between treatment and time period. Colony was included as a random factor in each model, to account for repeated measurements of the same colonies, which were assumed to be independent. For group-level analyses, response variables were transformed using Tukey's ladder of powers transformation ($\lambda = 0.450, 0.175, 0.925, 0.950$ and 0.775 respectively for mean, variance, transitivity, centrality and time off-nest transformations), and were thereafter confirmed to adhere adequately to the assumption of normality using the package *rcompanion* (Mangiafico 2018). In all cases, continuous variables were centred and scaled to facilitate model comparison.

2.3   Results

Of 76 colonies collected, nine were excluded due to colony failure: six did not produce the minimum required number of four adult workers within 21 days of collection, and in a further three the queen died before the emergence of the first brood. Additionally, I excluded 12 colonies that were heavily stylopised by strepsipteran flies (*Xenos vesparum*), defined as displaying >50% stylopisation among all offspring produced across the course of the experiment. Following these exclusions, final colony sample sizes for the experiment were n=55 colonies, with n=16 colonies each for treatments QR3, QR12 and QC3, and n=7 colonies for treatment QC12.

Approximately half of the remaining 55 colonies produced at least one stylopised individual (mean±SE 1.56±0.28 stylopised wasps/colony across all 55 colonies, representing 13.6±2.1% of workers in each colony). Stylopised individuals become asocial, fail to perform typical foraging and feeding behaviours, and eventually disperse from the nest (Hughes et al. 2004; Dapporto et al. 2007; Kathirithamby 2009; Beani et al.

2011; Geffre et al. 2017). Stylopised individuals do not typically engage in or respond to dominance interactions, and disperse within a few days of emergence (Hughes et al. 2004). Here, they comprised 12% (86/701) of experimental individuals but took part in a total of just 0.3% (56/1633) observed dominance interactions. Accordingly, I excluded stylopised individuals from further analysis.

A very small number of males emerged during the experiment (3/701 observed individuals). Since male *P. dominula* rarely engage in dominance behaviour (5/1633 observed dominance interactions in this experiment) and are known to disperse relatively quickly following eclosion in nature (Reeve 1991), I excluded males from all analyses.

Following these exclusions, behavioural and ovarian data for 55 queens and 557 workers from 55 colonies remained (mean±SE 12.2±0.66 wasps/colony; **Tables S2.2**-**S2.3**).

2.3.1 Hypothesis 1: Predictors of queen succession

Age was the sole significant predictive variable in each of the three complete models of individual phenotypes following queen removal, suggesting that age is a reliable predictor of both post-removal ovarian development (Est = 0.43, SE = 0.12, p = 0.0023) and post-removal dominance (Est. = 0.28, SE = 0.10, p = 0.008), and thus of post-removal queenness (Est. = 0.17, SE = 0.044, $p < 0.001$; **Table 2.2**). The slope of queenness on age was similar when models were constructed using data from QR3 colonies only (Est. = 0.18, SE = 0.15, p = 0.22) or from QR12 colonies only (Est. = 0.14, SE = 0.11, p = 0.21; **Table S2.4**).

	Queenness			Ovarian development			Elo rating at removal		
	Estimate	Std. Error	*p*-value	Estimate	Std. Error	*p*-value	Estimate	Std. Error	*p*-value
Intercept	**0.2335**	**0.0572**	**<0.001**	0.0440	0.1326	0.7480	-0.0671	0.1012	0.5092
Wing length	-0.0302	0.0374	0.4199	0.0392	0.1052	0.7102	-0.0767	0.1119	0.4951
Pre-removal dominance	0.0271	0.0276	0.3261	0.0458	0.0956	0.6330	0.1571	0.1033	0.1317
Age	**0.1693**	**0.0435**	**0.0001**	**0.4262**	**0.1228**	**0.0023**	**0.2815**	**0.1045**	**0.0084**
Wing length * Age	-0.0301	0.0278	0.2796	-0.0334	0.0941	0.7238	-0.1807	0.1001	0.0743
Wing length * Pre-removal dominance	-0.0197	0.0266	0.4582	0.1100	0.1017	0.2834	-0.0921	0.1140	0.4211
Pre-removal dominance * Age	-0.0392	0.0284	0.1673	0.0434	0.1028	0.6737	-0.0813	0.1141	0.4777

Table 2.2. Coefficient estimates with standard errors and p-values for three different linear models with queenness, ovarian development or dominance of individuals in post-removal colonies as the response variables, and nest ID as a random effect. Coefficient terms with p <0.05 are in bold.

Excluding two colonies which failed to produce any candidate queen replacements (**Section 2.3.3**), in 19/30 (63%) queen removal colonies the individual with the highest queenness was also the oldest individual (**Figure 2.2**). In a further 8/30 (27%) colonies

this individual was the second oldest, leaving three colonies in which neither of the two oldest individuals was the most queenlike. These proportions did not differ between QR3 and QR12 treatments ($\chi^2 = 0.89$, $p = 0.64$; **Table S2.5**). Thus, while age appears to act a strong predictor of succession, it does not perfectly explain variation between individuals' caste identities.

Neither body size (measured by wing length) nor pre-manipulation Elo rating emerged as significant terms in any of the models, indicating that queen succession in *P. dominula* is unlikely to be strongly dictated by physical strength. A possible exception appeared in the model with post-manipulation Elo rating as the response variable, in which there was a near-significant negative interaction between body size and age (Est. = -0.18, SE = 0.10, $p = 0.074$; **Table 2.2**). This might plausibly indicate that for particularly large individuals, age is a less important determinant of dominance following queen removal.

2.3.2 Hypothesis 2: Role of ovarian development in queen succession

Contrary to the hypothesis that post-eclosion queen succession in *Polistes* reflects inherent physiological constraints upon ovarian development in workers, when I included control colonies in my ovarian model, I found no evidence for a relationship between age and ovarian development on queenright colonies (Est. = 0.060, SE = 0.096, $p = 0.53$; **Figure 2.3**; **Table 2.3**). Only following queen loss did older individuals begin to display greater ovarian development than their younger sisters (Est. = 0.52, SE = 0.12, $p < 0.001$).

	Estimate	Std. Error	p-value
Intercept	**-0.2722**	**0.1050**	**0.0134**
Wing length	0.1092	0.0913	0.2327
Age	0.0598	0.0960	0.5342
Treatment	**0.4666**	**0.1409**	**0.0020**
Wing length * Age	-0.0013	0.0550	0.9813
Wing length * Treatment	-0.0485	0.1173	0.6793
Age * Treatment	**0.5164**	**0.1223**	**<0.001**

Table 2.3. Coefficient estimates with standard errors and p-values for a linear model with ovarian development as the response variable, and nest ID as a random effect. Coefficient terms with p <0.05 are in bold.
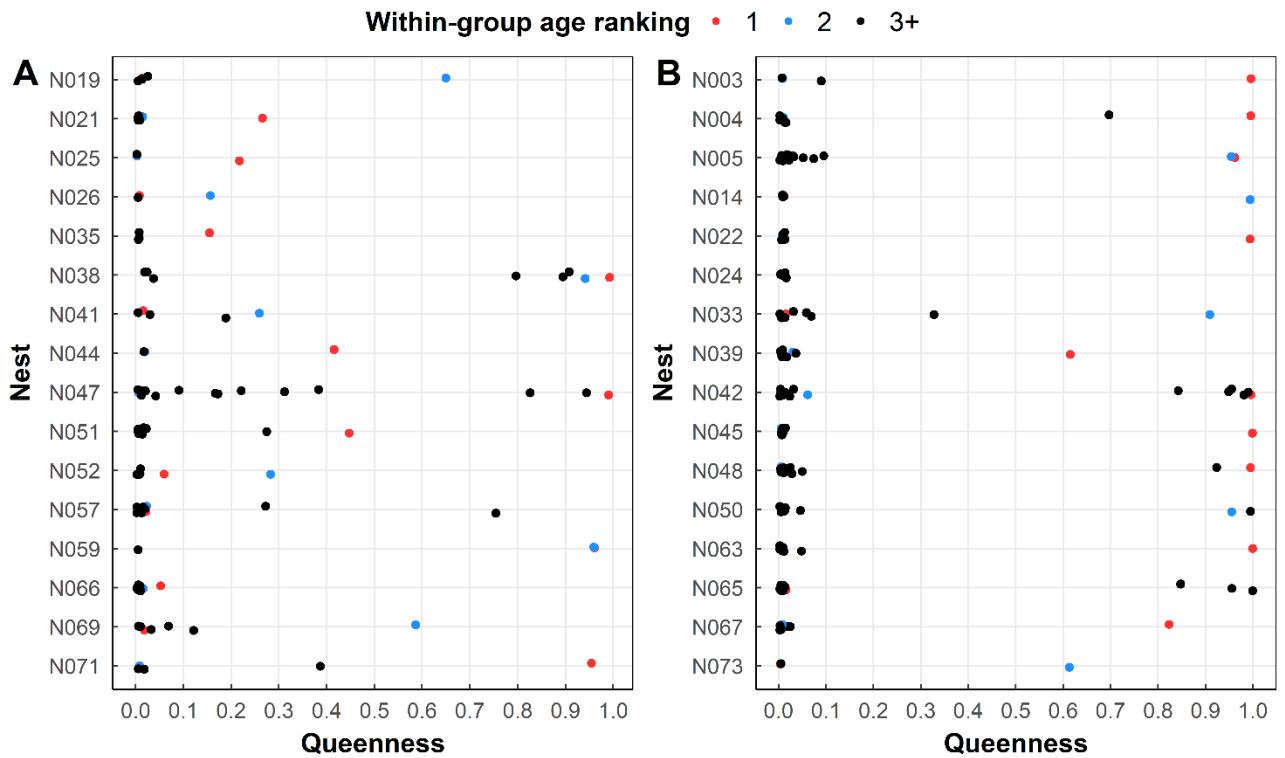
Figure 2.2. Queenness estimates for individuals from (A) QR3 and (B) QR12 nests. Coloured points represent the oldest (red) and second-oldest (blue) individuals from each nest.
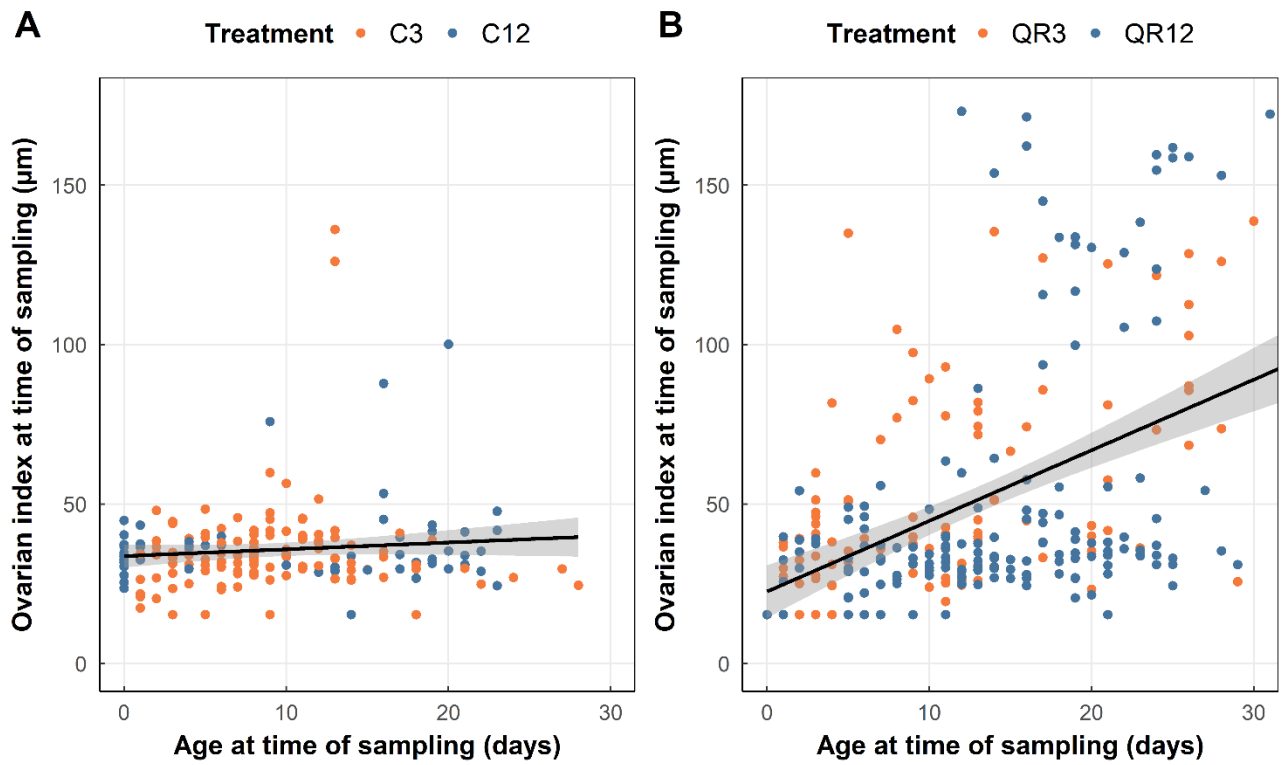
Figure 2.3. Ovarian index plotted against age for workers from (A) control and (B) queen removal colonies. Trend lines with 95% confidence intervals are shown in black.

2.3.3 Hypothesis 3: Group-level resilience to queen loss

I found little evidence that queen removal significantly altered the group-level social dynamics within colonies. Neither mean (Est. = 0.24, SE = 0.27, p = 0.37) nor variance (Est. = 0.12, SE = 0.11, p = 0.28) of antennation rates were significantly increased following queen removal, indicating that loss of the dominant does not result in a meaningful increase in the rate and variance of dominance behaviour (**Figure 2.4; Table 2.4**).
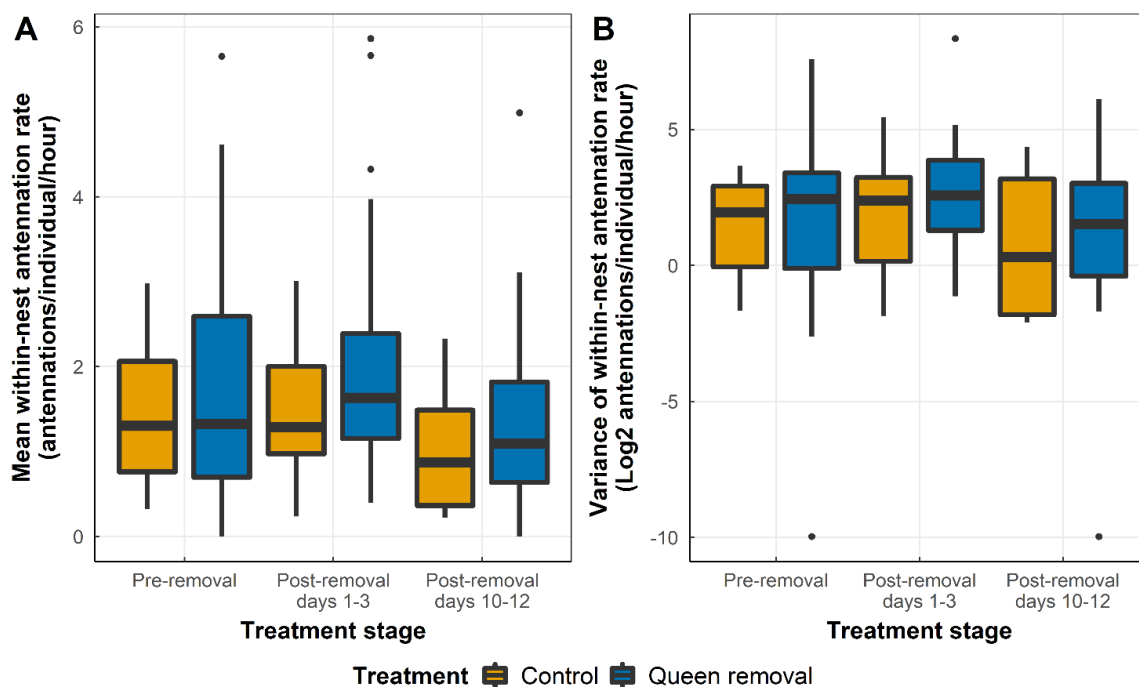
Figure 2.4. (A) Mean and (B) log2 variance of within-colony antennation rates. Neither mean nor variance of within-colony antennation rate vary significantly with stage or treatment.

	Antennation rate mean			Antennation rate variance		
	Estimate	Std. Error	*p*-value	Estimate	Std. Error	*p*-value
Intercept	**-1.1588**	**0.3005**	**<0.001**	**0.8120**	**0.1194**	**<0.001**
Post-manipulation days 1-3	0.1368	0.2426	0.5748	0.1082	0.1021	0.2934
Post-manipulation days 10-12	-0.1898	0.3783	0.6171	0.0331	0.1569	0.8334
Treatment	0.2444	0.2705	0.3687	0.1192	0.1086	0.2751
Colony size	**0.1598**	**0.0343**	**<0.001**	**0.0653**	**0.0135**	**<0.001**
Post-manipulation days 1-3 * Treatment	0.0623	0.3223	0.8473	-0.0546	0.1354	0.6880
Post-manipulation days 10-12 *Treatment	-0.2430	0.4563	0.5959	-0.2216	0.1897	0.2460

Table 2.4. Coefficient estimates with standard errors and p-values for two different mixed-effects linear models with mean and variance of within-colony antennation rate as the response variables. Coefficient terms with p <0.05 are in bold.

Both social network centrality (Est. = 0.49, SE = 0.28, p = 0.089) and clustering coefficient (Est. = 0.63, SE = 0.37, p = 0.091) exhibited borderline-significant increases in the three days following queen removal, but I otherwise found no evidence of group-level responses to queen loss (**Table 2.5**). By contrast, control colonies exhibited significantly decreased centrality (Est. = -0.76, SE = 0.33, p = 0.022) and increased mean time off-nest (Est. = 1.1, SE = 0.41, p = 0.0096) at days 10-12 following manipulation (**Table 2.5**). Network centrality (Est. = 0.20, SE = 0.028, p = <0.001) and mean (Est. = 0.16, SE = 0.034, p = <0.001) and variance (Est. = 0.065, SE = 0.014, $p < 0.001$) of antennation rate were all strongly positively correlated with colony size (**Table 2.5**).

	Network centrality			Network clustering			Mean time off-nest		
	Estimate	Std. Error	*p*-value	Estimate	Std. Error	*p*-value	Estimate	Std. Error	*p*-value
Intercept	**-1.5016**	**0.2456**	**<0.001**	-0.2801	0.3361	0.4072	0.0483	0.3296	0.8838
Post-manipulation days 1-3	-0.1614	0.2138	0.4533	0.0109	0.2740	0.9683	0.3043	0.2611	0.2484
Post-manipulation days 10-12	**-0.7566**	**0.3252**	**0.0223**	-0.4458	0.5480	0.4181	**1.0848**	**0.4091**	**0.0096**
Treatment	0.3405	0.2225	0.1291	0.3900	0.2927	0.1862	-0.2963	0.2959	0.3193
Colony size	**0.2042**	**0.0281**	**<0.001**	-0.0253	0.0385	0.5136	-0.0176	0.0377	0.6416
Post-manipulation days 1-3 * Treatment	0.4900	0.2840	0.0893	0.6284	0.3659	0.0913	-0.0890	0.3472	0.7984
Post-manipulation days 10-12 *Treatment	0.6061	0.3983	0.1319	0.8536	0.6187	0.1714	-0.6264	0.4930	0.2076

Table 2.5. Coefficient estimates with standard errors and p-values for three different mixed-effects linear models with group-level social network centrality, group-level social network clustering, and mean time off-nest as the response variables. Coefficient terms with p <0.05 are in bold.

One notable manner in which queen loss may have altered nests' social structures was in the production of multiple replacement queens. A large majority of individuals on post-removal colonies were assigned queenness values very close to zero, as expected if most individuals remain as un-reproductive, low-dominance workers. I thus considered any individual with a queenness estimate greater than 0.1 to be exhibiting significantly divergent caste expression relative to the normal worker profile; any such individual might represent a potential replacement queen. While the majority of nests produced either one (15/32 nests) or two (11/32) potential replacement queens, a small number (4/32) produced three or more replacements, and a further 2/32 failed entirely to produce a potential replacement queen. These numbers did not differ significantly between QR3 and QR12 conditions (χ^2 = 0.16, p = 0.98; **Table S2.6**).

2.3.4 Hypothesis 4: Within-group distribution of behavioural responses to queen removal

In accordance with the predictions of a convention-based succession model, the change in individuals' antennation rates between the three days preceding and the three days following queen removal was strongly predicted by age (Est. = 1.2, SE = 0.22, $p < 0.001$), although there was also a strongly significant negative interaction between the effects of body size and age (Est. = -0.69, SE = 0.22, p = 0.0018; **Table 2.6**). This negative interaction term indicates that age was a weaker predictor of antennation rate for larger individuals.

	Estimate	Std. Error	*p*-value
Intercept	0.3203	0.2304	0.1663
Wing length	0.2527	0.2560	0.3249
Age	**1.1874**	**0.2238**	**<0.001**
Wing length * Age	**-0.6888**	**0.2173**	**0.0018**

Table 2.6. Coefficient estimates with standard errors and p-values for a mixed-effects linear model with individual-level change in antennation rate following queen removal as the response variable, and nest ID as a random effect. Coefficient terms with p<0.05 are in bold.

The mean increase in dominance rates was close to zero for the youngest (latest-eclosing) individuals on each colony: an increase of 0.09 antennations/hour for individuals that were fifth or below in the order of eclosion vs an increase of 2.42 antennations/hour for individuals that eclosed first or second on their respective colonies, and an increase of 0.71 for those that eclosed third or fourth.

2.3.5 Hypothesis 5: Efficacy of gerontocracy in the presence of low age rank resolution

Age-based convention appears to act as an effective means of conflict mitigation during queen succession in *P. dominula.* Despite this, I found no evidence that a lack of resolution within a colony's age hierarchy results in the failure of this conflict-resolution mechanism, i.e. a reversion to a contest-based system. The age gap between the two oldest workers in a colony was not a significant predictor of the increase in antennation rate in that colony in the three days immediately following queen removal (Est. = 0.50, SE = 0.40, p = 0.22; **Table 2.7**), suggesting that colonies with multiple oldest individuals of similar ages are nonetheless able to transition to a successor queen without a significant increase in intra-group conflict. I also found no correlation between the difference in

antennation rates of the two most dominant workers on a colony prior to queen removal and increases in colony-wide antennation rate on that same colony following queen removal (Est. = -0.22, SE = 0.43, p = 0.62; **Table S2.7**).

	Estimate	Std. Error	*p*-value
Intercept	0.1407	0.3279	0.6701
Treatment	0.2890	0.4337	0.5089
Antennation gap	-0.1364	0.2771	0.6251
Treatment * Antennation gap	0.4977	0.4006	0.2213

Table 2.7. Coefficient estimates with standard errors and p-values for a mixed-effects linear model with group-level change in mean antennation rate in the three days before and after queen removal as the response variable.

2.4 Results

2.4.1 The role of age in *P. dominula* succession

Age acted as a strong predictor for queen succession in my focal colonies, both strongly predicting individuals' chances of inheriting the queen position and seemingly moderating social disruption following queen loss. Despite this, I did not find evidence that groups were less able to mitigate within-colony conflict when the strength of this cue was relatively weak. Groups with a poorly-resolved age ranking did not experience greater social disruption than ones in which the age gap between the oldest individuals was large. Moreover, the most queenlike individual on a given nest was not always the oldest, although in 90% of nests she was one of the two oldest individuals. Colonies were only maintained for a relatively short period of time in order to minimize any behavioural effects of the laboratory environment (Jandt et al. 2015), but as a result the age difference between individuals on any given colony was small. The influence of the gerontocratic convention identified here might be stronger on colonies that have undergone a larger number of brood cycles, with a larger range of ages therefore represented. My results nonetheless indicate that age is successfully employed as a mediator of succession conflict in *P. dominula*.

Intriguingly, I identified a strongly negative interaction between age and size in predicting the change in individuals' antennation rate following queen removal. Age was a weaker predictor of antennation rate increase for larger individuals, which might indicate that larger individuals were attempting to compete for the dominant position even while young. Despite this there was no meaningful influence of size upon ovarian development

or Elo rating. This suggests that the gerontocratic convention operates effectively even in the face of physically large competitors, at least for the early-season colonies described here. One possibility is that such 'queue jumpers' might become more aggressive in their efforts later in the season when the indirect fitness benefits of cooperation have declined, as occurs in certain other reproductively plastic social insects such as bumble bees (Rottler-Hoermann et al. 2016).

My results are consistent with the established notion that social hierarchies on queenright *P. dominula* colonies are age-based, with the oldest individuals being the most dominant (Pardi 1948; Pratte et al. 1990). The physiological bases of this age-based system have remained elusive, however. One established hypothesis is that gerontocracy in *P. dominula* reflects a physiological constraint, i.e. that younger individuals might have underdeveloped ovaries and so as a result are poorly positioned to transition to a reproductive role (Pardi 1948). Contrary to this, my data show that the positive relationship between workers' age and ovarian development is present only following queen removal. This is what we would expect to observe if gerontocracy is antecedent to, rather than a consequence of, variation in ovarian development. Moreover, while *P. dominula* expresses a positive relationship between age and reproductive dominance, this is not the case in all *Polistes* species: several species are thought to express the reverse relationship, with younger individuals more likely to inherit the queen role, possibly due to ecological variables that affect future fitness payoffs (Tsuji & Tsuji 2005). The existence of age-based conventions acting in opposite directions in different *Polistes* species seems incongruent with the idea that ovarian development is limited by age in this genus, but is unsurprising if age acts as a predominantly arbitrary signifier of dominance.

2.4.2 Explaining a rapid and pacific succession process

The very low level of group-level perturbation I observed following queen removal contrasts with results from Strassmann et al. (2014), who found significant increases in within-group conflict immediately following queen loss in mature nests outside the laboratory context. I consider three possible explanations for the discrepancy between these results and my own. First, colony size is known to predict conflict between dominant and subordinate individuals in *Polistes (*Cant et al. 2006), and thus it may be that the early-stage nests I observed were too small to merit conflict over succession. Second, it is possible that some aspect of the lab context, such as *ad libitum* access to

food, reduced the propensity of individuals to engage in conflict. Finally, the fact that I enforced a single-foundress context upon my colonies in question may have been a factor. From the perspective of an early-emerging worker, the absence of subordinate foundresses on the nest may act as an indicator of particularly high within-nest relatedness, since in a monogynous colony the workers are guaranteed to be fully matrilineally related.

The speed with which colonies appear to have generated replacement queens and the low-conflict nature of this transition indicate a remarkably robust and efficient conflict resolution mechanism operating within *P. dominula* colonies. Such robust mechanisms for the mitigation and resolution of intragroup conflict are essential components in the long-term maintenance of social groups that retain high degrees of reproductive plasticity (Aureli & de Waal 2000). Despite this, the nature of these mechanisms has been difficult to elucidate, especially outside of the non-human primates (Aureli et al. 2002). It has proven particularly challenging to separate different aspects of the phenotype in order to identify the specific cues that matter for reproductive succession in complex vertebrate societies within which measuring behaviour and physiology is difficult and time consuming. My ability to generate detailed phenotypic data for a large number of individuals over a short period of time was key in revealing the capacity for gerontocracy to act as a robust conflict-resolution system in a society with very high reproductive skew. Though labour-intensive to produce even in an invertebrate system, such in-depth data will be necessary to advance our understanding of social conflict resolution.

2.4.3 The role of seasonality

Age-based dominance structures appear to be common in polistine wasps, but the direction of these age conventions varies heavily: younger wasps tend to more dominant in tropical species, with the trend generally reversed in temperate species (Tsuji & Tsuji 2005). Because younger workers tend to be larger than their older sisters and are less likely to have experienced injury (Pardi 1949; Reeve 1991), and are therefore expected to win outright physical contests, the association between reverse gerontocracy and aseasonality may indicate that queen succession is more likely to be conflict-determined in tropical species.

Why should seasonality favour the evolution of alternative, non-aggressive succession mechanisms? *P. dominula* nests invariably decline over winter, and seasonality imposes

strong time constraints upon colony productivity as a result: the reproductive role holds no value if a queen is not able to lay eggs in time for them to emerge as reproductives. These constraints reduce the potential direct fitness benefits of inheriting the reproductive role within a colony, and also increase the relative costs to colony productivity that a protracted queen succession process would impose. If the loss in indirect fitness that competing workers experience as a result of disruption to the colony's productivity outweighs the expected direct fitness benefits of competing, then selection may favour a succession mechanism that facilitates a rapid return to normal colony functioning. This might explain the value of gerontocracy as a mediator of queen replacement: even if age confers no inherent benefits to replacement queens (as appears to be the case from my results), it may function as an honest signal that allows a new queen to be chosen rapidly and without costly conflict.

2.5   Conclusions

Understanding how social cohesion is maintained in the face of the within-group conflicts that exist in non-superorganismal social insect societies is an important step towards a fuller picture of the evolution of insect sociality. In this chapter, I have shown that age strongly predicts reproductive succession following the loss of a queen in *P. dominula*, despite the fact that older individuals do not appear to be inherently advantaged in terms of ovarian development or behavioural dominance prior to queen loss. Instead, age appears to act as a simple conventional mechanism that allows colonies to replace a lost queen rapidly and with little disruption to within-group social functioning. My results indicate that, even in the absence of obligate worker subfertility, convention-based mechanisms can successfully suppress within-group conflict. However, I speculate that the evolution of gerontocracy in *P. dominula* may partially reflect the constraints of seasonality, which impose high group-level costs on prolonged succession mechanisms. In aseasonal environments, where the colony cycle is indefinite and the benefits of succession are therefore proportionally greater, conflict-mediated succession may be the norm.

Chapter 3 – The molecular basis of socially-mediated phenotypic plasticity in *Polistes dominula*

The results presented in this chapter have been accepted for publication as: Taylor B. A., Cini A., Wyatt. C. D. R., Reuter M., & Sumner S. (in press). The molecular basis of socially-mediated phenotypic plasticity in a eusocial paper wasp. *Nature Communications*.

3.1 Introduction

3.1.1 Mechanisms of plasticity in social insects

The retention of reproductive plasticity is thought to be a major source of potential within-group conflict in less derived social insect taxa (Boomsma 2009; Boomsma & Gawne 2018). Social insects in taxa that retain high degrees of reproductive plasticity exhibit an unusual quality, namely that they are fully capable of switching between reproductive and non-reproductive roles but in practice rarely do so. In this respect, social insect taxa exhibit a similarity to other species with strongly distinct reproductive and non-reproductive cycles, such as progressively provisioning solitary wasps (Field 1992; Kelstrup et al. 2018). Which mechanisms underlie the stable but plastic division of labour exhibited by such species is a major question in the field of social insect research (Ferreira et al. 2013; Patalano et al. 2015; Kennedy et al. 2017).

The drive to explain the mechanisms of social insect plasticity is at least partially motivated by the fact that it is precisely the loss of this plasticity that is thought to underlie the ecological success of the most derived social insect species (as idea that I discuss further in Chapter 5 of this thesis). Loss of reproductive plasticity aligns the fitness interests of queens and workers, removing a source of within-group conflict that might otherwise place a burden on colony functioning (Boomsma & Gawne 2018). However, the question of how such 'superorganisms' have evolved can only be answered with an understanding of the mechanisms by which the expression of plastic castes is achieved in those social insects that do possess reproductive plasticity. A review of several prominent hypotheses regarding the evolution of social insect castes and the molecular basis of caste is given in Chapter 1 of this thesis.

3.1.2 Plasticity, gene expression and the phenotypic gambit

The relationship between gene expression and external phenotype is complex and unresolved. Much research in behavioural and evolutionary ecology is based on the implicit assumption that phenotypic traits can be modelled as though they directly reflect gene expression patterns, and that evolutionary trajectories can therefore be studied while remaining agnostic with regard to the underlying molecular mechanisms (Hadfield et al. 2007; Rittschof & Robinson 2014; Rubin 2016). This 'phenotypic gambit' has proven a useful rule of thumb, permitting the establishment of a rich body of literature surrounding the evolution of complex traits, despite a lack of data relating to the genetic basis of these traits (e.g. Réale et al. 2010; Chapman et al. 2011; Fowler-Finn & Rodríguez 2012). In the past decade, however, advances in the affordability of 'omic' data and availability of powerful bioinformatic methods have greatly enhanced our ability to assess the assumptions made by the phenotypic gambit (Rittschof & Robinson 2014; Heyes 2016). The time is right to disentangle the molecular foundations of complex phenotypic traits.

Phenotypic plasticity, the ability of an individual to effect phenotypic changes in response to external cues, is an ideal phenomenon with which to study the relationship between gene expression and phenotype because it involves the production of multiple phenotypes without gene sequence changes. Of particular value are species in which adult individuals can be experimentally induced to transition between distinct, measurable phenotypes. By comparing the gene expression profiles of groups of individuals that differ in their phenotypes as a result of plasticity rather than as a result of genetic differences, it is possible to isolate phenotypic effects of gene expression. Using this approach, significant progress has been made in unravelling the molecular underpinnings of sequential sex changes in hermaphroditic fish (Horiguchi et al. 2013; Casas et al. 2016; Todd et al. 2019), the distinct gregarious social phenotype of desert locusts (Cullen et al. 2017; Lo et al. 2018), and the reproductive castes of social insects (Simola et al. 2016; Gaspocic et al. 2017; Libbrecht et al. 2018; Rehan et al. 2018; Shell & Rehan 2018). Such studies typically rely on comparisons between groups of individuals with well-differentiated phenotypes, however. As a result, little is known about more subtle effects during the transition from one morph to the other, and the relationship between expression patterns and phenotypic traits at the individual level.

The reproductive castes found in the colonies of social insects provide excellent model systems for determining the extent to which fine-scale changes in phenotype are reflected

at the molecular level. Workers in many species can be experimentally induced to transition to a reproductive role in response to the removal of a colony's queen (e.g. Strassmann et al. 2004; Tibbetts & Huang 2010) or as a result of exposure to varying levels of brood (e.g. Chandra et al. 2018; Libbrecht et al. 2018), allowing changes in the behavioural, physiological and molecular traits that define caste identity to be tracked. An additional benefit—and challenge—of studying social insect colonies is that they involve complex social structures. Such interactions can be hard to study, but offer the opportunity to assess the effects of social interactions upon phenotypes and transcriptomes.

3.1.3 Support Vector Classification

Support vector classification is a powerful tool with which to transform complex patterns in multidimensional data into a continuous classification score, allowing the detection of subtle, widespread signals of differential expression between phenotypic states that are likely to be missed in conventional differential expression analyses. Support vector classification operates by taking multidimensional data (e.g. the expression levels of many different genes) pertaining to two or more classes (e.g. queens and workers) and identifying a kernel function that will transform the data such that the classes can be linearly separated. The choice of kernel function is optimized by performing a 'grid search' in which each combination of parameters within the kernel function is tested across a range of values in order to identify the set of parameters that best allow the data to be linearly separated between classes. Although support vector machines (SVMs) are mathematically more complex than e.g. logistic regression models, at its core the outcome of support vector classification is similar to that of logistic regression, namely the generation of a model that allows novel data to be placed into different classification groups.

Support vector machines (SVMs) have become a key tool in the early identification of phenotypically indistinguishable cancer subtypes (Segal et al. 2003; Abeel et al. 2010; Huang et al. 2018) and their potential value has recently been demonstrated in animal behaviour studies: Chakravarty et al. (2019), for example, show that an SVM trained using accelerometer data can reliably classify the behaviours of wild Kalahari meerkats. Given their ability to detect subtle changes in patterns of high-dimensional data, SVMs should be ideally suited to quantify gene expression variation across the spectrum between differentiated worker and queen roles.

3.1.4 Chapter overview

In this chapter, I apply an SVM approach to analyse brain transcriptomes from 96 of the individuals for which behavioural and ovarian data were presented in Chapter 2, including queens and workers from stable colonies and individuals from colonies that had their queens experimentally removed. Combining this approach with standard differential expression and gene co-expression analyses, I show that brain gene expression responses to queen removal in *P. dominula* include a colony-wide response that does not match that observed at the phenotypic level. My results indicate that gene expression in *P. dominula* colonies reflects both a generalized response to queen loss that is seemingly independent of phenotype, and a phenotype-specific response that tracks individuals' expression of plastic phenotypic changes. This study provides a comprehensive analysis of the ways that plastic phenotypes are reflected at the transcriptomic level; my results expose the complexity of the relationship between individual-level gene expression and individuals' outward phenotypes.

3.2   Methodology

3.2.1 Sample collection

Full details of the methodology by which individuals were collected, reared, and sampled are provided in Chapter 2. From among the focal individuals used in that experiment, I randomly selected 29 queens and 20 control workers for sequencing. I additionally selected 65 individuals from queen removal nests for sequencing, using stratified random sampling to cover a full range of values of ovarian development, Elo rating and queenness. A number of these samples failed at the point of transcriptome sequencing (next section), leaving me with final sample sizes of 12 control workers, 58 queen removal individuals, and 26 queens.

3.2.2 Gene expression quantification

Brain tissue was extracted from the heads of individual samples and RNA was extracted using the RNeasy Mini Kit (Qiagen) according to manufacturer's instructions. Library preparation was performed by Novogene Co. followed by sequencing on an Illumina HiSeq 2000 platform with 150-base pair paired-end reads to a depth of 10 million reads/sample; this relatively shallow sequencing depth was informed by the power analyses described in Appendix A. Reads were filtered with SortMeRNA (Kopylova et al. 2012) using default options to remove ribosomal sequences. Trimmomatic (Bolger et

al. 2014) was then used to perform quality trimming. First, I trimmed adapter sequences and leading and trailing bases with low phred scores (<3). I then used the MAXINFO option with target length 36 and strictness 0.7 to trim low-quality sequences from the remaining reads. Reads were next mapped to 11313 transcripts from the *P. dominula* genome annotation 1.0 (Standage et al. 2016) using STAR (Dobin et al. 2013) with default options. All 101 samples produced >85% uniquely mapped reads. Reads were assembled into transcripts using StringTie2 (Kovaka et al. 2019) before being passed on as raw counts to downstream analyses. Prior to downstream analysis, transcripts were filtered to remove any gene which did not have >20 counts across all samples in at least one of the experimental groups (queens, control workers, and day three and day twelve post-manipulation workers). Following this filtering, 10734/11313 (94.9%) of transcripts remained.

3.2.3 Support vector machine classification

Support vector classification was performed in R using the e1071 package (Meyer et al. 2017) using the gene expression profiles for all available queens (coded with a value of 1; n=26) and control workers (coded with a value of 0; n=12). Classifiers were assessed via their 3-fold cross-validation error rates; classifiers with lower classification errors were considered to be superior. Initially, I tested classifiers using a variety of kernel functions—radial, linear, sigmoid and polynomial—combined with grid searches across a wide range of parameters for each kernel (kernel parameters vary depending on the form of the kernel, but always include a cost parameter C which determines how strongly misclassifications are penalized). The kernel function that produced the lowest error rate was the radial function, which gives the distance between two samples x and y as:

$$\exp(-\boldsymbol{\gamma}|\boldsymbol{x} - \boldsymbol{y}|^2)$$

A radial kernel with $\gamma=10^{-6}$ and cost parameter $C=2^5$ was found to produce the lowest error rate of all combinations of parameters, so all subsequent classifiers were fit using this kernel and a more focused grid search in a parameter space of $2^4<C<2^6$ and $10^{-5}<\gamma<10^{-7}$ in order to minimize the processing power necessary to perform feature selection. For feature selection, I took the classifier fitted with all genes, and iteratively performed the following process: (1) The 3-fold cross-validation error of the model was calculated twenty times using randomly-assigned bins, and the mean of the resulting errors was recorded as the true validation error of the classifier; (2) The feature weights

of all genes were calculated by taking the matrix product of that classifier's coefficients with its support vectors; (3) The gene with the smallest absolute weight in the model was dropped; (4) A new classifier was calculated using the remaining set of genes. This process was repeated until just 100 genes remained, and the optimal support vector classifier was then taken as that for which the cross-validation error reached its minimum.

3.2.4 Differential expression analysis

Differential expression analyses were performed in R using the DESeq2 package (Love et al. 2014). DESeq2 was run on all groups and contrasts were then calculated for each pair of groups. Unless otherwise stated, differential expression was calculated relative to a baseline fold change of 1.5, i.e. p-values refer to the probability that absolute change between two groups was greater than 50%. Genes were considered differentially expressed between conditions if $p<0.05$ after false discovery rate correction according to the Benjamini-Hochberg procedure.

3.2.5 Gene co-expression network analysis

Weighted gene co-expression network analysis was performed in R using the WGCNA package (Langfelder & Horvath 2008). As WGCNA is particularly sensitive to genes with low expression, data were first subjected to a second round of filtering in which genes that had <10 reads in >90% of samples were removed, as recommended by the package authors. This second round of filtering removed an additional 1631 genes, leaving a total of 9103 genes. Counts were then subjected to a variance-stabilizing transformation prior to further analysis. Consensus gene modules across all samples were then constructed using a soft-threshold power of 9. Initially, 26 gene modules were identified. Modules whose eigengene correlation was >0.75 were subsequently merged, after which 22 consensus modules remained. Finally, the Pearson correlation of each module with each phenotypic trait within each group (queens, control workers and individuals from queen removal nests) was calculated and subjected to Benjamini-Hochberg FDR correction. Network summary measures and gene dendrograms for WGCNA are provided in **Figures S3.1-3.2.**

3.2.6 Gene ontology (GO) enrichment analysis

In order to perform GO enrichment analysis, I first used OrthoFinder (Emms & Kelly 2019) to identify orthologues for each *P. dominula* gene in *D. melanogaster*, a model species for which GO annotations are much more complete. GO annotations for each *D.*

melanogaster gene were acquired from BioMart (Smedley et al. 2015) and each *P. dominula* gene was then assigned GO terms permissively, i.e. a given *P. dominula* gene was assigned a GO term if that term appeared as an annotation to any of its orthologues. 6659/10734 (62.0%) of genes possessed at least one orthologue in *D. melanogaster*. GO enrichment analysis was then performed in R via the topGO package (Alexa & Rahnenfuhrer 2009) using TopGO's weight01 algorithm and Fisher's exact test to identify GO terms that were significantly overrepresented ($p<0.01$) in a focal set of genes against a background consisting of all genes that appeared in the relevant analysis. For plotting purposes, the semantic similarity of GO terms was determined using GOSemSim (Yu et al. 2020).

3.3 Results

3.3.1 Support vector classification reveals consistent patterns of caste gene expression differentiation involving many genes

Support vector machines operate in a similar fashion to multivariate linear regression models, estimating the relationship between a response variable (here, known caste identities coded as worker = 0 and queen = 1) and one or more independent variables (here, expression profile) in a training data set. The model derived in this way is then applied to the query data set to derive predictions of the response. In contrast to standard linear models, SVM models project input data into a higher-dimensional space, thereby making it possible to fit a linear relationship to what would otherwise be non-linear data (Noble 2006). Using this approach allowed me to reduce the brain gene expression data I had generated down to a single dimension of predicted caste identity (the classifier variable), analogous to the Bayesian logistic regression model I used in Chapter 2 to condense ovarian and behavioural data into a unidimensional metric of phenotypic caste ('queenness').

I trained an SVM using whole transcriptome data from 26 queens and 12 workers from stable, queenright colonies. A full model, based on all 10734 genes annotated in the experiment, achieved a root mean squared validation error of 0.065 in three-fold cross-validation, i.e. a model trained on a random subset of two thirds of workers and queens classified the remaining third within 25.5% of their true values (0 and 1 for workers and queens, respectively). This error rate suggests reliable classification, because it indicates

that any given control worker was likely to receive a classification that was closer to that of all other workers than it was to the classification of any given queen, and vice versa.

As many genes will not vary consistently in their expression between workers and queens, an SVM model fit to the entire transcriptome is likely to exhibit a significant degree of overfitting. In order to identify a minimal set of genes that were maximally predictive of caste identity, I applied a process of 'feature selection' in which uninformative genes were progressively dropped until an optimal model containing only caste-informative genes was achieved (**Figure 3.1; Table S3.1**). The model obtained in this way contained 1992 genes with a root mean squared classification error of 0.021, a substantial improvement over that achieved for the model containing all genes (**Table S3.2**). This model classified queens and workers very consistently, with strong separation of queens from workers (**Figure 3.2A**). Thirty-one gene ontology (GO) terms were significantly enriched among these 1992 genes, including a number of terms associated with translation such as *rRNA processing*, *tRNA aminoacylation*, and *ribosomal large subunit biogenesis* (**Figure S3.3**; **Table S3.3**).

When I applied a more standard differential expression approach based on DESeq2 analysis and a 1.5 fold-change threshold to the same set of queens and workers, I identified just 81 differentially expressed genes (with no associated GO terms), a number that is typical for similar analyses in *Polistes* (e.g. Toth et al. 2014; Patalano et al. 2015; Geffre et al. 2017). Of these genes, 77/81 (95%) were present in the larger SVM set (**Figure S3.4**; **Table S3.4**), suggesting that the SVM captures the information contained in the set of highly caste-differentiated genes identified using this standard method. Yet, an SVM trained using just these 81 differentially-expressed genes exhibited a cross-validation error rate of 0.065, significantly higher than that of the optimized model (0.021) and no better than the original un-optimized model containing all expressed genes. Thus, the picture of caste differentiation provided by standard differential expression analysis appears to miss a great number of subtle differences in individuals' gene expression profiles that contribute to caste differentiation.

The small set of differentially-expressed genes that I identified using DESeq2 in the above analysis might conceivably reflect the conventional fold change threshold that was applied as part of the approach. Indeed, when I omitted this cut-off, I found 2438 differentially-expressed genes, on the same order of magnitude as that found via SVM feature selection. This set of genes overlapped significantly with those identified by

feature selection (Jaccard index = 0.44; one-sided hypergeometric overlap $p < 0.001$).
However, an SVM model trained using the differentially-expressed genes exhibited a root
mean squared error rate of 0.030, and thus still performed substantially less well than the
model using the genes identified using feature selection (0.021).
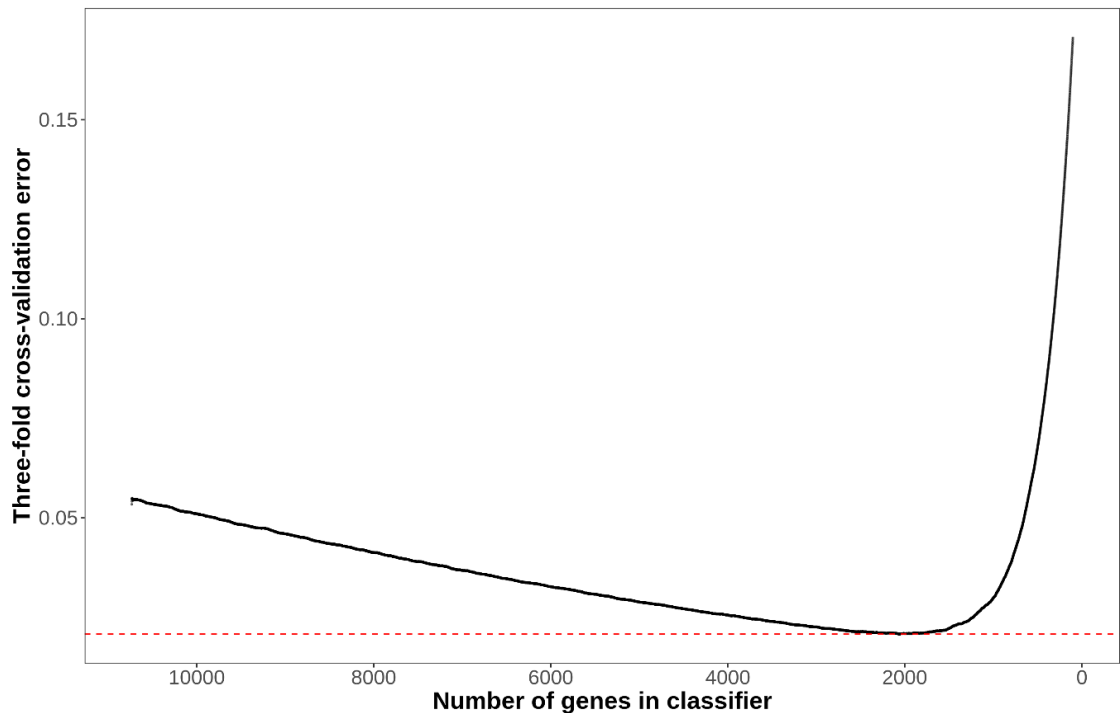
Figure 3.1. Root mean squared three-fold cross-validation error of SVMs trained using an
iteratively decreasing number of genes, showing a characteristic 'hockey stick' shape. Initially,
removing genes reduces error rate as a result of reduced overfitting, but once informative genes
start being removed, error rate increases exponentially. In each iteration, the lowest-weight gene
of the previous model was removed and a new model was trained with the remaining genes. For
each model, three-fold cross validation error was taken as the mean of 20 calculations using
randomly-selected validation bins. The error shown in figure is a moving average with a window
size of 50. Red dashed line shows the minimum validation error achieved.
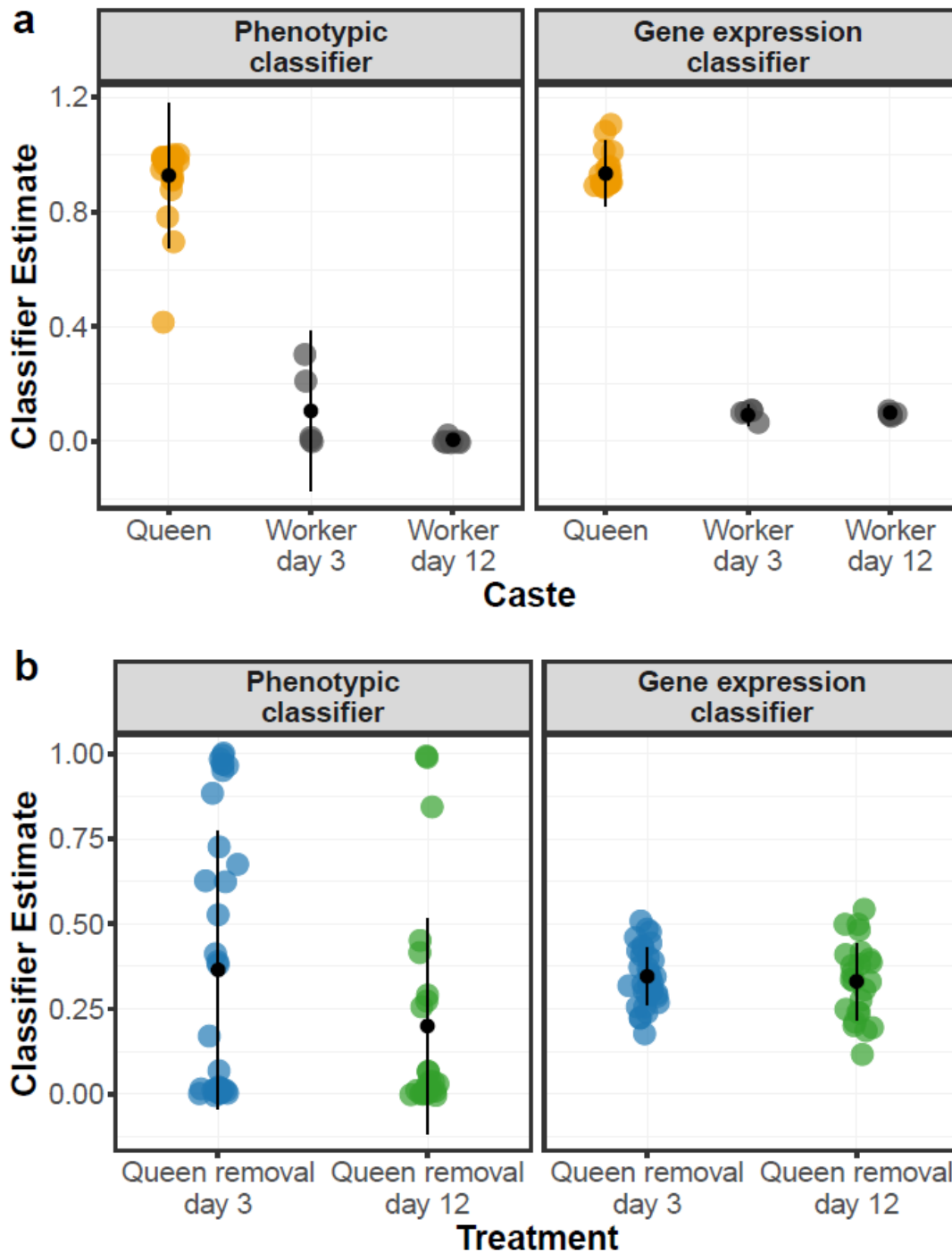
Figure 3.2. Classifier estimates generated by a Bayesian logistic regression model using ovarian and dominance data (left) and an SVM model using 1992 caste-informative genes (right) for (a) queens (orange; n = 26 biologically independent samples) and workers (black; n = 12 biologically independent samples) from control colonies and (b) individuals from experimental colonies either 3 days (blue; n = 24 biologically independent samples) or 12 days (green; n = 34 biologically independent samples) following queen removal. Mean and standard deviation for each group are shown in black. 0 = complete similarity to control workers; 1 = complete similarity to control queens.

3.3.2 Colony-wide brain gene expression responses to queen removal

Following the loss of a queen from a *Polistes* colony, typically one or a few individuals undergo a phenotypic transition to become a replacement queen while the rest of the colony members remain workers (Chapter 2; Miyano 1991; Dapporto et al. 2005). To capture this transition at the transcriptional level, I analysed individuals' gene expression profiles at three days after queen removal, when queen replacement is ongoing (n=24), and at twelve days after queen removal, when succession is largely settled at the phenotypic level (Strassmann et al. 2004; n=34). Individuals for sequencing were selected to cover a wide range of phenotypes, including those that remained entirely worker-like, those that had transitioned to highly queen-like phenotypes, and those with intermediate phenotypes at the time of sampling (**Figure S3.5**).

In order to assess changes in individuals' caste-specific gene expression profiles following queen removal, I applied the optimized SVM model described above to the gene expression profiles of these 62 individuals. Doing so generated an SVM classification for each individual that describes the degree to which its gene expression corresponds to the worker state (classifier=0) or queen state (classifier=1) as reflected in the expression profiles of the control workers and queens on which the SVM was trained.

Analysing shifts in SVM estimates of individual wasps then allowed me to assess the degree to which these concurred with the changes visible at the phenotypic level. Doing so, I found that the SVM estimates of individuals from post-removal nests were intermediate between those of queens and workers from control nests (**Figure 3.2B**), a finding which broadly concurs with the placement of these individuals according to principal component analysis (**Figure 3.3**). This result is surprising given that the majority of individuals on queen removal nests are phenotypically indistinguishable from workers on control nests in terms of ovarian development and behavioural dominance (Chapter 2). Thus, the large majority of individuals on queen removal nests exhibited perturbation of their caste-associated gene expression, even though only a few of these individuals exhibited responses to queen loss at the level of physiology or behaviour.
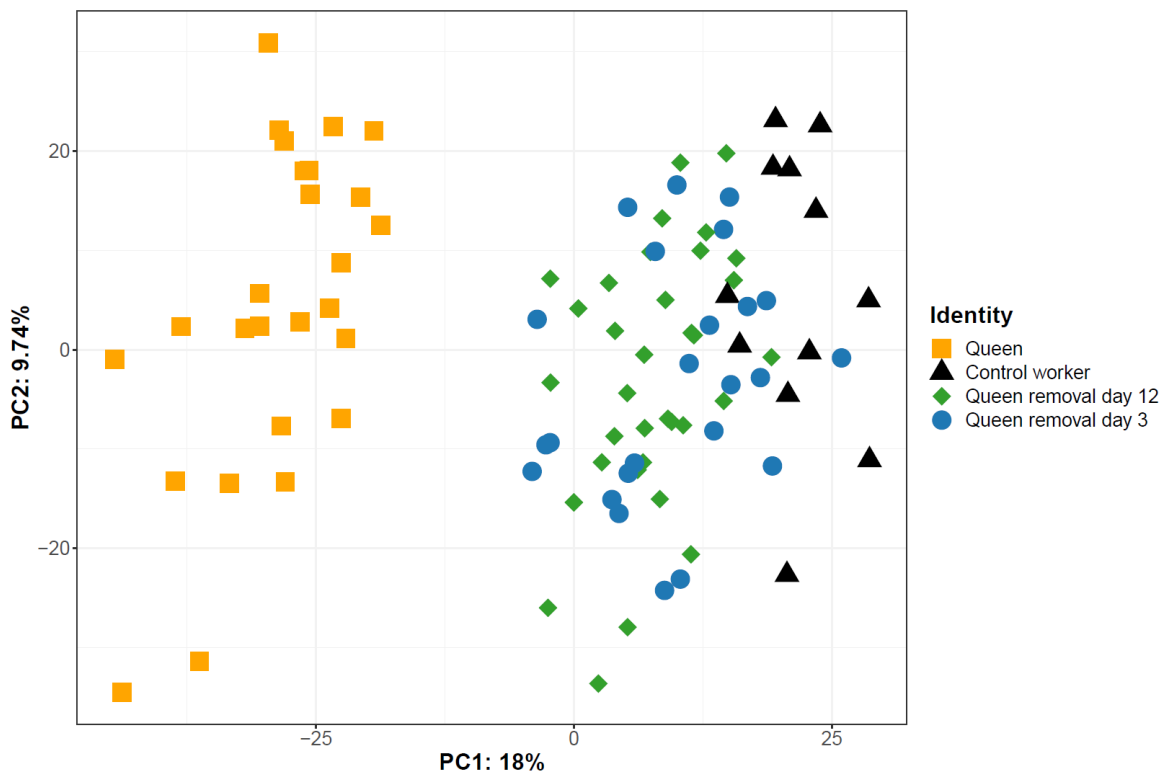
Figure 3.3. First two principal components generated by PC analysis of 1992 caste-associated genes identified by SVM classification.

Furthermore, I found no evidence that the degree of transcriptional perturbation declined over time following queen removal. SVM classification estimates of individuals from queen removal nests did not differ significantly between days three and twelve following queen removal (QR3 mean 0.331 ± 0.111; QR12 mean 0.345 ± 0.082; Wilcoxon W=369, p=0.55). The effects of queen removal on individuals' gene expression profiles therefore appear to be both widespread and persistent, affecting all individuals in a nest and lasting beyond the point at which a new queen has already become phenotypically established.

Interestingly, the strong colony-wide perturbation following queen loss that this SVM approach identifies would have been entirely missed using a standard differential expression approach: a DESeq2 analysis with a 1.5 fold-change threshold identified just five genes as differentially expressed between control workers and individuals from manipulated nests, with no associated GO enrichment (**Table S3.5**). Even omitting the fold-change threshold, the number of genes identified as being differentially expressed between control and queen removal workers only rose to 291, a small fraction of the 2438 genes identified as differentially expressed between control queens and workers.

3.3.3 Age and queenness explain variation in individual-level molecular responses to queen loss

While SVM classification indicates that queen removal causes colony-wide perturbation to brain expression profiles, classifier estimates varied substantially between individuals following queen removal, spanning a much greater range of values (0.116-0.540) than those of queens (0.900-1.100) or workers (0.057-0.100) from queenright colonies (**Figure 3.2**). To better understand this variation, I examined whether the classifier estimates for individuals from manipulated colonies were predicted by those individuals' phenotypic traits—specifically ovarian development, behavioural dominance, and age.

Our phenotypic measure of caste identity (queenness) was a significant predictor of expression-based SVM classifier estimates when fitted using a linear model (slope$\pm$SE = 0.0414$\pm$0.0114, $p = 6.0 \times 10^{-4}$; **Figure 3.4A**). This relationship did not change when using queenness values recalculated using a phenotypic training set consisting of only the individuals for which gene expression data were generated (slope$\pm$SE = 0.0407$\pm$0.0115, $p = 8.3 \times 10^{-4}$). The individual components of queenness, ovarian development and dominance were also significant or near-significant predictors of SVM classification individually (ovarian development: slope$\pm$SE = 0.0426$\pm$0.0113, $p = 4.0 \times 10^{-4}$; dominance: slope$\pm$SE = 0.0231$\pm$0.0122, $p = 0.06$). Notably, however, because phenotypic queenness was strongly correlated with age among post-removal individuals (cor = 0.4832, p = 8.0×10^{-5}; Chapter 2), age was an equally strong predictor of caste estimates when fitted in a separate linear model (slope$\pm$SE = 0.0516$\pm$0.0106, $p = 9.8 \times 10^{-6}$; **Figure 3.4B**). Thus, the significance of queenness as a predictor of caste estimates might have been an artefact of the fact that both are correlated with age. To test whether queenness had an effect over and above that accounted for by age, I calculated the residuals of phenotypic queenness on age and fitted the caste estimates against these. These residuals were not significantly predictive of SVM classification (slope$\pm$SE = 0.0202$\pm$0.0123, $p = 0.11$; **Figure 3.4C**), nor were the residuals of ovarian development alone on age (slope$\pm$SE = 0.0223$\pm$0.0123, $p = 0.07$) or the residuals of dominance alone (slope$\pm$SE = 0.0114$\pm$0.0126, $p = 0.37$). Age is thus the strongest determinant both of individuals' caste phenotypes and of their caste-associated gene expression.
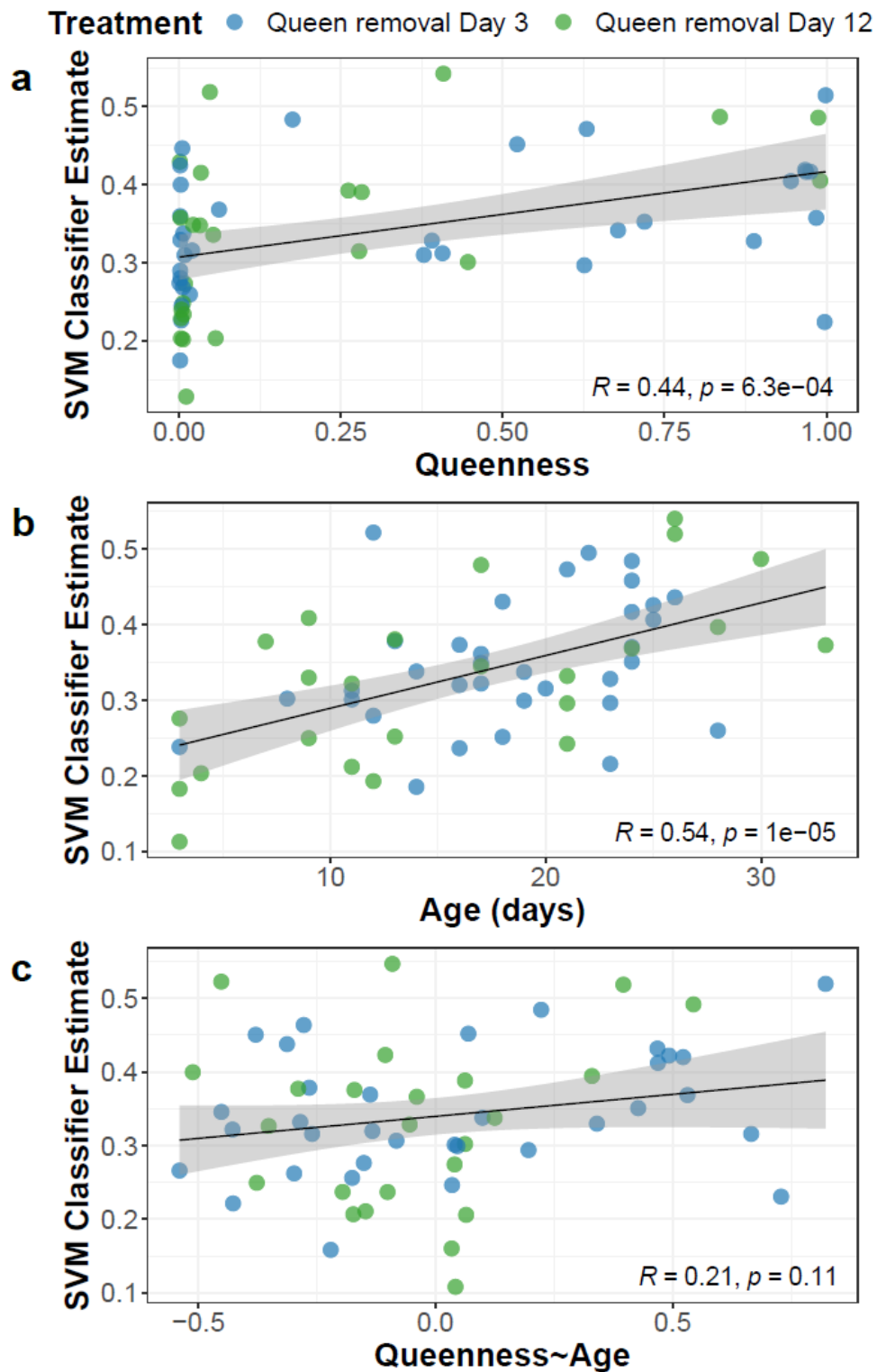
Figure 3.4. Scatterplots of SVM classifier estimates for individuals from post-removal colonies (n = 58 biologically independent samples) plotted against (a) phenotypic queenness, (b) age, and (c) the residuals of queenness on age. R = Pearson correlation; p = uncorrected significance value of linear models fitting SVM classification against the specified variables. Regression line and 95% confidence interval for each model are shown in black and grey respectively.

The control queens and workers that were used to train the SVM model differed not only in their caste, but also in their age—because foundress queens had overwintered, they were several months old at the time of sampling, while workers were weeks or days old. This fact, together with the observation that age predicts individuals' SVM classifications, could suggest that my SVM model is in fact a classifier for age rather than caste. Yet, three lines of evidence speak against this possibility. First, control workers and queen removal individuals were of comparable age but their SVM classifications were clearly distinct (but should have been very similar if based purely on age). Second, age was not predictive of variation in the SVM classifications among control workers (slope±SE = 0.0054±0.0033, $p = 0.14$). Third, the set of caste-informative genes identified by my SVM approach did not overlap significantly with a set of 625 age-biased genes identified using DESeq2 (**Table S3.6**; Jaccard index = 0.05; one-sided hypergeometric overlap $p = 0.12$). Taken together, these three results strongly suggest that the SVM model that I have produced classifies individuals by caste identity rather than by age.

In order to further discern the factors that shaped individuals' gene expression profiles following queen removal, I performed weighted gene co-expression network analysis (WGCNA) using the full set of annotated genes. I generated 22 consensus modules across all samples, and then determined which traits significantly predicted the expression of a given module within each group (**Figure 3.5**). Six modules exhibited a significant degree of overlap with the set of 1992 caste-predictive genes identified using SVM feature selection (**Table S3.7**), including three modules whose expression was significantly correlated with either phenotypic traits or SVM classification in at least one treatment group.

Module 2 consists of 102 genes and appears to be associated with many caste-related traits in the queen removal condition, being negatively correlated with queenness, age, SVM classification and (less strongly) with ovarian development among queen removal individuals. I also found weak evidence that this module is negatively associated with age among workers from control nests. Unexpectedly, despite the seeming importance of Module 2 in predicting caste identity among workers, not a single GO term was enriched among the genes in the module at p<0.01. A second module, Module 8, consists of 614 genes and has a strongly negative correlation with age among both control workers and queen removal individuals. This module is associated with 20 GO terms, including a

number of terms associated with molecular binding, such as protein binding, DNA binding and RNA binding (**Figure S3.6; Table S3.8**).

Finally, Module 11 consists of 519 genes and is notable in that its expression in individuals from queen removal colonies is positively associated with SVM classifications but not with phenotypic correlates of caste—significantly, chromatin remodelling is one of the 22 GO terms with which this module is associated (**Figure S3.7; Table S3.9**). Regulation of chromatin accessibility is one of several epigenetic processes that have been implicated in the control of caste expression in social insects (Simola et al. 2016; Wojciechowski et al. 2018; Duncan et al. 2020). The fact that Module 11 correlates with transcriptomic caste identity in post-removal nests might therefore suggest that disturbances to individuals' gene expression patterns following queen removal reflect some form of epigenetic reprogramming (Patalano et al. 2020). Genes that were more strongly related to phenotypic queenness also exhibited higher module significance (i.e. stronger correlation with the module eigengene) within Module 11 (cor = 0.6, $p = 1.0 \times 10^{-78}$), which was not the case for Module 2 (cor = 0.95, $p = 0.34$) or for Module 8 (cor = 0.077, $p = 0.057$). Taken together, these results strongly suggest a role for Module 11 in caste differentiation.
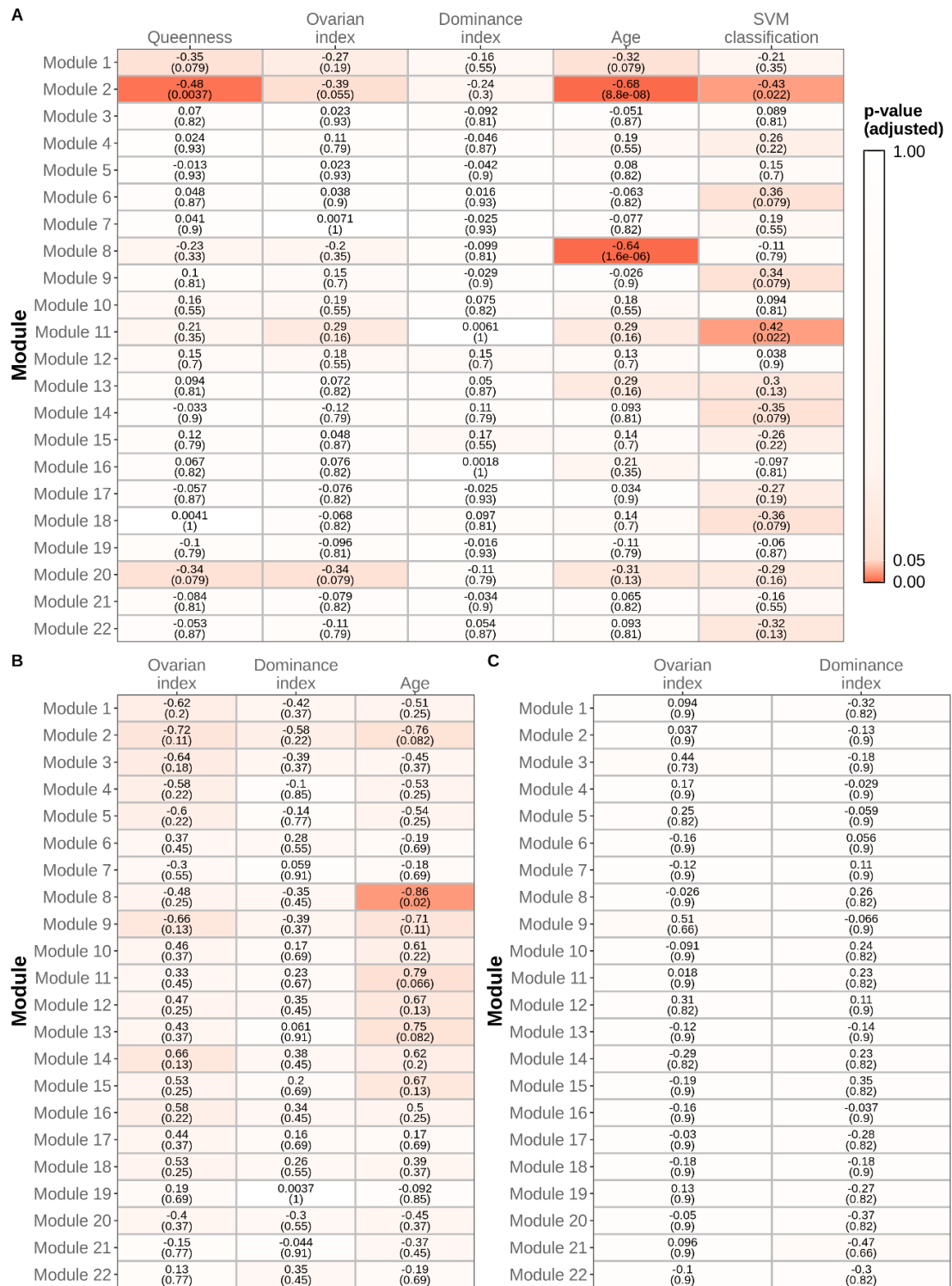
Figure 3.5. Module-trait correlations for consensus modules identified by WGCNA in (A) individuals from queen removal nests, (B) control workers and (C) queens. Within each cell, the Pearson correlation coefficient is given above and the FDR-adjusted p-value is given below in parentheses.

3.4   Discussion

3.4.1   Molecular correlates of caste plasticity in *P. dominula*

My discovery of colony-wide responses to queen loss suggests that this social perturbation provokes a significant reaction even from individuals that have little hope of attaining the vacant reproductive role. This is an unexpected finding given that *P. dominula* appears to express a 'conventional' gerontocratic mechanism of queen succession that mitigates the need for costly intragroup conflicts over the identity of the replacement queen (Pardi 1948; Tsuji & Tsuji 2005; Monnin et al. 2009), which should greatly reduce the need for young, low-ranking workers to respond to queen loss (Chapter 2). The gene expression responses of lower-ranked workers to queen removal might plausibly represent a form of safeguard against queen loss: if queen loss sometimes occurs multiple times in quick succession or is frequently associated with a general reduction of the nest population (i.e. through predation), there might be kin-selected benefits of a colony-wide 'de-differentiation' of individuals that facilitates a quicker succession process. This possibility is supported by the fact that one of the WGCNA modules that I identified was significantly enriched for the GO term *chromatin remodeling*. Regulation of chromatin accessibility, together with other epigenetic mechanisms, is a prime candidate mechanism for the regulation of caste identity in social insects (Simola et al. 2016; Wojciechowski et al. 2018; Duncan et al. 2020). The fact that this module was correlated with transcriptomic but not phenotypic caste identity among individuals from post-removal nests may therefore indicate that individuals respond to queen loss by priming their transcriptomes for further epigenetic changes, but that this priming only goes on to produce phenotypic effects if accompanied by a second trigger (such as the absence of an older sibling on the nest).

I identified 34 GO terms that were significantly enriched among the 1992 genes optimally separating queens and workers. Most of these terms were related to the regulation of gene expression or cell signalling, rather than to processes directly involved in reproduction or aggression. I also did not find a significant overlap between these 1992 genes and 1533 genes whose expression was correlated with vitellogenin in a previous study of caste in *P. dominula* (Standage et al. 2016; Jaccard index = 0.09; one-sided hypergeometric overlap $p = 0.64$), which reinforces the findings of my GO analysis because vitellogenin is thought to be a key regulator of aggression in *P. dominula* (Manfredini et al. 2018). While it might have been expected that GO terms associated with aggression or

reproduction would be among the most highly enriched terms separating queens from workers, my results do fit with previous studies of caste expression in *P. dominula*. For example, the most highly-enriched GO terms differentiating queens and workers in Standage et al. (2016) were involved in biosynthetic processes, which the authors argue to be evidence in favour of the idea that highly-conserved genes with basic biological functions may play a key role in the evolution of insect sociality (the 'genetic toolkit' hypothesis; Toth & Rehan 2017). While my 1992 caste-biased genes do not significantly overlap with 295 genes that Standage et al. identified as differentially expressed between queens and workers (Jaccard index = 0.02; one-sided hypergeometric overlap $p = 0.72$), the GO terms that I identified do correspond to highly-conserved functions such as transcription and biosynthesis. My results thus may be interpreted as adding to the increasing body of data supporting the genetic toolkit hypothesis.

Replacement queens in my colonies did not have access to males and therefore remained unmated even after queen succession. This may partially explain the fact that even individuals with fully developed ovaries and very high dominance ratings did not transition to a fully queen-like gene expression profile, as mating can induce significant gene expression changes in insects (e.g. Gomulski et al. 2012; Zhou et al. 2014; Manfredini et al. 2017). The lack of immediate mating opportunities for new queens in this experiment is not necessarily unrealistic, however: unmated Hymenopteran females can lay unfertilized eggs, which develop as males. Moreover, in naturally-occurring early *P. dominula* nests, replacement queens may be established a month or more before they are mated (Strassmann et al. 2004), presumably due to a scarcity of early males. The unmated replacement queens analysed here are therefore representative of those that would be present on wild nests shortly after queen loss.

3.4.2 Implications for the phenotypic gambit

This study contributes to the significant progress in our understanding of the relationship between molecular changes and changes in phenotypic expression that has been made in the past decade, facilitated by the increased availability of 'omic' data and complex bioinformatic analyses. Recent studies have started to challenge the view that there is a direct correspondence between transcriptomic states and external phenotypes. Libbrecht et al. (2018), for example, show that gene expression responses associated with a reversible phenotypic change differ qualitatively based on the directionality of the change (from reproductive to non-reproductive or vice versa). Meanwhile, molecular

manipulations have revealed a surprising degree of plasticity in canonically implastic traits such as mammalian sex (Matson et al. 2011) or ant castes (Simola et al. 2016). My results appear to go further, showing a shift in caste-specific brain gene expression profiles among individuals whose phenotypic caste expression remains otherwise apparently unchanged. If accurate, this result shows that the expectation of a close match between expression profiles and phenotypes is excessively simplistic, or at least that detecting such a match requires detailed knowledge of relevant genes and/or exhaustive phenotypic characterization.

While I failed to find a clear age-independent association between expression profiles and caste-related phenotypes, I cannot rule out the possibility that there are other, more subtle facets of caste identity in *P. dominula* that I failed to measure and that might explain variation in gene expression. For example, *Polistes* wasps are known to exhibit increased juvenile hormone titres following queen loss (Tibbetts & Huang 2010), and foundress queens may possess more substantial lipid stores than early season workers (Yoshimura & Yamada 2018). It is possible that changes in these traits or others would explain the large shifts in gene expression that I observed across all individuals following queen removal. I nonetheless consider it significant that the phenotypic changes associated with queen removal are not expressed at the level of ovarian development or dominance. These two traits are the ultimate determinants of caste identity: the fact that an individual has high JH levels does not matter to colony functioning if that same individual continues to occupy the social and reproductive role of a worker. Therefore, while I cannot unequivocally state that my results reflect a disconnect between phenotype and gene expression, it is certainly unexpected that changes in individuals' caste-specific gene expression profiles would not be reflected at the level of caste expression that matters.

It is also possible that the relationship between caste phenotypes and expression profiles would be more obvious in tissues that were not assayed in my study (such as ovaries). I concentrated on the brain because I was able to assess individuals only over a relatively brief period following queen removal and this tissue is known to show high short-term expression plasticity (Bell & Robinson 2011; Rittschof et al. 2014). Similarly, my choice to analyse transcriptomes of whole brain tissue rather than singling out individual tissues within the brain was driven by both practical and scientific concerns: I aimed to acquire enough tissue per sample to avoid pooling, and also to remain agnostic regarding the specific regions of the brain that could be influential upon reproductive phenotypes. Despite the limitations of focusing on a single, heterogenous tissue type, my analyses

revealed significant associations between measured expression patterns and ovarian development, indicating that brain gene expression at least partially reflects organismal physiology and the state of other tissue types.

A major advantage of this study is the use of individual-level gene expression data from a large number of subjects, including individuals reared in a shared social environment but exhibiting very different phenotypic responses to perturbation. By sequencing individuals rather than pools, I was able to match each gene expression profile to high-resolution phenotypic data that captures the scale of naturally-occurring variation in features such as age, ovarian development and dominance behaviour. This resolution allows me to address questions that are otherwise inaccessible in gene expression analyses. For example, I have been able to show that caste identity, but not the residuals of caste identity on age, are significantly predictive of individuals' change in transcriptomic caste identity following queen loss.

3.4.3 Support vector classification as a sociogenomic tool

Applying a support vector classification approach to behaviour-associated transcriptomic data, I identified a large group of genes as differing meaningfully between *Polistes* castes—over 10% of annotated genes, a similar number to that which I was able to identify using a standard differential expression program with no log fold change threshold. While the set of caste-biased genes identified by my SVM approach did not differ strongly in size from that generated using a standard approach, the use of SVMs nonetheless provides two clear advantages over standard analyses. First, SVM classification not only identified a set of genes that are predictive of caste identity, but the univariate classification that the model generated also allowed me to characterize the molecular caste identity of samples that did not fall squarely into the roles of control workers and queens. Standard differential expression analyses such as edgeR (Robinson et al. 2010), DESeq2 (Love et al. 2014), or NOISeq (Tarazona et al. 2011) assess differential expression at the level of individual genes. This focus makes these approaches well-suited to the identification and ranking of genes that distinguish pre-defined states, but of limited use when classifying intermediate or uncategorized samples, a purpose for which SVM analysis is ideally suited.

A second strength of the SVM classification approach is that it partially bypasses the requirement for judgements over which genes can be considered 'biologically

meaningful'. Differential expression studies almost always include a fold change threshold in their analyses to ensure that they do not include genes with very small fold changes in expression that may be statistically but not biologically significant. Choice of fold change threshold is largely arbitrary and can be hugely impactful upon the results achieved (Dalman et al. 2012). For example, in this study I identified just 81 differentially-expressed genes when using a 1.5 fold-change threshold, compared to 2438 genes when no threshold was applied. Both of these sets of genes exhibited reduced predictive ability compared to a model generated using SVM feature selection, suggesting that the standard differential expression approach either misses genes that are predictive of caste identity or includes genes that confound caste prediction (via overfitting). Indeed, the primary reason that the majority of standard differential expression analyses include a fold-change threshold is the concern that failure to do so will result in the identification of genes that are statistically differentially expressed but are biologically uninformative (Jung et al. 2011). My results reinforce this notion. By contrast, the decision of whether a given gene was included in my SVM model was based on a criterion that directly reflects biological relevance: did expression measures for that gene provide additional information about an individual's likely caste identity, and hence improve the predictive ability of my model?

3.5   Conclusions

Species that exhibit inducible phenotypic plasticity represent excellent models with which to examine the relationship between gene expression and phenotypes. Non-superorganismal social insect species exhibit a particularly tractable form of plasticity between distinct caste phenotypes, and yet the mechanisms that underlie this plasticity have received little attention. In this chapter, I have shown that gene expression responses to queen loss in *Polistes dominula* appear to be decomposed into two components. One component is shared by all individuals regardless of whether or not they exhibit actualised phenotypic plasticity, a surprising result that contrasts with the orderly phenotypic response to queen loss described in Chapter 2. The other component of gene expression response is strongly correlated with phenotype, and seems to consist primarily of changes in highly-conserved genes and functions. My results broadly support a role for conserved 'toolkit genes' in the control of social insect reproductive plasticity, but also highlight the complexity of the relationship between gene expression and phenotype.

Chapter 4 – Molecular signatures of alternative fitness strategies in *Liostenogaster flavolineata*

4.1 Introduction

Liostenogaster flavolineata is a facultatively social species in which the majority of individuals live in small groups with a clear division of labour between reproductive 'queens' and non-reproductive 'workers', despite being fully capable of nesting independently. More so than in obligately social taxa such as *Polistes*, queuing for direct fitness opportunities is thought to be a dominant factor incentivising the decision to remain in a non-reproductive role for *L. flavolineata* workers (Field et al. 2006; Bridge & Field 2007), although indirect fitness also plays a clear role in this species (Field et al. 2000). In this species, individuals that are of a higher rank in the group's social hierarchy, and therefore more likely to inherit the reproductive position, perform a significantly lower proportion of a group's foraging than do lower-ranked individuals, reflecting a continuum of reproductive investment: higher-ranked individuals are more strongly invested in direct fitness and less strongly invested in indirect fitness than lower-ranked individuals (Field et al. 2006; Bridge & Field 2007). In this chapter, I test whether the graduated differences in fitness investment in *L. flavolineata* groups are reflected at the level of gene expression. I find that the largest transcriptomic differences within a group are between queens and workers, as in obligately social species, although variation in foraging effort also explains a significant degree of variation in transcription among workers. I also provide evidence that genes associated with indirect fitness (as measured by foraging effort) in this species are more ancient and more conserved than genes that are associated with direct fitness, a surprising result that contradicts findings from more derived social insect species. The results presented herein represent the first detailed analysis to date of the molecular correlates of fitness strategies in a facultatively social wasp.

4.1.1 Alternative fitness strategies in facultatively social species

In socially complex insect species, most individuals spend the majority of their lives invested in a single role as either a queen or worker, even where they retain the ability to transition between roles (Fletcher & Ross 1985). For example, the fact that queen loss is a relatively rare occurrence means that most workers of a *Polistes* colony will never reproduce themselves (even though they have the capacity to do so) and are therefore

dedicated to an indirect fitness role. The relative stability of fitness strategies in such species is reflected in strong molecular differentiation: in Chapter 3, for example, I identified some 2000 genes differentiating *P. dominula* queens and workers. Moreover, such species are obligately social, such that individuals may occasionally transition between reproductive roles within a group, but solitary living does not occur.

In facultatively social species, however, the distinction between social and solitary living, and between direct and indirect fitness strategies, is often far less clear. In halictid sweat bees and some *Ceratina* carpenter bees, for example, females may or may not rear daughters that remain at the nest as non-reproductive workers (Eickwort et al. 1996; Smith et al. 2003; Wcislo et al. 2004; Kapheim 2010; Rehan & Richards 2010). Both sociality and reproductive role are highly plastic traits in these taxa: variation in the degree of sociality exhibited is predicted by ecological conditions (Cronin & Hirata 2003; Eickwort et al. 1996; Rehan et al. 2011; Mikát et al. 2017), and non-reproductive workers are able to mate and subsequently match the reproductive output of their mothers (i.e., workers are not subfertile; Smith et al. 2009).

As a result of the plastic nature of their social and fitness strategies, facultatively social bees have become models with which to test hypotheses regarding the earliest stages of social evolution (Shell & Rehan 2018). For example, facultatively social bee taxa have been used to provide evidence for 'molecular ground plan' hypotheses (Kapheim et al. 2012, 2020), and also to test the long-standing sociogenomic prediction that queen-biased genes should be relatively ancient and conserved compared to worker-biased genes (Jones et al. 2017). Such taxa are also ideal systems with which to study the respective contributions of social and reproductive context to individuals' molecular (e.g. gene expression) profiles, since individuals can be induced to occupy solitary or social and reproductive or non-reproductive roles separately. Studies of facultatively social insects outside of the Anthophila (bees) are rare, however.

The stenogastrine hover wasps, of which *L. flavolineata* is the best-studied example, represent an independent evolutionary origin of facultative sociality in the Vespidae (Bank et al. 2017). In *L. flavolineata*, as in other facultatively eusocial species, many individuals spend part of their lives as non-reproductive 'workers' but may transition at some point to a reproductive 'queen' role, shifting between indirect and direct fitness strategies (Field et al. 2006; Bridge & Field 2007). *L. flavolineata* groups consist of strictly age-determined dominance hierarchies, with the oldest non-reproductive individual in a group first in line to replace the reproductive. As a result of the fact that

they are more likely to attain a direct fitness role in the near future, older (higher ranked) individuals exhibit a reduced foraging effort relative to younger (lower ranked) individuals, indicating a shift away from indirect fitness investment as the chances of future direct fitness gains become greater (Field et al. 2006; Bridge & Field 2007). *L. flavolineata* colonies thus represent a particularly interesting model with which to study the molecular basis of insect sociality. At the level of realised direct fitness at any given time (that is, egg-laying; Sumner et al. 2002), groups exhibit the strong reproductive skew that is the defining characteristic of insect sociality; yet at the level of reproductive investment, there appears to be a relatively smooth gradation from indirect fitness investment to direct fitness investment as one moves up the dominance hierarchy. Whether gene expression best reflects actualised or potential fitness is an open question that analyses of *L. flavolineata* are well-suited to answer.

4.1.2 Evolutionary trajectories of genes associated with alternative fitness strategies

A long-standing hypothesis in the field of sociogenomics states that genes that are primarily involved in the regulation of indirect fitness should be less evolutionarily conserved than genes that are primarily involved in the regulation of direct fitness (Gadagkar 1997; Linksvayer & Wade 2009; Van Dyken & Wade 2010). Although queens in social insect societies may be larger, longer-lived and/or more fecund than females of solitary species (Roisin 2000; Jemielity et al. 2005; Holman 2014b; Friedman et al. 2019), the behaviours such individuals perform (e.g. egg-laying) are ones that necessarily existed in their solitary ancestors. By contrast, workers may exhibit highly derived traits, such as suicidal nest defence or hyperspecialised morphology (Shorter & Rueppell 2012; Boomsma & Gawne 2018), that are significantly apomorphic, and which may therefore be underlain by novel or fast-evolving molecular mechanisms. An additional reason to expect reduced conservation of 'indirect fitness genes' is that such genes are only indirectly related to fitness. The fitness of a non-reproductive worker is generated by effects of that worker's behaviour on a reproductive individual to whom the worker is usually not genetically identical; as a result, the trait-fitness covariance of worker-specific traits is reduced relative to traits that are expressed directly in reproductives (Marshall 2011), and this is expected to result in weakened selection on worker-specific traits (Linksvayer & Wade 2009).

Both of the factors that contribute to the theoretical differences in evolutionary trajectories for direct and indirect fitness-related genes are most prominent in highly-derived social species, especially 'superorganismal' species in which workers are obligately morphologically differentiated from queens. Accordingly, the strongest evidence for this theory has come from superorganisms (Harpur et al. 2014; Feldmeyer et al. 2014): for example, Warner et al. (2019) found that honey bee and pharaoh ant (*Monomorium pharaonic*) worker-biased genes are less ancient and more loosely-connected than queen-biased genes. By contrast, the few studies that have assessed evidence for differential conservation of indirect and direct fitness-related genes in non-superorganismal social insects have been more equivocal in their conclusions (Ferreira et al. 2013; Jones et al. 2017), which may be unsurprising given that alternative fitness strategies are far less canalised in such taxa. Nonetheless, the fact that these studies have been limited to a relatively small number of species mean that further tests are necessary before we can begin to establish a consensus on the validity or otherwise of this hypothesis in less-derived social taxa.

4.1.3 Chapter overview

In this chapter, I present the results of two experiments exploring the gene expression foundations of direct and indirect fitness strategies in *L. flavolineata*. The majority of the chapter involves an analysis of reproductive and non-reproductive social individuals from *L. flavolineata* groups, including individuals that were experimentally promoted from a lower to a higher within-group rank. I begin by examining the phenotypic correlates of rank: in doing so, I confirm previous work showing that individual-level foraging effort is strongly structured by rank, and that individuals that are promoted in rank decrease their foraging rate to match their new rank (Section 4.3.1). I next examine differential gene expression between individuals of different ranks to test the prediction that transcriptomic variation within groups will more strongly reflect potential than realised fitness, i.e. that the overall transcriptomic differences between reproductive 'queens' and non-reproductive 'workers' will be relatively minor in this species (Hypothesis 1; Section 4.3.2), while the differences that exist between individuals of different ranks overall will be more substantial (Hypothesis 2; Section 4.3.3). I also test the prediction that individuals that have been promoted in rank will shift their gene expression profiles to match their new rank, reflecting the shift towards direct fitness that occurs when an individual's rank increases (Hypothesis 3; Section 4.3.3). I then attempt to identify co-expressed gene modules whose expression correlates with investment into direct or indirect fitness strategies, hypothesising that if such co-expression modules exist, they

will overlap significantly in both composition and function with both caste-biased genes and rank-biased genes (Hypothesis 4; Section 4.3.4).

In the next two sections, I attempt to test the long-standing hypothesis that genes associated with direct fitness should be relatively conserved and slow-evolving relative to genes that are associated with indirect fitness. First, I compare sets of *L. flavolineata* fitness-biased genes to the *P. dominula* caste-biased genes identified in Chapter 3 to test whether direct fitness-biased genes are more likely than indirect fitness-biased genes to retain overlapping genetic signatures in separate lineages (Hypothesis 5; Section 4.3.5). I then compare the evolutionary ages of genes associated with different fitness strategies in *L. flavolineata*, predicting that direct fitness genes will be of relatively ancient origin compared to indirect fitness genes (Hypothesis 6; Section 4.3.6). Finally, I present the results of a secondary experiment decomposing the molecular effects of social and reproductive context, both of which may be significant determinants of gene expression in this facultatively social species (Hypothesis 7; Section 4.3.7).

4.2 Methods

4.2.1 A note on contributions

The work described in sections 4.2.2-4.2.4, as well as RNA extraction, was performed by Daisy Taylor. All other work described in this chapter is my own, but the *L. flavolineata* genome against which I align RNAseq reads (as described in section 4.2.5) was generated by collaborators: Alexandrina Bodrug, Nancy Stralis-Pavese and Heinz Himmelbauer (Department of Biotechnology, University of Natural Resources and Life Sciences, Vienna) contributed to the assembly; and Roderic Guigo, Francisco Câmara Ferreira and Anna Vlasova (Centre for Genomic regulation, Barcelona) contributed to functional annotation. Full details of the genome assembly are provided in **Appendix B**.

4.2.2 Wasp identification and marking

Fieldwork was undertaken in Fraser's Hill, Malaysia between January and April 2017. *L. flavolineata* nests from aggregations situated in under-road culverts were selected for further study on the basis of the presence of multiple pupal caps and observed egg-laying. Over two days, all individuals on each nest were given a unique combination of coloured paint marks to facilitate subsequent individual-level identification. Confirmation that all wasps had been successfully marked was achieved by censuses of group members at night, when all individuals are present on the nest. The presence of unmarked individuals

during night censuses was recorded, and such individuals were marked the following day (without removing them from the nest, in order to avoid interfering with orientation flights). Brood were mapped in each nest to confirm the presence of an egg-layer, and to identify pupae which would shortly hatch. Reproductives were identified by witnessing egg-laying or through censuses to determine foraging effort: reproductives rarely leave the nest in this species (Field & Foster 1999; Cant & Field 2001; Shreeves & Field 2002). Every day, nests were monitored in order to measure foraging effort and to identify newly emerging wasps. Newly emerged group members were identified by the co-appearance of an unmarked wasp and a hatched pupal cell (drifting between nests does not occur in this species). Once identified, newly emerged individuals were left on the nest for three days before marking, to avoid interfering with their nest orientation flights.

4.2.3 Experiment 1: Molecular basis of within-group dominance rank

To produce data that could be used to identify genes that are differentially expressed with shifts between direct and indirect fitness strategies, in this experiment nests with groups of 3-5 wasps of known age and rank were generated. Additionally, manipulations were performed in which the second-ranked (Rank 2) or third-ranked (Rank 3) individual from each nest was removed in order to promote lower-ranked individuals to a higher position within the group hierarchy and thus shift their fitness interests away from indirect fitness and towards future direct fitness, an experimental design that was based on that employed by Field et al. (2000). Doing so allowed me to compare gene expression between reproductives (Rank 1) and non-reproductives (Ranks 2-5) and also variation in gene expression among non-reproductives of different ranks (Ranks 2-5) and therefore with different fitness interests. I hypothesised that, reflecting the gradual shifts in fitness investment that exist across dominance hierarchies in this species, I would observe relatively little differentiation between Rank 1 and Rank 2 individuals (Hypothesis 1), but larger numbers of genes would be differentially expressed with rank when comparing across all individuals (Hypothesis 2). I also predicted that individuals that had been promoted from Rank 3 to Rank 2 would exhibit a concomitant shift in fitness-associated gene expression, exhibiting significant differentiation from Rank 3 individuals that did not change rank (those removed during manipulation), but not from Rank 2 individuals that had not changed rank (Hypothesis 3).

To generate the focal samples, 28 focal nests were identified and censused daily until 2-4 new individuals had emerged. These newly emerged individuals, together with the nest's reproductive (Rank 1), were chosen as the focal individuals for the experiment. In

order to generate nests of comparable group sizes and with non-reproductives of known ages, all other wasps were removed at dawn of the day following the emergence of the requisite number of focal wasps. For 10 days following this manipulation, nests were censused every ~30 minutes during peak foraging hours (07:00-11:00) in order to estimate foraging effort for each individual. From this initial manipulation until the end of the experiment, additional wasps that emerged from the nest were removed to ensure that colony sizes remained constant. Thus, at the end of the 10 day period, the rank and average time spent off-nest were known for each individual from 28 focal nests consisting of a single reproductive (of unknown age) and 2-4 non-reproductives of known ages (within-rank age range ≈ 5 days).

In order to promote lower-ranked wasps on each nest either from Rank 3 to Rank 2 or from Rank 4 to Rank 3, at dawn on the 11th day a single focal wasp was removed from each of the 28 nests and placed directly into RNAlater. The removed individual was always of either rank 2 (n = 15; age mean ± SD = 30.3 ± 3.1 days) or rank 3 (n = 7; age mean ± SD = 25.3 ± 1.4 days). In order to ensure that the ratio of helpers to brood remained constant, brood were also removed from the nests at this time. Brood was divided into three categories: eggs, small/medium larvae and large larvae. A proportion, R/N (where R is the number of adults removed and N is the original number of wasps), of each category was removed using fine tweezers. Pupae, which do not require feeding, were not removed. Nests were given 48 hours to settle, after which censuses were performed daily for 5 days using the same methodology as previously. At the end of the 5-day censusing period, all wasps were removed from the nest before dawn: individuals' heads were placed directly into RNAlater (Thermo Fisher Scientific) for gene expression analysis, and their bodies into 95% EtOH for dissection (focal collection 2). The ages of non-reproductives of the same rank collected at this stage that were subsequently sequenced were approximately age-matched across nests: Rank 2 (n = 21; age mean ± SD = 33.3 ± 4.3 days); Rank 3 (n = 15; age mean ± SD = 27.0 ± 1.6 days); Rank 4 (n = 6; age mean ± SD = 22.8 ± 1.5 days). The ovarioles of each individual were subsequently dissected and the number of developed eggs counted, and the mating status of each individual assessed by examining the spermatheca for the presence of sperm.

4.2.4 Experiment 2: Effects of social context on the molecular basis of caste-specific behaviours

In order to separate out the effects of social and reproductive context upon gene expression, in this secondary experiment data were generated for age-matched (28 day-old) wasps that were foraging either alone ('solitary foragers') or as part of a group ('social foragers'), and likewise individuals that were egg-laying either alone ('solitary reproductives') or as part of a group ('social reproductives'). I predicted that I would be able to identify distinct suites of genes differentiating social and solitary individuals from one another regardless of reproductive context, and likewise differentiating reproductive from non-reproductive individuals regardless of social context (Hypothesis 7).

To generate the samples necessary to test this hypothesis, nest membership on 30 focal nests was monitored and a single newly-emerged wasp from each was chosen as the focal individual. When the focal individual had reached an age of 10 days, nests were randomly assigned to one of four treatment groups and manipulated accordingly:

Social Forager (SocFor; n = 7): The nest was manipulated to ensure a nest size of 2-4 individuals, including both the reproductive and the focal individual. After 18 days, the focal individual was collected directly from the nest.

Social Reproductive (SocRep; n = 9): The reproductive was removed from nests of 3-5 individuals, leaving the focal wasp (who became the new reproductive) and 2-3 younger individuals. After 18 days, the focal individual was collected directly from the nest.

Solitary Forager (SolFor; n = 6) & *Solitary Reproductive* (SolRep; n = 8): All wasps except the focal individual were removed. After 18 days, the focal individual was collected either returning to the nest with forage (SolFor) or laying an egg (SolRep).

For all treatments, newly emerged wasps were removed daily following manipulation to ensure nest size remained constant. Egg-laying was confirmed through brood counts. Censuses were performed daily every ~2 hours during peak foraging times (07:00-11:00) to determine foraging effort for each focal individual. Regardless of treatment, after 18 days the focal individual (always aged 28 days) was removed from the nest using forceps. Focal individuals' heads were placed directly into RNAlater (Thermo Fisher Scientific) for gene expression analysis, and their bodies into 95% EtOH for dissection. The ovarioles of each focal individual were subsequently dissected and the number of developed eggs counted, and the mating status of each individual assessed by examining the spermatheca for the presence of sperm.

4.2.5 Gene expression quantification

Brain tissue was extracted from the heads of individual samples and RNA was extracted using the RNeasy Mini Kit (Qiagen) according to manufacturer's instructions. Library preparation was performed by Novogene Co. followed by sequencing on an Illumina HiSeq 2000 platform with 150-bp paired-end reads to a depth of 30 million reads/sample. Transcript filtering, trimming, alignment and quantification were performed with the aid of Nextflow using the default options provided by the nf-core/rnaseq pipeline v1.4.2 (Ewels et al. 2020): adapter and quality trimming were performed using TrimGalore (Kreuger 2015), followed by removal of ribosomal sequences using SortMeRNA (Kopylova et al. 2012); read alignment against the *L. flavolineata* genome (**Appendix B**) was performed using STAR (Dobin et al. 2013) followed by quantification using Salmon (Patro et al. 2017); finally reads were assembled into transcripts using StringTie2 (Kovaka et al. 2019) and then subjected to a round of filtering (separately for each experiment) in order to remove any gene that was not expressed at a minimum of 1 count/sample in at least one focal group. Following this filtering, 11258/14095 (79.9%) and 11590/14095 (82.2%) transcripts remained for Experiment 1 and Experiment 2 analyses respectively.

4.2.6 Differential expression analysis

Differential expression analyses were performed in R using the DESeq2 package (Love et al. 2014). Unless otherwise stated, differential expression was calculated relative to a baseline fold change of 0. Where a fold change threshold was used, this value became the baseline against which differential expression was measured (i.e., for a fold change threshold of 1.5, genes were considered differentially expressed if their absolute fold change was significantly greater than 1.5). Genes were considered differentially expressed between conditions if $p < 0.05$ after false discovery rate correction according to the Benjamini-Hochberg procedure.

4.2.7 Gene ontology (GO) enrichment analysis

In order to perform GO enrichment analysis, I first identified reciprocal BLAST best hits between the *L. flavolineata* and *Drosophila melanogaster* proteins using BLAST+ (Camacho et al. 2006). 8266/11258 (73.4%) *L. flavolineata* genes possessed a reciprocal best hit with *D. melanogaster*. GO annotations for each D. melanogaster gene were acquired from BioMart (Smedley et al. 2009) and each *L. flavolineata* gene was assigned the GO terms of its reciprocal best hit. GO enrichment analysis was then performed in R via the topGO package (Alexa & Rahnenfuhrer 2009) using TopGO's weight01 algorithm

and Fisher's exact test to identify GO terms that were significantly overrepresented (p<0.01) in a focal set of genes against a background consisting of all genes that appeared in the relevant analysis.

4.2.8 Gene co-expression network analysis

I sought to identify co-expressed modules of genes associated with correlates of direct and indirect fitness (Hypothesis 4). To achieve this, weighted gene co-expression network analysis was performed in R using the WGCNA package (Langfelder & Horvath 2008). As WGCNA is particularly sensitive to genes with low expression, data were first subjected to a second round of filtering in which genes that had <10 reads in >90% of Experiment 2 non-reproductives were removed, as recommended by the package authors. This second round of filtering removed an additional 1821 genes, leaving a total of 9437 genes. Counts were subjected to a variance-stabilizing transformation prior to further analysis. Consensus gene modules across all Experiment 2 non-reproductives were then constructed using a soft-threshold power of 6 and the signed hybrid adjacency criterion. Network summary measures and gene dendrograms for this analysis are provided in **Supplementary Figures S4.1-4.2**. Initially, 26 gene modules were identified. Modules whose eigengene correlation was >75% were subsequently merged, after which 20 consensus modules remained. Finally, the Pearson correlation of each module with each phenotypic trait was calculated and subjected to Benjamini-Hochberg FDR correction.

4.2.9 Gene list overlap tests

I compared the overlap between different sets of genes and GO terms both to identify whether the correlates of direct and indirect fitness were consistent between different comparisons within *L. flavolineata*, and also subsequently to compare suites of putatively fitness-related genes across *L. flavolineata* and *P. dominula* (Hypothesis 5). I chose to compare the *L. flavolineata* against the *P. dominula* dataset that I describe in Chapter because the two datasets are comparable in terms of sequencing depth, sampling size, and differential expression analysis methodology, with the additional advantage that the reader will already be familiar with the experimental design that was used to generate the *P. dominula* data.

In order to determine the significance of overlaps between different sets of genes and GO terms, I performed two-sided hypergeometric tests between each pair of sets. The background for each comparison was the intersection of the two lists from which the sets were drawn; for example, for a comparison between differentially-expressed genes differentiating *P. dominula* castes and differentiating *L. flavolineata* 'castes', the

background was the intersection of the full lists of genes that were originally subjected to differential expression analysis in each comparison (less any genes for which orthologues were not identified between the two groups).

4.2.10 Phylostratigraphy

I used phylostratigraphy to detect differences in the evolutionary ages of genes that were associated with direct and indirect fitness strategies (Hypothesis 6). Phylostratigraphy is a method by which the approximate age of genes is estimated by assessing the presence or absence of identifiable orthology for each gene among a large number of other species. Phylostratigraphy was performed in R using the phylostratr package (Arendsee et al. 2019). I began by downloading protein FASTAs from UniProt for 154 species evenly distributed between 19 phylostrata. I then used BLAST to match each *L. flavolineata* gene against each genome in order to identify the oldest phylostratum within which that gene possessed an identifiable orthologue. For simplicity, I assigned genes to four categories: 'ancient' genes possess orthologues in non-hexapod animals; 'insect' genes are found in non-hymenopteran insects but not non-insects; 'hymenopteran' genes are found in hymenopterans but no other insect; and 'aculeate' genes were found in aculeates but no other taxon.

4.3   Results

4.3.1 Phenotypic correlates of rank in *L. flavolineata*

Before testing the hypotheses outlined in section 4.1.3, I examined the foraging effort and ovarian development of wasps from Experiment 1 to determine whether individuals exhibited the behavioural and physiological correlates of reproductive rank found in previous studies (Shreeves & Field 2002; Sumner et al. 2002; Bridge & Field 2007).

Time off nest is thought to be a reliable proxy for foraging effort: older individuals in *L. flavolineata* are more likely to inherit the position of egglayer and are therefore less invested in risky foraging behaviour (Field & Cronin 2006; Bridge & Field 2007). In line with this, there was a strongly significant negative relationship between within-group age rank and the amount of time individuals from unmanipulated groups spent off the nest (linear regression of time off-nest on rank: slope$\pm$SE = -27.37 $\pm$ 1.95, $p = 5.6\times10^{-23}$, $t_{df} = 78$; **Figure 4.1A**). Additionally, there was a near-significant trend for individuals from smaller nests to spend a larger proportion of time off-nest relative to similarly ranked counterparts on larger nests (linear regression of time off-nest on nest size: slope$\pm$SE = -

5.084 ± 2.93, $p = 0.087$, $t_{df} = 77$), as expected if there are diminishing returns of helping effort with increasing nest size.
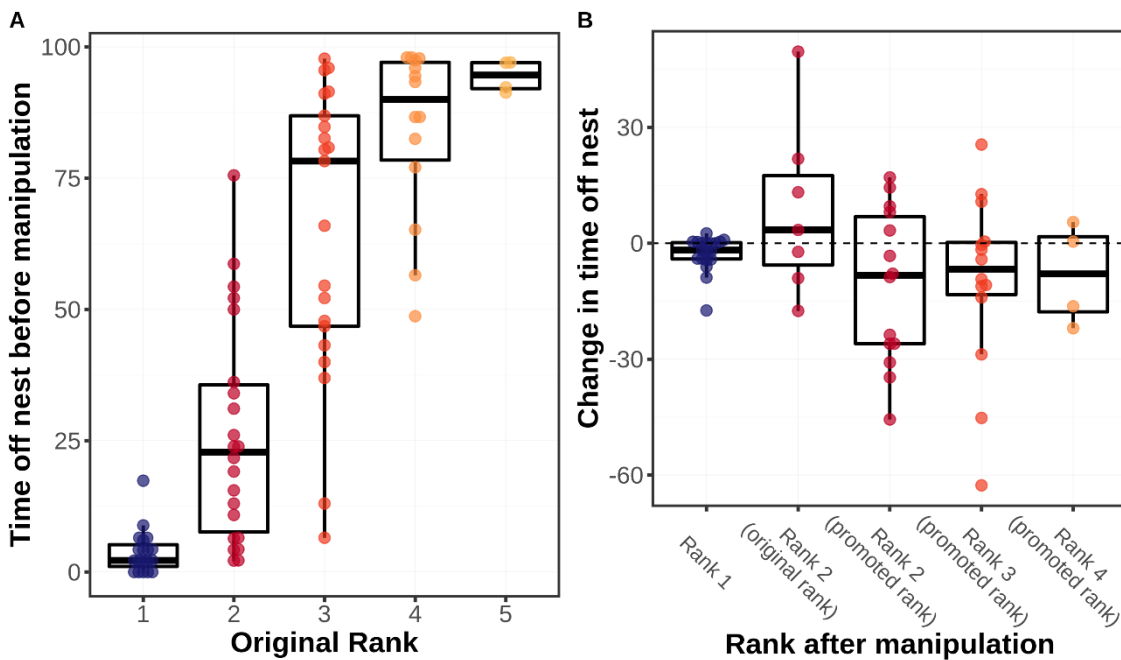
Figure 4.1. (A) Proportion of time spent off-nest before manipulation and (B) change in proportion of time spent off-nest following manipulation by individuals of different within-group ranks.

I observed the expected changes in foraging efforts in response to the manipulation of rank. Individuals that were promoted in rank spent less time off the nest (and therefore putatively less time foraging) following manipulation (n = 35; two-sided Wilcoxon W = 137, $p = 0.017$), although the dispersion of this change was high (mean±SD = -10.16 ± 20.25; **Figure 4.1B**). By contrast, individuals whose rank did not change showed no significant shift in their time spent off the nest (n = 26; mean±SD = 0.35 ± 12.68; two-sided Wilxocon V = 90, $p = 0.24$) despite the fact that group size was decreased for these individuals following manipulation, indicating that the removal of brood during manipulation was successful in maintaining a constant per-capita foraging requirement.

Ovarian development was strongly dependent on rank. The most dominant individual (Rank 1) within a nest was always inseminated and possessed several mature eggs in her ovarioles (n = 19; mean±SD = 12.63 ± 2.29 eggs). Meanwhile, individuals of Rank 2 and below (n = 64) almost never possessed developed eggs; the sole exceptions were two Rank 2 individuals, which possessed two and three developed eggs respectively in their ovarioles at the time of sampling. Neither these individuals nor any other individual below Rank 1 was inseminated.

Overall, my results are in line with previous work on this species: groups are defined by a strong reproductive division of labour between individuals at Rank 1 (who are the sole egg-layers) and an age-based queue of non-reproductives (Rank 2 and beyond), who show reduced investment in foraging effort the higher ranked they are in the queue (Field et al. 1999; Shreeves & Field 2002; Sumner et al. 2002; Bridge & Field 2007). Time off nest can therefore be used as a proxy for foraging effort, and rank as an indicator of investment into direct vs indirect fitness. Hereafter I use the term 'foraging effort' as a placeholder for the proportion of time spent off the nest.

4.3.2 Hypothesis 1: Molecular basis of reproductive division of labour

A principal component analysis of the 100 most variable genes revealed only weak clustering of individuals by rank or treatment (**Figure S4.3**), suggesting that there was little overall transcriptomic differentiation between individuals of different ranks. However, comparing individuals of different ranks in a pairwise fashion using DESeq2, I found strong signals of differential expression between reproductives (Rank 1 individuals) and all other treatment groups, with >300 differentially-expressed genes (DEGs) in each comparison (**Figure 4.2A**). This result held whether grouping individuals by their rank prior to manipulation or by their rank following manipulation (**Figure 4.2B**), suggesting that manipulation had not strongly disrupted the social order of the group.
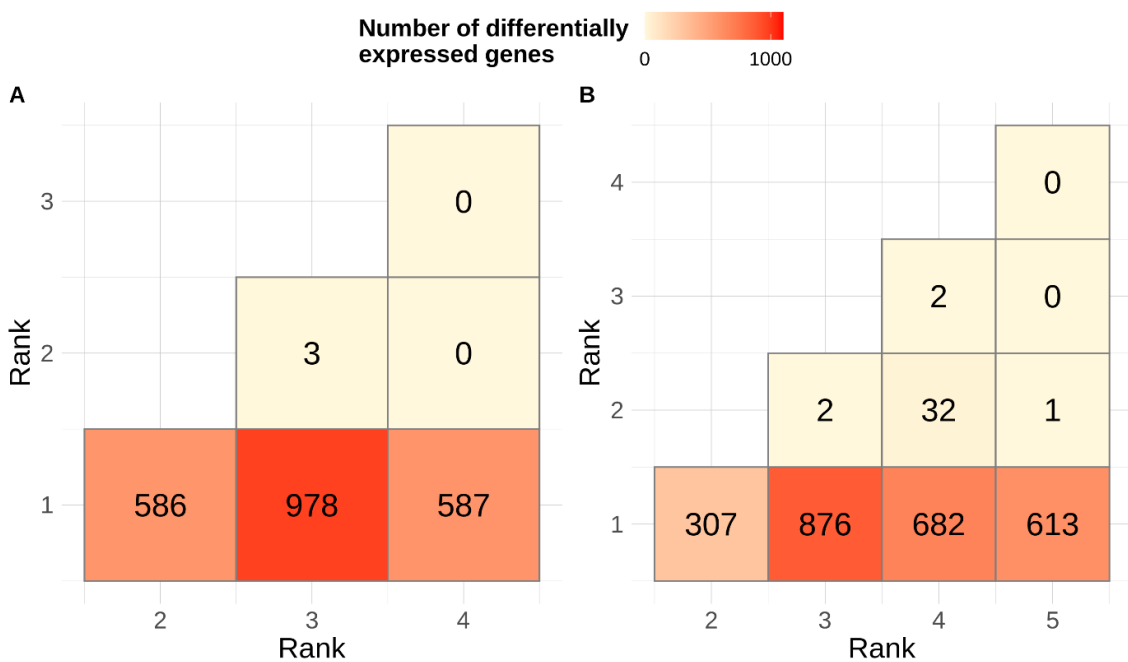
Figure 4.2. Number of genes differentially expressed between each pair of ranks. Ranks shown are those to which individuals belonged before (A) and after (B) manipulation.

Comparing all reproductives against all non-reproductives, I identified 1117 DEGs. The 469 queen-biased genes were associated with 18 significantly enriched GO terms (**Figure S4.4; Table S4.1**), including several terms related to DNA replication. The 648 genes that were worker-biased were associated with 73 significantly enriched GO terms (**Figure S4.5; Table S4.2**), most of them connected to respiration and metabolism. My results suggest a high degree of differentiation between 'queens' and 'workers' in this facultatively social species, comparable to that present in obligately social species such as *P. dominula* (as I found in Chapter 3).

4.3.3 Hypotheses 2 & 3: Molecular basis of fitness investment among non-reproductives

Over 1000 genes were identified as being differentially expressed with foraging effort and/or with rank; this was true regardless of whether the rank considered was that identified before or after manipulation (**Table 4.1**). However, these results appear to have been driven principally by the presence of reproductives: when I excluded Rank 1 individuals, the number of genes that were differentially expressed with continuous variables was greatly reduced (**Table 4.1**). Notably, when excluding reproductives, the number of genes differentially expressed with respect to foraging effort (n=256) was much larger than the number differentially expressed with rank (n=45). Moreover, I found that no genes were differentially expressed with rank after correcting for the confounding relationship between rank and foraging effort (by taking the residuals of a regression model with rank, either pre- or post-manipulation, on foraging effort), which suggests that it is foraging rate *per se* rather than rank that best predicts individuals' gene expression profiles. As such, I subsequently focused on the molecular signatures of foraging effort among non-reproductive individuals.

Trait	DEGs (Rank 1 included)	DEGs (Rank 1 excluded)
Original rank	1204	45
Final rank	1085	1
Foraging effort	1282	256
Age	NA	0

Table 4.1. Number of genes whose expression was correlated with continuous traits across all individuals or when excluding Rank 1 individuals. Age data were not available for Rank 1 individuals and so it was not possible to identify genes that were differentially expressed with age when including this group.

Among non-reproductives, 173 genes were upregulated with respect to foraging rate, and these were associated with 26 significantly enriched GO terms, many of which were associated with developmental and metabolic processes (**Figure S4.6; Table S4.3**). The 83 genes that were downregulated with respect to foraging rate were associated with just 10 GO terms, including several terms related to metabolism (**Figure S4.7; Table S4.4**). Although the number of DEGs associated with foraging rate among non-reproductives (256 DEGs) is not negligible, it is considerably smaller than the number differentiating 'queens' and 'workers' ('caste-biased' genes; 1117 DEGs). This suggests that it is realised rather than potential fitness that most strongly structures gene expression in this species.

Individuals that were promoted in rank by removing a higher-ranked wasp responded by modulating their foraging effort to match their new rank, and this change appears to have been matched at the level of gene expression. Individuals that were promoted from Rank 3 to Rank 2 showed a strong gene expression differentiation from Rank 3 individuals that had not changed their rank (152 DEGs). By contrast, individuals that were promoted to Rank 2 showed little differentiation from Rank 2 individuals that had not changed their rank (6 DEGs). These results suggest that the shifts in fitness interest and fitness-related behaviour (foraging effort) that accompany a change in rank are also reflected at the level of the transcriptome, with individuals adopting a transcriptional profile that is closer to their new rank than to the rank they possessed prior to promotion.

4.3.4 Hypothesis 4: Co-expressed modules underlying direct and indirect fitness strategies

In practise, genes rarely act individually, but rather form co-regulatory networks with one another. As such, it is frequently informative to identify suites of genes ('gene modules') whose expression is significantly correlated and which may therefore share a common function or regulatory basis. Given the substantial number of genes that are differentially expressed between individuals displaying differing foraging effort among *L. flavolineata* non-reproductives, I had predicted that I might be able to identify co-regulated gene modules whose expression was consistently associated with rank and/or foraging effort, and which therefore represent regulatory networks putatively underlying different fitness strategies. For example, if there is a set of genes that regulate investment into direct fitness in tandem, then we might expect to identify this set of genes as a co-expressed module, and the expression of genes in this module should be consistently correlated with

individuals' rank or foraging effort. I identified 20 distinct co-regulated gene modules ranging in size from 32 genes to 1831 genes. None of these modules were correlated with rank (either before or after manipulation) nor with age, but two exhibited strong correlation with foraging effort (**Figure 4.3**).

One module, Module 12, was significantly negatively associated with foraging rate, and therefore putatively associated with direct fitness. Genes that were more strongly correlated with foraging rate (either positively or negatively) were also more strongly correlated with the eigengene of Module 12 (i.e. they had higher 'module membership'), suggesting a meaningful relationship between the expression of this module and individuals' foraging effort (**Figure 4.4A**). The 83 genes contained within this module were enriched for 14 GO terms, many of them associated with visual processing (**Figure S4.8; Table S4.5**). Genes that were part of this module overlapped significantly with genes and that were found to be negatively correlated with respect to foraging rate when using DESeq2 (two-sided hypergeometric $p < 0.001$), and also with genes that were queen-biased (i.e. those that were upregulated in reproductives vs non-reproductives; two-sided hypergeometric $p = 0.016$), further strengthening the supposition that this module is meaningfully associated with a direct fitness strategy. However, I did not find this same pattern of overlap at the level of gene function: GO terms that were enriched in this module did not overlap significantly with GO terms that were queen-biased (two-sided hypergeometric $p = 1$), nor with GO terms that were negatively correlated with foraging rate (two-sided hypergeometric $p = 0.116$).

A second module, Module 16, exhibited a strongly positive association with foraging rate, and so putatively associated with indirect fitness. As with Module 12, genes that were more strongly associated (either positively or negatively) with foraging effort were also more strongly correlated with this module's eigengene (**Figure 4.4B**). This large module contained 606 genes associated with 90 GO terms, among them many terms associated with respiration and metabolic and biosynthetic processes (**Figure S4.9; Table S4.**). These genes and GO terms overlapped significantly with the set of genes and GO terms that were upregulated with foraging rate among non-reproductives as measured by DESeq2 (genes: two-sided hypergeometric $p < 0.001$; GO terms: two-sided hypergeometric $p < 0.001$), and also with worker-biased (i.e. upregulated in non-reproductives vs reproductives) genes and GO terms (genes: two-sided hypergeometric $p < 0.001$; GO terms: two-sided hypergeometric $p < 0.001$), supporting the possibility that this module is an important regulator of indirect fitness strategies.

Module	Time off nest	Rank (before manipulation)	Rank (after manipulation)	Age
Module 1	0.2 (0.26)	0.14 (0.43)	0.079 (0.67)	-0.12 (0.62)
Module 2	-0.19 (0.26)	-0.16 (0.4)	-0.2 (0.45)	0.094 (0.62)
Module 3	-0.25 (0.19)	-0.2 (0.36)	-0.14 (0.52)	0.14 (0.62)
Module 4	-0.25 (0.19)	-0.26 (0.25)	-0.29 (0.27)	0.26 (0.51)
Module 5	0.098 (0.6)	0.084 (0.64)	0.15 (0.52)	-0.11 (0.62)
Module 6	0.11 (0.54)	0.13 (0.48)	0.085 (0.67)	-0.12 (0.62)
Module 7	0.2 (0.26)	0.26 (0.25)	0.12 (0.58)	-0.2 (0.62)
Module 8	0.048 (0.73)	-0.0013 (0.99)	-0.14 (0.52)	-0.0021 (0.99)
Module 9	0.13 (0.47)	0.12 (0.48)	-0.02 (0.94)	-0.062 (0.77)
Module 10	0.32 (0.1)	0.24 (0.25)	0.11 (0.58)	-0.15 (0.62)
Module 11	-0.27 (0.17)	-0.17 (0.38)	-0.085 (0.67)	0.17 (0.62)
Module 12	-0.48 (0.0037)	-0.34 (0.13)	-0.22 (0.44)	0.23 (0.62)
Module 13	-0.24 (0.21)	-0.18 (0.38)	-0.011 (0.94)	0.12 (0.62)
Module 14	-0.17 (0.32)	-0.14 (0.43)	0.014 (0.94)	0.1 (0.62)
Module 15	0.053 (0.73)	0.046 (0.78)	0.18 (0.45)	-0.03 (0.87)
Module 16	0.4 (0.022)	0.33 (0.13)	0.37 (0.12)	-0.27 (0.51)
Module 17	0.23 (0.21)	0.18 (0.38)	0.28 (0.27)	-0.14 (0.62)
Module 18	0.3 (0.14)	0.24 (0.25)	0.26 (0.27)	-0.15 (0.62)
Module 19	0.05 (0.73)	0.059 (0.74)	0.14 (0.52)	-0.034 (0.87)
Module 20	-0.075 (0.69)	-0.23 (0.27)	-0.19 (0.45)	0.096 (0.62)

Figure 4.3. Association of phenotypic traits (time off nest, which is a proxy for foraging effort; rank assessed before manipulation; rank assessed after manipulation; and age) among non-reproductive individuals with co-expressed gene modules present in those individuals. Number outside parentheses: Pearson correlation. Number within parentheses: FDR-corrected *p*-value.
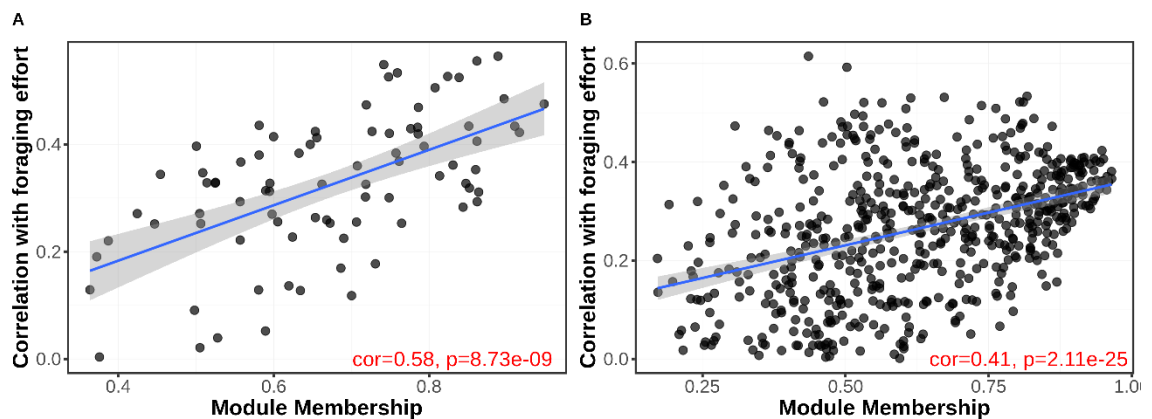
Figure 4.4. Strength of correlation of genes with foraging effort among non-reproductives, plotted against module membership of those genes within (A) Module 12 and (B) Module 16.

In summary, I identified two modules that were putatively associated with two complementary fitness strategies (direct and indirect). Genes that were part of Module 16, and GO terms that were enriched in this gene set, were significantly overrepresented among genes that were worker-biased, and likewise among genes that were positively

correlated with foraging rate among non-reproductives. In totality, these results appear to indicate that this module represents a regulatory network that is meaningfully associated with indirect fitness investment in *L. flavolineata*. Similarly, genes that were part of a separate module (Module 12) were significantly overrepresented among genes that were queen-biased and genes that were negatively correlated with foraging rate among non-reproductives, which suggests that this module may play a role in direct fitness investment. However, GO terms that were enriched in this module did not exhibit the same degree of overlap with other direct fitness-associated comparisons, which renders the link between this module and direct fitness more equivocal than that between Module 12 and indirect fitness.

4.3.5 Hypothesis 5: Molecular processes underlying two independent origins of insect sociality

In the analyses described above, I identified several overlapping sets of genes and GO terms associated with indirect fitness. I also identified several overlapping sets of genes that are associated with direct fitness (although the gene ontology overlap between these groups was less consistent). A long-standing prediction of sociogenomic theory is that genes and processes that are associated with direct fitness will be relatively evolutionarily conserved, while genes that are associated with indirect fitness will be faster-evolving and associated with relatively novel functions. To test this hypothesis, I assessed the degree to which these sets of fitness-associated genes overlapped with the sets of genes identified by DESeq2 as being caste-biased in *P. dominula* in Chapter 3.

I found that each set of genes and GO terms that was putatively associated with indirect fitness in *L. flavolineata* overlapped strongly with worker-biased genes in *P. dominula* (**Figure 4.5**). By contrast, gene sets and GO terms that I found to be putatively associated with direct fitness in *L. flavolineata* did not exhibit a significant degree of overlap with queen-biased *P. dominula* genes (**Figure 4.6**). These results run directly counter to my predictions: where theory predicts that direct fitness-associated genes, but not indirect fitness-associated genes, will be conserved among different origins of sociality, my results appear to support the opposite conclusion.
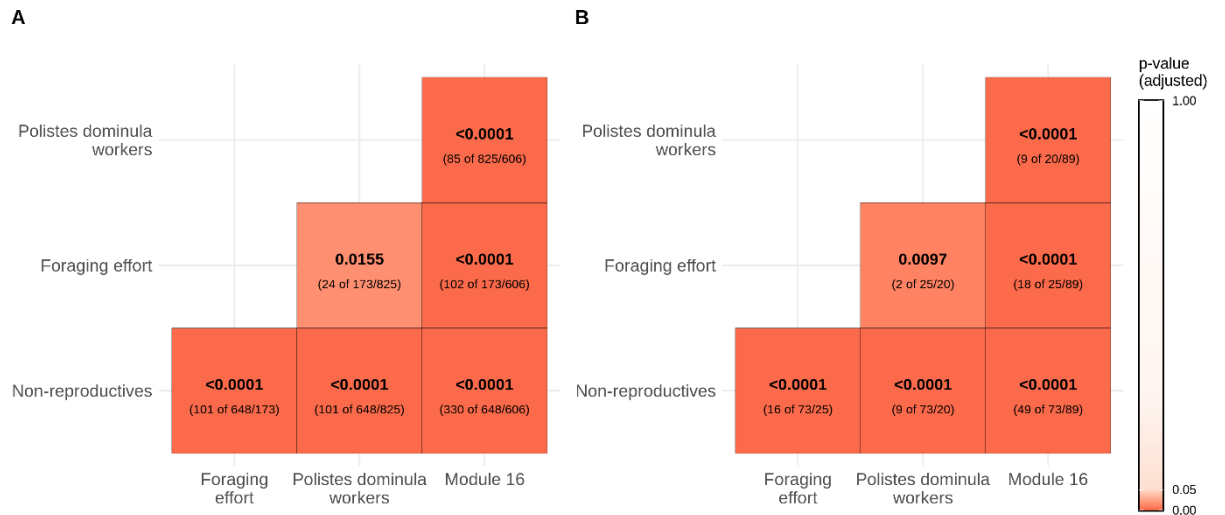
Figure 4.5. Hypergeometric overlap of (A) genes and (B) GO terms upregulated in treatments associated with indirect fitness strategies. Numbers outside parentheses: multiple comparisons-corrected hypergeometric p-value for the overlap. Numbers inside parentheses: number of features overlapping between the two groups, and the number of features in each group.
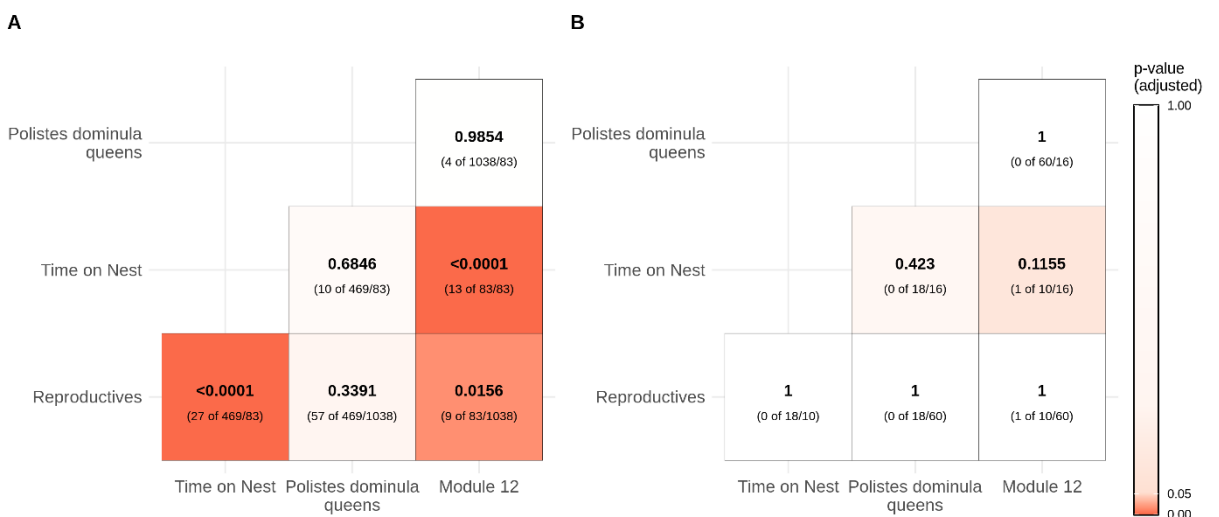
Figure 4.6. Hypergeometric overlap of (A) genes and (B) GO terms upregulated in treatments associated with direct fitness strategies. Numbers outside parentheses: multiple comparisons-corrected hypergeometric p-value for the overlap. Numbers inside parentheses: number of features overlapping between the two groups, and the number of features in each group.

Because of the large size of the background set of genes and GO terms, the expected overlap between sets of genes/terms was very small. Thus, although the overlaps between sets of genes and GO terms associated with indirect fitness were strongly significant (even after correction for multiple comparisons), the actual numbers of overlapping terms were relatively small—for example, the overlap between the 25 GO terms positively correlated

with foraging in *L. flavolineata* and the 20 GO terms upregulated in *P. dominula* workers consisted of just two common terms, yet was quite strongly significant ($p = 0.0049$) even after correction for multiple testing via the conservative Bonferroni method. Nonetheless, these overlap analyses appear to support the surprising conclusion that genes and pathways associated with indirect fitness, but not with direct fitness, are conserved between *L. flavolineata* and *P. dominula*.

4.3.6 Hypothesis 6: Evolutionary age of fitness-associated genes

To further test the hypothesis that genes associated with direct fitness should be more conserved than those associated with indirect fitness, I performed phylostratigraphic analysis to identify trends in the relative evolutionary ages of genes that were upregulated in different groups. Genes whose expression was queen-biased were significantly less likely to be of ancient origin than genes that were uncorrelated with time off-nest (**Figure 4.7A**, Fisher's exact test; odds ratio = 0.78, $p = 0.02$), whereas genes that were upregulated in non-reproductives were significantly less likely than neutral genes to be of recent (aculeate) origin (Fisher's exact test; odds ratio = 0.51, $p = 4.13 \times 10^{-7}$). These results therefore appear to corroborate the surprising result from my overlap analysis that 'indirect fitness genes' are more conserved than 'direct fitness genes' in *L. flavolineata*.
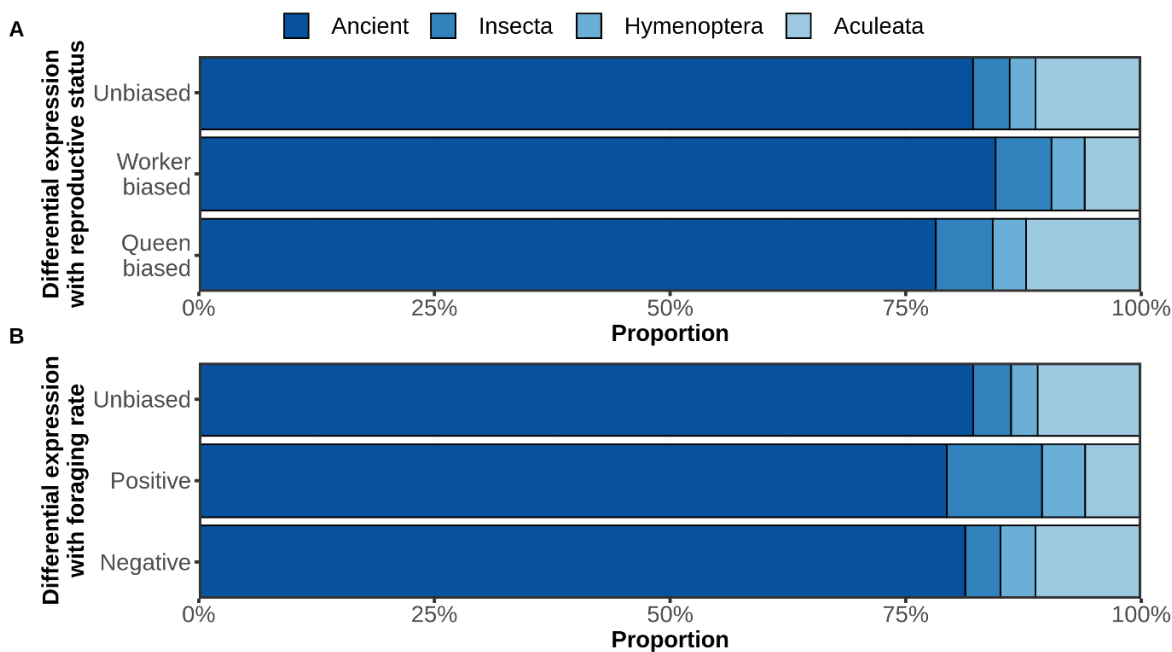
Figure 4.7. Proportion of genes of different estimated evolutionary ages among different sets of fitness-related genes. (A) Representation of phylostrata among genes that were upregulated in reproductives ('queen-biased'), non-reproductives ('worker-biased') and unbiased with respect to reproductive caste. (B) Representation of phylostrata among genes that were positively correlated, negatively correlated, and uncorrelated with foraging effort among non-reproductives.

98

I also compared the phylostratigraphic distribution of genes with respect to their correlation with foraging rate among non-reproductives. Genes that were positively correlated with foraging rate among non-reproductives were also significantly more likely to be worker-biased, and vice versa for genes that were negatively associated with foraging rate among non-reproductives. As such, I expected that foraging-associated genes would exhibit the same pattern of phylostratigraphic distribution as caste-biased genes, but this was not the case (**Figure 4.7B**): genes that were negatively correlated and uncorrelated with foraging effort showed extremely similar phylostratigraphic patterns, but genes that were positively correlated with foraging effort were significantly less likely to be of aculeate origin (Fisher's exact test; odds ratio = 0.51, p = 0.02), while being much more likely to have originated in non-hymenopteran insects (Fisher's exact test; odds ratio = 2.69, p = 9.1 x 10^{-5}) than genes that were unbiased with respect to foraging effort.

4.3.7 Hypothesis 7: Decomposing the effects of social and reproductive context

In a second experiment, I manipulated nests to generate individuals that were either solitary or social, and divided these individuals further in social foragers, social reproductives, solitary foragers, and solitary reproductives. Entering social context (social vs solitary) and reproductive status (reproductive vs forager) into an additive model, 643 genes were differentially expressed with reproductive status while just 4 were differentially expressed with social context, which suggests that social context is significantly weaker determinant of differential expression than reproductive context. The 352 genes and 62 GO terms that were upregulated in reproductives in this experiment overlapped strongly with those that were queen-biased in the experiment described in sections 4.3.1-4.3.4 (genes: two-sided hypergeometric p = 0.066; GO: two-sided hypergeometric p < 0.001), and likewise the 291 genes and 51 GO terms genes that were upregulated in foragers here overlapped strongly with those that were worker-biased in the previous experiment (genes: two-sided hypergeometric p < 0.001; GO: two-sided hypergeometric p < 0.001).

Comparing treatment groups in a pairwise fashion rather than in combined groups, however, the only strong signals of differential expression were between social foragers and the other three treatment groups, with >100 differentially-expressed genes in each comparison (**Figure 4.8**). These DEGs were of relatively small magnitude change in expression, however: a $\log_2$ fold-change threshold of 1.5 resulted in all signals of

differential expression being lost save for a single DEG between social foragers and social reproductives. The relative lack of differential expression between solitary foragers and solitary reproductives (who differed only in the behaviour that they were actively engaged in at the time of collection) suggests that reproductive context makes little difference in term of gene expression within a single individual that is plastically switching between reproductive and non-reproductive tasks. Similarly, solitary and social reproductives appear not to differ transcriptomically, which suggests that existing as part of a social group does not in and of itself induce changes in gene expression. Social foragers were the only individuals that differed substantially from any other treatment group in terms of gene expression; importantly, these were the only individuals in this experiment that lacked developed eggs, as well as exhibiting higher foraging rates than any other group (mean foraging effort = 0.6 vs. mean foraging effort ≈ 0.15 for the other three groups). This pattern of gene expression therefore indicates that the reproductive suppression induced by an individual's status as a non-reproductive worker within a group is a strong determinant of variation in gene expression in this species, but social context and reproductive context alone do not substantially affect brain transcriptomes.
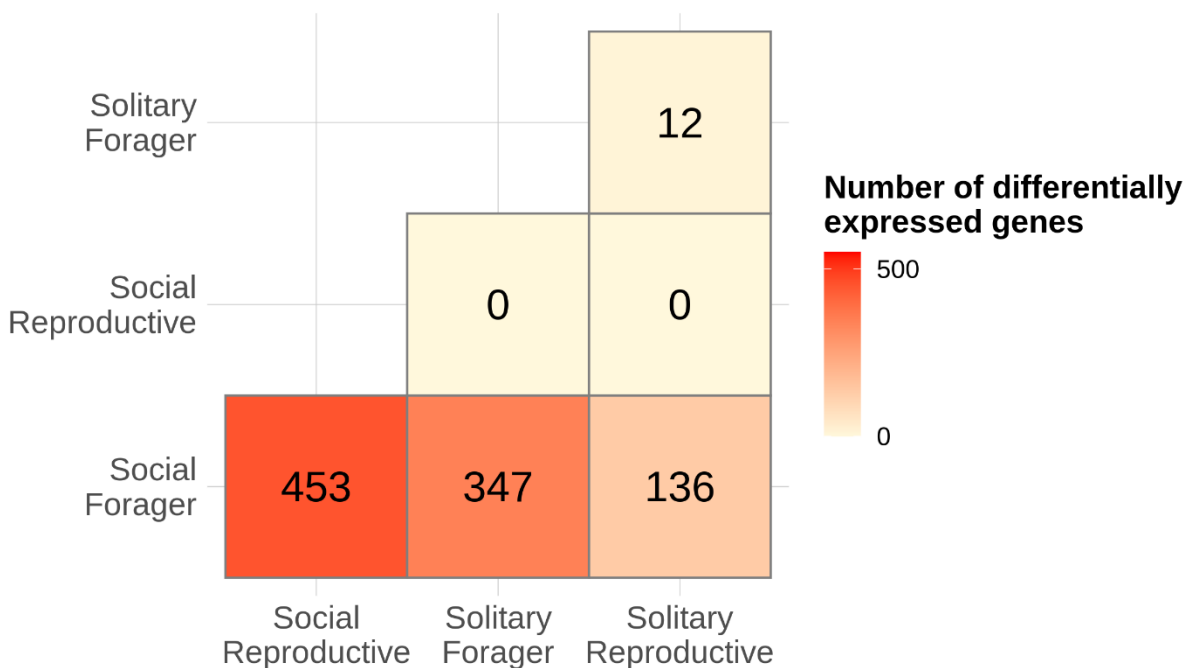
Figure 4.8. Number of genes differentially expressed between each pair of treatment groups, including individuals that were forced to breed solitarily.

4.4 Discussion

In this chapter, I have explored the brain gene expression profiles that underlie social organization in reproductively skewed groups of the facultatively social hover wasp *L. flavolineata*. The results presented here indicate that both caste and foraging effort structure brain gene expression variation in this species, the former much more strongly. Thus, although older workers in this species exhibit a substantially different phenotypic profile than that of younger workers, at the level of gene expression differences, workers and queens nonetheless form distinct groups at the transcriptomic level. Additionally, I identify two gene co-expression modules whose expression corresponds to investment into direct and indirect fitness respectively. My results also suggest an unexpected reversal of the predicted relationship between genes' fitness and their evolutionary age: I found that genes and processes that were associated with indirect fitness were substantially shared between *L. flavolineata* and *P. dominula*, which was not the case for genes and processes associated with direct fitness. The results that I have presented here represent the first in-depth study of the molecular basis of sociality in any vespid species outside of the Polistinae, and suggest potentially fruitful lines for future research in this independent origin of insect sociality.

4.4.1 Molecular signatures of alternative fitness strategies in *L. flavolineata*

L. flavolineata is a facultatively social species with societies governed by reproductive 'queues', and I therefore predicted that this species might lack the strong caste differentiation patterns that are typical of obligately social species. Contrary to this prediction, my results indicate that gene expression in *L. flavolineata* groups substantially resembles that in the social groups of obligately social insects, with the strongest signatures of within-group differentiation being found between queens and workers rather than being correlated with proxies for more graduated shifts, such as foraging rate or rank. While the strong signatures of differential expression that I identified between castes might be expected if I had sequenced ovarian tissue (which is expected to differ strongly based on whether an individual is mated and actively egg-laying), the fact that I found such large differences in brain transcriptomes is surprising. The behaviour of Rank 2 individuals in this species is intermediate between that of queens (Rank 1 individuals), which do not forage at all (Field & Foster 1999; Cant & Field 2001; Shreeves & Field 2002), and low-ranked workers, which perform the bulk of the group's foraging.

Nonetheless, the reproductive differences that separate queens and highly-ranked workers appear to be strongly reflected even at the level of the brain transcriptome. This surprisingly strong link between the brain and reproductive tissues is concordant with results obtained by previous studies that have found a similar link between ovarian development and brain gene expression in the facultatively social sweat bee *Megalopa genalis* (Jones et al. 2017; Kapheim et al. 2020). Feedback from reproductive tissues to the brain is also apparent in the honey bee literature (Amdam et al. 2006; Wang et al. 2009, 2010). Wang et al. (2010), for example, showed that surgical implantation of extra ovaries into honey bee workers' abdomens caused those workers to shift their caste-associated behaviour.

Although the strongest determinant of within-group variation in gene expression was caste, I was able to identify a substantial number (~250) of genes that were differentially expressed with respect to foraging effort. These genes, and the GO terms with which they were associated, overlapped significantly with caste-biased genes and GO terms as would be expected if decreasing foraging rate with increased rank reflects a shift toward direct fitness investment. Alternatively, it is possible that these genes instead reflect less abstract phenotypic differences—for example, differential expression with foraging rate might stem from the energetic demands of spending more time in flight or differences in age, rather than being directly linked to fitness *per se*. The fact that I was able to identify a larger number of genes differentially expressed with foraging rate than with within-group rank among non-reproductives (and that the latter set of genes was almost entirely subsumed by the former) may support this interpretation. Given that rank is (in theory) the strongest measure of how close an individual is to attaining a reproductive role, with foraging effort putatively a secondary consequence of rank, we might expect rank to be a stronger predictor of fitness-related gene expression than foraging itself.

An alternative interpretation, however, is that rank is in fact a relatively crude proxy of individual-level variation in proximity to the reproductive role—variation in the foraging rate of Rank 2 and Rank 3 individuals was very high (**Figure 4.1A**), which suggests that factors other than rank influence individuals' effort. Nest size is known to be one such factor (Shreeves & Field 2002; Cant & English 2006), but others might include the age, health and fecundity of the incumbent reproductive (Kokko & Johnstone 1999; Bridge & Field 2007), which are factors that I did not measure. If individuals modulate their foraging rate based on their estimation of expected future direct fitness gains, and such gains are affected by e.g. the projected mortality of the current reproductive, then foraging

rate might actually be a stronger proxy of individual-level alternative fitness strategies than rank.

Three more factors speak in favour of the hypothesis that gene expression variation with respect to foraging genuinely reflects individuals' fitness investment. First, the GO terms that were enriched among foraging-biased genes (both positively and negatively) related to relatively basal metabolic and developmental processes, a pattern similar to that observed in caste-biased GO terms in both *L. flavolineata* and *P. dominula* (Chapter 3; Standage et al. 2016), which seems to belie the possibility that these genes solely associated with such specific behaviours as e.g. flying and foraging. Second, individuals that were promoted from Rank 3 to Rank 2 shifted gene expression to match their new rank, which strongly suggests that gene expression variation is not primarily structured by independent factors such age. Third, the mean foraging rate for Rank 2 individuals was closer to that of reproductives than that of the lowest-ranked individuals (**Figure 4.5A**). This suggests that if within-group gene expression differences primarily reflect actualized foraging effort, then I should have observed stronger differences between Rank 2 individuals and Rank 4/Rank 5 individuals than between Rank 2 individuals and Rank 1 queens, which was not the case. Overall, these factors support the hypothesis that gene expression variation among *L. flavolineata* workers does reflect individual-level differences in direct vs indirect fitness investment.

4.4.2 Evolutionary conservation of genes associated with alternative fitness strategies

I found support for the rather counterintuitive conclusion that genes that are associated with indirect fitness in *L. flavolineata* are more likely to be conserved than genes that are associated with direct fitness, a result that runs directly counter to predictions and empirical data for other social insect species. This result is nonetheless corroborated by two separate findings: first, that worker-biased but not queen-biased genes and GO terms were conserved between *L. flavolineata* and *P. dominula*, which represent two independent origins of Vespid sociality; and second, that queen-biased genes were substantially more likely than worker-biased genes to be of recent (Aculeate) evolutionary origin, while worker-biased genes were more likely to be of ancient (Metazoan or earlier) origin.

When I applied the same phylostratigraphic analysis to the *P. dominula* data generated in Chapter 3, I generated the results predicted from the existing literature, with queen-biased genes more likely than neutral genes to be ancient, and worker-biased genes less likely to be ancient (**Figure S4.10**). This suggests that the unusual result that I obtained is not a methodological artefact. The completeness score of the *L. flavolineata* genome that I used for phylostratigraphy is also high (96.7% complete single-copy hymenopteran BUSCOs; Appendix B), which likewise suggests that these results are unlikely to stem from a failure to identify homologues. I therefore conclude that *L. flavolineata* seems to exhibit a genuinely unusual pattern of genomic conservation with respect to genes' correlation with caste. To my knowledge, no stenogastrine species other than *L. flavolineata* has been the subject of gene expression analysis or any other focused molecular study, which means that it is not currently possible to assess to what degree the results that I have generated here are general to other species sharing the same origin of sociality as *L. flavolineata*.

It is worth noting that the phylostratigraphic differences that I identified were somewhat different when assessing genes that were differentially expressed with respect to foraging rate rather than caste-biased genes, despite that the fact that these two sets of genes were significantly overlapping: genes that were downregulated with foraging rate ('direct fitness genes') among non-reproductives exhibited a phylostratigraphic distribution that was very similar to that of unbiased genes, while genes that were upregulated with foraging rate ('indirect fitness genes') were much more likely to be of Hexapod origin than would be expected by chance. While the significance of genes that are of Hexapod origin is difficult to interpret, it remains the case that genes associated with indirect fitness in this comparison were less likely than other genes to be of very recent origin, which again runs counter to the predicted pattern.

4.4.3 Molecular correlates of social and reproductive context

I found little evidence that social context significantly affects reproductive individuals' brain gene expression in this species: in Experiment 2, no genes were differentially expressed between social reproductives and solitary reproductives, nor between social reproductives and solitary foragers. It therefore seems to be the case that reproductive individuals in this species are relatively transcriptomically insensitive to social context, although such a result is surprising given that solitary reproductives must switch between reproductive and foraging tasks in a manner that social reproductives do not. By contrast, social foragers were fully non-reproductive and exhibited strong signals of differential

expression from all other treatment groups, and genes and GO terms that were upregulated in reproductives or foragers in this experiment overlapped strongly with those that were caste-biased in the nests studied in Experiment 1.

Surprisingly, I was only able to identify 12 DEGs between solitary *L. flavolineata* individuals that were actively foraging or egg-laying in the same experiment, and this set of genes did not overlap at all with caste- or foraging-biased genes. This result may represent tentative evidence against a role for a solitary ground-plan in this lineage; alternatively, it could be the case that acts of foraging and of egg-laying are not sufficiently temporally distinct in solitary *L. flavolineata* individuals for significant changes to occur in the brain transcriptome between behaviours. A possible future avenue of analysis in this species might be to obtain transcriptomes for solitary individuals that are in the process of nest-building. Nest founding is a non-reproductive activity that seems to require significant investment of time and energy in *L. flavolineata* (Field et al. 1998), and such individuals might therefore be better representatives for a 'solitary non-reproductive' phenotype.

4.4.4 Conclusions

Facultatively social insects are models for the earliest stages of insect social evolution. Social groups in the facultatively social hover wasp *L. flavolineata* are characterised by linear hierarchies, and individuals' positions in the resulting 'queue' dictates their investment into indirect vs direct fitness strategies, a system that should lend itself well to studies of the molecular correlates of alternative fitness strategies. In this chapter, I have shown that gene expression in *L. flavolineata* colonies is structured by both actualised and potential direct fitness, and that individuals are able to facultatively shift their gene expression profiles to match changes in direct fitness prospects. I have also provided evidence in favour of the surprising finding that genes associated with indirect fitness are more strongly conserved than genes associated with direct fitness in this species. The work I have presented here represents the first in-depth bioinformatic analysis of a social wasp outside of the Vespidae—my results indicate that this little-studied taxon could be a fruitful model for future studies of the genetic basis of reproductive investment.

Chapter 5 – A review of patterns of reproductive plasticity in the major evolutionary transition to superorganismality

The contents of this chapter have been published as: Taylor, B. A., Reuter, M., & Sumner, S. (2019). Patterns of reproductive differentiation and reproductive plasticity in the major evolutionary transition to superorganismality. *Current opinion in insect science, 34*, 40-47.

5.1   Introduction

In Chapters 2-4 of this thesis, I have examined the mechanisms and consequences of reproductive plasticity in those social insect species that retain the ability to plastically transition from workers to queen phenotypes. As stated in Chapter 1, a significant part of my rationale for focusing on the expression of reproductive plasticity is the fact that the most derived, well-studied, and ecologically successful social insect taxa are 'superorganisms' that are defined by the loss of this plasticity (Hölldobler & Wilson 2009; Kennedy et al. 2017; Boomsma & Gawne 2018). Here, I return to the topic of the importance of reproductive plasticity and its suppression in shaping social evolution.

The evolution of life on earth has been shaped by a number of major transitions in individuality (Szathmáry & Smith 1995; Szathmáry 2015; West et al. 2015). Each of these transitions has involved the formation of new, more complex individuals from the cooperation of previously independently-replicating units (Szathmáry & Smith 1995; West et al, 2015). Canonical major evolutionary transitions include the transition from independently-living unicells to multicellular organisms (Hanschen et al. 2015; Sebé-Pedrós et al. 2017), the formation of the eukaryotic cell from the conjunction of prokaryotic cells (Szathmáry 2015; Vellai & Vida, 1999), and the evolution of highly cohesive superorganismal insect societies from solitary ancestors (Kennedy et al. 2017; Boomsma & Gawne 2018).

Despite having played a large role in the evolutionary history of life, major evolutionary transitions in individuality have proven difficult to study. Transitions such as the formation of the genome and the formation of the eukaryote cell occurred in the distant past, and probably only once each. By contrast, the transition to complex insect sociality has occurred more recently, and multiple times independently, in the corbiculate bees,

termites, ants and vespine wasps (Hughes et al. 2008; Boomsma & Gawne 2018). Superorganismal colonies are relatively easily decomposed and manipulated, and therefore provide excellent models with which to study major evolutionary transitions and unravel the proximate and ultimate foundations that underlie major shifts in individuality (Kennedy et al. 2017).

In this chapter, I review the role that reproductive plasticity has played in the transition from a solitary to a superorganismal lifestyle. A fundamental aspect of each major transition is the functional differentiation and division of labour among the lower-level units that together constitute a new higher-level unit of individuality (Szathmáry & Smith 1995; Szathmáry 2015; West et al. 2015). The evolution of multicellularity, for example, has involved a transition from phenotypically flexible unicells to higher-level organisms with an obligate division of reproductive labour between germ and somatic cells (Hanschen et al. 2015). This is mirrored in the transition to insect superorganismality, which is defined by a reduction in phenotypic plasticity and a division of reproductive labour between reproductive queens and non-reproductive workers (Kennedy et al. 2017; Boomsma & Gawne 2018). I argue that the emergence of a fixed reproductive division of labour does not represent a loss of reproductive plasticity among lower-level units so much as a transfer of plasticity between levels of selection, from the individual to the colony.

5.2   Social organisation and reproductive plasticity

Insect sociality covers a broad spectrum of social systems of varying levels of social complexity. While a range of classification systems have been proposed (Sherman et al. 1995; Costa & Fitzgerald 2005; Boomsma & Gawne 2018), most insect species broadly fall into one of three categories: Solitary breeding represents the likely ancestral state for all social lineages; cooperative breeding (or simply 'sociality') involves the formation of social groups with division of reproductive labour but in which most individuals retain a significant degree of reproductive flexibility; finally, superorganismal species exhibit an irreversible division of reproductive labour, cementing the colony as the unit of selection. In Section 1.2.1 of this thesis, I outlined the parallelisation of reproductive and non-reproductive activities that defines the transition to insect sociality, which results in reproductives being freed to dedicate a larger proportion of their time and energy towards reproduction. However, as long as workers are reproductively plastic, the partitioning of

reproductive labour is not complete, meaning that workers may be incentivised to withhold labour or even attempt to usurp queens. Here, I outline the broad trends in reproductive plasticity across the major transition from non-superorganismal insect sociality to superorganismality (**Figure 5.1**).
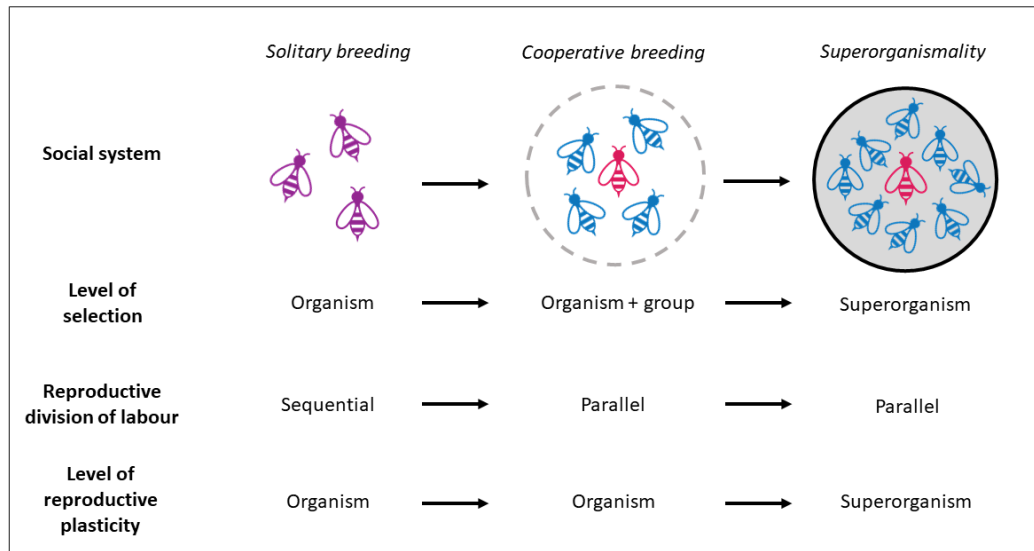
Figure 5.1. The evolutionary transition from solitary to cooperative breeding involves a shift from sequential execution of reproductive and non-reproductive tasks by individuals to parallel execution of tasks by groups of cooperating individuals, but with little or no loss of reproductive plasticity: at this stage in the major transition, most or all group members are able to switch between tasks if the opportunity arises. The transition from cooperative breeding to superorganismality is defined by a dramatic and putatively irreversible loss of reproductive plasticity at the level of the lower-level organism, but the simultaneous generation of a higher-level superorganism with its own colony-level plasticity. Arrows indicate the likely ancestral trajectory of modern superorganismal species; note that I do not mean to imply an inevitable evolutionary trajectory for extant solitary or cooperatively breeding species.

5.2.1 Losses of plasticity in non-superorganismal social insects

Reproductive and non-reproductive phenotypes within typical non-superorganismal insect societies are significantly more stable than those expressed within the lifetime of a solitary individual, yet substantial task plasticity persists in these societies. Non-reproductives are capable of upregulating reproductive traits in response to the loss of a dominant reproductive across cooperatively breeding bees (Michener & Brothers 1974; Breed & Gamboa 1977), wasps (Gadagkar 1987; Strassmann et al. 2004) and termites (Roisin 1990). Although there is some evidence that the rate at which individuals are able to transition between behavioural (Michener & Brothers 1974; Breed & Gamboa 1977)

or physiological (Roisin, 1990) roles may be limited in non-superorganismal social insect societies, the results I present in Chapter 2 suggest otherwise: there, I show that *P. dominula* successor queens are able to achieved an ovarian and behavioural profile nearly identical to that of their predecessors in a matter of days. There is also evidence that non-reproductive individuals may experience a loss of reproductive plasticity as they age, for example in *Polistes canadensis* (Sumner et al. 2010) and *Metapolybia cingulata* (West-Eberhard 1978) paper wasps. The transition from solitary to cooperative breeding thus may involve some loss in reproductive plasticity, but of a minor degree: the majority of individuals in a cooperatively breeding society will remain reproductively totipotent for the majority of their adult lives.

5.2.2 Transfer of plasticity from the individual level to the colony level in the superorganismal society

The major transition from cooperative breeding to superorganismality is defined by the evolution of obligate and irreversible differentiation of reproductive phenotypes and non-reproductive phenotypes, cementing the reproductive division of labour (Kennedy et al. 2017; Boomsma & Gawne 2018). Superorganismal castes are determined during development, after which non-reproductive individuals are unable to transition to a fully reproductive role. This reduction in reproductive plasticity commits non-reproductives to indirect fitness strategies and shifts the target of selection from the individual to the colony. Selection will then favour traits that maximise indirect fitness gains, opening the door for strong task specialisation. Accordingly, superorganismal castes are typically morphologically differentiated, exhibit a reversal of the usual negative trade-off between reproductive activity and longevity (Von Wyschetzki et al. 2015; Blacher et al. 2017), and may include specialised phenotypic subdivisions of the non-reproductive caste (Wills et al. 2018).

Adult individuals in superorganismal colonies are much less reproductively plastic than adult individuals in cooperative species, which themselves may be somewhat less reproductively plastic than solitary individuals. It is therefore tempting to infer a relatively straightforward pattern of decreasing reproductive plasticity across the spectrum of insect sociality from solitary to cooperatively breeding to superorganismal insect species. Such a pattern only holds, however, when one focuses exclusively on a single level of individuality, namely that of the individual adult organism. The nature of

a major evolutionary transition is that it results in the generation of a higher level of individuality: reproductive plasticity must therefore be considered at both the level of the individual organisms that constitute the superorganismal society and the higher level of the superorganism itself.

There are significant analogies between the organisation of complex multicellular organisms and the organisation of superorganisms: the queen is the reproductive tissue of a superorganism, while workers are its somatic tissue (Hölldobler & Wilson 2009; Gadau 2015; Helanterä 2016), and superorganismal societies exhibit analogues to individual-level multicellular traits such as homeostasis (Modlmeier et al. 2019); physiology (Johnson & Linksvayer 2010); and even 'cancer' (Oldroyd 2002). We can likewise understand superorganisms as exhibiting an analogous form of reproductive plasticity. Just as a solitary multicellular organism can be said to possess reproductive plasticity if it is able to facultatively vary its investment into reproductive and non-reproductive traits, the superorganism is reproductively plastic if it has the capacity to facultatively vary investment into the production of workers and sexual offspring. The ability of a superorganism to engage in this kind of plastic response is a function of its sociogenome, a composite of the genetic traits of all individuals that together constitute the superorganism and the interactions of those traits to produce colony-level feedback loops (Linksvayer et al. 2009; Johnson & Linksvayer 2010).

Sociogenome-mediated interactions between constituent individuals allow superorganisms to plastically mediate reproductive investment in much the same way that multicellular organisms do. Though the individuals within a superorganism are not perfectly analogous to the cells of a multicellular individual—the latter being fully clonal and therefore almost entirely insulated from internal reproductive conflict, while the former retain the potential for conflict since workers are usually not clonal and are often capable of producing males despite having limited fertility—this high degree of task parallelisation and specialisation has facilitated similar adaptations in each system. Like multicellular organisms, superorganismal colonies may upregulate investment into reproduction (i.e. the production of sexuals) in response to seasonality and may store energy within non-reproductives to later transfer to sexual brood (Tschinkel 1998; Hart & Tschinkel 2012). Both multicellular organisms and superorganisms typically transition from somatic to reproductive investment having reached some growth threshold (Smith et al. 2017). Ant colonies appear to possess a form of terminal investment, switching to the production of sexuals in response to queen mortality (Hölldobler & Wilson 1990;

110

Heinze & Schrempf 2012). Honeybee colonies respond to the loss of a queen by rearing replacement queens (Hatch et al. 1999). Thus, while superorganismal colonies lack the individual-level reproductive plasticity necessary for non-reproductives to directly replace reproductives, at the superorganism level these colonies are able to plastically modulate their investment into 'somatic' activities such as growth vs 'reproductive' activities such as the production of sexuals.

Viewed at a single level of selection, the transition to superorganismality might be generalised as a simple reduction in reproductive plasticity, producing individuals that cannot adapt their reproductive roles in response to changing environments. However, such an approach only holds if the 'individual' that matters remains the same throughout this major transition. In fact, the nature of major evolutionary transitions is that they produce new units of selection, and these units (once sufficiently coherent) can and should be considered individuals in their own right (West et al. 2015). As a result, a model of the role of reproductive plasticity in the evolution of insect sociality that focuses solely on the level of the individual multicellular organism is necessarily incomplete. Just as it would be wrong to argue that an adult mammal lacks phenotypic plasticity because it is largely comprised of unipotent cells, so too must we recognise that the transition to superorganismality can only be understood by considering the superorganism as a unified individual.

5.3   Conclusions

The major evolutionary transition to insect superorganismality is one of our best models for understanding the way in which lower-level units are able to come together to form new (higher) levels of biological complexity. Here I have briefly summarised how the shifts in reproductive division of labour have made this radical shift in complexity possible, while also emphasising the need to consider multiple levels of selection: while superorganismality represents a loss of plasticity at the level of one form of individual (the multicell), it involves a recapitulation of plasticity at a higher level of biological individuality (the superorganism itself). The patterns of differentiation involved in this transition are likely to be of general relevance to other shifts in individuality, such as the transition to multicellularity. In order to fully appreciate these commonalities, however, social insect researchers must embrace the status of the superorganismal colony as an individual in its own right.

Chapter 6 – Conclusions

6.1   Key findings

In this thesis, I have attempted to address two key questions regarding the evolution of social insect reproductive plasticity. First, how do reproductively skewed societies maintain social cohesion when most individuals retain the ability to assume a reproductive role? And second, how is reproductive plasticity expressed at the levels of phenotype and gene expression in such societies?

Chapter 2 addressed the question of how societies can avoid within-group conflict following a major social disruption using an in-depth analysis of the phenotypic process of queen succession in single-foundress *P. dominula* groups. This chapter concluded that in the context of competition between siblings over queen succession, selection has favoured the evolution of an age-based succession system, despite the high value of the queen role and the fact that older individuals appear not to be inherently physiologically advantaged to transition to a reproductive role. Chapter 3 examined the gene expression impacts of reproductive plasticity, asking how this plasticity is reflected in the transcriptomes of the individuals that express it, as well as the transcriptomes of those individuals' groupmates. That chapter concluded that the apparently highly efficient and pacific phenotypic process described in Chapter 2 is nonetheless accompanied by substantial colony-wide disruptions to individuals' gene expression profiles, which only partially track measured phenotypic traits. Taken together, these chapters provide two novel insights into the nature of phenotypic plasticity: first, they demonstrate the mechanisms by which species with reproductively plastic castes are able to maintain social cohesion in the face of high potential within-group conflict; and second, they reveal potential limitations to this apparent social cohesion which might have influenced the loss of reproductive plasticity in other social lineages.

Chapter 4 addressed the related question of how variation in individual-level investment into alternative reproductive strategies is expressed at the level of the transcriptome and the genome by investigating gene expression data from *L. flavolineata*, a facultatively social species in which 'workers' express a great deal of variation in their fitness-related behaviour. The results of this chapter produced some unexpected conclusions: first, that brain gene expression among group members in this species is explained better by reproductive status than by dominance rank or foraging activity; secondly, that gene

expression in 'queens' differs little between reproductive and non-reproductive behaviours; finally, and most surprisingly, that genes associated with indirect fitness in this species appear to be more strongly conserved than those associated with direct fitness. Finally, in Chapter 5, I returned to the topic of the importance of reproductive plasticity in social insect evolution, proposing that in superorganismal species, the plasticity that I have discussed throughout this thesis is not so much lost as it is delegated to a higher level of biological complexity.

Taken together, the results of this thesis shed light on the mechanisms, consequences and limitations of reproductive plasticity, a trait that is of central importance in our understanding of both insect social evolution and the evolution of cooperative living more generally.

6.1.1 Costs and benefits of reproductive plasticity

Full reproductive plasticity is the basal state for insect sociality, as well as for all solitary insect species. The ability to switch facultatively between reproductive and non-reproductive states allows individuals to modulate their reproductive output in response to social or environmental cues (Javoiš & Tammaru 2004; Koštál 2006; Cotter et al. 2011), but may impose costs in terms of time and energy (Gabriel 2006; Flatt & Heyland 2011; Murren et al. 2015). In social groups whose structure is reliant upon reproduction being monopolised by a small proportion of individuals, the retention of reproductive plasticity may also impose a group-level cost if it results in conflicts of interest between reproductive and non-reproductive individuals (Boomsma & Gawne 2018). Loss of reproductive plasticity is a key precursor to the evolution of the most complex forms of insect sociality, and thus understanding the costs and benefits of this trait is a prerequisite for any complete explanation of social insect evolution.

Several of my results seem to contradict the suggestion that reproductive plasticity is costly for social insect groups. I found that in *P. dominula*, queen succession is rapid and relatively conflict-free, with a simple gerontocratic convention seemingly obviating the need for costly fighting. A similar succession mechanism appears to exist in *L. flavolineata*, and my findings suggest that individuals are able to adapt more-or-less completely to the changes in foraging rate and gene expression that accompany a transition between within-group dominance ranks. My results therefore suggest that selection has produced robust mechanisms by which groups are able to cope with the

potential within-group conflicts that arise from reproductive plasticity. Such results should not be surprising. Though non-superorganismal social insect species are frequently referred to as 'simple' or 'primitive' (Linksvayer & Johnson 2019), the widespread distribution of such species speaks to the evolutionary viability of this form of insect sociality, a fact that is not invalidated by the emergence of superorganismality in other taxa.

Nonetheless, my results do suggest possible limitations to the mechanisms that insect societies use to mediate conflict in the presence of reproductive plasticity. One of these is the transcriptomic disruption following queen loss in *P. dominula* that I identified in Chapter 3. In that chapter, I speculated that this colony-wide response to queen loss might reflect either a form of cryptic stress or, alternatively, a pre-emption of future instances of queen loss, in which individuals become primed to cope with future disruption. In either case, the changes that I observed might well be expected to confer an energetic load upon the individuals in which they were exhibited, which might be reflected in terms of e.g. reduced future lifespan or reduced ability to suppress conflict. Intriguingly, unpublished data communicated to me by Joan Strassmann and David Queller indicate that in *Polistes*, colonies' ability to accommodate queen loss depletes over multiple rounds of succession, with groups descending into a fully conflict-driven succession process after losing several queens in a row (Strassmann, pers. comm.). This transition to a conflict-based succession process might indicate that repeated rounds of queen loss cause transcriptomic changes to accrue to the point that gerontocracy no longer functions as a mediating mechanism.

Another limitation of reproductive plasticity is exemplified by *L. flavolineata* societies, in which individuals invest less in foraging as their future direct fitness prospects increase (Field et al. 1998, 2000). This means that Rank 2 individuals, which perform only a small proportion of the group's foraging, contribute relatively little to social functioning. This situation contrasts to that in superorganismal societies wherein workers are sterile and therefore lack an incentive to withhold their labour, a fact that increases the efficiency of the colony as a whole (Hölldobler & Wilson 2009; Boomsma & Gawne 2018).

My results provide little evidence for substantial individual-level costs of expressing reproductive plasticity. The fact that the individuals I observed were able to transition between roles within a few days (from worker to queen in *P. dominula* and from lower to higher rank in *L. flavolineata*) seems to indicate that the expression of plasticity is not

especially costly. I also found few gene expression differences between solitary *L. flavolineata* individuals that were in the act of egg-laying vs those that were in the act of foraging, which suggests that the degree of molecular restructuring necessary for switching between these roles is minimal and therefore presumably of little energetic cost. On the other hand, the transcriptomes of *P. dominula* replacement queens that I analysed had only transitioned part of the way towards being fully queen-like, which might indicate that caste succession is a more involved process than is apparent from phenotypic data alone. Additionally, *P. dominula* workers are uniformly unmated (Turillazzi 1980; Strassmann et al. 2004), and my results indicate that the same is true for *L. flavolineata*. As a result, replacement queens must seek out a mate before they will be able to produce female offspring, a process that can take several weeks in *P. dominula* (Strassmann et al. 2004). Mating occurs off-nest in both *P. dominula* (Beani et al. 1992) and *L. flavolineata* (Truillazzi & Francescato 1990), and as a result replacement queens presumably suffer some risk of injury or predation when they leave the nest to mate, placing an additional constraint on the efficiency of queen replacement.

Overall, my results indicate the individual-level costs of expressing reproductive plasticity are probably minor. The individuals I observed were able to transition between roles rapidly and with little apparent cost in terms of time or social disruption. However, my results do expose possible group-level costs of plasticity, either in the form of reductions in workers' foraging effort as they become more likely to express a transition to the queen role, or in the form of transcriptomic disruption following the loss of a queen even in a low-conflict context. This work therefore supports the notion that the principal benefit of the transition to superorganismality is a reduction in within-group conflict. Notably, this is less likely to be the case for other major evolutionary transitions, such as the transition to multicellularity, because multicellular organisms are typically composed of clonal cells and so already exhibit complete alignment of within-group (i.e. between-cell) fitness interests. In superorganisms, however, relatedness between subunits is only partial, and therefore suppression of conflicts of interest is expected to be a precursor to major gains in reproductive efficiency, rather than such gains following directly from a loss of plasticity (Boomsma 2009, 2013). My result support this expectation.

6.1.2 The importance of basal and derived mechanisms in social evolution

A prominent question in the field of social evolution is the importance of conserved genes and mechanisms relative to those that are novel or derived in the evolution of social phenotypes (Simola et al. 2013; Rehan & Toth 2015). An emerging consensus is that more basal mechanisms may be of relatively greater importance in the expression of reproductive (queen-specific) phenotypes, while more derived mechanisms will be of greater importance in the expression of non-reproductive (worker-specific) phenotypes (Gadagkar 1997; Linksvayer & Wade 2009; Van Dyken & Wade 2010). The results I obtained for the two species upon which I focus in this thesis produced conflicting results as regards this hypothesis. In *P. dominula* I found the expected pattern of evolutionary ages, with worker-biased genes being more derived on average than non-biased genes and queen-biased genes being on average more conserved; this result is in line with previous studies in *Polistes* paper wasps (Ferreira et al. 2013). My analysis of *L. flavolineata*, by contrast, produced a result that was directly opposite to that predicted: worker-biased genes were more likely to be of ancient origin than queen-biased genes.

A corollary to the hypothesis that indirect fitness functions will be relatively derived, and direct fitness functions relatively ancient, is that queen-biased genes should share more commonalities than worker-biased genes across independent origins of insect sociality; that is, it is *prima facie* more likely that there will be a shared evolutionary 'toolkit' for direct fitness than indirect fitness (Toth & Rehan 2017). My results for *L. flavolineata* again contradict this prediction. While worker-biased genes and GO terms in this species overlapped significantly with those that were worker-biased in *P. dominula*, this was not the case for queen-biased genes and terms. Although not conclusive, my results therefore indicate that the relationship between the evolutionary rates of genes and those genes' association with alternative fitness strategies may be less consistent than previously assumed. Whether the unexpected pattern that I identified in *L. flavolineata* would also be found in others members of the Stenogastrinae is unclear, since no other stenogastrine genome has been sequenced; my results therefore also speak to the continued need for greater phylogenetic diversity in our studies of insect sociality.

Given that solitary breeding is the basal state for all origins of insect sociality, we might also reasonably expect that genes that contribute to social behaviour generally (i.e., genes whose expression is upregulated in social vs solitary contexts) should be less strongly

conserved than those that contribute to solitary behaviour in the same species. My attempt to assess this hypothesis in *L. flavolineata* was inconclusive, as I was unable to identify strong signatures of social context in reproductive individuals of this species. I was similarly unable to identify significant gene expression differences between reproductive and non-reproductive contexts in solitary individuals of this species. These results indicate that the magnitude of transcriptomic changes experienced by solitary individuals throughout their adult lives is probably very minor relative to the degree of transcriptomic variation experienced by social individuals as they as they move up social queues.

I also note here that in almost all the comparisons that I have performed in this thesis, either between castes or between individuals with different fitness interests (e.g. different ranks), the gene ontology results have been dominated by terms related to metabolic, developmental or biosynthetic processes. As all of these are functions that are expected to be relatively strongly conserved (Dearden et al. 2006; Peregrín-Alvarez et al. 2009; Hinman & Cheatle Jarvela 2014; Bannerman et al. 2018), these results might be considered evidence in favour of the idea that highly-conserved 'toolkit genes' are of particular importance in determining caste expression in non-superorganismal social insects such as *P. dominula* and *L. flavolineata*. However, given that such terms tend to dominate gene ontology results relating to caste differentiation even in superorganismal species (e.g. Elango et al. 2009; Morandin et al. 2015; Warner et al. 2019), this result appears to reflect a general feature of the mechanisms that underlie caste expression, rather than being specific to those taxa that possess reproductive plasticity.

Overall, my results support the notion that there is an association between the role of genes in direct or indirect fitness strategies and the evolutionary trajectory of those genes. The direction of this association was not consistent in my data, however, and in some cases was opposite to that expected. While this is not entirely surprising given that more basal taxa are expected to exhibit weaker genomic signals of sociality, my results nonetheless highlight the importance of testing sociogenomic hypotheses outside of the small group of taxa that have been the dominant focus of research to date. In particular, the generation of more vespid genomes from species (both solitary and social) outside of the Vespinae and Polistinae would not only have the potential to reveal unexpected genomic patterns such as those that I found in *L. flavolineata*, but would also strengthen existing analyses by providing much needed points of comparison with which to estimate e.g. rates of gene sequence evolution within this clade.

6.1.3 Age as a cue for reproductive dominance in social insect societies

Age-based division of labour appears to be a common feature of social insect societies. In superorganismal societies, where reproductive division of labour is fixed, age plays a role in the division of non-reproductive labour among workers across many taxa, including honey bees (Johnson 2008) and several species of ants (e.g. Camargo et al. 2015; Bernadou et al. 2015) and termites (e.g. Li et al. 2015; Yanagihara et al. 2018). In these species the relationship between age and division of labour is thought to stem from the fact that, for maximum colony efficiency, riskier activities that take place further from the shelter of the nest should be performed by older individuals that are closer to the end of their natural lifespans (Tofilski 2002; Camargo et al. 2015).

Age is also important in societies with reproductive plasticity, but the causes of this importance are less clear. In both *P. dominula* and *L. flavolineata*, for example, age reliably and positively predicts reproductive succession, with older individuals succeeding ahead of younger ones (Chapters 3 and 4). Why should age predict the capacity of individuals to assume a reproductive role in these species? One possibility is that younger individuals are subfertile and therefore cannot efficiently take on a reproductive role. Pardi (1948), for example, theorized that younger individuals in *P. dominula* are unable to compete for a reproductive role as a result of underdevelopment of their ovaries. The results I presented in Chapter 2, however, show that this is not the case: ovarian development did not correlate with age until after the loss of a queen, which indicates that age-based differences in reproductive physiology are the consequence, rather than the cause, of gerontocracy.

In many vertebrate taxa, older individuals are dominant as a result of greater size and/or experience, neither of which factors are likely to be applicable in social insects. Experience is not expected to play a significant role in queen fitness, and younger workers tend to be larger and healthier than older ones (Pardi 1949; Reeve 1991). Although there was not a strong relationship between size and age in the colonies I examined, this may have been because the overall age range that I assessed was relatively small. Notably, the social insect species in which younger individuals tend to inherit are those in which succession is conflict-determined, possibly because younger and therefore larger workers are able to physically dominate their older sisters (Tsuji & Tsuji 2005).

Intriguingly, in Chapter 2 I did identify a negative interaction between size and change in individual-level antennation rate following queen loss, which I interpreted as indicating

that age was a poorer predictor of antennation rate change for larger individuals. This might indicate that larger individuals attempted to 'queue-jump' by competing for dominance even while young, although I lack the data to substantiate this possibility. The fact that the successor queen of a colony was not always the oldest individual does suggest that factors other than age may influence the strength of gerontocracy in *P. dominula,* as is also thought to be the case in *L. flavolineata* (Bridge & Field 2007). However, my failure to identify a relationship between size and queenness makes the possibility that larger individuals are potential queue-jumpers seem unlikely.

If older individuals are not larger, more dominant, or more reproductively developed, why do both *P. dominula* and *L. flavolineata* exhibit a strongly gerontocratic succession mechanism? In Chapter 2, I speculated that age is a purely conventional marker, i.e. that age is not intrinsically advantageous but is nonetheless valuable as a signal that cannot easily be faked. In this context the value of age is that, as long as wasps are able to individually recognize their nestmates, gerontocracy is an easy rule to follow. While specific markers of age such as size might not always be easy to discriminate, a wasp can assess the age of her nestmates relative to herself by a simple heuristic: was this individual already present when I eclosed? This rule alone is enough to ensure that all individuals are clear as to their own position within the gerontocratic hierarchy. Indeed, if individuals in *P. dominula* employ a simple heuristic such as this (rather than retaining information about the precise ages of all other wasps), this would explain my finding that the process of reproductive succession is insensitive to the absolute age gap between the two oldest wasps in a group. The size of this gap might be expected to be influential upon succession conflict if individuals care about the absolute age of their sisters, but not if they only assess age relative to themselves.

While the simplicity of age as a marker by which to identify dominance rank explains why this trait might be a good candidate succession mechanism, it still does not explain the direction of this convention. Age is an equally honest signal whether succession is gerontocratic or 'infantocratic' (favouring succession of younger individuals), so why should gerontocracy be the apparent norm? Bridge & Field (2007) suggest an elegant solution to this question: because workers are always older than the larvae that they rear, in an infantocratic system workers might be incentivised to under-nourish (and thereby induce subfertility in) larvae, reducing the ability of those larvae to assume the reproductive role and thereby effectively removing competition for that role for the

provisioning workers themselves. By contrast, in a gerontocratic system workers always rear individuals that will enter the dominance hierarchy at a lower position than the workers themselves, thereby removing this conflict of interest between larvae and adult workers.

Taken together, then, my results and those from the existing literature present a satisfying explanation for the role of age in those reproductively plastic social insect societies that exhibit gerontocracy. Age does not provide inherent benefits to older individuals, but by acting as an honest and easily assessed signal it is nonetheless able to provide a consistent metric by which replacement queens can be chosen with little room for costly within-group disagreements over succession.

6.1.4 Integrating detailed phenotypic and bioinformatic data

A shared feature of the approaches I have taken to the *P. dominula* and *L. flavolineata* data presented in this thesis is the generation of detailed phenotypic data as a basis upon which to develop bioinformatic analyses. The results of Chapter 2 showed in detail how *P. dominula* is able to deal, at the level of the group and the level of the individual, with the potential social conflict that arises from the coexistence of strong reproductive skew and strong reproductive plasticity. Additional to this, however, the detailed phenotypic data that I generated in that chapter allowed me to identify gene expression patterns that would otherwise have been totally missed. The SVM approach that I took in Chapter 3 evinced significant disruption to individuals' gene expression profiles following queen removal, but only in combination with high-resolution behavioural and ovarian data from Chapter 2 was it possible to reveal the extent to which this disruption was independent of caste-specific phenotypic changes. Similarly, in Chapter 4 I took advantage of individual-level measures of age and foraging effort to answer questions about the molecular basis of fitness investment in *L. flavolineata* that would not have been tractable with a less phenotype-specific approach.

The majority of gene expression analyses in the field of social insect evolution have focused on relatively simple comparisons between well-defined groups, typically pooled samples of queens and workers (e.g. Feldmeyer et al. 2014; Berens et al. 2015; Morandin et al. 2016; Geffre et al. 2017; Qiu et al. 2018; He et al. 2019). Such analyses have high power but are limited in the questions they can answer: not all phenotypes exist in discontinuous binaries, and pooling discards many additional data that could be utilized if individuals were sequenced individually. While the appeal of such approaches is

understandable given the time and effort required to generate higher-resolution phenotypic data (especially in social insects, whose large groups are difficult to study), the results I have presented here show that overcoming these difficulties can yield substantial analytic returns. Indeed, given that RNAseq data generation remains expensive relative to the generation of e.g. behavioural data, high-effort experimental designs that maximise analytic yield are likely to be both scientifically fruitful and financially economical.

Notably, the type of low-depth, high-replicate experimental design that facilitated the *P. dominula* analyses described in Chapter 3 remains relatively rare despite the fact that the weight of evidence suggests that shallow sequencing should be preferred in most cases (Atallah et al. 2013; Liu et al. 2014; Milanez-Almeida 2020; Zhang et al. 2020). Except those cases where there is a specific reason to require highly accurate read counts for particular genes (for example when looking for biomarkers for well-defined diseases) sequencing depth is less important than sample size. My decision to use this type of experimental design was principally informed by the power simulations that I detail in Appendix A, which showed decreasing power returns from increasing sequencing depth beyond 10 million reads/sample, even when sample sizes were relatively low and gene expression was noisy. My results therefore reinforce the benefits of careful consideration of experimental design in bioinformatic experiments.

6.2   Final statement

Throughout this thesis, I have combined phenotypic and molecular data to elucidate the mechanisms and consequences of reproductive plasticity in two species of vespid wasp. The conclusions that I have drawn are multi-faceted: that it is possible to maintain a high degree of social cohesion even in a situation with potential for within-group conflict, using a simple conventional succession heuristic; that such cohesion may nonetheless belie substantial molecular disruption, revealing disconnects between gene expression and phenotypic expression; that the differences in patterns of gene expression within insect societies may share more commonalities than would be expected given differences in the expression of sociality at the level of phenotype; and, conversely, that emerging hypotheses regarding the evolutionary trajectory of genes associated with alternative fitness strategies may be more phylogenetically contingent than previously thought. This work nonetheless has a clear unifying message, which is that reproductive plasticity, and

the social insect taxa that express it, can teach us a great deal about the origins and control of sociality. In both their commonalities and their differences, these taxa are scientifically valuable models for the group-level, phenotypic, and molecular processes that underlie social evolution. Until we understand how reproductively plastic societies function, we will not have a full picture of the processes that have generated the extreme forms of sociality that have been the dominant focus of past studies into social insect evolution. This thesis contributes to that goal by exploring the mechanisms that allow reproductively plastic societies to function, while also shedding light on potential drawbacks of this form of social organisation. In doing so, it brings us a step closer to answering the major evolutionary question of when and why biological individuals choose cooperation over conflict.

Appendix A – Exploring RNAseq experimental design using simulated gene expression data

A.1 Introduction

Although the costs of generating sequencing data have declined substantially over the past two decades (Muir et al. 2016), financial constraints remain an important consideration for experimental design in 'omics'-oriented experiments. The two largest determinants of cost in the generation of RNA-seq data are the number of biological replicates and the number of reads per replicate (sequencing depth), and for a given budget these two parameters must be balanced against one another (Ching et al. 2015). Both sample size and sequencing depth affect analytic power: increasing either parameter allows us to more accurately and precisely estimate the 'true' expression means and dispersions present in the population. The practical difficulties of generating larger numbers of biological replicates, especially for difficult-to-source samples, have meant that researchers have frequently preferred to sacrifice sample size rather than sequencing depth. This is reflected in the design of modern differential gene expression programmes such as DESeq2 (Love et al. 2014) and edgeR (Robinson et al. 1010), which place an emphasis on the ability to identify differential expression even when sample sizes are very small (often much smaller than would be permissible for e.g. phenotypic or demographic data).

One method of compensating for the small sample sizes typical of gene expression analyses is to pool many samples of the same treatment in a given biological replicate, which reduces within-treatment variation and allows meaningful differences between groups to be identified more easily. For example, pooling queens of multiple ages and colonies and comparing these to pools of workers may allow a researcher to identify caste-biased gene expression differences while controlling for variation in age or colony. Pooling can greatly increase the power of an experiment, but necessarily discards information about potentially interesting properties of the individuals that are combined within pools (e.g. age, measures of phenotypic covariates, etc.; Konczal et al. 2014; Marinov et al. 2014).

A growing body of papers employing real and simulated transcriptomic data are examining the relative importance of sample size and sequencing depth (Tarazona et al.

2011; Kliebenstein 2012 Atallah et al. 2013; Liu et al. 2014; Milanez-Almeida et al. 2020; Zhang et al. 2020). An emerging theme from this literature is the possibility that the power to detect differential expression may be more sensitive to changes in sample size than to changes in sequencing depth, especially at higher sequencing depths. For example, Atallah et al. (2013) find that in honeybees, the majority of genes identified as being differentially expressed with caste at higher sequencing depths are still apparent at a relatively shallow depth (12 million reads/replicate), and the top 100 differentially expressed genes (DEGs) between queens and workers can be retrieved even with very shallow sequencing (1 million reads/replicate). Atallah et al. used large pools (30 samples/replicate), however, and it is not certain that these results would be applicable to un-pooled data. Moreover, most of the studies that have explored the power impacts of changing sample size and sequencing depth have looked at comparisons between relatively well-defined groups, such as the irreversibly morphologically differentiated castes of honeybees or highly-specific human tissue types. Whether the results of these studies would hold in the case of comparisons between more heterogeneous groups, such as the reproductively plastic castes of less derived social insects, remains an open question.

In this appendix, I use published gene expression data from *Dinoponera quadriceps*, a social insect with plastic castes, as a basis with which to simulate transcriptomic datasets for different combinations of sample sizes, sequencing depths, and genes expression variances. Doing so, I assess the relative impact of sample size and sequencing on power to detect differentially-expressed, in order to assess whether the importance of these factors in a system with high gene expression variance differs from that found in previous studies which used large pools and highly standardised treatment groups.

A.2 Methodology

A.2.1 Reference dataset

In order to simulate a distribution of expression values that was as close to reality as possible, I required transcriptomic data from a species with reproductively plastic castes, because caste expression is expected to be much more heterogenous in such species than in superorganismal taxa. Additionally, I wished to base my simulated data on a dataset with a reported and exceptionally high sequencing depth, since RNA-seq data more closely approximate true gene expression values as more reads are sequenced. One dataset that satisfied both of these criteria and for which read counts were provided was the

Dinoponera quadriceps dataset published by Patalano et al. (2015). *D. quadriceps* lives in small groups with no morphological differentiation and complete reproductive plasticity, making it a suitable model upon which to base simulations of gene expression differences between reproductively plastic castes. Importantly, the dataset published by Patalano et al. utilised un-pooled individuals that were sequenced to a depth of 50 million reads, which is significantly higher than the norm for social insect research.

A.2.2 Data simulation

My aim was to generate a simulated data set of very high quality (many sequenced samples and many reads/sample) gene counts from un-pooled individuals, with a distribution of expression means sampled from an existing social insect study. The purpose of this dataset was to allow me to identify genes that would be identifiable as differentially expressed under idealised conditions, in order to give me a reference dataset against which to assess the power of more realistic experimental designs.

I began by recording the mean expression value (number of reads) for each gene in un-pooled queens and workers in Patalano et al's *D.quadriceps* data and using these data to simulate an 'idealised' gene expression data set with a large sample size (30 queens and 30 workers). Read counts for each gene and un-pooled sample were simulated as being Poisson-distributed around the expected gene- and caste-specific read count taken from the *D. quadriceps* data, with the true mean for each sample being equal to the group mean. Real gene expression values tend to follow a negative binomial distribution (Anders & Huber 2010; Robinson et al. 2010), which is mathematically similar to an over-dispersed Poisson distribution, where variation in read counts is greater than expected from Poisson sampling. My approach thus mimicked an experiment with similar read depth to the original data, but with very low expression variance. I next used DESeq2 to identify genes differentiated between queens and workers in this simulated dataset against a log2 fold-change threshold of 1.5. As expected given the large sample size and low variance of the dataset I had generated, I was able to identify a large number of genes (3332/20140) that were differentially expressed between the simulated queens and workers.

Having generated an 'ideal' gene expression dataset and identified genes that were differentially expressed in these data, I next simulated more realistic datasets according to different experimental designs in order to ask what proportion of the information in the idealised dataset would be lost in the case of smaller sample sizes, smaller sequencing

depths, and/or greater gene expression variance. I constructed a simple R function (*genpowersim*) to assess the power of different experimental designs. This function simulates counts for a specified sample size of two treatment groups using the mean values from a real dataset using the same methodology described above, with two exceptions. First, counts are generated for each sample according to a negative binomial distribution, which is adequate to simulate the effects of both technical and biological variation (Anders & Huber 2010; Robinson et al. 2010), instead of a Poisson distribution. The degree of gene expression variance is set via a dispersion control parameter that is infinite when gene expression variance is equal to mean expression (i.e. Poisson-distributed, no variance added) and is 0 when the variance of each gene's expression is equal to its mean expression squared (amount of variance increased by a factor equal to the mean expression level). Second, after generating read counts for each gene and sample, counts are reduced by a factor equal to the simulated expression depth divided by the original expression depth of the data. For example, as my data were simulated using gene counts from an experiment with sequencing depth = 50 million reads/sample, *genpowersim* would simulate an experiment with sequencing depth = 10 million reads/replicate by generating gene expression counts and then multiplying these counts by 10 million/50 million = 0.2 (and rounding the result downwards to obtain the read count as an integer).

Having generated a gene expression dataset and corrected for simulated sequencing depth, *genpowersim* uses a custom DESeq2 wrapper function to identify differential expression in the simulated dataset. Finally, *genpowersim* compares the results of this DESeq2 analysis to 'true' calls of differential expression, based on the analysis of the idealised dataset generated at the start. The comparison between the more realistic and the idealised analyses can be based on several power metrics: the overall proportion of genes correctly assigned as differentially expressed or not; the proportion of genes that should have been identified as differentially expressed that were assigned correctly, (i.e. power ignoring false positives); and the Receiver Operating Characteristic Area Under Curve (ROC AUC), a more complex metric that assesses how well true positives and true negatives are distinguished. ROC is a method of assessing the power of a binary classifier in which each observation (e.g. each gene) is ranked by the confidence with which it was assigned as a positive result (e.g. its FDR-adjusted p-value). For a perfect classifier, all true positives will be higher rated than all true negatives. The quality of a classifier's ROC can be summarised by the AUC, a metric which can theoretically vary between 0 and 1.

When the AUC for a ROC curve is 1, then the classifier exhibits no error, while an AUC of 0.5 indicates that observations have been classified entirely randomly. Values lower than 0.5 only occur in the unusual case that a classifier is performing worse than randomly (i.e. it misclassifies samples more often than it classifies them correctly), which typically only occurs due to user error such as misidentification of the reference level in a differential expression analysis output. For this reason, I instead report a standardized AUC with values ranging from 0 (random assignment) to 1 (perfect discrimination).

Using *genpowersim* and the idealised dataset generated using Patalano et al's (2015) data, I generated power metrics for a range of different experimental designs. I assessed numbers of replicates per treatment group in the range [5, 10, 15, 20, 25] and sequencing depth (in million reads) in the range [5, 10, 15, 20, 30]. Additionally, I assessed power for each experimental design with different degrees of gene expression variance: high (dispersion control parameter = 10), intermediate (dispersion control parameter = 25), low (dispersion control parameter = 50) and very low (dispersion control parameter = 200).

A.3   Results

As expected, DESeq2's ability to correctly identify differentially-expressed genes increased substantially with increasing sample sizes and sequencing depths, and with decreasing gene expression variance (**Figure A1**). Unsurprisingly, DESeq2 performed very poorly when sample sizes and sequencing depths were both at their lowest (5 replicates/treatment with 5 million reads/sample): even when gene expression variance was very low, only an eighth (12.5%) of DEGs were correctly recovered, and at higher gene expression variances this number fell as low as 3%. At the highest sample size (25 replicates/treatment) and sequencing depth (30 million reads/sample) simulated, DESeq2 performed well when gene expression variance was low, recovering 78.8% of DEGs, but at higher variances this number fell below 50%.

Power increased substantially when increasing both sample size and sequencing depth from their lowest simulated values, including when gene expression variance was high. The initial power gains when increasing from 5 to 10 million reads were not sustained at higher sequencing depths, however: beyond 15 million reads, there was very little power gain from increased depth, especially for higher sample sizes (**Figure A2**). By contrast, while the power gains from increasing sample size were most dramatic when increasing

from 5 to 10 replicates, DESeq2's ability to correctly identify DEGs continued to improve substantially for all simulated increases in sample size and at all sequencing depths (**Figure A3**).

Figure A1. Variation in power to detect differentially-expressed genes in a simulated dataset of workers and queens with varying sequencing depth and number of replicates, and with variance of gene expression simulated as being high (A), intermediate (B), low (C) or very low (D). Bold numbers: Area Under Curve of the Receiver Operating Characteristic for the given set of conditions, standardised to vary between 0 and 1. Italicised numbers: percentage of genes that were differentially expressed in an idealised dataset that were identified as such by DESeq2 under the given set of conditions.
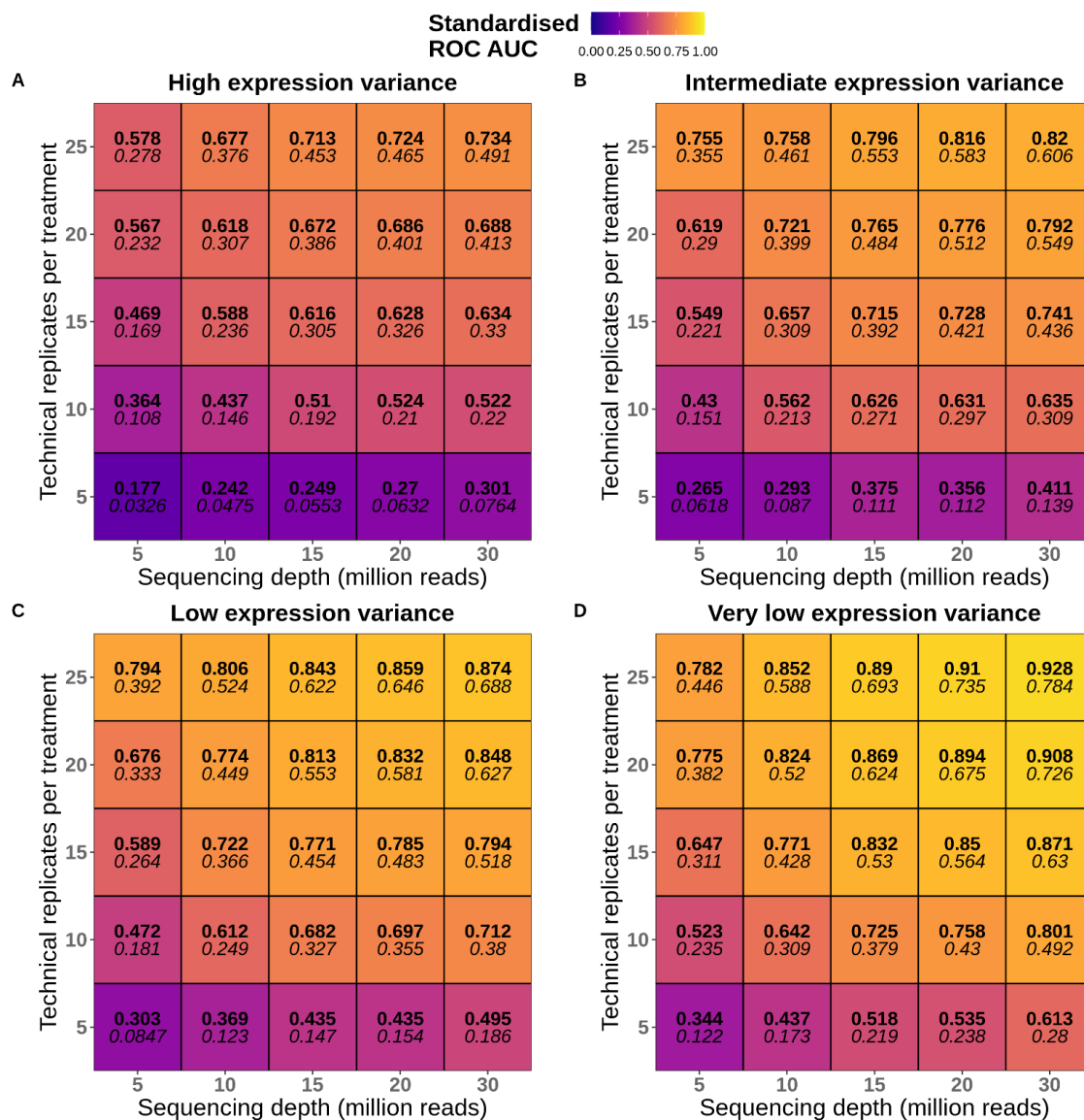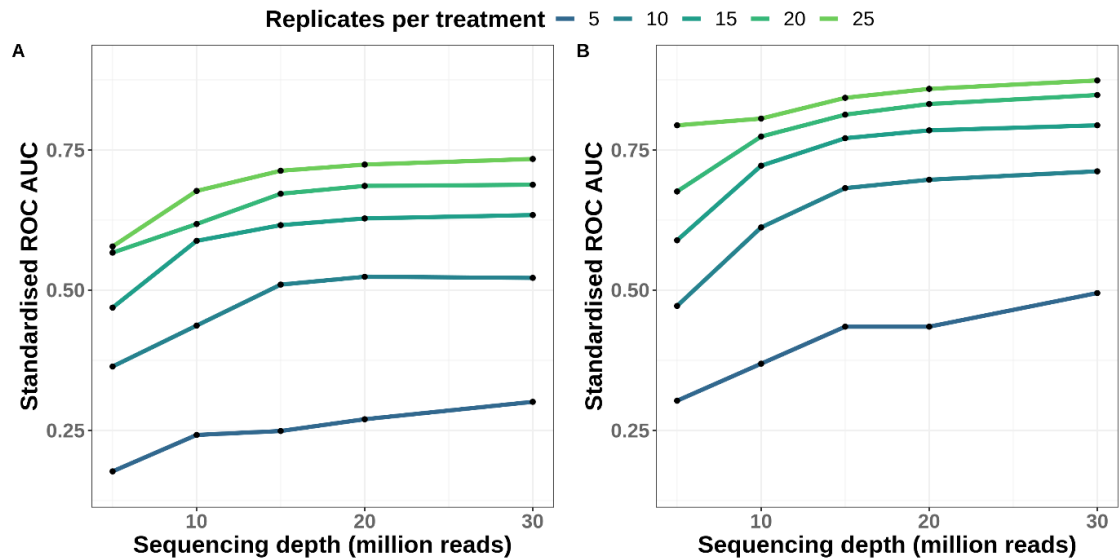
Figure A2. Variation in power to detect differentially-expressed genes in a simulated dataset of workers and queens with varying sequencing depth and with variance of gene expression simulated as being (A) high or (B) very low, for various numbers of technical replicates.
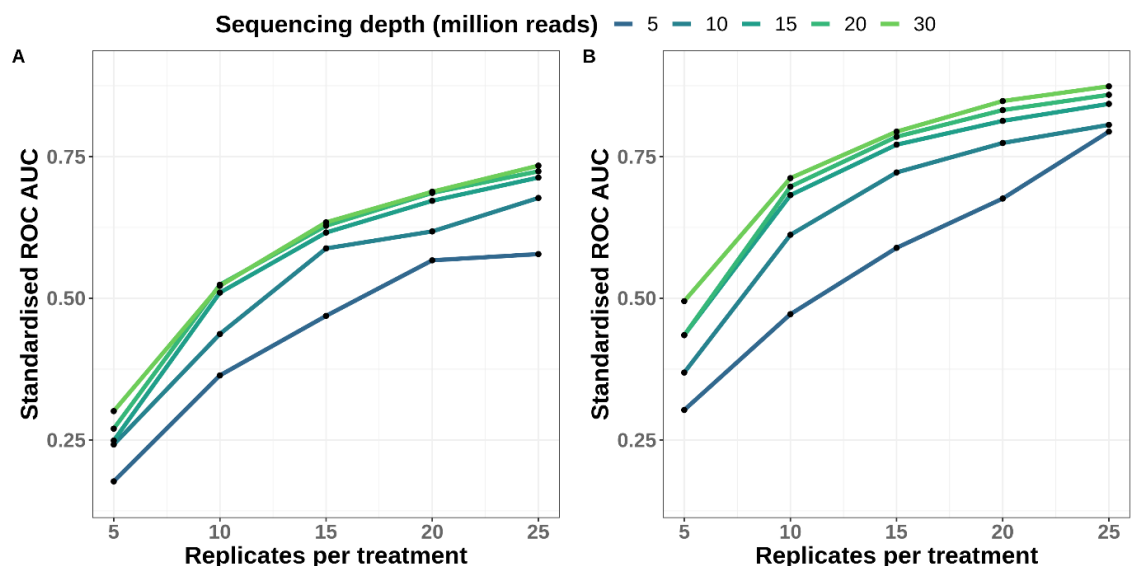
Figure A3. Variation in power to detect differentially-expressed genes in a simulated dataset of workers and queens with varying number of replicates and with variance of gene expression simulated as being (A) high or (B) very low, various levels of sequencing depth.

A.4 Discussion

Over the past decade, a substantial body of literature exploring the effects of varying experimental design on the power of transcriptomic analyses has emerged. Two consensus outcomes of these studies have been that (1) sample size is at least as important as sequencing depth in determining the power of differential expression analysis, and (2) beyond a relatively shallow depth, increases in sequencing depth generate decreasing returns in terms of power. However, most studies to date have focused on low-variance

expression data from highly specific tissues or from pooled samples. Here, I have shown that the broad results of the existing literature also apply to simulated datasets with high gene expression variance and based on the distribution of expression differences from relatively heterogenous groups such as the plastic castes of non-superorganismal social insects.

It is particularly notable that in my simulated data, diminishing returns became apparent when I increased sequencing depth from approximately 10-15 million reads/sample, because this is also the sequencing depth that previous papers using strongly pooled data have identified as being optimal. For example, Atallah et al. (2013) find diminishing returns beyond 12 million reads/replicate using pooled *Apis mellifera* caste gene expression data, while Liu et al. (2014) identify 10 million reads/replicate as the point at which power gains begin levelling off when examining clonal human cells. The fact that approximately the same sequencing depth optimum appears in my simulated data, even when gene expression variance was modelled as being very high, suggests that this result is applicable beyond highly-controlled clinical contexts. The consistency with which the specific figure of 10-15 million reads/sample appears in the literature is particularly surprising given that whole transcriptome sizes are thought to vary strongly between species and even between tissues within individuals (Reef et al. 2010; Coate & Doyle 2015). An important future direction for research in this area might be comparing the point at which power gains from sequencing depth begin to diminish in species with large known differences in whole transcriptome sizes, such as recently-formed polyploids (Coate & Doyle 2010).

For the purposes of this thesis, my finding that a shallow sequencing depth of 10 million reads/sample allows the recovery of the majority of differentially-expressed genes even in the higher-variance context of unpooled data and heterogenous treatment groups informed the experimental design that I describe in Chapter 3. My decision to sequence unpooled data turned out to be instrumental in obtaining several of the results presented in that chapter, thus demonstrating the benefits of careful and informed decision-making when planning bioinformatic experiments.

Appendix B – *Liostenogaster flavolineata* genome assembly

B.1 Introduction

This appendix details the methods used to generate the *Liostenogaster flavolineata* genome assembly against which the transcriptomes analysed in Chapter 4 were aligned. As this genome is currently unpublished, I provide the details here as context for that chapter. The work described in this appendix was performed by Alexandrina Bodrug, Nancy Stralis-Pavese and Heinz Himmelbauer of the Department of Biotechnology, University of Natural Resources and Life Sciences (BOKU), Vienna, Austria.

B.2 Methods

DNA was extracted from a single haploid *L. flavolineata* male using a DNeasy Blood & Tissue Kit (Qiagen) according to the manufacturers' instructions. DNA quantification was performed with a Qubit 3.0 fluorometer using a dsDNA BR assay kit (Thermo Fisher, Waltham, MA, USA) and DNA integrity was monitored on an agarose gel. The preparation of sequencing libraries was performed using DNA isolated from haploid males. A sequencing library with a peak insert size of 535 bp was constructed using 200 ng of genomic DNA with a TruSeq Nano LT library preparation kit (Cat # FC-121-4002, Illumina, San Diego, CA, USA) according to the kit supplier´s instructions. A mate-pair (MP) library with a peak span size of 1450 bp and a mean span size of 1027 bp was prepared from 570 ng of genomic DNA by tagmentation using an MP library preparation kit (Cat # FC-132-1001, Illumina, San Diego, CA, USA), without size selection. The MP library was amplified using 12 cycles of PCR. The quality and quantity of the libraries was checked on a DNA 1000 chip on the Agilent Bioanalyzer 2100 (Agilent, Santa Clara, CA, USA).

Illumina sequencing was performed at the Vienna BioCenter Core Facilities (VBCF), Vienna, Austria, on a HiSeq 2500 instrument utilising v4 Illumina sequencing chemistry, combined with a 2x125 cycle sequencing recipe. Raw sequencing data underwent quality control with FastQC (Andrews 2010), thereafter, trimmomatic (Bolger et al. 2014) was employed for data filtering based on phred scores, using the following parameters: LEADING:25 TRAILING:25 SLIDINGWINDOW:10:25 MINLEN:36. Genome assemblies were preformed using SOAPdenovo_v2.04 (Luo et al. 2012). Pre-assemblies

were first calculated based on paired-end (PE) reads which were assembled either as single reads or as PE reads (Dohm et al. 2014), in order to assess the insert size distribution (PE reads) and span size distribution (MP reads) of the sequencing libraries, respectively. Bowtie2 (Langmead & Salzberg 2012) was used with an insert size interval between 100 and 1200 (bowtie2 --fr -I 100 -X 1200 -1 pe_reads1.fastq -2 pe_reads2.fastq -x preassembly_1) and 1 million PE read-pairs were sampled to estimate the library insert size. To estimate the span size of the MP read-pairs, bowtie2 was used on an assembly version using the paired-end library as pairs and an insert size interval between 100 and 20000 (bowtie2 --rf -I 100 -X 20000 -1 mp_reads1.fastq -2 mp_reads2.fastq -x preassembly_2), and sampled 1 million MPs. Using the determined library insert size and MP span size as parameters for the assembly run, several assemblies were calculated from the quality-filtered sequencing reads by varying the k-mer size parameter between 23 and 125. An assembly calculated with k-mer size 69 (SOAPdenovo-63mer all -s assembly.config -K 69 -R -o assembly.K69) was the best performing in terms of assembly metrics as assessed by QUAST (Gurevitch et al. 2013).

BUSCOv3 (Simão et al. 2015; Waterhouse et al. 2017) was used in genome mode with the hymenoptera_odb9 lineage and honeybee1 species to assess assembly completeness (blast 2.2.30, AUGUSTUS 3.2.1). Metrics of the final assembly were determined with custom scripts, taking only sequences larger than 500 bp into account. Jellyfish 2.2.10 (Marçais & Kingsford 2011) was used to determine genome size based on the quality-filtered Illumina PE sequencing reads. Bioawk was used to retrieve the GC content of all the reads as well as non-overlapping 125 nt segments of the final assembly lacking undetermined bases (no unknown nucleotide "N").

B.3   Results

The genome of *L. flavolineata* was assembled from DNA extracted from a single haploid male. From this individual, both a paired-end sequencing library and a mate-pair library were prepared (see methods for details). By means of Illumina sequencing,163 million pairs of genomic paired-end reads and 102 million pairs of mate-pairs were obtained. The size of the genome was estimated by counting k-mers, using the unassembled paired-end sequencing data as input for an analysis using jellyfish (Marçais & Kingsford 2011). In this way, a genome size of 373 Mbp was calculated for *L. flavolineata* based on 17-mers. The genome sequence that was assembled using SOAPdenovo_v2.04 was smaller, i.e. 291 Mbp, taking account only of sequences > 500 bp. The longest scaffold in the *L.*

flavolineata assembly Lifl-v1.0 had a length of 5.22 Mbp, and the N50 scaffold length was 1.5 Mbp (**Table B1**). A fraction of the total assembly was contained within sequences <= 500 nt, i.e. 243,289 sequences (30 Mbp). The GC content distribution had a single peak in both in the PE sequencing data that were used as input for the assembly and the assembly itself, which contrasts with findings of bimodal or trimodal GC content distributions in other wasps. The completeness of the *L. flavolineata* genome assembly was assessed with respect to conserved hymenopteran genes, using the Benchmarking Universal Single-Copy Orthologs (BUSCO) approach (Simão et al. 2015; Waterhouse et al. 2017). Of 4,415 BUSCO groups searched, 97.9% were found in the assembly, and 96.9% were complete (**Table B2**). We may therefore conclude that the Lifl-v1.0 genome assembly is a highly comprehensive representation of the *L. flavolineata* genome.

Assembly size	291.28 Mbp
N50 size	1.50 Mbp
% GC	41.65
% unspecified bases (N)	5.1
Largest scaffold	5.22 Mbp
Number of scaffolds + contigs	3,541

Table B1. *L. flavolineata* genome assembly metrics based on sequences > 500 bp.

BUSCO category	Number	Percentage
Complete BUSCOs	4,277	96.9
Complete Single-Copy BUSCOs	4,268	96.7
Complete Duplicated BUSCOs	9	0.2
Fragmented BUSCOs	46	1
Missing BUSCOs	92	2.1
Total BUSCO number of groups	4,415	100

Table B2. BUSCO completeness metrics for the *L. flavolineata* genome assembly.

Appendix C – Supplementary figures

Figure S3.1. WGCNA summary network measures for each *P. dominula* treatment group against soft thresholding power. Numbers in the plots indicate the corresponding soft thresholding powers. A soft thresholding power of 9 was chosen to balance model fit against connectivity.
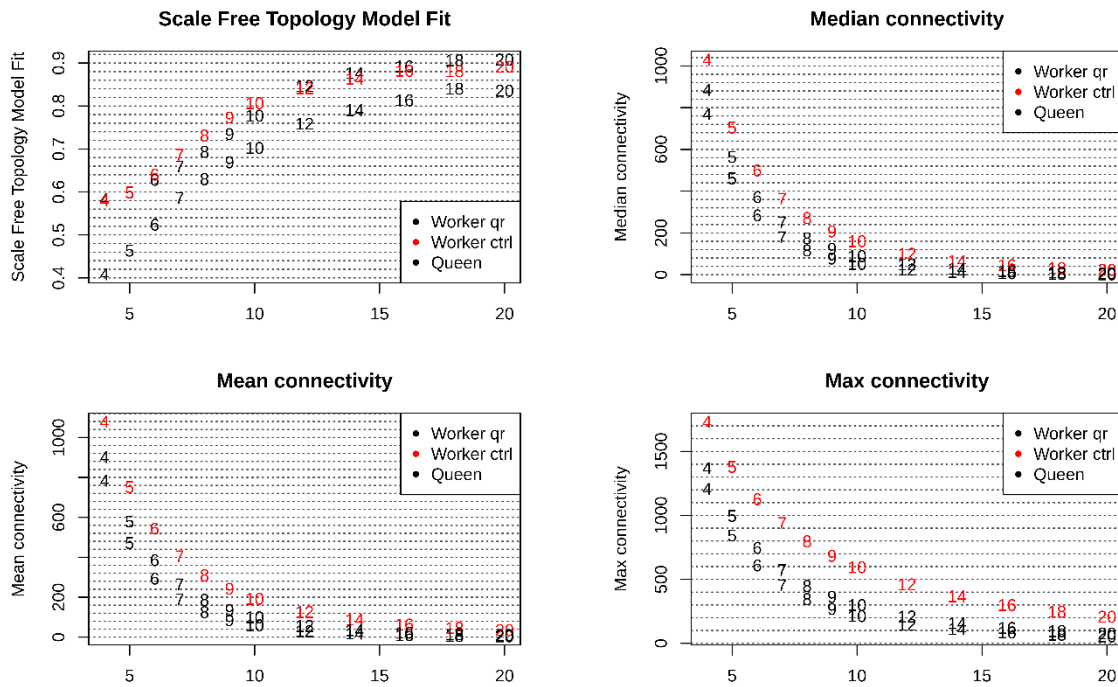
Figure S3.2. *P. dominula* gene dendrogram with clustering based on consensus topological overlap. Upper colour row: consensus module assignments prior to merging of modules with similar expression profiles. Lower colour row: consensus modules following merging of similar modules.
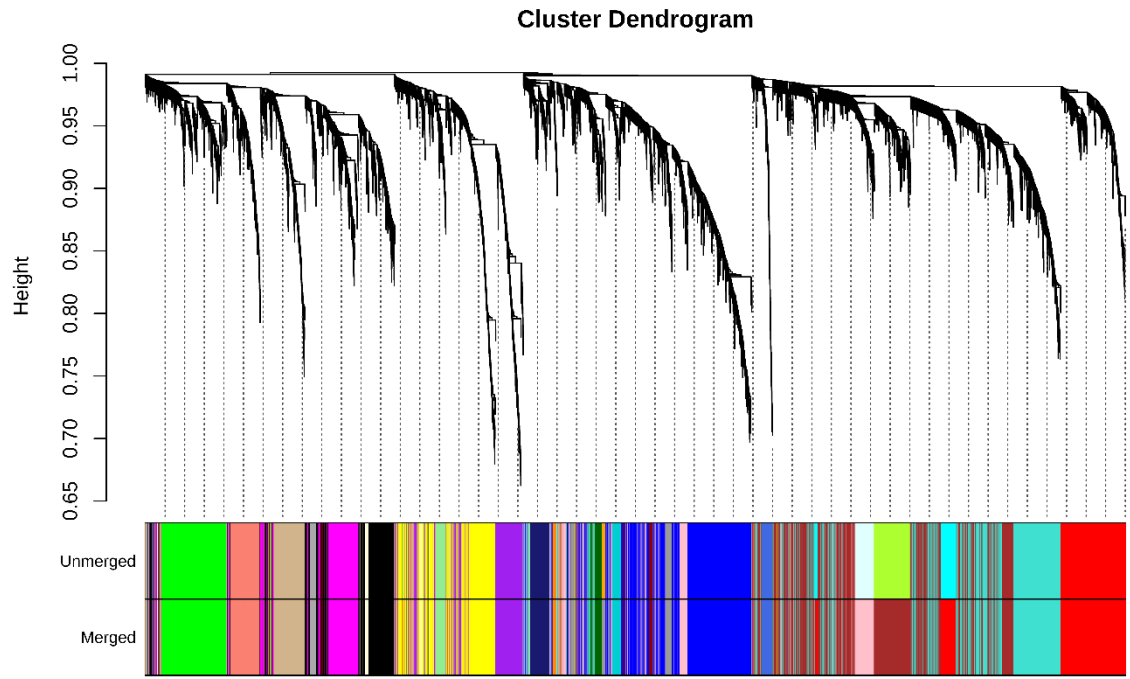
Figure S3.3 GO terms enriched among the 1992 *P. dominula* genes in an optimised SVM classifier separating queens and workers from control colonies. GO terms with a semantic similarity greater than 0.9 are grouped; bold labels for each group indicate the term with the greatest significance (lowest p-value) within that group.
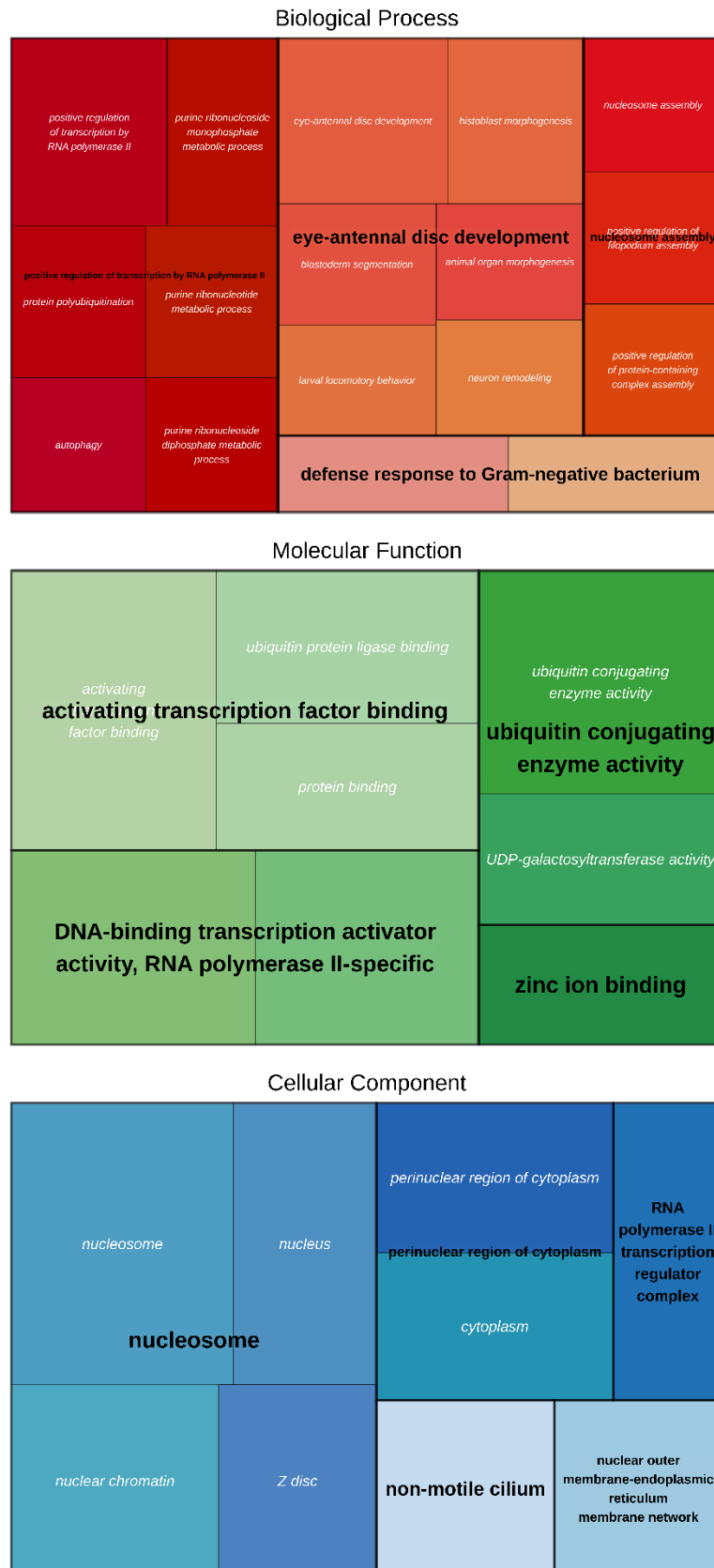
Figure S3.4 Absolute feature weights of the 1992 *P. dominula* genes in the optimised SVM. 81 genes identified by DESeq2 as being differentially expressed between queens and control workers with a baseline log2 fold-change of 1.5 are marked in blue.
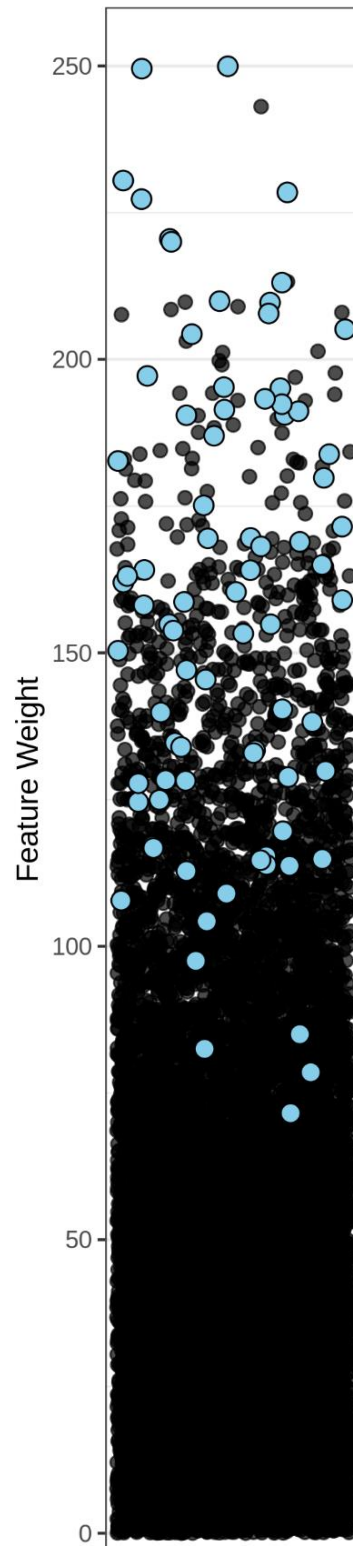
Figure S3.5. Distribution of phenotypic traits of sequenced *P. dominula* individuals from queen removal colonies. Where possible, individuals were selected to represent as wide as possible a range of values for (A) Ovarian development; (B) Dominance; and (C) Phenotypic caste identity ('queenness').
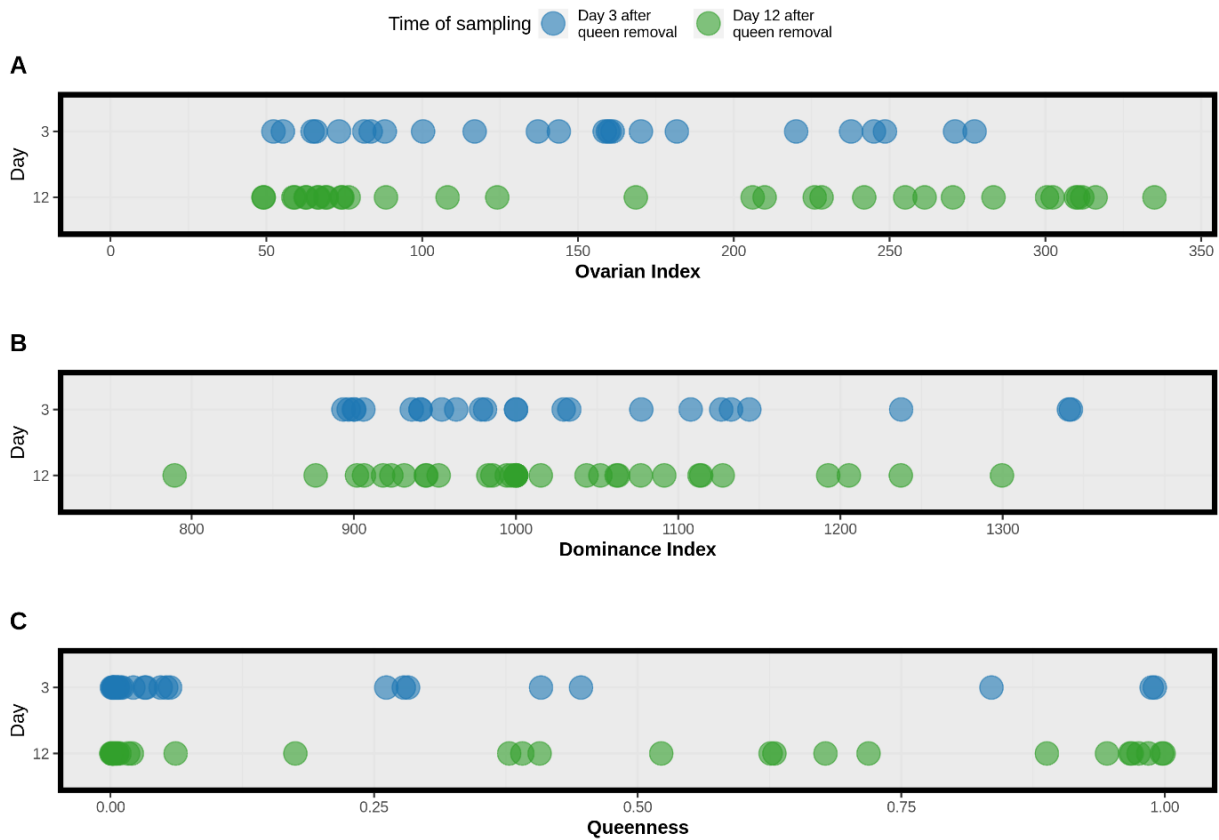
Figure S3.6. GO terms enriched among genes in *P. dominula* WGCNA Module 8. GO terms with a semantic similarity greater than 0.9 are grouped; bold labels for each group indicate the term with the greatest significance (lowest p-value) within that group.

Figure S3.7. GO terms enriched among genes in *P. dominula* WGCNA Module 8. GO terms with a semantic similarity greater than 0.9 are grouped; bold labels for each group indicate the term with the greatest significance (lowest p-value) within that group.
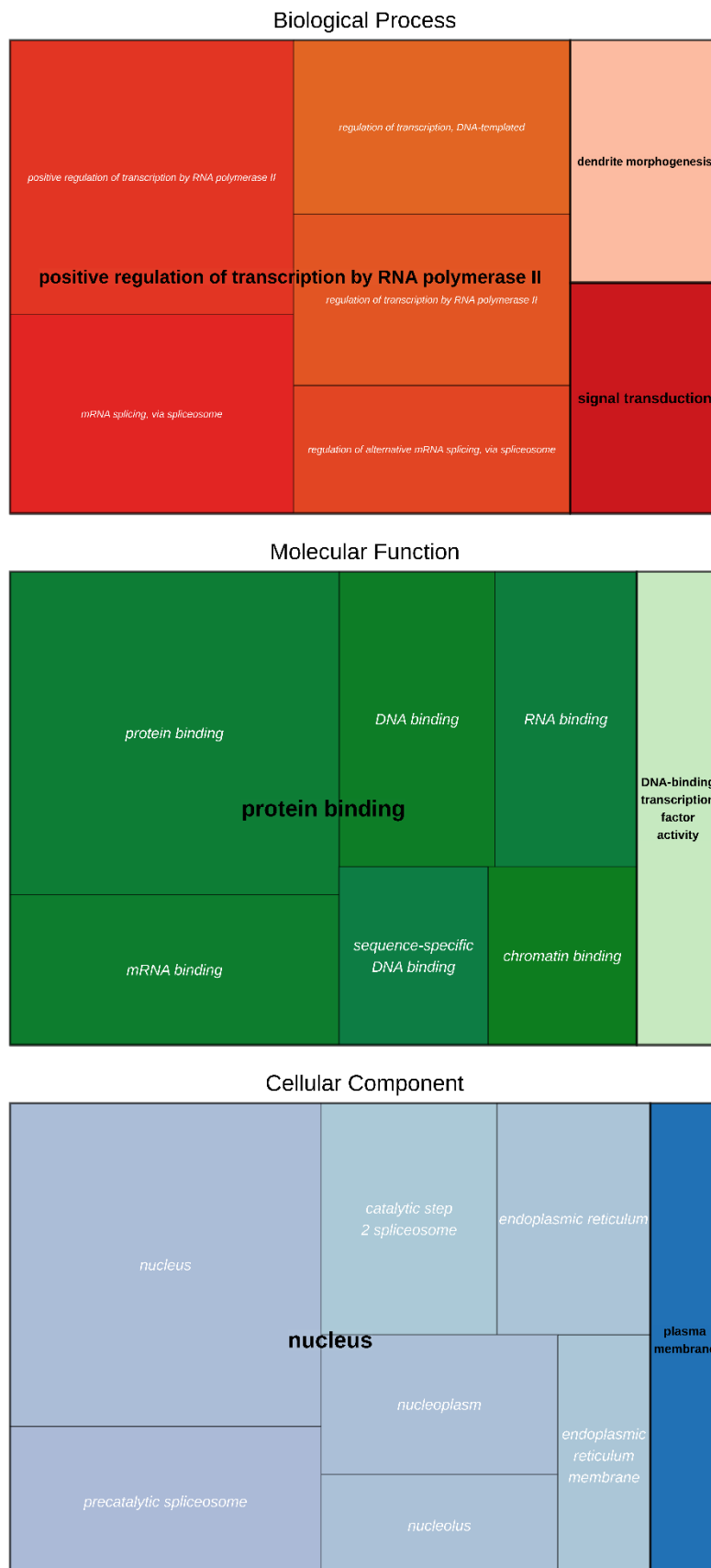
Figure S4.1. *L. flavolineata* WGCNA summary network measures against soft thresholding power. Numbers in the plots indicate the corresponding soft thresholding powers. A soft thresholding power of 6 was chosen to balance model fit against connectivity.
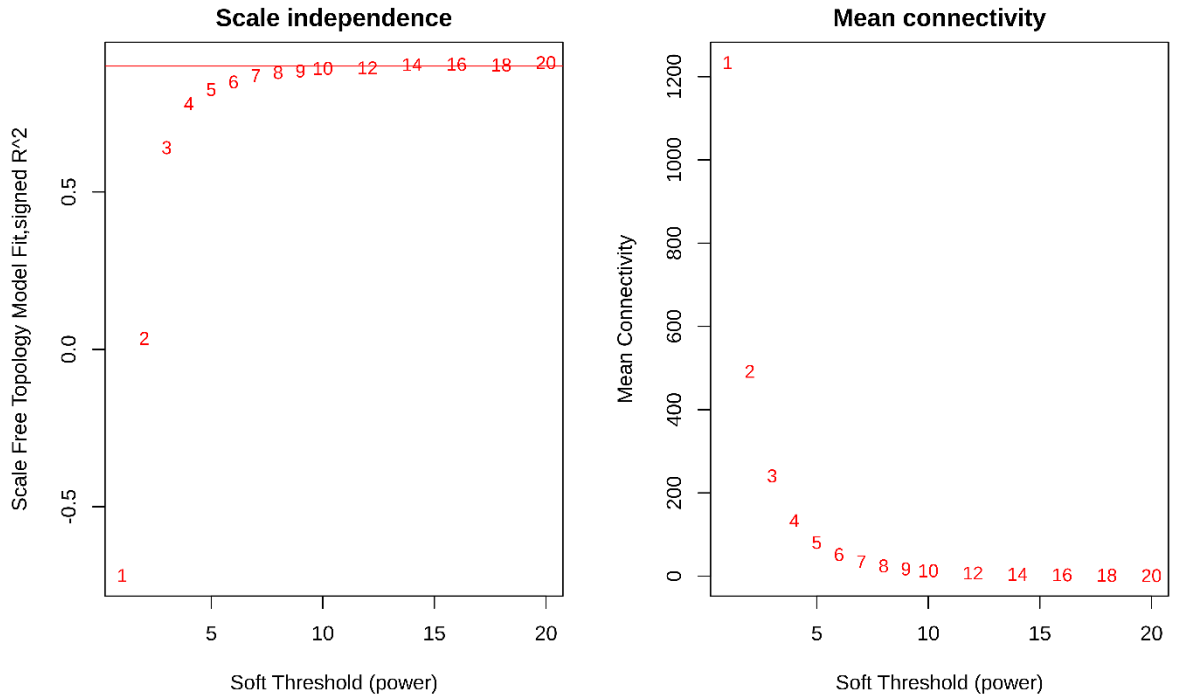
Figure S4.2. *L. flavolineata* gene dendrogram with clustering based on consensus topological overlap. Upper colour row: consensus module assignments prior to merging of modules with similar expression profiles. Lower colour row: consensus modules following merging of similar modules.
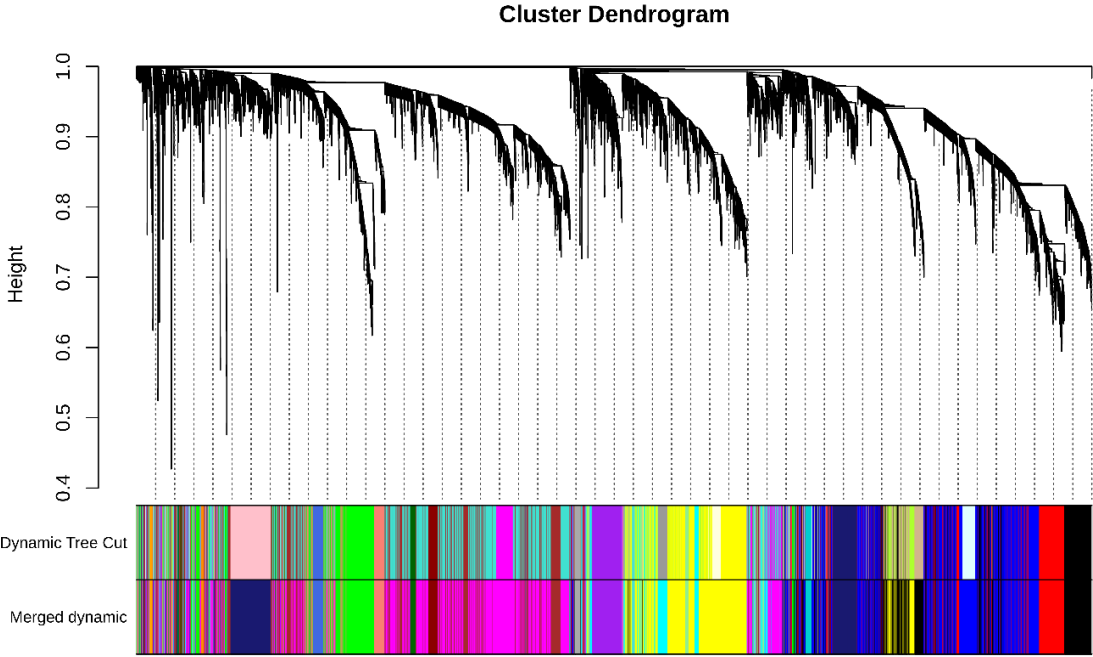
Figure S4.3. First two PCs of a principal components analysis for the 100 *L. flavolineata* genes with highest expression variance among individuals sequenced in Experiment 1.
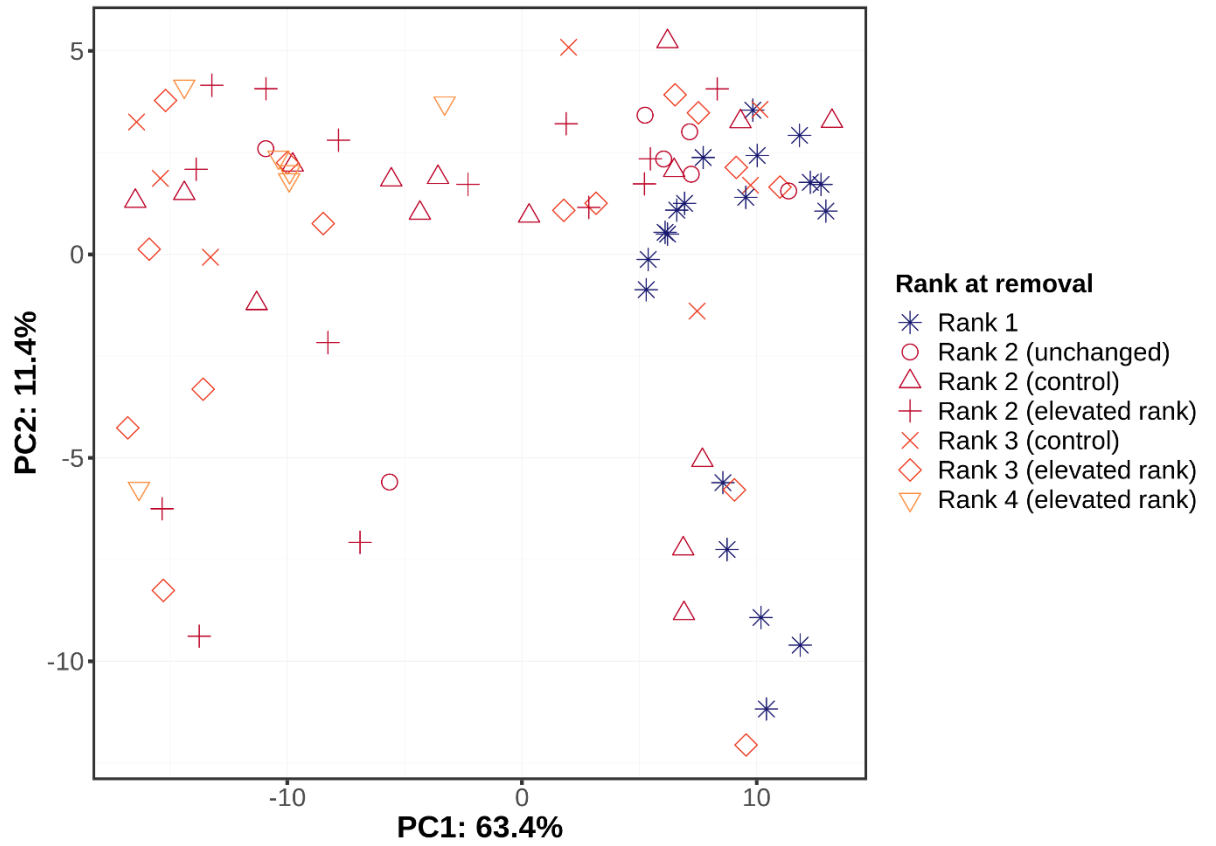
Figure S4.4. GO terms enriched among queen-biased genes in *L. flavolineata*. GO terms with a semantic similarity greater than 0.9 are grouped; bold labels for each group indicate the term with the greatest significance (lowest p-value) within that group.
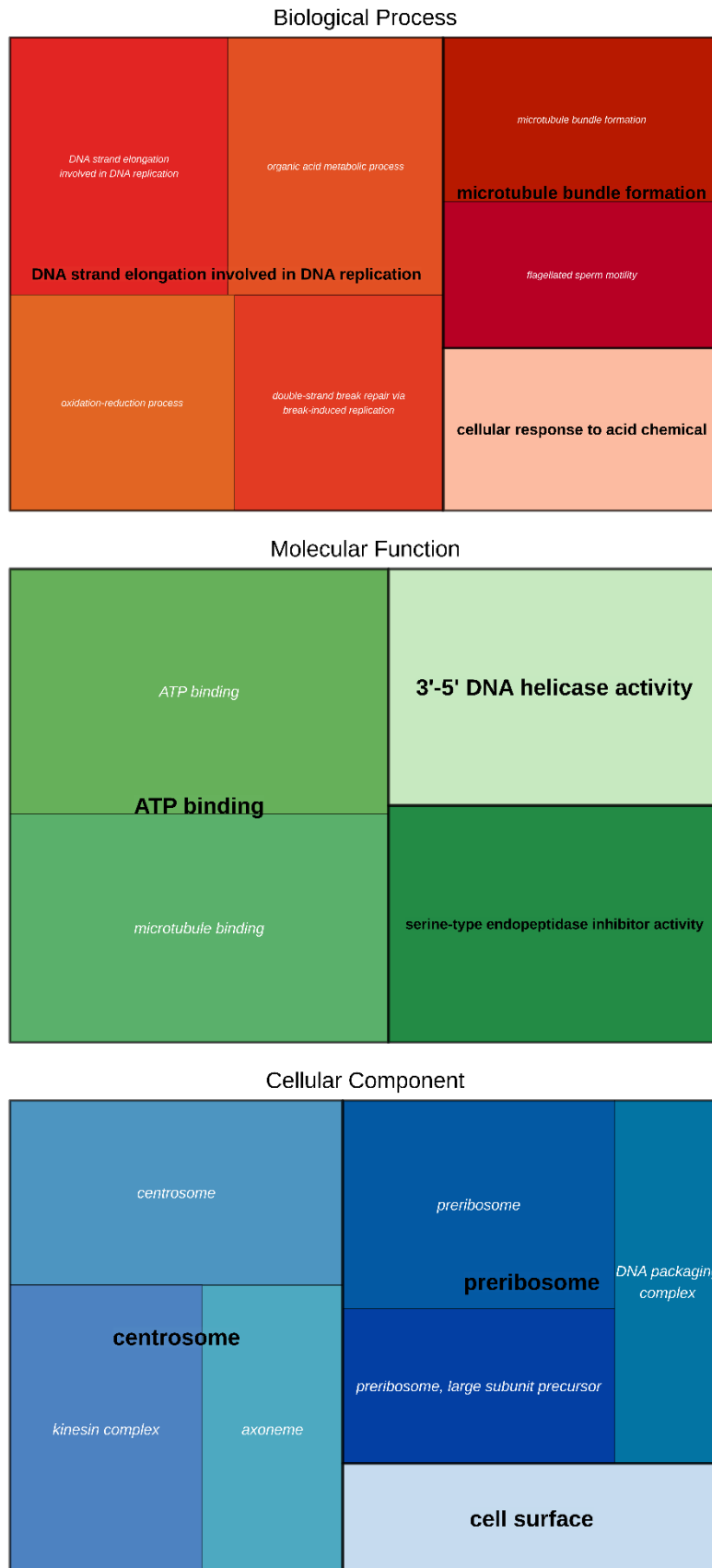
Figure S4.5. GO terms enriched among worker-biased genes in *L. flavolineata*. GO terms with a semantic similarity greater than 0.9 are grouped; bold labels for each group indicate the term with the greatest significance (lowest p-value) within that group.
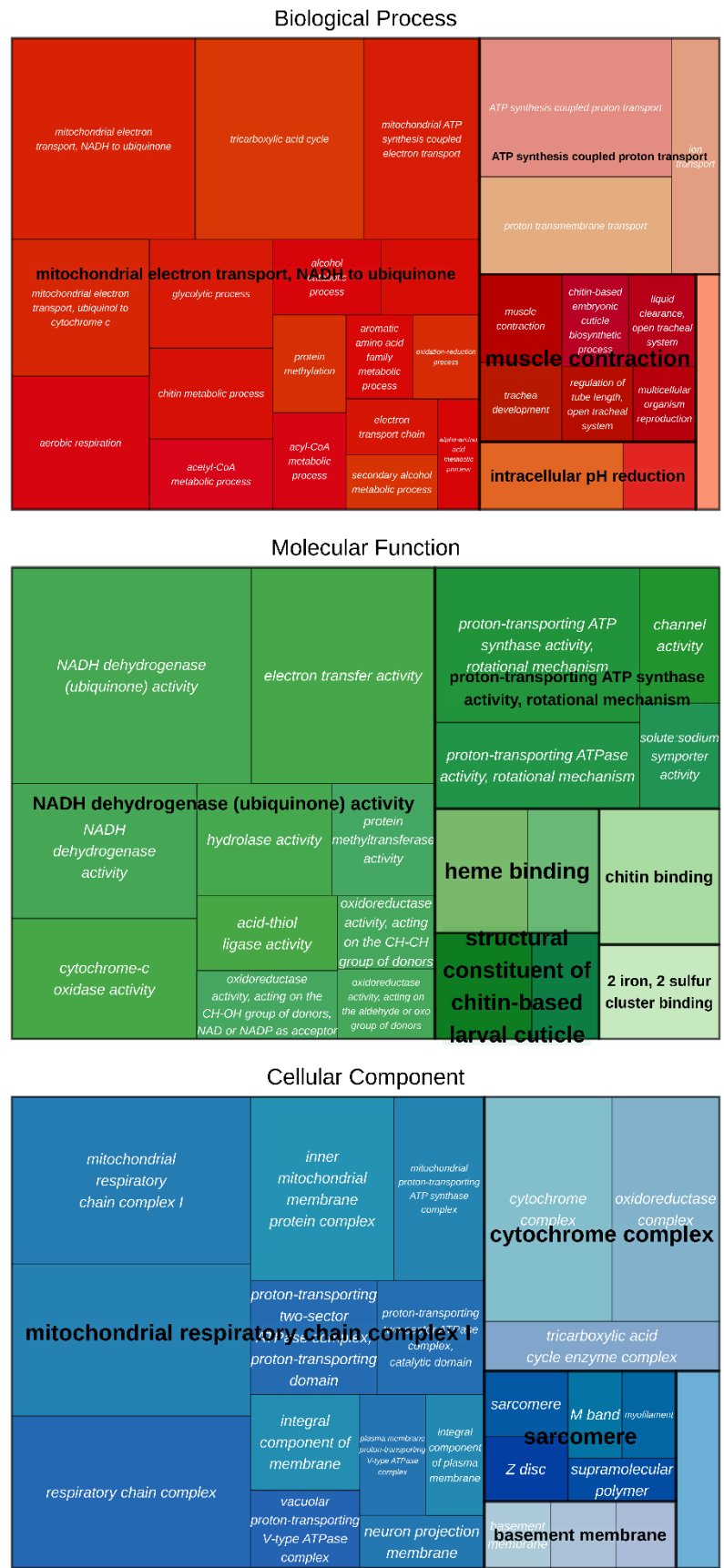
Figure S4.6. GO terms enriched among genes whose expression was positively correlated with foraging effort in *L. flavolineata* workers. GO terms with a semantic similarity greater than 0.9 are grouped; bold labels for each group indicate the term with the greatest significance (lowest p-value) within that group.
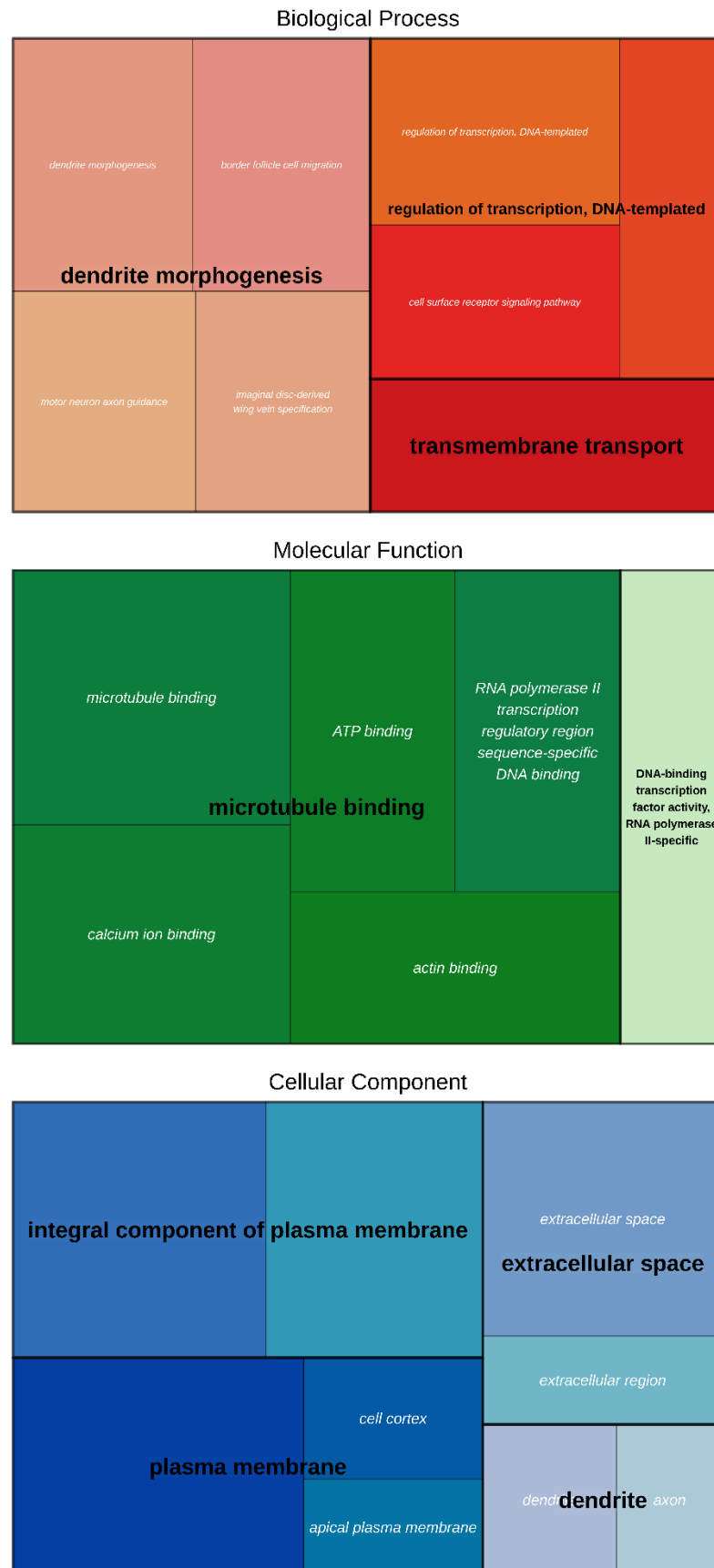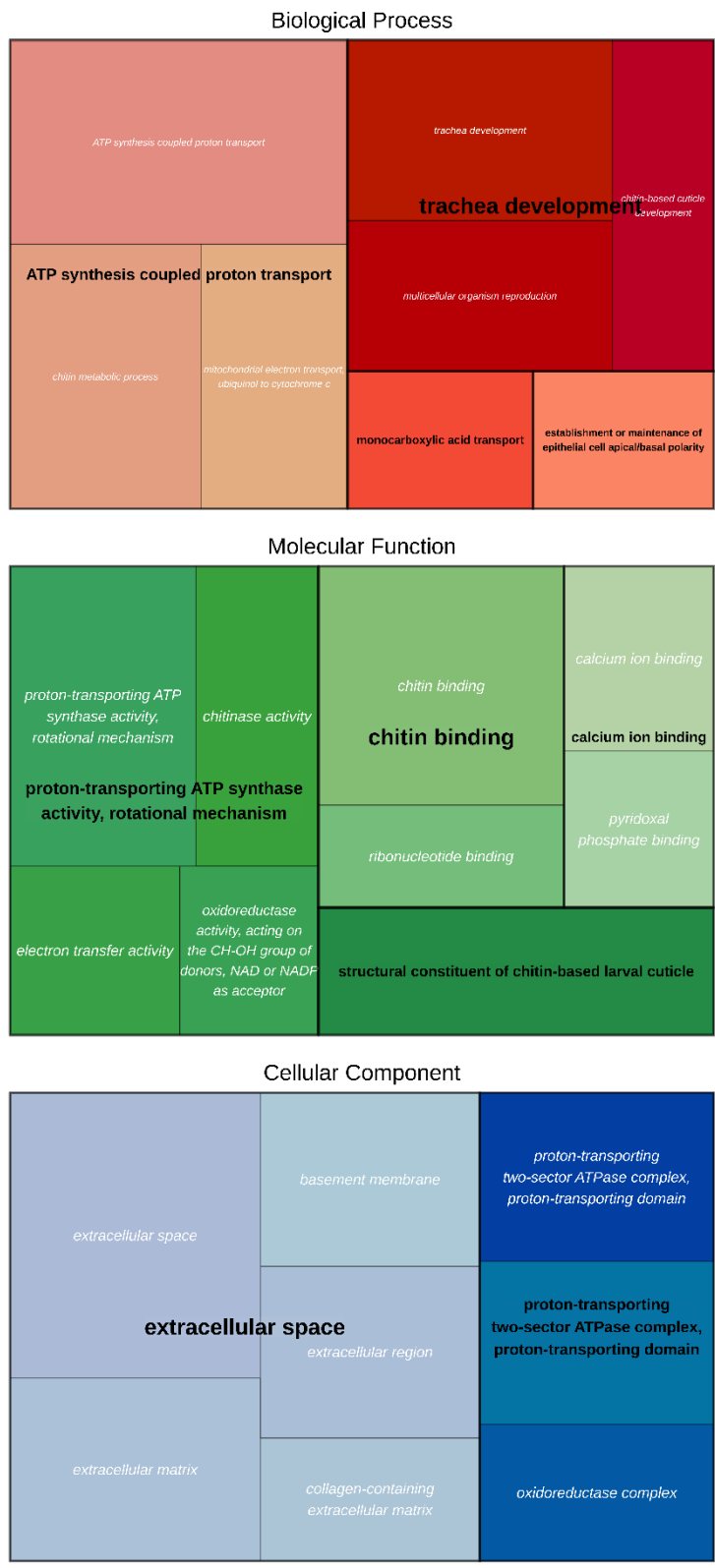
Figure S4.7. GO terms enriched among genes whose expression was negatively correlated with foraging effort in *L. flavolineata* workers. GO terms with a semantic similarity greater than 0.9 are grouped; bold labels for each group indicate the term with the greatest significance (lowest p-value) within that group.
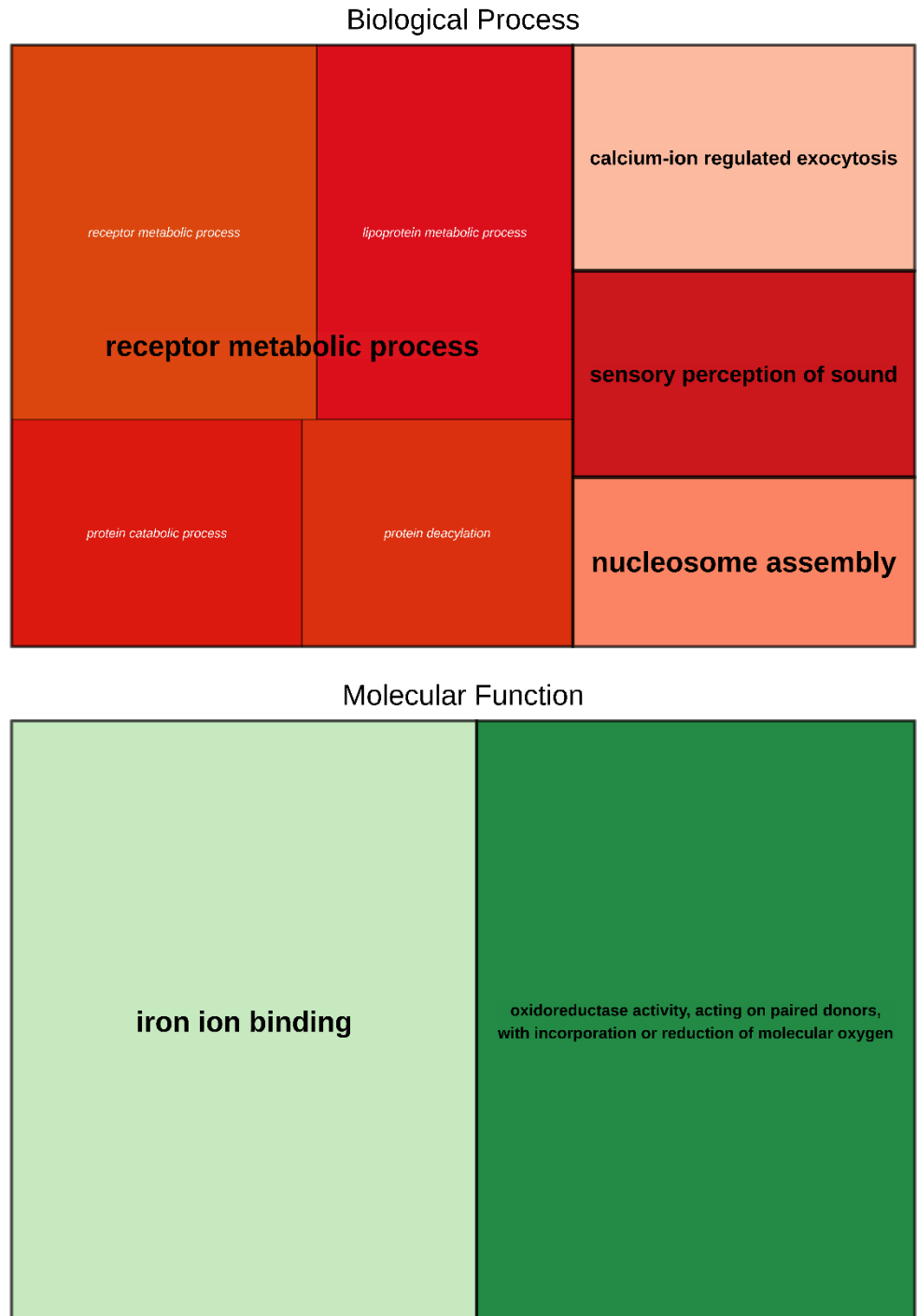
Figure S4.8. GO terms enriched among genes that were members of *L. flavolineata* Gene Module 12. GO terms with a semantic similarity greater than 0.9 are grouped; bold labels for each group indicate the term with the greatest significance (lowest p-value) within that group.
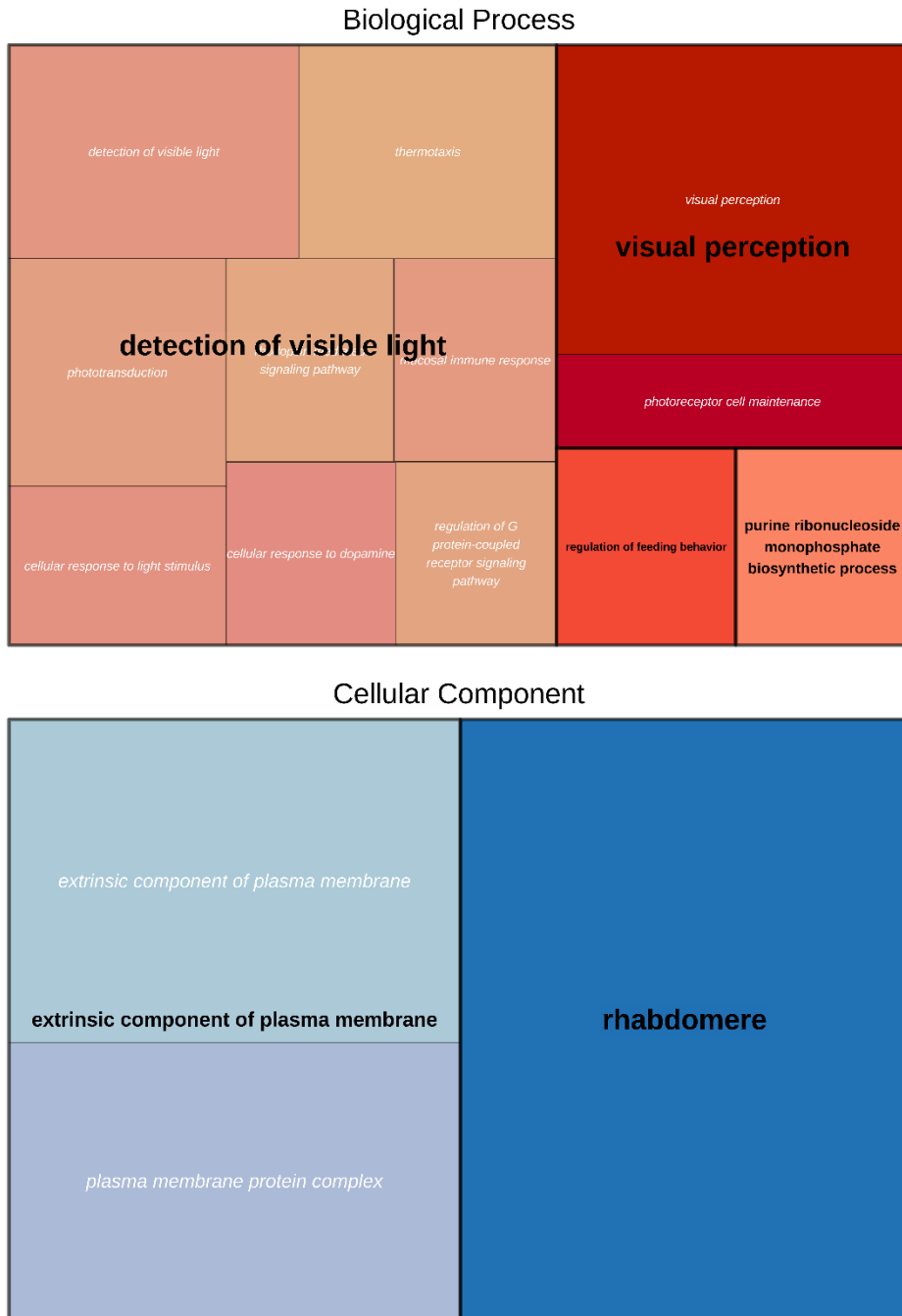
Figure S4.9. GO terms enriched among genes that were members of *L. flavolineata* Gene Module 16. GO terms with a semantic similarity greater than 0.9 are grouped; bold labels for each group indicate the term with the greatest significance (lowest p-value) within that group.
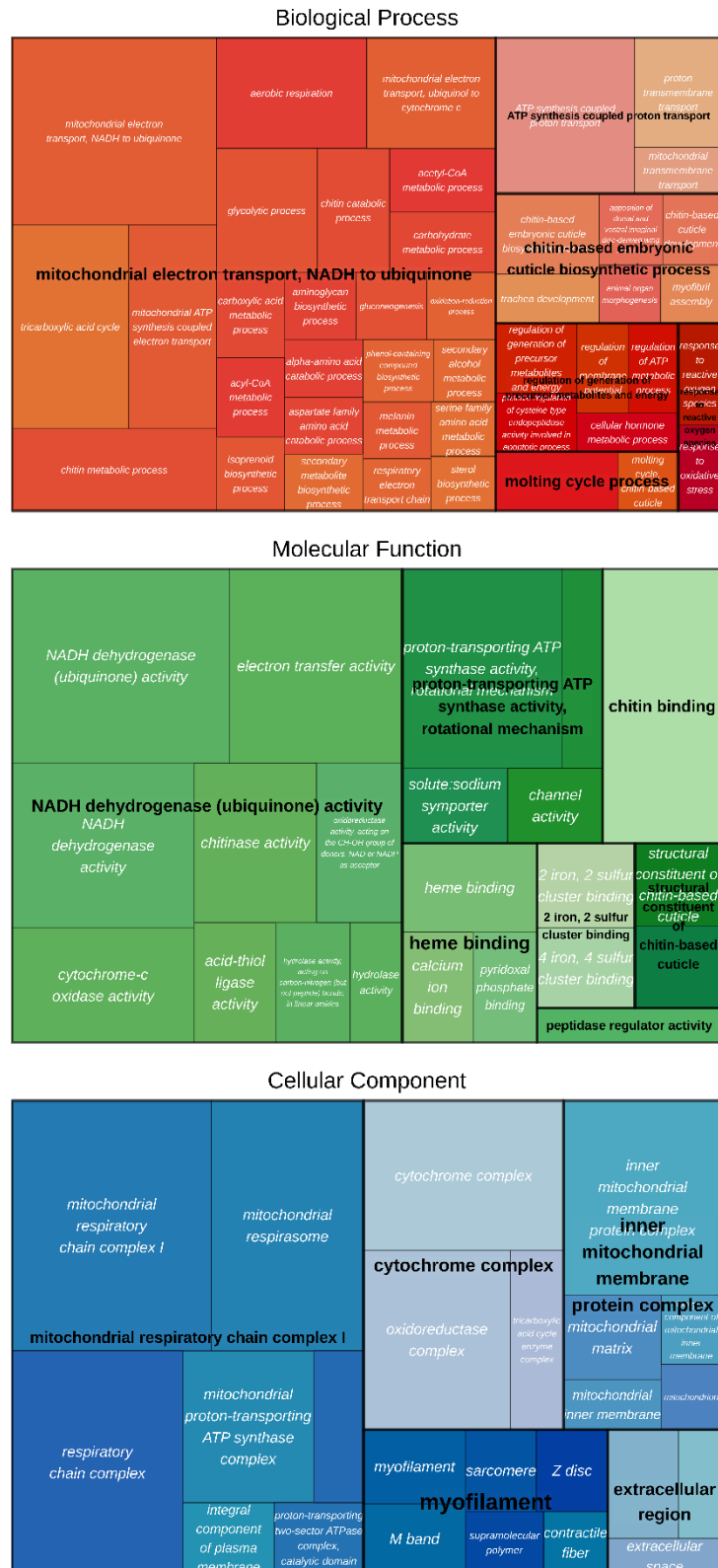
Figure S4.10. Proportion of genes of different estimated evolutionary ages among genes that were upregulated in reproductives ('queen-biased'), non-reproductives ('worker-biased') and unbiased with respect to reproductive caste in *P. dominula*.
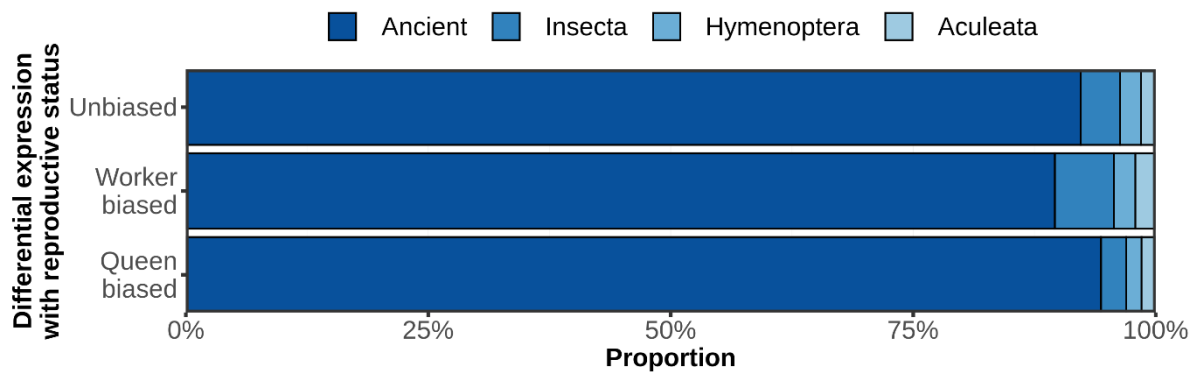
References

Abeel, T., Helleputte, T., Van de Peer, Y., Dupont, P., & Saeys, Y. (2010). Robust biomarker identification for cancer diagnosis with ensemble feature selection methods. *Bioinformatics*, 26(3), 392-398.

Albers, P. C., & Vries, H. D. (2001). Elo-rating as a tool in the sequential estimation of dominance strengths. *Animal Behaviour*, 489-495.

Alexa, A., & Rahnenführer, J. (2009). Gene set enrichment analysis with topGO. *Bioconductor Improv*, 27.

Amdam, G. V., Csondes, A., Fondrk, M. K., & Page, R. E. (2006). Complex social behaviour derived from maternal reproductive traits. *Nature*, *439*(7072), 76-78.

Amdam, G. V., Norberg, K., Fondrk, M. K., & Page, R. E. (2004). Reproductive ground plan may mediate colony-level selection effects on individual foraging behavior in honey bees. *Proceedings of the National Academy of Sciences*, 101(31), 11350-11355.

Anders, S., & Huber, W. (2010). Differential expression analysis for sequence count data. *Nature Precedings*, 1-1.

Andrew G, Yu-Sung S. (2018). arm: Data Analysis Using Regression and Multilevel/Hierarchical Models. R package version 1.10-2/r276.

Andrews S (2010) FastQC: A quality control tool for high throughput sequence data. URL https://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

Ashby, R., Forêt, S., Searle, I., & Maleszka, R. (2016). MicroRNAs in honey bee caste determination. *Scientific Reports*, *6*, 18794.

Atallah, J., Plachetzki, D. C., Jasper, W. C., & Johnson, B. R. (2013). The utility of shallow RNA-seq for documenting differential gene expression in genes with high and low levels of expression. *PLoS One*, *8*(12), e84160.

Aureli, F., Cords, M., & Van Schaik, C. P. (2002). Conflict resolution following aggression in gregarious animals: a predictive framework. *Animal behaviour*, 64(3), 325-343.

Aureli, F., de Waal, F., & Waal, F. B. (Eds.). (2000). Natural conflict resolution. Univ of California Press.Baracchi D. 2017. The reproductive division of labour but not worker

age affects spatial sorting within the nest in a paper wasp. *Insectes Sociaux*. 64(3):379–385.

Baerends, G. P. (1941). On the Life-History of *Ammophila campestris* Jur. *Proceedings of the Koninklijke Nederlandse Akademie van Wetenschappen*, 44, 483–488.

Bank, S., Sann, M., Mayer, C., Meusemann, K., Donath, A., Podsiadlowski, L., ... & Niehuis, O. (2017). Transcriptome and target DNA enrichment sequence data provide new insights into the phylogeny of vespid wasps (Hymenoptera: Aculeata: Vespidae). *Molecular phylogenetics and evolution*, 116, 213-226.

Bannerman, B. P., Kramer, S., Dorrell, R. G., & Carrington, M. (2018). Multispecies reconstructions uncover widespread conservation, and lineage-specific elaborations in eukaryotic mRNA metabolism. *PLoS One, 13*(3), e0192633.

Baracchi, D., Petrocelli, I., Chittka, L., Ricciardi, G., & Turillazzi, S. (2015). Speed and accuracy in nest-mate recognition: a hover wasp prioritizes face recognition over colony odour cues to minimize intrusion by outsiders. *Proceedings of the Royal Society B: Biological Sciences*, 282(1802), 20142750.

Barker, J. S. S. (2009). Defining fitness in natural and domesticated populations. In *Adaptation and fitness in animal populations* (pp. 3-14). Springer, Dordrecht.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1-48.

Beani, L., Cervo, R., Lorenzi, C. M., & Turillazzi, S. (1992). Landmark-based mating systems in four *Polistes* species (Hymenoptera: Vespidae). *Journal of the Kansas Entomological Society*, 211-217.

Beani, L., Dallai, R., Mercati, D., Cappa, F., Giusti, F., & Manfredini, F. (2011). When a parasite breaks all the rules of a colony: morphology and fate of wasps infected by a strepsipteran endoparasite. *Animal Behaviour*, 82(6), 1305-1312.

Beekman, M., Komdeur, J., & Ratnieks, F. L. (2003). Reproductive conflicts in social animals: who has power?. *Trends in Ecology & Evolution*, 18(6), 277-282.

Bell, A. M., & Robinson, G. E. (2011). Behavior and the dynamic genome. *Science*, 332(6034), 1161-1162.

Benowitz, K. M., & Moore, A. J. (2016). Biparental care is predominant and beneficial to parents in the burying beetle *Nicrophorus orbicollis* (Coleoptera: Silphidae). *Biological Journal of the Linnean Society*, 119(4), 1082-1088.

Berens, A. J., Hunt, J. H., & Toth, A. L. (2015). Comparative transcriptomics of convergent evolution: different genes but conserved pathways underlie caste phenotypes across lineages of eusocial insects. *Molecular biology and evolution*, *32*(3), 690-703.

Bernadou, A., Busch, J., & Heinze, J. (2015). Diversity in identity: behavioral flexibility, dominance, and age polyethism in a clonal ant. *Behavioral ecology and sociobiology*, *69*(8), 1365-1375.

Bernasconi, G., & Strassmann, J. E. (1999). Cooperation among unrelated individuals: the ant foundress case. *Trends in Ecology & Evolution*, 14(12), 477-482.

Bewick, A. J., Vogel, K. J., Moore, A. J., & Schmitz, R. J. (2017). Evolution of DNA methylation across insects. *Molecular biology and evolution*, 34(3), 654-665.

Blacher, P., Huggins, T. J., & Bourke, A. F. (2017). Evolution of ageing, costs of reproduction and the fecundity–longevity trade-off in eusocial insects. *Proceedings of the Royal Society B: Biological Sciences*, 284(1858), 20170380.

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15), 2114-2120.

Bomtorin, A. D., Mackert, A., Rosa, G. C. C., Moda, L. M., Martins, J. R., Bitondi, M. M. G., ... & Simões, Z. L. P. (2014). Juvenile hormone biosynthesis gene expression in the corpora allata of honey bee (*Apis mellifera*) female castes. *PloS one*, 9(1), e86923.

Bonasio, R., Li, Q., Lian, J., Mutti, N. S., Jin, L., Zhao, H., ... & Jin, Z. (2012). Genome-wide and caste-specific DNA methylomes of the ants *Camponotus floridanus* and *Harpegnathos saltator*. *Current Biology*, 22(19), 1755-1764.

Boomsma, J. J. (2009). Lifetime monogamy and the evolution of eusociality. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1533), 3191-3207.

Boomsma, J. J. (2013). Beyond promiscuity: mate-choice commitments in social breeding. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 368(1613), 20120050.

Boomsma, J. J., & Gawne, R. (2018). Superorganismality and caste differentiation as points of no return: how the major evolutionary transitions were lost in translation. *Biological Reviews*, 93(1), 28-54.

Bourke, A. F. (1999). Colony size, social complexity and reproductive conflict in social insects. *Journal of evolutionary Biology*, 12(2), 245-257.

Breed, M. D., & Gamboa, G. J. (1977). Behavioral control of workers by queens in primitively eusocial bees. *Science*, 195(4279), 694-696.

Brian, M. V. (2012). Social insects: ecology and behavioural biology. *Springer Science & Business Media.*

Bridge, C., & Field, J. (2007). Queuing for dominance: gerontocracy and queue-jumping in the hover wasp *Liostenogaster flavolineata*. *Behavioral Ecology and Sociobiology*, 61(8), 1253-1259.

Camargo, R. S., Forti, L. C., Lopes, J. F. S., Andrade, A. P. P., & Ottati, A. L. T. (2007). Age polyethism in the leaf-cutting ant *Acromyrmex subterraneus brunneus* Forel, 1911 (Hym., Formicidae). *Journal of Applied Entomology*, *131*(2), 139-145.

Cant, M. A., & English, S. (2006). Stable group size in cooperative breeders: the role of inheritance and reproductive skew. *Behavioral Ecology*, 17(4), 560-568.

Cant, M. A., & Field, J. (2001). Helping effort and future fitness in cooperative animal societies. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 268(1479), 1959-1964.

Cant, M. A., & Young, A. J. (2013). Resolving social conflict among females without overt aggression. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 368(1631), 20130076.

Cant, M. A., English, S., Reeve, H. K., & Field, J. (2006). Escalated conflict in a social hierarchy. *Proceedings of the Royal Society B: Biological Sciences*, 273(1604), 2977-2984.

Capella, I. C. S., & Hartfelder, K. (2002). Juvenile-hormone-dependent interaction of actin and spectrin is crucial for polymorphic differentiation of the larval honey bee ovary. *Cell and tissue research*, 307(2), 265-272.

Carpenter, J. M. (1996). Distributional checklist of species of the genus *Polistes* (Hymenoptera, Vespidae, Polistinae, Polistini). *American Museum novitates*; no. 3188.

Carpenter, J. M. (2003). On "Molecular phylogeny of Vespidae (Hymenoptera) and the evolution of sociality in wasps". *American Museum Novitates*, 2003(3389), 1-20.

Carrete, M., Donázar, J. A., Margalida, A., & Bertran, J. (2006). Linking ecology, behaviour and conservation: does habitat saturation change the mating system of bearded vultures?. *Biology Letters*, 2(4), 624-627.

Casas, L., Saborido-Rey, F., Ryu, T., Michell, C., Ravasi, T., & Irigoien, X. (2016). Sex change in clownfish: molecular insights from transcriptome analysis. *Scientific reports*, 6, 35461.

Cervo, R., Dapporto, L., Beani, L., Strassmann, J. E., & Turillazzi, S. (2008). On status badges and quality signals in the paper wasp *Polistes dominulus*: body size, facial colour patterns and hierarchical rank. *Proceedings of the Royal Society B: Biological Sciences*, 275(1639), 1189-1196.

Chakravarty, P., Cozzi, G., Ozgul, A., & Aminian, K. (2019). A novel biomechanical approach for animal behaviour recognition using accelerometers. *Methods in Ecology and Evolution*, 10(6), 802-814.

Chandra, V., Fetter-Pruneda, I., Oxley, P. R., Ritger, A. L., McKenzie, S. K., Libbrecht, R., & Kronauer, D. J. (2018). Social regulation of insulin signaling and the evolution of eusociality in ants. *Science*, 361(6400), 398-402.

Chapman, B. B., Brönmark, C., Nilsson, J. Å., & Hansson, L. A. (2011). The ecology and evolution of partial migration. *Oikos*, 120(12), 1764-1775.

Ching, T., Huang, S., & Garmire, L. X. (2014). Power analysis and sample size estimation for RNA-Seq differential expression. *Rna*, *20*(11), 1684-1696.

Cini, A., Meconcelli, S., & Cervo, R. (2013). Ovarian indexes as indicators of reproductive investment and egg-laying activity in social insects: a comparison among methods. *Insectes sociaux*, 60(3), 393-402.

Clarke, F. M., & Faulkes, C. G. (1997). Dominance and queen succession in captive colonies of the eusocial naked mole–rat, *Heterocephalus glaber*. *Proceedings of the Royal Society of London. Series B: biological sciences*, 264(1384), 993-1000.

Coate, J. E., & Doyle, J. J. (2010). Quantifying whole transcriptome size, a prerequisite for understanding transcriptome evolution across species: an example from a plant allopolyploid. *Genome biology and evolution*, *2*, 534-546.

Coate, J. E., & Doyle, J. J. (2015). Variation in transcriptome size: are we getting the message?. *Chromosoma*, *124*(1), 27-43.

Cornwallis, C. K., Botero, C. A., Rubenstein, D. R., Downing, P. A., West, S. A., & Griffin, A. S. (2017). Cooperation facilitates the colonization of harsh environments. *Nature Ecology & Evolution*, 1(3), 1-10.

Cortes-Avizanda, A., Almaraz, P., Carrete, M., Sanchez-Zapata, J. A., Delgado, A., Hiraldo, F., & Donazar, J. A. (2011). Spatial heterogeneity in resource distribution promotes facultative sociality in two trans-Saharan migratory birds. *PLoS One*, 6(6), e21016.

Costa, J. T., & Fitzgerald, T. D. (2005, January). Social terminology revisited: Where are we ten years later?. In *Annales Zoologici Fennici* (Vol. 42, No. 6, pp. 559-564).

Cotter, S. C., Ward, R. J., & Kilner, R. M. (2011). Age-specific reproductive investment in female burying beetles: independent effects of state and risk of death. *Functional Ecology*, 25(3), 652-660.

Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T. L. (2008). BLAST+: architecture and applications. *BMC Bioinformatics*, 10, 421.

Creel, S., Creel, N., Wildt, D. E., & Monfort, S. L. (1992). Behavioural and endocrine mechanisms of reproductive suppression in Serenge dwarf mongooses. *Animal Behaviour*, 43(2), 231-245.

Croft, D. P., James, R., & Krause, J. (2008). Exploring animal social networks. *Princeton University Press.*

Cronin, A. L., & Hirata, M. (2003). Social polymorphism in the sweat bee *Lasioglossum (Evylaeus) baleicum* (Cockerell)(Hymenoptera, Halictidae) in Hokkaido, northern Japan. *Insectes Sociaux*, 50(4), 379-386.

Cronin, A. L., Bridge, C., & Field, J. (2011). Climatic correlates of temporal demographic variation in the tropical hover wasp *Liostenogaster flavolineata*. *Insectes sociaux*, 58(1), 23-29.

Csardi, G., & Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal complex systems,* 1695(5), 1-9.

Cullen, D. A., Cease, A. J., Latchininsky, A. V., Ayali, A., Berry, K., Buhl, J., ... & Rogers, S. M. (2017). From molecules to management: mechanisms and consequences of locust

phase polyphenism. In *Advances in Insect Physiology* (Vol. 53, pp. 167-285). Academic Press.

Dalman, M. R., Deeter, A., Nimishakavi, G., & Duan, Z. H. (2012, December). Fold change and p-value cutoffs significantly alter microarray interpretations. In *BMC bioinformatics* (Vol. 13, No. S2, p. S11). BioMed Central.

Damien, R. F., Alfredo, S. T. (2018). aniDom: Inferring Dominance Hierarchies and Estimating Uncertainty. R package version 0.1.3.

Dapporto, L., Bruschini, C., Cervo, R., Petrocelli, I., & Turillazzi, S. (2010). Hydrocarbon rank signatures correlate with differential oophagy and dominance behaviour in *Polistes dominulus* foundresses. *Journal of Experimental Biology*, 213(3), 453-458.

Dapporto, L., Cini, A., Palagi, E., Morelli, M., Simonti, A., & Turillazzi, S. (2007). Behaviour and chemical signature of pre-hibernating females of *Polistes dominulus* infected by the strepsipteran Xenos vesparum. *Parasitology*, 134(4), 545.

Dapporto, L., Sledge, F. M., & Turillazzi, S. (2005). Dynamics of cuticular chemical profiles of *Polistes dominulus* workers in orphaned nests (Hymenoptera, Vespidae*). Journal of insect physiology*, 51(9), 969-973.

Dearden, P. K., Wilson, M. J., Sablan, L., Osborne, P. W., Havler, M., McNaughton, E., ... & Beye, M. (2006). Patterns of conservation and change in honey bee developmental genes. *Genome Research*, *16*(11), 1376-1384.

Djawdan, M., Sugiyama, T. T., Schlaeger, L. K., Bradley, T. J., & Rose, M. R. (2004). Metabolic aspects of the trade-off between fecundity and longevity in *Drosophila melanogaster*. In *Methuselah Flies: A Case Study in the Evolution of Aging* (pp. 145-164).

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., ... & Gingeras, T. R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, 29(1), 15-21.

Dohm, J. C., Minoche, A. E., Holtgräwe, D., Capella-Gutiérrez, S., Zakrzewski, F., Tafer, H., ... & Himmelbauer, H. (2014). The genome of the recently domesticated crop plant sugar beet (*Beta vulgaris*). *Nature*, 505(7484), 546-549.

Duncan, C., Gaynor, D., & Clutton-Brock, T. (2018). The importance of being beta: female succession in a cooperative breeder. *Animal Behaviour*, 146, 113-122.

East, M. L., & Hofer, H. (1991). Loud calling in a female-dominated mammalian society: II. Behavioural contexts and functions of whooping of spotted hyaenas, *Crocuta crocuta*. *Animal Behaviour*, 42(4), 651-669.

Eickwort, G. C., Eickwort, J. M., Gordon, J., Eickwort, M. A., & Wcislo, W. T. (1996). Solitary behavior in a high-altitude population of the social sweat bee *Halictus rubicundus* (Hymenoptera: Halictidae). *Behavioral Ecology and Sociobiolog*y, 38(4), 227-233.

Elango, N., Hunt, B. G., Goodisman, M. A., & Soojin, V. Y. (2009). DNA methylation is widespread and associated with differential gene expression in castes of the honeybee, Apis mellifera. Proceedings of the National Academy of Sciences, 106(27), 11206-11211.

Elo, A. E. (1978). The Rating Of Chess Players, Past & Present. New York: Arco.

Emms, D. M., & Kelly, S. (2019). OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome biology*, 20(1), 1-14.

Evans, H. E. (1958). The evolution of social life in wasps. *Proceedings of the 10th International Congress of Entomology*, 449-457.

Ewels, P. A., Peltzer, A., Fillinger, S., Patel, H., Alneberg, J., Wilm, A., Garcia, M. U., Di Tommaso, P. & Nahnsen, S. (2020). The nf-core framework for community-curated bioinformatics pipelines. *Nature Biotechnology*, 38(3), 276-278.

Faulkes, C. G., & Bennett, N. C. (2013). Plasticity and constraints on social evolution in African mole-rats: ultimate and proximate factors. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 368(1618), 20120347.

Feldmeyer, B., Elsner, D., & Foitzik, S. (2014). Gene expression patterns associated with caste and reproductive status in ants: worker-specific genes are more derived than queen-specific ones. *Molecular ecology*, *23*(1), 151-161.

Ferreira, P. G., Patalano, S., Chauhan, R., Ffrench-Constant, R., Gabaldon, T., Guigo, R., & Sumner, S. (2013). Transcriptome analyses of primitively eusocial wasps reveal novel insights into the evolution of sociality and the origin of alternative phenotypes. *Genome biology*, 14(2), R20.

Field, J. (1992). Patterns of nest provisioning and parental investment in the solitary digger wasp *Ammophila sabulosa*. *Ecological Entomology*, 17(1), 43-51.

Field, J., & Foster, W. (1999). Helping behaviour in facultatively eusocial hover wasps: an experimental test of the subfertility hypothesis. *Animal behaviour*, 57(3), 633-636.

Field, J., & Leadbeater, E. (2016). Cooperation between non-relatives in a primitively eusocial paper wasp, *Polistes dominula*. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371(1687), 20150093.

Field, J., Cant, M. A. (2009). Reproductive skew in primitively eusocial wasps: How useful are current models? In: Hager, R., Jones, C. B., editors. *Reproductive skew in vertebrates: proximate and ultimate causes*. Cambridge: Cambridge University Press. p. 305–334.

Field, J., Cronin, A., & Bridge, C. (2006). Future fitness and helping in social queues. *Nature*, 441(7090), 214-217.

Field, J., Foster, W., Shreeves, G., & Sumner, S. (1998). Ecological constraints on independent nesting in facultatively eusocial hover wasps. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 265(1400), 973-977.

Field, J., Shreeves, G., & Sumner, S. (1999). Group size, queuing and helping decisions in facultatively eusocial hover wasps. *Behavioral Ecology and Sociobiology*, 45(5), 378-385.

Field, J., Shreeves, G., Sumner, S., & Casiraghi, M. (2000). Insurance-based advantage to helpers in a tropical hover wasp. *Nature*, 404(6780), 869-871.

Flatt, T., & Heyland, A. (Eds.). (2011). *Mechanisms of life history evolution: the genetics and physiology of life history traits and trade-offs*. OUP Oxford.

Fletcher, D. J., & Ross, K. G. (1985). Regulation of reproduction in eusocial Hymenoptera. *Annual review of entomology*, 30(1), 319-343.

Foret, S., Kucharski, R., Pellegrini, M., Feng, S., Jacobsen, S. E., Robinson, G. E., & Maleszka, R. (2012). DNA methylation dynamics, metabolic fluxes, gene splicing, and alternative phenotypes in honey bees. *Proceedings of the National Academy of Sciences*, 109(13), 4968-4973.

Fowler-Finn, K. D., & Rodríguez, R. L. (2012). The evolution of experience-mediated plasticity in mate preferences. *Journal of Evolutionary Biology*, 25(9), 1855-1863.

Friard, O., & Gamba, M. (2016). BORIS: a free, versatile open-source event-logging software for video/audio coding and live observations. *Methods in Ecology and Evolution*, 7(11), 1325-1330.

Friedman, N. R., Remeš, V., & Economo, E. P. (2019). A morphological integration perspective on the evolution of dimorphism among sexes and social insect castes. *Integrative and comparative biology*, 59(2), 410-419.

Gabriel, W. (2006). Selective advantage of irreversible and reversible phenotypic plasticity. *Archiv für Hydrobiologie*, 167(1-2), 1-20.

Gadagkar, R. (1987). Social structure and the determinants of queen status in the primitively eusocial wasp *Ropalidia cyathiformis*. *Chemistry and biology of social insects*, 377-378.

Gadau, J. (2015). Genetic architecture of key social trait differs significantly between primitive and advanced eusocial species. *Proceedings of the National Academy of Sciences*, 112(45), 13755-13756.

Gamboa, G. J., Noble, M. A., Thom, M. C., Togal, J. L., Srinivasan, R., & Murphy, B. D. (2004). The comparative biology of two sympatric paper wasps in Michigan, the native Polistes fuscatus and the invasive *Polistes dominulus* (Hymenoptera, Vespidae). *Insectes Sociaux*, 51(2), 153-157.

Geffre, A. C., Liu, R., Manfredini, F., Beani, L., Kathirithamby, J., Grozinger, C. M., & Toth, A. L. (2017). Transcriptomics of an extended phenotype: parasite manipulation of wasp social behaviour shifts expression of caste-related genes. *Proceedings of the Royal Society B: Biological Sciences*, 284(1852), 20170029.

Gelman, A., Jakulin, A., Pittau, M. G., & Su, Y. S. (2008). A weakly informative default prior distribution for logistic and other regression models. *The annals of applied statistics*, 2(4), 1360-1383.

Giray, T., Giovanetti, M., & West-Eberhard, M. J. (2005). Juvenile hormone, reproduction, and worker behavior in the neotropical social wasp *Polistes canadensis*. *Proceedings of the National Academy of Sciences*, 102(9), 3330-3335.

Gobin, B., Heinze, J., Strätz, M., & Roces, F. (2003). The energetic cost of reproductive conflicts in the ant *Pachycondyla obscuricornis*. *Journal of Insect Physiology*, 49(8), 747-752.

Gomulski, L. M., Dimopoulos, G., Xi, Z., Scolari, F., Gabrieli, P., Siciliano, P., ... & Gasperi, G. (2012). Transcriptome profiling of sexual maturation and mating in the Mediterranean fruit fly, *Ceratitis capitata*. *PloS one*, 7(1), e30857.

Gospocic, J., Shields, E. J., Glastad, K. M., Lin, Y., Penick, C. A., Yan, H., ... & Bonasio, R. (2017). The neuropeptide corazonin controls social behavior and caste identity in ants. *Cell*, 170(4), 748-759.).

Griffin, A. S., West, S. A., & Buckling, A. (2004). Cooperation and competition in pathogenic bacteria. *Nature*, 430(7003), 1024-1027.

Guo, X., Su, S., Skogerboe, G., Dai, S., Li, W., Li, Z., ... & Zhang, S. (2013). Recipe for a busy bee: microRNAs in Honey Bee caste determination. *PloS one*, 8(12), e81661.

Gurevich, A., Saveliev, V., Vyahhi, N., & Tesler, G. (2013). QUAST: quality assessment tool for genome assemblies. *Bioinformatics*, 29(8), 1072-1075.

Hadfield, J. D., Nutall, A., Osorio, D., & Owens, I. P. F. (2007). Testing the phenotypic gambit: phenotypic, genetic and environmental correlations of colour. *Journal of evolutionary biology*, 20(2), 549-557.

Haggard, C. M., & Gamboa, G. J. (1980). Seasonal variation in body size and reproductive condition of a paper wasp, *Polistes metricus* (Hymenoptera: Vespidae). *The Canadian Entomologist*, 112(3), 239-248.

Hahn, D. A., & Denlinger, D. L. (2011). Energetics of insect diapause. *Annual review of entomology*, 56, 103-121.

Hamilton, W. D. (1964a). The genetical evolution of social behaviour. I. *Journal of theoretical biology*, 7(1), 1-16.

Hamilton, W. D. (1964b). The genetical evolution of social behaviour. II. *Journal of theoretical biology*, 7(1), 17-52.

Hanschen, E. R., Shelton, D. E., & Michod, R. E. (2015). Evolutionary transitions in individuality and recent models of multicellularity. In *Evolutionary transitions to multicellular life* (pp. 165-188). Springer, Dordrecht.

Harpur, B. A., Kent, C. F., Molodtsova, D., Lebon, J. M., Alqarni, A. S., Owayss, A. A., & Zayed, A. (2014). Population genomics of the honey bee reveals strong signatures of positive selection on worker traits. *Proceedings of the National Academy of Sciences*, 111(7), 2614-2619.

Hart, L. M., & Tschinkel, W. R. (2012). A seasonal natural history of the ant, *Odontomachus brunneus*. *Insectes sociaux*, 59(1), 45-54.

Hartfelder, K., Makert, G. R., Judice, C. C., Pereira, G. A., Santana, W. C., Dallacqua, R., & Bitondi, M. M. (2006). Physiological and genetic mechanisms underlying caste development, reproduction and division of labor in stingless bees. *Apidologie*, 37(2), 144-163.

Hatch, S., Tarpy, D. R., & Fletcher, D. J. C. (1999). Worker regulation of emergency queen rearing in honey bee colonies and the resultant variation in queen quality. *Insectes Sociaux*, 46(4), 372-377.

He, X. J., Jiang, W. J., Zhou, M., Barron, A. B., & Zeng, Z. J. (2019). A comparison of honeybee (*Apis mellifera*) queen, worker and drone larvae by RNA-Seq. *Insect science*, *26*(3), 499-509.

Heinze, J., & Schrempf, A. (2012). Terminal investment: individual reproduction of ant queens increases with age. *PLoS One*, 7(4), e35201.

Helanterä, H. (2016). An organismal perspective on the evolution of insect societies. *Frontiers in Ecology and Evolution*, 4, 6.

Heyes, C. (2016). Blackboxing: social learning strategies and cultural evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371(1693), 20150369.

Hines, H. M., Hunt, J. H., O'Connor, T. K., Gillespie, J. J., & Cameron, S. A. (2007). Multigene phylogeny reveals eusociality evolved twice in vespid wasps. *Proceedings of the National Academy of Sciences*, 104(9), 3295-3299.

Hinman, V. F., & Cheatle Jarvela, A. M. (2014). Developmental gene regulatory network evolution: insights from comparative studies in echinoderms. *genesis*, *52*(3), 193-207.

Hölldobler, B., & Wilson, E. O. (1990). *The ants*. Harvard University Press.

Hölldobler, B., & Wilson, E. O. (2009). *The superorganism: the beauty, elegance, and strangeness of insect societies*. WW Norton & Company.

Holman, L. (2014a). Conditional helping and evolutionary transitions to eusociality and cooperative breeding. *Behavioral Ecology*, 25(5), 1173-1182.

Holman, L. (2014b). Caste load and the evolution of reproductive skew. *The American Naturalist*, 183(1), 84-95.

Horiguchi, R., Nozu, R., Hirai, T., Kobayashi, Y., Nagahama, Y., & Nakamura, M. (2013). Characterization of gonadal soma-derived factor expression during sex change in the protogynous wrasse, *Halichoeres trimaculatus*. *Developmental Dynamics*, 242(4), 388-399.

Huang, S., Cai, N., Pacheco, P. P., Narrandes, S., Wang, Y., & Xu, W. (2018). Applications of support vector machine (SVM) learning in cancer genomics. *Cancer Genomics-Proteomics*, 15(1), 41-51.

Hughes, C. R., & Strassmann, J. E. (1988). Age is more important than size in determining dominance among workers in the primitively eusocial wasp, *Polistes instabilis*. *Behaviour*, 107(1-2), 1-14.

Hughes, D. P., Kathirithamby, J., Turillazzi, S., & Beani, L. (2004). Social wasps desert the colony and aggregate outside if parasitized: parasite manipulation?. *Behavioral Ecology*, 15(6), 1037-1043.

Hughes, W. O., Oldroyd, B. P., Beekman, M., & Ratnieks, F. L. (2008). Ancestral monogamy shows kin selection is key to the evolution of eusociality. *Science*, 320(5880), 1213-1216.

Jandt, J. M., Thomson, J. L., Geffre, A. C., & Toth, A. L. (2015). Lab rearing environment perturbs social traits: a case study with *Polistes* wasps. *Behavioral Ecology*, 26(5), 1274-1284.

Jandt, J. M., Tibbetts, E. A., & Toth, A. L. (2014). *Polistes* paper wasps: a model genus for the study of social dominance hierarchies. *Insectes Sociaux*, 61(1), 11-27.

Javoiš, J., & Tammaru, T. (2004). Reproductive decisions are sensitive to cues of life expectancy: the case of a moth. *Animal Behaviour*, 68(2), 249-255.

Jemielity, S., Chapuisat, M., Parker, J. D., & Keller, L. (2005). Long live the queen: studying aging in social insects. Age, 27(3), 241-248.

Johnson, B. R. (2008). Within-nest temporal polyethism in the honey bee. *Behavioral Ecology and Sociobiology*, 62(5), 777-784.

Johnson, B. R., & Linksvayer, T. A. (2010). Deconstructing the superorganism: social physiology, groundplans, and sociogenomics. *The Quarterly Review of Biology*, 85(1), 57-79.

Johnson, R. N., & Starks, P. T. (2004). A surprising level of genetic diversity in an invasive wasp: *Polistes dominulus* in the northeastern United States. *Annals of the Entomological Society of America*, 97(4), 732-737.

Joiner, T. E., Hom, M. A., Hagan, C. R., & Silva, C. (2016). Suicide as a derangement of the self-sacrificial aspect of eusociality. *Psychological Review*, 123(3), 235.

Jones, B. M., Kingwell, C. J., Wcislo, W. T., & Robinson, G. E. (2017). Caste-biased gene expression in a facultatively eusocial bee suggests a role for genetic accommodation in the evolution of eusociality. *Proceedings of the Royal Society B: Biological Sciences*, 284(1846), 20162228.

Jung, K., Friede, T., & Beißbarth, T. (2011). Reporting FDR analogous confidence intervals for the log fold change of differentially expressed genes. *BMC bioinformatics*, 12(1), 288.

Kapheim, K. M. (2010). Maternal effects and the evolution of sociality in a facultatively eusocial sweat bee, *Megalopta genalis*. University of California, Los Angeles.

Kapheim, K. M., Bernal, S. P., Smith, A. R., Nonacs, P., & Wcislo, W. T. (2011). Support for maternal manipulation of developmental nutrition in a facultatively eusocial bee, *Megalopta genalis* (Halictidae). *Behavioral ecology and sociobiology*, 65(6), 1179-1190.

Kapheim, K. M., Jones, B. M., Pan, H., Li, C., Harpur, B. A., Kent, C. F., ... & Wcislo, W. T. (2020). Developmental plasticity shapes social traits and selection in a facultatively eusocial bee. *Proceedings of the National Academy of Sciences*.

Kapheim, K. M., Nonacs, P., Smith, A. R., Wayne, R. K., & Wcislo, W. T. (2015). Kinship, parental manipulation and evolutionary origins of eusociality. *Proceedings of the Royal Society B: Biological Sciences*, 282(1803), 20142886.

Kapheim, K. M., Smith, A. R., Ihle, K. E., Amdam, G. V., Nonacs, P., & Wcislo, W. T. (2012). Physiological variation as a mechanism for developmental caste-biasing in a facultatively eusocial sweat bee. *Proceedings of the Royal Society B: Biological Sciences*, 279(1732), 1437-1446.

Kathirithamby, J. (2009). Host-parasitoid associations in Strepsiptera. *Annual Review of Entomology*, 54, 227-249.

164

Kelstrup, H. C., Hartfelder, K., & Wossler, T. C. (2015). *Polistes smithii* vs. *Polistes dominula*: the contrasting endocrinology and epicuticular signaling of sympatric paper wasps in the field. *Behavioral ecology and sociobiology*, 69(12), 2043-2058.

Kelstrup, H. C., Hartfelder, K., Lopes, T. F., & Wossler, T. C. (2018). The behavior and reproductive physiology of a solitary progressive provisioning vespid wasp: evidence for a solitary-cycle origin of reproductive castes. *The American Naturalist*, 191(2), E27-E39.

Kennedy, P., Baron, G., Qiu, B., Freitak, D., Helanterä, H., Hunt, E. R., ... & Sumner, S. (2017). Deconstructing superorganisms and societies to address big questions in biology. *Trends in Ecology & Evolution*, 32(11), 861-872.

Kennedy, P., Higginson, A. D., Radford, A. N., & Sumner, S. (2018). Altruism in a volatile world. *Nature*, 555(7696), 359-362.

Kikuta, N., & Tsuji, K. (1999). Queen and worker policing in the monogynous and monandrous ant, *Diacamma* sp. *Behavioral Ecology and Sociobiology*, 46(3), 180-189.

Klein, A., Schultner, E., Lowak, H., Schrader, L., Heinze, J., Holman, L., & Oettler, J. (2016). Evolution of social insect polyphenism facilitated by the sex differentiation cascade. *PLoS genetics*, *12*(3), e1005952.

Kliebenstein, D. J. (2012). Exploring the shallow end; estimating information content in transcriptomics studies. *Frontiers in plant science*, *3*, 213.

Kokko, H., & Johnstone, R. A. (1999). Social queuing in animal societies: a dynamic model of reproductive skew. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 266(1419), 571-578.

Konczal, M., Koteja, P., Stuglik, M. T., Radwan, J., & Babik, W. (2014). Accuracy of allele frequency estimation using pooled RNA-Seq. *Molecular ecology resources*, *14*(2), 381-392.

Kopylova, E., Noé, L., & Touzet, H. (2012). SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics*, 28(24), 3211-3217.

Korb, J. (2015). Juvenile hormone: a central regulator of termite caste polyphenism. In *Advances in insect physiology* (Vol. 48, pp. 131-161). Academic Press.

Koštál, V. (2006). Eco-physiological phases of insect diapause. *Journal of insect physiology*, 52(2), 113-127.

Kovaka, S., Zimin, A. V., Pertea, G. M., Razaghi, R., Salzberg, S. L., & Pertea, M. (2019). Transcriptome assembly from long-read RNA-seq alignments with StringTie2. *Genome biology*, 20(1), 1-13.

Krause, J., Ruxton, G. D. (2002). *Living in groups*. Oxford: Oxford University Press.

Kronforst, M. R., Gilley, D. C., Strassmann, J. E., & Queller, D. C. (2008). DNA methylation is widespread across social Hymenoptera. *Current Biology*, 18(7), R287-R288.

Krueger, F. (2015). Trim galore: A wrapper tool around Cutadapt and FastQC to consistently apply quality and adapter trimming to FastQ files. Available from: https://github.com/FelixKrueger/TrimGalore.

Kucharski, R., Maleszka, J., Foret, S., & Maleszka, R. (2008). Nutritional control of reproductive status in honeybees via DNA methylation. *Science*, 319(5871), 1827-1830.

Langfelder, P., & Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. *BMC bioinformatics*, 9(1), 559.

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature methods*, 9(4), 357.

Leadbeater, E., Carruthers, J. M., Green, J. P., Rosser, N. S., & Field, J. (2011). Nest inheritance is the missing source of direct fitness in a primitively eusocial insect. *Science*, 333(6044), 874-876.

Li, B., Hou, L., Zhu, D., Xu, X., An, S., & Wang, X. (2018). Identification and caste-dependent expression patterns of DNA methylation associated genes in *Bombus terrestris*. *Scientific Reports*, 8(1), 1-10.

Li, H., Yang, M., Chen, Y., Zhu, N., Lee, C. Y., Wei, J. Q., & Mo, J. (2015). Investigation of age polyethism in food processing of the fungus-growing termite Odontotermes formosanus (Blattodea: Termitidae) using a laboratory artificial rearing system. *Journal of economic entomology*, *108*(1), 266-273.

Libbrecht, R., Corona, M., Wende, F., Azevedo, D. O., Serrão, J. E., & Keller, L. (2013). Interplay between insulin signaling, juvenile hormone, and vitellogenin regulates maternal effects on polyphenism in ants. *Proceedings of the National Academy of Sciences*, 110(27), 11050-11055.

Libbrecht, R., Oxley, P. R., & Kronauer, D. J. (2018). Clonal raider ant brain transcriptomics identifies candidate molecular mechanisms for reproductive division of labor. *BMC biology*, 16(1), 89.

Libbrecht, R., Oxley, P. R., Keller, L., & Kronauer, D. J. C. (2016). Robust DNA methylation in the clonal raider ant brain. *Current Biology*, 26(3), 391-395.

Linksvayer, T. A. (2010). Subsociality and the evolution of eusociality. In *Encyclopedia of animal behavior* (Vol. 3, pp. 358-362). Academic Press Oxford, UK.

Linksvayer, T. A., & Johnson, B. R. (2019). Re-thinking the social ladder approach for elucidating the evolution and molecular basis of insect societies. *Current opinion in insect science*, 34, 123-129.

Linksvayer, T. A., & Wade, M. J. (2005). The evolutionary origin and elaboration of sociality in the aculeate Hymenoptera: maternal effects, sib-social effects, and heterochrony. *The Quarterly Review of Biology*, *80*(3), 317-336.

Linksvayer, T. A., Fondrk, M. K., & Page Jr, R. E. (2009). Honeybee social regulatory networks are shaped by colony-level selection. *The American Naturalist*, 173(3), E99-E107.

Liu, Y., Ferguson, J. F., Xue, C., Silverman, I. M., Gregory, B., Reilly, M. P., & Li, M. (2013). Evaluating the impact of sequencing depth on transcriptome profiling in human adipose. *PLoS One*, *8*(6), e66883.

Lo, N., Simpson, S. J., & Sword, G. A. (2018). Epigenetics and developmental plasticity in orthopteroid insects. *Current opinion in insect science*, 25, 25-34.

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology*, 15(12), 550.

Luo, R., Liu, B., Xie, Y., Li, Z., Huang, W., Yuan, J., ... & Wang, J. (2012). SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. *Gigascience*, 1(1), 2047-217X.

Maleszka, R., Mason, P. H., & Barron, A. B. (2014). Epigenomics and the concept of degeneracy in biological systems. *Briefings in functional genomics*, 13(3), 191-202.

Manfredini, F., Brown, M. J., & Toth, A. L. (2018). Candidate genes for cooperation and aggression in the social wasp *Polistes dominula*. *Journal of Comparative Physiology A*, 204(5), 449-463.

Manfredini, F., Romero, A. E., Pedroso, I., Paccanaro, A., Sumner, S., & Brown, M. J. (2017). Neurogenomic signatures of successes and failures in life-history transitions in a key insect pollinator. *Genome biology and evolution*, 9(11), 3059-3072.

Mangiafico, S. 2018. rcompanion: Functions to Support Extension Education Program Evaluation. R package version 2.0.0.

Marçais, G., & Kingsford, C. (2011). A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. *Bioinformatics*, 27(6), 764-770.

Marinov, G. K., Williams, B. A., McCue, K., Schroth, G. P., Gertz, J., Myers, R. M., & Wold, B. J. (2014). From single-cell to cell-pool transcriptomes: stochasticity in gene expression and RNA splicing. *Genome research*, *24*(3), 496-510.

Marshall, J. A. (2011). Group selection and kin selection: formally equivalent approaches. *Trends in Ecology & Evolution*, 26(7), 325-332.

Matson, C. K., Murphy, M. W., Sarver, A. L., Griswold, M. D., Bardwell, V. J., & Zarkower, D. (2011). DMRT1 prevents female reprogramming in the postnatal mammalian testis. *Nature*, 476(7358), 101-104.

McAuliffe, K., & Whitehead, H. (2005). Eusociality, menopause and information in matrilineal whales. *Trends in Ecology & Evolution*, 20(12), 650.

Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., Leisch, F., Chang, C. C., & Lin, C. (2015). Misc functions of the department of statistics, probability theory group (formerly: E1071). R package version 1.6-8. https://CRAN.R-project.org/package=e1071.

Michener, C. D., & Brothers, D. J. (1974). Were workers of eusocial Hymenoptera initially altruistic or oppressed?. *Proceedings of the National Academy of Sciences*, 71(3), 671-674.

Mikát, M., Franchino, C., & Rehan, S. M. (2017). Sociodemographic variation in foraging behavior and the adaptive significance of worker production in the facultatively social small carpenter bee, *Ceratina calcarata. Behavioral Ecology and Sociobiology*, 71(9), 135.

Milanez-Almeida, P., Martins, A. J., Germain, R. N., & Tsang, J. S. (2020). Cancer prognosis with shallow tumor RNA sequencing. *Nature Medicine*, *26*(2), 188-192.

Miyano, S. (1986). Colony development, worker behavior and male production in orphan colonies of a Japanese paper wasp, *Polistes chinensis antennalis* Pérez (Hymenoptera: Vespidae). *Researches on Population Ecology*, 28(2), 347.

Miyano, S. (1991). Worker reproduction and related behavior in orphan colonies of a Japanese paper wasp, *Polistes jadwigae* (Hymenoptera, Vespidae). *Journal of Ethology*, 9(2), 135-146.

Modlmeier, A. P., Colman, E., Hanks, E. M., Bringenberg, R., Bansal, S., & Hughes, D. P. (2019). Ant colonies maintain social homeostasis in the face of decreased density. *Elife*, 8, e38473.

Monnin, T., Cini, A., Lecat, V., Fédérici, P., & Doums, C. (2009). No actual conflict over colony inheritance despite high potential conflict in the social wasp *Polistes dominulus*. *Proceedings of the Royal Society B: Biological Sciences*, 276(1662), 1593-1601.

Morandin, C., Dhaygude, K., Paviala, J., Trontti, K., Wheat, C., & Helanterä, H. (2015). Caste-biases in gene expression are specific to developmental stage in the ant *Formica exsecta. Journal of Evolutionary Biology*, 28(9), 1705-1718.

Morandin, C., Tin, M. M., Abril, S., Gómez, C., Pontieri, L., Schiøtt, M., ... & Mikheyev, A. S. (2016). Comparative transcriptomics reveals the conserved building blocks involved in parallel evolution of diverse phenotypic traits in ants. *Genome biology*, *17*(1), 1-19.

Muir, P., Li, S., Lou, S., Wang, D., Spakowicz, D. J., Salichos, L., ... & Gerstein, M. (2016). The real cost of sequencing: scaling computation to keep pace with data generation. *Genome biology*, *17*(1), 1-9.

Murren, C. J., Auld, J. R., Callahan, H., Ghalambor, C. K., Handelsman, C. A., Heskel, M. A., ... & Pfennig, D. W. (2015). Constraints on the evolution of phenotypic plasticity: limits and costs of phenotype and plasticity. *Heredity*, 115(4), 293-301.

Nattrass, S., Croft, D. P., Ellis, S., Cant, M. A., Weiss, M. N., Wright, B. M., ... & Franks, D. W. (2019). Postreproductive killer whale grandmothers improve the survival of their grandoffspring. *Proceedings of the National Academy of Sciences*, 116(52), 26669-26673.

Neumann, C., Duboscq, J., Dubuc, C., Ginting, A., Irwan, A. M., Agil, M., ... & Engelhardt, A. (2011). Assessing dominance hierarchies: validation and advantages of progressive evaluation with Elo-rating. *Animal Behaviour*, 82(4), 911-921.

Nichols, H. J., Bell, M. B. V., Hodge, S. J., & Cant, M. A. (2012). Resource limitation moderates the adaptive suppression of subordinate breeding in a cooperatively breeding mongoose. *Behavioral Ecology*, 23(3), 635-642.

Nijhout, H. F. (1998). *Insect hormones*. Princeton University Press.

Noble, W. S. (2006). What is a support vector machine?. *Nature biotechnology*, 24(12), 1565-1567.

Nonacs, P., & Hager, R. (2011). The past, present and future of reproductive skew theory and experiments. *Biological Reviews*, 86(2), 271-298.

Nowak, M. A., Tarnita, C. E., & Wilson, E. O. (2010). The evolution of eusociality. *Nature*, 466(7310), 1057-1062.

Oldroyd, B. P. (2002). The Cape honeybee: an example of a social cancer. *Trends in Ecology & Evolution*, 17(6), 249-251.

Olson, R. S., Haley, P. B., Dyer, F. C., & Adami, C. (2015). Exploring the evolution of a trade-off between vigilance and foraging in group-living organisms. *Royal Society open science*, 2(9), 150135.

Pardi, L. (1942) Ricerche sui Polistini. 5. La poliginia iniziale di *Polistes gallicus. Bollettino dell'Istituto di Entomologia dell'Universita di Bologna* 14: 1-106.

Pardi, L. (1946) Ricerche sui Polistini. 7. La «dominazione» e il ciclo ovarico annuale in *Polistes gallicus. Bollettino dell'Istituto di Entomologia dell'Universita di Bologna* 15: 25-84.

Pardi, L. (1948). Dominance order in *Polistes* wasps. *Physiological Zoology*, 21(1), 1-13.

Patalano, S., Alsina, A., Gregorio-Rodriguez, C., Bachman, M., Dreier, S., Hernando-Herraez, I., ... & Rulands, S. (2020). Specialisation and plasticity in a primitively social insect. *bioRxiv*. Preprint at https://www.biorxiv.org/content/10.1101/2020.03.31.007492v1 (2020).

Patalano, S., Hore, T. A., Reik, W., & Sumner, S. (2012). Shifting behaviour: epigenetic reprogramming in eusocial insects. *Current opinion in cell biology*, 24(3), 367-373.

Patalano, S., Vlasova, A., Wyatt, C., Ewels, P., Camara, F., Ferreira, P. G., ... & González-Navarrete, I. (2015). Molecular signatures of plastic phenotypes in two eusocial insect species with simple societies. *Proceedings of the National Academy of Sciences*, *112*(45), 13970-13975.

Patro, R., Duggal, G., Love, M. I., Irizarry, R. A., & Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. *Nature methods*, 14(4), 417-419.

Peeters, C., & Molet, M. (2010). Colonial reproduction and life histories. *Ant ecology*, 159-176.

Penick, C. A., Prager, S. S., & Liebig, J. (2012). Juvenile hormone induces queen development in late-stage larvae of the ant *Harpegnathos saltator*. *Journal of insect physiology*, 58(12), 1643-1649.

Peregrín-Alvarez, J. M., Sanford, C., & Parkinson, J. (2009). The conservation and evolutionary modularity of metabolism. *Genome biology*, *10*(6), 1-17.

Pratte, M., Gervet, J., & Theraulaz, G. (1990). Behavioural profiles in *Polistes dominulus* (Christ) wasp societies: a quantitative study. *Behaviour*, 113(3-4), 223-249.

Qiu, B., Larsen, R. S., Chang, N. C., Wang, J., Boomsma, J. J., & Zhang, G. (2018). Towards reconstructing the ancestral brain gene-network regulating caste differentiation in ants. *Nature ecology & evolution*, *2*(11), 1782-1791.

Queller, D. C., Zacchi, F., Cervo, R., Turillazzi, S., Henshaw, M. T., Santorelli, L. A., & Strassmann, J. E. (2000). Unrelated helpers in a social insect. Nature, 405(6788), 784-787.

Quiñones, A. E., & Pen, I. (2017). A unified model of Hymenopteran preadaptations that trigger the evolutionary transition to eusociality. *Nature communications*, 8(1), 1-13.

Ragsdale, J. E. (1999). Reproductive skew theory extended: the effect of resource inheritance on social organization. *Evolutionary Ecology Research*, 1(7), 859-874.

Ratnieks, F. L., & Wenseleers, T. (2008). Altruism in insect societies and beyond: voluntary or enforced?. *Trends in ecology & evolution*, 23(1), 45-52.

Ratnieks, F. L., Foster, K. R., & Wenseleers, T. (2006). Conflict resolution in insect societies. *Annu. Rev. Entomol.*, 51, 581-608.

Réale, D., Dingemanse, N. J., Kazem, A. J., & Wright, J. (2010). Evolutionary and ecological approaches to the study of personality.

Reeve, H. K. (1991). *Polistes*. In: Ross, K. G., Matthews, R. W., editors. *The social biology of wasps*. Ithaca: Cornell University Press.

Reeve, H. K., & Gamboa, G. J. (1987). Queen regulation of worker foraging in paper wasps: a social feedback control system (*Polistes fuscatus*, Hymenoptera: Vespidae). *Behaviour*, 102(3-4), 147-167.

Reeve, H. K., Starks, P. T., Peters, J. M., & Nonacs, P. (2000). Genetic support for the evolutionary theory of reproductive transactions in social wasps. Proceedings of the Royal Society of London. *Series B: Biological Sciences*, 267(1438), 75-79.

Rehan, S. M., & Richards, M. H. (2010). Nesting biology and subsociality in *Ceratina calcarata* (Hymenoptera: Apidae). *The Canadian Entomologist*, 142(1), 65-74.

Rehan, S. M., & Toth, A. L. (2015). Climbing the social ladder: the molecular evolution of sociality. *Trends in ecology & evolution*, 30(7), 426-433.

Rehan, S. M., Berens, A. J., & Toth, A. L. (2014). At the brink of eusociality: transcriptomic correlates of worker behaviour in a small carpenter bee. *BMC evolutionary biology*, 14(1), 1-11.

Rehan, S. M., Glastad, K. M., Steffen, M. A., Fay, C. R., Hunt, B. G., & Toth, A. L. (2018). Conserved genes underlie phenotypic plasticity in an incipiently social bee. *Genome biology and evolution*, 10(10), 2749-2758.

Rehan, S. M., Schwarz, M. P., & Richards, M. H. (2011). Fitness consequences of ecological constraints and implications for the evolution of sociality in an incipiently social bee. *Biological Journal of the Linnean Society*, 103(1), 57-67.

Rittschof, C. C., & Robinson, G. E. (2014). Genomics: moving behavioural ecology beyond the phenotypic gambit. *Animal Behaviour*, 92, 263-270.

Rittschof, C. C., Bukhari, S. A., Sloofman, L. G., Troy, J. M., Caetano-Anollés, D., Cash-Ahmed, A., ... & Stubbs, L. (2014). Neuromolecular responses to social challenge: Common mechanisms across mouse, stickleback fish, and honey bee. *Proceedings of the national Academy of Sciences*, 111(50), 17929-17934.

Robinson, M. D., McCarthy, D. J., & Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, 26(1), 139-140.

Roisin, Y. (1990). Queen replacement in the termite *Microcerotermes papuanus*. *Entomologia experimentalis et applicata*, 56(1), 83-90.

Roisin, Y. (2000). *Diversity and evolution of caste patterns in termites: evolution, sociality, symbioses, ecology* (pp. 95-119). Springer, Dordrecht.

Rottler-Hoermann, A. M., Schulz, S., & Ayasse, M. (2016). Nest wax triggers worker reproduction in the bumblebee *Bombus terrestris. Royal Society open science*, 3(1), 150599.

Roy, S., Saha, T. T., Zou, Z., & Raikhel, A. S. (2018). Regulatory pathways controlling female insect reproduction. *Annual review of entomology*, 63, 489-511.

Rubenstein, D. R. (2012). Family feuds: social competition and sexual conflict in complex societies. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1600), 2304-2313.

Rubin, H. (2016). The phenotypic gambit: selective pressures and ESS methodology in evolutionary game theory. *Biology & Philosophy*, 31(4), 551-569.

Sasaki, K., & Harada, M. (2020). Dopamine production in the brain is associated with caste-specific morphology and behavior in an artificial intermediate honey bee caste. *PloS one*, *15*(12), e0244140.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., ... & Cardona, A. (2012). Fiji: an open-source platform for biological-image analysis. *Nature methods*, 9(7), 676-682.

Schmitz, J., & Moritz, R. F. (1998). Molecular phylogeny of Vespidae (Hymenoptera) and the evolution of sociality in wasps. *Molecular Phylogenetics and Evolution*, 9(2), 183-191.

Schneider, C. A., Rasband, W. S., & Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nature methods*, 9(7), 671-675.

Sebé-Pedros, A., Degnan, B. M., & Ruiz-Trillo, I. (2017). The origin of Metazoa: a unicellular perspective. *Nature Reviews Genetics*, 18(8), 498.

Segal, N. H., Pavlidis, P., Antonescu, C. R., Maki, R. G., Noble, W. S., DeSantis, D., ... & Cordon-Cardo, C. (2003). Classification and subtype prediction of adult soft tissue sarcoma by functional genomics. *The American journal of pathology*, 163(2), 691-700.

Shell, W. A., & Rehan, S. M. (2018). Behavioral and genetic mechanisms of social evolution: insights from incipiently and facultatively social bees. *Apidologie*, 49(1), 13-30.

Sherman, P. W., Lacey, E. A., Reeve, H. K., & Keller, L. (1995). The eusociality continuum. *Behavioral Ecology*, 6(1), 102-108.

Shorter, J. R., & Rueppell, O. (2012). A review on self-destructive defense behaviors in social insects. *Insectes sociaux*, 59(1), 1-10.

Shreeves, G., & Field, J. (2002). Group size and direct fitness in social queues. *The American Naturalist*, 159(1), 81-95.

Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., & Zdobnov, E. M. (2015). BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*, 31(19), 3210-3212.

Simola, D. F., Graham, R. J., Brady, C. M., Enzmann, B. L., Desplan, C., Ray, A., ... & Berger, S. L. (2016). Epigenetic (re) programming of caste-specific behavior in the ant *Camponotus floridanus. Science*, 351(6268).

Simola, D. F., Wissler, L., Donahue, G., Waterhouse, R. M., Helmkampf, M., Roux, J., ... & Gadau, J. (2013). Social insect genomes exhibit dramatic evolution in gene composition and regulation while preserving regulatory features linked to sociality. *Genome research*, 23(8), 1235-1247.

Smedley, D., Haider, S., Ballester, B., Holland, R., London, D., Thorisson, G., & Kasprzyk, A. (2009). BioMart–biological queries made easy. *BMC genomics*, 10(1), 22.

Smedley, D., Haider, S., Durinck, S., Pandini, L., Provero, P., Allen, J., ... & Kasprzyk, A. (2015). The BioMart community portal: an innovative alternative to large, centralized data repositories. *Nucleic acids research*, 43(W1), W589-W598.

Smith, A. R., Kapheim, K. M., O'Donnell, S., & Wcislo, W. T. (2009). Social competition but not subfertility leads to a division of labour in the facultatively social sweat bee *Megalopta genalis* (Hymenoptera: Halictidae). *Animal Behaviour*, 78(5), 1043-1050.

Smith, C. R., Morandin, C., Noureddine, M., & Pant, S. (2018). Conserved roles of Osiris genes in insect development, polymorphism and protection. *Journal of evolutionary biology*, *31*(4), 516-529.

Smith, A. R., Wcislo, W. T., & O'donnell, S. (2003). Assured fitness returns favor sociality in a mass-provisioning sweat bee, *Megalopta genalis* (Hymenoptera: Halictidae). *Behavioral Ecology and Sociobiology*, 54(1), 14-21.

Smith, M. L., Koenig, P. A., & Peters, J. M. (2017). The cues of colony size: how honey bees sense that their colony is large enough to begin to invest in reproduction. *Journal of Experimental Biology*, 220(9), 1597-1605.

Spannhoff, A., Kim, Y. K., Raynal, N. J. M., Gharibyan, V., Su, M. B., Zhou, Y. Y., ... & Bedford, M. T. (2011). Histone deacetylase inhibitor activity in royal jelly might facilitate caste switching in bees. *EMBO reports*, 12(3), 238-243.

Spencer, K. A., Buchanan, K. L., Goldsmith, A. R., & Catchpole, C. K. (2004). Developmental stress, social rank and song complexity in the European starling (*Sturnus vulgaris*). *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 271, S121-S123.

Standage, D. S., Berens, A. J., Glastad, K. M., Severin, A. J., Brendel, V. P., & Toth, A. L. (2016). Genome, transcriptome and methylome sequencing of a primitively eusocial wasp reveal a greatly reduced DNA methylation system in a social insect. *Molecular Ecology*, 25(8), 1769-1784.

Starks, P. T., Turillazzi, S., & West-Eberhard, M. J. (2006). *Polistes* paper wasps: emergence of a model genus. In *Annales Zoologici Fennici* (Vol. 43, No. 5-6, pp. 385-386). Helsinki: Suomen Biologian Seura Vanamo, 1964-.

Strassmann, J. E. (1981). Parasitoids, predators, and group size in the paper wasp, *Polistes exclamans. Ecology*, 62(5), 1225-1233.

Strassmann, J. E., & Meyer, D. C. (1983). Gerontocracy in the social wasp, *Polistes exclamans. Animal Behaviour*, 31(2), 431-438.

Strassmann, J. E., Fortunato, A., Cervo, R., Turillazzi, S., Damon, J. M., & Queller, D. C. (2004). The cost of queen loss in the social wasp *Polistes dominulus* (Hymenoptera: Vespidae). *Journal of the Kansas Entomological Society*, 77(4), 343-355.

Sumner, S. (2014). The importance of genomic novelty in social evolution. *Molecular ecology*, *23*(1), 26-28.

Sumner, S., Casiraghi, M., Foster, W., & Field, J. (2002). High reproductive skew in tropical hover wasps. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 269(1487), 179-186.

Sumner, S., Kelstrup, H., & Fanelli, D. (2010). Reproductive constraints, direct fitness and indirect fitness benefits explain helping behaviour in the primitively eusocial wasp, *Polistes canadensis. Proceedings of the Royal Society B: Biological Sciences*, 277(1688), 1721-1728.

Suzuki, S. (2020). Biparental negotiation or larval begging? Determinant of male provisioning in a burying beetle (*Nicrophorus quadripunctatus*). *Behaviour*, 1(aop), 1-11.

Szathmáry, E. (2015). Toward major evolutionary transitions theory 2.0. *Proceedings of the National Academy of Sciences,* 112(33), 10104-10111.

Szathmáry, E., & Smith, J. M. (1995). The major evolutionary transitions. *Nature*, 374(6519), 227-232.

Tarazona, S., García, F., Ferrer, A., Dopazo, J., & Conesa, A. (2011). NOIseq: a RNA-seq differential expression method robust for sequencing depth biases. *EMBnet. journal,* 17(B), 18-19.

Taylor B. A., Cini A., Wyatt. C., Reuter M., & Sumner S. (in press). The molecular basis of socially-mediated phenotypic plasticity in a eusocial paper wasp. *Nature Communications*. Pre-print available online at: www.biorxiv.org/content/10.1101/2020.07.15.203943v1.

Taylor, B. A., Cini, A., Cervo, R., Reuter, M., & Sumner, S. (2020). Queen succession conflict in the paper wasp *Polistes dominula* is mitigated by age-based convention. *Behavioral Ecology*.

Taylor, B. A., Reuter, M., & Sumner, S. (2019). Patterns of reproductive differentiation and reproductive plasticity in the major evolutionary transition to superorganismality. *Current opinion in insect science*, 34, 40-47.

Tibbetts, E. A., & Huang, Z. Y. (2010). The challenge hypothesis in an insect: juvenile hormone increases during reproductive conflict following queen loss in *Polistes* wasps. *The American Naturalist*, 176(2), 123-130.

Tibbetts, E. A., & Shorter, J. R. (2009). How do fighting ability and nest value influence usurpation contests in *Polistes* wasps?. *Behavioral Ecology and Sociobiology*, 63(9), 1377-1385.

176

Tibbetts, E. A., Mettler, A., & Donajkowski, K. (2013). Nutrition-dependent fertility response to juvenile hormone in non-social *Euodynerus foraminatus* wasps and the evolutionary origin of sociality. *Journal of insect physiology*, 59(3), 339-344.

Todd, E. V., Ortega-Recalde, O., Liu, H., Lamm, M. S., Rutherford, K. M., Cross, H., ... & Gemmell, N. J. (2019). Stress, novel sex genes, and epigenetic reprogramming orchestrate socially controlled sex change. *Science advances*, 5(7), eaaw7006.

Tofilski, A. (2002). Influence of age polyethism on longevity of workers in social insects. *Behavioral Ecology and Sociobiology*, *51*(3), 234-237.

Toth, A. L., & Rehan, S. M. (2017). Molecular evolution of insect sociality: an eco-evo-devo perspective. *Annual Review of Entomology*, *62*, 419-442.

Toth, A. L., & Robinson, G. E. (2007). Evo-devo and the evolution of social behavior. *Trends in Genetics*, *23*(7), 334-341.

Toth, A. L., Tooker, J. F., Radhakrishnan, S., Minard, R., Henshaw, M. T., & Grozinger, C. M. (2014). Shared genes related to aggression, rather than chemical communication, are associated with reproductive dominance in paper wasps (*Polistes metricus*). *BMC genomics*, *15*(1), 75.

Toth, A. L., Varala, K., Henshaw, M. T., Rodriguez-Zas, S. L., Hudson, M. E., & Robinson, G. E. (2010). Brain transcriptomic analysis in paper wasps identifies genes associated with behaviour across social insect lineages. *Proceedings of the Royal Society B: Biological Sciences*, *277*(1691), 2139-2148.

Toth, A. L., Varala, K., Newman, T. C., Miguez, F. E., Hutchison, S. K., Willoughby, D. A., ... & Robinson, G. E. (2007). Wasp gene expression supports an evolutionary link between maternal behavior and eusociality. *Science*, *318*(5849), 441-444.

True, J. R., & Carroll, S. B. (2002). Gene co-option in physiological and morphological evolution. *Annual review of cell and developmental biology*, *18*(1), 53-80.

Tschinkel, W. R. (1998). Sociometry and sociogenesis of colonies of the harvester ant, *Pogonomyrmex badius*: worker characteristics in relation to colony size and season. *Insectes Sociaux*, 45(4), 385-410.

Tsuji, K., & Tsuji, N. (2005). Why is dominance hierarchy age-related in social insects? The relative longevity hypothesis. *Behavioral Ecology and Sociobiology*, 58(5), 517-526.

Turillazzi, S. (1980). Seasonal variations in the size and anatomy of *Polistes gallicus* (Hymenoptera Vespidae). *Monitore Zoologico Italiano-Italian Journal of Zoology*, 14(1-2), 63-75.

Turillazzi, S., & Francescato, E. (1990). Patrolling behaviour and related secretory structures in the males of some stenogastrine wasps (Hymenoptera, Vespidae). *Insectes sociaux,* 37(2), 146-157.

Turillazzi, S., & Pardi, L. (1977). Body size and hierarchy in polygynic nests of *Polistes gallicus* (L.)(Hymenoptera Vespidae). *Monitore Zoologico Italiano-Italian Journal of Zoology*, 11(1-2), 101-112.

Uematsu, K., Shimada, M., & Shibao, H. (2013). Juveniles and the elderly defend, the middle-aged escape: division of labour in a social aphid. *Biology letters*, 9(2), 20121053.

Van Dyken, J. D., & Wade, M. J. (2010). The genetic signature of conditional expression. *Genetics*, 184(2), 557-570.

Vellai, T., & Vida, G. (1999). The origin of eukaryotes: the difference between prokaryotic and eukaryotic cells. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 266(1428), 1571-1577.

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of statistical software*, 36(3), 1-48.

Vojvodic, S., Johnson, B. R., Harpur, B. A., Kent, C. F., Zayed, A., Anderson, K. E., & Linksvayer, T. A. (2015). The transcriptomic and evolutionary signature of social interactions regulating honey bee caste development. *Ecology and evolution*, *5*(21), 4795-4807.

Von Wyschetzki, K., Rueppell, O., Oettler, J., & Heinze, J. (2015). Transcriptomic signatures mirror the lack of the fecundity/longevity trade-off in ant queens. *Molecular biology and evolution*, 32(12), 3173-3185.

Wang, Y., Amdam, G. V., Rueppell, O., Wallrichs, M. A., Fondrk, M. K., Kaftanoglu, O., & Page Jr, R. E. (2009). PDK1 and HR46 gene homologs tie social behavior to ovary signals. *PloS one*, 4(4), e4899.

Wang, Y., Kaftanoglu, O., Siegel, A. J., Page Jr, R. E., & Amdam, G. V. (2010). Surgically increased ovarian mass in the honey bee confirms link between reproductive physiology and worker behavior. *Journal of Insect Physiology*, 56(12), 1816-1824.

Warner, M. R., Mikheyev, A. S., & Linksvayer, T. A. (2017). Genomic signature of kin selection in an ant with obligately sterile workers. *Molecular biology and evolution*, *34*(7), 1780-1787.

Warner, M. R., Qiu, L., Holmes, M. J., Mikheyev, A. S., & Linksvayer, T. A. (2019). Convergent eusocial evolution is based on a shared reproductive groundplan plus lineage-specific plastic genes. *Nature communications*, 10(1), 1-11.

Waterhouse, R. M., Seppey, M., Simão, F. A., Manni, M., Ioannidis, P., Klioutchnikov, G., ... & Zdobnov, E. M. (2018). BUSCO applications from quality assessments to gene prediction and phylogenomics. *Molecular biology and evolution*, 35(3), 543-548.

Wcislo, W. T., Arneson, L., Roesch, K., Gonzalez, V., Smith, A., & Fernández, H. (2004). The evolution of nocturnal behaviour in sweat bees, *Megalopta genalis* and *M. ecuadoria* (Hymenoptera: Halictidae): an escape from competitors and enemies?. *Biological Journal of the Linnean Society*, 83(3), 377-387.

Weiner, S. A., & Toth, A. L. (2012). Epigenetics in social insects: a new direction for understanding the evolution of castes. *Genetics research international*, 2012.

Weiner, S. A., Galbraith, D. A., Adams, D. C., Valenzuela, N., Noll, F. B., Grozinger, C. M., & Toth, A. L. (2013). A survey of DNA methylation across social insect species, life stages, and castes reveals abundant and caste-associated methylation in a primitively social wasp. *Naturwissenschaften*, 100(8), 795-799.

Weiner, S. A., Noble, K., Upton, C. T., Woods, W. A., & Starks, P. T. (2011). A role for thermoregulation in the *Polistes dominulus* invasion: a comparison of the thermoregulatory abilities of the invasive wasp *P. dominulus* and the native wasp *P. fuscatus*. *Insectes sociaux*, 58(2), 185-190.

West, S. A., Fisher, R. M., Gardner, A., & Kiers, E. T. (2015). Major evolutionary transitions in individuality. *Proceedings of the National Academy of Sciences*, 112(33), 10112-10119.

West, S. A., Griffin, A. S., & Gardner, A. (2007). Evolutionary explanations for cooperation. *Current biology*, 17(16), R661-R672.

West-Eberhard, M. J. (1969). The Social Biology of Polistine Wasps. *Miscellaneous Publications of the Museum of Zoology, University of Michigan*. 140, 1-101.

West-Eberhard, M. J. (1978). Temporary queens in *Metapolybia* wasps: nonreproductive helpers without altruism?. *Science*, 200(4340), 441-443.

West-Eberhard, M. J. (1996). Wasp societies as microcosms for the study of development and evolution. *Natural history and evolution of paper wasps*, *290*, 317.

West-Eberhard, M. J. (2003). *Developmental plasticity and evolution*. Oxford University Press.

Wills, B. D., Powell, S., Rivera, M. D., & Suarez, A. V. (2018). Correlates and consequences of worker polymorphism in ants. *Annual review of entomology,* 63, 575-598.

Wojciechowski, M., Lowe, R., Maleszka, J., Conn, D., Maleszka, R., & Hurd, P. J. (2018). Phenotypically distinct female castes in honey bees are defined by alternative chromatin states during larval development. *Genome research*, 28(10), 1532-1542.

Woodard, S. H., Bloch, G. M., Band, M. R., & Robinson, G. E. (2014). Molecular heterochrony and the evolution of sociality in bumblebees (*Bombus terrestris*). *Proceedings of the Royal Society B: Biological Sciences*, *281*(1780), 20132419.

Wyatt, C. D., Bentley, M., Taylor, D., Brock, R. E., Taylor, B. A., Bell, E., ... & Sumner, S. (2020). Genetic toolkit for sociality predicts castes across the spectrum of social complexity in wasps. *bioRxiv*.

Yan, H., Simola, D. F., Bonasio, R., Liebig, J., Berger, S. L., & Reinberg, D. (2014). Eusocial insects as emerging models for behavioural epigenetics. *Nature Reviews Genetics*, 15(10), 677-688.

Yanagihara, S., Suehiro, W., Mitaka, Y., & Matsuura, K. (2018). Age-based soldier polyethism: old termite soldiers take more risks than young soldiers. *Biology letters*, *14*(3), 20180025.

Yoshimura, H., & Yamada, Y. Y. (2018). Caste-fate determination primarily occurs after adult emergence in a primitively eusocial paper wasp: significance of the photoperiod during the adult stage. *The Science of Nature*, 105(1-2), 15.

Yu, G., Li, F., Qin, Y., Bo, X., Wu, Y., & Wang, S. (2010). GOSemSim: an R package for measuring semantic similarity among GO terms and gene products. *Bioinformatics*, 26(7), 976-978.

Zhang, M. J., Ntranos, V., & Tse, D. (2020). Determining sequencing depth in a single-cell RNA-seq experiment. *Nature communications*, *11*(1), 1-11.

Zhou, S., Mackay, T. F., & Anholt, R. R. (2014). Transcriptional and epigenetic responses to mating and aging in *Drosophila melanogaster. BMC genomics*, 15(1), 927.