Title: Functional genomics of parental care of insects.

Author: Christopher B. Cunningham

Correspondence: c.b.cunningham@swansea.ac.uk

Institutional Address: Department of Biosciences, Swansea University, Swansea SA2 8PP, Wales, UK

Abstract

Parental care was likely the first step most lineages made towards sociality. However, the molecular mechanisms that generate parental care are not broadly characterized. Insects are important as an evolutionary independent group from classic models of parental care, such as, house mice. They provide an opportunity to test the generality of our understanding. With this review, I survey the functional genomics of parental care of insects, summarize several recent advances in the broader framework for studying and understanding parental care, and finish with suggested priorities for further research. Although there are too few studies to draw definitive conclusions, I argue that natural selection appears to be rewiring existing gene networks to produce parental care, that the epigenetic mechanisms influencing parental care are not well understood, and, as an interesting early consensus, that genes strongly associated with carer/offspring interactions appear biased toward proteins that are secreted. I summarize the studies that have functionally validate candidate genes and highlight the increasing need to perform this work. I finish with arguments for both conceptual and practical changes moving forward. I argue that future work can increase the use of predictive frameworks, broaden its definition of conservation of mechanism to gene networks rather than single genes, and increase the use of more established comparative methods. I further highlight the practical considerations of standardizing analyses and reporting, increasing the sampling of both carers and offspring, better characterizing gene regulatory networks, better characterizing taxonomically restricted genes and any consistent role they have underpinning parental care, and using factorial designs to disentangle the influence of multiple variables on the expression of parental care.

Keywords: Epigenetics, Sexual Selection, Social Behavior, RNA-sequencing

Highlights

- Current studies suggest evolution is rewiring existing gene networks
- Thus far, DNA methylation has no strong role regulating parental care of insects.
- The field almost completely lacks validated genes with causal links to phenotypes.
- Standardized analyses and reporting are needed to make studies more comparable.
- Conservation of mechanism at gene networks rather than single genes is likely beneficial.
- Sampling both carers and offspring and better characterizing gene regulatory networks are needed.

Introduction.

Understanding sociality is a major goal of evolutionary biology (*sensu* Maynard Smith & Szathmáry, 1995; NESCent Working Group, 2014; Kramer & Meunier, 2018). The evolution of family units and parental behavior is generally considered to be a common first transition into sociality for many animal lineages (Kramer & Meunier, 2018). As lineages transition from solitary to social, they must overcome the costs of sociality (Kramer & Meunier, 2018). They must also generate parental phenotypes from standing genetic variation, new interactions between existing gene pathways, novel expression patterns of existing gene pathways, or *de novo* genes (reviewed in Korb & Heinz, 2016; Jones & Robinson, 2018). Therefore, parental care likely holds a pivotal position for understand sociality, as well as generally helping us better understand the mechanisms that underpin complex behaviors. This review provides a general summary of the functional genomics (differential gene expression and its regulation) of the parental care of insects from the last five years, since RNA-sequencing has become more accessible. It will include species with alloparental care where adult siblings care for offspring rather than parents (e.g., honey bees). To further focus this review and where possible, only research that directly observed parental care and sampled specific tissues of interest will be included (i.e., little focus on studies that used entire animals; see Johnson et al. (2013) or Montgomery and Mank (2016) for further discussion). For Hymenopteran species, I will bias discussion towards research that looked at difference of behavior within one caste (i.e., little focus on queen vs worker comparisons). Although there are a limited number of studies, some possible early conclusions will be highlighted.

I will first comment on how parental care will be defined here (i.e., as a summation of many individual traits), why invertebrates are well suited to answer some questions, and why functional genomics is a relevant area to understand the mechanics of parental care. I then survey the current state of the field and then will propose considerations for how the field can move forward. My survey will provide a concise review of 1) experimental transcriptomic work of carers, 2) experimental epigenetic work of carers, 3) experimental work when both carer and offspring are assessed, and 4) the functional validation of targets through genetic manipulations. I finish with arguments for future directions. Theses highlight conceptual advances that should be integrated into future studies and also highlight some practical considerations.

Parental Care

As a higher order trait, parental care is produced from an integration of many external and internal signals with an individual's internal state. It is a categorization and not a behavior *per se* with actual phenotypes being the individual behaviors combining to produce the overall "parenting" state (i.e., parenting is produced from individuals modifying their feeding, sociality, aggression, reproductive, etc. behaviors; Cunningham et al., 2017; Kohl et al., 2018; Moore and Benowitz, 2019). Here, parental care will be defined as a behavior that is directed towards offspring, that increases offspring fitness, and that originated and is maintained for that purpose (Klug et al., 2012).

Why invertebrates?

Understanding the evolution of parental care requires sampling animals of many lineages with highly analogous behaviors and using a strict comparative framework to evaluate patterns (NESCent Working

Group, 2014; Fisher et al., 2019). Beyond just providing a more complete understanding of the evolution of any trait, comparative work is of increasing importance as we seek to test the generality of our understanding from classical models with novel species expressing analogous behaviors (e.g., Phelps et al., 2010; Kelly & Ophir, 2018; Fischer et al., 2019). Invertebrates offer a great opportunity to test our understanding as they have many independent lineages that have evolved parental care, from very simple indirect care (food provisioning, e.g., dung or carrion beetles) to elaborate and extended direct parental care (feeding offspring from specialized tissues, e.g., honey bees; Trumbo, 2012). Furthermore, both the ultimate causes and proximate mechanisms must be evaluated to completely understand the evolution of parental care (Boake et al., 2002; Bateson and Laland, 2013; NESCent Working Group, 2014, 2015). Ultimate and proximate causes underpin the evolution of a phenotype and the mechanisms producing a phenotype, respectively (Bateson and Laland, 2013). Much work pursues the evolutionary causes and mechanisms of parental care of vertebrates (Dulac et al., 2014; Fisher et al., 2019). This is mirrored for invertebrates for the ultimate causes (Kramer & Meunier, 2018), but only recently have more than a handful of species been evaluated for the proximate mechanisms that underpin parental care.

Why functional genomics?

Parental care can be studied at many levels (e.g., selective forces, ecological, behavioral, physiological, molecular) or by assaying many molecular mechanisms (e.g., physiological, endocrine, neural, genetic). Here, I will focus on gene expression, both the gene networks that are (de)activated or modulated during parental care and the mechanisms performing these actions, namely epigenetic mechanisms (DNA methylation, histone tail post-translational modifications, non-coding RNA). This focus is appropriate for several reasons, some of which are conceptual and some of which are practical. First, behavior, like any other trait, is ultimately traceable to when and what genes are collectively expressed (Boake, et al., 2002; Cardoso et al., 2015). Second, differences of gene expression are an excellent predictor of differences of protein abundance after responding to a stimulus to change a cellular state (e.g., Jovanovic, et al., 2015; Koussounadis et al., 2015). For example, gene expression changes explain more than 90% of the variation of protein abundance changes in response to environmental stimuli (Jovanovic, et al., 2015). This highly predictive association is important to note as proteins are the actual direct effectors of cellular functions (Evans, 2015). This contrasts with gene expression being a relatively poor predictor of standing or maintenance protein abundance (e.g., Vogel & Marcotte, 2012; Franks et al., 2016). Third, although other mechanisms that influence behavior are also important (e.g., neurotransmitters, hormones), studies of gene expression are found among many taxa as they are well suited to capturing global genome dynamics that control many phenotypes (Evans, 2015; Ritschoff & Hughes, 2018; Stark et al., 2019). Studies of gene expression are also particularly well suited to capture the broad temporal range over which behavior can shift (Ritschoff & Hughes, 2018). Fourth, it is efficient way to gain maximum comparative insight when adding a new species to examine the mechanisms of a phenotype due to its tractability within non-model organisms. More practically, gene expression investigated with RNA-sequencing will likely continue to be the go-to first step when interrogating the molecular mechanisms of a phenotype. This is due to its ease of use, standardized practices (Conesa et al., 2016), and well-characterized technical issues allowing for efficient experimental design (Todd et al., 2016). These factors make a review based on this technology and other next-generation sequencing technologies useful for the future.

Experimental work investigating parental care - Transcriptomes.

There are seven species of insects that have been made molecularly tractable and have been used to explore the transcriptomic basis of parental of insects; the burying beetles *Nicrophorus vespilloides* and *Nicrophorus orbicollis*, the pharaoh ant *Monomorium pharaonis*, the clonal raider ant *Ooceraea biroi*, the carpenter bee *Ceratina calcarata*, the ant *Temnothorax logispinosus*, the European earwig *Forficula auricularia*, and the Western honey bee *Apis mellifera*. The honey bee is the most well established and several other reviews of their behavior exists (e.g., Zayed and Robinson, 2012). Therefore, this review will only highlight more recent work of Western honey bee transcriptomics. Where there is enough work, I group my discussion by species; otherwise, topics are reviewed with species aggregated.

Transcriptomic architecture of carers. Distinct transcriptional profiles are associated with distinct behavioral states (reviewed in Zayed and Robinson, 2012; Cardoso et al., 2015). This empirically supported thesis is the basis of functional genomic studies of behavior over the past decade. A take home message is that there are studies of the genomic basis of parental care of insects, but only preliminary conclusions are possible. This is mostly driven because studies are spread over many species and life history stages (larvae, adults, castes). It is difficult to know what differences are attributable to lineage, life history, or style of behavioral transition (permanent or reversible). This makes rigorous testing of alternative explanations premature, but some early consensus conclusions can be proposed. There are some similarities with metabolic pathways being repeatedly detected and with specific genes or gene families being repeatedly detected, such as, *vitellogenin* (discussed below).

Another take-home message is that selection seems to be coopting existing gene networks to produce parental care. In this way, parental care is mirroring the conclusions of evo-devo studies of morphological traits (e.g, Wittkopp and Kalay, 2012). However, a rigorous test of the proportion of conserved vs non-conserved genes underpinning parental care is not currently possible despite this being a central question of the field for many years (e.g., Robinson and Ben-Shahar, 2002)

Burying Beetles. Burying beetles are subsocial beetles that bury carcasses, keep them from decaying, and directly regurgitate food to offspring. Two species are currently molecularly tractable, *N. vespilloides* and *N. orbicollis*. Both males and females of these species care for offspring. There are a number of studies that have looked at the molecular mechanisms of this parental care. The first contrasted uniparental males and females to biparental males and females (Parker et al., 2015). They showed that there was ~750 genes differentially expressed between uniparental caring and before caring of both sexes with surprisingly little overlap (26.9%) given that each sex was preforming the same behaviors. Differentially expressed genes during parental care were enriched for food processing and included several *vitellogenin*'s (Parker et al., 2015). Benowitz and colleagues (2017) looked for a quantitative association between levels of direct care (high or low regurgitation of food to offspring) and gene expression. They used both sexes under uniparental conditions and two species, *N. vespilloides* and *N. orbicollis*. They did not find a strong transcriptional signature between sexes, species, or levels of care. However, using gene set enrichment analysis, they did find that the top 100 differentially expressed genes of *N. vespilloides* females were also enriched in *N. vespilloides* and *N. orbicollis* males. They also found that the differentially expressed genes of the Parker et al. (2017) study associated between non-care and care were enriched in both sexes of *N. vespilloides* between the two level of provisioning. They suggested that variation of care is likely linked to many subtle transcriptional differences and that genes responsible for behavioral shifts (non-caring vs caring) are also involved in variation of care behavior. Cunningham and colleagues (2019) used the ability of males to take over direct care when their female mate separate

the influence of care, social context, and behavioral flexibility on gene expression. They found ~550, ~100, and ~20 genes associated with these variables, respectively, showing the greatest influence is from behavior. Biological function of these genes was associated with metabolism (Cunningham et al., 2019). The general trends here are broadly consistent with reuse of established gene networks produce parental phenotypes.

Pharaoh Ant. The pharaoh ant (*M. pharaonis*) is a eusocial species that displays task specialization based on age, from brood care to foraging, very similar to Western honey bees. Several studies have been conducted on the molecular underpinnings of these tasks. We will focus on gene expression associated with brood care. Mikheyev and Linksvayer (2015) studied the transition from brood care to foraging and found ~2,500 genes differed between the tasks. Nurse up-regulated genes were highly enriched for metabolic functions, while forager genes were underrepresented for metabolic functions. Fourteen of their gene co-expression networks switched the direction of their association within three days with the task changes as workers aged and switched tasks suggesting gene networks were being turned on and off across the transition. They also showed that nurse upregulated genes had higher rates of molecular evolution and were less well connected within their gene co-expression networks (Mikheyev and Linksvayer, 2015). Further concentrating on just brood care, Walsh and colleagues (2018) showed that there are behavioral specializations of nurses based on larval developmental stage and this is mirrored with a transcriptomic signature of this specialization. They found ~210 genes differentially expressed between nurses specializing on old vs young larvae and, as before, the genes were highly enriched for metabolic functions and for secreted proteins, proteins made specifically to be exported from cells (Walsh et al., 2018). Warner and colleagues (2019a) tried to better understand the basis of age-based task specialization in both pharaoh ant and honey bees. However, they found few differentially expressed genes shared between the heads of nurses and foragers of each species (405 for *M. pharaonis* and 927 for *A. mellifera*). Abdominal gene expression between the castes of the two species was more similar (Warner et al., 2019a). Warner and colleagues (2019b) also collected just nurses and the larvae they were caring for through a developmental time-series. For the nurses, they found ~2,650 differentially expressed genes through the developmental time-series. This study reinforced earlier findings of nursing genes being enriched for genes that produce secreted proteins and being less evolutionarily constrained (i.e., higher rates of non-synonymous substitutions; Warner et al., 2019b). It might be that during a major evolutionary transition, such as, a transition from solitary to social, there is an increased likelihood of evolving novel genes as novel master regulators of established regulatory network. These studies are broadly consistent with the reuse of gene networks underpinning brood care but do suggest that some lineage specific genes contribute.

Clonal Raider Ant. The clonal raider ant (*O. biroi*) is a eusocial species whose colonies alternate between asexual reproductive and brood care/foraging phases and these phases can be manipulated through the introduction or removal of eggs or larvae. Libbrecht and colleagues (2018) conducted an experiment where they added larvae or pupae to broods that had just initiated the opposite phase. They sampled the manipulated broods at 6, 12, 24, 48, and 96 hours after manipulation. They showed differences in the timing of gene expression changes transitioning into and out of brood care, gene co-expression networks associated with each transition (27 and 35 gene co-expression networks, respectively), and a core set of known transcription factors enriched within these networks. The transcription factor *forkhead* was a noted candidate master regulator being enriched within some networks for both types of phase transitions (Libbrecht et al., 2018). In the reproductive to brood care phase

transition, the top 40 differentially expressed genes contained many genes linked to neural functions and metabolism.

The carpenter bee *C. calcarata*. *Ceratina calcarata* is a subsocial bee. The founding mother will guard, forage, and groom offspring until late into the reproductive season when the oldest and morphologically distinct offspring is forced into taking over care of siblings. Rehan and colleagues (2014) sampled many stages, but for our concerns here, sampled caring mothers and caring oldest daughters over the course of a breeding cycle. They found few genes associated with care (~180) and most were unannotated. However, enriched gene ontology terms were associated with neural function, protein regulation, and gene regulation. In this case, it is possible or likely the lack of annotation is driven by this being an early study. One gene of interest was an odorant binding protein and that is consistent with the role of olfaction being an important mediator of social behavior (e.g., Trible et al., 2017). In a similar analysis, Shell and Rehan (2019) assayed mothers over a breeding season. Differentially expressed genes for caring behaviors were identified (guarding/grooming: ~1,140; nesting: ~330) and were well conserved across bees. They noted that genes differentially expressed during guarding were well conserved compared other behavioral states. Enriched biological processes were metabolic, immune, and neural functions.

The ant *T. longispinosus*. *Temnothorax longispinosus* is an eusocial species of long-lived individuals whose workers specialize on brood care or foraging based on age. Kohlmeier and colleagues (2019) were able to manipulate age and task and control for fertility status (workers are not completely sterile) to disentangle the influence of behavior from other variables. They found thousands of genes associated with brood care (~3,600 genes), including *vitellogenin*'s. These gene were enriched for biological functions of DNA integrity and metabolic functions (oxidative stress response, lipid processing). Interestingly, behavior had an outsized influence on the transcriptome with only ~370 and ~70 genes associated with age and fertility status, respectively. This mirrors the same conclusion from Cunningham and colleagues (2019) who contrasted behavior, social context, and behavioral flexibility.

The European earwig. *Forficula auriculari* is a sub-social species of earwig. Females guard eggs and provision food to hatched offspring. Contrasting mothers that were removed from eggs, allowed to guard eggs, or allowed to guard eggs and interact with offspring, Wu and colleagues (2020) found ~1,600 genes associated with care across four tissues and ~700 were within the head. The differentially expressed genes included *vitellogenin*'s, many metabolic genes, but the majority were uncharacterized. Here, there seems to be a mix of conserved and uncharacterized genes underpinning care. The uncharacterized genes might truly be novel or might be driven by annotation problems as this is a new species to this field.

The Western honey bee. *Apis mellifera* is a eusocial species whose workers show age-based task specialization with young workers providing brood care and older workers foraging. This is the best-studied and characterized species reviewed here. Hamilton and colleagues (2019) reanalyzed much of the functional genomics data to produce a more complete picture of how several behaviors are regulated (i.e., brood care, foraging, aggression). They found twenty, known transcription factors could explain much of the transcriptional variation associated with behavioral state. They further showed that for five of the twenty transcription factors variation in downstream differentially expressed was an interaction between the expression of the transcription factors and the specific behavior expressed. These genes showed strong changes in their quantitative relationship to target genes or changes in the direction of their relationship with target genes dependent on the behavioral state (Hamilton et al., 2019). This is strong evidence of

context specific cooption of transcription factor to regulate new behaviors, but again suggests that there is a rewiring of gene networks rather than new genes producing parental care.

Experimental work investigating parental care - Epigenetics.

Social behavior is highly responsive to changes of social context and an individual's internal state (Bailey et al., 2018; Bludau et al., 2019). Across the animal kingdom, much of this responsiveness is mediated through epigenetic mechanisms, including DNA methylation, histone post-translational modifications, and regulatory RNAs (Bludau et al., 2019). There are also several lesser-studied mechanisms of epigenetic mechanisms, such as RNA modifications, RNA-directed DNA methylation, and genome architecture (Robine et al., 2009; Lewis et al., 2018; Bludau et al., 2019; Rubenstein et al., 2019). They mediate gene expression by either modifying the access of transcription factors to regulatory elements or the abundance of mRNAs through degradation (Bludau et al., 2019). Due to their influence on gene expression, studies of the epigenetic regulation of parental care has increased quickly in insects.

The role of DNA methylation regulating behavior is the most studied epigenetic mechanism within insects. The take home message is that there is currently little evidence that DNA methylation is associated with changes of behavior or with parental care within insects. Between genes, high DNA methylation is associated with broad and high levels of gene expression, a difference from vertebrates (Libbrecht et al., 2016; Cunningham et al., 2019). Within a gene, difference of DNA methylation are not associated with difference of gene expression, including for Western honey bees. However, it is an emerging field so strong conclusions are not yet possible.

Burying Beetles & Clonal Raider Ants. I am grouping these two species as their studies both produced the same conclusion that differences of DNA methylation were not associated with differences of behavior. Cunningham and colleagues (2019), using *N. vespilloides*, found no difference of DNA methylation associated with males that either directly feed or did not feed offspring. Libbrecht and colleagues (2016) investigated the association of DNA methylation between the reproductive and brood care phases of *O. biroi* and found no link between DNA methylation and the two phases.

Buff-tailed Bumblebee. The buff-tailed bumblebee (*Bombus terrestis*) is an eusocial species whose workers preform both brood care or forage along a gradient, with individuals biasing their behavior between the two tasks over a period of days. There is also a set of reproductive individuals which produce haploid sons (Marshall et al., 2019). Porath and colleagues (2019) looked at the A-to-I RNA editing and found ~150 sites that had high levels of evidence of editing. Editing was biased towards nursing state at both the individual level and with averaged values; although, very few individual sites were statistically different between nurses and foragers (Porath et al., 2019). Protein-coding genes with RNA editing were enriched for ion channels lending credence to this mechanism possibly influencing behavior. The influence this has on behavior is unknown currently. Further, but not directly associated with brood care, Marshall and colleagues (2019) using whole heads found no association between differential methylation of genes and differential expression of genes between reproductive and sterile workers. Marshall and colleagues (2020) also found little overlap between differences of allele-specific DNA methylation and difference of allele-specific gene expression between reproductive and sterile workers.

The carpenter bee *C. calcarata*. Understanding offspring is also a part of understanding the mechanistic basis of parental care. To understand the response of offspring to parental care, Arsenault and colleagues (2018) removed the mother of *C. calcarata* offspring and investigated their behavioral and genomic responses. Their analysis of DNA methylation reported detectable difference of DNA

methylation between care and no-care offspring; however, these differences were not associated with differences of gene expression or associated with detected alternative splicing events. Therefore, if DNA methylation is important to behavior in this species, its importance not mediated through its traditional role of regulating of gene expression.

Western Honey Bee. There is some evidence that DNA methylation is associated with the behavior of honey bees. Herb and colleagues (2012) induced foraging bees to revert to nursing by removing all the nurses of a colony. They found ~150 differentially methylated regions during the standard nurse to forager transition and ~110 differentially methylated regions during the forager to nurse reversion. Fifty-seven of the regions overlapped and 45 of which could be recovered with whole genome bisulfite sequencing on replicated samples. However, Libbrecht and colleagues (2016) performed a reanalysis of their data using the methods of other studies that suggested DNA methylation was associated with behavior of honey bees and ants, and did find significant differences between life-history phases, and convincingly suggested many of the differences detected between worker castes were due to the use of single biological replicates (*Camponotus floridanus* - Bonasio et al., 2012; honey bee – Lyko et al., 2010). Removal of DNA methylation did impact alternative splicing of 87 genes for honey bees (Li-Byarlay et al., 2013). Recently, Harris and colleagues (2019) showed that differences of cytosine methylation do not associate with difference of gene expression in the Western honey bee. This further complicates the biological interpretation of difference of cytosine methylation between behavioral states. A survey of long non-coding RNAs (lncRNAs) between nurses and foragers under standard and age-controlled conditions was performed by Liu and colleagues (2019a). They found ~50 lncRNAs were consistently differentially expressed between their multiple pairwise comparisons. The genes targeted by these lncRNAs were developmental pathways. Circular RNAs (circRNAs) are non-coding RNAs produced from cell circularizing protein-coding exons with no firmly established function; however, the expression of several circRNAs correlate with the nurse to forager transition in honey bees. (Thölken et al., 2019).

Experimental work investigating parental care of insects – Transcriptomic.

Transcriptomic architecture of carers and offspring. Behavior has historically been suggested to have a special set of evolutionary rules due to its flexibility and its rapid rate of evolution, although behavior does not seriously challenge the classical laws of evolution (reviewed in Bailey et al., 2018). However, social behavior is unique in that its environment contains the heritable genes of others that can interact with the focal individual making the environment itself evolvable, which can lead to unusual and accelerated rates of evolution (Moore et al., 1997; Linksvayer, 2015; Schneider et al., 2017; Bailey et al., 2018). The phenomenon where genes expressed in one individual affect the phenotype of another individual is termed indirect genetic effects (IGEs; Moore et al., 1997). Little is known about how these evolutionary dynamics play out at the molecular level or through gene networks (Bailey and Hoskins, 2014; Warner et al., 2019b). This is despite the evolutionary and quantitative genetics of social behavior being well established at the theoretical level (e.g., McGlothlin et al., 2010).

A few studies have approached the genes producing IGEs by sampling interacting individuals at discrete time points and correlating gene expression to gain a better picture of the "interactome" and help us construct the gene networks of carer and offspring interactions. These studies have used the pharaoh ant *M. pharaonis* (Warner et al., 2019b), the honey bee *A. mellifera* (Vojvodic et al., 2015), and the European earwig (Wu et al., 2020). Warner and colleagues (2019b) looked at correlated gene expression

between nurses that specialized caring for worker-destined larvae at specific developmental stages and compared them with nurses that cared for larvae of all developmental stages. They found ~2,050 differentially expressed genes and a broad pattern of correlated transcriptional modules between stage specific larvae and their stage specific nurses (65% of genes expressed in nurses were contained in modules that were significantly correlated with stage-specific larvae). Vojvodic and colleagues (2015) looked at differences between nurses of queen-destined and worker-destined larvae, as well as, differences between queen and worker larvae. They detected 105 genes that were differentially expressed in the heads of different nurses. What is experimentally difficult but likely informative is to assay the numbers of genes differentially expressed in larvae that are taken care of by transplanted workers specializing on a different larval stage. This would address how plastic nurses are and the gene networks nurses to larvae and larvae to nurses. Wu and colleagues (2020) also assessed differentially expressed genes in nymphs with ~110 differentially expressed, but none overlapped with mothers. Under relaxed statistical criteria, they found one overlapping gene, *tyrosine hydroxylase*, that might influence care. For the discussion here, it is worth noting that it is also a known (i.e., conserved) gene. They followed up with an experimental manipulation detailed below.

This is a recent and emerging area of study that needs further work; however, there are some shared conclusions that can inform future studies. Two of these studies found tens to hundreds of genes that were differentially expressed between carers of different castes or different stages of larval development. Genes involved in interactions are not highly evolutionarily constrained for both the carer and the offspring (i.e., showed higher rate of nonsynonymous nucleotide substitutions within exons), which suggests a symmetrical selection on genes underpinning interactions (Vojvodic et al., 2015; Warner et al., 2019b). Additionally, genes that have secreted products are overrepresented within differentially expressed gene cohorts, with those genes also being some of the highest differentially expressed in carers (Vojvodic et al., 2015; Warner et al., 2019b*f*a). Genes with strong putative, regulatory connections for the nurse-larvae gene "interactome" had low connectivity to within tissue networks, an estimate possible because the abdominal tissues of nurses and larvae were also samples. They were also evolutionarily less constrained and there was a trend for these being evolutionarily young (phylostratigraphic analysis: young genes were only present within ants (LCA: ~140 MYA) – 31 vs ancient genes present within non-insect animals as well – 1,120; Warner et al., 2019b). Enriched biological processes of these regulatory genes were linked to metabolism. These are both generally expected trends as genes associated with derived phenotypes are expected to undergo more selection than genes not associated with derived phenotypes (Molodtsova et al., 2014). They are also expected to be under less pleiotropic constraint (i.e., at the periphery of gene networks), which are both general features of gene network evolution (Molodtsova et al., 2014).

Functional validation of targets from transcriptomic studies.

While genes are being discovered and quantified in an ever-increasing number of species, experimental validation of gene function is greatly lagging (Chang et al., 2016). Direct experimental validation is the only definitive way to establish a causal link between a gene and a phenotype of interest. Tests are complicated because many molecules associated with behavior (e.g., neuropeptides and neurotransmitters) are highly potent and tend to control physiology as well. This limits the use of traditional null alleles due to pleiotropic effects (e.g., Sokolowski, 2001).

For parental care, there are a few studies that have functionally validated roles of genes for particular aspects of parental care. *odorant receptor coreceptor* (*orco*) is an obligate co-receptor for all olfaction of insects and without this protein the nurses of two ant species display reduced sociality (aggregation with others; *O. biroi* - Trible et al., 2017; *Harpegnathos saltator* - Yan et al., 2017). For many species, there is parental investment in the innate immunity of offspring. An RNA-seq study of the burying beetle *N. vespilloides* gut found a lysozyme gene (*lysozyme 6*, *lys6*) upregulated during parental care and followed this up by correlating the mRNA levels of *lys6* with the antimicrobial action of the exudates among individual in a separate sample (Palmer et al., 2016). There have also been several studies that have tried to pharmacologically manipulate caring behavior. Juvenile hormone (JH) application does not manipulate task specialization – nursing or foraging – among the adult buff-tailed bumblebees *B. terrestris* (Shpigler et al., 2016). In the Western honey bee, a juvenile hormone analogue accelerated the transition of nurses to foragers (Chang et al., 2015). To date, the best functional validations of a gene action underpinning insect parental care were conducted by Kohlmeier and colleagues (2018) and Wu and colleagues (2020). Kohlmeier and colleagues (2018) used RNA interference (RNAi) to knockdown expression of one of the vitellogenin genes, *vitellogenin-like A* (*vg-like A*), in the ant *T. longispinosus*. They found that RNAi of *vg-like A* through spiked food reduced brood care; however, under extended exposure it also prolonged the care of nest mates, an effect they suggested was mediated through decreased responsiveness to larvae sensory cues (Kohlmeier et al., 2018). They were also able to show that this is likely a direct genetic effect within nurses as RNAi of *vg-like A* within larvae did not alter the care larvae received. Wu and colleagues (2020) using RNAi to knockdown expression of *tyrosine hydroxylase* (*Th*) in mothers of the European earwig. This knockdown lowered maternal provisioning of offspring. In a *tour de force*, Hamilton and colleagues (2019) used RNAi with honey bee workers to knockdown expression of the transcription factors *broad* and *fushi-tarazu transcription factor 1* (*ftz-f1*). Of relevance to this review, these knockdowns increased brood care; however, there was some between colony variability for the RNAi effects. They also used a pharmacological treatment to show a juvenile hormone analog reversed the effect of *broad* RNAi (i.e., decreased brood care), but not *ftz-f1* RNAi (Hamilton et al., 2019). They also confirmed these knockdowns had effects on the expression of direct target genes. This allowed them to construct the regulatory network of these two transcription factors with *broad* likely upstream of *ftz-f1* (Hamilton et al., 2019). Another study using honey bees artificially selected for increased royal jelly production by Wu and colleagues (2019) used *in vitro* measurements of protein-odorant bindings to verify that several candidate chemosensory genes were targets of brood pheromones that stimulate the antennae of selected lines more than controls and increase provisioning of royal jelly to queen-destined larvae.

The successful studies here are encouraging for our ability to more regularly establish a causal relationship between genes and behavioral phenotypes. Our ability to perform genetic manipulations will only increase as the resources needed for them increase for non-model systems (e.g., as well assembled and annotated genomes become easier to produce). New genetic manipulation technologies will also provide extra tools to establish causal relationships, such as, CRIPR-Cas9 (Bono et al., 2015).

Future Directions. Conceptual Shifts & Practical Considerations

In this final section I will highlight some considerations for the field moving forward. I will first outline some conceptual considerations and suggest how some new ideas can better inform our experiments and conclusions. Of particular note is the suggestion that functional genomics increase its exploration of the

441 genes that underpin indirect genetic effects. Second, I will discuss some practical and experimental
442 considerations. Within the practical concerns, I would like to highlight the call to standardize the
443 reporting of biological measures (e.g., GO term enrichment).
444
445 **Conceptual Shifts.**
446 **Predictive Frameworks.** Functional genomics is by its relative youth and nature usually
447 exploratory. Through its comprehensive broadness it exceeds at many "look and see" exercises. This
448 contrasts with behavioral ecology, which is a very mature field that usually demands well-defined
449 hypotheses and predictions from its experimenters. Both fields can learn from the other. Many functional
450 genomics studies of well-characterized behaviors can have specific hypotheses within a larger framework
451 of an exploratory analysis (e.g., Cunningham et al., 2019). Predictions within functional genomic studies
452 of behavior lead on from three established ideas – Wright's theory of near universal pleiotropy (Wright,
453 1968; mechanistically formalized within the "Omnigenetic model" by Pritchard and colleagues (Boyle et
454 al., 2017; Liu et al., 2019b)), the Phenotypic Gambit (Grafen, 1984; empirically formalized within
455 Cheverud's Conjecture (Cheverud, 1988)), and higher order behaviors being synthesized from multiple
456 lower order behaviors (e.g., Székely et al., 2013) that are extensions of existing behaviors (Tallamy,
457 1984) and of existing gene networks (Alberts and Kruglyak, 2015; Halfon, 2017; Fig. 1). It also relies on
458 the aggregate knowledge of genes functions generated over the past fifty years of molecular biology. It
459 was formalized as the Precursor Hypothesis within a recent review by Moore and Benowitz (2019). As a
460 single statement, the precursor hypothesis for parental care would posit the gene networks that underpin
461 parental care are predictable given strong pleiotropy shown by many genes influencing the specific
462 behaviors that integrate to produce full parental care, such as, a modification of feeding habits (Moore and
463 Benowitz, 2019; Fig 1). For our purposes here, this line of thinking allows for the prediction of the gene
464 networks that should be co-opted for each behavioral component that underlies parental care in a novel
465 species, at both a single gene level (Cunningham et al., 2016; Benowitz et al., 2019) or for classes of
466 genes (Cunningham et al., 2017; 2019). It can efficiently take the experimenter from phenotypic space,
467 which is usually the first thing characterized, to a prediction of "transcriptional" space, which is more
468 technically difficult to access. This view of the evolutionary-genetic trajectories of derived phenotypes
469 promotes that natural selection will usually act on conserved gene networks (Moore and Benowitz, 2019).
470 The Precursor Hypothesis differs from the classic and successful evo-devo view of candidate genes (i.e,
471 genetic-toolkits; Toth and Robinson, 2007) in that it takes a behavior led view for predictions of genes
472 networks, analogous to the difference between top-down and bottom-up genetics within behavior (Boake
473 et al., 2002; Rubenstein and Hofmann, 2015). Also, to become more predictive, I argue that we also need
474 to become more explicit about defining our phenotypes. This also means being more explicit about the
475 phenotypes we are comparing to and why those phenotypes are directly comparable.
476 **Conservation of Mechanisms.** Leading on from a discussion of pleiotropy within modern
477 genetics is a related idea about how to quantify the conservation of mechanisms. As information on the
478 genes used for social behavior first became available, there was a strong sense that repeated evolution of
479 behavioral phenotypes could be achieved by use of homologous genes, mainly due to the success of
480 candidate gene approaches from fruit flies (*Drosophila* spp.) to honey bees, such as *foraging* (Ben-Shahar
481 et al., 2002; reviewed in Anreiter and Sokolowski, 2019). However, as the number of species that we have
482 used for these questions continues to increase, this strict construction of what is conserved has been
483 challenged with thinking now generally broadened to gene networks and pathways (Berens et al., 2015;
484 Kapheim et al., 2015; Ritschoff and Robinson, 2016; Moore and Benowitz, 2019). An empirical test of

485 this idea is difficult in a behavior as complex as parental care, but tests of single behaviors have been
486 performed. Shorter and colleagues (2015) leveraged *Drosophila melanogaster* to perform a highly
487 powered and focused genome-wide association study for male-male aggression between a huge panel of
488 inbreed lines and followed this up by intercrossing the lines that demonstrated the extreme high and low
489 values of aggression. The population used has greatly reduced genetic variation compared with wild
490 populations and the authors were not able to recapitulate the effect of individual genes on aggression, but
491 did readily and strongly recapitulate selection on gene networks to produce the same phenotype from,
492 importantly, a map of gene networks from independent data (Shorter et al., 2015). A similar result was
493 also found by Berens and colleagues (2015) for the genes underlying castes' behaviors within eusocial
494 Hymenoptera when they contrasted a primitively eusocial wasp *Polites metricus* to honey bees and a fire
495 ant. Individual genes important for caste behavior of honey bees were poor predictors of behavior within
496 *P. metricus*. However, they found a much greater overlap when testing if the same pathways (Kyoto
497 Encyclopedia of Genes and Genomes (KEGG) pathways or Gene Ontology (GO) terms), rather than
498 individual genes, were shared between the castes of the species (Berens et al., 2015). They also
499 empirically demonstrated an under-representation of taxon restricted genes being involved in caste
500 differentiation in contrast to previous reports. Functional genomics might be a poor place to look for
501 conservation of some derived phenotype as different levels of biological organization might have
502 different predispositions to be used for different behaviors (Fischer et al., 2019). Although it is unlikely
503 that this is the case as variation within higher levels of biological organization (e.g., rewiring a neural
504 circuit) remains fundamentally linked to gene expression. While we will not know a general conclusion
505 until we have more data from more species, it is clear that not every gene will have similar expression or
506 associations analogous phenotypes.
507 The empirical work discussed above links well with the suggestion that the broader field of social
508 behavior moves towards network-based thinking, both for gene networks and behaviors themselves
509 (Linksvayer et al., 2012; Akçay et al., 2015; Linksvayer, 2015; Silk et al., 2018; Fig. 2). This suggestion
510 is driven by the want to explicitly link the well-established theory of social evolution (e.g., inclusive
511 fitness, indirect genetic effects) with genes and their networks, the actual substrate of evolution. The
512 benefit is reciprocal with social evolution theory benefiting from having to consider the proximate causes
513 of evolution and their dynamics, while functional genomics will benefit by moving away from verbal
514 models to explicit expectations of what the gene networks underpinning social behavior should look like.
515 From the work reviewed here, we can find some consensus for gene networks that naturally align with
516 behavior (e.g., genes involved in neural function); however, we can also find genes involved in social
517 interactions that might not align directly with an individual's behavior (e.g., secreted proteins; Warner et
518 al., 2019a). This is a relatively unexplored area with few studies attempting to estimate the genetic loci
519 underpinning IGEs (Wang et al., 2008; Bailey & Hoskins, 2014). Understanding that social behavior is a
520 product two individuals and their interaction, we might predict that there are two different types of gene
521 networks needed to understand social behavior; the direct genetic effects within an individual and the
522 indirect genetic effects of an individual onto the interacting partner (Fig. 3C). Questions that remain open
523 are many, but fundamentally are about the size and topological features of these gene networks, what
524 level of regulation and overlap do they share, and are IGE gene networks expected to be as relatively
525 conserved as the gene networks that underpin direct genetic effect?
526 **Comparative Frameworks.** A very natural question after elucidating the genetic basis of some
527 phenotype is – how general is the pattern I am observing? This is a place where the evolutionary biology
528 within behavioral ecology can help inform functional genomics as more datasets become available.

However, with more data comes the need to be more explicit and rigorous about the comparative framework the work is placed into. This consideration has both a conceptual and an experimental aim. At a conceptual level, we need to be clear about what RNA-seq data from any one species can and cannot tell us about the general trends for evolution of a phenotype. For example, there is little reason to believe extant subsocial species occupy the complete and necessary phenotypic/genotypic space for lineages that have become eusocial because not all species are steadily marching towards eusociality and each lineage is constrained by lineage-specific factors (*sensu* Linksvayer and Johnson, 2019; *contrario sensu* Rehan and Toth, 2015; Kronauer and Libbrecht, 2018). However, that does not invalidate all data gathered from extant subsocial species to inform the evolution and mechanisms of other parental care species. It just means that we have to be more careful about how definitive our conclusions are. Questions such as "what is the complete set of genes used to produce parental care on a species' way to eusociality" might be overly simple compared with questions about trends, such as, "to what extent does a reproductive network need to be rewired to produce a sterile, alloparental caring workers?" A direct counter-argument to the limit of predictability between species is the success of some candidate genes, such as, *foraging* and *neuropeptide F*. A candidate gene is a specific gene-phenotype association hypothesized to have a continued association in all/most new species, even if the direction of influence is not conserved (reviewed in Fitzpatrick et al., 2005). This framework has strong historical roots from the study of the evolution and development (evo-devo) of morphological traits (Fitzpatrick et al., 2005; Toth and Robinson, 2007). In fact, the success of candidate genes as a whole has led to the enduring success of the "genetic toolkits," that evolution uses certain genes preferentially to generate novel behavioral phenotypes, particularly for some behaviors, such as, a link between feeding and parental care (Rittschof and Robinson, 2014; 2016; Fischer and O'Connell, 2017). A weakness of this specific framework is that there is no *a priori* way to guarantee which individual gene actions will be conserved, even if strong arguments can be made for specific genes, such as, *neuropeptide F*'s role in modifying feeding behavior (Cunningham et al., 2016). It is therefore likely advantageous to think of the conservation of specific pathways or gene networks (see previous section).

Practical & Experimental Considerations.

Standardized analyses and reporting. I would like to argue the field needs to improve and standardize its analyses and reporting. There are technical checklist for reporting qRT-PCR results (MIQE; Bustin et al., 2009) and RNA-seq experimental guidelines (Conesa et al., 2016), but no such guidelines exist for biological measures. First, I think it is important for authors to estimate how whole transcriptomes cluster or support their hypotheses. This requires some test of overall association or clustering of sample based on treatment, such as, Principal Component Analysis (PCA), hierarchical clustering, or Multidimensional Scaling (MDS). Testing for a statistically significant association between treatment and association/clustering requires care, but is possible (e.g., PCA - Peres-Neto et al., 2003; hierarchical clustering - Kimes et al., 2017). Differential gene expression is the standard analysis of gene expression and something that all RNA-seq studies preform (Conesa et al., 2016). This should also be followed by an analysis of gene co-expression. This can help find gene network influenced by treatment not detectable at the single gene level (Saelens et al., 2018). These networks can also be used to infer regulatory structure within the gene networks, with genes at the center of these networks assumed to have regulatory roles do to their high connectivity. In this field, weighted gene co-expression network analysis (WGCNA) protocol seems to be a popular tool (Langfelder and Horvath, 2008) and one that still performs

well (Saelens et al., 2018). Gene ontology (GO) term enrichment for both differentially expressed genes and gene networks associated with treatments should be analyzed. GO term enrichment allows authors to assign known biological functions to genes and then test for overrepresentation of terms with any cohort (Ashburner et al., 2000). These terms, rather than solely the number and idenity of differentially expressed genes, improve our ability to broadly interpret and compare molecular mechanisms across species. GO term enrichment has noticeable inconsistency with reporting, even if some studies state that these analyses done.

Permutation/randomization tests are an underutilized tool for this field to produce null expectations. Permutation tests can reveal how likely estimated values are given the data; however, they cannot tell one anything outside of that particular data set. That means they cannot be used as an *a priori* expectation for another data set. An example is the number of methylation differences between care and non-care states (Libbrecht and Kronauer, 2016; Cunningham et al., 2019). Even though both studies found consistent difference of DNA methylation between care and non-care states, they each showed using permutation/randomization tests this was not more than expected by chance for the samples. This completely changed the biological interpretation of the data; from DNA methylation being associated with behavior to there being no association of DNA methylation with behavior.

Under sampled research areas. There are several obvious areas where there is a lack of understanding for the mechanistic basis of parental care. First, I would highlight the need to sample both carers and offspring. Social behavior requires more than one individual and that the other individual also has evolvable genes that can lead to unique and sometimes rapid evolution of social behavior (Bailey et al., 2018). This dynamic can perpetuate co-evolution of loci between interacting individuals. However, there are several basic questions that have no answers: is the transcriptional architecture of an IGE more or less complex than the behavioral response to abiotic or physiological stimuli, how many loci mediate IGEs on average, is a particular class of gene more likely to be used to mediate an IGE social interaction? Second, I would highlight the need to better validate the genes that are associated with parental care. *vitellogenin* was an obvious candidate gene family to directly experimentally validate for its role in parental care as it comes up in almost all analyses of differentially expressed genes between caring and non-caring individuals.

Evolution and mechanisms of gene regulation. *cis*-regulatory elements are DNA sequence motifs that are binding sites for transcription factors or other proteins that regulate transcription, both within and outside of the core promoters of genes (Halfon et a., 2008). *cis*-regulatory elements coupled with transcription factors and epigenetic mechanisms (chromatin state, regulatory RNAs (microRNAs and lncRNAs), and DNA methylation) regulate gene expression (Jaenisch and Bird, 2003). Although, DNA methylation looks to play a reduced role regulating the gene expression of insects (see discussion above). Understanding how these factors interact *in vivo* is necessary to fully characterize the regulation of gene expression through transcriptional regulatory networks (TRNs), or gene regulatory networks (Halfon, 2017; Lu et al., 2017). TRNs are a major target to evolve new phenotypes as they define the sets of genes expressed within a cell and therefore the identity and function of a cell (Simola et al., 2013; Halfon, 2017). Some work has detailed transcription factor networks of parental are and other behaviors in the honey bee (Chandrasekaran et al., 2011; Ament et al., 2012; Khamis et al., 2015; Hamilton et al., 2019) and carpenter bee (Shell and Rehan, 2019). That research has produced some interesting conclusions, such as, some transcription factors are shared across behaviors, relatively few transcription factors can regulate behavioral transitions, and that there seems to be a strong hierarchical nature to the TRNs in honey bee behavior (Chandrasekaran et al., 2011; Khamis et al., 2015; Hamilton et al., 2019). Shell and

Rehan (2019) also found conserved transcription factors regulating guarding/grooming behaviors. We need to understand if this is a general feature of all behavior, reversible behaviors, or just of behavior that evolved to one-way, permanent transitions under standard conditions, such as, the transition from nursing to foraging of honey bees. What are the properties of a TRN of a reversible behavior where transitions are not predictable based on or evolved to be linked with age? Bringing in dedicated systems biology approaches to parental care research will likely produce a much stronger and harmonized understanding of how this behavior is orchestrated (Fernald, 2011).

At a practical level, construction of TRNs can be achieved with either computational or experimental tools. Combining these approaches represents a particularly powerful approach to construct TRNs (Song et al., 2016). The transcription factor component of TRNs can be constructed with gene expression data and transcription factor binding site motif information computationally (Ament et al., 2012). A method used by several in the field is ASTRIX (Chandrasekaran et al., 2011; Shell and Rehan, 2019), but other methods do exist (Song et al., 2016). These methods search for over-represented transcription factor binding sites around differentially expressed genes to reconstruct regulatory networks. Active binding of transcriptional factors can be investigated at an experimental level with Chromatin Immunoprecipitation followed by Sequencing (ChIP-seq; Johnson et al., 2007), which uses antibodies to capture the bound protein-DNA complexes. There are also ways to discover novel binding sites of transcription factors (e.g., MEME-ChIP; Machanick and Bailey, 2011). Chromatin state is regulated by histone tail post translational modifications (histone tail PTMs), nucleosome positioning, and chromatin accessibility through chromosomal packing openness. Histone tail PTMs can be investigated with ChIP-seq, or for any protein-DNA interaction that an antibody is available. The openness of chromatin, a necessary condition for gene expression, can be investigated with the small samples of many non-model insects using an Assay for Transposase Accessible Chromatin using Sequencing (ATAC-seq; Buenrostro et al., 2013). The technique can help to identify 1) *cis*-regulatory regions to better understand promoter and enhancer architecture, 2) bound TFs within *cis*-regulatory regions using "DNA footprints," and 3) nucleosome positioning, and 4) open regions of chromatin that indicate proximal genes are active (Buenrostro et al., 2013; Lu et al., 2017). ATAC-seq is well suited for use with emerging-model species because it needs little *a priori* knowledge, small tissue amounts are permissible, and is a relatively simple protocol to better understand genomic regions of regulatory importance (Lu et al., 2017). Distal regulatory elements, such as, enhancers and silencers, are also part of TRNs and difficult to investigate due to their distance from the genes they regulate. Hi-C is a next generation sequencing based extension of chromosome conformation capture (Dekker et al., 2002) that gives information about the distal elements of the genome that interact, such as enhancer-promoter interactions, through direct contact or co-contact with protein complexes providing information on another layer of TRNs (Belton et al., 2012). Regulatory RNAs (microRNAs and long non-coding RNAs) can be investigated with standard RNA-seq protocols that differ in pre-processing of total RNA (rRNA depletion rather than polyA selection) or post-library construction size-selection (due to their small sizes). Likely important given their impact on mRNA levels (Flynt and Lai, 2008; Kornienko et al., 2013), the contribution of these regulatory RNAs to parental care of insects is withstanding.

TRNs are an accessible new frontier to understand the functional genomics of insect parental care. However, the view taken here is very "up-stream" of gene expression. The regulation "down-stream" of gene expression; regulation of proteins in it many forms, is also likely of equal importance as gene expression regulation *per se* (e.g., Vogel and Marcotte, 2012). Some common types of regulation are post-translational modifications, co-factor presence, degradation, and allosteric regulation. However, I

have not discussed it here on mainly practical grounds, such as, kits are not widely available to robustly measure proteins, larger amounts of starting material are generally needed, very well constructed and annotated genomes are needed, and highly specialized equipment with highly specialized technicians to run them are needed to make progress in this area.

Taxonomically Restricted Genes (TRGs). While it is likely true that conserved gene network underpin many new behaviors, as evidenced by the success of candidate gene pathways, it does not preclude novel genes arising to underpin the evolution of derived phenotypes. Genes that have been categorized, hereto, as TRGs have been documented to play causal role in some derived social behaviors (reviewed in Taborsky and Taborsky, 2015), although this is sometimes questioned by re-analyses (Berens et al., 2015) and is completely dependent on the quality and number of annotated genomes of closely and moderately related species. How commonly TRGs underpin the evolution of behavior within the central nervous system might be an outstanding question, but their role in the generation of novel tissues has more support. Jasper and colleagues (2015) used a very strong predictive framework for the pattern of TRGs, their connectedness within gene networks, and the amount of coding vs. regulatory changes associated with genes differentially expressed within novel tissues of honey bees (e.g., hypopharyngeal gland and Nasonov's gland). They found that TRGs were highly enriched among differentially expressed genes within the novel tissues of honey bees; but, were not enriched among differentially expressed genes within conserved tissues, such as, brain or muscle tissue. Among these different tissues they also showed that TRGs are not centrally within gene networks (i.e., had statistically significantly lower degrees of connectedness within constructed gene networks) and had very high rate of coding sequence evolution (Jasper et al., 2015). This line of thinking extended to behavior would suggest that as the novelty of a behavior increases, so might the influence of TRGs. An example, regurgitation of food directly to offspring is highly unusual, but provisioning offspring with a food source is more common. Would we therefore expect a higher chance of a TRG to underpin regurgitation or a gene with more coding sequence change than those that underpin food provisioning? While this prediction does directly follow on from the Jasper and colleagues (2015), I would like to comment on what I view as a potential pitfall. This type of prediction brings up the issues of what is "novel." Is the evolutionary change from no provisioning to provisioning offspring any less novel than provisioning food to direct regurgitation of food to offspring? As posed this line of thinking also relates a highly step-wise vision of evolution. Nevertheless, it remains with investigators to carefully articulate their expected evolutionary scenario and therefore the basis of their predictions of type of genes, and their novelty, underpinning a behavior.

Factorial designs to disentangle multiple variables. With the decreasing costs of sequencing, there is a possibility to design experiments that test multiple variables and combinations of variables using factorial design with adequate sample sizes. While there will always be a place for experiments that standardize as much as possible experimentally, it opens the door for experiments that can purposely assay multiple variables, their interactions, and begin to understand the magnitude of influence that each variable has on transcriptional architecture. This experimental design is already beginning to enter our field (e.g., Cunningham et al., 2019; Kohlmeier et al., 2019). Both these studies were able to show that the behavior an individual was expressing was more influential on transcriptional architecture than other environmental or physiological variables within the brains of those individuals. With increasing complexity of experimental design, one needs to remember that there will more uncertainty about the influence of any one variable and the genes associated with that variable; however, a more complete

understanding will be achieved. This a natural and expected consequence of multi-variable analyses (Morissey and Ruxton, 2018).

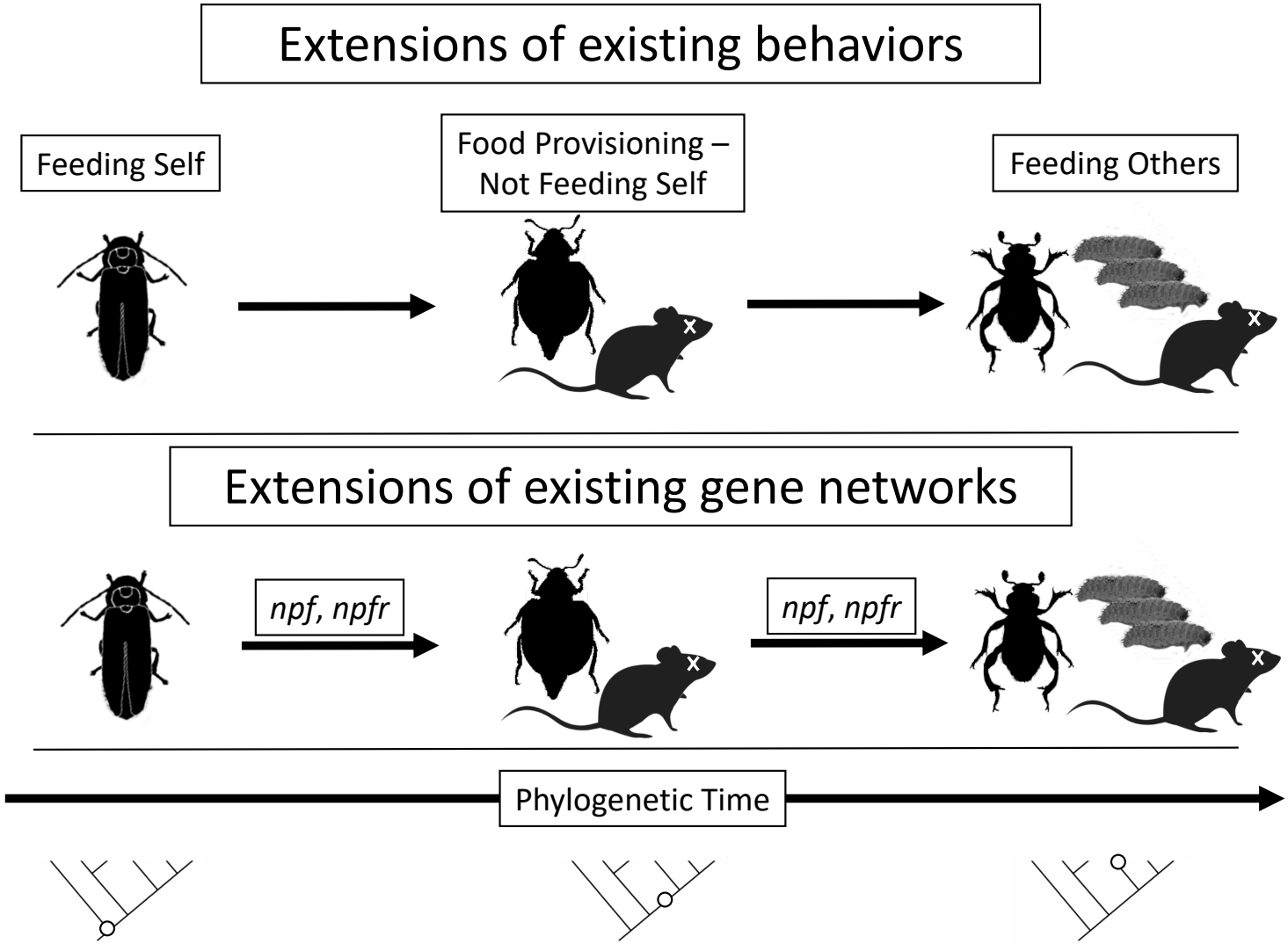
Data availability. There is no data to publicly archive from this work.

Conflicts of Interest. I declare I have no conflict of interests for this work or its conclusions.

Acknowledgements. I would like to thank all of the behavioral ecologist/genetists I have meet at conferences and seminars over the past five years that have taken time to discuss, correct, extend, and refine my arguments over the best way to view parental care, its mechanistic underpinnings, and the best ways forward for the field with me. I would like to thank A.J. Moore, C.J. Coates, and two anonymous reviewers for thoughtful comments on drafts of this manuscripts.

Funding: This work was supported by funds from Swansea University's College of Science.

720 **Figures**
721



722
723
724 Figure 2. Schematic of the Precursor Hypothesis through one evolutionary lineage, *N. vespilloides*, for one trait, feeding offspring. The evolution of
725 parental care is expected to be produced from an extension of existing behaviors as a given species' lineage evolvs parental care. The hypothesized
726 transitions from feeding oneself to not eating food provisioned to offspring to direct regurgitation of processed food to offspring is shown (Top Panel).

The gene networks that underpin extended behaviors are expected to be extensions of existing gene networks. Here, this is hypothesized to be novel regulation of the same gene pathway, neuropeptide F and its receptor (*npf* and *npfr*) and the gene network it regulates (not shown) (Middle Panel). A schematic of phylogenetic time and trait appearance (hollow white circles) on a phylogeny for the evolution of feeding offspring is shown (Bottom Panel). Silhouettes are from PhyloPic.com or the author's research group.

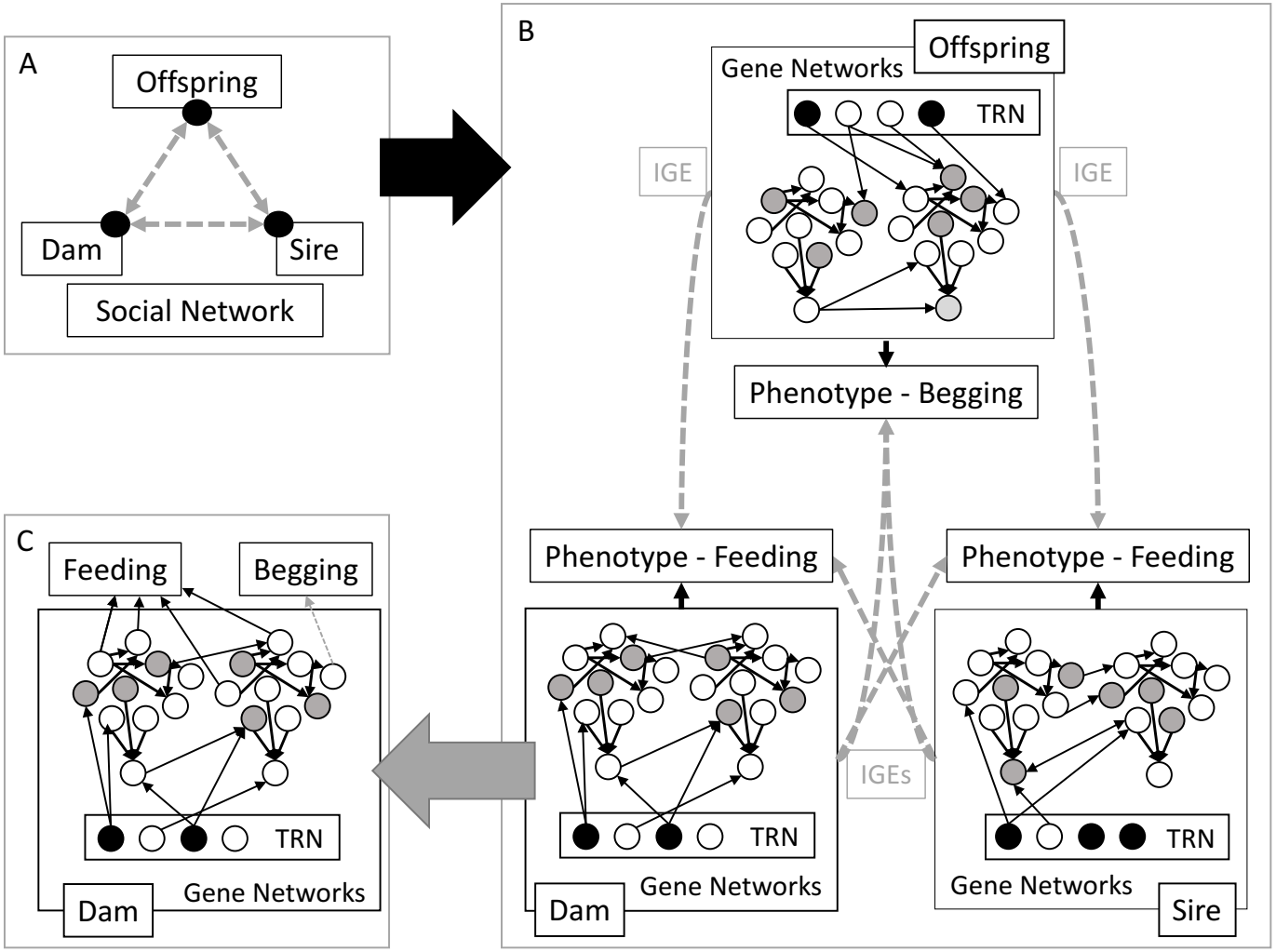


Figure 3. Integration of social networks with gene networks and indirect genetic effects (Moore et al., 1997; Linksvayer et al., 2012; Akçay et al., 2015; Linksvayer, 2015; Silk et al., 2018). (A) The social network of a simple family; dam, sire, and one offspring; is shown. (B) Gene networks produce the additive genetic effects on an individual's behavior (small, solid black arrows), which in this specific case is shown as parental feeding and begging of offspring. One level of transcriptional regulatory network (TRN) is shown regulating the expression of gene networks to display the hierarchical nature of these networks. Indirect genetic effects (IGEs; long, dashed grey arrows) describe the additive genetic variance of one individual's genes onto the phenotype of another individual. (C) The familiar direct genetic effects of a dam onto her own feeding behavior that generate the additive genetic

variance component (multiple black arrows onto the Feeding box) and one indirect genetic effect onto the begging behavior of her offspring to visualize an IGE at the level of a single gene (single grey dashed arrow onto the begging box). One level of transcriptional regulatory network (TRN) is again portrayed to represent the hierarchical nature of transcriptional control.

References

Akçay, E., Linksvayer, T.A., Van Cleve, J., 2015. Bridging social evolution theory and emerging empirical approaches to social behavior. Curr. Opin. Behav. Sci. 6, 59-64.

Albert, F.W., Kruglyak, L., 2015. The role of regulatory variation in complex traits and disease. Nature Rev. Genet. 16, 197-212.

Ament, S.A., Blatti, C.A., Alaux, C., Wheeler, M.M., Toth, A.L., Le Conte, Y., Hunt, G.J., Guzmán-Novoa, E., DeGrandi-Hoffman, G., Uribe-Rubio, J.L., Amdam, G.V., Page Jr., R.E., Rodriguez-Zas, S.L., Robinson, G.E., Sinha, S., 2012. New meta-analysis tools reveal common transcriptional regulatory basis for multiple determinants of behavior. Proc, Nat. Acad. Sci. USA 109, 10162-10163.

Anreiter, I., Sokolowski, M.B., 2019. The *foraging* gene and its behavioural effects: pleiotropy and plasticity. Ann. Rev. Genet. 53, 373-392.

Arsenault, S.V., Hunt, B.G., Rehan, S.M., 2018. The effect of maternal care on gene expression and DNA methylation in a subsocial bee. Nature Comm. 9, 3468.

Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., Harris, M.A., Hill, D.P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J.C., Richardson, J.E., Ringwald, M., Rubin, G.M., Sherlock, G., 2000. Gene ontology: tool for the unification of biology. Nature Genet. 25, 25-29.

Bailey, N.W., Hoskins, J.L., 2014. Detecting cryptic indirect genetics effects. Evolution 68, 1871-1882.

Bailey, N.W., Marie-Orleach, L., Moore, A.J., 2018. Indirect genetic effects in behavioral ecology: does behavior play a special role in evolution? Behav. Ecol. 29, 1-11.

Belton, J.M., McCord, R.P., Gibcus, J.H., Naumova, N., Zhan, Y., Dekker, J., 2012. Hi-C: a comprehensive technique to capture the conformation of genomes. Methods 58, 268-276.

Ben-Shahar, Y., Robichon, A., Sokolowski, M.B., Robinson, G.E., 2002. Influence of gene action across different time scales on behavior. Science 296, 741-744.

Bateson, P., Laland, K.N., 2013. Tinbergen's four questions: an appreciation and an update. Trends Ecol. Evol. 28, 713-718.

Benowitz, K.M., McKinney, E.C., Cunningham, C.B., Moore, A.J., 2017. Relating quantitative variation within a behavior to variation in transcription. Evolution 71, 1999–2009.

Benowitz, K.M., McKinney, E.C., Cunningham, C.B., Moore, A.J., 2019. Predictable gene expression related to behavioral variation in parenting. Behav. Ecol. 30, 402-407.

Berens, A.J., Hunt, J.H., Toth, A.L., 2014. Comparative transcriptomics of convergent evolution: different genes but conserved pathways underlie caste phenotypes across lineages of eusocial insects. Mol. Biol. Evol. 32, 690–703.

Bludau, A., Royer, M., Meister, G., Neumann, I.D., Menon, R., 2019. Epigenetic regulation of the social brain. Trends Neurosci. 42, 471-484.

Boake, C.RB., Arnold, S.J., Breden, F., Meffert, L.M., Ritchie, M.G., Taylor, B.J., Wolf, J.B., Moore, A.J., 2002. Genetic tools for studying adaptation and the evolution of behavior. Am. Nat. 160, S143-S159.

Bonasio,R., Li, Q., Lian, J., Mutti, N.S., Jin, L., Zhao, H., Zhang, P., Wen, P., Xiang, H., Ding, Y., Jin, Z., Shen, S.S., Wang, Z., Wang, W., Wang, J., Berger, S.L., Liebig, J., Zhang, G., Reinberg, D., 2012. Genome-wide and caste-specific DNA methylomes of the ants *Camponotus floridanus* and *Harpegnathos saltator*. Curr. Biol. 22, 1755-1764.

Bono, J.M., Olesnicky, E.C., Matzkin, L.M., 2015. Connecting genotypes, phenotypes and fitness: harnessing the power of CRISPR/Cas9 genome editing. Mol. Ecol. 24, 3810-3822.

Boyle, E.A., Li, Y.I., Pritchard, J.K., 2017. An expanded view of complex traits: From polygenic to omnigenic. Cell 169, 1177-1186.

Buenrostro JD, Giresi, P.G., Zaba, L.C., Chang, H.Y., Greenleaf, W.J., 2013. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nature Meth. 10, 1213–1218.

Bustin, S.A., Benes, V., Garson, J.A., Hellemans, J., Huggett, J., Kubista, M., Mueller, R., Nolan, T., Pfaffl, M.W., Shipley, G.L., Vandesompele, J., Wittwer, C.T., 2009. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. Clin. Chem. 55, 611-622.

Cardoso, S.D., Teles, M.C., Oliverio, R.F., 2015. Neurogenomic mechanisms of social plasticity. J. Exp. Biol. 218, 140-149.

Chandrasekaran, S., Ament, S.A., Eddy, J.A., Rodriguez-Zas, S.L., Schatz, B.R., Price, N.D., Robinson, G.E., 2011. Behavior-specific changes in transcriptional modules lead to distinct and predictable neurogenomic states. Proc. Nat. Acad. Sci. USA 108, 18020-18025.

Chang, L.-H., Barron, A.B., Cheng, K., 2015. Effects of the juvenile hormone analogue methoprene on rate of behavioural development, foraging performance and navigation in honey bees (*Apis mellifera*). J. Exp. Biol. 218, 1715-1742.

Chang, Y.C., Hu, Z., Rachlin, J., Anton, B.P., Kasif, S., Roberts, R.J., Steffen, M., 2016. COMBREX-DB: An experiment centered database of protein function: knowledge, predictions and knowledge gaps. Nuc. Acids Res. 44, D330-D335.

831 Cheverud, J.M., 1988. A comparison of genetic and phenotypic correlations. Evolution 42, 958-968.
832
833 Conesa, A., Madrigal, P., Tarazona, S., Gomez-Cabrero, D., Cervera, A., McPherson, A., Szcześniak,
834 M.W., Gaffney, D.J., Elo, L.L., Zhang, X., Mortazavi, A., 2016. A survey of best practices for RNA-seq
835 data analysis. Genome Biol. 17, 13.
836
837 Cunningham, C.B., VanDenHeuvel, K., Khana, D.B., McKinney, E.C., Moore, A.J., 2016. The role of
838 *neuropeptide F* in a transition to parental care. Biol. Lett. 12, 20160158.
839
840 Cunningham, C.B., Badgett, M.J., Meagher, R.J., Orlando, R., Moore, A.J., 2017. Ethological principles
841 predict the neuropeptides co-opted to influence parenting. Nature Comm. 8, 14225.
842
843 Cunningham, C.B., Ji, L., McKinney, E.C., Benowitz, K.M., Schmitz, R.J., Moore, A.J., 2019. Changes
844 of gene expression but not cytosine methylation are associated with male parental care reflecting
845 behavioural state, social context and individual flexibility. J. Exp. Biol. 222, jeb188649.
846
847 Dekker, J., Rippe, K., Dekker, M. & Kleckner, N., 2002. Capturing chromosome conformation. Science
848 295, 1306–1311.
849
850 Dulac, C., O'Connell, L.A., Wu, Z., 2014. Neural control of maternal and paternal behaviors. Science
851 345, 756-770.
852
853 Evans, T.G., 2015. Considerations for the use of transcriptomics in identifying the 'genes that matter' for
854 environmental adaptation. J. Exp. Biol. 218, 1925-1935.
855
856 Fernald, R.D., 2011. Systems biology meets behaviour. Proc. Nat. Acad. Sci. USA 108, 17861-17862.
857
858 Flynt, A.S., Lai, E.C., 2008. Biological principles of microRNA-mediated regulation: shared themes amid
859 diversity. Nature Rev. Genet. 9, 831-842.
860
861 Fischer, E.K., O'Connell, L.A., 2017. Modification of feeding circuits in the evolution of social behavior
862 J. Expt. Biol. 220, 92-102.
863
864 Fischer, E.K., Nowicki, J.P., O'Connell, L.A., 2019. Evolution of affiliation: patterns of convergence
865 from genomes to behaviour. Phil. Trans. Roy. Soc. B 374: 20180242.
866
867 Fitzpatrick, M.J., Ben-Shahar, Y., Smid, H.M., Vet, L.E.M., Robinson, G.E., Sokolowski, M.B., 2005.
868 Candidate genes in behavioural ecology. Trends Ecol. Evol. 20, 96-104.
869
870 Franks, A., Airoldi, E., Slavov, N., 2016. Post-transcriptional regulation across human tissues. PLoS
871 Comp. Biol. 13, e1005535.
872
873 Grafen, A., 1984. Natural selection, kin selection and group selection, in: Krebs, J.R., Davies, N.B.
874 (Eds.), Behavioural Ecology: An Evolutionary Approach, 2nd ed. Blackwell Scientific, Oxford, pp. 62–84.

Halfon, M.S., 2017. Perspectives on gene regulatory network evolution. Trends Genet. 33, 436-447.

Halfon, M.S., Gallo, S.M., Bergman, C.M., 2008. REDfly 2.0: an integrated database of *cis*-regulatory modules and transcription factor binding sites in *Drosophila*. Nucl. Acids Res. 36, D594–D598.

Hamilton, A.R., Traniello, I.M., Ray, A.M., Caldwell, A.S., Wickline, S.A., Robinson, G.E., 2019. Division of labor in honey bees is associated with transcriptional regulatory plasticity in the brain. J. Exp. Biol. 222, jeb200196.

Harris, K.D., Lloyd, J.P.B., Domb, K., Zilberman, D., Zemach, A., 2019. DNA methylation is maintained with high fidelity in the honey bee germline and exhibits global non-functional fluctuations during somatic development. Epigenet. Chromatin 12, 62.

Herb, B.R., Wolschin, F., Hansen, K.D., Aryee, M.J., Langmead, B., et al., 2012. Reversible switching between epigenetic states in honeybee behavioral subcastes. Nature Neurosci. 15, 1371–73.

Jaenisch, R., Bird, A., 2003. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. Nature Genet. 33, 245–254.

Jasper, W.C., Linksvayer, T.A., Atallah, J., Friedman, D., Chiu, J.C., Johnson, B.R., 2015. Large-scale coding sequence change underlies the evolution of postdevelopmental novelty in honey bees. Mol. Ecol. Evol. 32, 334-346.

Johnson, D.S., Mortazavi, A., Myers, R.M., Wold, B., 2007. Genome-wide mapping of *in vivo* protein-DNA interactions. Science 316, 1497-1502.

Johnson, B.R., Arallah, J., Plachetzki, D.C., 2013. The importance of tissue specficificty for RNA-seq: highlighting the errors of composite structure extractions. BMC Genom. 14, 586.

Jones, B.M, Robinson, G.E., 2018. Genetic accommodation and the role of ancestral plasticity in the evolution of insect eusociality. J. Exp. Biol. 221, jeb153163.

Jovanovic, M, Rooney M.S., Mertins, P., Przybylski, D., Chevrier, N., Satija, R., Rodriguez, E.H., Fields, A.P., Schwartz, S., Raychowdhury, R., Mumbach, M.R., Eisenhaure, T., Rabani, M., Gennert, D., Lu, D., Delorey, T., Weissman, J.S., Carr, S.A., Hacohen, N., Regev, A., 2015. Dynamic profiling of the protein life cycle in response to pathogens. Science 347, 1259038.

Kapheim, K.M., Pan, H., Li, C., Salzberg, S.L., Puiu, D., Magoc. T., Robertson, H.M., Hudson, M.E., Venkat, A., Fischman, B.J., Hernandez, A., Yandell, M., Ence, D., Holt, C., Yocum, G.D., Kemp, W.P., Bosch, J., Waterhouse, R.M., Zdobnov, E.M., Stolle, E., Kraus, F.B., Helbing, S., Moritz, R.F., Glastad, K.M., Hunt, B.G., Goodisman, M.A., Hauser, F., Grimmelikhuijzen, C.J., Pinheiro, D.G., Nunes, F.M., Soares, M.P., Tanaka, É.D., Simões, Z.L., Hartfelder, K., Evans, J.D., Barribeau, S.M., Johnson, R.M., Massey, J.H., Southey, B.R., Hasselmann, M., Hamacher, D., Biewer, M., Kent, C.F., Zayed, A., Blatti,

C., Sinha, S., Johnston, J.S., Hanrahan, S.J., Kocher, S.D., Wang, J., Robinson, G.E., Zhang G., 2015. Genomic signatures of evolutionary transitions from solitary to group living. Science 348, 1139-1143.

Kelly, A.M., Ophir, A.C., 2015. Compared to what: what can we say about nonapeptide function and social behavior without a frame of reference? Curr. Opin. Behav. Sci. 6, 97–103.

Khamis, A.M., Hamilton, A.R., Medvedeva, Y.A., Alam, T., Alam, I., Essack, M., Umylny, B., Jankovic, B.R., Naeger, N.L., Suzuki, M., Harbers, M., Robinson, G.E., Bajic, V.B., 2015. Insights into the transcriptional architecture of behavioral plasticity in the honey bee *Apis mellifera*. Sci. Rep. 5, 11136.

Kimes, P.K., Liu, Y., Hayes, D.N., Marron, J.S., 2017. Statistical significance for hierarchical clustering. Biometrics 73, 811-821.

Klug, H., Alonzo, S.H., Bonsall, M.B., 2012. Theoretical foundations of parental care, in: Royle, N.J., Smiseth, P.T., Kölliker, M. (Eds.), The evolution of parental care. Oxford University Press, Oxford, pp. 21-36.

Kohl, J., Babayan, B.M., Rubinstein, N.D., Autry, A.E., Marin-Rodriguez, B., Kapoor, V., Miyamishi, K., Zweifel, L.S., Luo, L., Uchida, N., Dulac, C., 2018. Functional circuit architecture underlying parental behaviour. Nature 556, 326-331.

Kohlmeier, P., Feldmeyer, B., Foitzik, S., 2018. Vitellogenin-like A-associated shifts in social cue responsiveness regulate behavioral task specialization in an ant. PLoS Biol. 16, e2005747.

Kohlmeier, P., Alleman, A.R., Libbrecht, R., Foitzik, S., Feldmeyer, B., 2019. Gene expression is more strongly associated with behavioural specialization than with age or fertility in ant workers. Mol. Ecol. 28, 658–670.

Korb, J., Heinz, J., 2016. Major Hurdles for the evolution of sociality. Annu. Rev. Entomol. 61, 297–316.

Kornienko, A.E., Guenzl, P.M., Barlow, D.P., Pauler, F.M., 2013. Gene regulation by the act of long non-coding RNA transcription. BMC Biol. 11, 59.

Koussounadis, A., Langdon, S.P., Um, I.H., Harrison, D.J., Smith, V.A., 2015. Relationship between differentially expressed mRNA and mRNA-protein correlations in a xenograft model system. Sci. Rep. 5, 10775.

Kramer, J., Meunier, J., 2018. The other facets of family life and their role in the evolution of animal sociality. Biol. Rev. 94, 199-215.

Kronauer D.J.C., Libbrecht, R., 2018. Back to the roots: the importance of using simple insect societies to understand the molecular basis of complex social life. Curr. Opin. Insect Sci. 28, 33–39.

962 Langfelder, P., Horvath, S., 2008. WGCNA: an R package for weighted correlation network analysis.
963 BMC Bioinformatics 9, 559.
964
965 Lewis, S.H., Quarles, K.A., Yang, Y., Tanguy, M., Frézal, L., Smith, S.A., Sharma, P.P., Cordaux, R.,
966 Gilbert, C., Giraud, I., Collins, D.H., Zamore, P.D., Miska, E.A., Sarkies, P., Jiggins, F.M., 2018. Pan-
967 arthropod analysis reveals somatic piRNAs as an ancestral defence against transposable elements. Nature
968 Ecol. Evol. 2, 174-181.
969
970 Li-Byarlay, H., Li, Y., Stroud, H., Feng, S., Newman, T.C., Kaneda, M., Hou, K.K., Worley, K.C., Elsik,
971 C.G., Wickline, S.A., Jacobsen, S.E., Ma, J., Robinson, G.E., 2013. RNA interference knockdown of
972 DNA methyltransferase 3 affects gene alternative splicing in the honey bee. Proc. Nat. Acad. Sci. USA
973 110, 12750-12755.
974
975 Libbrecht, R., Oxley, P.R., Keller, L., Kronauer, D.J.C., 2016. Robust DNA methylation in the clonal
976 raider ant brain. Curr. Biol. 26, 391–395.
977
978 Libbrecht, R., Oxley, P.R., Kronauer, D.J.C., 2018. Clonal raider ant brain transcriptomics identifies
979 candidate molecular mechanisms for reproductive division of labor. BMC Biol. 16, 89.
980
981 Linksvayer, T.A., 2015. The molecular and evolutionary genetic implications of being truly social for the
982 social insects. Advan. Insect Physiol. 48, 271-291.
983
984 Linksvayer, T.A., Fewell, J.H., Gadau, J., Laubichler, M.D., 2012. Developmental evolution in social
985 insects: regulatory networks from genes to society. J. Exp. Zool. B Mol. Develop. Evol. 318, 159-169.
986
987 Linksvayer, T.A., Johnson, B.R., 2019. Re-thinking the social ladder approach for elucidating the
988 evolution and molecular basis of insect societies. Curr. Opin. Insect Sci. 34, 123-129.
989
990 Liu, F., Shi, T., Qi, L., Su, X., Wang, D., Dong, J., Huang, Z.Y., 2019a. lncRNA profile of *Apis mellifera*
991 and its possible role in behavioural transition from nurses to foragers. BMC Genom. 20, 393.
992
993 Liu, X., Li, Y.I., Pritchard, J.K., 2019b. Trans effects on gene expression can drive omnigenic inheritance.
994 Cell 177, 1022-1034.
995
996 Lu Z., Hofmeister, B.T., Vollmers, C., DuBois, R.M., Schmitz, R.J., 2017. Combining ATAC-seq with
997 nuclei sorting for discovery of cis-regulatory regions in plant genomes. Nuc. Acids Res., gkw1179.
998
999 Lyko, F., Foret, S., Kucharski, R., Wolf, S., Falckenhayn, C., Maleszka, R., 2011. The honey bee
1000 epigenomes: differential methylation of brain DNA in queens and workers. PLoS Biol. 9, 10.1371.
1001
1002 Machanick, P., Bailey, T.L., 2011. MEME-ChIP: motif analysis of large DNA datasets. Bioinfor. 27,
1003 1696-1697.
1004

Marshall, H., Lonsdale, Z.N., Mallon, E.B., 2019. Methylation and gene expression differences between reproductive and sterile bumblebee workers. Evol. Lett. 3, 485-499.

Marshall, H., Jones, A.R.C., Lonsdale, Z.N., Mallon, E.B., 2020. Bumblebee worker castes show differences in allele-specific DNA methylation and allele-specific expression. bioRxiv, doi.org/10.1101/2020.02.07.938423.

Maynard-Smith, J., Szathmáry, E., 1995. The Major Transitions in Evolution. Oxford University Press, Oxford.

McGlothlin, J.W., Moore, A.J., Wolf, J.B., Brodie III, E.D., 2010. Interacting phenotypes and the evolutionary process. III. Social evolution. Evolution 64, 2558-2574.

Mikheyev, A.S., Linksvayer, T.A., 2015. Genes associated with ant social behavior show distinct transcriptional and evolutionary patterns. eLife 4, e04775.

Molodtsova, D., Harpur, B.A, Kent, C.F., Seevananthan, K., Zayed, A., 2014. Pleiotropy constrains the evolution of protein but not regulatory sequences in a transcription regulatory network influencing complex social behaviours. Front. Genet. 5, 431.

Montgomery, S.H., Mank, J.E., 2016. Inferring regulatory change from gene expression: the confounding effect of tissue scaling. Mol. Ecol. 25, 5114-5128.

Moore, A.J., Benowitz, K.M., 2019. From phenotype to genotype: the precursor hypothesis predicts genetic influences that facilitate transitions in social behavior. Curr. Opin. Insect Sci. 34, 91–96.

Moore, A.J., Brodie III, E.D., Wolf, J.B., 1997. Interacting phenotypes and the evolutionary process. I. Direct and indirect genetic effects of social interactions. Evolution 51, 1352-1362.

Morrissey, M.B., Ruxton, G.E., 2018. Multiple regression is not multiple regressions: the meaning of multiple regression and the non-problem of collinearity. Phil. Thero. Prac. Biol. 10, 3.

NESCent Working Group on Integrative Models of Vertebrate Sociality: Evolution, Mechanisms, and Emergent Properties, 2014. An evolutionary framework for studying mechanisms of social behavior Trends Ecol. Evol. 29, 581-589.

NESCent Working Group on Integrative Models of Vertebrate Sociality: Evolution, Mechanisms, and Emergent Properties, 2015. Taxon matters: promoting integrative studies of social behavior. Trends Neurosci. 38, 189-191.

Palmer, W.J., Duarte, A., Schrader, M., Day, J.P., Kilner, R., Jiggins, F.M., 2016. A gene associated with social immunity in the burying beetle *Nicrophorus vespilloides*. Proc. Roy. Soc. B 283, 20152733.

Parker, D.P., Cunningham, C.B., Walling, C.A., Stamper, C.E., Head, M.L., Roy-Zokan, E.M., McKinney, E.C., Ritchie, M.G., Moore, A.J., 2015. Transcriptomes of parents identify parenting strategies and sexual conflict in a subsocial beetle. Nature Comm. 6, 8449.

Peres-Neto, P.R., Donald, A.J., Keith, M.S., 2003. Giving meaningful interpretation to ordination axes: assessing loading significance in principal component analysis. Ecol. 84, 2347–2363.

Phelps, S.M., Campbell, P., Zheng, D., Ophir, A.G., 2010. Beating the boojum: Comparative approaches to the neurobiology of social behavior. Neuropharmocol. 58, 17–28.

Porath, H.T., Hazan, E., Shpigler, H., Cohen, M., Band, M., Ben-Shahar, Y., Levanon, E.Y., Eisenberg, E., Bloch, G., 2019. RNA editing is abundant and correlates with task performance in a social bumblebee. Nature Comm. 10, 1605.

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 Evol. Biol. 14, 260.

Rehan, S.M, Toth, A.L., 2015. Climbing the social ladder: the molecular evolution of sociality. Trends Ecol. Evol. 30, 426-433.

Rittschof, C.C., Robinson, G.E., 2014. Genomics: moving behavioural ecology beyond the phenotypic gambit. Anim. Behav. 92, 263-270.

Rittschof, C.C., Robinson, G.E., 2016. Behavioral genetic toolkits: Toward the evolutionary origins of complex phenotypes. Curr. Top. Devel. Biol. 119, 157-204.

Rittschof, C.C., Hughes, K.A., 2018. Advancing behavioural genomics by considering timescale. Nature Comm. 9, 489.

Robine, N., Lau, N.C., Balla, S., Jin, Z., Okamura, K., Kuramochi-Miyagawa, S., Blower, M.D., Lai, E.C., 2009. A broadly conserved pathway generates 3'UTR-directed primary piRNAs. Curr. Biol. 19, 2066-2076.

Robinson, G.E., Ben-Shahar, Y., 2002. Social behavior and comparative genomics: new genes or new gene regulation? Genes Brain Behav. 1: 197–203.

Rubenstein, D.R., Hofmann, H.A., 2015. Proximate pathways underlying social behavior. Curr. Opin. Behav. Sci. 6, 154-159.

Rubenstein, D.R., Ågren, J.A., Carbone, L., Elde, N.C., Hoekstra, H.E., Kapheim, K.M., Keller, L., Moreau, C.S., Toth, A.L., Yeaman, S., Hofmann, H.A., 2019. Coevolution of Genome Architecture and Social Behavior, Trends Ecol. Evol. 34, 844-855.

Saelens, W., Cannoodt, R., Saeys, Y., 2018. A comprehensive evaluation of module detection methods for gene expression data. Nature Comm. 9, 1090.

Schneider, J., Atallah, J., Levine, J.D., 2015. Social structure and indirect genetic effects: genetics of social behaviour. Biol. Reviews 92, 1027-1038.

Shpigler.H.Y., Siegel, A.J., Huang, Z.Y., Bloch, G., 2016. No effect of juvenile hormone on task performance in a bumblebee (*Bombus terrestris*) supports an evolutionary link between endocrine signalling and social complexity. Horm. Behav. 85, 67-75.

Shell, W.A, Rehan, S.M., 2019. Social modularity: conserved genes and regulatory elements underlie caste-specific antecedent behavioural states in an incipiently social bee. Proc. Roy.Soc. B Biol. Sci. 286, rspb.2019.1815.

Sokolowski, M.B., 2001. *Drosophila*: genetics meets behaviour. Nature Rev. Genet. 2, 879–890.

Shorter, J., Couch, C., Huang, W., Carbone, M.A., Peigger, J., Anholt, R.R.H., Mackay, T.F.C., 2015. Genetic architecture of natural variation in *Drosophila melanogaster* aggressive behaviour. Proc. Nat. Acad. Sci. USA Plus, E3555–E3563.

Silk, M.J., Finn, K.R., Porter, M.A., Pinter-Wollman, N., 2018. Can multilayer networks advance animal behavior research. Trends Ecol. Evol. 33, 376-378.

Simola, D.F., et al., 2013. Social insect genomes exhibit dramatic evolution in gene composition and regulation while preserving regulatory features linked to sociality. Genome Res. 23, 1235-1247.

Song, L., Huang, S.C., Wise, A., Castanon, R., Nery, J.R., Chen, H., Watanabe, M., Thomas, J., Bar-Joseph, Z., Ecker, J.R., 2016. A transcription factor hierarchy defines an environmental stress response network. Science 354, aag1550.

Stark, R., Grzelak, M., Hadfield, J., 2019. RNA sequencing: the teenage years. Nature Rev. Genetics, 20, 631-656.

Székely, T., Remeš, V., Freckleton, R.P., Liker, A., Why care? Inferring the evolution of complex social behaviour. J. Evol. Biol. 26, 1381-1391.

Taborsky, M., Taborsky, B., 2015. Evolution of genetic and physiological mechanisms of cooperative behaviour. Curr. Opin. Behav. Sci. 6, 132-138.

Tallamy, D.W., 1984. Insect parental care. 34, 20-24.

Thölken, C., Thamm, M., Erbacher, C., Lechner, M., 2016. Sequence and structural properties of circular RNAs in the brain of nurse and forager honey bees (*Apis mellifera*). BMC Genomics 20, 88.

1135 Todd, E.V., Black, M.A., Gemmell, N.J., 2016. The power and promise of RNA-seq in ecology and
1136 evolution. Mol. Ecol. 25, 1224-1241.
1137
1138 Toth, A., Robinson, G.E., 2007. Evo-devo and the evolution of social behaviour. Trends Genet. 23, 334-
1139 341.
1140
1141 Trible, W., Olivos-Cisneros, L., McKenzie, S.K., Saragosti, J., Chang, N., Matthews, B.J., Oxley, P.R.,
1142 Kronauer, D.J.C., 2017. *orco* mutagenesis causes loss of antennal lobe glomeruli and impaired social
1143 behavior in ants. Cell 170, 727-735.
1144
1145 Trumbo, S.T., 2012. Patterns of parental care in invertebrates, in: Royle, N.J., Smiseth, P.T., Kölliker, M.
1146 (Eds.), The evolution of parental care. Oxford University Press, Oxford, pp. 81-100.
1147
1148 Vogel, C., Marcotte, E.M., 2012. Insights into the regulation of protein abundance from proteomic and
1149 transcriptomic analyses. Nature Rev. Genetics 13, 227-232.
1150
1151 Vojvodic, S., Johnson, B.R., Harpur, B.A., Kent, C.F., Zayed, A., Anderson, K.E., Linksvayer, T.A.,
1152 2015. The transcriptomic and evolutionary signature of social interactions regulating honey bee caste
1153 development. Ecol. Evol. 5, 4795-4807.
1154
1155 Wang, J., Ross, K.G., Keller, L., 2008. Genome-Wide Expression Patterns and the Genetic Architecture
1156 of a Fundamental Social Trait. PLoS Genetics 4, e1000127.
1157
1158 Warner, M.R., Qiu, L., Homes, M.J., Mikheyev, A.S., Linksvayer, T.A., 2019a. Convergent eusocial
1159 evolution is based on a shared reproductive groundplan plus lineage-specific plastic genes. Nature Comm.
1160 10, 2651.
1161
1162 Warner, M.R., Mikheyev, A.S., Linksvayer, T.A., 2019b. Transcriptomic basis and evolution of the ant
1163 nurse-larval social interactome. PLoS Genet. 15, e1008156.
1164
1165 Walsh, J.T., Warner, M.R., Kase, A., Cushing, B.J., Linksvayer, T.A., 2018. Ant nurse workers exhibit
1166 behavioural and transcriptomic signals of specializations on larval stage. Anim. Behav. 141, 161-169.
1167
1168 Wittkopp, P.J., Kalay, G., 2012. *cis*-regulatory elements: molecular mechanisms and evolutionary
1169 processes underlying divergence. Nature Rev. Genetics 13, 59-69.
1170
1171 Wright, S., 1968. Evolution and the Genetics of Populations Volume 1: Genetic and Biometric
1172 Foundations. University of Chicago Press, Chicago.
1173
1174 Wu, F., Ma, C., Han, B., Meng, L., Hu, H., Fang, Y., Feng, M., Zhang, X., Rueppell, O., Li, J., 2019.
1175 Behavioural, physiological, and molecular changes in alloparental caregivers may be responsible for
1176 selection response for female reproductive investment in honey bees. Mol. Ecol. 28, 4212-4227.
1177
1178 Wu, M., Walser, J.-C., Sun, L., Kölliker, M., 2020. The genetic mechanism of selfishness and altruism

in parent-offspring coadaptation. Science Advances 6, eaaw0070.

Yan, H., Opachaloemphan, C., Mancini, G., Yang, H., Gallitto, M., Mlejnek, J., Leibholz, A., Haight, K., Ghaninia, M., Huo, L., Perry, M., Slone, J., Zhou, X., Traficante, M., Penick, C.A., Dolezal, K., Gokhale, K., Stevens, K., Fetter-Pruneda, I., Bonasio, R., Zwiebel, L.J., Berger, S.L., Liebig, J., Reinberg, D., Desplan, C., 2017. An engineered *orco* mutation produces aberrant social behavior and defective neural development in ants. Cell 170, 736-747.

Zayad, A., Robinson, G.E., 2012. Understanding the relationship between brain gene expression and social behaivor: lessons from the honey bee. Ann. Rev. Genetics 46, 591-615.