Immunological and Gene Regulatory Functions of the Protein Vitellogenin in Honey Bees (*Apis mellifera*)

by

Gyan Harwood

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved November 2019 by the
Graduate Supervisory Committee:

Gro Amdam, Chair
Kenro Kusumi
Yung Chang
Christian Rabeling

ARIZONA STATE UNIVERSITY

December 2019

ABSTRACT

Vitellogenin (Vg) is an ancient and highly conserved multifunctional protein. It is primarily known for its role in egg-yolk formation but also serves functions pertaining to immunity, longevity, nutrient storage, and oxidative stress relief. In the honey bee (*Apis mellifera*), Vg has evolved still further to include important social functions that are critical to the maintenance and proliferation of colonies. Here, Vg is used to synthesize royal jelly, a glandular secretion produced by a subset of the worker caste that is fed to the queen and young larvae and which is essential for caste development and social immunity. Moreover, Vg in the worker caste sets the pace of their behavioral development as they transition between different tasks throughout their life. In this dissertation, I make several new discoveries about Vg functionality. First, I uncover a colony-level immune pathway in bees that uses royal jelly as a vehicle to transfer pathogen fragments between nestmates. Second, I show that Vg is localized and expressed in the honey bee digestive tract and suggest possible immunological functions it may be performing there. Finally, I show that Vg enters to nucleus and binds to deoxyribonucleic acid (DNA), acting as a potential transcription factor to regulate expression of many genes pertaining to behavior, metabolism, and signal transduction pathways. These findings represent a significant advance in the understanding of Vg functionality and honey bee biology, and set the stage for many future avenues of research.

ACKNOWLEDGMENTS

I would first like to thank all of my friends, family, and colleagues for their unwavering support and inspiration over the years. Nothing is achieved alone, and I would not be where I am today if it was not for all of these fantastic people.

Second, I would like to thank my dissertation committee for all of their professional support and constructive feedback that has helped me grow as a scientist. In particular, I would like to thank my advisor and mentor, Gro Amdam. Her wealth of knowledge, compassion, and enthusiasm is truly inspirational and something I will strive to emulate in my life.

Finally, I would like to thank all of my funding sources that helped make this research possible. This includes the ASU School of Life Sciences and the Graduate and Professional Student Association, as well as the Natural Science and Engineering Research Council of Canada.

TABLE OF CONTENTS

Page

CHAPTER Page

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

Figure Page

CHAPTER 1

INTRODUCTION

Eusociality represents a major evolutionary transition in history, in which individuals forgo direct reproductive output and instead live in colonies of close relatives that display separate reproductive and worker castes, cooperative care of young, and overlapping generations (Michener 1969a; Crespi and Yanega 1995; Wilson and Hölldobler 2005). This phenomenon can be observed in a few crustacean (Duffy 1996) and mammalian species (Jarvis 1981; Burda et al. 2000), but it is predominantly found in social insects like bees, ants, wasps, and termites (Wilson 1971a). Eusociality has proven to be a successful evolutionary strategy for social insects, as they now dominate many terrestrial ecosystems in terms of abundance and biomass (King et al. 2013). The maintenance and proliferation of social insect colonies requires them to overcome several challenges, including the need to combat pathogens, and the need to regulate division of labor among nestmates.

In this first challenge, pathogen defense is required by all organisms, but social insect colonies can face heightened pressure owing to their dense populations of genetically similar individuals and their homeostatic conditions within the nest (Wilson 1971a). In response, social insects have evolved "social immunity", a suite of behavioral and physiological phenotypes that impedes pathogen transmission within the colony (Traniello et al. 2002; Cremer 2019). This includes allogrooming, removal of dead or infected individuals, increasing internal nest temperatures by bodily vibrations to kill pathogens (i.e., a "behavioral fever"), and sharing antimicrobial compounds among nestmates (Traniello et al. 2002; Cremer et al. 2007; Meunier 2015; Cremer 2019). In the

second challenge, division of labor allows for task specialization among workers, which enable colonies to efficiently exploit resources and flexibly respond to changing colony needs and shifting resource availability (Robinson 1992; Beshers and Fewell 2001). Some species create morphologically distinct workers for specific tasks (e.g., major and minor ant workers), while others have their workers progress through an age-dependent series of tasks, including nest construction, brood rearing, colony defense, and foraging. Many of the genetic, physiological, and chemical regulatory mechanisms that control worker behavioral development are co-opted from ancient regulatory networks still found in solitary animals. For example, shifts from brood rearing to foraging tasks are controlled by reproductive regulatory networks, even though the worker caste does not typically reproduce (Amdam, Norberg, et al. 2004; Amdam, Csondes, et al. 2006).

Advances in molecular techniques have increased our understanding and appreciation for just how integral some of these ancient regulatory mechanisms have been to the evolution of eusocial species. One key player that has critical functions in reproduction, social immunity, and division of labor is Vitellogenin (Vg), an ancient and highly conserved protein found in nearly all animals (see Background). In this dissertation, I use a plethora of molecular biology, imaging, genetic, and bioinformatical approaches to uncover novel functions of Vg pertaining to the maintenance of social immunity and division of labor in one of the preeminent social insect model systems, the honey bee (*Apis mellifera*). Specifically, I investigate Vg's role in an immune system pathway allowing for the transfer of immune elicitors (i.e., pathogen fragments) between nestmates, and I uncover Vg's hitherto unknown function as a possible transcription factor regulating genes involved in complex phenotypes, including division of labor.

BACKGROUND

Evolutionary transitions do not necessarily require the development of new biological processes or novel genes but can instead be achieved by re-organizing the regulatory pathways that control already-extant genes. In social insects like honey bees, processes and genes that exist in solitary species can undergo selection and lead to the emergence of complex phenotypes like social immunity and division of labor (Amdam, Norberg, et al. 2003; Amdam, Csondes, et al. 2006). In this regard, much of honey bee biology makes sense in light of Vg and how it has evolved over time.

Reproductive functions of Vg

Vg is a glycolipophosphoprotein, meaning it is a protein with properties of being a carbohydrate and a phosopholipid. It was first identified in the late 1960s for its role in egg production, where it delivers lipids and other nutrients to the embryo and serves as a yolk protein precursor (Pan et al. 1969). Vg is primarily synthesized in non-ovarian tissues such as the liver (Wang et al. 2005), adipose tissue (fat) (Brookes 1969; Pan et al. 1969), or the hepatopancreas (Guan et al. 2016), and then secreted into the blood or hemolymph where it can be taken up by the ovaries and other tissues (Noah Koller et al. 1989; Raikhel and Dhadialla 1992). In insects, Vg is primarily synthesized in the fat body, an organ analogous to the vertebrate liver that plays important roles in insect nutrition, metabolism, immunology, and pheromone production (Wigglesworth 1988; de Oliveira and Cruz-Landim 2003; Oliveira and Cruz-Landim 2006; Wicker-Thomas et al. 2009; Arrese and Soulages 2010; Makki et al. 2014). Given its central role in egg production, Vg is common to nearly every extant egg-laying species. It is part of a large

family of proteins known as Large Lipid Transfer Proteins (LLTPs), which are used to transport lipids throughout the body (Shikina et al. 2013; Wu et al. 2013). In non-egg-laying animals such as therian mammals, LLTPs include important cholesterol transporters like apolipoprotein-B (ApoB) and microsomal triglyceride transporter (MTP) (Babin et al. 1999). LLTPs first evolved roughly 750 million years ago around the time that metazoans (animals) first appeared, and multiple studies have shown Vg to be the oldest member of this family (Baker 1988a; Hayward et al. 2010). It is unknown whether Vg's original function was for egg production, but regardless, it has had ample time to evolve numerous non-reproductive functions.

Non-reproductive functions of Vg

Some of Vg's non-reproductive functions pertain to immunity, inflammation, and longevity. Vg is a pathogen pattern recognition receptor, allowing it to bind to numerous bacterial and fungal pathogens (Zhang et al. 2005; Shi et al. 2006; Li et al. 2008, 2009; Liu et al. 2009). It does this my recognizing molecular moieties that are not found in animal cells, so called pathogen-associated molecular patterns (PAMPs). These PAMPs include lipopolysaccharide and peptidoglycan, which are predominant components in the cell walls of Gram negative and Gram positive bacteria, respectively, as well as zymosan, a key component of fungal cell walls (Salmela et al. 2015). Once bound to a pathogen, Vg can kill pathogens directly, in the case of some bacteria (Li et al. 2009), or it can act as an opsonin to guide other humoral components of the immune system to destroy the pathogen (Li et al. 2008; Zhang et al. 2011). Vg further helps immune cells in insects by transporting the zinc necessary for their proper functioning (Amdam, Simões, et al.

2004). Non-cellular pathogens, i.e., viruses, can also be bound by Vg (Garcia et al. 2010; Huo et al. 2014; Whitfield et al. 2015), and in some cases neutralized (Garcia et al. 2010). Vg also plays key roles in the body's anti-inflammatory response. *Vg* expression (the gene name for the same protein) is upregulated in response to wounding, and the Vg protein binds to phosphatidylserine, a lipid that is exposed on damaged or necrotic cells (Havukainen et al. 2013). This helps to protect vulnerable cells from further cellular damage and may promote cellular clearance. Vg's mammalian orthologs, like ApoB, are also known to suppress inflammation by binding to necrotic cells (Cho and Seong 2009). What's more, Vg acts as an antioxidant by binding to and neutralizing reactive oxygen species, thereby protecting host cells from cellular damage (Seehuus et al. 2006; Havukainen et al. 2013; Sun and Zhang 2015a; Salmela et al. 2016). Vg's anti-pathogenic, anti-inflammatory, and antioxidant properties, along with its role in nutrient storage, are believed to contribute to organismal longevity. For example, honey bee queens have very high levels of Vg and live for several years, quite old for an insect, whereas the worker caste has much lower levels of Vg and only lives a couple of months. Compare this with so-called winter bees, a group of worker bees that live for several months inside the hive during the nutritionally barren winter months, and which also possess very high titers of Vg (Amdam and Omholt 2002; Amdam, Simões, et al. 2004; Amdam, Norberg, et al. 2005a; Seehuus et al. 2006; Corona et al. 2007a; Ihle et al. 2015; Münch et al. 2015). Interestingly, Vg's anti-pathogenic and antioxidant properties are seen across a wide array of taxa including fish and coral (Zhang et al. 2005; Li et al. 2008, 2009; Du et al. 2017), suggesting that such functions evolved early in Vg's history contemporary with its reproductive functions.

Protein structure

Vg's multifunctionality stem from its ability to bind to numerous different ligands, and this ability arises from its biochemical structure. Vg is a relatively large protein (~200kDa) that can undergo numerous post-translational modifications like glycosylation, phosphorylation, and lipidation, as well as cleavage at various sites in different species (Tufail and Takeda 2008). For example, the 180 kDa honey bee Vg is cleaved into smaller units of 150 kDa and 40 kDa (Havukainen, Halskau, Skjaerven, et al. 2011). There are several structural domains that are conserved across species, both at the sequence and structure level. This includes the N-terminus β-barrel (also known as the N-sheet), the α-helical domain, the von Willebrand factor type D domain (VWD) and the domain of unknown function (DUF) 1943. The N-sheet contains the purported receptor-binding domain as well as a positively charged lipophilic cavity used to transport lipids (Havukainen, Halskau, Skjaerven, et al. 2011; Roth et al. 2013), while the α-helical domain is known to mediate Vg's anti-inflammatory actions (Havukainen, Halskau, Skjaerven, et al. 2011; Salmela et al. 2016). Both the VWD and DUF1943 contribute to Vg's functions as a pathogen pattern recognition receptor and opsonin (Sun et al. 2013a). These conserved domains allow for Vg to carry out multiple functions in diverse animal taxa.

Vg in honey bees

In honey bees, Vg has evolved still further to include several sophisticated functions pertaining to social immunity and the division of labor. First, Vg is a key component of a phenomenon called trans-generational immune priming (TGIP). TGIP

allows a female that survives a pathogen attack to produce offspring that are more resistant to that pathogen. In animals with an *acquired* immune system, like mammals, this can be achieved when pathogen-specific antibodies produced in the mother are transferred to her offspring via the placenta or breast milk (Shahid et al. 2002). However, animals that only possess an *innate* immune system, like insects, cannot produce antibodies and yet a surviving female can still produce more pathogen-resistant offspring. Here, TGIP is achieved when fragments of the destroyed pathogen that contain PAMPs (e.g., bits of the cell wall) are transferred into the female's eggs, where the immune system of developing embryos becomes more activated in response (Freitak et al. 2009; López et al. 2014a; Knorr et al. 2015). In honey bees, Vg binds to pathogens and carries them across the cell membrane into developing eggs via receptor-mediated endocytosis (Salmela et al. 2015). Second, Vg is used as an amino acid donor in the production of royal jelly, a protein-rich food source with important roles in both social immunity and division of labor (Amdam, Norberg, et al. 2003). Royal jelly is a glandular secretion produced by a subset of the worker caste known as nurses, who feed it to the queen and young larvae. This food not only contains many antimicrobial compounds that protect the queen and larvae from disease (Blum et al. 1959; Bíliková et al. 2001; Bachanová et al. 2002a; Fontana, Mendes, et al. 2004; Klaudiny et al. 2012; Sugiyama et al. 2012; Bucekova et al. 2014), but it also alters signaling pathways in larvae that regulate their development into either a queen or a worker (Evans and Wheeler 1999). Nurses feed royal jelly to all young larvae, but after 3 days, larvae that are switched to a diet of pollen and honey develop into workers, while those that continue to be fed royal jelly develop into queens. Queens continue to feed exclusively on royal jelly throughout their adult

lives (Haydak 1970). Finally, Vg helps regulate behavioral development in the worker caste. Workers transition between different tasks as they age, and the timing of these transitions is determined, in part, by Vg titers in their haemolymph: Newly emerged workers have low Vg titers and spend their first few days cleaning comb cells, but soon transition into nurses as their Vg production increases greatly. After a couple weeks, a drop in Vg titers and a concomitant increase in juvenile hormone prompt nurses to transition into foragers (Amdam and Omholt 2003; Guidugli, Nascimento, et al. 2005; Amdam, Csondes, et al. 2006; Nelson et al. 2007; Antonio et al. 2008). Vg titers at different developmental stages can also influence other worker behaviors such as their responsiveness to sucrose and their foraging preference for nectar or pollen (Amdam, Norberg, et al. 2006; Nelson et al. 2007). However, the molecular mechanisms by which Vg regulates such behavioral changes are not fully understood.

PURPOSE

The purpose of this research project is to discover novel molecular mechanisms by which Vg contributes to social immunity and division of labor in honey bees. First, I build on research originally undertaken by Schmid-Hempel and colleagues (Sadd et al. 2005) showing social insects display TGIP, and later expanded upon by my own collaborators elucidating Vg's role in this pathway (Salmela et al. 2015). Here, I uncover a potential colony-level immune pathway in honey bees that extends TGIP to include the worker caste. Queens have been shown to transfer pathogen particles to offspring (Hernández López Javier et al. 2014; Salmela et al. 2015), but these studies involved injecting queens with an inoculum and such scenarios do not occur in nature. There are

few opportunities for queens to be inoculated naturally because they rarely leave the nest (mating flights and swarming events being the exception) and they do not consume the potentially-contaminated nectar or pollen collected by foragers. However, queens feed exclusively on the royal jelly produced by nurses, and this may serve as a vehicle for transferring pathogen particles from workers to queen, or from workers to young larvae directly. In this dissertation, I show that nurses that ingest bacteria can transport these bacteria from their gut to their glands where royal jelly is produced, and that this phenomenon is not observed in nurses with experimentally-reduced *vg* expression. I also show that ingested bacteria are ultimately incorporated into the royal jelly, and that royal jelly from nurses fed pathogenic bacteria has higher concentrations of a potent antimicrobial peptide. Furthermore, I show for the first time that Vg is localized and expressed in worker gut issue, a key organ in insect immunity, where it may carry out antipathogenic and antioxidant functions.

Second, I build on recent research showing that Vg-knockdown affects expression in thousands of genes (Wheeler et al. 2013), as well as research by my collaborator showing Vg capable of entering the nucleus (Salmela et al, *submitted*). Both of these finds suggested that Vg may be entering the nucleus to directly regulate gene expression. I make a major new discovery that Vg is able to bind to DNA and potentially serve as a transcription factor to regulate genes involved in complex phenotypes, like behavior. I show that Vg binds at hundreds of DNA loci, and that the identity of these loci shift as workers transition from different tasks like cell cleaning, nursing, and foraging. Furthermore, I identify other proteins bound to the Vg-DNA complex, which shed light on the gene regulatory networks that Vg may be a part of.

SIGNIFICANCE

The work presented in this dissertation greatly enhances our understanding of Vg multifunctionality and the mechanisms by which it carries out such functions. I make two important discoveries regarding a colony-level immune pathway in honey bees, and transcriptional regulatory properties of Vg. By understanding how a colony-level immune pathway operates in honey bees, researchers can exploit this mechanism to deliver "vaccines" to honey bees and make them resistant to deadly and economically costly diseases. That is, nurses can be fed with an inert form of a pathogen and deliver it to the queen and larvae via royal jelly and thus prime the immune system of future generations. This discovery has already been put into action by my collaborators, who are developing and marketizing an edible vaccine for honey bees against certain pathogens (Salmela and Freitak 2017).

The discovery that Vg is a DNA-binding protein and a likely transcription factor has the potential to be paradigm shifting in terms of how we understand Vg to regulate so many complex phenotypes. We have known that Vg titers are associated with different worker task groups, and that it mutually regulates other behavior-related biochemicals like juvenile hormone, but we have lacked a clear mechanistic understanding of how it regulates behavior on a molecular level. The discoveries outlined in this dissertation suggest that Vg may be acting to directly regulate expression of genes related to behavior and other phenotypes. In at least a dozen cases, Vg-bound genes are associated with differing gene expression levels between different worker task groups, including at least one gene known to be critical for division of labor in an ant species (Gospocic et al. 2017). Many Vg-bound genes play key roles in signal transduction pathways known to

affect behavior, as well as important genes in the immune system. Furthermore, this is the first documentation of any protein from the LLTP superfamily having DNA-binding properties, and given that Vg is a ubiquitous and highly conserved protein across metazoan taxa, it may be performing similar functions in a large number of animals. The discoveries made here may spur other research to investigate Vg's gene regulatory abilities in other organisms.

APPROACH

This dissertation utilized a number of methods to further elucidate Vg's role in immunity and gene regulation, including immunohistochemistry, confocal microscopy, mass spectrometry, and DNA- and RNA-sequencing. In **chapter 2**, I sought to determine whether ingested bacteria can be transported to the hypopharyngeal glands (the site of royal jelly synthesis), and to understand the role that Vg may play in this. To achieve this, my collaborators and I fed fluorescently-labelled bacteria to nurse bees and used confocal microscopy to detect the bacteria in several tissues: the hypopharyngeal glands, the fat body, and the midgut. To understand whether bacteria transport is affected by Vg titers, we subjected some of the nurses to RNA-interference (RNAi)-mediated gene knockdown by injecting them with double stranded RNA (dsRNA) against a portion of the *vg* gene. I used Reverse Transcription Quantitative Polymerase Chain Reaction (RT-qPCR) to quantify transcript abundance for *vg*, as well as a housekeeping gene and other genes that could potentially be affected by the dsRNA injection. This chapter has been published in *The Journal of Insect Physiology*, for which I am the first author. My contributions included rearing bees and administering a bacterial diet, staining and imaging tissue,

providing funds for microscopy work, analyzing image data, performing RT-qPCR experiments and statistical analysis, and primary manuscript drafting.

In **chapter 3**, I sought to determine not only whether pathogenic bacteria can be incorporated into royal jelly produced by nurses, but also whether pathogen exposure leads to higher levels of other immunological proteins in the jelly. To do this, I fed either a pathogen diet (containing fluorescently-labelled *Paenibacillus larvae*) or a control diet to small colonies of nurses and harvested the royal jelly they produced. I used fluorescent microscopy to detect bacteria in the royal jelly, and used mass spectrometry to determine the identity and relative abundance of other constituent proteins in royal jelly samples. This chapter is being prepared for submission to the journal *PLoS Pathogens*, and I am the first author. My contributions included experimental design, establishing colonies, harvesting royal jelly, performing fluorescent microscopy and analyzing images, analyzing mass spectrometry results, and primary manuscript drafting.

In **chapter 4**, I sought to build on a discovery made in chapter 2, namely, that Vg appeared to be present in worker midgut tissue. This is an organ critical to initial immune response against ingested pathogens, but Vg has not been documented here before. I sought to compare and contrast how Vg is localized in this tissue for nurses and foragers, and to determine whether *vg* is transcribed therein. To do this, I used a combination of immunohistochemistry, confocal microscopy, and RT-qPCR. This chapter is being prepared for submission to the journal *Apidologie*. My contributions amount to all aspects of experimental design, data collection, statistical analysis, and primary manuscript drafting.

In **chapter 5**, my collaborators and I sought to greatly expand our understanding of Vg's protein structure, its structural response to pathogen challenge, its translocation into the nucleus and subsequent DNA binding there, and the classes of genes to which it binds in differently aged bees. This study involved over half a dozen researchers employing a multitude of molecular, theoretical, and computational methods. For my part, I used a method called chromatin immunoprecipitation followed by high-throughput sequencing (ChIP-seq), which is the standard approach for examining protein-DNA interactions. I then performed a Gene Ontology analysis to look for significant enrichment of functional terms for Vg-DNA binding sites in newly emerged and nurse bees. This chapter will be submitted shortly to *PLoS Biology*, and I am listed as the second author. In addition to my experimental and analytical contributions, I wrote a significant portion of the manuscript and took the lead in the subsequent edits made after the first draft. I also provided partial funding to obtain and sequence samples in the ChIP-seq protocol.

In **chapter 6**, I sought to determine whether Vg-DNA binding could potentially regulate gene expression in nurses and foragers, and to determine what other proteins may be interacting with Vg at the DNA in a regulatory complex. To do this, I used (1) ChIP-seq to map out Vg-DNA binding sites in same-age nurses and foragers, (2) RNA-seq to measure transcript abundance at these loci, and (3) co-immunoprecipitation paired with mass spectrometry to identify other nuclear proteins that are bound to Vg. This chapter is in preparation for submission to one of *Nature*, *PNAS*, or *PLoS Genetics*. Apart from designing and performing all molecular laboratory experiments, and performing

Gene Ontology and other post-hoc analyses after sequencing and mass spectrometry, I am also the lead author on this manuscript.

CHAPTER 2

THE ROLE OF VITELLOGENIN IN THE TRANSFER OF IMMUNE ELICITORS FROM GUT TO HYPOPHARYNGEAL GLANDS IN HONEY BEES

Gyan Harwood, Gro Amdam, and Dalial Freitak

Published in the *Journal of Insect Physiology* (112) 2019

ABSTRACT

Female insects that survive a pathogen attack can produce more pathogen-resistant offspring in a process called trans-generational immune priming. In the honey bee (*Apis mellifera*), the egg-yolk precursor protein Vitellogenin transports fragments of pathogen cells into the egg, thereby setting the stage for a recruitment of immunological defenses prior to hatching. Honey bees live in complex societies where reproduction and communal tasks are divided between a queen and her sterile female workers. Worker bees metabolize Vitellogenin to synthesize royal jelly, a protein-rich glandular secretion fed to the queen and young larvae. We ask if workers can participate in trans-generational immune priming by transferring pathogen fragments to the queen or larvae via royal jelly. As a first step toward answering this question, we tested whether worker-ingested bacterial fragments can be transported to jelly-producing glands, and what role Vitellogenin plays in this transport. To do this, we fed fluorescently labelled *Escherichia coli* to workers with experimentally manipulated levels of Vitellogenin. We found that bacterial fragments were transported to the glands of control workers, while they were not detected at the glands of workers subjected to RNA interference-mediated Vitellogenin gene knockdown, suggesting that Vitellogenin plays a role in this transport.

Our results provide initial evidence that trans-generational immune priming may operate at a colony-wide level in honey bees.

INTRODUCTION

Efficient anti-pathogen defense mechanisms support the survival of individuals. Various physiological and behavioral mechanisms have evolved to maximize organismal fitness as part of the immune system (Schmid-Hempel 2001). Immune responses must act quickly and target bacterial and fungal cells, as well as viruses. In vertebrates, a complex antibody-based immunological memory has evolved, which renders individuals and initially also their offspring immune to the same pathogen (Hasselquist and Nilsson 2009). Invertebrates, although lacking antibody-based immunological memory, can prime both themselves (Sadd and Schmid-Hempel 2006; Roth et al. 2009; Tidbury et al. 2011) and offspring (Little et al. 2003; Sadd et al. 2005; Moret 2006; Freitak et al. 2009; Tidbury et al. 2011; López et al. 2014b) against pathogens. The phenomenon is called trans-generational immune priming. Trans-generational immune priming occurs by females transferring ingested pathogen fragments to their developing eggs, where they elicit an immune response in the developing embryo (Freitak et al. 2014).

We found that Vitellogenin (Vg) may facilitate trans-generational immune priming in honey bees, as this protein transports immune elicitors into developing eggs (Salmela et al. 2015). Vg is an egg yolk precursor protein, essential in delivering nutrients into the eggs of most oviparous species. Moreover, it can bind pathogen-associated molecular patterns (PAMPs) (Li et al. 2008, 2009; Liu et al. 2009; Salmela et al. 2015). PAMPs include molecular motifs of bacterial and fungal cells like lipopolysaccharides and peptidoglycans. The ability of Vg to bind PAMPs, and the protein's transport into eggs, suggests a central role for Vg in immune priming.

Honey bee Vg has evolved to have important social functions (Amdam, Norberg, et al. 2003; Nelson et al. 2007). Colonies have a division of labor, with the queen responsible for reproduction and sterile workers responsible for colony maintenance. Workers further undergo an age-associated behavioral maturation that depends on titers of Vg in their blood (haemolymph) (Amdam and Omholt 2003; Guidugli, Nascimento, et al. 2005; Nelson et al. 2007). Young workers remain in the nest as nurses to clean cells, rear brood, and feed the queen, and have high titers of Vg that peak around age 5-15 days (Fluri et al. 1982; Engels et al. 1990). A sharp drop in Vg titers and a concomitant increase in juvenile hormone titers prompt workers to transition into foragers and leave the nest to collect nectar and pollen (Amdam and Omholt 2003; Nelson et al. 2007). Additionally, Vg is used by the nurse bees as an amino acid donor for producing royal jelly (Amdam, Norberg, et al. 2003), a protein-rich food synthesized in their hypopharyngeal (head) glands that they feed orally to the queen and young larvae (Snodgrass 1956). The hypopharyngeal glands express Vg receptors, presumably to allow uptake of Vg circulating in the haemolymph (Guidugli-Lazzarini et al. 2008).

The immune elicitors used to initiate trans-generational immune priming may come from environmental pathogens encountered and ingested by a female. However, honey bee queens have limited exposure to environmental pathogens since they reside entirely in the nest except for their mating flight (or flights) early in life and possible swarming flight later in life (Michener 1969b). For trans-generational immune priming to be effective in honey bees, a queen must prime her offspring against pathogens that her workers encounter. As we suggested previously (Salmela et al. 2015), queen exposure may occur via consuming contaminated food. As queens feed exclusively on worker-

produced royal jelly (Haydak 1970), this potential pathway would require workers to transfer ingested pathogen fragments to their hypopharyngeal glands. Here, we test if that requirement can be met by observing whether bacterial fragments fed to workers are transferred to their hypopharyngeal glands.

In this study, we fed heat-killed fluorescently-labelled *Escherichia coli* particles to worker honey bees and used immunohistochemistry to localize Vg and bacterial particles in three tissues: the midgut, the fat body, and the hypopharyngeal glands. The midgut is where ingested pathogen cells are broken down and potentially absorbed through the epithelium (Buchon et al. 2013). The fat body regulates the metabolic and immunological state of the organism and plays a central role in immune response (Bulet and Stöcklin 2005; Stokes et al. 2015). It is also the primary site of Vg synthesis and storage (Pan et al. 1969; Isaac and Bownes 1982; Raikhel and Lea 1983; Bownes 1986). We contrasted control workers and RNA interference (RNAi)-mediated Vg-knockdown workers to examine what role Vg plays in this pathway. We also measured how Vg RNAi affected expression of two other control genes that either share similar sequence homology or function as Vg: *Vitellogenin-like-C* (Salmela et al. 2016) and *apolipophorin-III* (Whitten et al. 2004a; Weers and Ryan 2006), respectively. These controls are used to validate the specificity of the double-stranded RNA (dsRNA) used in the RNAi protocol, and to ensure that any difference between treatments in *E. coli* tissue localization is due to Vg-knockdown and not due to inadvertent effects to other genes that share similar immunological functions as Vg.

We found that *E. coli* particles were transferred to the hypopharyngeal glands of control workers, but not to the glands of Vg-knockdown workers. This observation

confirms that ingested pathogen fragments can be transported to the production-site of royal jelly, and that Vg plays a role in this process in worker bees. As an additional finding, we detected Vg protein immunoreactivity in the midgut and hypopharyngeal glands of both controls and Vg-knockdowns. This result suggests that these organs can maintain Vg stores after the rate of *de novo* synthesis has been suppressed by RNAi in the fat body.

METHODS

Bees

Stock colonies of honey bees were maintained at the Arizona State University (ASU) Bee Research Facility in Mesa, Arizona. Frames of sealed brood from two hives were placed overnight in a 34°C incubator with 80% relative humidity. Newly emerged workers (up to 24 h old) were then secured to a wax-filled dissecting dish with two crossed needles and immobilized in a refrigerator. Following oft-used protocols first established in Amdam et al. 2003b, bees were next subjected to one of three injection treatments: i) a 1 µl injection of 10 ng/µl double-stranded RNA (dsRNA) of the *vitellogenin* (*vg*) gene to achieve RNAi (N = 89); ii) a 1 µl sham injection of nuclease-free water (vehicle, Ambion #AM9938) to serve as an injection control (N = 85); or iii) no injection to serve as a handling control (N=139). For injection controls, injecting the vehicle is a thoroughly established procedure that is frequently performed and widely accepted (e.g., (Amdam, Simões, et al. 2003; Guidugli, Nascimento, et al. 2005; Seehuus et al. 2006; Antonio et al. 2008; Ihle et al. 2010, 2015; Ament et al. 2011; Wheeler et al. 2013). Injections were made between for 5^{th} and 6^{th} abdominal segment using a 10 µl

Hamilton syringe with a G30 needle (BD). Bees that showed signs of bleeding after an injection were discarded and omitted from the experiment. The remaining bees were paint marked according to injection treatment and placed in two established host colonies at the main ASU campus in Tempe, Arizona. After 7 days we collected 158 bees (32 Vg dsRNA-injected, 41 sham-injected, 85 control-handled) and placed them into feeding cages (13 x 8 x 7 cm). The cages received 1 of 2 feeding treatments: either 30% sucrose in distilled water (control feeding), or the same food but with 0.5mg/mL of *E. coli* (K – 12 strain) BioParticles® with Texas Red® conjugate (Molecular Probes #E2863) (bacteria feeding). Food was provided in 10 mL aliquots via a 30 mL syringe, which was replenished daily. Cages were also provided with water and the bees remained caged in the 34°C incubator with 80% humidity for 48 h before being dissected.

Immunohistochemistry

Bees were anesthetized on ice and pinned to a wax-filled dissecting dish. The midgut, hypopharyngeal glands, and dorsal fat body were dissected and fixed separately in 4% paraformaldehyde for 48 h at 4°C, while the ventral fat body was prepared separately to confirm RNAi-mediated gene knockdown. Tissues were washed three times in 1X phosphate-buffered saline (PBS) before the midgut and hypopharyngeal glands were embedded in agarose gel and sectioned into 100 µm sections using a Leica VT1000s vibratome. The fat body was not embedded in agarose, but instead remained attached to the cuticle for the staining and washing procedures before being dissected from the cuticle and whole-mounted onto a slide. This was for practical purposes, as the fat body is a 1-cell thick sheet of tissue that does not require sectioning and would likely

be damaged with other processing protocols. All tissue samples were incubated overnight at 4°C in 1X PBS containing 1:1000 polyclonal rabbit-anti-Vg 1° antibodies (raised against 180 kDa honey bee vitellogenin; Pacific Immunology, Ramona, CA), 5% goat serum (Jackson ImmunoResearch #005-000-121), and 0.1% Triton X-100 (Sigma #T8787). Antibody specificity has been tested and confirmed in previous studies (Seehuus et al. 2007). Samples were then washed five times before being incubated at room temperature for 3 h in 1X PBS containing 1:1000 goat-anti-rabbit 2° antibodies conjugated with Alexa Fluor® 488 (Jackson ImmunoResearch #111-545-047), 5% goat serum, and 0.1% Triton X-100. The samples were washed three more times before being incubated for 15 min at room temperature in 1X PBS containing 1:30000 DAPI (Molecular Probes®D1306) and 0.1% Triton X-100. After a final five washes the samples were mounted in glycerol on glass slides. We also performed negative staining controls on all tissue-types to again confirm the specificity of the 1° antibodies and to look for any issues of autofluorescence that may occur in the range of wavelengths covered by Alexa Fluor® 488. For this, tissues were prepared and stained exactly as described above, except that in the first incubation the tissue was bathed in PBS with goat serum and triton but no 1° antibodies. Slides were imaged on a Leica TCS SP5 confocal microscope using a 40X oil-immersion objective. We imaged tissues from 22 bees, representing 3-4 individuals for all 6 treatments (3 injections treatments x 2 feeding treatments).

Preparation of dsRNA

Double-stranded RNA was prepared as previously described (Amdam, Simões, et al. 2003; Guidugli, Nascimento, et al. 2005; Amdam, Norberg, et al. 2006; Nelson et al. 2007; Antonio et al. 2008; Ihle et al. 2010, 2015). Primers were designed against the honey bee *vitellogenin* gene (GenBank number: AJ517411) cDNA clone *AP4a5* and fused with the T7 promotor sequence (underlined).

Forward: 5' - TAATACGACTCACTATAGGGCGAACGACTCGACCAACGACTT – 3'.

Reverse: 5' – TAATACGACTCACTATAGGGCGAAACGAAAGGAACGGTCAATTCC – 3'

PCR amplification was performed under normal conditions using Illustra™ PuReTaq Ready-to-Go™ PCR beads (GE Healthcare # 27955701) and the *AP4a5* clone as a template. PCR produced a product (excluding the T7 promotor) with a size of 504bp. The product was purified using a QIAquick PCR purification kit (Qiagen #28104) and the RNA was prepared using a RiboMAX™ Large Scale RNA Production System (Promega #P1300). The RNA was then extracted via phenol-chloroform extraction using TRIzol® LS reagent (Invitrogen #10296028) and re-suspended in nuclease-free water to a final concentration of 10 ng/µl.

RNAi knockdown validation

Gene expression was compared using fat body tissue, as this is the primary site of *vg* expression (Pan et al. 1969; Isaac and Bownes 1982; Raikhel and Lea 1983; Bownes

1986) and the target tissue for *vg* RNAi in honey bees (Amdam, Simões, et al. 2003; Guidugli, Nascimento, et al. 2005; Amdam, Norberg, et al. 2006; Nelson et al. 2007; Antonio et al. 2008; Ihle et al. 2010, 2015). RNA from the ventral fat body was extracted with TRIzol® LS reagent and re-suspended in nuclease-free water to a concentration of 100 ng/µl. A 1-step RT-qPCR was performed in triplicate with an ABI Prism 7500 (AppliedBiosystems) using a Quantitech SYBR® Green RT-PCR kit (Qiagen #204243). Actin was used as a housekeeping gene because it is stably expressed across honey bee tissues (Lourenço et al. 2008; Scharlaken et al. 2008) and is commonly used for knockdown validation (Amdam, Simões, et al. 2004; de Azevedo and Hartfelder 2008; Wang et al. 2012, 2013; Ihle et al. 2015). Data were analyzed using the ΔΔCT method (Schmittgen and Livak 2008). The effect of treatment on *vg* expression was determined using a one-way ANOVA, and differences between individual treatment groups were calculated with Tukey's honest significant difference test. Data were log-transformed to achieve normality. Analyses were performed in R (v3.3.2). Negative controls (no template) were used to rule out DNA contamination. Primer sequences used for RT-qPCR reactions were as follows:

Vg Forward: 5' – GTTGGAGAGCAACATGCAGA - 3'

Vg Reverse: 5' – TCGATCCATTCCTTGATGGT – 3'

Actin Forward: 5' – TGCCAACACTGTCCTTTCTG – 3'

Actin Reverse: 5' – AGAATTGACCCACCAATCCA – 3'

Effect of Vg RNAi on expression of control genes

RNAi can have unintended consequences by affecting expression of off-target genes, thereby confounding the interpretation of results. In these cases, it becomes unclear whether the phenomena observed are due to successful knockdown of the target gene, or altered expression of some off-target genes. For example, injected dsRNA is broken down into small interfering RNAs (siRNAs) that target transcripts with complementary sequences, but if the dsRNA lacks sufficient specificity then the resulting siRNAs can also knockdown other genes with a similar complementary sequences (Jackson and Linsley 2010). To validate the specificity of our *vg* dsRNA we measured expression of *vg-like-C*, as this *vg* homolog shares a similar sequence as *vg* and would likely be knocked down if our dsRNA were insufficiently specific (Morandin et al. 2014; Salmela et al. 2016). We can validate the specificity of our dsRNA if *vg-like-C* is equally expressed in individuals across all three injection and control treatments. Additionally, confounding results in RNAi can arise by the unintended activation of the immune system. Injected dsRNA activates the antiviral state in honey bees and alters expression of hundreds of genes (Flenniken and Andino 2013). If expression is altered in an immune gene with similar functions as *vg*, then any differing pattern of *E. coli* tissue localization observed between treatment groups could be due to this immune response to dsRNA and not due to the Vg-knockdown. *ApoLp-III* is a good candidate for an off-target immune-related control gene because of its functional similarity to *vg*: it performs several innate immunity functions including binding many PAMPs (Whitten et al. 2004a; Weers and Ryan 2006), it circulates in the haemolymph (Kawooya et al. 1984), it is expressed in the fat body (Cole and Wells 1990) and in the hypopharyngeal glands (Corby-Harris et al.

2016), and it is present in royal jelly (Han et al. 2011a). We contrasted *apoLp-III* expression in individuals injected with nuclease-free water (sham) and those injected with *vg* dsRNA to confirm that observed patterns of *E. coli* tissue localization were due to the effect of Vg-knockdown and not due to an immunological response to dsRNA. We measured *vg-like-C* and *apoLp-III* expression using the same procedure described above, and used honey bee primer sequences that were previously used and published by other (Lourenço et al. 2009; Salmela et al. 2016)

Vg-like-C Forward: 5' - AACGCGATCACATCAGTCGT - 3'

Vg-like-C Reverse: 5' - CGTGCCGCCAACAGATATGG - 3'

ApoLp-III Forward: 5' - TCTGACAAAGCTGCGAAATC - 3'

ApoLp-III Reverse: 5' - AGTTGCGGCAGTTTGAAGTT - 3'

RESULTS

Gene expression

Individuals injected with *vg* dsRNA (hereafter referred to as Vg-knockdown bees) showed successful *vg* gene expression knockdown compared with control-handled and sham-injected bees (Fig. 2.1A). Treatment groups differed significantly in *vg* expression (1-way ANOVA, $F_{2, 45} = 16.00$, $p = 5.6E\text{-}6$). Control-handled bees had significantly higher *vg* expression than both the sham-injected ($p=0.026$) and Vg-knockdown ($p<0.001$) bees, while sham-injected bees had significantly higher *vg* expression than Vg-knockdown bees ($p=0.014$) (Fig. 2.1A). Injection treatments did not affect expression of the first control gene, *vg-like-C*, as there was no significant difference in expression

among bees from any treatment (1-way ANOVA, $F_{2, 44} = 1.28$, P = 0.288)(Fig. 2.1B). For the second control gene, *apoLp-III*, there was no difference in expression between individuals injected with nuclease-free water (sham) and those injected with *vg* dsRNA (P = 0.151), nor between sham-injected and control-handled individuals (P = 0.830) (Fig. 2.1C). There were small but significant differences in expression between Vg-knockdown and control-handled bees (P = 0.040) (1-way ANOVA, $F_{2,44} = 3.47$, P = 0.040).

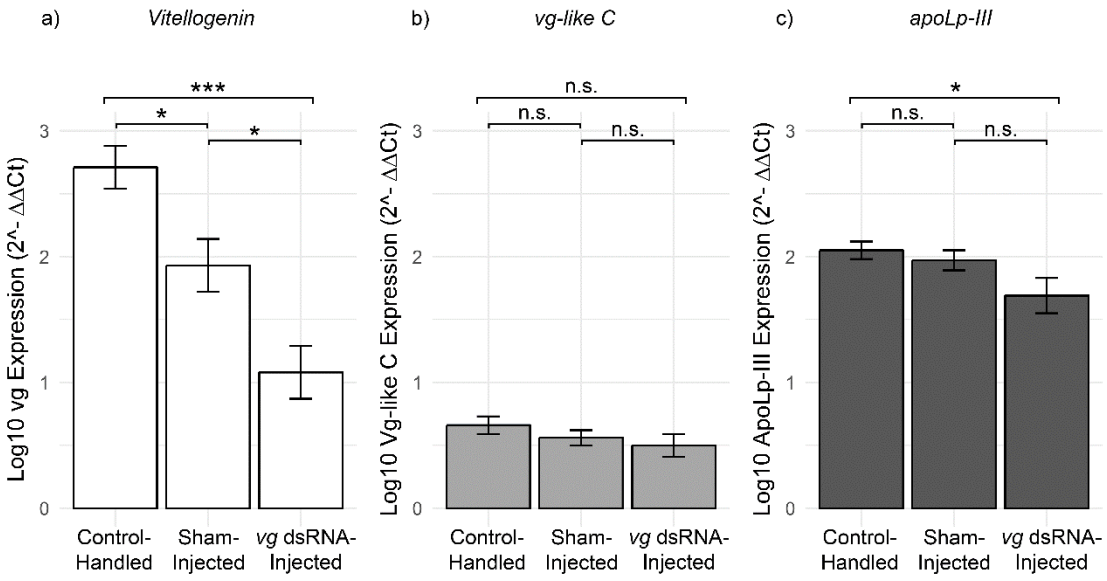

Fig. 2.1: Relative expression levels of three genes following *vg* dsRNA-injection, sham-injection, or control-handling. Bar heights represent the mean $\log_{10}$ gene expression level of N=16 individuals sampled, while error bars represent ± 1 standard error. Expression levels were determined using the ΔΔCt method. Data were normalized via log-transformation prior to analysis. Pairs of treatment groups with stars above were deemed significantly different from one another via a Tukey HSD test. A: Relative *vg* expression differed significantly between all treatment groups (1-way ANOVA: $F_{2,45} = 16$, $p<0.001$).

B: Relative *vg-like-C* expression did not differ between treatment groups (1-way ANOVA, $F_{2, 44} = 1.28$, $P = 0.288$). C: Relative *apoLp-III* expression significantly differed between treatment groups (1-way ANOVA, $F_{2, 44} = 3.47$, $P = 0.040$), but this difference was only observed between control-handled and *vg* dsRNA-injected individuals ($P = 0.040$).

Histology

Despite this effect of RNAi-mediated *vg* knockdown, bees from all treatments showed positive signal for Vg protein in their midguts (Fig. 2.2). Moreover, bees from all treatments that were fed fluorescently labelled *E. coli* particles showed uptake of label into their midgut. Bees subjected to the control feeding (Fig. 2.2M-N) did not show this signal. The Vg and *E. coli* particles appeared in close proximity or even overlapped in a few instances, but mostly the two signals were separate.

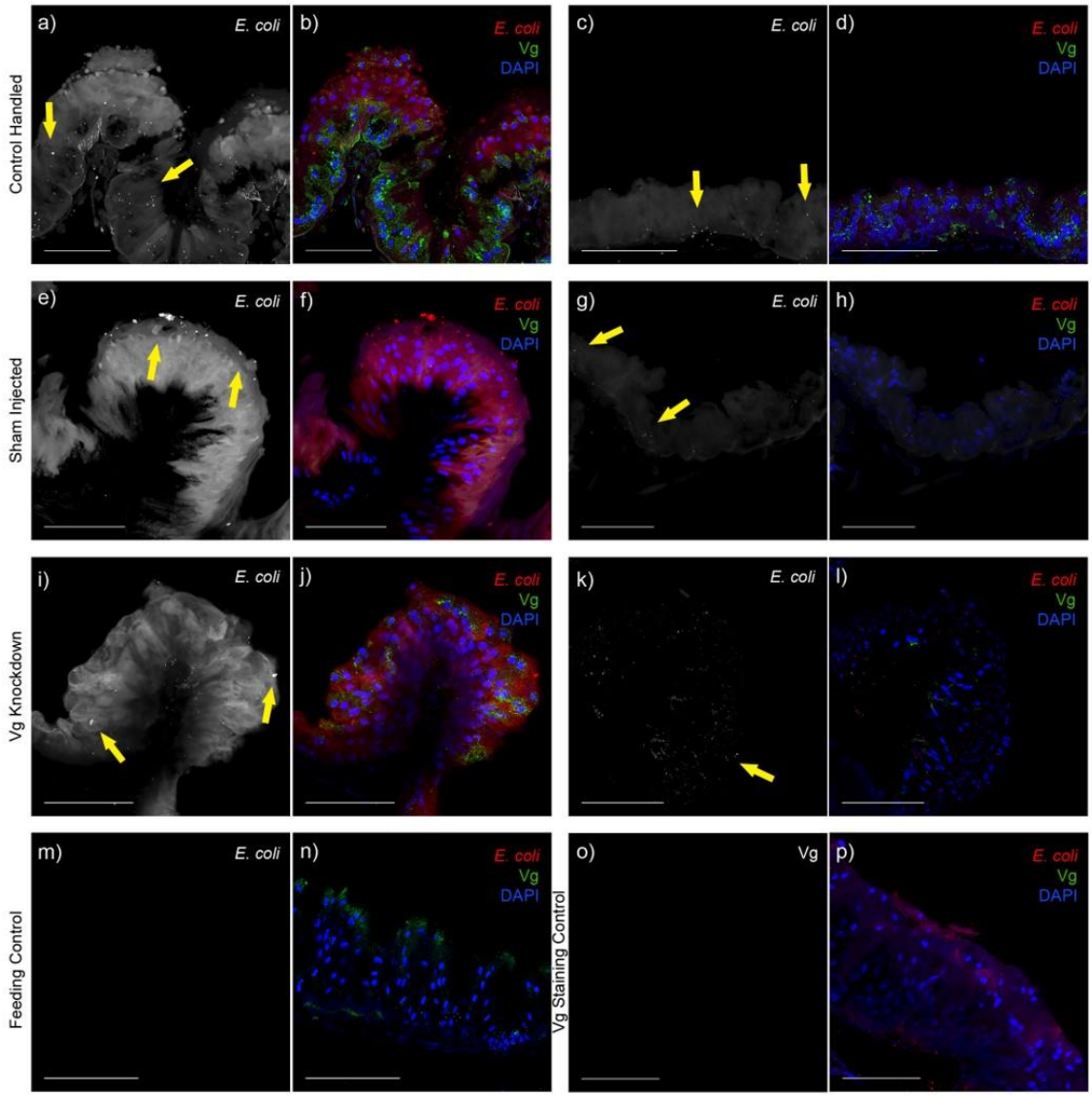

Fig. 2.2: Midguts of *E. coli*-fed and control-fed workers. The organs were sectioned longitudinally, and each set of images shows a collection of epithelial cells protruding into the lumen. Cell nuclei are stained blue with DAPI, *E. coli* particles are stained red with Texas Red, and Vitellogenin protein is stained green with Alexa 488. White scalebars are 100 µm in length. Yellow arrows indicate examples of positive *E. coli* signal. A-D: representative samples of control-handled bees fed with *E. coli*; E-H:

representative samples of sham-injected individuals fed with *E. coli*; I-L: representative samples of Vg-knockdown individuals fed with *E. coli*; M-N: representative sample of a control-handled individual fed with a control diet; O-P: representative sample of Vg control staining (incubated with Alexa 488, but without Vg 1° antibody).

As expected, fat body from control-handled and sham-injected bees showed positive Vg signal while that from Vg-knockdown bees showed little to no signal (Fig. 2.3). Honey bee fat body contains two cell types, trophocytes and oenocytes, with the former the site of Vg synthesis and the latter performing lipid metabolism functions (see section 4)). Trophocytes can be identified by their large irregularly-shaped nuclei, while oenocytes have smooth round nuclei. Positive signal for Vg was restricted to trophocytes and absent in oenocytes. Bees from all treatment groups that were fed *E. coli* particles showed positive fluorescent *E. coli* signal in their fat body. The particles appeared in both the trophocytes and oenocytes, located on the cell membrane or in the cytoplasm (Fig. 2.3). We observed small granular structures in the oenocyte cytoplasm that faintly autofluoresce in the same Texas Red channel as the *E. coli* , but these spots were substantially dimmer that the positive *E. coli* signal. These granules are likely fat droplets that are known to autofluoresce (Fletcher et al. 1973; Clokey and Jacobson 1986; Le et al. 2010) (see section 4).

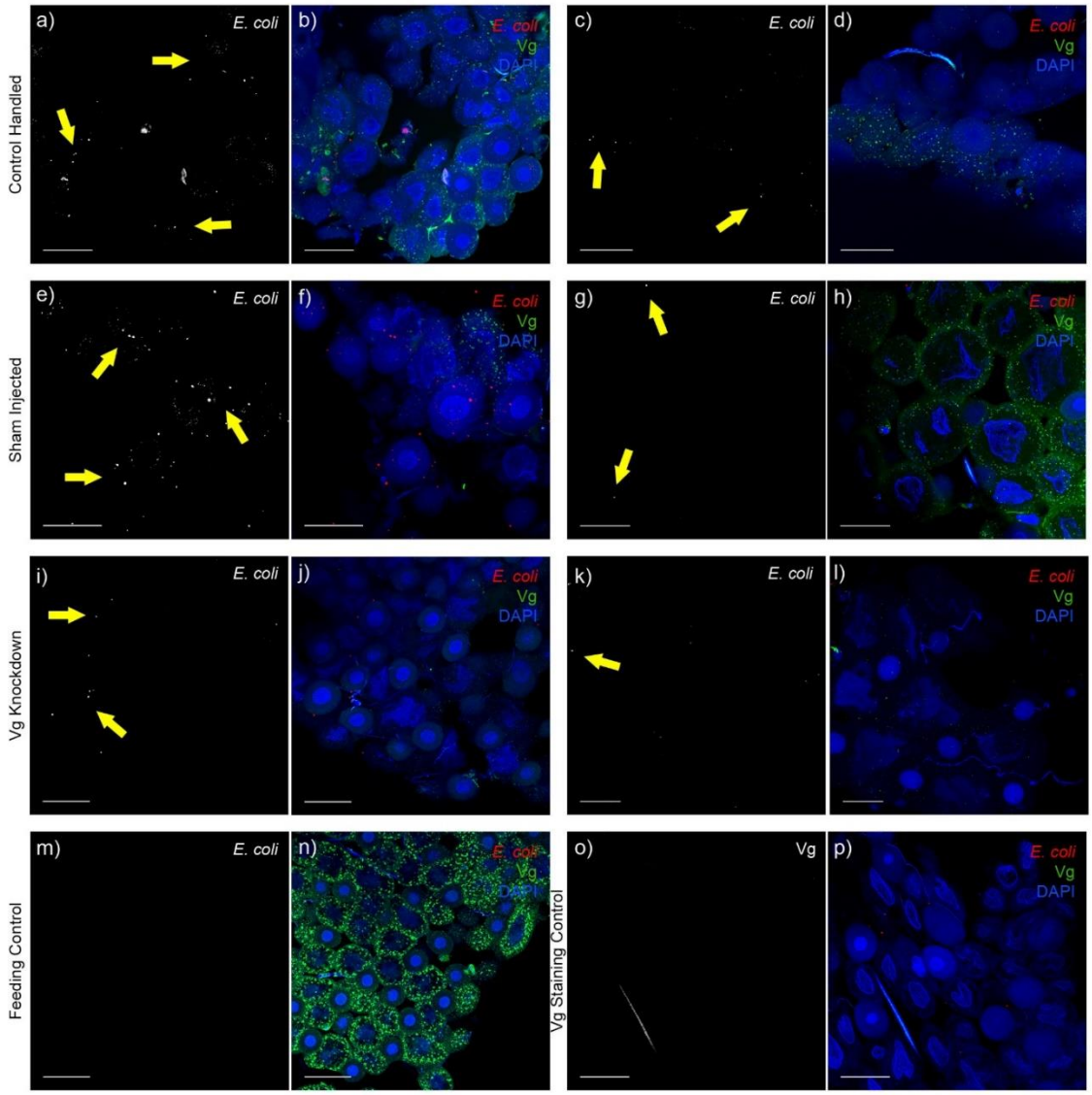

Fig. 2.3: Fat body cells of *E. coli*-fed and control-fed workers. Each sample shows a collection of the two cell-types found in the honey bee fat body, the trophocytes (with large, irregularly shaped nuclei) and the oenocytes (with rounded, smooth nuclei). Cell nuclei are stained blue with DAPI, *E. coli* particles are stained red with Texas Red, and Vitellogenin protein is stained green with Alexa 488. White scalebars are 50 μm in length. Yellow arrows indicate examples of positive *E. coli* signal. A-D: representative

samples of control-handled bees fed with *E. coli*; E-H: representative samples of sham-injected individuals fed with *E. coli*; I-L: representative samples of Vg-knockdown individuals fed with *E. coli*; M-N: representative sample of a control-handled individual fed with a control diet; O-P: representative sample of Vg control staining (incubated with Alexa 488, but without Vg 1° antibody).

Bees from all treatments showed positive signal for Vg in their hypopharyngeal glands, even Vg-knockdown bees. *E. coli* particles, however, were observed in the glands of control-handled and sham-injected bees, but not in the glands of Vg-knockdown bees (Fig. 2.4). The *E. coli* signal was mostly confined to the surface of the glands but was also observed in the collecting duct leading away from the glands (Fig. 2.4C-D). In several instances, *E. coli* and Vg appear to co-localize in the glands, which resolve as yellow pixels in the micrographs (e.g., Fig. 2.4 B, D, & F).

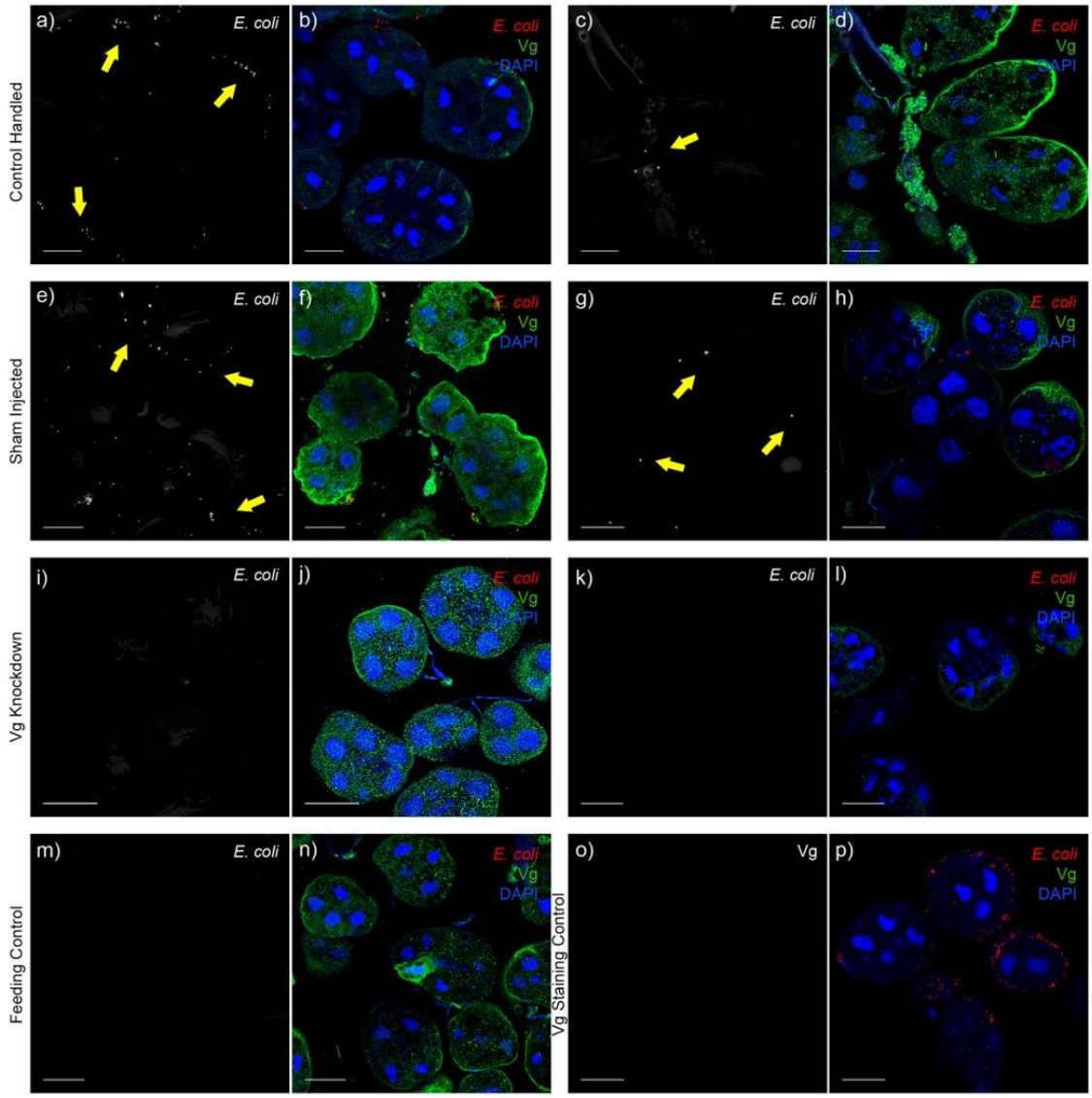

Fig. 2.4: The hypopharyngeal glands of *E. coli*-fed and control-fed workers. Within each acinus, cell nuclei are stained blue with DAPI, *E. coli* particles are stained red using Texas Red, and Vitellogenin protein is stained green with Alexa 488. White scalebars are 50µm in length. Individuals from the control handling and sham-injection treatments showed positive signal for *E. coli* in their glands (A-H) but individuals that received an injection of Vg dsRNA lacked *E. coli* signal in their glands (I-L). A-D: representative

samples of control-handled bees fed with *E. coli*; E-H: representative samples of sham-injected individuals fed with *E. coli*; I-L: representative samples of Vg-knockdown individuals fed with *E. coli*; M-N: representative sample of a control-handled individual fed with a control diet; O-P: representative sample of Vg control staining (incubated with Alexa 488, but without Vg 1° antibody).

DISCUSSION

This experiment has demonstrated two aspects of honey bee worker physiology not previously known. First, ingested bacterial fragments are transported to workers' hypopharyngeal glands, and second, this phenomenon is absent or greatly diminished when *vg* expression is down-regulated with RNAi (Fig. 2.4). Hypopharyngeal glands are an intriguing destination for bacterial fragments because of the important social function the glands play in food production by workers (Snodgrass 1956; Patel et al. 1960; Amdam, Norberg, et al. 2003). The hypopharyngeal glands absorb Vg in the haemolymph and appear to metabolize the protein into amino acids that are used during the synthesis of royal jelly (Amdam, Norberg, et al. 2003; Seehuus et al. 2007). Workers feed this protein-rich jelly to the queen and young larvae (Snodgrass 1956). In this way, the hypopharyngeal glands use Vg much the same way that some other insects use Vg to produce trophic eggs (Amdam, Norberg, et al. 2003), which are non-viable eggs laid by females to nourish young offspring or to boost overall reproductive output (Engels et al. 1990; Brian 2012). Converting Vg into royal jelly instead of trophic eggs is a novel adaptation of honey bees and may have allowed for more efficient brood rearing (Amdam, Norberg, et al. 2003). Our current work further suggests that this route of social

nutrient transfer incorporating immune elicitors could facilitate more efficient protection of honey bee brood via mechanisms of trans-generational immune priming.

The bacterial fragments that we observe at the glands of control-handled and sham-injected workers may be part of a trans-generational immune priming pathway whereby nestmates share immunological signals via food secretions. This could occur in a couple of ways. Firstly, bacterial fragments may be transported with Vg across the gland membrane and incorporated into royal jelly and then orally transferred directly to larvae or to the queen. As we suggested previously (Salmela et al. 2015), one potential source of pathogen exposure for queens is through ingesting contaminated food. It is well documented that female insects that ingest bacteria transport the pathogens from their midgut to their ovaries for trans-generational immune priming (Freitak et al. 2009; Knorr et al. 2015). The same pathway would operate for honey bee queens, except their food (with bacterial fragments) would come from worker-produced royal jelly. Alternatively, the bacterial fragments may remain along the gland membrane and initiate a molecular signal cascade that produces an immune response, perhaps influencing specific products secreted in the royal jelly. Royal jelly contains several components with immunological functions, including *defensin1* and major royal jelly protein 3 (Blum et al. 1959; Fujiwara et al. 1990a; Okamoto et al. 2003; Fontana, Mendes, et al. 2004; Vucevic et al. 2007; Romanelli Alessandra et al. 2011; Klaudiny et al. 2012; Sugiyama et al. 2012; Bucekova et al. 2017). Possibly, the presence of bound bacterial fragments on the gland surface can induce more of such immunomodulators to be incorporated into the jelly. New experiments are required to determine whether bacterial fragments that bind to honey bee

hypopharyngeal glands are incorporated into jelly and/or induce changes to jelly composition.

The second key finding of this study is that *E. coli* is present at the glands of control-handled and sham-injected bees, but not detected at glands of Vg-knockdown bees. This finding suggests that Vg may be an *E. coli* transporter in this pathway. That Vg binds to *E. coli* is well established (Shi et al. 2006; Li et al. 2008; Tong et al. 2010; Salmela et al. 2015), and we have shown previously that Vg is necessary and sufficient to transport *E. coli* into the ovaries (Salmela et al. 2015). Vg binds to a substantial array of different ligands, including both gram-negative and gram-positive bacteria (Shi et al. 2006; Li et al. 2009; Tong et al. 2010; Zhang et al. 2011; Salmela et al. 2015), fungal cells (Li et al. 2008), viruses (Huo et al. 2014; Whitfield et al. 2015), damaged and apoptotic host cells (Havukainen et al. 2013), reactive oxygen species (Nakamura et al. 1999; Seehuus et al. 2006; Havukainen et al. 2013), and zinc (Amdam, Simões, et al. 2004). Moreover, these binding abilities are conserved in Vitellogenins across disparate taxa, including insects (for review, see (Salmela and Sundström 2017)), fish (for review, see (Sun and Zhang 2015a), (Zhang et al. 2015)), and corals (Du et al. 2017), suggesting that a variety of binding abilities evolved early in Vg's evolutionary history. Strong binding ability stems from Vg's molecular structure, and includes a positively charged region in the protein's α-helical domain (Havukainen et al. 2013). In nurse-age bees like those used in this experiment, Vg makes up 30-50% of all haemolymph proteins (Engels and Fahrenhorst 1974; Fluri et al. 1982), which would provide ample binding-opportunity for *E. coli* that is digested and enters the haemolymph.

That *E. coli* is not detected at the glands of Vg-knockdown bees is likely due, at least partly, to a reduction in Vg molecules in the haemolymph available to bind *E. coli.* However, this is likely an incomplete explanation, as Vg titers are *reduced* but not eliminated by RNAi: In a previous study by our group using the same Vg-knockdown and control protocols, 10-day old Vg-knockdowns had an 80% reduction in median Vg titers compared with nurse controls (1.89 µg/µL and 9.24 µg/µL, respectively) (Nelson et al. 2007). An additional contributing factor to *E. coli* absence at the glands of Vg-knockdowns is likely Vg's lower binding affinity to *E. coli* than other ligands. As we showed previously, honey bee Vg has a lower binding affinity to *E. coli* than to the honey bee pathogen *Paenibaccilus larvae*, a gram-positive bacteria responsible for American foulbrood disease (Salmela et al. 2015). Vg also has lower binding affinity to lipopolysaccharide than to peptidoglycan (Salmela et al. 2015). Lipopolysaccharide comprises the outer membrane of gram-negative bacteria like *E. coli* (Beveridge 1999) and is absent in gram-positive bacteria, while peptidoglycan is present in both bacterial types but is substantially more abundant in gram-positive bacteria (Salton and Kim 1996). Vg may have higher binding affinity to peptidoglycan because it is common to both bacterial types and thus represents a more general bacteria marker. Alternatively, this higher binding affinity to peptidoglycan may result from selective pressures on Vg to combat specific gram-positive bacterial diseases such as American foulbrood and European foulbrood (*Melissococcus plutonius*). These diseases are not only deadly (Morse and Flottum 1997), but they are also the only widespread bacterial pathogens that target developing honey bee larvae (Evans and Schwarz 2011). In contrast, *E. coli* is not a honey bee pathogen. It was used as a model in this study because it is can be reliably

labeled. Thus, it is not expected that selection has acted on Vg to confer effective binding to this bacterium. Vg's lower binding affinity for *E. coli* coupled with reduced Vg abundance in the haemolymph after *vg* RNAi may explain why *E. coli* is not detected in the glands of Vg knockdowns. For control-handled and sham-injected bees, Vg is available in abundance, allowing it to still effectively transport the labeled bacterial particles to the glands.

Alternatively, *E. coli*'s lack of detection at glands of Vg-knockdown workers may reflect changes in worker physiology. Typically, Vg-knockdown results in workers prematurely transitioning from nurses to foragers (Nelson et al. 2007; Antonio et al. 2008). Foragers no longer need to produce royal jelly to feed larvae, so their hypopharyngeal glands begin to atrophy (Milojevic 1940; Huang and Robinson 1996; Amdam, Aase, et al. 2005). This process may result in fewer Vg receptors, or receptors capable of binding bacterial fragments, being present along the gland membrane. This altered membrane interface, with fewer transport- or binding-opportunities for *E. coli*, could explain why we do not see *E. coli* particles at the glands of Vg knockdowns. This alternative interpretation implies that there can still be sufficient Vg to transport labeled bacterial fragments to the glands of Vg knockdowns. Future molecular studies of the hypopharygeal glands are required to test this possibility.

Another alternative explanation is that *E. coli*'s presence or absence at the glands is not due to Vg-knockdown, per se, but instead due to off-target effects of injecting dsRNA. These off-target effects might arise from using insufficiently specific *vg* dsRNA that knocks down other genes with sequence homology to *vg*, or from an immunological response to dsRNA that alters expression of other genes which may facilitate *E. coli*

transport. However, these alternative explanations seem unlikely based on the data presented here. First, the *vg* dsRNA appears to be highly specific, as the homologous control gene *vg-like-C* showed no difference in expression among all treatment groups. This further reduces the likelihood that genes analogous to *vg* would be subject to off-target RNAi effects. Second, there was no difference in expression of *apoLp-III* between bees that received a sham injection (nuclease-free water) and those that received a *vg* dsRNA injection, suggesting that exposure to dsRNA per se does not alter expression of this functional control gene. *ApoLp-III* circulates in the haemolymph and can bind to the PAMPs of *E. coli* cell walls (Weers and Ryan 2006), and so could theoretically transport *E. coli* to the hypopharyngeal glands. However, *E. coli* is only visible at the glands of sham-injected individuals and not *vg* dsRNA-injected individuals, despite these treatment groups expressing similar levels of *apoLp-III.* Therefore, it is unlikely that *apoLp-III* is involved in *E. coli* transport to the glands. Finally, Vg and *E. coli* appear to co-localize in some instances at the glands, where the red pixels of *E. coli* and the green pixels of Vg overlap to resolve into yellow pixels (Fig. 2.4 B, D, & F). This may indicate the Vg-bacteria complex interacting with Vg receptors on the gland surface, and lends further support to the argument that Vg plays a central role in bacteria transport to the glands.

In addition to the main findings, this study also made several intriguing observations that warrant further research. While we were not able to determine the role of Vg in transporting labeled *E. coli* particles from gut lumen to hemocoel, Vg was observed in the epithelial cells lining the midgut interior of all the treatments (Fig. 2.2). These are the most predominant cells in the insect midgut and are responsible for digestion and absorption (Dow 1987; Buchon et al. 2013). They extend from the basal

lamina in towards the lumen and secrete digestive enzymes and the peritrophic membrane, a chitinous material that protects the cells from abrasion and damage from ingested food (Brandt et al. 1978; Lehane 1997; Hegedus et al. 2009; Lehane and Billingsley 2012). They also produce and secrete antimicrobial peptides (AMPs) and reactive oxygen species (ROS) (Ha et al. 2005; Buchon et al. 2009, 2013; Kumar et al. 2010) into the lumen in response to ingested pathogens. However, an excessive immune response damages the epithelial cells, which are quickly replaced by midgut stem cells that proliferate and differentiate (Micchelli and Perrimon 2006; Ward et al. 2008; Buchon et al. 2013). Given what we know about Vg's properties, there may be several functions it performs in this organ. For example, Vg may intercept and destroy pathogen cells in the midgut, owing to its ability to bind to many pathogens (Li et al. 2008, 2009; Liu et al. 2009; Salmela et al. 2015) and its bactericidal properties (Li et al. 2009; Tong et al. 2010; Zhang et al. 2011). It may also act as an opsonin when bound to pathogens and recruit other immunological factors to destroy the pathogen (Li et al. 2008, 2009; Liu et al. 2009; Zhang et al. 2011). Additionally, Vg is an antioxidant, and it has ability to recognize and bind to damaged host cells and protect them from further ROS damage (Havukainen et al. 2013). In this capacity, Vg may serve to protect epithelial cells from the host's own immunological arsenal and prolong their life until they can be replaced by the differentiating stem cells. Midgut cells express *vg* (Harwood & Amdam, unpublished) and more Vg is likely absorbed from the haemolymph, as the midgut expresses low levels of the Vg receptor (Guidugli-Lazzarini et al. 2008).

In the fat body, bees from all bacteria feeding treatments showed positive *E. coli* signal. Again, we were unable to link presence or absence of immune elicitors to Vg, as

we did not account for Vg circulating in the haemolymph, and high levels of Vg in fat body may be insensitive to RNAi treatments in some individuals. The fat body regulates the metabolic and immunological status of the organism by monitoring the haemolymph (Chapman 1998; Oliveira and Cruz-Landim 2006; Arrese and Soulages 2010). It lines the abdominal cuticle surrounding the digestive tract and is well-positioned to detect digested pathogens that enter the haemolymph. It stores and releases lipids, proteins, and carbohydrates (Chapman 1998, reviewed in de Oliveira and Cruz-Landim 2006), and plays a key role in immunity by producing many AMPs (Bulet and Stöcklin 2005; Stokes et al. 2015). Our data show *E. coli* particles to be associated with both of the two honey bee fat body cell types: trophocytes and oenocytes. Trophocytes possess large irregularly shaped nuclei and regulate metabolism (Chapman 1998), and are also the primary site of Vg synthesis (Pan et al. 1969; Isaac and Bownes 1982; Raikhel and Lea 1983; Bownes 1986). Oenocytes possess rounded nuclei and contain highly developed smooth endoplasmic reticulum in their cytoplasm (Martins et al. 2011; Martins and Romalho-Ortigao 2012). They perform functions in lipid metabolism (Gutierrez et al. 2007), and as such, contain lipid droplets (Gutierrez et al. 2007; Kühnlein 2011; Brasaemle and Wolins 2012; Makki et al. 2014). They are also involved in cuticle formation (Wigglesworth 1988), hormone production (Wicker-Thomas et al. 2009), and detoxification of xenobiotics (Lycett et al. 2006; Martins et al. 2011). In this latter function, they express many key immune genes, including cytochrome p450, dehydrogenase, catalase and lysosome P (Martins et al. 2011). Our results do not explain the role of effects of *E. coli* binding to the fat body cells in honey bees, but it may be interacting with pattern

recognition receptors as part of the innate immune response to trigger the production of AMPs and other defenses.

We also found Vg inside the hypopharyngeal glands, consistent with previous findings (Seehuus et al. 2007). A novel finding here is that Vg remains in the hypopharyngeal glands even after Vg knockdown (Fig. 2.4), and the same pattern is observed in the midgut (Fig. 2.2). The reason for this pattern remains unclear. The Vg receptors in both the hypopharyngeal glands and midgut may be very efficient at extracting Vg from the haemolymph, even after RNAi knocks down the Vg titer to a low level. Alternatively, Vg may only be slowly metabolized in the hypopharyngeal glands of Vg-knockdown workers, which are known to transition from nursing to foraging activities (Nelson et al. 2007; Antonio et al. 2008). Foragers do not synthesize royal jelly, and an associated reduction in the consumption of Vg could lead to a reduced turnover and prolonged presence of Vg in the glands.

Overall, this study demonstrates that ingested bacterial fragments are transported to the hypopharyngeal glands of worker honey bees, and that Vg likely plays a role in this transport. We cannot conclude whether bacterial fragments are secreted directly into jelly or whether they only bind to the gland surface to elicit an immune-related response by the glands. But either mechanism could result in trans-generational immune priming aimed at protecting the brood. This protection could occur by workers feeding bioactive components (bacteria fragments or immune-related molecules) to the queen, or potentially also via direct feeding of the brood, which receive jelly during the first days of larval life (Haydak 1970). Both mechanisms can be interpreted as an adaptation for a social insect colony to respond to pathogen threats in real-time.

CHAPTER 3

ROYAL JELLY AS A VEHICLE IN TRANSFERING BACTERIAL FRAGMENTS BETWEEN NESTMATES

Gyan Harwood, Heli Salmela, Dalial Freitak, Gro Amdam

ABSTRACT

Social immunity is a suite of behavioral and physiological traits that allow colony members to protect one another from pathogens and includes the oral transfer of immunological compounds between nestmates. In honey bees, royal jelly is a glandular secretion produced by a subset of workers that is fed to the queen and young larvae, and which contains many antimicrobial compounds. A related form of social immunity, transgenerational immune priming (TGIP), allows queens to transfer pathogen fragments into their developing eggs where they are recognized by the embryo's immune system and induce higher pathogen-resistance in the new offspring. These pathogen fragments are transported by vitellogenin (Vg), an egg-yolk precursor protein that is also used by nurses to synthesize royal jelly. Therefore, royal jelly may serve as a vehicle to transport pathogen fragments between nestmates. To investigate this, we recently showed that ingested bacteria are transported to nurses' jelly-producing glands, and here, we show that pathogen fragments are incorporated into the royal jelly. Moreover, we show that consuming pathogen cells induces higher levels an antimicrobial peptide found in royal jelly, *defensin-1*.

INTRODUCTION

Royal jelly has long fascinated biologists because of the key role it plays in caste development in honey bees. This glandular secretion is produced by a subset of the worker caste known as nurses and fed to the queen throughout the duration of her development and adult life, while worker-destined larvae are fed royal jelly for the first 3 days of their lives before being switched to a more pollen-based diet (Townsend and Lucas 1940; Johansson 1955; Haydak 1970). As analytical chemistry methods advanced, scientists began to study the nutritional components of royal jelly and found it contained many pathogen-killing compounds that protect the queen and young larvae from disease (Blum et al. 1959; Fujiwara et al. 1990a; Okamoto et al. 2003; Fontana, Mendes, et al. 2004; Vucevic et al. 2007; Romanelli Alessandra et al. 2011; Sugiyama et al. 2012). The transfer of anti-pathogenic compounds between nestmates is a form of *social immunity*, a suite of behavioral and physiological traits that help colony members protect one another from pathogens (Cremer et al. 2007). A related form of social immunity, transgenerational immune priming (TGIP), has also garnered much attention over the past decade (Sadd et al. 2005; Moret 2006; Freitak et al. 2009, 2014; Zanchi et al. 2011; López et al. 2014b; Knorr et al. 2015; Salmela et al. 2015). Here, female insects that survive a pathogen attack can transfer pathogen fragments to their eggs and produce offspring that are more disease-resistant. These pathogen fragments contain pathogen-associated molecular patterns (PAMPs), structural components found in the cell walls of non-animal cells, which trigger an immune response in the developing offspring. While TGIP has been demonstrated in honey bees, the studies relied on *injecting* queens with an inoculum and this presented a problem: how would a queen be inoculated with

pathogens under natural conditions given that she feeds exclusively on royal jelly and has no opportunity to consume potentially-contaminated nectar and pollen collected by her foragers? A major discovery by our group hinted at a new possible pathway. Queens transfer pathogen fragments into their eggs using vitellogenin (Vg)(Salmela et al. 2015), an egg-yolk precursor protein that is also used by nurses to synthesize royal jelly (Amdam, Norberg, et al. 2003). Thus, honey bees may use royal jelly as a vehicle for transferring pathogen fragments between nurses and queens and larvae as part of a colony-level immune pathway. In our first step to elucidating this pathway, we showed that nurses that ingested bacteria were able to transfer it from their midgut to their jelly-producing glands, and that knocking down *vg* expression impeded this process (Harwood et al. 2019). Now, we seek to determine whether nurses that ingest bacterial pathogens can incorporate them into their royal jelly, and whether ingesting pathogens also increases the concentration of other antimicrobial components in the royal jelly.

Understanding immune pathways in honey bees is not only important for the field of organismal biology, but also for human food security. Honey bees are the premier insect pollinator in agriculture, adding at least $15 billion annually to the value of crops in the United States alone (USDA press release). However, high annual colony losses continue to plague American beekeepers (USDA honey bee health report 2019), owing to multiple stressors like pesticide exposure and poor nutrition. But bee pests and pathogens are also a major contributor and are often detected in colonies that have perished (USDA honey bee health report 2019). These pathogens include a multitude of bacteria, fungus, and viruses, many of which have no pharmacological treatments. Understanding how

social immunity mediates pathogen infections could allow for more comprehensive treatment regimes and help reduce the disease burden that contributes to colony losses.

It is important to understand how mechanisms of social immunity complement the immunological defenses present in individual bees. Honey bees have several layers of defense, starting with structural barriers like a water-tight cuticle that block pathogen entry (Moret and Moreau 2012). Pathogens that are ingested will end up in the midgut, the organ responsible for digestions and absorption, and here they face another physical barrier called the peritrophic matrix (Brandt et al. 1978; Lehane 1997; Hegedus et al. 2009). This is a chitinous substance secreted from the midgut epithelial cells that acts as a sieve to block out large particles like pathogen cells from being absorbed. If pathogens breach these physical barriers, they are detected by pathogen pattern recognition receptors that activate several cellular and humoral immune responses. Cellular defenses can include phagocytosis by hemocytes (Lavine and Strand 2002; Evans and Spivak 2010; Marringa et al. 2014), while humoral responses can include the production of melanin to encapsulate foreign particles (González-Santoyo and Córdoba-Aguilar 2012) and the production of antimicrobial peptides that directly kill pathogens (Bulet et al. 1999; Bulet and Stöcklin 2005; Evans et al. 2006). The melanization response is induced when bee venom serine protease (*Bi-VSP*) activates the phenoloxidase cascade that leads to the production of melanin (Choo et al. 2010). Antimicrobial peptides, such as *lysozyme* and *defensin-1*, are induced following activation of the Toll or Immune Deficiency (IMD) pathways (De Gregorio et al. 2002).

Despite a robust anti-pathogen arsenal, honey bees actually possess fewer immune related genes than solitary bee species (Evans et al. 2006), thanks largely to the evolution of social immunity. Royal jelly contains many of these enzymes and peptides just discussed, including *Bi-VSP*, *lysozyme*, and *defensin-1*, as well as *glucose oxidase*, which produces hydrogen peroxide that also kills pathogens (Fujiwara et al. 1990b; Fontana, Mendes, et al. 2004; Furusawa et al. 2008; Han et al. 2011b; Romanelli Alessandra et al. 2011; Fujita et al. 2012). This sterilized food is particularly important for the most vulnerable colony members, the young larvae, whose immune defenses are still developing. For example, young larvae have lower levels of pro-phenoloxidase (Chan and Foster 2008; Chan et al. 2009) (important in the melanization cascade) and antimicrobial peptides (Chan and Foster 2008), and their still-maturing midgut and peritrophic matrix leaves their midgut epithelium vulnerable to breach (Yue et al. 2008; Garcia-Gonzalez and Genersch 2013; Riessberger-Gallé et al. 2016). As a result, larvae are vulnerable to some diseases that are fairly innocuous for adults. This includes American foulbrood, caused by the spore-forming Gram-positive bacteria *Paenibacillus larvae* (Hansen and Brødsgaard 1999; Genersch 2010). American foulbrood infection is not only lethal for colonies, but costly for beekeepers because spores are resilient and remain viable for years, meaning beekeepers are forced to destroy infected hives. Thus, larvae stand to benefit from food supplemented with compounds that kill pathogens, or with pathogen particles that can elicit immune priming.

In this study, our aim was twofold. First, we sought to determine whether nurses that consume American foulbrood vegetative cells could incorporate pathogen fragments into their royal jelly, and second, to determine if this induces higher levels of immune

proteins found in royal jelly. To this end, we fed nurses with fluorescently-labelled and heat-killed *P. larvae* cells and examining the royal jelly they produced thereafter. We used fluorescent microscopy to confirm the presence of bacteria cell fragments in the royal jelly, and mass spectrometry to compare the proteomic profile of royal jelly from challenged and control colonies. In particular, we tested the hypothesis that royal jelly from challenged colonies would be higher in the immune proteins *glucose oxidase*, *bi-SVP*, *lysozyme,* and *defensin-1*.

METHODS

Culturing pathogenic bacteria

Vegetative cells of *P. larvae* were obtained from the Göteburg University culture collection and cultured under normal bacteria growing conditions. Briefly, cells were suspended in MYPG broth and kept at 37°C overnight in a shaking incubator, before being plated on MYPG agar and left to incubate for 1 week at 37°C. When enough bacteria were grown, vegetative cells were harvested and suspended in 1X phosphate buffered saline (PBS), before being heat-killed in an autoclave. The resulting non-viable bacteria cells were then conjugated to a fluorescent dye using pHrodo™ Red Phagocytosis Particle Labeling Kit (Invitrogen #a10026) following the manufacturer's instructions. Once dyed, *P. larvae* cells were suspended in 1X PBS to a concentration of 10 mg/mL and stored at 4°C until ready for use.

Bees and feeding experiment

Naturally-mated European honey bees were maintained in standard Langstroth hive boxes at the University of Helsinki on the Vikki campus in Finland. These standard Langstroth hives served as donors for smaller queenless colonies housed in mini mating nuc hives. To establish these smaller queenless colonies, a section of brood comb containing 1-day-old larvae and measuring 12 cm x 12 cm was excised from the donor hive and transferred into a mini mating nuc box. Roughly 100-200 nurses were also transferred from the same donor hive and placed in the mini mating nuc. Nurses were identified as those individuals that were seen entering brood cells and feeding larvae in the donor hive. In total, N = 6 small queenless colonies were established, each from a separate donor hive. The colonies were made queenless so as to stimulate nurses to produce more royal jelly in an effort to make a new queen (Sahinler and Kaftanoglu 1997). The nurses and young larvae were sealed in the mini mating nuc boxes for 3 days and provided with a 30% sucrose solution in a syringe-feeder suspended from the ceiling. N = 3 colonies received a control diet, while N = 3 had fluorescently-labeled *P. larvae* added to their sucrose to a final pathogen concentration of 0.6 mg/mL. Control and pathogen-laced diets were replenished with fresh food daily. After 3 days, newly deposited royal jelly was harvested from the brood combs of each colony, and divided into 1 of 2 workflows.

Fluorescent microscopy

Royal jelly from each comb was transferred into 200 µl reaction tubes. To estimate the volume of jelly obtained from each hive, we pre-filled 30 reaction tubes with

increasing volumes of water ranging from 1 to 30 µl and had two observers independently and blindly assess which water volume matched the royal jelly volume. In all 6 cases, each royal jelly sample received the same volume estimate from both observers. Next, each royal jelly sample was diluted with 1X PBS to a final concentration of 10% by volume and vortexed. Samples were then applied to a hemocytometer and examined under a Leica DM6000 fluorescent microscope. On the hemocytometer, the field of view is evenly divided into 9 sections, and we made observations on 5 sections in total: the 4 corner sections, and the center section. This allowed us to calculate the concentration of fluorescent particles per µl of royal jelly. We used the following formula, $C = (P \div [S \times V]) \times D$, where *C* is the concentration of fluorescent particles per µl of royal jelly, *P* is the total number of positive fluorescent particles observed in each sample, *S* is the total number of squares observed on the hemocytometer (N=5), *V* is the volume of each square observed on the hemocytometer (0.1 µl), and *D* is the dilution factor of the royal jelly (10x dilution). All samples were observed under identical microscope settings.

The camera was set at 16 bit resolution (Waters 2009), which yields a brightness intensity value ranging from 0 – 65535. To identify positive fluorescent signals in our pathogen-diet samples, we first needed to establish a baseline level of background noise, above which we could determine whether a signal was real fluorescence or merely an artefact. To do this, we examined the intensity values from each image of our control-diet samples, and in each we subtracted the minimum intensity value from the maximum intensity value to obtain a range of background noise (Waters 2009). We then compared all of our control-diet sample images to obtain a mean and standard deviation of the range

of background noise. Thereafter, we deemed any signal we observed in our pathogen-diet samples as being a positive fluorescent signal above background noise if it was at least 3 standard deviations ($p \leq 0.01$) above the mean background intensity range. The mean background intensity range was 329 ± 264 (mean ± s.d.), giving us a positive fluorescent intensity threshold of 1121.

Sample preparation for proteomics analysis

The royal jelly samples were diluted by addition of 150 µl 8 M urea and sonicated in a water bath. The protein concentration was adjusted to 2.1 mg/ml using the Bradford assay with final volume of 100 µl. Cystein bonds were reduced with 45mM dithiothreitol (#D0632 Sigma-Aldrich, USA) for 20 min at 37°C and alkylated with 0.1M iodoacetamide (#57670 Fluka, Sigma-Aldrich, USA) at room temperature. Samples were digested by adding 0.75 µg trypsin (Sequencing Grade Modified Trypsin, V5111, Promega) for overnight at 37°C. After digestion peptides were purified with C18 microspin columns (Harvard Apparatus) according to the manufacture's protocol. The dried peptides were reconstituted in 30 µl 0.1% trifluoroacetic acid (TFA) in 1% acetonitrile (ACN), buffer A.

Liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS)

The analysis was carried out on an EASY-nLC1000 (Thermo Fisher Scientific, Germany) connected to a Velos Pro-Orbitrap Elite hybrid mass spectrometer (Thermo Fisher Scientific, Germany) with nano electrospray ion source (Thermo Fisher Scientific, Germany). The LC-MS/MS samples were separated using a two-column setup consisting

of a 2 cm C18-Pepmap trap column (Thermo Fisher Scientific, Germany), followed by 15 cm C18-Pepmap analytical column (Thermo Fisher Scientific, Germany). The linear separation gradient consisted of 5% buffer B (0.1% TFA acid in 98% acetonitrile) in 5 min, 35% buffer B in 60 min, 80% buffer B in 5 min and 100% buffer B in 10 min at a flow rate of 0.3 µl/min. 6 ul of sample was injected per LC-MS/MS run and analyzed. Full MS scan was acquired with a resolution of 60 000 at normal mass range in the orbitrap analyzer and followed with CID –MS2 top 20 most intense precursor ions with in ion trap (energy 35). Data was acquired using LTQ Tune software.

Protein identification

In a shotgun proteomics approach, proteins are enzymatically digested into smaller peptides, and these peptides are then matched to proteins in an annotated genome. Some of these peptide sequences may be shared by several proteins, while some are unique to just a single protein in the proteome. Our analysis yielded a total of 496 protein hits containing at least 1 peptide spectrum match (PSM) and 1 unique peptide. To further rule out false positives we only considered proteins with at least 2 unique peptides as being reliable protein hits, as has been done in other royal jelly proteomics studies (Zhang et al. 2014; Hu, Bezabih, et al. 2019), and in part because experimental variation in sample preparation steps affects most greatly the least abundant proteins in mass-spectrometric samples (Zhang et al. 2010). Proteins were annotated to the *Amel_4.5* honey bee genome, and we used Uniprot and Genbank to confirm their identities and functions. We then compared the abundance of 4 immune-related proteins in control-diet

and pathogen-diet royal jelly: *Glucose oxidase*, *venom serine protease* (*Bi-VSP*), *lysozyme*, and *defensin-1*.

Statistical analysis

To evaluate the abundance of immune-related proteins between control-diet and pathogen-diet royal jelly, we compared the peptide spectrum match values (PSM) of our chosen proteins. We used non-parametric Wilcoxon Rank Sum tests to address our limited sample size (N=3 per treatment). We used one-tailed comparisons as our planned hypotheses predicted that these select immune proteins would be upregulated in royal jelly from pathogen-fed colonies. Fold-changes in protein abundance were calculated as (PSM [pathogen] – PSM [control]) / PSM [control]. All analyses were performed in R (version 3.5.2).

RESULTS

Pathogen fragments in royal jelly

We found that royal jelly samples from all pathogen-diet colonies contained particles that fluoresced at significantly higher intensity than our established background fluorescent threshold, suggesting that the bacterial fragments were incorporated into the royal jelly. No control-diet royal jelly samples contained particles that surpassed this threshold (Fig. 3.1). On average, pathogen-diet colonies had 88.3 ± 23.6 (mean ± s.d.) fluorescent particles per µl of royal jelly (N=3) while control fed colonies had 0 fluorescent particles per µl of royal jelly (N=3) (Fig. 3.1).

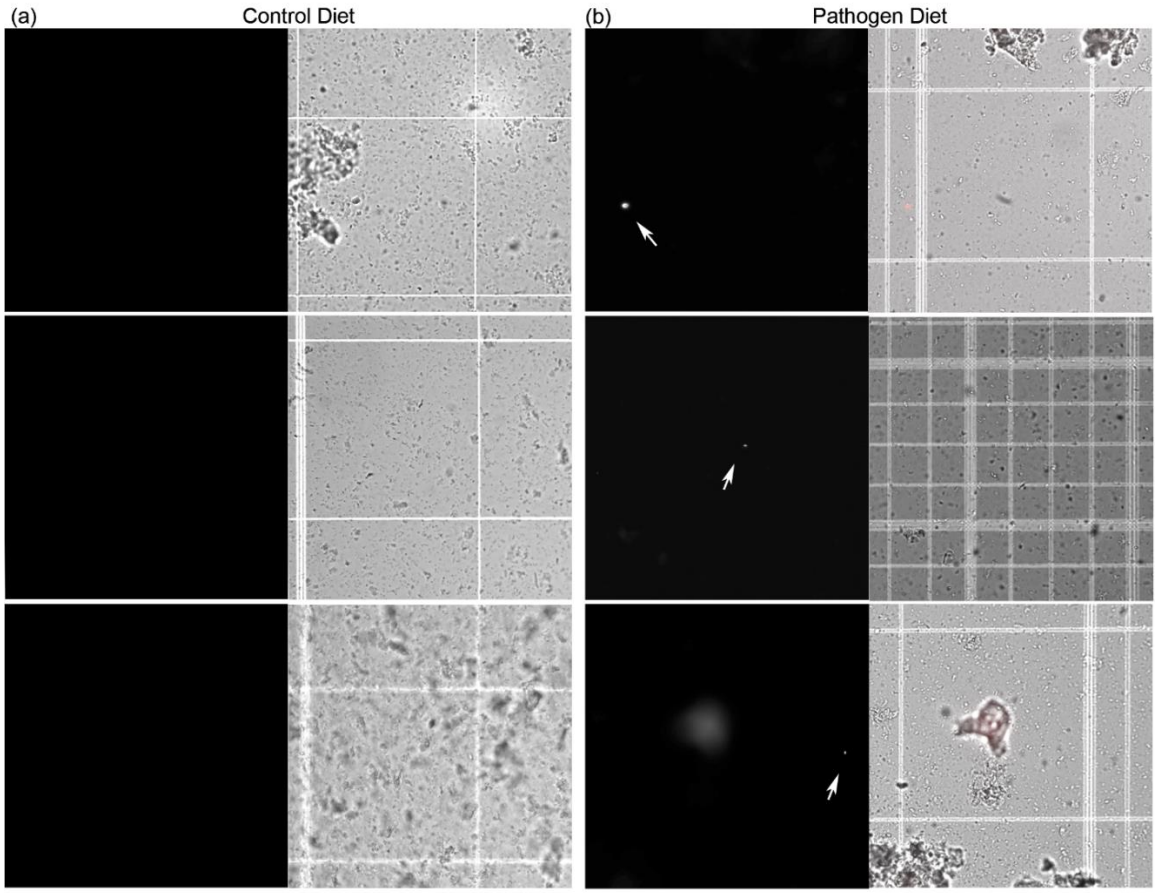

Fig. 3.1: Fluorescent and brightfield micrographs of royal jelly samples from control-diet (a) and pathogen-diet (b) treatments observed on a hemocytometer. All samples were prepared together using identical protocols and observed under identical microscope settings using a Leica DM6000. Dark images in the left-hand columns are taken with a TRITC laser to identify fluorescence in the Texas Red wavelength, while images on the right-hand columns are brightfield images of the same area to show the deposition of royal jelly on the hemocytometer. Positive fluorescent spots (white arrows) were only observed in the pathogen-diet samples (b) and not in the control-diet samples (a).

Proteomics

We found a total of 496 protein hits, and of these, 44 had at least 2 unique peptides assigned to them (Table 3.1). The most abundant proteins in all samples were major royal jelly proteins, as was expected. In general, the proteins we observe in our samples are quite consistent with other recent proteomics studies on royal jelly (Furusawa et al. 2008; Han et al. 2011b; Fujita et al. 2012, 2013; Zhang et al. 2014; Hu, Bezabih, et al. 2019), but there are a few proteins identified in this study that have not been reported elsewhere. These include: *Transferrin*, an iron transport molecule (Kucharski and Maleszka 2003) that is also found in honey bee venom gland tissue (Peiren et al. 2008); *Serpin-5*, a serine protease inhibitor known to regulate proPO activation in other insects (Li, Ma, et al. 2016); *Artichoke*, a chemosensory protein with functions in larval locomotion in *Drosophila melanogaster* (Andrés et al. 2014); and *NPC2-like*, a cholesterol transporter with implied alloparental functions in honey bees (Thompson et al. 2006). We also found 3 additional proteins in our pathogen-diet royal jelly samples: *Glutathione peroxidase*, *Peptidyl-prolyl cis-trans isomerase*, and an uncharacterized protein (LOC725202). These proteins were also found in our control-diet samples, but they failed to meet our criteria of having 2 unique peptides (Table 3.1).

Table 3.1: List of proteins found in royal jelly of control-diet and pathogen-diet samples. Columns for control-diet and pathogen-diet peptide spectrum match values (PSM) display the mean across3 samples, ± 1 standard deviation. Fold change is the relative change in pathogen-diet PSM compared to control-diet PSM ([pathogen – control] / control). Columns for the number of unique peptides show the mean across 3 samples.

Accession	Control-Diet PSM	Pathogen-Diet PSM	Fold Change	Control-Diet Unique Peptides	Pathogen-Diet Unique Peptides	Protein Name
GB55205	1115.3 [±227.1]	905.7 [±99.4]	-0.19	39	34.33	Major Royal Jelly Protein 1
GB55204	526.7 [±147.3]	441.7 [±40.3]	-0.16	23.33	20.33	Major Royal Jelly Protein 3
GB55212	304.3 [±77.5]	257.7 [±43.4]	-0.15	32.67	30	Major Royal Jelly Protein 2 precursor
GB55206	170.7 [±62.9]	238.3 [±37.5]	0.4	27	26.67	Major Royal Jelly Protein 4
GB45796	293.3 [±124.1]	213.0 [±14.2]	-0.27	2.67	2.33	Uncharacterized protein (LOC727045)
GB44549	169.7 [±11.3]	158.7 [±9.0]	-0.06	30	29	Glucose oxidase
GB55208	160.0 [±25.2]	134.0 [±10.0]	-0.16	18	15.67	Major Royal Jelly Protein 5 precursor
GB55213	134.0 [±20.9]	125.0 [±6.5]	-0.07	25	23.33	Major Royal Jelly Protein 7 precursor
GB43247	120.0 [±29.2]	99.0 [±50.1]	-0.18	31.33	27.33	Alpha-glucosidase
GB41428	48.0 [±17.7]	80.7 [±6.9]	0.68	3.67	3.33	Defensin-1
GB53578	44.0 [±9.4]	54.3 [±7.7]	0.23	19.67	20	Glucosylceramidase
GB55207	52.0 [±5.7]	52.0 [±5.4]	0	11	10.67	Major Royal Jelly Protein 6 precursor
GB55209	42.0 [±5.7]	41.7 [±6.8]	-0.01	8.67	9	Uncharacterized protein
GB43908	49.7 [±2.4]	41.3 [±3.3]	-0.17	40.67	32.33	Aldhedyde oxidase / Xanthine dehydrogenase
GB44548	37.3 [±5.0]	38.3 [±6.1]	0.03	20.33	19.33	Glucose dehydrogenase isoform X1

GB48020	15.3 [±1.2]	22.0 [±8.8]	0.43	10.33	8.67	Uncharacterized protein
GB48969	13.7 [±0.9]	19.3 [±6.2]	0.41	5.67	5.67	Uncharacterized protein (LOC408608)
GB50226	20.3 [±3.3]	18.3 [±3.3]	-0.1	17.67	16	Transferrin
GB41326	18.0 [±5.0]	18.0 [±2.2]	0	12	12.67	Venom acid phosphatase Acph1
GB49854	18.0 [±5.7]	18.0 [± 13.5]	0	13.67	11	Alpha amylase
GB41777	13.7 [±1.7]	17.3 [±2.9]	0.27	5.67	5.67	Uncharacterized protein
GB44223	11.0 [±5.4]	16.3 [±5.2]	0.48	10.33	15.33	Lysosomal alpha-mannosidase
GB51783	14.0 [±2.2]	15.7 [±1.7]	0.12	8.33	9.67	Carboxylpeptidase Q
GB55452	11.3 [±0.5]	15.3 [±1.7]	0.35	6.67	7.67	Apolipophorin III-like
GB53579	15.0 [±2.9]	15.0 [±2.4]	0	11.67	11.33	Putative glucosylceramidase 4
GB54611	10.7 [±3.9]	13.7 [±0.5]	0.28	8.67	10	Serpin-5 (serine protease inhibitor)
GB41776	9.7 [±0.5]	11.7 [±2.1]	0.21	4	3.67	Uncharacterized protein
GB43708	10.7 [±0.9]	9.3 [±1.7]	-0.13	7.67	6.67	Ferritin heavy polypeptide-like 17
GB49552	4.7 [±2.5]	8.7 [±3.7]	0.86	4.67	6.67	Venom serine protease Bi-VSP
GB51613	4.7 [±2.4]	8.3 [±1.9]	0.79	4	7	Uncharacterized protein (LOC408570)
GB53830	4.3 [±0.9]	7.0 [±0.8]	0.62	3.67	6	Protein Artichoke
GB50012	6.3 [±0.5]	7.0 [±2.2]	0.11	5.33	5	Uncharacterized protein (LOC726323)
GB42800	6.0 [±1.4]	6.7 [±0.9]	0.11	6	6	Protein Takeout-like
GB40759	6.7 [±0.9]	6.3 [±1.9]	-0.05	4.67	4.33	Icarapin-like
GB43731	8.0 [±1.4]	5.7 [±2.9]	-0.29	5.33	5.5	Ferritin subunit
GB44533	4.0 [±1.2]	5.0 [±0]	0.15	4.33	5	Uncharacterized protein (LOC408851)
GB40758	4.0 [±0.8]	5.0 [±0]	0.25	2	2	Icarapin-like precursor
GB47104	3.0 [±0.8]	4.0 [±0.8]	0.33	3	4	Lysozyme

GB50115	4.0 [±0]	4.0 [±0.8]	0		2.33	Uncharacterized protein (LOC725202)
GB48634	2.7 [±0.9]	3.3 [±0.5]	0.25		2.67	Glutathione peroxidase
GB43823	6.0 [±1.6]	3.3 [±1.2]	-0.44	4	2.33	Chemosensory Protein 1 precursor
GB44564	4.0 [±1.4]	3.0 [±0.8]	-0.25	4	3	Protein NPC2-like
GB55451	3.0 [±0]	2.7 [±0.5]	-0.11	3	2.67	Uncharacterized protein
GB46652	2.3 [±0.5]	2.7 [±0.5]	0.14		2.67	Peptidyl-prolyl cis-trans isomerase

When comparing control-diet and pathogen-diet samples, we observed very little change in relative protein abundance within the royal jelly (Table 3.1). Among our focal immune proteins, we found no significant difference in protein levels between treatments for *glucose oxidase* (*U* = 7, P = 0.9), *bee venom serine protease* (*Bi-VSP*) (*U* = 1.5, P = 0.134) or *lysozyme* (*U* = 2, P = 0.184). However, we did find significantly more of the antimicrobial peptide *defensin-1* in pathogen-diet samples than control-diet samples (*U* = 0, P = 0.05) (Fig. 3.2). It had a fold change of +0.68.

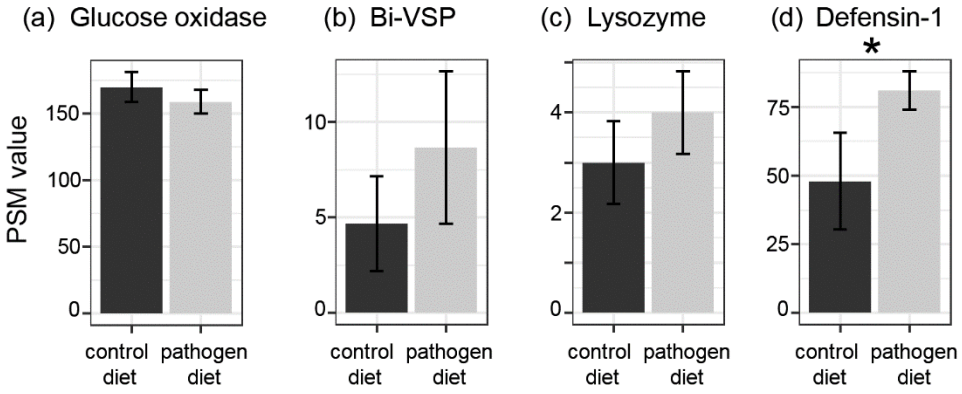

Fig. 3.2: The abundance of select immune proteins in control-diet and pathogen-diet royal jelly samples. Bar heights represent the mean peptide spectrum match values over three samples, with error bars denoting ± 1 s.d.. Of the 4 proteins compared, only the antimicrobial peptide *defensin-1* was significantly upregulated in pathogen-diet samples (d).

DISCUSSION

This study shows that nurses that ingest cells of the pathogen *P. larvae* appear to incorporate fragments of the cells into the royal jelly they produce (Fig. 3.1). We have shown recently that worker-ingested bacteria are transported to the hypopharyngeal glands, the site of royal jelly synthesis (Harwood et al. 2019), and here we show that they are incorporated into the royal jelly. This finding represents on overlooked social immunity pathway that can allow nestmates to share immunological memory. Our results also suggest that royal jelly composition may be sensitive to foreign material circulating systemically in the hemolymph. In a similar recent study, researchers found that nurses fed with double-stranded RNA (dsRNA) were able to incorporate these molecules into the royal jelly and transfer them to larvae, where the dsRNA remained biologically active (Maori et al. 2019). dsRNA activates the antiviral RNA-interference (RNAi) pathway that elicits sequence-specific gene silencing, and synthetic dsRNA is frequently used for targeted gene knockdown studies (Fire et al. 1998). Further still, another study showed that nurses fed with pesticides can also incorporate trace amounts of these agrochemicals into their royal jelly (Böhme et al. 2018), although the emphasis here is how well the hypopharyngeal glands prevent systemic pesticides from entering the royal jelly.

Nevertheless, these results indicate that royal jelly composition may be sensitive to the condition of nurses. These findings imply that royal jelly can serve as a conduit for sharing immune elicitors like pathogen fragments, dsRNA, and other foreign substances between colony members.

The incorporation of pathogen fragments into royal jelly suggests that trans-generational immune priming may operate at the colony-level in honey bees. Unlike females from solitary species, honey bee queens do not normally venture outside their nest (Maeterlinck 1901), nor do they ingest food collected from their environment (Haydak 1970), so they have fewer opportunities to accumulate immune elicitors from pathogens that currently threaten the colony. Pathogen fragments in royal jelly may help queens inoculate her offspring against a wider array of immune elicitors, and hence protect them from pathogens they are likely to encounter as adults. This colony-level pathway may be particularly suited to pathogens like American foulbrood, since larvae are highly susceptible to the disease but adult workers are immune. Larvae only need to consume as few as 10 American foulbrood spores for an infection to set in (Shimanuki 1997; Brødsgaard et al. 1998), but spores fail to germinate in the digestive tract of adults (Wilson 1971b). Nurses may also be able to directly inoculate larvae as well with royal jelly, meaning this pathway between nestmates would operate both vertically between offspring and parent, and horizontally between siblings.

This study also showed that exposing nurses to American foulbrood does not induce large changes in the proteomic profile of the royal jelly they produce. Between samples from control-diet and pathogen-diet colonies, there is minimal change in the relative abundance of most proteins, including immunological proteins like *glucose*

oxidase, *venom serine protease*, and *lysozyme* (Table 3.1). However, the antimicrobial peptide *defensin-1* appears to be an exception, with a modest relative fold change of +0.68 in royal jelly from pathogen-diet colonies compared to control-diet colonies. This finding raises two immediate questions. First, how does exposure to a pathogen lead to nurses producing elevated levels of *defensin-1* in the royal jelly? One possibility is that *P. larvae* fragments bind to pathogen pattern recognition receptors in the Toll pathway (for Gram positive bacteria) and induce greater gene transcription of *defensin-1*, but the evidence for this is somewhat inconsistent. Studies have shown that adult workers exposed to bacteria, including American foulbrood, will increase expression of *defensin-1* (Casteels-Josson et al. 1994; Evans et al. 2006), but a recent study failed to observe significant upregulation *defensin-1* expression in nurse head tissue after exposure to American foulbrood (López-Uribe et al. 2017). Nevertheless, insect AMPs like *defensin-1* can be observed in hemolymph for weeks after pathogen exposure (Casteels 1998), even as transcription of such AMP genes subsides during that time (Uttenweiler-Joseph et al. 1998), so only measuring gene expression may miss important information about the availability of the final peptide products. Alternatively, pathogen exposure may upregulate the translation of *defensin-1* mRNA into peptides, or increase the rate of post-translational modifications used to activate the peptide (Casteels-Josson et al. 1994), but these will have to be studied further.

The second question raised by these results is whether a relative fold change of +0.68 in *defensin-1* is biologically relevant. Conventionally, many human proteomics studies opt for a +1.2 threshold for significant upregulation (Keenan et al. 2009; Serang et al. 2013), so *defensin-1* would not meet this criterion. However, even modest increases

in *defensin-1* concentration in royal jelly may improve larval resistance to American foulbrood. Not only has *defensin-1* been shown to directly inhibit American foulbrood (Bíliková et al. 2001; Bachanová et al. 2002b), but it is also effective against Gram positive bacteria at concentrations as low as low as 1 μM (Fujiwara et al. 1990a). Furthermore, several components of the innate immune system are known to work synergistically, whereby two components working together have greater pathogen inhibitory effects than the sum of each component working independently. For example, bacterial growth inhibition assays have shown that antimicrobial peptides and *lysozyme* work together to greatly enhance one another's effectiveness (Chalk et al. 1994). *Lysozyme* has also been shown to combat pathogens synergistically with *apolipophorin-III* (*ApoLp-III*) (Zdybicka-Barabas et al. 2013), a lipid-transporting protein with anti-pathogenic functions in insects (Whitten et al. 2004b; Kim and Jin 2015) that is found in royal jelly (Furusawa et al. 2008; Han et al. 2011b; Fujita et al. 2013; Zhang et al. 2014). Finally, young honey bee larvae's maturing immune systems may struggle to produce their own *defensin-1* when attacked by American foulbrood, as studies have shown that such a challenge induces little to no *defensin-1* expression in larvae (8,11,66,67, although see 68 for exception). Thus, additional *defensin-1* from exposed nurses may supplement the larvae's dearth of this crucial antimicrobial peptide. Taken together, these findings suggest that even modest increases in *defensin-1* in royal jelly may be biologically relevant in protecting the colony's most vulnerable members.

This study also further exemplified critical role that the protein Vg plays in honey bee immunity. Owing to its role as a pathogen pattern-recognition receptor, which has been documented in a broad range of animal taxa (Li et al. 2008, 2009; Liu et al. 2009;

Garcia et al. 2010; Zhang et al. 2011; Salmela et al. 2015; Du et al. 2017; Knight 2019), as well as its role in egg formation, Vg has been identified as the carrier protein that transports immune elicitors into the queen's eggs (Salmela et al. 2015). Likewise in nurses, recent findings suggest that Vg plays a key role in transporting ingested pathogens to the hypopharyngeal glands (Harwood et al. 2019). Vg is taken up by the nurses' hypopharyngeal glands to be used as an amino acid donor in the production of royal jelly (Amdam, Norberg, et al. 2003), and it's possible that Vg transports bacteria into the glands to be incorporated into this jelly. Thus, Vg may play a role in all stages of colony-level TGIP in honey bees, from first inoculation and incorporation into the royal jelly, to delivery to the queen or larvae, and finally in the transport into the queen's ovaries.

This study has shown that royal jelly composition can be altered by a pathogen challenge, both in the corporation of pathogen fragments and through increased levels of the AMP *defensin-1*. The ability to transfer pathogen fragments in royal jelly means that trans-generational immune priming may function at a colony-wide level in honey bees. In addition to a vertical transfer of immune elicitors from a reproductive female to her offspring, honey bees may have a more complex pathway that goes from adult offspring (nurses) to reproductive female to future offspring, as well as horizontal transmission from adult offspring to immature offspring. The benefit here would be that larvae are inoculated with a larger repertoire of immune elicitors accumulated from the collective immunological experience of workers that gather resources from outside the nest, making larvae more resistant against pathogens they are likely to encounter when they themselves venture out of the nest.

CHAPTER 4

VITELLOGENIN EXPRESSION IN THE HONEY BEE MIDGUT

Gyan Harwood and Gro Amdam

ABSTRACT

The alimentary canal is an important organ that carries out many functions critical to insect physiology, including digesting and absorbing consumed nutrients, retaining water, eliciting immunological responses, and housing important endosymbionts. The midgut, in particular, is a compartment where digestion and absorption take place, and serves as a first point of contact between potential ingested pathogens and the insect immune system. Recently, we found the protein Vitellogenin (Vg) localized in midgut cells of some honey bee workers. Vg is an important egg-yolk precursor protein in nearly all oviparous animals, but it also has immunological functions shared across a wide-array of taxa. Additional and unexpected Vg functions have been characterized in honey bees, but none of these functions involve the midgut directly. Therefore, we sought to map out how Vg is localized and expressed in this organ across the two most common worker bee behavioral groups, namely nurses and foragers. We used quantitative reverse transcription PCR and immunohistochemistry to show that Vg is expressed in the midgut of both nurse and forager bees, and that the two groups have somewhat different protein localization in this organ. Our study provides a platform for building a more detailed understanding of the possible roles of Vg in insect midgut cells, and it adds to the current knowledge-base on this fascinating, multi-functional protein.

INTRODUCTION

The insect alimentary canal has been the subject of intense scientific inquiry. These efforts have led to the discovery and characterization of pluripotent stem cells (Hakim et al. 2001; Ohlstein and Spradling 2006; Ward et al. 2008; Buchon et al. 2009; Lehane and Billingsley 2012), a greater understanding of the molecular pathways activated in the alimentary canal's responses against ingested pathogens (Ha et al. 2005; Buchon et al. 2009; Kumar et al. 2010; Lehane and Billingsley 2012), and the classification and functional characterization of the endosymbionts living therein (Kikuchi 2009; Anderson et al. 2013; Engel and Moran 2013). The insect alimentary canal is compartmentalized into the foregut, midgut, and hindgut, with the foregut and hindgut originating from the ectodermal layer and thus being replaced at each molt (Snodgrass 2018). In most adult insects, the foregut contains the crop which is used to store food, while the hindgut contains the rectum which absorbs water and ions, as well as the ileum which houses many endosymbionts (Chapman 1998; Martinson et al. 2012). The midgut is located between the foregut and hindgut and originates from the endodermal layer, and so it is retained each time the insect molts (Snodgrass 2018). The midgut is the site of digestion and absorption and plays a critical role in insect immune physiology.

The insect midgut is composed primarily of elongated epithelial cells called enterocytes that protrude into the lumen and which carry out several important functions (Snodgrass 2018). They are critical for digesting and absorbing nutrients from consumed food as they produce and secrete digestive enzymes into the lumen, including proteinases, lipases and amylases (Lehane and Billingsley 2012). Enterocytes also secrete a chitinous

material called the peritrophic membrane (also called the peritrophic matrix or envelope) which protects the enterocytes from abrasion from ingested food, and which serves as a sieve to prevent pathogens, toxins, and large molecules from contacting the enterocytes (Brandt et al. 1978; Lehane 1997; Hegedus et al. 2009; Lehane and Billingsley 2012; Teixeira et al. 2015). Finally, enterocytes will produce and secrete an arsenal of immunological defenses when a pathogen is detected in the lumen. This arsenal includes antimicrobial peptides (AMPs) and reactive oxygen species (ROS) (Lehane et al. 1997; Nakajima et al. 2002; Ha et al. 2005; Buchon et al. 2009; Kumar et al. 2010; Dussaubat et al. 2012; Lehane and Billingsley 2012; Moreno-García et al. 2014).

Recently, we discovered that the protein Vitellogenin (Vg) is present in midgut cells of at least some honey bee workers (Harwood et al. 2019). Vg is an egg-yolk precursor protein synthesized in the insect fat body that is used by females to deliver nutrients to developing eggs in most oviparous animals (Pan et al. 1969; Isaac and Bownes 1982; Raikhel and Lea 1983; Bownes 1986). However, this reproductive protein also has functions in immunity and organismal health: It binds to and eliminates pathogenic bacterial and fungal cells by recognizing pathogen-associated molecular patterns (PAMPs)(Li et al. 2008, 2009; Liu et al. 2009; Zhang et al. 2011, 2015; Salmela et al. 2015); It protects host cells from oxidative stress by binding to and neutralizing ROS (Havukainen et al. 2013; Sun and Zhang 2015a; Salmela et al. 2016); It binds to damaged host cells and protects them from further injury by recognizing damage-associated molecular patterns (DAMPs)(Havukainen et al. 2013); and it transports the zinc required to maintain circulating hemocytes, which are cells of the innate immune system (Amdam, Simões, et al. 2004). None of these functions have been documented to

involve midgut tissue directly. Therefore, after finding Vg in worker midgut cells (Harwood et al. 2019), we sought to map out how Vg is expressed and localized in this organ. The rational for this work was to start building a foundation for understanding the possible roles of Vg in the gut of honey bee nurses and foragers.

In honey bees, workers are functionally sterile females that undergo an age-associated task specialization focusing on brood care (nursing) vs. foraging for resources outside the colony. Vg appears to play roles in the control of transitions between maternal behaviors involving brood care and foraging (Amdam, Norberg, et al. 2004): Nurse bees have high Vg titers while forager bees have low titers, and reducing Vg titers in young worker bees by means of RNA-interference (RNAi)-mediated *vg* gene knockdown triggers precocious foraging behavior. The transition from nursing to foraging is associated with sweeping changes in worker gene expression, behavior, and physiology, including changes in immunity (Fluri et al. 1977, 1982; Bedick et al. 2001; Amdam, Simões, et al. 2004; Schmid et al. 2008), gut function (Crailsheim et al. 1992), and dietary preferences (Brodschneider and Crailsheim 2010). To decouple behavior-associated patterns from age-associated patterns (foragers are typically older than nurses and thus behavior is naturally confounded by age), we compared nurses and foragers of the same age using single-cohort colonies (see Methods). Here, we tested whether *vg* is expressed the gut of both honey bee nurses and foragers using quantitative reverse transcription PCR (RT-qPCR), and we asked whether Vg localization in midgut tissue differs between the two behavioral groups using immunohistochemistry. These results allow us to make progress towards understanding Vg's function in the gut, and some of these possible functions are discussed here.

METHODS

Bees

Honey bee colonies were maintained at the Arizona State University Bee Lab in Mesa, Arizona. In order to obtain nurses and foragers of the same age, we used single-cohort colonies. Two such colonies were established in which newly emerged workers (<24h old) were housed with a queen and several frames of honey, pollen, and empty comb. After several days, many of these workers prematurely transition from nurses into foragers to meet the resource needs of the colony. On the seventh day after establishing the colonies, we paint marked foragers that we observed returning to the nest with pollen. On the fourteen day after establishing the colonies, we collected 25 marked foragers and an equal number of nurses. This ensured that all nurses and foragers had been performing their given task for a minimum of seven days. Nurse bees were selected by observing workers that entered brood cells to feed the larvae.

Gene Expression

A total of 16 nurses and 16 foragers were used for gene expression comparison. Midguts were dissected from anesthetized bees and RNA was extracted via phenol-extraction using TRIzol® reagent (Invitrogen #15596018). RNA was diluted to a concentration of 200 ng/µl and DNA was removed via DNase I treatment (Thermo Scientific™ #EN0525). One-step RT-qPCR was performed using a QuantiTect SYBR® Green RT-PCR kit (Qiagen #204245) and performed on an ABI Prism 7500 (Applied Biosystems). *Actin* was used as a reference gene, as it is stably expressed across honey bee tissues and commonly used in honey bee research (Lourenço et al. 2008; Scharlaken

et al. 2008). Each sample was prepared in duplicate for both *actin* and *vg*. The PCR program began with a reverse transcription step at 50°C for 30 minutes, followed by an activation step at 95° for 15 minutes. It then repeated 40 cycles of 94°C for 15 seconds, 52°C for 31 seconds, and 72°C for 32 seconds. Primer sequences used in the RT-qPCR reactions were as follows:

Vg Forward: 5' – GTTGGAGAGCAACATGCAGA - 3'

Vg Reverse: 5' – TCGATCCATTCCTTGATGGT – 3'

Actin Forward: 5' – TGCCAACACTGTCCTTTCTG – 3'

Actin Reverse: 5' – AGAATTGACCCACCAATCCA – 3'

Immunohistochemistry

Bees were anesthetized on ice and pinned to a dissecting tray. The midgut was dissected and cold 1X phosphate buffered saline (PBS) was applied during this procedure to maintain tissue freshness. The organ was fixed in 4% paraformaldehyde overnight at 4°C then washed 3 times for 10 minutes in cold 1X PBS before being embedded in an agarose medium. The embedded tissue was cut into 200 µm thick longitudinal sections using a Leica VT1000s vibratome. Our longitudinal sections offered us two views of midgut structure depending on where in the organ the section was made: Sections made through the center of the lumen provide a cross section where one can observe how the epithelial cells protrude into the gut lumen (Fig. 4.2), while sections made at the external surface of the organ allows one to observe how midgut cells are arranged in two-

dimensional space when attached to the basement membrane and surrounding muscles (Fig. 4.3). All samples received the same staining treatment. First, sections were blocked for 10 minutes with 5% goat serum (Jackson ImmunoResearch #005-000-121) in a solution of 1X PBS and 0.1% Triton X-100 (Sigma #T8787). We then added a 1:1000 dilution of polyclonal rabbit-anti-Vg primary antibodies (raised against 180 kDa honey bee vitellogenin; Pacific Immunology, Ramona, CA) and incubated the tissue overnight at 4^{o}C. The tissue was then washed 3 times for 10 minutes and incubated for an additional 3 hours at room temperature in 1X PBS with 5% goat serum and 0.1% Triton X-100, this time with a 1:1000 dilution of goat-anti-rabbit secondary antibodies tagged with Alexa Fluor® 488 (Jackson ImmunoResearch #111-545-047) and a 1:500 dilution of rhodamine phalloidin (Invitrogen™ #R415). Tissue was then washed 3 times for 10 minutes in cold 1X PBS and incubated an additional 15 minutes at room temperature in 1X PBS with 0.1% Triton X-100 and a 1:30000 dilution of DAPI (4′,6-diamidino-2-phenylindole. Invitrogen™ #D3571) before being washed for a final 5 times for 10 minutes in 1X PBS. The tissue was then mounted on slides and imaged on a Leica SP5 confocal microscope with a 20X oil-immersed lens. Separate 10 µm z-stack image series were taken sequentially for each fluorophore to avoid cross-talk between excitation and emission spectrums. We imaged tissues from 3-6 individuals from each behavioral group. In order to confirm the binding specificity of the secondary antibodies, we also created antibody control staining slides. In this case, staining procedure was the same except the rabbit-anti-Vg 1^{o} antibodies were excluded from the first incubation step. Control slides were made using the longitudinal sections through the center of the lumen (i.e., cross section images, Fig. 4.2).

Statistics and Analysis

Relative gene expression levels were calculated using the ΔΔCt method (Schmittgen and Livak 2008). Differences between nurses and foragers were determined using a T-test on log10-transformed data. All statistical analysis was performed in R (3.5.2). We assessed differences in Vg tissue localization visually between nurses and foragers. The aim was to look for general patterns relating to Vg localization in different regions of the midgut or within different regions of enterocytes.

RESULTS

Gene expression

We found that *vg* was expressed in the midguts of both nurses and foragers from single-cohort colonies. For both groups, *vg* transcript amplification occurred at later cycles (PCR cycles 27.9 and 28.5, respectively), indicating low transcript abundance. The relative expression did not differ significantly between nurses and foragers when compared via their 2^-ΔΔCt values ($t = -0.71$, $df = 30$, $p = 0.48$) (Fig. 4.1).

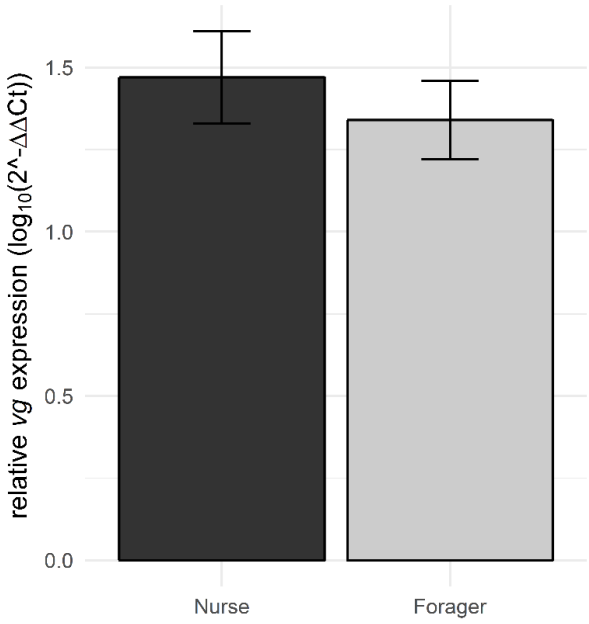

Fig. 4.1: *vg* expression in the midguts of same-aged nurses and foragers. Bar heights represent the mean log10 *vg* expression in the two groups, as calculated by the ΔΔ-Ct method, with error bars representing +/- 1 standard error. Statistical comparison between the two groups was performed via a Student's T-test on log10-transformed data. Expression levels were found to not significantly differ between the two behavioral groups.

General observations of structure

In the longitudinal cross sections, we observed enterocytes protruding into the lumen that measure approximately 100-200 µm in length (Fig. 4.2). Along the length of the organ, the midgut folds inward to increase surface area. These folds can be seen in Fig. 4.2B, D, E, & F, with a closeup of the folds seen in Fig. 4.2A & C. The enterocytes are connected to one another at their base via a basement membrane, and in order to facilitate gut motility, the exterior of the midgut is wrapped in a lattice of muscles that either run the length of the organ (i.e., longitudinal muscles) or which wrap around the organ (i.e., circular muscles). These muscles are clearly visible with the phalloidin stain in Fig. 4.2B & D, which appear as bright red lines and dots running along the bottom of the images. The enterocytes respond to ingested material by secreting a chitinous

substance into the lumen called the peritrophic membrane. This peritrophic membrane is visible above the enterocytes in Fig. 4.2A and B, and to a lesser extent in Fig. 4.2C, and reaches a maximum thickness of approximately 50 µm in our samples.

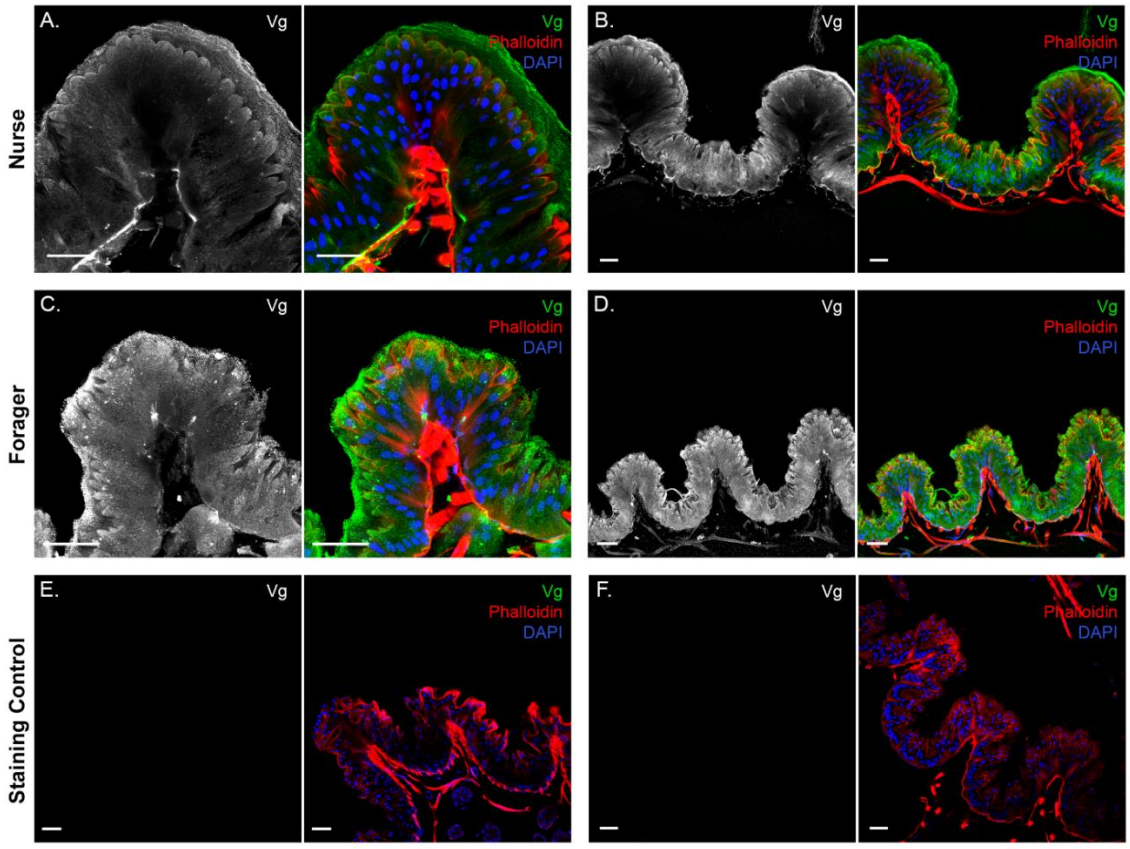

Fig. 4.2: Confocal micrographs showing longitudinal cross-sections of the honey bee worker midgut. The protein Vg is stained with Alexa Fluor 488 (green), nuclei are stained with DAPI (blue), and F-actin is stained with rhodamine phalloidin (red). In each sample, images are oriented with the midgut lumen towards the top and the midgut exterior towards the bottom. Scale bars represent 50 µm. Positive signal for Vg is observed in both the nurses (A-B) and foragers (C-D), but not in the Vg control stained samples (E-

F). Vg is seen not only in the gut epithelial cells, but also in the peritrophic membrane that shields the epithelial cells in the lumen. The peritrophic membrane is more prominent in nurses (A-B) than foragers (C-D). These are representative images of observations made on N=3 and N=6 nurses and foragers, respectively.

In the sections made at the external surface of the midgut, we observed that cells attach to the basement membrane in a cluster formation, and that cells towards the centers of the clusters have their nuclei placed closer to the basement membrane (Fig. 4.3 B, E). Based on their structure, we interpret these clusters as being regenerative nidi (Jimenez and Gilliam 1990; Silva-Olivares et al. 2003; Illa-Bochaca and Montuenga 2006; Park et al. 2009): That is, cells at the center of the cluster are undifferentiated stem cells that divide and differentiate into enterocytes and other midgut cells.

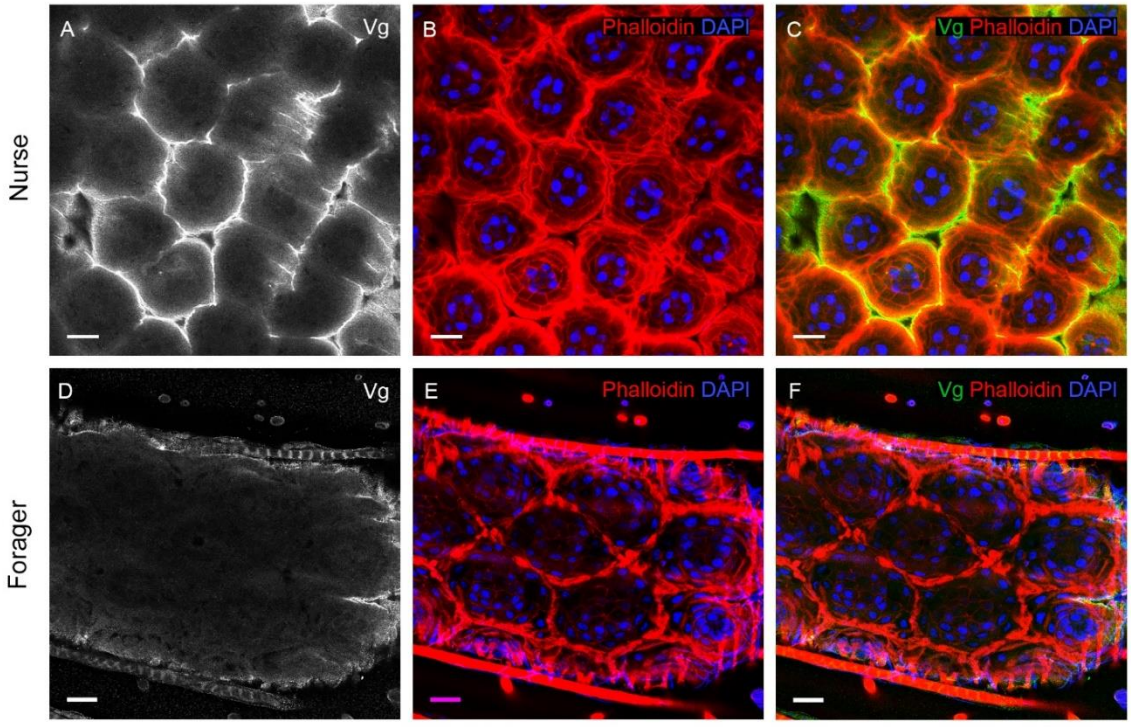

Fig. 4.3: Confocal micrographs looking at the outside surface of the honey bee worker midgut. Vg is stained with Alexa Fluor 488 (green), nuclei are stained with DAPI (blue), and F-actin is stained with rhodamine phalloidin (red). Scalebars represent 25 μm. Cells are arranged in cluster formations, with cells at the center of the clusters having their nuclei closer to the basement membrane. Vg appears around these clusters in nurse bees (A), but not in foragers (D).

Vg localization

We found that Vg protein was localized throughout the enterocytes of same-aged nurses and forages (Fig. 4.2). For both groups, Vg does not appear to be restricted to certain regions of the midgut or compartmentalized in particular areas of the enterocytes: it is found throughout the cytoplasm from the base of the cells to the apex. However, when observing clusters of regenerative cells (Fig. 4.3), Vg appears to surround these clusters in nurses but the pattern is not observed in foragers. This Vg may be deposited extracellularly around these cell clusters, or this phenomenon could be due to microscope edge effects if cells towards the outside of the cluster contain a higher abundance of Vg than cells towards the center (Fig. 4.3 A, B). Nevertheless, Vg does appear to be secreted out of enterocytes and into the peritrophic membrane along the luminal edge of the enterocytes (Fig. 4.2 A, B). We found that the peritrophic membrane was consistently present in all nurse samples imaged but was absent or greatly reduced in all forager samples (Fig. 4.2).

DISCUSSION

This research yielded several intriguing results that expand our understanding of honey bee physiology. First, we show that Vg protein is not only localized in midgut epithelial cells (Figs. 4.2, 4.3), but it is also apparently synthesized there based on local gene expression (Fig. 4.1). This finding confirms a previously overlooked organ for *vg* expression in insects in general, and honey bees in particular (Mao et al. 2013). As with other insects, honey bees primarily synthesize Vg in their fat body tissue (Sappington and S. Raikhel 1998; Tufail and Takeda 2008), but the protein can also be synthesized in queen ovaries (Guidugli, Piulachs, et al. 2005). In some blood-feeding arthropods like ticks, *vg* is expressed in the midgut to activate the ovaries and transport required nutrients from the blood meal (Thompson et al. 2007; Boldbaatar et al. 2010), while in mosquitos a blood meal stimulates fat body production of Vg to facilitate egg production (Bonizzoni et al. 2011). In these examples, Vg appears to be carrying out its typical reproductive function in response to a nutrient stimulus, but roles of Vg in the honey bee worker midgut may be less clear, since these workers are functionally sterile and only rarely activate their ovaries. In honey bees, worker reproduction primarily occurs after irreversible loss of the dominant reproductive queen, with some exceptions (see (Barron et al. 2001)).

We observed that *vg* is expressed at similar levels for same-age nurses and foragers (Fig. 4.1). This result was unexpected, given that *vg* expression in the fat body and Vg titers in the haemolymph are substantially different between nurses and foragers. It has been well-established over decades of research that worker behavior and *vg* expression are intricately linked: nurses have high titers of circulating Vg and foragers

have low titers (Nelson et al. 2007), and conversely, knocking down Vg via RNAi in young adults causes workers to prematurely transition into foragers (Amdam and Omholt 2003; Antonio et al. 2008; Harwood et al. 2016). In the midgut, however, behavioral maturation and *vg* expression appear to have been decoupled. The exact reason for this decoupling remains unclear, but it could be that the midgut maintains steady (albeit low) expression of *vg* in order to carry out some function(s) that are important for both nurses and foragers. Alternatively, midgut *vg* expression may still vary with age as it does with fat body expression and circulating titers of the protein (Nelson et al. 2007), but our use of same-age nurses and foragers will not reveal this pattern. This latter point will require further investigation.

Finally, we found that the peritrophic membrane is consistently present in nurses but is absent or greatly reduced in foragers (Fig. 4.2). This may be due to dietary and physiological changes that coincide with the transition from nurse to forager, as nurses feed primarily on protein-rich pollen in order to produce royal jelly to feed the larvae and queen, while foragers feed on carbohydrate-rich honey (Brodschneider and Crailsheim 2010). Pollen is much more abrasive than honey and may require a more robust peritrophic membrane to protect midgut cells. In ants, carnivorous species are observed with larger peritrophic membranes than their nectar- or honeydew-feeding counterparts (Cook and Davidson 2006). Different diets also require different digestive machinery. The peritrophic membrane contains many digestive enzymes secreted from the enterocytes (Brandt et al. 1978; Lehane 1997; da Cruz-Landim and Cavalcante 2003; Hegedus et al. 2009) and nurses have higher enzymatic activity in their midgut than foragers (Moritz and Crailsheim 1987; Jimenez and Gilliam 1989). The level of

enzymatic activity is directly related to the amount of protein consumed (Jimenez and Gilliam 1989), and foragers have a reduced ability to digest protein (Crailsheim et al. 1992). Alternatively, these results could simply be an artifact of our samples, wherein the nurses had more recently consumed food to trigger the secretion of a peritrophic membrane. However, given the consistent pattern across multiple samples this explanation seems unlikely. Intriguingly, we found that Vg is secreted out of the enterocytes and into the peritrophic membrane, which suggests that it is intended to interact with the ingested materials and/or microorganisms in the gut lumen.

The evidence gathered in this and previous studies points to an as-yet-unknown function of Vg in the midgut: Vg is stably expressed in nurses and foragers, it is secreted into the peritrophic membrane, it is resilient against systemic *vg* knockdown via RNAi (Harwood et al. 2019), and the Vg receptor is expressed in midgut tissue (Guidugli-Lazzarini et al. 2008). The precise nature of Vg's role in the midgut remains unclear, but it's other non-reproductive functions may give a clue. First, we propose that Vg can play a role in antimicrobial activity and cellular maintenance in the insect midgut. For example, Vg may be functioning as part of the anti-microbial defense mechanism against ingested pathogens in worker bees. The midgut is a key organ for fighting off infections (Chapman 1998), and the enterocytes produce AMPs to kill pathogenic cells detected in the lumen. Likewise, Vg can bind pathogenic cells via PAMPS (Li et al. 2009; Liu et al. 2009; Salmela et al. 2015) that include lipopolysaccharides and peptidoglycans found in Gram negative and Gram positive bacteria cell walls, respectively. Vg can either destroy such cells directly (Li et al. 2009) or opsonize them to recruit other components of the innate immune system to destroy them (Li et al. 2008). Interestingly, the AMPs produced

by the enterocytes activate the Immune Deficiency (IMD) pathway which combats Gram negative bacteria, but it is believed that enterocytes do not respond to Gram positive bacteria (Chapman 1998). Several deadly honey bee pathogens like American foulbrood (*Paenibacillus larvae*) and European foulbrood (*Melissococcus plutonius*) are Gram positive bacteria, and their point of entry is the larval gut. We do not yet have data on Vg localization in the larval gut of the honey bee, but larvae are known to express the *vg* gene with functions currently unknown (Guidugli, Piulachs, et al. 2005). Vg may be deployed to combate Gram positive bacterial pathogens.

Second, we propose that Vg can protect midgut enterocytes from the host's own immune defenses. Enterocytes secrete ROS to kill pathogens detected in the midgut (Ha et al. 2005; Buchon et al. 2009, 2013; Kumar et al. 2010), but ROS can also inflict cellular damage on the host's cells. Vg is an antioxidant (Seehuus et al. 2006; Havukainen et al. 2013; Sun and Zhang 2015a) and so may act as a buffer to protect the host's cells from autotoxicity. Moreover, Vg can recognized damaged host cells and protect them against further injury and ROS damage (Havukainen et al. 2013). With abrasive damage from ingested food and ROS damage from immune responses, the insect midgut is a harsh environment that requires pluripotent stem cells to regenerate the population of enterocytes (Micchelli and Perrimon 2006; Ward et al. 2008; Buchon et al. 2014). Cellular turnover of enterocytes is very quick (about 1 – 2 weeks in drosophila (Chapman 1998)), and Vg may help to protect damaged enterocytes and extend the life of each cell in the midgut.

Further studies are required to properly elucidate Vg's function in the midgut, but this will require novel experimental approaches to overcome existing challenges. As

noted, Vg in the midgut appears to be resilient against standard systemic injections of double-stranded RNA (dsRNA) used to elicit RNAi gene knockdown (Harwood et al. 2019), making it difficult to compare control individuals with gene-knockdown individuals. One can administer *vg* dsRNA orally (e.g., (Nunes and Simões 2009), but this introduces issues of inconsistent dosing since individuals consume variable amounts of food (Araujo et al. 2006). Furthermore, ingested dsRNA will be excreted or absorbed through the midgut and into the haemocoel, but the dsRNA that transiently passes through the midgut epithelial cells may have limited efficacy against target transcripts therein. Other studies that tried to target gut-specific genes with orally administered dsRNA failed to knockdown the target gene (Rajagopal et al. 2002). Thus, without an adequate experimental reduction of Vg in the midgut, it is challenging to test for functional roles of this protein.

Nevertheless, the need to understand Vg's role in this important organ remains. This study has expanded our understanding of Vg genetic expression and tissue localization in honey bee workers, and provided important new findings. We have shown that *vg* is not only expressed in honey bee workers, but also that expression is steady between same-age nurses and foragers, and that Vg is secreted from enterocytes into the peritrophic membrane. Further functional experiments will be required to ascertain Vg's precise role in this organ.

CHAPTER 5

NUCLEAR TRANSLOCATION OF VITELLOGENIN IN THE HONEY BEE

Heli Salmela, **Gyan Harwood**, Daniel Münch, Christine Elsik, Elías Herrero-Galán, Maria K. Vartiainen, Gro V. Amdam

ABSTRACT

The Large Lipid Transfer Protein vitellogenin is known for a broad phylogenetic role in yolk formation. A vitellogenin-encoding gene was likely present in the last common ancestor of the majority of animals 700 million years ago, representing the ancestral gene in this protein family. In recent years, additional functions of vitellogenin were discovered in oxidative stress resistance and immunity, in groups as diverse as fish and social insects. It has remained enigmatic how vitellogenin affects multiple traits. We asked whether vitellogenin enters the nucleus and acts via DNA-binding. We used immunohistology, cell fractionation and cell culture to show that an N-terminal part of honey bee vitellogenin translocates into cell nuclei. We used prediction tools to show that vitellogenin can bind to DNA, then demonstrated this empirically with chromatin immunoprecipitation followed by sequencing (ChIP-seq). We found robust vitellogenin-DNA binding loci at genes that influence immunity and behavior. Thereafter, we did an initial test of dynamic properties by means of immunological challenge with *Escherichia coli*, revealing variation in cutting- and nuclear translocation- patterns of vitellogenin *in vitro* and *in vivo*. Finally, we inspected the associated enzymatic modifications with a protease inhibition assay, and summarized findings by building the most complete three-dimensional vitellogenin structure for insects so far. Our data represent the first demonstration of nuclear translocation and DNA binding of a Large Lipid Transfer

Protein. We suggest that a fundamental regulatory role of this ancient protein has been overlooked, with possible implications for a large number of animal species.

INTRODUCTION

Vitellogenin (Vg) is the oldest member of the Large Lipid Transfer Protein (LLTP) family (Hayward et al. 2010). Vg evolved early in the animal lineage and is currently present in nearly all oviparous species as an essential yolk precursor protein. While Vg diversified in some groups to include the cytosolic large subunit of microsomal triglyceride transfer protein (MTP) and apolipoproteins (Apo)(Hayward et al. 2010), the Vg molecule itself may have evolved to support multiple functions. In fish and some insects, for example, Vg is highly pleiotropic (Havukainen, Halskau, and Amdam 2011); acting as an immunomodulatory protein (Li et al. 2008; Liu et al. 2009; Garcia et al. 2010; Tong et al. 2010; Sun et al. 2013b), a behavior regulator (Amdam, Norberg, et al. 2006; Nelson et al. 2007; Antonio et al. 2008; Ihle et al. 2010; Gospocic et al. 2017), an antioxidant with potential impact on rates of aging (Seehuus et al. 2006; Corona et al. 2007b; Havukainen et al. 2013; Sun and Zhang 2015b; Salmela et al. 2016), and as a resource for the feeding of young (Kishida and Specker 2000; Amdam, Norberg, et al. 2003). We have detailed understanding on how Vg forms egg yolk (Tufail and Takeda 2008), and, more recently, how it acts in innate immunity as a pathogen pattern recognition receptor (Zhang et al. 2011; Sun et al. 2013b), but the molecular mechanism(s) by which Vg can influence multiple traits remain unknown.

Pleiotropic proteins can implement their multitude of effects in various ways. One route is by acting as transcription factors or participating in gene-regulatory complexes that affect the expression of many genes (Chesmore et al. 2016). Interestingly, down-regulation of Vg by means of RNA interference mediated *vg* gene knockdown alters the expression of thousands of genes, as exemplified by the honey bee (*Apis mellifera*)

(Wheeler et al. 2013). However, no previous research has addressed or established whether Vg, or any protein homologous to Vg including MTP or Apo (Baker 1988a), can translocate into the cell nucleus and bind DNA.

This gap in our knowledge is self-evident given the molecular size of the LLTPs: The proteins are typically around 200 kDa (Finn 2007; Tufail and Takeda 2008), while nuclear proteins are below 60 kDa in general (although nuclear transfer of one 110 kDa protein was reported (Wang and Brattain 2007). In other words, it seems very unlikely that a LLTP can enter a nucleus. Yet in many animals, Vg is cut and reassembled prior to secretion (Finn 2007; Tufail and Takeda 2008). In invertebrates such as insects, Vg is cut in the vicinity of a multiply phosphorylated polyserine linker sequence stretch that resides between two evolutionarily conserved Vg protein domains called the N-sheet and the α-helical domain (Baker 1988a; Mann et al. 1999; Tufail and Takeda 2008; Havukainen et al. 2012) . In honey bees, this cutting occurs *in vivo* in fat body tissue, which is functionally homologous to the liver and white fat in vertebrates and the primary production-site of Vg (Tufail and Takeda 2008). The cut is between amino acid residues 335 and 427 within the honey bee Vg polyserine linker, resulting in a detached 40 kDa N-sheet of unknown function (Havukainen, Halskau, Skjaerven, et al. 2011).

In this study, we asked whether pleiotropic effects of Vg may be partly explained by nuclear translocation and DNA-binding of the conserved N-sheet, and we used the honey bee as our study organism. First, we tested whether the N-sheet subunit can translocate into the nucleus of fat body cells using an antibody targeting the honey bee N-sheet domain. We verified the results using fluorescently-labeled Vg in cell culture. Second, we predicted the DNA-binding ability of honey bee Vg using sequence- and

structure-based programs, and then used chromatin immunoprecipitation followed by sequencing (ChIP-seq) to identify Vg binding sites in honey bee DNA. We searched for *de novo* binding motifs and performed additional functional analysis of Vg-DNA binding sites *in silico*, identifying many sites associated with immune- and behavior- related genes. These data motivated a functional response-to-challenge assay using *E. coli* to reveal if behavior of Vg cutting and translocation is in fact dynamic. Finally, we established the enzymatic conditions required for Vg cutting, and we developed a three-dimensional structure model to elucidate the critical N-terminal area of the protein.

METHODS

Antibody against honey bee Vg 40 kDa fragment

The honey bee Vg 40 kDa fragment (Uniprot ID Q868N5; the amino acid residues 24-360) was produced in *E. coli* by GenScript (Piscataway, NJ, USA); it was subcloned into pUC57 vector, and an N-terminal hexahistidine tag was used for one-step affinity purification. The polyclonal serum antibody was raised in a rat by Harlan Laboratories (Boxmeer, the Netherlands), and tested by Western blotting against proteins extracted from the honey bee (as in (Havukainen, Halskau, Skjaerven, et al. 2011)) (Fig. S1).

Western blotting

The gel and Western blot reagents were purchased from Bio-Rad, and the protocol for all of the blots was as follows. The blotted nitrocellulose membrane was incubated with PBS containing 0.5% Tween with 2.5% bovine serum albumin overnight.

The membrane was incubated for 1 h with the N-sheet antibody (1:2000) and for 1 h with a horseradish peroxidase-conjugated secondary antibody (1:5000) prior to imaging using an Immun-Star kit. All gels and blots were imaged, and band intensities were measured using a ChemiDoc XRS imager (Bio-Rad).

Vg purification

Vg was purified from wintertime worker honey bee hemolymph as in Salmela *et al.* 2015.

Immunohistology

N=6 winter worker bees were collected from two hives at Norwegian University of Life Sciences, Aas, Norway. The gut was removed, and the abdomen detached and placed for overnight fixing in 4% paraformaldehyde in 4 ºC, followed by three PBS washes, 10 min each. The fat body was dissected in PBS and de/rehydrated with a full ethanol series. For each individual, one tissue sample was used as test (full set of antibodies) and another one as control (no primary antibody). The samples were incubated with the antibody against 40 kDa Vg (see above) overnight in 4°C (1:50 in 2 % BSA-PBS-Triton X-100). DAPI was used as a nuclear marker. The anti-rat secondary antibody was Alexa-568 nm that has no emission spectrum overlap with DAPI. For qualitative anatomical descriptions, a high NA (=high resolution) objective was chosen (40x immersion oil; NA 1.25). Scans were taken with zoom 1.0 and zoom 2.0. All images were taken in sequential acquisition mode to minimize crosstalk between the two channels for detecting DAPI and the Vg signal.

Cell fractionation

Nine winter worker honey bee individuals were anesthetized in cold and the fat body tissue was dissected in ice cold PBS. The fat body tissues were placed in three tubes with 50 µl hypotonic buffer (20 mM Tris-HCl pH 7, 10 mM NaCl and 3 mM $MgCl_2$) pooling three individuals per tube. 2.5 µl 10 % NP-40 was added, followed by 10 s vortex and centrifugation for 10 min 3000 g in 4°C. The supernatant was collected as the cytosolic fraction. The pellet was washed once with 500 µl hypotonic buffer, and then suspended in 30 µl hypotonic buffer with 5 mM EDTA, 1 % Tween-20 and 0.5 % SDS and vortexed shortly. All samples were then centrifuged for 10 min 15000 g in 4°C. 15 µl of each sample was run on SDS-PAGE gel and blotted.

Labeled Vg in cell culture

Pure Vg was labeled with Alexa Fluor 488 protein labeling kit (Invitrogen, Carlsbad, CA, USA). HighFive cells were grown on a 8-well chamber slide (Thermo Fisher) over night (number of slides one, repeated three times). The media was replaced with 20 µg/µl Vg-488 in PBS and incubated in dark for 1 h. The cells were washed twice with PBS, fixed with 4 % paraformaldehyde and the washes were repeated. DAPI was used as a nuclear marker. The cells were imaged the following day with Leica TCS SP5 MP (63x).

DNA-binding prediction

The following prediction tools available for protein-DNA binding were tested with default settings: Sequence-based DNABIND (Liu and Hu 2013), DP-BIND (Hwang

et al. 2007) (run twice due to the input limitation of 1000 amino acids) and DRNApred (Yan and Kurgan 2017), and structure-based DNABIND (Liu and Hu 2013) and DISPLAR (Tjong and Zhou 2007). The structure used was the honey bee Vg N-sheet homology model published earlier (Havukainen, Halskau, Skjaerven, et al. 2011).

ChIP-seq

Samples of newly emerged and 7-day older worker bees were created by pooling fat body tissue from 10 individuals from each age group, all originating from the same hive. Freshly harvested samples were flash frozen in liquid nitrogen and homogenized. Chromatin immuno-precipitation was carried out using established protocols (Bai et al. 2013) using Dynabeads™ Protein G (Invitrogen). We opted to use polyclonal antibodies raised in rabbits against the whole 180 kDa Vg molecule (Pacific Immunology, Ramona, CA) (Jensen and Børgesen 2000) rather than the rat-origin antibodies against the 40 kDa Vg subunit used elsewhere in this study because of their superior immunoprecipitation performance. In a preliminary study, the rabbit-origin antibodies consistently pulled down more chromatin than multiple batches of the rat-origin antibodies, which failed to retrieve sufficient chromatin for sequencing. This is likely due to Protein G (and A) having a greater affinity for rabbit-origin than rat-origin antibodies, as per the manufacturer. In a preliminary study we had compared immuno-precipitation efficacy using antibodies against the whole 180 kDa Vg molecule with antibodies against the 40 kDA N-terminal domain used elsewhere in this study. The antibody against the 180 kDa molecule consistently pulled down more chromatin than multiple batches of antibodies against the 40 kDa domain, so we elected to use the former in this procedure. Chromatin

samples were sequenced at the DNASU lab at Arizona State University. The raw Illumina 2x75bp pair-end reads were quality checked using FastQC v0.10.1 (Andrews 2010), followed by adapter trimming and quality clipping by Trimmomatic 0.35 (Bolger et al. 2014). Any reads with start, end or the average quality within 4bp window falling below quality scores 18 were trimmed. The clean reads were aligned to reference genome *Apis mellifera* Amel_4.5 (https://www.ncbi.nlm.nih.gov/genome/48?genome_assembly_id=22683) by Bowtie 2 version 2.2.9 (Langmead and Salzberg 2012). Another round of QC was performed on bam files. Library complexity was checked by NRF (nonredundancy fraction), defined as the number of unique start positions of uniquely mappable reads divided by number of uniquely mappable reads. All the samples passed the threshold 0.8 recommended by ENCODE. IGVtools and bamCompare from deepTools (Ramírez et al. 2014) were employed to compare two BAM files based on the number of mapped reads. First the genome is partitioned into bins of equal size and then the number of reads in each bin is counted. The log2 value for the ratio of number of reads per bin was reported for IGV visualization. MACS2 was used for peaks calling with 0.01 FDR cutoff. Narrowpeak files as MACS2 output were annotated by HOMER (Heinz et al. 2010). It first determines the distance to the nearest TSS and assigns peak to that gene. Then it determines the genomic annotation of the region covered by the center of the peak, including TSS, TTS, Exon (coding), 5' UTR Exon, 3' UTR Exon, Intronic, or Intergenic.

To test for non-random occurrence of peaks within genome features, we used 1000 random peak datasets from the 7-day old worker data set. To create the random peak datasets, we used the shuffle program of the BEDTools package (Aronesty 2011) on

the 782 peaks from the 7-day data set that were located on full chromosome assemblies to generate 1000 bed files with peak locations that were shuffled within chromosomes, such that shuffled peak locations were non-overlapping and did not occur in assembly gaps. The annotatePeaks.pl program from the Homer package (Heinz et al. 2010) was then used to annotate the 1000 shuffled peak data sets and the 7-day peak data set with respect to genome features in the NCBI *Apis mellifera* RefSeq annotation. We used chi-squared tests to determine whether the observed numbers of peaks overlapping promoter plus TSS regions (-1 kb to +100 bp from TSS), TTS, exons, introns and intergenic regions were greater or less than expected by chance.

We performed a gene functional annotation and cluster analysis using DAVID 6.8 (Huang et al. 2009a, 2009b) to look for common gene ontology (GO) terms among genes pertain to *biological function*, *molecular function*, and *cellular component*, as well as KEGG pathway terms. We used default settings on DAVID in which clusters were 3 or more genes that shared a common GO term. This cluster analysis uses a modified Fisher's Exact Test to look for enriched terms in the gene list, as well as provides a more conservative FDR statistic using a Benjamini-Hochberg procedure. Finally, it provides an estimate of fold enrichment to further test whether certain gene annotations are proportionally more represented in the ChIP data set than with the honey bee genome as a whole.

Vg-DNA binding motif search

For *de novo* motif identification, we created a non-redundant dataset of 790 peaks by combining peaks identified in the newly emerged and 7-day old samples and merging

peaks with overlapping regions between the two dataset. We used GimmeMotifs (Heeringen et al. 2011)ich ran and combined results for ten motif prediction packages – Mdmodule , Weeder, GADEM, trawler, Improbizer, BioProspector, Posmo, ChIPMunk, JASPAR, AMD, HMS and Homer. GimmeMotifs clusters the results, performs enrichment and computes receiver operator characteristic (ROC) and mean normalized conditional probability (MNCP). Half of the peaks (395) were used to train each algorithm and the other half was used for testing. Since we did not have *a priori* knowledge of the motif length, we ran GimmeMotifs four times with different size range options (small 5-8 bp, medium 5-12 bp, large 6-15 bp, xl 6-20 bp). Since the peaks were located throughout the genome, we used sequences randomly chosen from the genome with GC content to the peak sequences as the background for the enrichment tests.

Vg cutting and nuclear translocation in response to E. coli In vitro

K-12 strain of *E. coli* were grown to $1x10^8$ cells/ml density, washed with PBS and heat-killed by shaking them in 95 °C for 30 min (Groh et al. 1996). This thermal treatment also permanently inactivates the *E. coli* proteases (Moran et al. 2001). The bacteria were then diluted 1:5 and 1:25 in PBS. 6 µl of bacteria and equal volume of 8.5 mg/ml fat body protein extract (see above) were mixed and incubated in +26°C for 30 min and then blotted using the Vg-N-sheet antibody. *In vivo*: Worker honey bees tending brood were captured. This behavior is performed by 4- to 12-day-old nurse bees (Johnson 2010); this age/behavior group was chosen because they are known to have elevated levels of Vg compared to younger or older summer workers and to have better age control compared to winter bees (Amdam, Norberg, et al. 2005a). The bees were caged in

groups of 14 and fed *ad libitum* either with sterile 50 % sucrose solution or the same solution mixed with *E. coli* K-12 strain BioParticles, Alexa Fluor 488 conjugate #E13231 (Invitrogen) in final concentration of 0.5 mg/ml. The bees were kept in +34°C in dark for 24 h. Subsequently, the bees were immobilized by cold treatment and their fat body tissue was prepared for confocal microscopy as above (see: Immunohistology). This protocol resulted in 8 successful control samples and 9 treatment samples, in addition to one antibody control sample per treatment group. The presence of the labeled *E. coli* particles in the fat body tissue was confirmed by detection of their 488 nm fluorescence signal. For each individual, we zoomed in three different fat body tissue areas with approximately 15 cells and counted the number of cells with visible nuclear localization and the number of cells without a sign of nuclear localization. Logistic regression with quasibinomial distribution was used for statistical analysis.

Electron microscopy

Aliquots of pure Vg (1.1 mg/ml in 20 mM Tris, 150 mM NaCl, pH 7.5) were applied to carbon-coated copper grids (30 s) and stained with 2% uranyl acetate. Micrographs were taken in minimal dose conditions in a JEOL 1230 transmission electron microscope operated at 100 kV, using a low-dose protocol and a 4k x 4k TVIPS CMOS detector under the control of EM-TOOLS software (TVIPS). Final magnification of the CMOS images was 54,926. A total of 15,000 particles were selected, normalized, and CTF-corrected using procedures implemented in the XMIPP package (Scheres et al. 2008). For three-dimensional reconstructions, different starting templates were generated using the EMANstartcsym program, by common lines or using artificial noisy models

and Gaussian blobs with the rough dimension of the proteins (Ludtke et al. 1999), in both cases with identical results, which confirms the robustness of the structure obtained. The structure of the homologous protein lipovitellin from silver lamprey (1lsh) was fitted into the final vitellogenin volume and used to determine the handedness of the structure.

RESULTS

Immunohistology

Worker honey bee fat body tissue was observed with confocal microscopy using an antibody targeting the Vg N-terminus and a fluorescently labeled secondary antibody (Fig. 5.1A-E). Honey bee fat body consists of trophocyte and oenocyte cells (Chapman 1998). Trophocytes are responsible for Vg production and storage while oenocytes do not produce or contain Vg (Pan et al. 1969; Isaac and Bownes 1982; Raikhel and Lea 1983; Bownes 1986). We observed two types of Vg localization patterns in trophocytes: i) Vg signal co-localized with DAPI in the nucleus, and also found spread throughout the cytosol (Fig. 5.1A-B), and ii) Vg signal absent in the nucleus and instead restricted to granule-like formations in the cytosol (Fig. 5.1C, see also (Havukainen, Halskau, Skjaerven, et al. 2011) for observations of this pattern). Controls for autofluorescence and unspecific secondary antibody staining were included (Fig. 5.1D-E), which indicated that the immunosignals present in Fig. 5.1A-D are specific for Vg antibody incubation.

Cell fractionation

To further verify the nuclear location of the N-terminal Vg fragment, we separated fat body cells into cytosolic and nuclear compartments by cell fractionation

followed by western blot (Alberts et al. 2002). Full-length Vg (180 kDa) and a ~75 kDa band were found in the cytosolic component, while the nuclear component only contained fragments below 75 kDa, including the 40 kDa N-sheet domain (Fig. 5.1F). We have shown previously that the 40 kDa Vg subunit is a specific cleavage product of fat body cells and not simply a biproduct of unspecific degradation (Havukainen, Halskau, Skjaerven, et al. 2011; Havukainen et al. 2012). We also observed 70, 37 and 25 kDa fragments in the nuclear fraction, but these are likely degradation products caused by the method, as the bands are faint or non-existent in untreated Western blot samples (Supplementary data SI1).

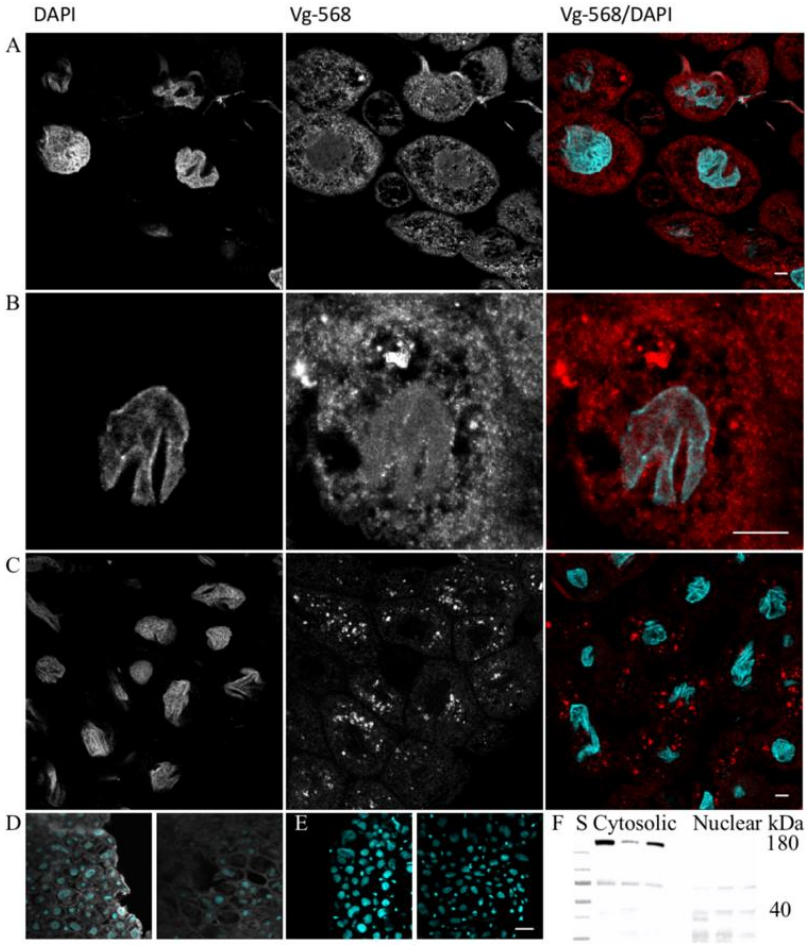

Fig. 5.1. Localization of Vg signals in the honey bee fat body using an antibody targeting the N-terminal 40 kDa domain. A-E Confocal images of fat body trophocyte cells with a characteristic irregularly shaped nuclei representing six samples. The last panel shows the superposition

of DAPI (cyan) and Vg signal (red). A Vg signal co-localizes with DAPI in cell nuclei. B Zoom-in of a single cell showing Vg nuclear translocation. C Cells that do not show Vg co-localization with DAPI, instead, Vg signal is found in granules in the cytosol. The scale bar = 10 µm. D-E show representative images for controls where the primary antibody was applied (D) or omitted (E). Note that when omitting Vg antibody incubation, no immunosignal was detected (grayscale), while only the DAPI signal (cyan) was present. Scalebar = 50 µm. Magnification 40 x. F Western blot of cytosolic and nuclear fractions of honey bee fat body tissue. S = size standard. The full-length (180 kDa) Vg is dominating the cytosolic fraction of the fat body proteins, whereas the 40-kDa N-terminal fragment mostly localizes in the nuclear fraction. Also other fragments of approximately 70 and 25 kDa were visible in the nuclear fraction. Three individuals were pooled for each of lane. N = 3.

Antibody-free, labeled Vg in cell culture

To rule out that the nuclear signal was an artefact caused by Vg primary antibody, we incubated cell cultures (Lepidopteran ovarian HighFive cells) with pure fluorescently-labelled Vg. During incubation, the fluorescent Vg was imported into cells and was visible in the cell nuclei co-localizing with DAPI, and also visible in the cytosol in granule-like formations (Fig. 5.2).

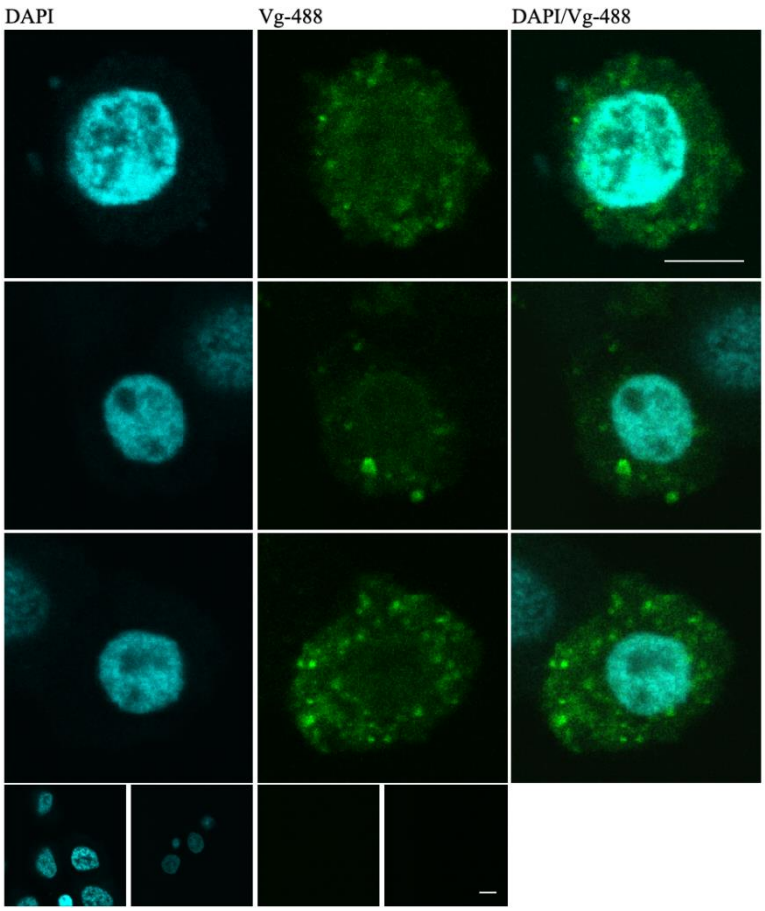

Fig. 5.2. Localization of Alexa-488-labelled purified Vg in cultured insect HighFive cells. Vg was observed both in cytosol mostly in bright granules and as haze in the nucleus in the confocal sections. The lowest row shows two Vg-negative controls imaged with the same settings. Scale bar = 5 µm. Magnification x 63

Predicting Vg-DNA-binding

Using the whole Vg amino acid sequence, two separate programs with different search algorithms, DP-Bind (Hwang et al. 2007) and DRNApred (Yan and Kurgan 2017), both identified the same amino acid residue stretch as a putative DNA binding domain: SRSSTSR in the position of 250-256 in the N-sheet domain of Vg. Another sequence-based program, DNABIND (Liu and Hu 2013), also identified N-terminal Vg as a DNA-binding protein. This program predicts a protein's DNA binding probability and sets a

threshold probably of 0.5896, above which a protein is statistically likely to be able to bind DNA. Vg scored a 0.6267 probability, indicating it can bind DNA. Additionally, the structure-based prediction software DISPLAR (Tjong and Zhou 2007) identified the SRSSTSR stretch as capable of binding DNA using a published honey bee Vg N-sheet model (Havukainen, Halskau, Skjaerven, et al. 2011) as input. There were another 12 and 3 additional stretches identified by DP-BIND and DRNApred, respectively, which did not overlap between the programs, whereas DISPLAR identified two additional stretches that did not overlap with predictions made by the two sequence-based programs. Taken together, results from multiple prediction software platforms support the hypothesis that Vg can bind to DNA, and this capability is likely restricted to the N-sheet.

ChIP-seq

To test empirically whether Vg binds to DNA, and to determine what types of genes and genomic regions Vg is bound to, we performed ChIP-seq on newly emerged (<24 hours) and 7-day old worker bees. Newly emerged workers will take on tasks such as cleaning comb cells, while workers at 7 days old tend to nourish the queen and developing larvae. Thus, workers in these two age groups span an important behavioral transition or maturation that is integral to a colony's division of labor (Seeley 1982). When comparing ChIP DNA with input DNA (i.e., DNA not immunoprecipitated with antibodies), we found significant enrichment (FDR < 0.01) at 90 and 782 putative Vg-DNA binding sites in newly emerged and 7-day old workers, respectively. We found a greater number of binding sites than expected by chance within promoter-Transcription Start Site (TSS) regions, Transcription Termination Sites (TTS), and intergenic regions,

while we found fewer binding sites than expected by chance in exons and none in introns (Table 5.1). Of the 90 binding sites in newly emerged bees, 83 (92%) were also present in 7-day old bees, illustrating an expansion of binding sites as workers age. Many of the binding sites (66 and 559, respectively) were found in close proximity to coding regions, being located either in promotor regions, exons, TTSs, or within 2kb from a transcription start site.

Table 5.1. Tests for nonrandom occurrence of peak in genome feature

genome feature	total 7D peak on chromo-somes	obs overlap	obs not overlap	exp overlap	exp not overlap	chi-squared	P-value
promoter	782	162	620	62	720	175.1792115	5.47106E-40
tts	782	109	673	46	736	91.67527174	2.08916E-24
exon	782	81	701	99	683	3.747105018	0.052899057
intron	782	222	560	439	343	244.5499512	4.00545E-55
intergenic	782	208	574	136	646	46.14241486	1.09962E-11

We pooled all Vg-DNA binding sites from both age groups, and we performed a gene functional annotation clustering analysis to better understand the biological role of genes at these loci. We performed this analysis using DAVID Bioinformatics Resources (Huang et al. 2009a, 2009b) with default settings, drawing from Gene Ontology terms for *biological processes*, *molecular function*, and *cellular component*, as well as KEGG pathway terms. This yielded a total of 26 functional clusters (Table 2), with several clusters significantly enriched when compared to the honey bee genome as a whole, including “cation channel activity” (N=11, P<0.001), “G protein-coupled receptor activity (N=11, P=0.004), “cell adhesion” (N=6, P=0.009), “gated channel activity”

(N=9, P=0.009), “apoptotic process” (N=4, P=0.03), and “nucleic acid binding” (N=54, P=0.054). We found Vg-DNA binding sites at several key innate immune system genes that were not assigned to a cluster, including *Toll-like receptor 4*, *defensin-1*, *autophagy protein 5*, and *autophagy-related protein 9*. Within the “G protein-coupled receptor activity” cluster, we also found genes coding for receptors that are known to play a role in complex phenotypes, like behavior, including receptors for corazonin, glutamate, acetylcholine, and octopamine (see Dataset S1 for complete list of Vg binding sites). These results demonstrate that Vg binds to DNA at hundreds of loci in the honey bee genome, that several are behavior-relevant, and that many of these Vg-DNA interactions persist as adult workers transition between tasks.

Table 5.2. Gene function cluster analysis of gene-associated Vg-DNA binding sites

Cluster	GO Term	Number of Genes	Modified Fisher's Exact Test (P-value)	Fold Enrich-ment	FDR (B-H adjusted P-value)
cation channel activity	GO:0005261	11	< 0.001	4.587	0.025
G-protein coupled receptor activity	GO:0004930	11	0.004	2.940	0.217
cell adhesion	GO:0007155	6	0.009	4.537	0.990
gated channel activity	GO:0022836	9	0.009	2.978	0.351
apoptotic process	GO:0006915	4	0.030	5.694	1.000
nucleic acid binding	GO:0003676	54	0.054	1.239	0.722
extracellular matrix	GO:0031012	4	0.099	3.476	0.830
calcium channel activity	GO:0005262	3	0.109	5.212	0.863
actin filament-based process	GO:0030029	4	0.131	3.122	1.000
monosaccharide metabolic process	GO:0005996	3	0.166	4.033	1.000
cellular component assembly	GO:0022607	8	0.217	1.613	1.000
regulation of cell communication	GO:0010646	4	0.245	2.305	1.000
sequence-specific DNA binding	GO:0043565	10	0.248	1.448	0.967
Pyruvate metabolism	ame00620	3	0.321	2.573	0.994
protein tyrosine phosphatase activity	GO:0004725	3	0.337	2.502	0.974
membrane organization	GO:0061024	3	0.364	2.342	1.000
phosphatidylinositol binding	GO:0035091	3	0.372	2.317	0.979
transcription factor activity, transcription factor binding	GO:0000989	3	0.406	2.157	0.985
serine hydrolase activity	GO:0017171	4	0.455	1.604	0.992
chitin metabolic process	GO:0006030	3	0.493	1.815	1.000
motor activity	GO:0003774	3	0.562	1.604	0.996
metal ion binding	GO:0046872	42	0.679	0.968	0.998
structural molecule activity	GO:0005198	8	0.756	0.937	0.999
cellular catabolic process	GO:0044248	4	0.819	0.913	1.000
nucleotide metabolic process	GO:0009117	3	0.867	0.864	1.000
GTP binding	GO:0005525	4	0.964	0.609	1.000

Binding motif identification

We ran four analyses using GimmeMotif (Heeringen et al. 2011), and we selected the top motif predictions, based on a combination of percent enrichment, P-value, ROC-AUC and MNCP (Table 3) (Fig. 5.3). The top motif prediction for the small motif run

was a sub-motif of the large motif run, so it is was discarded. These results suggest that there are specific sequence motifs to which Vg binds.

Table 5.3. Top de novo motif predictions.

Motif	Enrich-ment	P-value	ROC-AUC *	MNCP*	Best Known Match	P-value for Match	No. Peaks with motif	Analysis
TCAAGAGATGGCGC	35.75	6.36E-122	0.737	5.44	C2H2_ZF_M2196_1.01	5.27E-04	300	large
nAkyrCCATCTnTyGrwwAn	34.5	2.12E-116	0.788	5.58	bHLH_Average_34	2.70E-02	310	medium
TyAGCGCCATCT	33.1	2.81E-116	0.725	5.239	C2H2_ZF_M2196_1.01	8.10E-07	314	xl

*ROC-AUC = area under ROC curve, MNCP = mean normalized conditional probability

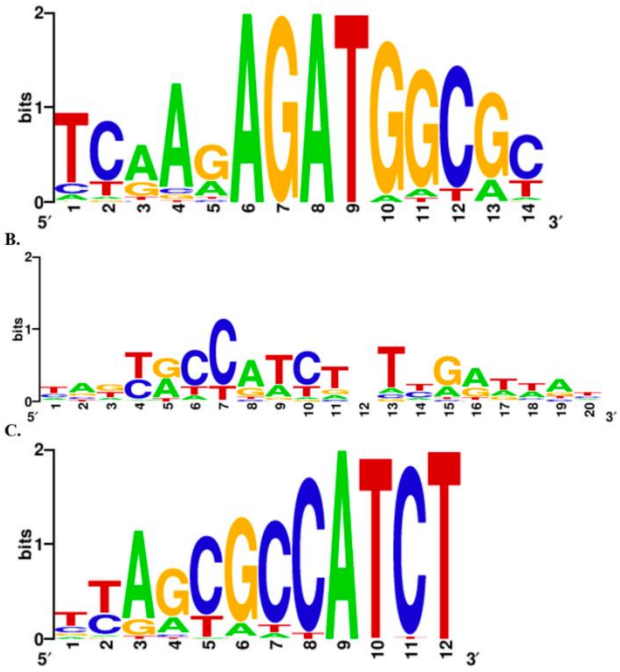

Fig. 5.3. Logo representations for the top motif predictions. A) Top motif prediction from analysis with “large” motif setting (35.75% enrichment). B) Top motif prediction from analysis with “medium” motif setting (34.5% enrichment). C) Top motif prediction from analysis with “xl” motif setting (33.1% enrichment).

In vitro E. coli incubation

As Vg is a pathogen pattern recognition receptor (Li et al. 2008; Salmela et al. 2015), including vertebrates like fish (Liu et al. 2009; Zhang et al. 2011), and we detected Vg binding at several innate immunity genes (Results), we did a bacterial challenge for initial exploration of Vg translocational mechanisms. We incubated fat body protein extract rich in Vg with an increasing concentration of heat-killed *E. coli* whose bacterial proteases were inactivated by the thermal treatment (Moran et al. 2001). We measured the level of Vg fragmentation by Western blotting (four concentrations of *E. coli*, N = 3 each). The presence of the bacteria appeared to enhance Vg-cutting, as the amount of N-terminal fragmentation products increased in response to the number of bacteria (Fig. 5.4A).

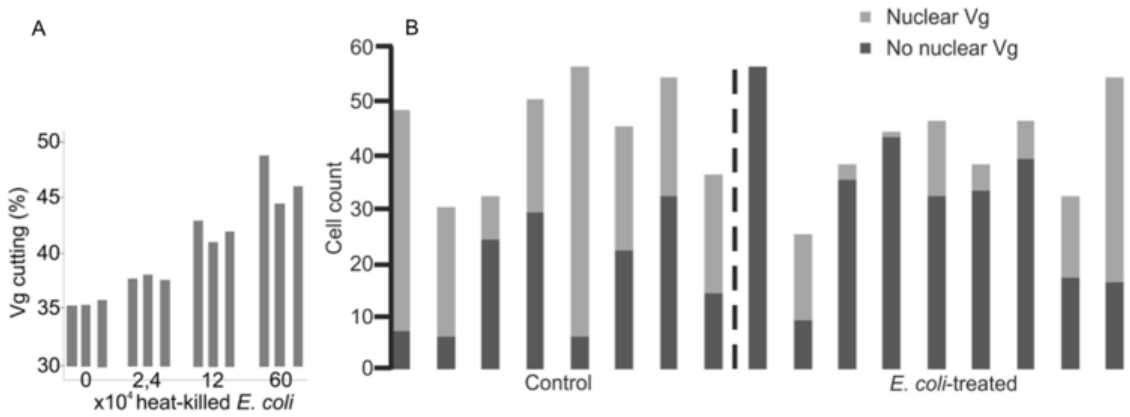

Fig. 5.4. Effect of exposure to bacterial material on Vg cutting and nuclear localization. A. Vg cutting in response to *E. coli*. Honey bee fat body protein extract was subjected to a dilution series of heat-killed *E. coli* for 30 min in replicates of three and blotted using the Vg-N-terminal antibody. The level of Vg cutting was determined by dividing the total Vg protein signal by the fragments below 75 kDa (the nuclear fragments, see Fig. 5.1F).

B. Vg localization response to orally consumed bacteria in honey bee workers. Presence of Vg in the nucleus of fat body trophocyte cells was detected using confocal microscope in a control group (N = 8) and in a group fed with fragments of killed *E. coli* (N = 9). Each bar represents the number of cells counted in a honey bee individual's fat body tissue. See representative images in Fig. 5.2.

In vivo E. coli consumption

We fed caged honey bee worker nurses (age 4 – 12 days) with control food or food supplemented with killed fragmented *E. coli* overnight and detected the level of nuclear translocation of Vg using confocal microscopy. In the control group (N = 8), the proportion of fat body trophocyte cells showing nuclear translocation of Vg ranged from 25.00 % to 89.29 % (Fig. 5.4B). In the *E. coli* treatment group (N = 9), the range was from 0 % to 64.00 %. Regardless of the high level of individual variation in both the control and the *E. coli*-fed group, the proportion of cells with Vg signal in the nucleus (versus cells with signal in the cytosol) was significantly higher in the control group than in the group that received bacterial fragments in their diet ($b = -1.4499$, $z = -2.501$, $df = 16$, $p = 0.0245$). These two experiments, together, indicate that the regulation of Vg cutting and nuclear translocation responds dynamically after exposure to bacterial material.

Vg cutting site inspection via 3D structure and cutting assay

Few structural models exist for Vg (Roth et al. 2013) and none of them focus on the dynamics at the N-sheet domain. To help bridge this gap, we used electron

microscopy to create the most complete 3D model of insect Vg so far and, in addition, we establish the enzymatic conditions required for cutting the 40 kDa N-sheet domain from the rest of the honey bee Vg molecule using a western blot -based inhibition assay (Fig. S2).

A low-resolution 3D electron microscopy reconstruction of the purified honey bee Vg was carried out using negatively stained specimens (Fig. 5.5A), after multiple X-ray crystallization attempts failed. The structure has the general shape of a cylinder of ~140 Å high and 80 Å wide, with two small masses protruding from the top and bottom of the structure. The existing X-ray structure of the homologous protein lipovitellin from lamprey (1lsh) (27) was fitted into the vitellogenin volume (Fig. 5.5B). The N-terminal domain (orange mass in Fig. 5B), together with both the β-barrel and the α-helical domains of lipovitellin (yellow mass in Fig. 5B), were easily fitted into the Vg volume, and the lipid-binding cavity was easily identified. The unresolved fragments of the lipovitellin structure seem to match the position of the extra masses observed in the Vg EM volume (Fig. 5.5B). The N-sheet followed by the linker stretch appears to protrude from the structure (red mass in Fig. 5.5C). These results suggest that enzymes have clear access to the N-sheet cutting site (see Fig. S2 for enzymatic cutting assays). Moreover, the following treatments were found to prevent the cutting of Vg: leupeptin, YVAD-AOMK, EDTA and PhosSTOP. This indicates that, among possibly several other molecular mechanisms, dephosphorylation (PhosSTOP inhibits dephosphorylation) and caspase-type activity (YVAD-AOMK is a specific and leupeptin a broad-range caspase inhibitor) are important for Vg cutting to occur (Fig. S2).

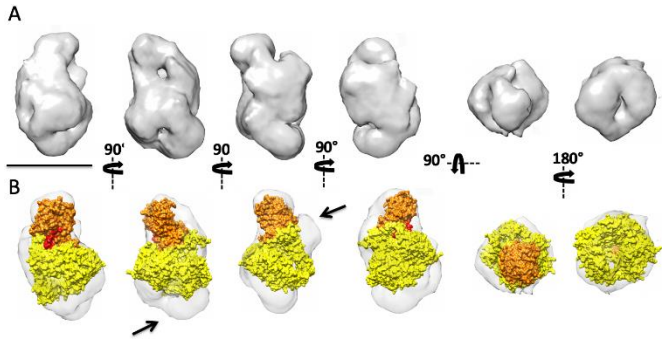

Fig. 5.5. The 3D reconstruction of the honey bee Vg reveals the exposure of the cutting site. (A) Different views of the 3D reconstruction of the Vg from honey bee. The four on the left are orthogonal side views of the volume, whereas the two on the right correspond to the two end-on views of the 3D reconstruction. Bar indicates 100 Å. (B) The same views with the atomic structure of lipovitellin from lamprey (pdb 1lsh) docked into the EM 3D reconstruction. The N-sheet domain, which protrudes from the main body of the structure, is highlighted in orange. The domain colored red points to the linker that connects the N-sheet domain to the lipid cavity (yellow mass).

DISCUSSION

Vg is the oldest member of the Large Lipid Transport Protein family, dating back at least 700 million years (Hayward et al. 2010), and is common to nearly all oviparous animals owing to its central role in egg-yolk production. With few exceptions, it is found across the entire animal kingdom, ranging from some of the simplest multicellular animals, like Placazoans (Hayward et al. 2010), to some of the most complex, like mammals (Warren et al. 2008) (Fig. 6). It performs numerous functions, including many pertaining to complex traits such as behavior (Amdam, Norberg, et al. 2003, 2006; Nelson et al. 2007; Antonio et al. 2008; Corona et al. 2013), longevity (Amdam and Omholt 2002; Corona et al. 2007b; Amdam et al. 2012), and immunity (Amdam, Simões, et al. 2004; Shi et al. 2006; Li et al. 2008; Zhang et al. 2011; Salmela et al. 2015; Sun and

Zhang 2015b). The Vg molecule contains several highly conserved domains, and thus many of its functions are shared by diverse taxa. For example, its role as a pathogen pattern recognition receptor is observed in coral (Du et al. 2017), fish (Li et al. 2008; Liu et al. 2009), and insects (Salmela et al. 2015).

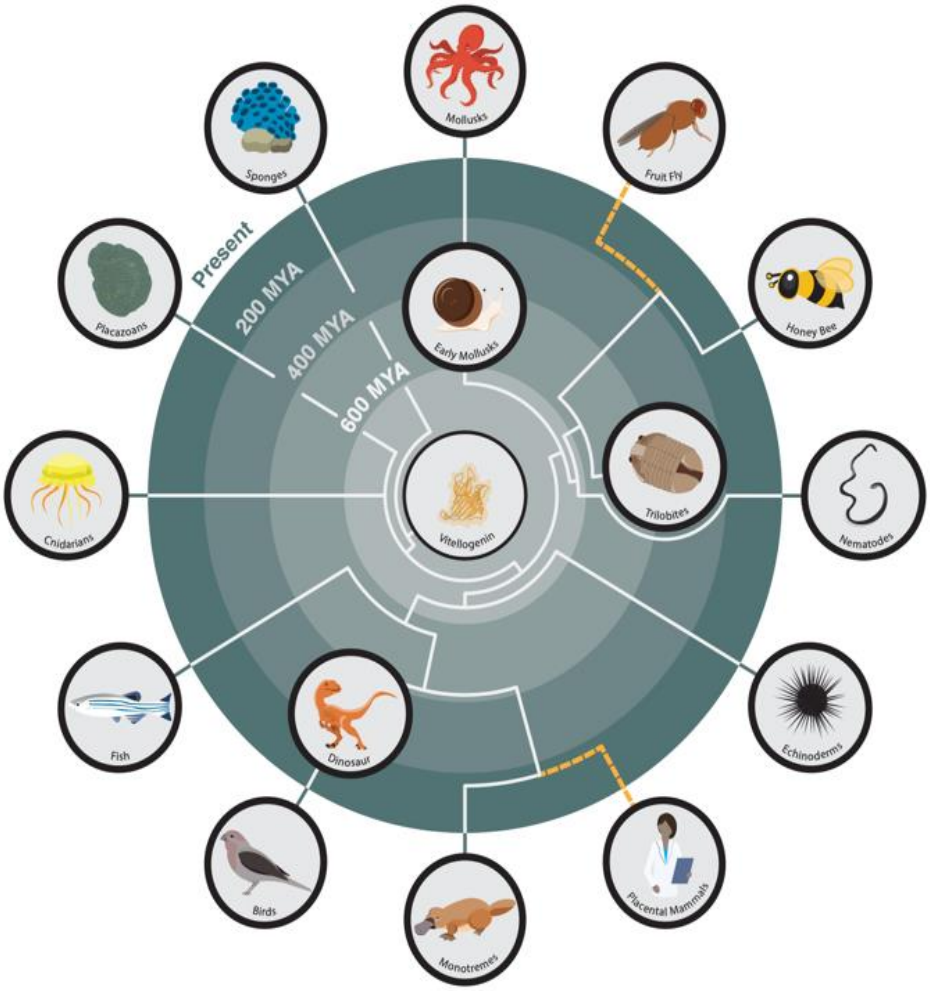

Fig. 5.6 The deep phylogenetic history of Vitellogenin. Vg is the oldest member of the LLTP family of proteins, and first evolved around the time Metazoans (animals) appeared, more than 700 million years ago (1, 97). It is present in all extant Metazoan

phyla, from earliest animals like sponges and cnidarians to the more recently evolved chordates, like fish, birds, and monotreme mammals (48, 98–108). Although nearly ubiquitous, Vg has been lost in several lineages that are the focus of much scientific research, including placental mammals and higher dipterans like *Drosophila melanogaster* (109, 110). Vg's earliest known functions pertain to egg-yolk formation and immunity, but it remains to be seen when DNA binding evolved.

Here, we reveal a major and hitherto unknown ability of Vg that can provide a molecular explanation for how Vg regulates so many complex traits: nuclear translocation and DNA binding. Vg may be directly involved in modulating gene expression, and given its ubiquity in the animal kingdom and its conserved molecular structure, it is possible that it performs a similar function across a wide array of oviparous animals.

In our study, we found strong evidence of Vg translocating to the nucleus using three methods: immunohistology, cell fractioning into nuclear and cytosolic compartments and anti-body-free localization of Vg in cell culture. The two first approaches confirmed the Vg N-sheet domain to be present in both nucleus and cytosol, with the latter result reaffirming previous findings (Smedal et al. 2009; Havukainen, Halskau, Skjaerven, et al. 2011). Moreover, the second approach specifies that the cell nucleus excludes the full length 180 kDA Vg molecule and the previously reported specific Vg fragmentation product of 150 kDa size (i.e., the Vg molecule without the N-sheet domain) (Havukainen, Halskau, Skjaerven, et al. 2011). The cell culture approach makes it less likely that antibody artefacts explain the outcomes of the first experiments.

We found theoretical support for Vg's DNA binding ability using several prediction software platforms, and this prompted us to test this ability empirically. Our ChIP-seq analysis confirmed that honey bee Vg binds to many loci in fat body DNA of newly emerged and 7-day old workers. Despite there being far fewer Vg-DNA binding sites in newly emerged than 7-day old workers (150 vs 927 loci, respectively), there is robust overlap in binding sites between these two age groups as over 75% of sites observed in newly emerged workers are still present in 7-day older workers. This suggests that Vg not only maintains long-term associations with specific loci, but also that the repertoire of Vg-DNA binding sites expands as workers age and behaviorally transition from cell cleaning to nursing.

Gene functional annotations of these Vg-DNA binding sites hint at the types of biological functions that Vg may be targeting for regulation. For example, two of the larger functional clusters have GO terms "nucleic acid binding" and "transcription factor activity" (Table 5.1), suggesting that Vg interacts with molecular machinery regulating gene expression. These clusters include genes coding for transcription factors, histones, methyltransferases, and zinc finger proteins (Dataset S1), all important players in gene regulation. Additionally, we found significantly enriched functional clusters for GO terms "ion channel activity" and "G protein-coupled receptor activity" (Table 5.1), suggesting that Vg regulates components of signal transduction pathways. Interestingly, newly emerged and 7-day old workers perform different tasks, and many genes in the "G protein-coupled receptor activity" cluster are known to affect how an individual behaves and responds to stimuli. These include receptors for glutamate (Kucharski et al. 2007), acetylcholine (Eiri and Nieh 2012), corazonin (Gospocic et al. 2017), and octopamine

(Grohmann et al. 2003). One possible explanation is that as workers age and transition between tasks they respond to different signals from nestmates and the environment, and thus must be equipped with relevant molecular machinery to receive and respond to those signals.

A potential shortcoming here is that these Vg-DNA binding loci are from fat body cells rather than neurons, so the effect on behavior might appear limited. However, Vg is neither found nor expressed in honey bee neurons (Münch et al. 2015), limiting its likelihood of binding to DNA therein. Fat body, on the other hand, not only produces and stores Vg, but also plays a central role in regulating metabolism and homeostasis. In this capacity, it produces many products that affect a wide range of insect behaviors, including courtship (Lazareva et al. 2007), host-seeking (Klowden et al. 1987), feeding (Zinke et al. 1999), and onset of foraging (Antonio et al. 2008). Moreover, the fat body is a key player in innate immunity (Fehlbaum et al. 1994; Morishima et al. 1997; Lycett et al. 2006), and our data show Vg to be bound to several key immune-related genes in this tissue, including toll-like receptors (reviewed in (Akira et al. 2001; Medzhitov 2001)), *defensin-1* (reviewed in (Ganz 2003)), and autophagy proteins (reviewed in (Levine and Deretic 2007; Xu et al. 2007; Nakahira et al. 2011; Deretic 2012)) (Dataset S1). Taken together, the fat body's central role in honey bee biology means that any protein-DNA binding herein could have wide-ranging effects on numerous physiological pathways. This work also revealed several putative *de novo* DNA binding motifs for Vg. The aim of this ChIP-seq work here was to determine whether Vg binds to DNA, and if so, at which loci. Our next step is to further our understanding of Vg's actions inside the nucleus and to determine how it regulates gene expression across different caste types and populations

in honey bees. This is a multi-faceted approach that will involve performing ChIP-seq on workers, drones, and queens from multiple colonies to determine how Vg differentially binds to DNA, and performing gene expression analyses such as RNA-seq to elucidate how Vg-DNA binding up- or down-regulates gene expression at various loci. Furthermore, we can distinguish whether Vg is a transcription factor or co-regulator by determining what other proteins it interacts with in the nucleus via co-immunoprecipitation and mass spectrometry (Li, Collins, et al. 2016). This will greatly enhance our knowledge of Vg's regulatory properties and can potentially unveil co-evolutionary relationships between Vg and other proteins. Nevertheless, the findings presented in this study here represent a major new discovery of Vg function.

Given Vg's numerous immunological functions and that fact that it binds to several immune-related genes, we showed that Vg cutting is sensitive to infection and that the system dynamically responds to immunological perturbations. Interestingly, Vg cutting is enhanced by incubation with heat-killed *E. coli*, but ingestion of *E. coli* by workers *reduces* the amount of Vg N-sheet subunit translocated to the nucleus. This finding suggests that bacterial material provokes Vg instability, but that Vg nuclear translocation is a tightly regulated process. From an immunological standpoint, it may be that Vg responds to bacterial infection by directly binding to and eliminating bacterial cells rather than up-regulating the transcription of additional immune-related genes, as Vg is known to bind to bacteria such as *E. coli* (Salmela et al. 2015) and act as a bactericidal molecule (Zhang et al. 2011). Follow-up studies are needed to fully elucidate Vg's immune response to bacterial infection, as the current study setting does not capture the acute changes caused by the bacterial challenge.

Vg's roles in immunity and reproduction seem to be interlinked, and we have shown before that honey bee Vg binds to bacterial surface molecules and carries them to the queen's eggs in the process of maternal immune priming (Salmela et al. 2015). Injection of dead bacteria negatively impacts oogenesis in the mosquito *Anopheles gambiae* due to altered *vg* expression (Rono et al. 2010). The injected bacteria activate the Immune deficiency (IMD) pathway, which is the key cascade in the defense against Gram-negative bacterial infection. Hence, it has been suggested that the IMD-Vg association provides the molecular basis for the trade-off between reproduction and immunity (Rono et al. 2010). In *Drosophila*, a caspase called Dredd is required for the activation of IMD (Leulier et al. 2000). Dredd has YVAD-activity (Kim et al. 2014), which makes it a candidate caspase to study further in the context of the immunological and reproductive implications of Vg cutting.

In the *in vivo* setup, the level of nuclear translocation of Vg was highly variable among individuals in both the control and the *E. coli*-fed groups. Similarly, Western blots of individual bees show variation in the strength of the 40 kDa N-sheet band (Fig S1). The haemolymph titers of full-length Vg is known to have a high level of individual variation in honey bees, which is associated with variation in stress tolerance (Seehuus et al. 2006). Similarly, variation in the nuclear Vg might be linked to certain physiological features, but this speculation needs further investigations.

Finally, we present the first 3D structure of a full-length invertebrate Vg, albeit at low resolution. The previous structural biology carried out on the Vg protein family includes an X-ray structure of lamprey lipovitellin purified from eggs (Raag et al. 1988a; Anderson et al. 1998), an NMR structure of the honey bee Vg polyserine linker

(Havukainen et al. 2012), and homology modeling of honey bee Vg (Havukainen, Halskau, Skjaerven, et al. 2011; Havukainen et al. 2013)and mammalian apolipoprotein B by others (Mann et al. 1999). The lamprey structure lacks four C-terminal regions because they are disordered and cannot be resolved with the X-ray diffraction (Anderson et al. 1998). Our electron microscopy structure shows the whole molecule, where the missing regions appear as protrusions surrounding the C-terminal area when compared to the lamprey X-ray structure. The electron microscopy structure reveals, for the first time, that the honey bee Vg is a monomer (see previous speculations in (Wheeler and Kawooya 1990)), unlike most insect (Tufail and Takeda 2008)or vertebrate Vgs (Finn 2007).Comparison between the lamprey lipovitellin X-ray structure and our insect Vg 3D electron microscopy reconstruction reveals surprisingly little differences in the N-sheet or the linker region. The heavily phosphorylated honey bee polyserine linker is 70 amino acid residues longer than the corresponding non-phosphorylated linker in the lamprey protein (Havukainen et al. 2012). In addition, there are two insect specific loops of unknown function in the N-sheet domain (11 and 19 amino acid residues long in the *A. mellifera* Vg) (Havukainen, Halskau, Skjaerven, et al. 2011). In spite of the great differences in sequence and posttranslational modifications between insect and vertebrate Vgs, they appear structurally conserved. The lamprey X-ray structure docks well into the 3D reconstruction of the honey bee Vg, and shows that both N-sheet and the polyserine linker are well-exposed to solvent (Orange and red domains in Fig. 5.4B). Such easily accessible sites are often preferred by proteases (Fontana, Laureto, et al. 2004), unless they are protected by some other means, for example, by phosphate groups (Cohen 2000). Our Vg-cutting inhibition assays suggest that dephosphorylation, indeed, may be

important for Vg cutting to occur. Furthermore, the cutting inhibition assay suggests the Vg N-sheet domain is cleaved via caspase activity (see Supplementary Information). However, we highlight that the cutting and nuclear translocation of Vg might be linked to many more traits than tested or hypothesized here. We conclude that Vg cutting might not need major conformational changes, but the phosphorylation status of Vg may play a crucial role instead.

Our discovery that the Vg N-sheet subunit binds to DNA in fat body cell nuclei raises a number of questions that warrant further research. First, does Vg translocate into the nucleus in other tissues in addition to the fat body? Vg has been verified in eggs and ovaries (Seehuus et al. 2007), in hypopharyngeal glands (Seehuus et al. 2007), in immune cells (Hystad et al. 2017), and in glial cells of the honey bee brain (Münch et al. 2015). In this latter observation, it is specifically the Vg N-sheet that is localized in glial cells, and such subcellular localization should prompt a highly relevant research subject since honey bee *vg* knock-downs show an altered brain gene expression pattern (Wheeler et al. 2013) and major behavioral changes (Nelson et al. 2007). Second, does Vg naturally translocates to the nucleus in other animals? The N-terminal Vg cutting pattern is similar in most insects studied (Tufail and Takeda 2008), but the possibility of N-sheet translocation remains speculative before it is experimentally tested in another species. Finally, what role do other Vg fragments play in nuclear translocation? In addition to the most prominent 40 kDa N-sheet fragment (Havukainen, Halskau, Skjaerven, et al. 2011), Vg is prone to what appears to be unspecific degradation (Wheeler and Kawooya 1990). Both the N-sheet-specific and whole-Vg-specific antibodies, as well as Vg purified from fat body (Havukainen et al. 2012), show weak protein bands smaller in size to full-length

Vg, most notably, bands of ~75 and ~125 kDa in size. It is unclear if these are functional Vg fragments or simply the result of unspecific fragmentation. In general, we have observed that Vg fragment number grows in harsh sample treatment conditions. We suspect that at least a 25 kDa fragment detected by the Vg antibody used here in our tissue fractioning assay is a degradation product caused by the assay protocol. However, it is not ruled out that other Vg fragments in addition to N-sheet play a role in nuclear localization of Vg.

Vg was first discovered as an egg-yolk protein, but the protein is so ancient that we cannot be sure of its original function, and it's possible that it took on a gene regulatory role early in its history. While we only touch the functional link between Vg and immunity in this study, our results will, hopefully, spark diverse future studies on the role of Vg as a putative transcription factor in honey bees and other animal species.

CHAPTER 6

VITELLOGENIN IS A DNA-BINDING PROTEIN IN HONEY BEES

Gyan Harwood, Chris Elsik, Shanshan Yang, Gro Amdam

ABSTRACT

Many proteins can bind various ligands to perform multiple different functions, including some that bind to DNA and regulate gene expression. Vitellogenin (Vg) is a highly conserved multifunctional protein that is primarily known for its role in egg formation, but which also has functions pertaining to immunity, oxidative stress relief, longevity, and behavior. In the honey bee (*Apis mellifera*), gene knockdown of Vg elicits expression changes in many other genes, suggesting that it may play a more direct role in gene regulation. We have recently shown that a structural subunit of Vg translocates to the nucleus and binds DNA, potentially acting as a transcription factor. Here, we expand on this finding by examining how Vg-DNA binding may elicit gene expression changes in the honey bee worker caste, and by identifying other nuclear proteins bound to the Vg-DNA complex that may indicate signaling pathways involved in Vg gene regulation. We find that Vg-DNA binding is associated with expression changes in dozens of genes and that the Vg-DNA complex interacts with dozens more nuclear proteins. Our results suggest Vg-DNA binding may regulate several important processes in honey bee workers, including energy metabolism, behavior, and signaling.

INTRODUCTION

It was long thought that all proteins performed a single given function (Horowitz 1995). Advances in molecular biology have shifted our understanding, and it is now believed that many proteins, if not the majority, are multifunctional. Such multifunctional proteins perform different tasks depending a number of contexts including their extracellular or subcellular localization, concentration relative to binding partners, and post-translational modifications they have undergone (Volz 2008; Gurevich and Gurevich 2015; Faust et al. 2017). Proteins' functions stem from their molecular structure (shape), which allow them to bind specifically to and interact with various ligands, including other proteins, lipids, and nucleic acids. Proteins that localize in the cell nucleus and bind nucleic acids can achieve multifunctionality by regulating expression of genes involved in many biological pathways (Chesmore et al. 2016).

Vitellogenin (Vg) is a well-documented multifunctional protein. This ancient and highly conserved protein (Baker 1988a, 1988b; Babin et al. 1999; Smolenaars et al. 2007; Wu et al. 2013) is primarily known for its role in egg-yolk formation in oviparous animals, where it transports lipids and other nutrients into developing eggs and serves as a yolk protein precursor (Pan et al. 1969; Engelmann 1979). In addition, Vg also acts as a pathogen pattern recognition receptor (Zhang et al. 2005, 2011; Li et al. 2008; Liu et al. 2009; Tong et al. 2010; Salmela et al. 2015), an antioxidant (Seehuus et al. 2006; Havukainen et al. 2013; Sun and Zhang 2015a), and a nutrient storage protein (Amdam, Norberg, et al. 2003), and plays key roles in phenotypes like behavior (Amdam and Omholt 2003; Amdam, Norberg, et al. 2003; Ihle et al. 2010; Roy-Zokan et al. 2015; Dittmer et al. 2019) and longevity (Amdam, Norberg, et al. 2005b; Ihle et al. 2015;

Salmela et al. 2016). Many of these functions are shared across a diverse collection of organisms including corals (Du et al. 2017), fish (Shi et al. 2006; Li et al. 2008, 2009; Liu et al. 2009), and insects (Havukainen et al. 2013; Salmela et al. 2015; Harwood et al. 2019). Vg is localized both intracellularly and extracellularly, being first synthesized in the liver (Wang et al. 2005), adipose tissue (fat) (Brookes 1969), or hepatopancreas (Guan et al. 2016) (depending on the organism) before being secreted into the blood or hemolymph. It can then be taken up by the ovaries and other tissues via receptor-mediated endocytosis (Noah Koller et al. 1989; Raikhel and Dhadialla 1992). Vg's multitude of functions have been the subject of much study, and there is now a fairly good understanding of how Vg's molecular structure enables it to interact with specific ligands to facilitate discrete functions like lipid transport (Raag et al. 1988b), pathogen recognition (Liu et al. 2009; Salmela et al. 2015), and oxidative stress relief (Seehuus et al. 2006; Havukainen et al. 2013). However, evidence also points to Vg playing a more direct role in regulating gene expression and molecular pathways. For example, knocking down *vg* gene expression via RNA interference (RNAi) results in expression changes to thousands of genes in the brain (Wheeler et al. 2013). While the mechanism by which Vg can affect gene expression has thus far remained a mystery, we have recently made a major discovery that sheds new light on Vg's ability to affect so many functions: In honey bees (*Apis mellifera)*, a highly conserved structural subunit of Vg is cleaved and translocated into the nucleus of fat body cells where it binds to DNA at hundreds of loci (Salmela et al, *submitted*). The consequences of this Vg-DNA binding are not yet understood, but Vg may be acting as a transcription factor or transcriptional co-regulator to affect genes involved in a broad range of biological functions.

Honey bees are a premier insect model for Vg studies owing to the key roles it plays in social behavior, immunity, nutrient storage, and ageing (see references above). A honey bee colony is composed of a reproductive queen and a functionally sterile worker caste that performs a series of age-dependent tasks required for colony survival and proliferation (Seeley 1982). Vg sets the pace of worker behavioral development, with younger workers aged 1-2 weeks having high titers of Vg and performing brood-rearing tasks (i.e., nursing), while a subsequent decline in Vg and a concomitant increase in juvenile hormone prompt workers to become foragers that collect resources required by the colony (Amdam and Omholt 2003). This behavioral change from nursing to foraging is also accompanied by a change in diet from lipid- and protein-rich pollen to carbohydrate-rich honey. Vg titers at different worker life stages can also affect behaviors like responsiveness to sucrose and preference for collecting nectar or pollen as foragers (Amdam, Norberg, et al. 2006; Wang et al. 2012, 2013). Honey bees have also been used to study Vg's effects on ageing, as both queens and overwintering workers have high Vg titers and live a relatively long time (several years and several months, respectively), while summer workers have much lower titers and typically live just 1-2 months (Amdam, Norberg, et al. 2005a; Salmela et al. 2016). Vg's role in extending lifespan stems not only from its ability to fight pathogens and act as a nutrient store, but also from its ability to neutralize reactive oxygen species (ROS) and lessen the cellular damage that leads to senescence (Seehuus et al. 2006; Havukainen et al. 2013). With a new understanding of Vg nuclear translocation and DNA binding, we can begin to investigate how Vg's role as a potential transcription factor may contribute to differences observed among the honey bee worker caste.

In this study, we used a multi-pronged approach to compare Vg's actions in the cell nuclei of nurses and foragers. We limited confounding age-effects by comparing age-matched nurses and foragers sampled from single-cohort colonies. Specifically, 1-day old workers are paired with a queen, and a subset of these workers prematurely transition into foragers as the nutritional demands of the colony grow. Our aims in this study were threefold. First, we mapped out Vg-DNA binding sites using chromatin immunoprecipitation followed by sequencing (ChIP-seq) to determine which sites are shared or unique to nurses and foragers. We analyzed the genomic distribution of these binding sites to determine whether they are preferentially found in promotor regions (i.e., where transcription factors typically bind), and we performed a motif analysis to look for specific Vg-DNA binding sequences. We compared these motifs with known transcription factor binding motifs from *Drosophila melanogaster,* as the genetic architecture of this marquee insect model is more fully worked out. Second, we used RNA-seq to investigate whether Vg-DNA binding corresponds to gene expression differences in nurses and foragers, and we performed Gene Ontology (GO) term analyses to determine whether ChIP-seq and RNA-seq gene lists are enriched for specific biological functions. Finally, we used co-immunoprecipitation followed by mass spectrometry to identity additional nuclear proteins that are bound to the Vg-DNA complex. Given that transcriptional regulation can involve a complex interaction between transcription factors, co-regulators, enzymes, signaling molecules (Bondos and Tan 2001), the aim here is to elucidate regulatory pathways that may contributing to Vg-DNA binding and subsequent gene expression changes.

Our results suggest that Vg's DNA- and protein-binding in the nucleus regulates many genes involved in energy metabolism, behavior, and signal transduction pathways. This study represents the first comprehensive investigation of Vg's potential gene regulatory roles in honey bees, particularly as it relates to differences among different task groups in the worker caste. Given Vg's ubiquity in metazoans and its conserved sequence and structure, this study should open up a broad new field of research to examine similar functions in Vg and other related proteins across a large spectrum of animal taxa.

METHODS

Bees

Bee stocks were maintained at the Arizona State University Bee Research Facility in Mesa, Arizona, USA. We established 3 single-cohort colonies and treated each as a separate biological replicate owing to the high degree of relatedness among the nestmates. To make these colonies, brood frames from established hives were placed in an incubator overnight at 34°C and 50% humidity, and the next morning roughly 2,000-3,000 newly emerged were transferred into each of 3 "nuc" hives along with a caged queen and frames of honey, pollen, and empty comb (Scheiner and Amdam 2009; Amdam et al. 2010). The bees in each of the single-cohort colonies originated from different donor hives. The hives were sealed for 3 days, after which they were opened and the queen was released from her cage. In single cohort colonies, some of the workers will prematurely transition into foragers in order to collect resources for the colony, and on the 7th day after establishing the colonies we paint marked new foragers that were

returning to the hive with pollen. On the 14th day, we collected N=50 paint marked foragers and an equal number of nurses from each colony, meaning all foragers and nurses had been performing their given task for at least 7 days. Nurses were identified when they entered brood cells to feed larvae. Collected bees were anesthetized on ice before having their fat body dissected, flash frozen in liquid nitrogen, and stored in -80°C. From each of the three single-cohort colonies, we pooled fat body samples from 25 nurses or 25 foragers. Pooled samples were homogenized in a mortar and pestle with liquid nitrogen, and the resulting homogenate was divided into two workflows for ChIP-seq and RNA-seq.

Chromatin immunoprecipitation

For the ChIP-seq pipeline, we followed previously established protocols (Bai et al. 2013). Briefly, the homogenate was transferred to a 15mL dounce homogenizer tube on ice with 4mL of 1% formaldehyde in 1x phosphate-buffered saline (PBS) to crosslink DNA and proteins in the sample. The crosslinking was quenched after 20 mins by adding glycine (final concentration 125mM). The homogenate was transferred to a 15 mL tube and centrifuged at 1500g for 3 mins at 4°C, and the resulting pellet was washed 3 times with 1x PBS and protease inhibitor cocktail (Roche cOmplete™). We then washed the pellet once with cell lysis buffer (NaCl 100mM, HEPES [pH 7.6] 5mM, EDTA 1mM, NP-40 0.5%) and protease inhibitor and centrifuged it at 1500g for 5 mins at 4°C. After discarding the supernatant, we resuspended the pellet in 600 µl nuclear lysis buffer (HEPES [pH 7.6] 50mM, EDTA 10mM, Na-deoxycholate 0.1%, N-lauroylcosanine 0.5%) and sonicated the sample with a QSonica Q800R2 sonicator for a total of 5 mins

(15 s on, 20 s off, amplitude = 20%). We checked the sonicated sample on an agarose gel to confirm that the DNA was sheared to a size of 300-500 bp.

The Vg-DNA complex was precipitated using antibodies specific to the whole honey bee Vg protein (Pacific Immunology, Ramona, CA) and whose specificity to Vg has been verified in several studies (Münch et al. 2015; Harwood et al. 2019). We first conjugated these antibodies to magnetic Dynabeads™ (Invitrogen™ #10001D, Protein A). To do so, we removed the magnetic beads from their buffer and washed them 4 times with 1 mL blocking buffer (1x PBS with 5% bovine serum albumin [BSA, Jackson ImmunoResearch #001-000-161]) before adding 500 µl block solution and 20 µl antibodies to the beads and rotating overnight at 4°C. We then washed the antibody-bead complex 5 times with blocking buffer, then added 500 µl of the chromatin extracts and 500 µl of dilution buffer with protease inhibitor (SDS 0.01%, Triton x100 1%, EDTA 1.2mM, Tris-HCl [pH 8.0] 16.7mM, NaCl 167 mM) and rotated at 4°C overnight. Separately, we used 50 µl of chromatin extracts + dilution buffer as *input DNA*. These samples were incubated with magnetic beads that had undergone the same washing steps, except they had not been conjugated with the antibodies. The next day, all samples were washed with a low-salt buffer (SDS 0.01%, Triton x100 1%, EDTA 2mM, Tris-HCl [pH 8.0] 20mM, NaCl 150mM), a high-salt buffer (SDS 0.01%, Triton x100 1%, EDTA 2mM, Tris-HCl [pH 8.0] 20mM, NaCl 500mM), a LiCl buffer (LiCl 0.25M, NP40 1%, Na-deoxycholate 1%, EDTA 1mM, Tris-HCl [pH 8.0] 10mM), and TE buffer (EDTA 1mM, Tris-HCl 10mM). The samples were then eluted by suspending the beads in 200 µl elution buffer (EDTA 10mM, Tris-HCl [pH 8.0] 50mM, SDS 1%) and heating in a water bath at 65°C for 15 mins. The magnetic beads were discarded and the remaining solution

was reverse cross-linked by incubating at 65°C overnight on a heating rack. To purify the DNA in the samples, we added 200 µl of TE buffer and 8 µl of RNase A (LifeTech #12091021) and incubated at 37°C for 30 mins before adding 8 µl of Proteinase K (LifeTech #25530049) and incubating at 55°C for 1 hr. We then used phenol:chloroform:isoamyl (LifeTech #15593-031) and a heavy phase lock gel tube (5 PRIME #2302810) to separate the aqueous DNA from any contaminants. Finally, we used 16 µl glycogen (LifeTech #AM9510), 40 µl of sodium acetate (LifeTech #R1181), and 800 ul of 100% EtOH (Fisher #04355222) to precipitate the DNA for 1 hr at -80°C before centrifuging at 16,000g for 30 mins at 4°C. We washed the resulting pellet with 70% EtOH and allowed it to airdry before resuspending it with 50 µl of TE buffer. We measured the concentration of DNA on an Invitrogen Qubit™ Fluormenter 2.0.

DNA library prep, sequencing, and annotation

DNA samples (ChIP DNA + Input DNA) were submitted to the Biodesign DNASU Sequencing Core at Arizona State University. Illumina compatible libraries were generated on the Apollo 384 liquid handler using KAPA Biosystem's LTP library preparation kit (KAPA KK8232). Genomic DNA was sheared to approximately 400-600bp fragments using Covaris M220 ultrasonicator, then all samples were end repaired and A-tailed as described in the KAPA protocol. Illumina-compatible adapters with unique indexes (IDT #00989130v2) were ligated on each sample individually. The adapter ligated molecules were cleaned using Kapa pure beads (Kapa Biosciences, KK8002), and amplified with Kapa's HIFI enzyme (KK2502). Each library was then analyzed for fragment size on an Agilent's Tapestation, and quantified by qPCR (KAPA

Library Quantification Kit, KK4835) on Thermo Fisher Scientific's Quantstudio 5. Libraries were then multiplexed and sequenced on 2x75 flow cell on the NextSeq500 platform (Illumina) at the ASU's Genomics Core facility.

The raw Illumina 2x75bp pair-end reads were quality checked using FastQC v0.10.1, followed by adapter trimming and quality clipping by Trimmomatic 0.35. Any reads with start, end or the average quality within 4bp window falling below quality scores 18 were trimmed. The clean reads were aligned to reference genome Apis mellifera Amel_HAv3.1 (ht https://www.ncbi.nlm.nih.gov/assembly/GCF_003254395.2/) by Bowtie 2 version 2.2.9. Library insert size was checked by Picard Tool (https://broadinstitute.github.io/picard/). Library complexity was checked by NRF (nonredundancy fraction), defining as number of unique start positions of uniquely mappable reads divided by number of uniquely mappable reads. IGVtools and bamCompare from deepTools were employed for compare two BAM files based on the number of mapped reads. First the genome is partitioned in to bins of equal size and then the number of reads in each bin is counted. The log2 value for the ratio of number of reads per bin of each sample was reported for IGV visualization and compared between each pair. With 95% correlation, three biological replicates were combined for peak identification. MACS2 was used for peaks calling with 0.05 FDR cutoff. Narrowpeak files as MACS2 output for each individual replicates and combined samples were annotated by HOMER. It first determines the distance to the nearest transcription start site (TSS) and assigns the peak to that gene. Then it determines the genomic annotation of the region covered by the center of the peak, including promotor (1kb

upstream to 100bp downstream of TSS), transcription termination site (TTS), Exon (Coding), 5' UTR Exon, 3' UTR Exon, Intronic, or Intergenic.

Genome region and motif analysis of ChIP-seq binding sites

To determine whether Vg-DNA binding sites are more likely than chance to be located in promotor regions, we first created a null distribution of ChIP loci. To do this, we took 1000 hypothetical ChIP peaks of 200 bp long and randomly distributed them throughout the genome using HOMER, and then repeated this procedure for 1000 iterations. We then examined which types of genomic regions were at the center of these peaks using the same default classifications in HOMER as above. A Chi-square test was then used to determine whether to compare this null distribution with the distributions found in our ChIP-seq data.

Motif analysis

For motif discovery, we used the DREME feature (Bailey 2011) within the MEME suite of tools (Bailey et al. 2009) to look for short, ungapped motifs of 8bp within samples. We then compared the these motifs to known *Drosophila* motifs using TOMTOM (Gupta et al. 2007).

RNA extraction

Using the same homogenized samples as with the ChIP procedure, a portion of the homogenate went into an RNA extraction workflow. Here, use followed standard phenol RNA extraction protocols using TRIzol™ Reagent (Invitrogen™ #15596026),

chloroform (Alfa Aesar #32614) and isopropanol (LabChem #LC157501). The resulting RNA pellets were air-dried, re-suspended with nuclease-free water, and subject to DNAse treatment (Invitrogen™ Turbo DNA-free™ #AM1907). We measured the RNA concentration using a NanoDrop and then diluted each sample a working concentration of 200 ng/µl.

RNA library prep, sequencing, and annotation

Using KAPA's mRNA HyperPrep Kit (KAPA #KK8580), mRNA sequencing libraries were generated from total RNA. Magnetic oligo-dT beads were used to capture mRNA specifically, and the mRNA was sheared to approximately 150-200bp in length using heat and magnesium. The 1st strand of the mRNA fragments were reverse transcribed using random priming. The 2nd strand was generated with incorporated dUTP molecules, and dAMP was added to the 3' ends of the double-stranded cDNA molecules. Illumina-compatible adapters with unique indexes (IDT #00989130v2) were ligated on each sample individually. The adapter ligated molecules were cleaned using KAPA Pure beads (KAPA #KK8002), and amplified with Kapa's HIFI enzyme (KAPA KK2502). The strand marked with dUTP is not amplified, allowing for strand-specificity. Each library was then analyzed for fragment size on an Agilent Tapestation, and quantified by qPCR (KAPA KK4835) on Thermo Fisher Scientific's Quantstudio 5 before multiplex pooling and sequencing a 1x75 flow cell on the NextSeq500 platform (Illumina) at the ASU Genomics Core facility.

The *Apis mellifera* transcriptome and Annotation Release 104 from NCBI (derived from genome Amel_HAv3.1) were used for quasi-mapping and count

generation. FASTQC1 (version 0.11.8) was used on each sample for quality check. Average per-base read quality scores were over 30 in all samples and no adapter sequences were found, indicating high quality reads. We used Salmon3 version 0.13.1 to quasi-map reads to the transcriptome and quantify abundance of each transcript. The transcriptome was first indexed, then quasi-mapping was performed to map reads to transcriptome using additional arguments --seqBias and --gcBias to correct sequence-specific and GC content biases and --numBootstraps=30 to compute bootstrap transcript abundance estimates. Gene-level counts were then estimated based on transcript-level counts using the "bias corrected counts without an offset" method from the tximport package. 84 – 90% of reads mapped to the transcriptome (41.1-62.2 million per sample) and were kept for statistical analysis. We used the TMM (trimmed mean of M values) normalization in the edgeR package to adjust for possible biased in RNA composition, such as reads mapping to viral genomes. Normalization factors ranged from 0.52 and 1.35, but variation was between individuals, not caste or hives, suggesting no group-level difference in RNA composition. Samples with a higher proportion of reads mapping to viral genes tended to have lower TMM normalization factors, to account for the smaller number of reads mapping to the *Apis* transcriptome. The NCBI Amel_v3.1 Annotation Release 104 contains 12,090 genes, but not all are expressed in our samples at detectable levels. We set a detection threshold of 0.5 CPM (counts per million) in each sample, resulting in 9,134 genes detected that accounted for 99.95% of all reads. After filtering, TMM normalization was performed again and normalized log2-based CPM values were calculated using edgeR's cpm() function with prior.count = 3 to help stabilize fold-changes of extremely low expression genes. Multidimensional scaling in the limma7

package was used to identify potential treatment effects at higher level. Testing for differentially expressed genes (DEGs) was performed using Limma-trend methods. We identified 468 DEGs between nurses and foragers using a (Benjamini-Hochberg) FDR < 0.05 as a cutoff.

Protein immunoprecipitation and mass spectrometry

To pull down other proteins that are interacting with Vg and DNA, we integrated our ChIP procedure with a protocol for immunoprecipitation of chromatin-interacting protein complexes (Mohammed et al. 2016). As with the ChIP protocol, we pooled bee tissue (N=10) from each caste within colonies and crosslinked the homogenate with 1% PFA. We quenched the homogenate, washed it, lysed the cellular membranes, lysed the nuclear membranes, and then sheared the samples via sonication. We bound our anti-Vg antibodies (20 µl) to 100 µl of magnetic Dynabeads™ (Protein A) then incubated them with our samples. As a control, we incubated samples with Dynabeads™ that had not been bound with anti-Vg antibodies, so any proteins precipitated in these samples would be due to proteins interacting directly with the Dynabeads™. If any protein was identified in both the treatment and control samples it was deemed a false positive and removed from further analysis. The remainder of the procedure followed Mohammed et al.'s protocol (Mohammed et al. 2016), whereby the bead samples were washed first with RIPA buffer (HEPES [pH 7.6] 50 mM, EDTA 1mM, Na-deoxycholate 0.7%, NP-40 1%, and LiCl 0.5M) then with ammonium bicarbonate, then subject to trypsin digestion, solid phase extraction, and washing, before being loaded onto an AB Sciex 4800 mass

spectrometer for MALDI-TOF/TOF-MS (matrix-assisted laser desorption/ionization – tandem time of flight – mass spectrometry).

Proteomics analysis was performed on an Orbitrap Fusion Lumos Tribrid mass spectrometer with an Ultimate 3000 nano-LC and nanoelectrospray ionization. Peptides were separated with a nC18 analytical column (C18 Pepmap 100, 3 µm particle, 100 Å pore, 75 µm i.d. ×150 mm) using 150 min buffer gradient a low flow rate at 300 nL/min. Data-dependent acquisition in positive mode was performed for data collection. Acquired data was searched with Proteome Discoverer 2.2 using the SEQUEST search engine with label-free quantification workflow against the UniProt database of Apis mellifera (http://www.uniprot.org; Protome ID: UP000005203). Search parameters was trypsin cleavage sites with a 2 missed cleavage site allowance, precursor and fragment mass tolerance was set at ±10ppm and 0.6 Da. Carbamidomethyl of cysteine was set as a fixed modification, and oxidation of methionine as a variable modification. To consolidate a master list of Vg-interacting proteins, we first removed any proteins identified in the control precipitations, and restricted our analyses to proteins identified in at least 2 of the 3 samples per caste. We also removed any classes of proteins identified as potential contaminants in previous studies (Guo et al. 2009; Mellacheruvu et al. 2013).

Gene ontology

All gene lists from ChIP-seq, RNA-seq, and proteomics analyses were searched for enrichment of specific Gene Ontology terms with HymenopteraMine (v1.4) (Elsik et al. 2016) using Official Gene Set 3.2 (OGSv3.2) as the background. Specifically, we looked for enrichment of biological processes, molecular functions, and KEGG

pathways. We used a Benjamini-Hochberg-adjusted $P < 0.05$ to determine significantly enriched terms.

RESULTS

Vg-DNA binding

We found an average of 944 Vg-DNA binding sites in nurses (range: 901-1020) and 930 in foragers (range: 867-1049) that were aligned to the honey bee genome. These binding sites were found in promotor regions with a higher probability than chance, with an average of 25% and 27%, respectively, found within a 2kb window upstream of a transcription start site compared to 12.5% of sites observed here in a null distribution of randomly generated ChIP peaks ($\chi^2 = 193.92$, df = 1, $p < 0.001$, $\chi^2 = 137.47$, df = 1, $p < 0.001$) (Fig. 6.1A) [*Disclaimer: the null distribution was generated using the previous honey bee genome assembly Amel_4.5 and is currently being regenerated to the latest Amel_HAv3.1 genome by a collaborator*]. We trimmed away redundant binding sites annotated to the same gene and only considered genes that were shared by at least 2 of the 3 biological replicates per caste, leaving a total of 596 genes, hereafter referred to as *ChIP genes* for clarity. Of these ChIP genes, the majority were shared between nurses and foragers, while 188 and 96 were found only in nurses or foragers, respectively (Fig. 6.1B). A GO term analysis did not find significant enrichment at the BH-adjusted P-value level for any biological processes, molecular functions, or known pathways, but there was enrichment at the unadjusted P-value level for "signaling receptor activity" in nurses (GO: 0038023, unadjusted P = 0.024) and "methyltransferase activity" in foragers (GO: 0008168, unadjusted P<0.001).

Binding motifs

We next performed a motif analysis to find specific DNA sequences to which Vg binds, and to compare these sequences to known *Drosophila melanogaster* transcription factor binding motifs. We found significant enrichment for 55.3 motifs in nurses (range: 40-65 per sample) and 59 motifs in foragers (range: 45-69), indicating common DNA sequences to which Vg binds. The most highly enriched motifs for nurses and foragers are depicted in Fig. 6.1C. These motifs were similar to several known drosophila transcription factor binding sites. [*Disclaimer: These analyses are still being performed and a more comprehensive report will be included in the published manuscript*].

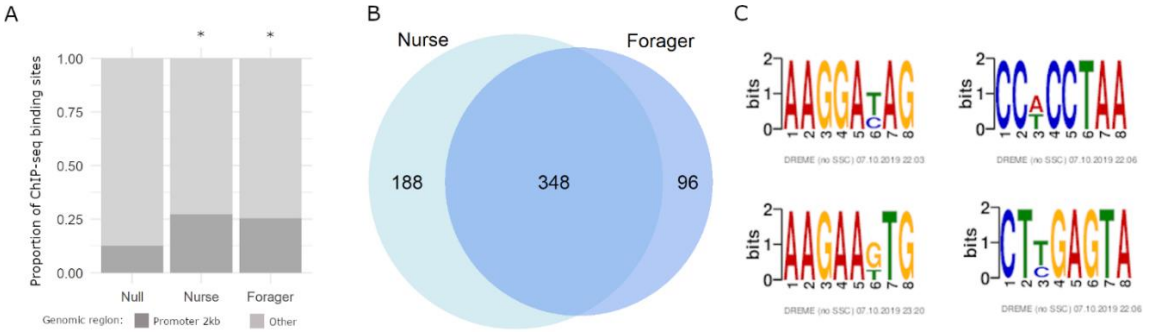

Fig. 6.1 ChIP-seq revealed 944 and 930 Vg-DNA binding sites in nurses and foragers, respectively. **A**: The distribution of ChIP-seq binding sites in promoter regions compared to other genomic regions. A promoter region is defined as a window 2kb upstream to 100bp downstream from a transcription start site. Both nurses and foragers had significantly more sites in promoter regions compared to a null distribution of randomized ChIP-seq peaks, as determined by χ^2 Goodness of Fit tests. **B**: ChIP-seq sites

were annotated to 596 separate genes. Nurse and forager samples shared peaks in a majority of these genes, while some were restricted to only nurses or foragers. **C**: Logos for the 4 most common binding motifs observed in all nurse and forager samples. The motifs are 8 bp long, with letters denoting the nucleotide sequence found therein (A, T, C, G).

Differential gene expression

RNA-seq analysis generated 41 – 62 million mappable reads per sample, detecting 9,134 genes above a threshold cutoff of 0.5 CPM. Variation between castes and hives was examined using multidimensional scaling of logCPM values of the 5000 most variable genes. All nurse samples were found to be similar to each other and clustered closely together, while forager samples showed higher inter-colony variation (Fig. 6.2A). The cause of this increased variation among foragers is unknown but may be due to viral infection as foragers had a higher proportion of their counts map to two known viral genomes (Table S1). Differential expression testing yielded 468 DEGs between nurses and foragers using a 1-way ANOVA with an FDR $P<0.05$, of which 342 genes had a $\log_2$ fold change greater than 1 (Fig. 6.2B). As expected, Vg was one of the most differentially expressed, showing levels nearly 8x higher in nurses than foragers ($\log_2$FC 2.97, BH-adj. $P = 0.009$). GO term analysis of all DEGs showed enrichment for several biological processes, molecular functions, and KEGG pathways, including "oxidation-reduction process" (GO:0055114, BH-adj. $P < 0.001$), "catalytic activity" (GO:0003824, BH-adj. $P < 0.001$), and many metabolic pathways (Table 6.1).

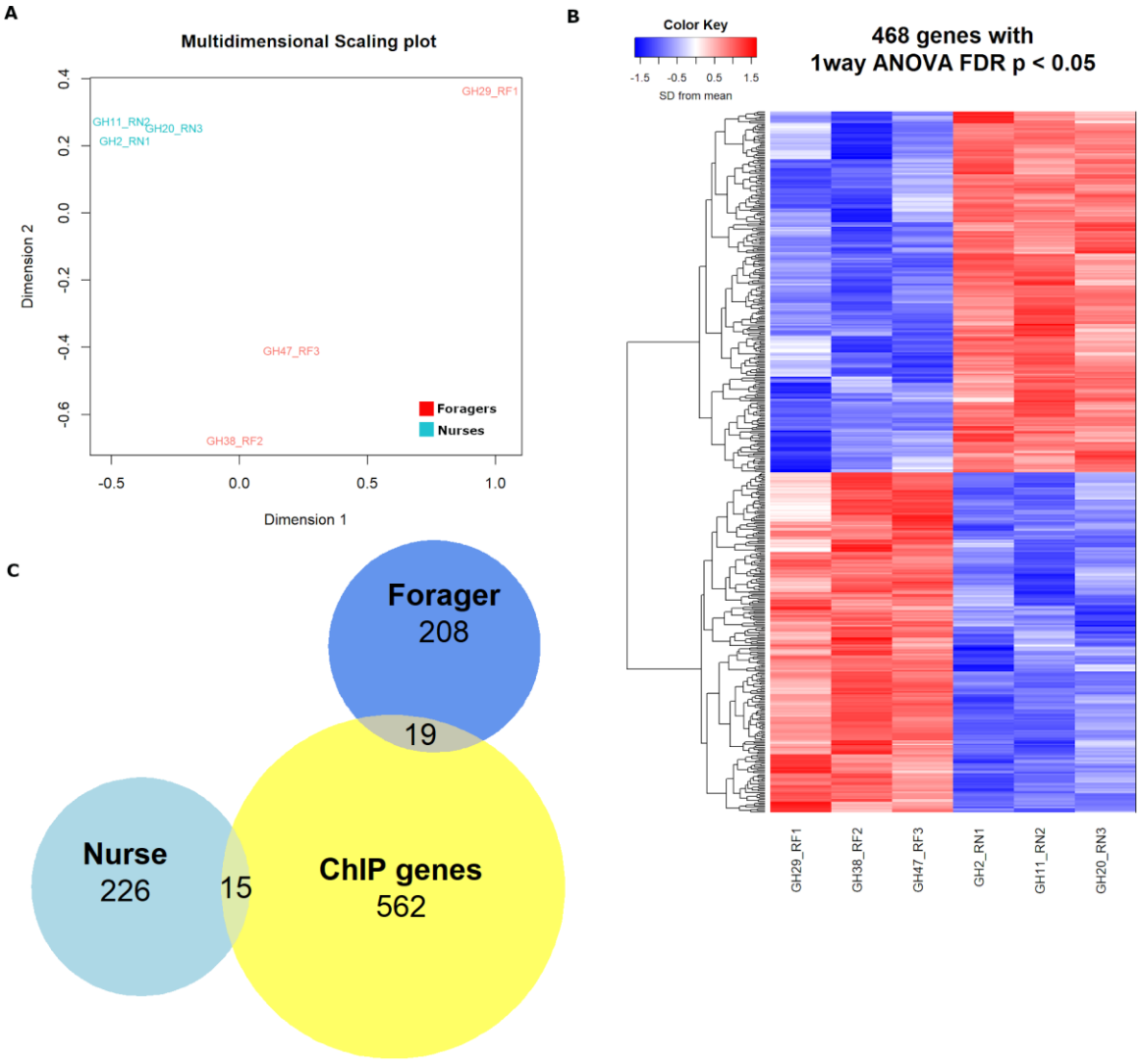

Fig. 6.2 RNA-seq expression results from nurses and foragers, and their overlap with ChIP-seq results. **A**: Multidimensional scaling of nurse and forager samples comparing variance of logCPM values for the 5000 most variable genes. The first dimension explains 47% of the variance and separates nurses from foragers, as well as accounts for some variance between forager samples. The second dimension explains 25% of the variance and accounts for additional variance among forager samples. **B**: Heat map of genes that were differentially expressed (DEGs) between nurses and foragers (FDR<0.05). The three columns on the left are from foragers samples, the three on the

right are from nurses. Blue and red coloration indicate the degree to which a given gene was up- or down-regulated (s.d. from mean). **C**: Euler plot depicting the overlap in DEGs upregulated in nurses or foragers with genes identified as Vg binding targets via ChIP-seq. Approximately 7% of genes bound by Vg show differential expression between nurses and foragers.

Overlapping ChIP genes and DEGs

We found that 34 genes bound by Vg also showed differential expression (Fig. 6.2C), indicating these as candidate genes for direct transcriptional regulation by Vg (Fig. 6.3). This overlap of ChIP genes and DEGs represents ~7% of all Vg-DNA binding sites, which is in line with other studies of transcription factors (Pilon et al. 2011; Pfeiffer et al. 2014; Zhan et al. 2018). Overlapping ChIP genes and DEGs showed enrichment for the GO term "heparan sulfate sulfotransferase activity" (GO:0034483, BH-adj. P = 0.005) (Table 6.1). Heparan sulfate is polysaccharide that binds to a variety of proteins and performs many functions pertaining to signal transduction, genomic regulation, and viral infection (Schubert et al. 2004; Christianson and Belting 2014; Stewart and Sanderson 2014; Gao et al. 2019). This term was also enriched when only considering genes upregulated in foragers, but there were no enriched terms for genes upregulated in nurses only. The Vg-bound genes that show differential expression in nurses or foragers are involved in a broad set of biological functions. These include several receptors with known effects on behavior and signaling, a number of enzymes involved in oxidation-reduction pathways, and an important antimicrobial peptide used to fight infection (*Defensin-1*).

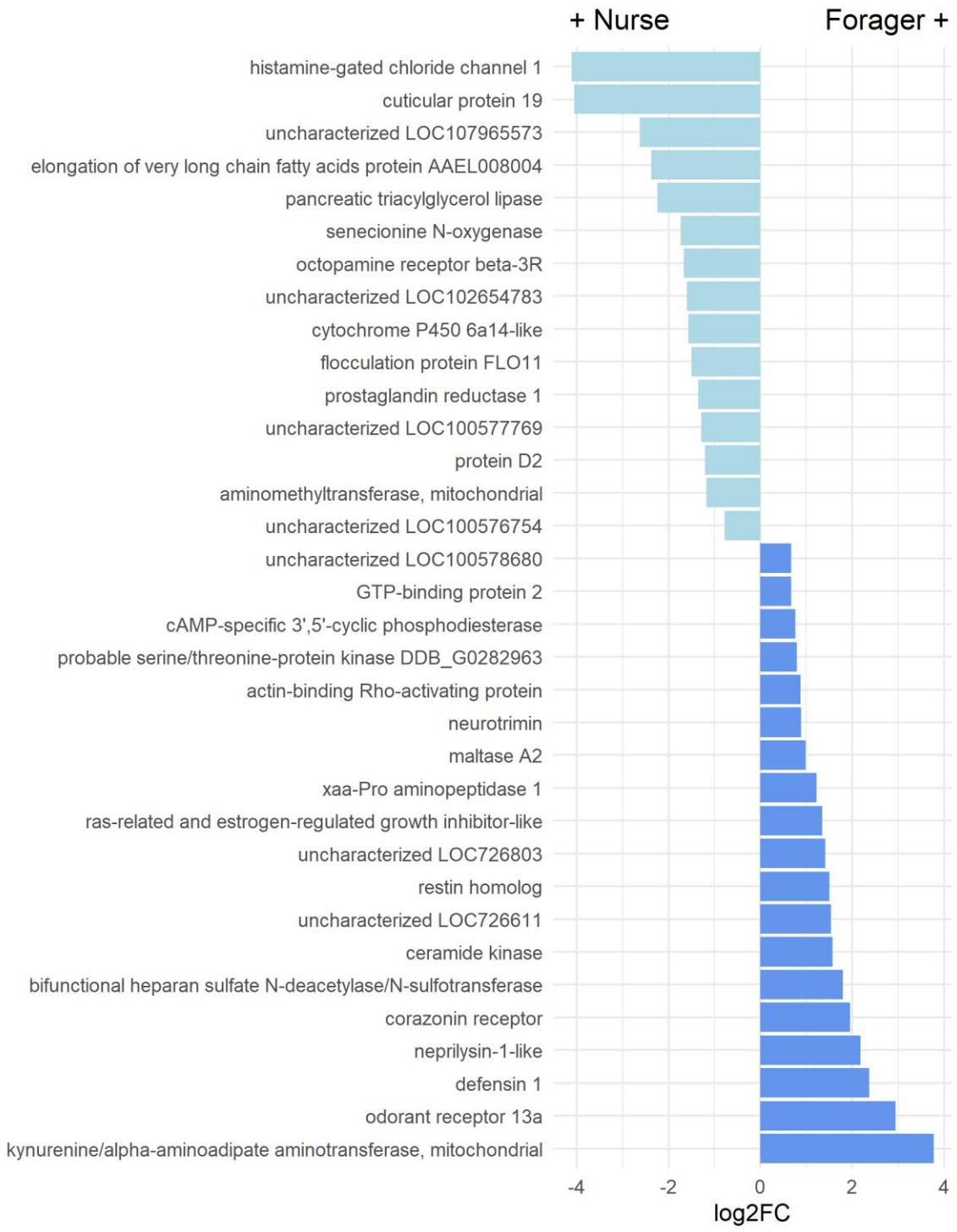

Fig. 6.3 Candidate genes for Vg transcriptional regulation. These are ChIP-genes that show differential expression between nurses and foragers (FDR<0.05). The x-axis shows

the $\log_2$ fold change, with nurse-upregulated genes to the left and forager-upregulated genes to the right.

Mass spectrometry of Vg-bound nuclear proteins

Our co-immunoprecipitation assay identified a total of 163 Vg-bound nuclear proteins that were not observed in our control precipitations. We restricted our downstream analysis to a proteins identified in at least 2 of the 3 biological replicates per caste and removed classes of proteins identified in previous studies as potential false positives, including ribonucleoproteins, histones, and ribosomal proteins (Mellacheruvu et al. 2013), leaving a total of 43 proteins. Here, we again found significant enrichment for the biological process "oxidation-reduction process" (GO: 0055114, BH-adj. $P < 0.001$), along with several metabolic processes and a tryptophan metabolic pathway (ame00380, BH-adj. $P = 0.028$) (Table 6.1). Additionally, we found another connection between Vg-bound genes differentially expressed in nurses or foragers and the nuclear proteins bound to Vg: the most abundant protein identified was β-glucuronidase, an enzyme involved in the hydrolysis of heparan sulfate that is known to co-localize with heparan sulfate in the nucleus and be enzymatically active there (Schubert et al. 2004) (Fig. 6.4). The other most abundant proteins include melittin, the pain-producing component of bee venom that has been shown to be upregulated during pathogen infection and display antimicrobial properties (Doublet et al. 2017), and ferritin, a multifunctional protein known to play important roles in storing and transporting iron, regulating apoptosis (Cozzi et al. 2003), and mediating the antioxidant and protective properties of an important immune-related transcription factor NF- κB (Pham et al.

2004). In total we identified 7 mitochondrial proteins bound to Vg in the nucleus, all but one of which are directly involved in oxidation-reduction processes.

Fig. 6.4 Identification and abundance of Vg-bound proteins from the nucleus of fat body cells. Abundance is quantified using the Normalized Spectral Abundance Factor (NSAF), which compares the number of peptide spectrum matches over the amino acid length of the protein, with the assumption that more lengthy proteins are more likely to have more matches.

Table 6.1: GO term enrichment for ChIP-seq, RNA-seq, and nuclear proteins

Description	Term ID	Adj. P Value	Number of Genes or Proteins
RNA-seq (all DEGs)			
oxidation-reduction process	GO:0055114	2.26E-13	57
aromatic amino acid family metabolic process	GO:0009072	4.68E-02	5
cofactor binding	GO:0048037	8.12E-08	37
iron ion binding	GO:0005506	1.95E-05	16
catalytic activity	GO:0003824	2.33E-05	151
oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen	GO:0016705	5.47E-05	16
coenzyme binding	GO:0050662	1.89E-04	21
monooxygenase activity	GO:0004497	6.48E-04	12
heme binding	GO:0020037	1.66E-03	14
tetrapyrrole binding	GO:0046906	1.68E-03	14
flavin adenine dinucleotide binding	GO:0050660	1.41E-02	10
oxidoreductase activity, acting on CH-OH group of donors	GO:0016614	1.97E-02	10
chitin binding	GO:0008061	3.04E-02	9
Valine, leucine and isoleucine degradation	ame00280	6.25E-05	9
Tryptophan metabolism	ame00380	2.05E-04	7
Fatty acid degradation	ame00071	2.64E-04	7
Phenylalanine metabolism	ame00360	1.05E-03	4
Glycine, serine and threonine metabolism	ame00260	4.12E-03	6
beta-Alanine metabolism	ame00410	5.99E-03	5
Drug metabolism - other enzymes	ame00983	6.37E-03	6
Fatty acid elongation	ame00062	9.26E-03	4
Glyoxylate and dicarboxylate metabolism	ame00630	1.04E-02	5
Tyrosine metabolism	ame00350	1.12E-02	4
Pyruvate metabolism	ame00620	3.66E-02	5
Lysine degradation	ame00310	4.41E-02	5
Overlapping ChIP-seq sites and DEGs			
heparan sulfate sulfotransferase activity	GO:0034483	5.42E-03	2
Vg-bound nuclear proteins			
oxidation-reduction process	GO:0055114	3.02E-05	13
organic acid metabolic process	GO:0006082	2.97E-02	7
oxoacid metabolic process	GO:0043436	2.97E-02	7
generation of precursor metabolites and energy	GO:0006091	4.20E-02	5
carboxylic acid metabolic process	GO:0019752	4.33E-02	7
Tryptophan metabolism	ame00380	2.88E-02	3

DISCUSSION

Vitellogenin has been the subject of much scientific interest over the last 50 years not only for its numerous functions, but also for its highly conserved structure across disparate taxa. The 40 kDa Vg subunit that we have examined here contains the N-terminal domain, which is present in Vg-orthologs of non-oviparous species, such as mammalian apolipoprotein B (ApoB) and microsomal triglyceride transfer protein (MTP) (Baker 1988b; Segrest et al. 1999). Results from earlier gene knockdown studies indicated that Vg may play a more direct role in regulating gene expression (Wheeler et al. 2013). In this study we sought to further elucidate Vg's potential gene regulatory role by examining its binding to DNA and nuclear proteins and to explore how these binding interactions may contribute to differences observed in the honey bee worker caste. We found that Vg binds to over 600 genes and is more likely than chance to be found in promotor regions, where transcription factors typically bind (Gill 2001). We also found that Vg-DNA binding is associated with differential gene expression in 34 genes of nurses and foragers, and that over 40 additional nuclear proteins bind to the Vg-DNA complex. While large data sets from sequencing and mass spectrometry studies will inevitably yield a plethora of potential biological relationships to explore, there are several commonalities shared among our data sets that suggest possible biological processes under Vg control, some of which may contribute to differences observed between honey bee nurses and foragers. These include energy metabolism and subsequent ROS production, behavior, and signaling.

Both the list of all DEGs between nurses and foragers and the list of nuclear proteins bound to the Vg-DNA complex are most highly enriched for the term

"oxidation-reduction process". Oxidation-reduction processes, or *redox* reactions, are enzymatic reactions used to transfer electrons between chemicals and are often associated with cellular respiration in mitochondria, where nutrients are oxidized in the production of ATP (Handy and Loscalzo 2012). That nurses and foragers would show differential expression of redox genes is certainly expected given that these castes consume different diets and so must employ different enzymatic machinery to metabolize them. Another recent study found that diet-induced DEGs in workers are also most highly enriched for "oxidation-reduction" processes (Azzouz-Olden et al. 2018). However, what is far less expected is the high number of redox proteins that are found bound to the Vg-DNA complex. This suggests that enzymes used in redox reactions may also interact with Vg as part a signaling pathway that influences gene expression. Half of the redox proteins we found bound to the Vg-DNA complex are also mitochondrial proteins, which suggests this signaling might originate from the mitochondria. Mitochondria are essential organelles for maintaining homeostasis and mitochondria-to-nucleus signaling can be used to affect expression of nuclear genes that reconfigure metabolic pathways (Cantó et al. 2015; Smith et al. 2018) and reduce oxidative stress from reactive oxygen species (ROS) (Khan et al. 2018), among other functions. There are several mitochondria-to-nucleus signaling pathways and one of the ChIP genes we identified that was upregulated in foragers, *ceramide kinase*, is part of a retrograde signaling pathway controlling apoptosis (Jazwinski 2015). ROS also play a prominent role in signaling. While the ROS produced from cellular respiration can cause cellular damage and ultimately shorten lifespan, they can also activate various signaling molecules, such as transcription factors like NF-κB that upregulate antioxidant and DNA-repair genes (Storz et al. 2005; Storz

2006). One possible explanation for this apparent relationship between Vg and redox enzymes is that Vg plays some as-yet-undermined role in mitochondria-to-nucleus signaling, though there is no solid evidence yet. Still, when one considers the mitochondria's central roles in energy metabolism, ROS production, nuclear signaling, and ageing in light of Vg's roles in those same processes, it hints there may be a closer functional and evolutionary relationship between Vg and the mitochondria than previously thought. Other proteins known to play important roles in both mitochondrial function and gene regulation include STAT3, which can localize in mitochondria and perform oxidative phosphorylation functions and also translocate to the nucleus as a transcription factor to regulate many processes like cell proliferation and apoptosis (Wegrzyn et al. 2009). This possible connection between Vg and the mitochondria should warrant further research.

Our results show that Vg-DNA binding is associated with expression changes in several behavior-related genes. *Histamine-gated chloride channel 1*, which is highly upregulated in nurses, has been shown to synchronize activity during light-dark cycles in *drosophila* (Alejevski et al. 2019), while *odorant receptor 13a* is speculated to play a role in detecting brood pheromone in honey bees (Oxley et al. 2008). But perhaps the most intriguing finding was that foragers upregulated expression of the corazonin receptor. Corazonin is a neuropeptide that, in *Harpagnathus* ants, has a co-repressive regulatory relationship with Vg to control caste identity and behavior (Gospocic et al. 2017). In *Harpagnathus* colonies, when a queen dies workers compete to become reproductives, or *gamergates*, and as the victors transition from hunting activity to egg-laying activity a drop in their corazonin expression causes an increase in *vg* expression. Conversely,

hunting behavior can be rescued in reproductives by injecting corazonin or knocking down *vg*, while knocking down the corazonin receptor decreases hunting behavior. In honey bees, a similar relationship may occur in reverse as workers transition from a quasi-reproductive state (i.e., exhibiting brood care) to foraging tasks. Here, high levels of Vg in nurses may repress corazonin receptor expression, while a reduction of Vg in foragers may release or promote expression of the receptor. What is revealed by our study is that Vg appears to play a direct role in regulating expression of one of its interactive partners.

In terms of signaling, there are several pieces of evidence suggesting Vg plays a role. First, "receptor activity" is enriched in the list of nurse ChIP genes, albeit at the unadjusted P-value level. Still, this suggests that many genes targeted for Vg binding code for cell membrane proteins and components of cell signaling pathways. Second, there is significant enrichment for heparan sulfotransferase activity in the list of overlapping ChIP genes and DEGs, and the most abundant nuclear protein we identified, β-glucuronidase, is an enzyme that interacts with heparan sulfate in the nucleus (Schubert et al. 2004). Heparan sulfate often binds with proteins to form heparan sulfate proteoglycans (HSPGs), which are active on the plasma membrane and act as receptors for many ligands (Sarrazin et al. 2011). It is also of note that heparan sulfate in the nucleus is believed to play several roles in regulating gene expression and proliferation, as well as transport of molecules into the nucleus (reviewed in (Stewart and Sanderson 2014). Interestingly, a relationship between Vg-like proteins and HSPGs has been documented in *Drosophila*: here, *crossveinless d* (*csv-d*) codes a Vg homolog that contains the same functional domain as the Vg subunit we investigated here, and in

drosophila it binds to HSPGs and plays a role in the bone morphogenic protein (BMP) signaling pathway (Chen et al. 2012). The relationship we have documented here between Vg, heparan sulfate, and other mediating proteins, should warrant further investigation into the types of signaling pathways they influence in honey bees.

One potential shortcoming of this study is the lack comparison between Vg-wildtype bees and Vg-knockout or -knockdown bees, as is a common approach for investigating how transcription factor binding affects expression. However, such comparisons may be unfeasible or unnecessary for our purposes. For one, CRISPR/Cas-9 gene knockout in honey bees is still in its infancy, as it has only been demonstrated in a limited number of genes (Kohno et al. 2016; Hu, Zhang, et al. 2019; Roth et al. 2019) and still has challenges to overcome regarding efficacy and the effects of rearing larvae in hives vs *in vitro* (Souza et al. 2018). Moreover, even if the technology were sufficient, Vg-knockouts are likely lethal given the key role it plays in development. The first Vg-knockout to date used the vertebrate model *Danio rerio*, which has several Vg genes, and showed offspring of Vg-knockout parents displayed mutated and usually lethal phenotypes (Yilmaz et al. 2019). That leaves RNAi-mediated knockdowns as the remaining option in our study, but this approach presents another issue: Vg-knockdown accelerates worker development into foragers (Nelson et al. 2007; Antonio et al. 2008), so the contrast between Vg-wildtype and Vg-knockdown workers is essentially the same as the same-age nurse and forager contrast we have created here using single cohort colonies. In our same-age workers, nurses *vg* expression was ~8 times higher than in foragers, which is a similar degree of difference observed in Vg RNAi studies (Nunes et al. 2013; Wheeler et al. 2013).

This study has helped shed light on Vg's actions in the nucleus and its potential gene regulatory pathways. While many questions remain, including what signals prompt Vg nuclear translocation and how Vg crosses the nuclear envelope, the data presented here offer insight into the types of biological processes that Vg may influence. They also offer an opportunity for new avenues for Vg research that were previously unknown or overlooked, including possible signaling pathways involving Vg and the mitochondria. The extent of Vg's multifunctionality has always been impressive, and the discovery that its N-terminal region can also bind DNA and act as a possible transcription provides a mechanistic understanding of how it affects so many traits. Given that the Vg N-terminal region examined here is so conserved across metazoan taxa, our hope is that our findings spur other researchers to examine Vg's gene regulatory functions in other organisms.

CHAPTER 7

DISCUSSION

Vitellogenin is an ancient and highly conserved protein, and over its evolutionary history it has gained numerous functions. These multitude of functions have been studied in many organisms, but perhaps none more so than in honeybees. Here, it has evolved beyond its role in egg production and plays important roles in social behavior and social immunity. In this dissertation, I have made several important advances in our understanding of Vg in honey bees that may also be applicable to other species.

MAJOR FINDINGS AND IMPLICATIONS

First, in chapters 2 and 3 I have shown that Vg is part of a colony-level immune pathway that for transferring immune elicitors between colony members. Here, worker bees transported ingested bacteria from their midgut to their hypopharyngeal glands, and Vg appears to play in this transport as Vg-knockdown individuals do not show bacteria at their glands. Vg is a key component in the synthesis of royal jelly, and as it enters the hypopharyngeal glands via receptor-mediated endocytosis it may carry with it across to plasma membrane to be incorporated into the royal jelly. Nurse bees can then deliver these pathogen fragments to the queen or young larvae. Transferring pathogen fragments between colony members may serve to prime their immune systems and make them more resistant to disease. Royal jelly is not only a vehicle for transferring pathogens, but it is also sensitive to the immunological state of the nurses producing it, as royal jelly from pathogen-fed nurses have higher concentrations of the antimicrobial peptide *defensin*-1. Uncovering this colony-level immune pathway also has important practical applications

for beekeepers. This pathway can now be exploited to deliver immune elicitors to queens in order to produce disease-resistant bees. Indeed, my collaborator is marketizing this discovery with a now-patented honey bee vaccine for the deadly American foulbrood disease (Salmela and Freitak 2017).

In chapter 4, I have shown that Vg is localized in the worker midgut and possibly synthesized there, as we detected *vg* mRNA transcripts in this tissue. While Vg in known to synthesize in the digestive tract or hepatopancreas of several species, this is the first time to my knowledge that it has been shown in the midgut of honey bees. This has been an overlooked organ for Vg synthesis and function, and this discovery should spark future research. The midgut plays key roles in pathogen defense, digestion, and absorption, and Vg may be playing several roles here. Given Vg's roles in binding pathogens and acting as an antioxidant, Vg may be mediated the midgut's immune response to ingested pathogens, either by binding and neutralizing pathogens directly, or by buffering against the reactive oxygen species that midgut cells produce to kill pathogens. More research will be needed to determine the nature of Vg's role in this organ.

In chapters 5 and 6, I have made a major new discovery of Vg's binding abilities that has the potential to re-frame our understanding of Vg's multifunctionality. Here, I have shown that Vg binds to DNA in the nucleus and likely acts as a transcription factor or co-regulator. In the worker caste, Vg binds to hundreds of genes, and the identity of these binding sites shift as workers transition between different tasks. The types of genes Vg binds to gives a hint as to the types of biological processes being regulated by Vg. This includes many genes responsible for signal transduction pathways, metabolic

processes, and behavior. Our results also suggest that Vg is closely involved in enzymatic reactions, as its binding is not only associated with gene expression changes in enyzmes, but also because many enzymes in turn bind to the Vg-DNA complex. These results point to deeper evolutionary relationships between Vg and certain organismal processes. Vg is the oldest member of a large family of proteins known as Large Lipid Transfer Proteins (LLTPs), and this is the first known case of an LLTP functioning as a transcription factor. Given Vg's conservation across many taxa, this discovery here should spark research to determine if Vg and its orthologs also perform similar functions in other species. This could potentially reveal a new class of transcription factors build from ancient proteins. Vg plays a key role in so many important biological processes, and my findings here suggest that is directly involved in regulating genes pertaining to those biological processes.

LIMITATIONS

The work presented does contain a few limitations. In looking at a colony-level immune pathway, our findings show Vg and pathogen fragments in specific localizations such as midgut, hypopharyngeal glands, and royal jelly, but all of the intermediate steps remain a mystery. That is to say, we are still in the dark about certain stages of the transportation process. For example, how do pathogen fragments exit the midgut, enter the hemocoel, and finally enter the hypopharyngeal glands? What types of receptors is it binding do to cross tissue barriers? Does Vg facilitate this transport from the beginning in the midgut, does it intercept the fragments in the hemolymph, or is the knockdown of *vg* and the absence of pathogens at the glands merely coincidental? These are specific

questions about this pathway that have not been answered here and which will require a more concerted effort to map them out.

In looking at Vg as a transcription factor, next-gen sequencing and mass spectrometry methods present some limitations as well. For one, our capabilities to produce large data sets and then them is constantly improving. So, what's current and state of the art today is likely to be dated within a few years. Uncovering transcription factor targets is still mostly reliant on correlating protein-DNA binding sites with differential expression at those sites. But this approach poses a few problems. First, gene transcription is controlled by numerous regulatory mechanisms and typically involves many transcription factors working in concert. So, just because a protein-DNA binding site does not show differential expression between individuals from different treatment groups, it does not necessarily mean that that given protein does not help regulate expression of that gene. One approach researchers use to increase confidence that a protein actually controls gene expression at a given loci is to knockout the purported transcription factor and see how this affects expression of its target gene. This can still be problematic because as there are multiple transcription factors regulating that gene, so too can there be compensatory regulatory mechanisms that keep that gene expressing at a similar level between treatments. Knockouts also pose a particular problem for our study organism because these types of molecular approaches are still being developed in honey bees. Moreover, even if the methods were developed, Vg-knockouts would likely be lethal since Vg is so important to development. Thus, despite the major discovery I have made regarding Vg's DNA binding ability and its *likely* role as a transcription factor, we

will have to continue using new molecular techniques as they develop in order to more precisely elucidate Vg's gene regulatory functions.

There are also some limitations in characterizing the additional nuclear proteins that are bound to the Vg-DNA complex. For one, the approach used in this dissertation pulls down a broad range of *all* the proteins that bind to the Vg-DNA complex, but it does not give information about which particular combinations of those proteins are bound at specific loci. Thus, we are left with a very general list of proteins from which we can examine common functions between them to try to ascertain information about potential signaling pathways. But, understanding how particular combinations of proteins work together will require development of more precise methods.

FUTURE DIRECTIONS

The findings outlined in this dissertation provide ample opportunity for future research. First, the colony level immune pathway I've laid out shows the transfer of bacterial pathogens, and a collaborator has turned this finding into a commercial vaccine product. However, it remains to be seen whether this same pathway may work for viral pathogens as well. When I've presented these findings to various beekeeper groups, this is a question they often raise since there are several deadly viruses common to US beekeeping operations. My work in this dissertation shows that Vg binds to nucleic acids, and there is some evidence that it binds viruses in fish (Garcia et al. 2010). Now, the next steps are to confirm Vg-virus binding, and then determine the consequences of this binding. Does Vg inhibit viral replication? Are viruses binding to Vg to gain entry into additional tissues? Can viral particles be incorporated into royal jelly to pass between

individuals, and if so, what are the consequences for the recipient? These questions will need to be answered before bee vaccines against viruses can be developed.

Finding Vg in the midgut is a curious discovery, as this organ has been overlooked for examining Vg's tissue-specific functions. The next step here is to determine Vg's role. It seems plausible that Vg acts in response to ingested food, acting to bind and/or kill pathogens, act as an antioxidant to protect epithelial cells from oxidative stress, or perhaps in capacity as a nutrient transport protein. Numerous functional assays can be performed to better understand Vg function here.

The discovery that Vg may be a transcription factor in honey bee workers is a major discovery, and there are plenty more questions to answer in follow up. I have shown how Vg binds DNA across different behavioral groups in the worker caste, but it still remains to be seen how Vg-DNA binding occurs in other castes like queens and drones. Knowing this will provide a more complete picture of how Vg regulates honey bee biology. Given Vg's highly conserved structure across so many organisms, studies should also examine whether it plays gene regulatory roles in other species as well. Finally, it would be useful to know how dynamic this Vg transcription factor activity is in response to different perturbations. For example, how might nutritional stress brought about through seasonal changes in flower availability alter the profile of Vg-DNA binding sites? Vg interacts in the nucleus with many genes and other proteins that function in metabolic pathways, so response to nutritional stress might reconfigure some of these signaling pathways. As another perturbation, how does immune challenge alter Vg-DNA interactions? My current research position examines the interplay between viral infection, behavior, and physiology in honey bees, and knowing how Vg responds to such

an infection would be useful, especially considering how many important functions seem to be influenced by Vg's functions in the nucleus.

CONCLUSIONS

Vg plays a remarkable number of important functions in organisms, and my research has revealed at least 2 more important functions. These include its role in a colony-level immune pathway in honey bees, and its functional as a potential transcription factor. Each discovery should pave the way for future avenues of research as we continue to unlock the mysteries of this ancient multifunctional protein.

REFERENCES

Akira S, Takeda K, Kaisho T. 2001. Toll-like receptors: critical proteins linking innate and acquired immunity. *Nat Immunol*. 2:675–680.

Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P. 2002. Fractionation of Cells. *Mol Biol Cell 4th Ed*.

Alejevski F, Saint-Charles A, Michard-Vanhée C, Martin B, Galant S, Vasiliauskas D, Rouyer F. 2019. The HisCl1 histamine receptor acts in photoreceptors to synchronize Drosophila behavioral rhythms with light-dark cycles. *Nat Commun*. 10.

Amdam GV, Aase ALTO, Seehuus S-C, Kim Fondrk M, Norberg K, Hartfelder K. 2005. Social reversal of immunosenescence in honey bee workers. *Exp Gerontol*. 40:939–947.

Amdam GV, Csondes A, Fondrk MK, Page RE. 2006. Complex social behaviour derived from maternal reproductive traits. *Nature*. 439:76–78.

Amdam GV, Fennern E, Baker N, Rascón B. 2010. Honeybee Associative Learning Performance and Metabolic Stress Resilience Are Positively Associated. *PLOS ONE*. 5:e9740.

Amdam GV, Fennern E, Havukainen H. 2012. Vitellogenin in Honey Bee Behavior and Lifespan. In: Galizia CG,, Eisenhardt D,, Giurfa M, editors. Honeybee Neurobiology and Behavior: A Tribute to Randolf Menzel. Dordrecht: Springer Netherlands. p. 17–29.

Amdam GV, Norberg K, Fondrk MK, Page RE. 2004. Reproductive ground plan may mediate colony-level selection effects on individual foraging behavior in honey bees. *Proc Natl Acad Sci U S A*. 101:11350–11355.

Amdam GV, Norberg K, Hagen A, Omholt SW. 2003. Social exploitation of vitellogenin. *Proc Natl Acad Sci*. 100:1799–1802.

Amdam GV, Norberg K, Omholt SW, Kryger P, Lourenço AP, Bitondi MMG, Simões ZLP. 2005a. Higher vitellogenin concentrations in honey bee workers may be an adaptation to life in temperate climates. *Insectes Sociaux*. 52:316–319.

Amdam GV, Norberg K, Omholt SW, Kryger P, Lourenço AP, Bitondi MMG, Simões ZLP. 2005b. Higher vitellogenin concentrations in honey bee workers may be an adaptation to life in temperate climates. *Insectes Sociaux*. 52:316–319.

Amdam GV, Norberg K, Page Jr. RE, Erber J, Scheiner R. 2006. Downregulation of vitellogenin gene activity increases the gustatory responsiveness of honey bee workers (Apis mellifera). *Behav Brain Res*. 169:201–205.

Amdam GV, Omholt SW. 2002. The Regulatory Anatomy of Honeybee Lifespan. *J Theor Biol*. 216:209–228.

Amdam GV, Omholt SW. 2003. The hive bee to forager transition in honeybee colonies: the double repressor hypothesis. *J Theor Biol*. 223:451–464.

Amdam GV, Simões ZL, Guidugli KR, Norberg K, Omholt SW. 2003. Disruption of vitellogenin gene function in adult honeybees by intra-abdominal injection of double-stranded RNA. *BMC Biotechnol*. 3:1.

Amdam GV, Simões ZLP, Hagen A, Norberg K, Schrøder K, Mikkelsen Ø, Kirkwood TBL, Omholt SW. 2004. Hormonal control of the yolk precursor vitellogenin regulates immune function and longevity in honeybees. *Exp Gerontol*. 39:767–773.

Ament SA, Chan QW, Wheeler MM, Nixon SE, Johnson SP, Rodriguez-Zas SL, Foster LJ, Robinson GE. 2011. Mechanisms of stable lipid loss in a social insect. *J Exp Biol*. 214:3808–3821.

Anderson KE, Sheehan TH, Mott BM, Maes P, Snyder L, Schwan MR, Walton A, Jones BM, Corby-Harris V. 2013. Microbial Ecology of the Hive and Pollination Landscape: Bacterial Associates from Floral Nectar, the Alimentary Tract and Stored Food of Honey Bees (Apis mellifera). *PLOS ONE*. 8:e83125.

Anderson T, Levitt D, Banaszak L. 1998. The structural basis of lipid interactions in lipovitellin, a soluble lipoprotein. *Structure*. 6:895–909.

Andrés M, Turiégano E, Göpfert MC, Canal I, Torroja L. 2014. The Extracellular Matrix Protein Artichoke Is Required for Integrity of Ciliated Mechanosensory and Chemosensory Organs in Drosophila Embryos. *Genetics*. 196:1091–1102.

Andrews S. 2010. FastQC A Quality Control tool for High Throughput Sequence Data [WWW Document]. URL https://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Antonio DSM, Guidugli-Lazzarini KR, Nascimento AM do, Simões ZLP, Hartfelder K. 2008. RNAi-mediated silencing of vitellogenin gene function turns honeybee (Apis mellifera) workers into extremely precocious foragers. *Naturwissenschaften*. 95:953–961.

Araujo RN, Santos A, Pinto FS, Gontijo NF, Lehane MJ, Pereira MH. 2006. RNA interference of the salivary gland nitrophorin 2 in the triatomine bug Rhodnius prolixus (Hemiptera: Reduviidae) by dsRNA ingestion or injection. *Insect Biochem Mol Biol*. 36:683–693.

Aronesty E. 2011. Ea-utils: Command-line tools for processing biological sequencing data.

Arrese EL, Soulages JL. 2010. Insect fat body: energy, metabolism, and regulation. *Annu Rev Entomol*. 55:207–225.

Azzouz-Olden F, Hunt A, DeGrandi-Hoffman G. 2018. Transcriptional response of honey bee (Apis mellifera) to differential nutritional status and Nosema infection. *BMC Genomics*. 19:628.

Babin PJ, Bogerd J, Kooiman FP, Van Marrewijk WJA, Van der Horst DJ. 1999. Apolipophorin II/I, Apolipoprotein B, Vitellogenin, and Microsomal Triglyceride Transfer Protein Genes Are Derived from a Common Ancestor. *J Mol Evol*. 49:150–160.

Bachanová K, Klaudiny J, Kopernickỳ J, Šimúth J. 2002a. Identification of honeybee peptide active against Paenibacillus larvae larvae through bacterial growth-inhibition assay on polyacrylamide gel. *Apidologie*. 33:259–269.

Bachanová K, Klaudiny J, Kopernickỳ J, Šimúth J. 2002b. Identification of honeybee peptide active against Paenibacillus larvae larvae through bacterial growth-inhibition assay on polyacrylamide gel. *Apidologie*. 33:259–269.

Bai H, Kang P, Hernandez AM, Tatar M. 2013. Activin Signaling Targeted by Insulin/dFOXO Regulates Aging and Muscle Proteostasis in Drosophila. *PLOS Genet*. 9:e1003941.

Bailey TL. 2011. DREME: motif discovery in transcription factor ChIP-seq data. *Bioinformatics*. 27:1653–1659.

Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS. 2009. MEME Suite: tools for motif discovery and searching. *Nucleic Acids Res*. 37:W202–W208.

Baker ME. 1988a. Is vitellogenin an ancestor of apolipoprotein B-100 of human low-density lipoprotein and human lipoprotein lipase? *Biochem J*. 255:1057–1060.

Baker ME. 1988b. Invertebrate vitellogenin is homologous to human von Willebrand factor. *Biochem J*. 256:1059–1061.

Barron AB, Oldroyd BP, Ratnieks FL. 2001. Worker reproduction in honey-bees (Apis) and the anarchic syndrome: a review. *Behav Ecol Sociobiol*. 50:199–208.

Bedick JC, Tunaz H, Nor Aliza AR, Putnam SM, Ellis MD, Stanley DW. 2001. Eicosanoids act in nodulation reactions to bacterial infections in newly emerged adult honey bees, Apis mellifera, but not in older foragers. *Comp Biochem Physiol Part C Toxicol Pharmacol*. 130:107–117.

Beshers SN, Fewell JH. 2001. Models of division of labor in social insects. *Annu Rev Entomol*. 46:413–440.

Beveridge TJ. 1999. Structures of gram-negative cell walls and their derived membrane vesicles. *J Bacteriol*. 181:4725–4733.

Bíliková K, Wu G, Šimúth J. 2001. Isolation of a peptide fraction from honeybee royal jelly as a potential antifoulbrood factor. *Apidologie*. 32:275–283.

Blum MS, Novak AF, Taber S. 1959. 10-Hydroxy-Δ2-Decenoic Acid, an Antibiotic Found in Royal Jelly. *Science*. 130:452–453.

Böhme F, Bischoff G, Zebitz CPW, Rosenkranz P, Wallner K. 2018. From field to food—will pesticide-contaminated pollen diet lead to a contamination of royal jelly? *Apidologie*. 49:112–119.

Boldbaatar D, Umemiya-Shirafuji R, Liao M, Tanaka T, Xuan X, Fujisaki K. 2010. Multiple vitellogenins from the Haemaphysalis longicornis tick are crucial for ovarian development. *J Insect Physiol*. 56:1587–1598.

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 30:2114–2120.

Bondos SE, Tan X-X. 2001. Combinatorial Transcriptional Regulation: The Interaction of Transcription Factors and Cell Signaling Molecules with Homeodomain Proteins in Drosophila Development. *Crit Rev Eukaryot Gene Expr*. 11.

Bonizzoni M, Dunn WA, Campbell CL, Olson KE, Dimon MT, Marinotti O, James AA. 2011. RNA-seq analyses of blood-induced changes in gene expression in the mosquito vector species, Aedes aegypti. *BMC Genomics*. 12:82.

Bownes M. 1986. Expression of the genes coding for vitellogenin (yolk protein). *Annu Rev Entomol*. 31:507–531.

Brandt CR, Adang MJ, Spence KD. 1978. The peritrophic membrane: Ultrastructural analysis and function as a mechanical barrier to microbial infection in Orgyia pseudotsugata. *J Invertebr Pathol*. 32:12–24.

Brasaemle DL, Wolins NE. 2012. Packaging of fat: an evolving model of lipid droplet assembly and expansion. *J Biol Chem*. 287:2273–2279.

Brian MV. 2012. Social Insects: Ecology and Behavioural Biology. Springer Science & Business Media.

Brodschneider R, Crailsheim K. 2010. Nutrition and health in honey bees. *Apidologie*. 41:278–294.

Brødsgaard CJ, Ritter W, Hansen H. 1998. Response of in vitro reared honey bee larvae to various doses of Paenibacillus larvae larvae spores. *Apidologie*. 29:569–578.

Brookes VJ. 1969. The induction of yolk protein synthesis in the fat body of an insect, Leucophaea maderae, by an analog of the juvenile hormone. *Dev Biol*. 20:459–471.

Bucekova M, Sojka M, Valachova I, Martinotti S, Ranzato E, Szep Z, Majtan V, Klaudiny J, Majtan J. 2017. Bee-derived antibacterial peptide, defensin-1, promotes wound re-epithelialisation in vitro and in vivo. *Sci Rep*. 7:7340.

Bucekova M, Valachova I, Kohutova L, Prochazka E, Klaudiny J, Majtan J. 2014. Honeybee glucose oxidase—its expression in honeybee workers and comparative analyses of its content and H2O2-mediated antibacterial activity in natural honeys. *Naturwissenschaften*. 101:661–670.

Buchon N, Broderick NA, Lemaitre B. 2013. Gut homeostasis in a microbial world: insights from Drosophila melanogaster. *Nat Rev Microbiol*. 11:615–626.

Buchon N, Broderick NA, Poidevin M, Pradervand S, Lemaitre B. 2009. Drosophila intestinal response to bacterial infection: activation of host defense and stem cell proliferation. *Cell Host Microbe*. 5:200–211.

Buchon N, Silverman N, Cherry S. 2014. Immunity in Drosophila melanogaster — from microbial recognition to whole-organism physiology. *Nat Rev Immunol*. 14:796–810.

Bulet P, Hetru C, Dimarcq J-L, Hoffmann D. 1999. Antimicrobial peptides in insects; structure and function. *Dev Comp Immunol*, Invertebrate Immunity. 23:329–344.

Bulet P, Stöcklin R. 2005. Insect antimicrobial peptides: structures, properties and gene regulation. *Protein Pept Lett*. 12:3–11.

Burda H, Honeycutt RL, Begall S, Locker-Grütjen O, Scharff A. 2000. Are naked and common mole-rats eusocial and if so, why? *Behav Ecol Sociobiol*. 47:293–303.

Cantó C, Menzies KJ, Auwerx J. 2015. NAD+ Metabolism and the Control of Energy Homeostasis: A Balancing Act between Mitochondria and the Nucleus. *Cell Metab*. 22:31–53.

Casteels P. 1998. Immune response in hymenoptera. In: Brey PT,, Hultmark D, editors. Molecular mechanisms of immune responses in insects. London: Chapman & Hall. p. 92–110.

Casteels-Josson K, Zhang W, Capaci T, Casteels P, Tempst P. 1994. Acute transcriptional response of the honeybee peptide-antibiotics gene repertoire and required post-translational conversion of the precursor structures. *J Biol Chem*. 269:28569–28575.

Chalk R, Townson H, Natori S, Desmond H, Ham PJ. 1994. Purification of an insect defensin from the mosquito, Aedes aegypti. *Insect Biochem Mol Biol*. 24:403–410.

Chan QW, Foster LJ. 2008. Changes in protein expression during honey bee larval development. *Genome Biol*. 9:R156.

Chan QW, Melathopoulos AP, Pernal SF, Foster LJ. 2009. The innate immune and systemic response in honey bees to a bacterial pathogen, Paenibacillus larvae. *BMC Genomics*. 10:387.

Chapman RF. 1998. The Insects: Structure and Function. Cambridge University Press.

Chen J, Honeyager SM, Schleede J, Avanesov A, Laughon A, Blair SS. 2012. Crossveinless d is a vitellogenin-like lipoprotein that binds BMPs and HSPGs, and is required for normal BMP signaling in the Drosophila wing. *Dev Camb Engl*. 139:2170–2176.

Chesmore KN, Bartlett J, Cheng C, Williams SM. 2016. Complex Patterns of Association between Pleiotropy and Transcription Factor Evolution. *Genome Biol Evol*. 8:3159–3170.

Cho N-H, Seong S-Y. 2009. Apolipoproteins inhibit the innate immunity activated by necrotic cells or bacterial endotoxin. *Immunology*. 128:e479–e486.

Choo YM, Lee KS, Yoon HJ, Kim BY, Sohn MR, Roh JY, Je YH, Kim NJ, Kim I, Woo SD, Sohn HD, Jin BR. 2010. Dual Function of a Bee Venom Serine Protease: Prophenoloxidase-Activating Factor in Arthropods and Fibrin(ogen)olytic Enzyme in Mammals. *PLoS ONE*. 5.

Christianson HC, Belting M. 2014. Heparan sulfate proteoglycan as a cell-surface endocytosis receptor. *Matrix Biol J Int Soc Matrix Biol*. 35:51–55.

Clokey GV, Jacobson LA. 1986. The autofluorescent "lipofuscin granules" in the intestinal cells of Caenorhabditis elegans are secondary lysosomes. *Mech Ageing Dev*. 35:79–94.

Cohen P. 2000. The regulation of protein function by multisite phosphorylation – a 25 year update. *Trends Biochem Sci*. 25:596–601.

Cole KD, Wells MA. 1990. A comparison of adult and larval Manduca sexta apolipophorin-III. *Insect Biochem*. 20:373–380.

Cook SC, Davidson DW. 2006. Nutritional and functional biology of exudate-feeding ants. *Entomol Exp Appl*. 118:1–10.

Corby-Harris V, Meador CAD, Snyder LA, Schwan MR, Maes P, Jones BM, Walton A, Anderson KE. 2016. Transcriptional, translational, and physiological signatures of undernourished honey bees (Apis mellifera) suggest a role for hormonal factors in hypopharyngeal gland degradation. *J Insect Physiol*. 85:65–75.

Corona M, Libbrecht R, Wurm Y, Riba-Grognuz O, Studer RA, Keller L. 2013. Vitellogenin Underwent Subfunctionalization to Acquire Caste and Behavioral Specific Expression in the Harvester Ant Pogonomyrmex barbatus. *PLOS Genet*. 9:e1003730.

Corona M, Velarde RA, Remolina S, Moran-Lauter A, Wang Y, Hughes KA, Robinson GE. 2007a. Vitellogenin, juvenile hormone, insulin signaling, and queen honey bee longevity. *Proc Natl Acad Sci*. 104:7128–7133.

Corona M, Velarde RA, Remolina S, Moran-Lauter A, Wang Y, Hughes KA, Robinson GE. 2007b. Vitellogenin, juvenile hormone, insulin signaling, and queen honey bee longevity. *Proc Natl Acad Sci*. 104:7128–7133.

Cozzi A, Levi S, Corsi B, Santambrogio P, Campanella A, Gerardi G, Arosio P. 2003. Role of iron and ferritin in TNFα-induced apoptosis in HeLa cells. *FEBS Lett*. 537:187–192.

Crailsheim K, Schneider LHW, Hrassnigg N, Bühlmann G, Brosch U, Gmeinbauer R, Schöffmann B. 1992. Pollen consumption and utilization in worker honeybees (Apis mellifera carnica): Dependence on individual age and function. *J Insect Physiol*. 38:409–419.

Cremer S. 2019. Social immunity in insects. *Curr Biol*. 29:R458–R463.

Cremer S, Armitage SAO, Schmid-Hempel P. 2007. Social Immunity. *Curr Biol*. 17:R693–R702.

Crespi BJ, Yanega D. 1995. The definition of eusociality. *Behav Ecol*. 6:109–115.

da Cruz-Landim C, Cavalcante VM. 2003. Ultrastructural and Cytochemical Aspects of Metamorphosis in the Midgut of Apis mellifera L. (Hymenoptera: Apidae: Apinae). *Zoolog Sci*. 20:1099–1107.

de Azevedo SV, Hartfelder K. 2008. The insulin signaling pathway in honey bee (Apis mellifera) caste development — differential expression of insulin-like peptides and insulin receptors in queen and worker larvae. *J Insect Physiol*. 54:1064–1071.

De Gregorio E, Spellman PT, Tzou P, Rubin GM, Lemaitre B. 2002. The Toll and Imd pathways are the major regulators of the immune response in Drosophila. *EMBO J*. 21:2568–2579.

de Oliveira VP, Cruz-Landim C. 2003. Morphology and function of insect fat body cells: a review. *Biociências*. 11:195–205.

Deretic V. 2012. Autophagy as an innate immunity paradigm: expanding the scope and repertoire of pattern recognition receptors. *Curr Opin Immunol*, Innate immunity / Antigen processing. 24:21–31.

Dittmer J, Alafndi A, Gabrieli P. 2019. Fat body–specific vitellogenin expression regulates host-seeking behaviour in the mosquito Aedes albopictus. *PLOS Biol*. 17:e3000238.

Doublet V, Poeschl Y, Gogol-Döring A, Alaux C, Annoscia D, Aurori C, Barribeau SM, Bedoya-Reina OC, Brown MJF, Bull JC, Flenniken ML, Galbraith DA, Genersch E, Gisder S, Grosse I, Holt HL, Hultmark D, Lattorff HMG, Le Conte Y, Manfredini F, McMahon DP, Moritz RFA, Nazzi F, Niño EL, Nowick K, van Rij RP, Paxton RJ, Grozinger CM. 2017. Unity in defence: honeybee workers exhibit conserved molecular responses to diverse pathogens. *BMC Genomics*. 18.

Dow JAT. 1987. Insect Midgut Function. In: Evans PD,, Wigglesworth VB, editors. Advances in Insect Physiology. Academic Press. p. 187–328.

Du X, Wang X, Wang S, Zhou Y, Zhang Y, Zhang S. 2017. Functional characterization of Vitellogenin_N domain, domain of unknown function 1943, and von Willebrand factor type D domain in vitellogenin of the non-bilaterian coral Euphyllia ancora: Implications for emergence of immune activity of vitellogenin in basal metazoan. *Dev Comp Immunol*. 67:485–494.

Duffy JE. 1996. Eusociality in a coral-reef shrimp. *Nature*. 381:512.

Dussaubat C, Brunet J-L, Higes M, Colbourne JK, Lopez J, Choi J-H, Martín-Hernández R, Botías C, Cousin M, McDonnell C, Bonnet M, Belzunces LP, Moritz RFA, Le Conte Y, Alaux C. 2012. Gut Pathology and Responses to the Microsporidium Nosema ceranae in the Honey Bee Apis mellifera. *PLoS ONE*. 7:e37017.

Eiri DM, Nieh JC. 2012. A nicotinic acetylcholine receptor agonist affects honey bee sucrose responsiveness and decreases waggle dancing. *J Exp Biol*. 215:2022–2029.

Elsik CG, Tayal A, Diesh CM, Unni DR, Emery ML, Nguyen HN, Hagen DE. 2016. Hymenoptera Genome Database: integrating genome annotations in HymenopteraMine. *Nucleic Acids Res*. 44:D793–D800.

Engel P, Moran NA. 2013. The gut microbiota of insects – diversity in structure and function. *FEMS Microbiol Rev*. 37:699–735.

Engelmann F. 1979. Insect Vitellogenin: Identification, Biosynthesis, and Role in Vitellogenesis. In: Treherne JE,, Berridge MJ,, Wigglesworth VB, editors. Advances in Insect Physiology. Academic Press. p. 49–108.

Engels W, Fahrenhorst H. 1974. Alters- und kastenspezifische Veränderungen der Haemolymph-Protein-Spektren bei Apis mellificia. *Wilhelm Roux Arch Für Entwicklungsmechanik Org*. 174:285–296.

Engels W, Kaatz H, Zillikens A, Simões ZLP, Trube A, Braun R, Dittrich F. 1990. Honey bee reproduction: vitellogenin and caste-specific regulation of fertility. *Honey Bee Reprod Vitellogenin Caste-Specif Regul Fertil*. 495–502.

Evans JD. 2004. Transcriptional immune responses by honey bee larvae during invasion by the bacterial pathogen, Paenibacillus larvae. *J Invertebr Pathol*. 85:105–111.

Evans JD, Aronstein K, Chen YP, Hetru C, Imler J-L, Jiang H, Kanost M, Thompson GJ, Zou Z, Hultmark D. 2006. Immune pathways and defence mechanisms in honey bees Apis mellifera. *Insect Mol Biol*. 15:645–656.

Evans JD, Lopez DL. 2004. Bacterial probiotics induce an immune response in the honey bee (Hymenoptera: Apidae). *J Econ Entomol*. 97:752–756.

Evans JD, Schwarz RS. 2011. Bees brought to their knees: microbes affecting honey bee health. *Trends Microbiol*. 19:614–620.

Evans JD, Spivak M. 2010. Socialized medicine: individual and communal disease barriers in honey bees. *J Invertebr Pathol*. 103:S62–S72.

Evans JD, Wheeler DE. 1999. Differential gene expression between developing queens and workers in the honey bee, Apis mellifera. *Proc Natl Acad Sci*. 96:5575–5580.

Faust TB, Binning JM, Gross JD, Frankel AD. 2017. Making Sense of Multifunctional Proteins: Human Immunodeficiency Virus Type 1 Accessory and Regulatory Proteins and Connections to Transcription. *Annu Rev Virol*. 4:241–260.

Fehlbaum P, Bulet P, Michaut L, Lagueux M, Broekaert WF, Hetru C, Hoffmann JA. 1994. Insect immunity. Septic injury of Drosophila induces the synthesis of a potent antifungal peptide with sequence homology to plant antifungal peptides. *J Biol Chem*. 269:33159–33163.

Finn RN. 2007. Vertebrate Yolk Complexes and the Functional Implications of Phosvitins and Other Subdomains in Vitellogenins. *Biol Reprod*. 76:926–935.

Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC. 1998. Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature*. 391:806.

Flenniken ML, Andino R. 2013. Non-Specific dsRNA-Mediated Antiviral Response in the Honey Bee. *PLOS ONE*. 8:e77263.

Fletcher BL, Dillard CJ, Tappel AL. 1973. Measurement of fluorescent lipid peroxidation products in biological systems and tissues. *Anal Biochem*. 52:1–9.

Fluri P, Lüscher M, Wille H, Gerig L. 1982. Changes in weight of the pharyngeal gland and haemolymph titres of juvenile hormone, protein and vitellogenin in worker honey bees. *J Insect Physiol*. 28:61–68.

Fluri P, Wille H, Gerig L, Lüscher M. 1977. Juvenile hormone, vitellogenin and haemocyte composition in winter worker honeybees (Apis mellifera). *Experientia*. 33:1240–1241.

Fontana A, Laureto D, Polverino P, Spolaore B, Frare E, Picotti P, Zambonin M. 2004. Probing protein structure by limited proteolysis. *Acta Biochim Pol*. 51:299–321.

Fontana R, Mendes MA, Souza BM de, Konno K, César LMM, Malaspina O, Palma MS. 2004. Jelleines: a family of antimicrobial peptides from the Royal Jelly of honeybees (Apis mellifera). *Peptides*. 25:919–928.

Freitak D, Heckel DG, Vogel H. 2009. Dietary-dependent trans-generational immune priming in an insect herbivore. *Proc R Soc B Biol Sci*. 276:2617–2624.

Freitak D, Schmidtberg H, Dickel F, Lochnit G, Vogel H, Vilcinskas A. 2014. The maternal transfer of bacteria can mediate trans-generational immune priming in insects. *Virulence*. 5:547–554.

Fujita T, Kozuka-Hata H, Ao-Kondo H, Kunieda T, Oyama M, Kubo T. 2012. Proteomic analysis of the royal jelly and characterization of the functions of its derivation glands in the honeybee. *J Proteome Res*. 12:404–411.

Fujita T, Kozuka-Hata H, Ao-Kondo H, Kunieda T, Oyama M, Kubo T. 2013. Proteomic Analysis of the Royal Jelly and Characterization of the Functions of its Derivation Glands in the Honeybee. *J Proteome Res*. 12:404–411.

Fujiwara S, Imai J, Fujiwara M, Yaeshima T, Kawashima T, Kobayashi K. 1990a. A potent antibacterial protein in royal jelly. Purification and determination of the primary structure of royalisin. *J Biol Chem*. 265:11333–11337.

Fujiwara S, Imai J, Fujiwara M, Yaeshima T, Kawashima T, Kobayashi K. 1990b. A potent antibacterial protein in royal jelly. Purification and determination of the primary structure of royalisin. *J Biol Chem*. 265:11333–11337.

Furusawa T, Rakwal R, Nam HW, Shibato J, Agrawal GK, Kim YS, Ogawa Y, Yoshida Y, Kouzuma Y, Masuo Y, Yonekura M. 2008. Comprehensive Royal Jelly (RJ) Proteomics Using One- and Two-Dimensional Proteomics Platforms Reveals Novel RJ Proteins and Potential Phospho/Glycoproteins. *J Proteome Res*. 7:3194–3229.

Ganz T. 2003. Defensins: antimicrobial peptides of innate immunity. *Nat Rev Immunol*. 3:710–720.

Gao H, Lin Y, He J, Zhou S, Liang M, Huang C, Li X, Liu C, Zhang P. 2019. Role of heparan sulfate in the Zika virus entry, replication, and cell death. *Virology*. 529:91–100.

Garcia J, Munro ES, Monte MM, Fourrier MCS, Whitelaw J, Smail DA, Ellis AE. 2010. Atlantic salmon (Salmo salar L.) serum vitellogenin neutralises infectivity of infectious pancreatic necrosis virus (IPNV). *Fish Shellfish Immunol*. 29:293–297.

Garcia-Gonzalez E, Genersch E. 2013. Honey bee larval peritrophic matrix degradation during infection with Paenibacillus larvae, the aetiological agent of American foulbrood of honey bees, is a key step in pathogenesis. *Environ Microbiol*. 15:2894–2901.

Genersch E. 2010. American Foulbrood in honeybees and its causative agent, Paenibacillus larvae. *J Invertebr Pathol*. 103:S10–S19.

Gill G. 2001. Regulation of the initiation of eukaryotic transcription. *Essays Biochem*. 37:33–43.

González-Santoyo I, Córdoba-Aguilar A. 2012. Phenoloxidase: a key component of the insect immune system. *Entomol Exp Appl*. 142:1–16.

Gospocic J, Shields EJ, Glastad KM, Lin Y, Penick CA, Yan H, Mikheyev AS, Linksvayer TA, Garcia BA, Berger SL, Liebig J, Reinberg D, Bonasio R. 2017. The Neuropeptide Corazonin Controls Social Behavior and Caste Identity in Ants. *Cell*. 170:748–759.e12.

Groh CD, MacPherson DW, Groves DJ. 1996. Effect of Heat on the Sterilization of Artificially Contaminated Water. *J Travel Med*. 3:11–13.

Grohmann L, Blenau W, Erber J, Ebert PR, Strünker T, Baumann A. 2003. Molecular and functional characterization of an octopamine receptor from honeybee (Apis mellifera) brain. *J Neurochem*. 86:725–735.

Guan Z-B, Yin J, Chen K, Shui Y, Cai Y-J, Liao X-R. 2016. The Hepatopancreas and Ovary are the Sites of Vitellogenin Synthesis in Female Red Swamp Crayfish (Procambarus Clarkii ()) (Decapoda: Astacoidea: Cambaridae). *J Crustac Biol*. 36:637–641.

Guidugli KR, Nascimento AM, Amdam GV, Barchuk AR, Omholt S, Simões ZLP, Hartfelder K. 2005. Vitellogenin regulates hormonal dynamics in the worker caste of a eusocial insect. *FEBS Lett*. 579:4961–4965.

Guidugli KR, Piulachs M-D, BelléS X, LourençO AP, Simões ZLP. 2005. Vitellogenin expression in queen ovaries and in larvae of both sexes of Apis mellifera. *Arch Insect Biochem Physiol*. 59:211–218.

Guidugli-Lazzarini KR, do Nascimento AM, Tanaka ÉD, Piulachs MD, Hartfelder K, Bitondi MG, Simões ZLP. 2008. Expression analysis of putative vitellogenin and lipophorin receptors in honey bee (Apis mellifera L.) queens and workers. *J Insect Physiol*. 54:1138–1147.

Guo R, Xu D, Wang W. 2009. Identification and analysis of new Proteins involved in the DNA Damage Response network of Fanconi anemia and Bloom syndrome. *Methods San Diego Calif*. 48:72–79.

Gupta S, Stamatoyannopoulos JA, Bailey TL, Noble WS. 2007. Quantifying similarity between motifs. *Genome Biol*. 8:R24.

Gurevich VV, Gurevich EV. 2015. Analyzing the roles of multi-functional proteins in cells: the case of arrestins and GRKs. *Crit Rev Biochem Mol Biol*. 50:440–452.

Gutierrez E, Wiggins D, Fielding B, Gould AP. 2007. Specialized hepatocyte-like cells regulate Drosophila lipid metabolism. *Nature*. 445:275–280.

Ha E-M, Oh C-T, Bae YS, Lee W-J. 2005. A Direct Role for Dual Oxidase in Drosophila Gut Immunity. *Science*. 310:847–850.

Hakim RS, Baldwin KM, Loeb M. 2001. The role of stem cells in midgut growth and regeneration. *Vitro Cell Dev Biol - Anim*. 37:338–342.

Han B, Li C, Zhang L, Fang Y, Feng M, Li J. 2011a. Novel Royal Jelly Proteins Identified by Gel-Based and Gel-free Proteomics. *J Agric Food Chem*. 59:10346–10355.

Han B, Li C, Zhang L, Fang Y, Feng M, Li J. 2011b. Novel Royal Jelly Proteins Identified by Gel-Based and Gel-free Proteomics. *J Agric Food Chem*. 59:10346–10355.

Handy DE, Loscalzo J. 2012. Redox Regulation of Mitochondrial Function. *Antioxid Redox Signal*. 16:1323–1367.

Hansen H, Brødsgaard CJ. 1999. American foulbrood: a review of its biology, diagnosis and control. *Bee World*. 80:5–23.

Harwood G, Amdam G, Freitak D. 2019. The role of Vitellogenin in the transfer of immune elicitors from gut to hypopharyngeal glands in honey bees (Apis mellifera). *J Insect Physiol*. 112:90–100.

Harwood GP, Ihle KE, Salmela H, Amdam G. 2016. Regulation of Honeybee Worker (Apis mellifera) Life Histories by Vitellogenin. *Non-Mamm Horm-Behav Syst*. 403–420.

Hasselquist D, Nilsson J-Å. 2009. Maternal transfer of antibodies in vertebrates: trans-generational effects on offspring immunity. *Philos Trans R Soc Lond B Biol Sci*. 364:51–60.

Havukainen H, Halskau Ø, Amdam GV. 2011. Social pleiotropy and the molecular evolution of honey bee vitellogenin. *Mol Ecol*. 20:5111–5113.

Havukainen H, Halskau Ø, Skjaerven L, Smedal B, Amdam GV. 2011. Deconstructing honeybee vitellogenin: novel 40 kDa fragment assigned to its N terminus. *J Exp Biol*. 214:582–592.

Havukainen H, Münch D, Baumann A, Zhong S, Halskau Ø, Krogsgaard M, Amdam GV. 2013. Vitellogenin Recognizes Cell Damage through Membrane Binding and Shields Living Cells from Reactive Oxygen Species. *J Biol Chem*. 288:28369–28381.

Havukainen H, Underhaug J, Wolschin F, Amdam G, Halskau Ø. 2012. A vitellogenin polyserine cleavage site: highly disordered conformation protected from proteolysis by phosphorylation. *J Exp Biol*. 215:1837–1846.

Haydak MH. 1970. Honey bee nutrition. *Annu Rev Entomol*. 15:143–156.

Hayward A, Takahashi T, Bendena WG, Tobe SS, Hui JHL. 2010. Comparative genomic and phylogenetic analysis of vitellogenin and other large lipid transfer proteins in metazoans. *FEBS Lett*. 584:1273–1278.

Heeringen V, J S, Veenstra GJC. 2011. GimmeMotifs: a de novo motif prediction pipeline for ChIP-sequencing experiments. *Bioinformatics*. 27:270–271.

Hegedus D, Erlandson M, Gillott C, Toprak U. 2009. New Insights into Peritrophic Matrix Synthesis, Architecture, and Function. *Annu Rev Entomol*. 54:285–302.

Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, Glass CK. 2010. Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities. *Mol Cell*. 38:576–589.

Hernández López Javier, Wolfgang S, Crailsheim Karl, Riessberger-Gallé Ulrike. 2014. Trans-generational immune priming in honeybees. *Proc R Soc B Biol Sci*. 281:20140454.

Horowitz NH. 1995. One-gene-one-enzyme: Remembering biochemical genetics. *Protein Sci*. 4:1017–1019.

Hu H, Bezabih G, Feng M, Wei Q, Zhang X, Wu F, Meng L, Fang Y, Han B, Ma C. 2019. In-depth Proteome of the Hypopharyngeal Glands of Honeybee Workers Reveals Highly Activated Protein and Energy Metabolism in Priming the Secretion of Royal Jelly. *Mol Cell Proteomics*. mcp–RA118.

Hu XF, Zhang B, Liao CH, Zeng ZJ. 2019. High-Efficiency CRISPR/Cas9-Mediated Gene Editing in Honeybee (Apis mellifera) Embryos. *G3 Bethesda Md*. 9:1759–1766.

Huang DW, Sherman BT, Lempicki RA. 2009a. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc*. 4:44–57.

Huang DW, Sherman BT, Lempicki RA. 2009b. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res*. 37:1–13.

Huang Z-Y, Robinson GE. 1996. Regulation of Honey Bee Division of Labor by Colony Age Demography. *Behav Ecol Sociobiol*. 39:147–158.

Huo Y, Liu W, Zhang F, Chen X, Li L, Liu Q, Zhou Y, Wei T, Fang R, Wang X. 2014. Transovarial Transmission of a Plant Virus Is Mediated by Vitellogenin of Its Insect Vector. *PLOS Pathog*. 10:e1003949.

Hwang S, Gou Z, Kuznetsov IB. 2007. DP-Bind: a web server for sequence-based prediction of DNA-binding residues in DNA-binding proteins. *Bioinformatics*. 23:634–636.

Hystad EM, Salmela H, Amdam GV, Münch D. 2017. Hemocyte-mediated phagocytosis differs between honey bee (Apis mellifera) worker castes. *PLOS ONE*. 12:e0184108.

Ihle KE, Fondrk MK, Page RE, Amdam GV. 2015. Genotype effect on lifespan following vitellogenin knockdown. *Exp Gerontol*. 61:113–122.

Ihle KE, Page RE, Frederick K, Fondrk MK, Amdam GV. 2010. Genotype effect on regulation of behaviour by vitellogenin supports reproductive origin of honeybee foraging bias. *Anim Behav*. 79:1001–1006.

Illa-Bochaca I, Montuenga LM. 2006. The regenerative nidi of the locust midgut as a model to study epithelial cell differentiation from stem cells. *J Exp Biol*. 209:2215–2223.

Isaac PG, Bownes M. 1982. Ovarian and Fat-Body Vitellogenin Synthesis in Drosophila melanogaster. *Eur J Biochem*. 123:527–534.

Jackson AL, Linsley PS. 2010. Recognizing and avoiding siRNA off-target effects for target identification and therapeutic application. *Nat Rev Drug Discov*. 9:57–67.

Jarvis JU. 1981. Eusociality in a mammal: cooperative breeding in naked mole-rat colonies. *Science*. 212:571–573.

Jazwinski SM. 2015. Mitochondria to Nucleus Signaling and the Role of Ceramide in Its Integration into the Suite of Cell Quality Control Processes during Aging Mitochondria to Nucleus Signaling and the Role of Ceramide in Its Integration into the Suite of Cell Quality Control Processes during Aging. *Ageing Res Rev*. 23:67–74.

Jensen PV, Børgesen LW. 2000. Regional and functional differentiation in the fat body of pharaoh's ant queens, Monomorium pharaonis (L.). *Arthropod Struct Dev*. 29:171–184.

Jimenez DR, Gilliam M. 1989. Age-related changes in midgut ultrastructure and trypsin activity in the honey bee, Apis mellifera. *Apidologie*. 20:287–303.

Jimenez DR, Gilliam M. 1990. Ultrastructure of the ventriculus of the honey bee, Apis mellifera (L.): cytochemical localization of acid phosphatase, alkaline phosphatase, and nonspecific esterase. *Cell Tissue Res*. 261:431–443.

Johansson TSK. 1955. Royal Jelly. *Bee World*. 36:21–32.

Johnson BR. 2010. Division of labor in honeybees: form, function, and proximate mechanisms. *Behav Ecol Sociobiol*. 64:305–316.

Kawooya JK, Keim PS, Ryan RO, Shapiro JP, Samaraweera P, Law JH. 1984. Insect apolipophorin III. Purification and properties. *J Biol Chem*. 259:10733–10737.

Keenan J, Murphy L, Henry M, Meleady P, Clynes M. 2009. Proteomics analysis of multidrug-resistance mechanisms in adriamycin-resistant variants of dlkp, a squamour lung cancer cell line. *Proteomics*. 9:1556–1566.

Khan AUH, Allende-Vega N, Gitenay D, Garaude J, Vo D-N, Belkhala S, Gerbal-Chaloin S, Gondeau C, Daujat-Chavanieu M, Delettre C, Orecchioni S, Talarico G, Bertolini F, Anel A, Cuezva JM, Enriquez JA, Cartron G, Lecellier C-H, Hernandez J, Villalba M. 2018. Mitochondrial Complex I activity signals antioxidant response through ERK5. *Sci Rep*. 8:1–14.

Kikuchi Y. 2009. Endosymbiotic Bacteria in Insects: Their Diversity and Culturability. *Microbes Environ*. 24:195–204.

Kim BY, Jin BR. 2015. Apolipophorin III from honeybees (Apis cerana) exhibits antibacterial activity. *Comp Biochem Physiol B Biochem Mol Biol*. 182:6–13.

Kim C-H, Paik D, Rus F, Silverman N. 2014. The caspase-8 homolog Dredd cleaves Imd and Relish but is not inhibited by p35. *J Biol Chem*. jbc.M113.544841.

King JR, Warren RJ, Bradford MA. 2013. Social Insects Dominate Eastern US Temperate Hardwood Forest Macroinvertebrate Communities in Warmer Regions. *PLOS ONE*. 8:e75843.

Kishida M, Specker JL. 2000. Paternal mouthbrooding in the black-chinned tilapia, Sarotherodon melanotheron (Pisces: cichlidae): changes in gonadal steroids and potential for vitellogenin transfer to larvae. *Horm Behav*. 37:40–48.

Klaudiny J, Bachanová K, Kohútová L, Dzúrová M, Kopernický J, Majtán J. 2012. Expression of larval jelly antimicrobial peptide defensin1 in *Apis mellifera* colonies. *Biologia (Bratisl)*. 67:200–211.

Klowden MJ, Davis EE, Bowen MF. 1987. Role of the fat body in the regulation of host-seeking behaviour in the mosquito, Aedes aegypti. *J Insect Physiol*. 33:643–646.

Knight K. 2019. Firebugs fight infection with egg yolk protein. *J Exp Biol*. 222:jeb206185.

Knorr E, Schmidtberg H, Arslan D, Bingsohn L, Vilcinskas A. 2015. Translocation of bacteria from the gut to the eggs triggers maternal transgenerational immune priming in Tribolium castaneum. *Biol Lett*. 11:20150885.

Kohno H, Suenami S, Takeuchi H, Sasaki T, Kubo T. 2016. Production of Knockout Mutants by CRISPR/Cas9 in the European Honeybee, Apis mellifera L. *Zoolog Sci*. 33:505–512.

Krongdang S, Evans JD, Chen Y, Mookhploy W, Chantawannakul P. 2019. Comparative susceptibility and immune responses of Asian and European honey bees to the American foulbrood pathogen, Paenibacillus larvae. *Insect Sci*. 26:831–842.

Kucharski R, Maleszka R. 2003. Transcriptional profiling reveals multifunctional roles for transferrin in the honeybee, Apis mellifera. *J Insect Sci*. 3.

Kucharski R, Mitri C, Grau Y, Maleszka R. 2007. Characterization of a metabotropic glutamate receptor in the honeybee (<Emphasis Type="Italic">Apis mellifera</Emphasis>): implications for memory formation. *Invert Neurosci*. 7:99–108.

Kühnlein RP. 2011. The contribution of the Drosophila model to lipid droplet research. *Prog Lipid Res*. 50:348–356.

Kumar S, Molina-Cruz A, Gupta L, Rodrigues J, Barillas-Mury C. 2010. A Peroxidase/Dual Oxidase System Modulates Midgut Epithelial Immunity in Anopheles gambiae. *Science*. 327:1644–1648.

Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods*. 9:357–359.

Lavine MD, Strand MR. 2002. Insect hemocytes and their role in immunity. *Insect Biochem Mol Biol*. 32:1295–1309.

Lazareva AA, Roman G, Mattox W, Hardin PE, Dauwalder B. 2007. A Role for the Adult Fat Body in Drosophila Male Courtship Behavior. *PLOS Genet*. 3:e16.

Le TT, Duren HM, Slipchenko MN, Hu C-D, Cheng J-X. 2010. Label-free quantitative analysis of lipid metabolism in living Caenorhabditis elegans. *J Lipid Res*. 51:672–677.

Lehane M, Billingsley P. 2012. Biology of the Insect Midgut. Springer Science & Business Media.

Lehane MJ. 1997. Peritrophic Matrix Structure and Function. *Annu Rev Entomol*. 42:525–550.

Lehane MJ, Wu D, Lehane SM. 1997. Midgut-specific immune molecules are produced by the blood-sucking insect Stomoxys calcitrans. *Proc Natl Acad Sci*. 94:11502–11507.

Leulier F, Rodriguez A, Khush RS, Abrams JM, Lemaitre B. 2000. The Drosophila caspase Dredd is required to resist Gram-negative bacterial infection. *EMBO Rep*. 1:353–358.

Levine B, Deretic V. 2007. Unveiling the roles of autophagy in innate and adaptive immunity. *Nat Rev Immunol*. 7:767–777.

Li J, Ma L, Lin Z, Zou Z, Lu Z. 2016. Serpin-5 regulates prophenoloxidase activation and antimicrobial peptide pathways in the silkworm, Bombyx mori. *Insect Biochem Mol Biol*. 73:27–37.

Li Y, Collins M, An J, Geiser R, Tegeler T, Tsantilas K, Garcia K, Pirrotte P, Bowser R. 2016. Immunoprecipitation and mass spectrometry defines an extensive RBM45 protein–protein interaction network. *Brain Res*, RNA Metabolism in Neurological Disease. 1647:79–93.

Li Z, Zhang S, Liu Q. 2008. Vitellogenin Functions as a Multivalent Pattern Recognition Receptor with an Opsonic Activity. *PLoS ONE*. 3:e1940.

Li Z, Zhang S, Zhang J, Liu M, Liu Z. 2009. Vitellogenin is a cidal factor capable of killing bacteria via interaction with lipopolysaccharide and lipoteichoic acid. *Mol Immunol*. 46:3232–3239.

Little TJ, O'Connor B, Colegrave N, Watt K, Read AF. 2003. Maternal Transfer of Strain-Specific Immunity in an Invertebrate. *Curr Biol*. 13:489–492.

Liu Q-H, Zhang S-C, Li Z-J, Gao C-R. 2009. Characterization of a pattern recognition molecule vitellogenin from carp (Cyprinus carpio). *Immunobiology*. 214:257–267.

Liu R, Hu J. 2013. DNABind: A hybrid algorithm for structure-based prediction of DNA-binding residues by combining machine learning- and template-based approaches. *Proteins Struct Funct Bioinforma*. 81:1885–1899.

López JH, Schuehly W, Crailsheim K, Riessberger-Gallé U. 2014a. Trans-generational immune priming in honeybees. *Proc R Soc Lond B Biol Sci*. 281:20140454.

López JH, Schuehly W, Crailsheim K, Riessberger-Gallé U. 2014b. Trans-generational immune priming in honeybees. *Proc R Soc B*. 281:20140454.

López-Uribe MM, Fitzgerald A, Simone-Finstrom M. 2017. Inducible versus constitutive social immunity: examining effects of colony infection on glucose oxidase and defensin-1 production in honeybees. *R Soc Open Sci*. 4:170224.

Lourenço AP, Mackert A, Cristino A dos S, Simões ZLP. 2008. Validation of reference genes for gene expression studies in the honey bee, Apis mellifera, by quantitative real-time RT-PCR. *Apidologie*. 39:372–385.

Lourenço AP, Martins JR, Bitondi MMG, Simões ZLP. 2009. Trade-off between immune stimulation and expression of storage protein genes. *Arch Insect Biochem Physiol*. 71:70–87.

Ludtke SJ, Baldwin PR, Chiu W. 1999. EMAN: Semiautomated Software for High-Resolution Single-Particle Reconstructions. *J Struct Biol*. 128:82–97.

Lycett GJ, McLaughlin LA, Ranson H, Hemingway J, Kafatos FC, Loukeris TG, Paine MJI. 2006. Anopheles gambiae P450 reductase is highly expressed in oenocytes and in vivo knockdown increases permethrin susceptibility. *Insect Mol Biol*. 15:321–327.

Maeterlinck M. 1901. The Life of the Bee. Dodd, Mead.

Makki R, Cinnamon E, Gould AP. 2014. The Development and Functions of Oenocytes. *Annu Rev Entomol*. 59:405–425.

Mann CJ, Anderson TA, Read J, Chester SA, Harrison GB, Köchl S, Ritchie PJ, Bradbury P, Hussain FS, Amey J, Vanloo B, Rosseneu M, Infante R, Hancock JM, Levitt DG, Banaszak LJ, Scott J, Shoulders CC. 1999. The structure of vitellogenin provides a molecular model for the assembly and secretion of atherogenic lipoproteins11Edited by A. R. Fersht. *J Mol Biol*. 285:391–408.

Mao W, Schuler MA, Berenbaum MR. 2013. Honey constituents up-regulate detoxification and immunity genes in the western honey bee Apis mellifera. *Proc Natl Acad Sci*. 201303884.

Maori E, Garbian Y, Kunik V, Mozes-Koch R, Malka O, Kalev H, Sabath N, Sela I, Shafir S. 2019. A Transmissible RNA Pathway in Honey Bees. *Cell Rep*. 27:1949–1959.e6.

Marringa WJ, Krueger MJ, Burritt NL, Burritt JB. 2014. Honey Bee Hemocyte Profiling by Flow Cytometry. *PLOS ONE*. 9:e108486.

Martins GF, Romalho-Ortigao JM. 2012. Oenocytes in insects. *Invertebr Surviv J*. 9:139–152.

Martins GF, Serrão JE, Ramalho-Ortigão JM, Pimenta PFP. 2011. Histochemical and ultrastructural studies of the mosquito Aedes aegypti fat body: effects of aging and diet type. *Microsc Res Tech*. 74:1032–1039.

Martinson VG, Moy J, Moran NA. 2012. Establishment of Characteristic Gut Bacteria during Development of the Honeybee Worker. *Appl Environ Microbiol*. 78:2830–2840.

Medzhitov R. 2001. Toll-like receptors and innate immunity. *Nat Rev Immunol*. 1:135–145.

Mellacheruvu D, Wright Z, Couzens AL, Lambert J-P, St-Denis N, Li T, Miteva YV, Hauri S, Sardiu ME, Low TY, Halim VA, Bagshaw RD, Hubner NC, al-Hakim A, Bouchard A, Faubert D, Fermin D, Dunham WH, Goudreault M, Lin Z-Y, Badillo BG, Pawson T, Durocher D, Coulombe B, Aebersold R, Superti-Furga G, Colinge J, Heck AJR, Choi H, Gstaiger M, Mohammed S, Cristea IM, Bennett KL, Washburn MP, Raught B, Ewing RM, Gingras A-C, Nesvizhskii AI. 2013. The CRAPome: a Contaminant Repository for Affinity Purification Mass Spectrometry Data. *Nat Methods*. 10:730–736.

Meunier J. 2015. Social immunity and the evolution of group living in insects. *Philos Trans R Soc B Biol Sci*. 370:20140102.

Micchelli CA, Perrimon N. 2006. Evidence that stem cells reside in the adult Drosophila midgut epithelium. *Nature*. 439:475–479.

Michener CD. 1969a. Comparative social behavior of bees. *Annu Rev Entomol*. 14:299–342.

Michener CD. 1969b. Comparative Social Behavior of Bees. *Annu Rev Entomol*. 14:299–342.

Milojevic BD. 1940. A New Interpretation of the Social Life of the Honey Bee. *Bee World*. 21:39–41.

Mohammed H, Taylor C, Brown GD, Papachristou EK, Carroll JS, D'Santos CS. 2016. Rapid immunoprecipitation mass spectrometry of endogenous proteins (RIME) for analysis of chromatin complexes. *Nat Protoc*. 11:316–326.

Moran AJ, Hills M, Gunton J, Nano FE. 2001. Heat-labile proteases in molecular biology applications. *FEMS Microbiol Lett*. 197:59–63.

Morandin C, Havukainen H, Kulmuni J, Dhaygude K, Trontti K, Helanterä H. 2014. Not only for egg yolk--functional and evolutionary insights from expression, selection, and structural analyses of Formica ant vitellogenins. *Mol Biol Evol*. 31:2181–2193.

Moreno-García M, Recio-Tótoro B, Claudio-Piedras F, Lanz-Mendoza H. 2014. Injury and immune response: applying the danger theory to mosquitoes. *Front Plant Sci*. 5.

Moret Y. 2006. "Trans-generational immune priming": specific enhancement of the antimicrobial immune response in the mealworm beetle, Tenebrio molitor. *Proc R Soc B Biol Sci*. 273:1399–1405.

Moret Y, Moreau J. 2012. The immune role of the arthropod exoskeleton. *Invertebr Surviv J*. 9:200–206.

Morishima I, Yamano Y, Inoue K, Matsuo N. 1997. Eicosanoids mediate induction of immune genes in the fat body of the silkworm, Bombyx mori. *FEBS Lett*. 419:83–86.

Moritz B, Crailsheim K. 1987. Physiology of protein digestion in the midgut of the honeybee (Apis mellifera L.). *J Insect Physiol*. 33:923–931.

Morse R, Flottum K. 1997. Honey bee pests, predators, and diseases. Ithica, NY: A.I. Root Company.

Münch D, Ihle KE, Salmela H, Amdam GV. 2015. Vitellogenin in the honey bee brain: Atypical localization of a reproductive protein that promotes longevity. *Exp Gerontol*, Aging in the Wild: Insights from Free-Living and Non-Model organisms. 71:103–108.

Nakahira K, Haspel JA, Rathinam VAK, Lee S-J, Dolinay T, Lam HC, Englert JA, Rabinovitch M, Cernadas M, Kim HP, Fitzgerald KA, Ryter SW, Choi AMK. 2011. Autophagy proteins regulate innate immune responses by inhibiting the release of mitochondrial DNA mediated by the NALP3 inflammasome. *Nat Immunol*. 12:222–230.

Nakajima Y, Van Der Goes van Naters-Yasui A, Taylor D, Yamakawa M. 2002. Antibacterial peptide defensin is involved in midgut immunity of the soft tick, Ornithodoros moubata. *Insect Mol Biol*. 11:611–618.

Nakamura A, Yasuda K, Adachi H, Sakurai Y, Ishii N, Goto S. 1999. Vitellogenin-6 Is a Major Carbonylated Protein in Aged Nematode, Caenorhabditis elegans. *Biochem Biophys Res Commun*. 264:580–583.

Nelson CM, Ihle KE, Fondrk MK, Page RE Jr, Amdam GV. 2007. The Gene vitellogenin Has Multiple Coordinating Effects on Social Organization. *PLoS Biol*. 5:e62.

Noah Koller C, Dhadialla TS, Raikhel AS. 1989. Selective endocytosis of vitellogenin by oocytes of the mosquito, Aedes aegypti: An in vitro study. *Insect Biochem*. 19:693–702.

Nunes FMF, Ihle KE, Mutti NS, Simões ZLP, Amdam GV. 2013. The gene vitellogenin affects microRNA regulation in honey bee (Apis mellifera) fat body and brain. *J Exp Biol*. 216:3724–3732.

Nunes FMF, Simões ZLP. 2009. A non-invasive method for silencing gene transcription in honeybees maintained under natural conditions. *Insect Biochem Mol Biol*. 39:157–160.

Ohlstein B, Spradling A. 2006. The adult Drosophila posterior midgut is maintained by pluripotent stem cells. *Nature*. 439:470–474.

Okamoto I, Taniguchi Y, Kunikata T, Kohno K, Iwaki K, Ikeda M, Kurimoto M. 2003. Major royal jelly protein 3 modulates immune responses in vitro and in vivo. *Life Sci*. 73:2029–2045.

Oliveira VTP de, Cruz-Landim C da. 2006. Morphology and function of insect fat body cells: a review. *Biociências -Line*. 11.

Oxley PR, Thompson GJ, Oldroyd BP. 2008. Four Quantitative Trait Loci That Influence Worker Sterility in the Honeybee (Apis mellifera). *Genetics*. 179:1337–1343.

Pan ML, Bell WJ, Telfer WH. 1969. Vitellogenic Blood Protein Synthesis by Insect Fat Body. *Science*. 165:393–394.

Park MS, Park P, Takeda M. 2009. Starvation induces apoptosis in the midgut nidi of Periplaneta americana: a histochemical and ultrastructural study. *Cell Tissue Res*. 335:631–638.

Patel N, Haydak MH, Gochnauer T. 1960. Electrophoretic components of the proteins in honey bee larval food. *Nature*. 186:633–634.

Peiren N, de Graaf DC, Vanrobaeys F, Danneels EL, Devreese B, Van Beeumen J, Jacobs FJ. 2008. Proteomic analysis of the honey bee worker venom gland focusing on the mechanisms of protection against tissue damage. *Toxicon*. 52:72–83.

Pfeiffer A, Shi H, Tepperman JM, Zhang Y, Quail PH. 2014. Combinatorial Complexity in a Transcriptionally Centered Signaling Hub in Arabidopsis. *Mol Plant*. 7:1598–1618.

Pham CG, Bubici C, Zazzeroni F, Papa S, Jones J, Alvarez K, Jayawardena S, De Smaele E, Cong R, Beaumont C. 2004. Ferritin heavy chain upregulation by NF-κB inhibits TNFα-induced apoptosis by suppressing reactive oxygen species. *Cell*. 119:529–542.

Pilon AM, Ajay SS, Kumar SA, Steiner LA, Cherukuri PF, Wincovitch S, Anderson SM, Mullikin JC, Gallagher PG, Hardison RC, Margulies EH, Bodine DM. 2011. Genome-wide ChIP-Seq reveals a dramatic shift in the binding of the transcription factor erythroid Kruppel-like factor during erythrocyte differentiation. *Blood*. 118:e139–e148.

Raag R, Appelt K, Xuong N-H, Banaszak L. 1988a. Structure of the lamprey yolk lipid-protein complex lipovitellin-phosvitin at 2.8 Å resolution. *J Mol Biol*. 200:553–569.

Raag R, Appelt K, Xuong N-H, Banaszak L. 1988b. Structure of the lamprey yolk lipid-protein complex lipovitellin-phosvitin at 2.8 Å resolution. *J Mol Biol*. 200:553–569.

Raikhel AS, Dhadialla TS. 1992. Accumulation of Yolk Proteins in Insect Oocytes. *Annu Rev Entomol*. 37:217–251.

Raikhel AS, Lea AO. 1983. Previtellogenic development and vitellogenin synthesis in the fat body of a mosquito: An ultrastructural and immunocytochemical study. *Tissue Cell*. 15:281–299.

Rajagopal R, Sivakumar S, Agrawal N, Malhotra P, Bhatnagar RK. 2002. Silencing of Midgut Aminopeptidase N of Spodoptera litura by Double-stranded RNA Establishes Its Role asBacillus thuringiensis Toxin Receptor. *J Biol Chem*. 277:46849–46851.

Ramírez F, Dündar F, Diehl S, Grüning BA, Manke T. 2014. deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res*. 42:W187–W191.

Riessberger-Gallé U, Hernández-López J, Rechberger G, Crailsheim K, Schuehly W. 2016. Lysophosphatidylcholine acts in the constitutive immune defence against American foulbrood in adult honeybees. *Sci Rep*. 6:1–10.

Robinson GE. 1992. Regulation of division of labor in insect societies. *Annu Rev Entomol*. 37:637–665.

Romanelli Alessandra, Moggio Loredana, Montella Rosa Chiara, Campiglia Pietro, Iannaccone Marco, Capuano Federico, Pedone Carlo, Capparelli Rosanna. 2011. Peptides from Royal Jelly: studies on the antimicrobial activity of jelleins, jelleins analogs and synergy with temporins. *J Pept Sci*. 17:348–352.

Rono MK, Whitten MMA, Oulad-Abdelghani M, Levashina EA, Marois E. 2010. The Major Yolk Protein Vitellogenin Interferes with the Anti-Plasmodium Response in the Malaria Mosquito Anopheles gambiae. *PLOS Biol*. 8:e1000434.

Roth A, Vleurinck C, Netschitailo O, Bauer V, Otte M, Kaftanoglu O, Page RE, Beye M. 2019. A genetic switch for worker nutrition-mediated traits in honeybees. *PLOS Biol*. 17:e3000171.

Roth O, Sadd BM, Schmid-Hempel P, Kurtz J. 2009. Strain-specific priming of resistance in the red flour beetle, Tribolium castaneum. *Proc R Soc Lond B Biol Sci*. 276:145–151.

Roth Z, Weil S, Aflalo ED, Manor R, Sagi A, Khalaila I. 2013. Identification of Receptor-Interacting Regions of Vitellogenin within Evolutionarily Conserved β-Sheet Structures by Using a Peptide Array. *ChemBioChem*. 14:1116–1122.

Roy-Zokan EM, Cunningham CB, Hebb LE, McKinney EC, Moore AJ. 2015. Vitellogenin and vitellogenin receptor gene expression is associated with male and female parenting in a subsocial insect. *Proc R Soc B*. 282:20150787.

Sadd BM, Kleinlogel Y, Schmid-Hempel R, Schmid-Hempel P. 2005. Trans-generational immune priming in a social insect. *Biol Lett*. 1:386–388.

Sadd BM, Schmid-Hempel P. 2006. Insect Immunity Shows Specificity in Protection upon Secondary Pathogen Exposure. *Curr Biol*. 16:1206–1210.

Sahinler N, Kaftanoglu O. 1997. Effects of Feeding, Age of the Larvae, and Queenlessness on the Production of Royal Jelly. In: Mizrahi A,, Lensky Y, editors. Bee Products: Properties, Applications, and Apitherapy. Boston, MA: Springer US. p. 173–178.

Salmela H, Amdam GV, Freitak D. 2015. Transfer of Immunity from Mother to Offspring Is Mediated via Egg-Yolk Protein Vitellogenin. *PLoS Pathog*. 11:e1005015.

Salmela H, Freitak D. 2017. Edible vaccination against microbial pathogens.

Salmela H, Stark T, Stucki D, Fuchs S, Freitak D, Dey A, Kent CF, Zayed A, Dhaygude K, Hokkanen H, Sundström L. 2016. Ancient Duplications Have Led to Functional Divergence of Vitellogenin-Like Genes Potentially Involved in Inflammation and Oxidative Stress in Honey Bees. *Genome Biol Evol*. 8:495–506.

Salmela H, Sundström L. 2017. Vitellogenin in inflammation and immunity in social insects. *Inflamm Cell Signal*.

Salton MRJ, Kim K-S. 1996. Structure. In: Baron S, editor. Medical Microbiology. 4th ed. Galveston (TX): University of Texas Medical Branch at Galveston.

Sappington TW, S. Raikhel A. 1998. Molecular characteristics of insect vitellogenins and vitellogenin receptors. *Insect Biochem Mol Biol*. 28:277–300.

Sarrazin S, Lamanna WC, Esko JD. 2011. Heparan Sulfate Proteoglycans. *Cold Spring Harb Perspect Biol*. 3.

Scharlaken B, de Graaf DC, Goossens K, Brunain M, Peelman LJ, Jacobs FJ. 2008. Reference Gene Selection for Insect Expression Studies Using Quantitative Real-Time PCR: The Head of the Honeybee, Apis mellifera, After a Bacterial Challenge. *J Insect Sci*. 8:1–10.

Scheiner R, Amdam GV. 2009. Impaired tactile learning is related to social role in honeybees. *J Exp Biol*. 212:994–1002.

Scheres SHW, Núñez-Ramírez R, Sorzano COS, Carazo JM, Marabini R. 2008. Image processing for electron microscopy single-particle analysis using XMIPP. *Nat Protoc*. 3:977–990.

Schmid MR, Brockmann A, Pirk CWW, Stanley DW, Tautz J. 2008. Adult honeybees (Apis mellifera L.) abandon hemocytic, but not phenoloxidase-based immunity. *J Insect Physiol*. 54:439–444.

Schmid-Hempel P. 2001. On the evolutionary ecology of host–parasite interactions: addressing the question with regard to bumblebees and their parasites. *Naturwissenschaften*. 88:147–158.

Schmittgen TD, Livak KJ. 2008. Analyzing real-time PCR data by the comparative CT method. *Nat Protoc*. 3:1101–1108.

Schubert SY, Ilan N, Shushy M, Ben-Izhak O, Vlodavsky I, Goldshmidt O. 2004. Human heparanase nuclear localization and enzymatic activity. *Lab Invest*. 84:535–544.

Seehuus S-C, Norberg K, Gimsa U, Krekling T, Amdam GV. 2006. Reproductive protein protects functionally sterile honey bee workers from oxidative stress. *Proc Natl Acad Sci U S A*. 103:962–967.

Seehuus S-C, Norberg K, Krekling T, Fondrk K, Amdam GV. 2007. Immunogold Localization of Vitellogenin in the Ovaries, Hypopharyngeal Glands and Head Fat Bodies of Honeybee Workers, Apis Mellifera. *J Insect Sci*. 7:1–14.

Seeley TD. 1982. Adaptive significance of the age polyethism schedule in honeybee colonies. *Behav Ecol Sociobiol*. 11:287–293.

Segrest JP, Jones MK, Dashti N. 1999. N-terminal domain of apolipoprotein B has structural homology to lipovitellin and microsomal triglyceride transfer protein: a "lipid pocket" model for self-assembly of apoB-containing lipoprotein particles. *J Lipid Res*. 40:1401–1416.

Serang O, Cansizoglu AE, Käll L, Steen H, Steen JA. 2013. Nonparametric Bayesian evaluation of differential protein quantification. *J Proteome Res*. 12.

Shahid NS, Steinhoff MC, Roy E, Begum T, Thompson CM, Siber GR. 2002. Placental and breast transfer of antibodies after maternal immunization with polysaccharide meningococcal vaccine: a randomized, controlled evaluation. *Vaccine*. 20:2404–2409.

Shi X, Zhang S, Pang Q. 2006. Vitellogenin is a novel player in defense reactions. *Fish Shellfish Immunol*. 20:769–772.

Shikina S, Chen C-J, Chung Y-J, Shao Z-F, Liou J-Y, Tseng H-P, Lee Y-H, Chang C-F. 2013. Yolk Formation in a Stony Coral Euphyllia ancora (Cnidaria, Anthozoa): Insight Into the Evolution of Vitellogenesis in Nonbilaterian Animals. *Endocrinology*. 154:3447–3459.

Shimanuki H. 1997. Bacteria. In: Morse RA,, Flottum K, editors. Honey bee pests, predators, and diseases. Medina, OH: A.I. Root Company. p. 33–54.

Silva-Olivares A, Díaz E, Shibayama M, Tsutsumi V, Cisneros R, Zúñiga G. 2003. Ultrastructural study of the midgut and hindgut in eight species of the genus Dendroctonus Erichson (Coleoptera: Scolytidae). *Ann Entomol Soc Am*. 96:883–900.

Smedal B, Brynem M, Kreibich CD, Amdam GV. 2009. Brood pheromone suppresses physiology of extreme longevity in honeybees (Apis mellifera). *J Exp Biol*. 212:3795–3801.

Smith RL, Soeters MR, Wüst RCI, Houtkooper RH. 2018. Metabolic Flexibility as an Adaptation to Energy Resources and Requirements in Health and Disease. *Endocr Rev*. 39:489–517.

Smolenaars MMW, Madsen O, Rodenburg KW, Horst DJV der. 2007. Molecular diversity and evolution of the large lipid transfer protein superfamily. *J Lipid Res*. 48:489–502.

Snodgrass RE. 1956. Anatomy of the Honey Bee. Cornell University Press.

Snodgrass RE. 2018. Principles of Insect Morphology. Cornell University Press.

Souza DAD, Kaftanoglu O, Jong DD, Page RE, Amdam GV, Wang Y. 2018. Differences in the morphology, physiology and gene expression of honey bee queens and workers reared in vitro versus in situ. *Biol Open*. 7:bio036616.

Stewart MD, Sanderson RD. 2014. Heparan sulfate in the nucleus and its control of cellular functions. *Matrix Biol*, Proteoglycan Biology. 35:56–59.

Stokes BA, Yadav S, Shokal U, Smith LC, Eleftherianos I. 2015. Bacterial and fungal pattern recognition receptors in homologous innate signaling pathways of insects and mammals. *Front Microbiol*. 6.

Storz P. 2006. Reactive Oxygen Species–Mediated Mitochondria-to-Nucleus Signaling: A Key to Aging and Radical-Caused Diseases. *Sci STKE*. 2006:re3-re3.

Storz P, Döppler H, Toker A. 2005. Protein Kinase D Mediates Mitochondrion-to-Nucleus Signaling and Detoxification from Mitochondrial Reactive Oxygen Species. *Mol Cell Biol*. 25:8520–8530.

Sugiyama T, Takahashi K, Mori H. 2012. Royal Jelly Acid, 10-Hydroxy-trans-2-Decenoic Acid, as a Modulator of the Innate Immune Responses. *Endocr Metab Immune Disord - Drug TargetsFormerly Curr Drug Targets - Immune Endocr Metab Disord*. 12:368–376.

Sun C, Hu L, Liu S, Gao Z, Zhang S. 2013a. Functional analysis of domain of unknown function (DUF) 1943, DUF1944 and von Willebrand factor type D domain (VWD) in vitellogenin2 in zebrafish. *Dev Comp Immunol*. 41:469–476.

Sun C, Hu L, Liu S, Gao Z, Zhang S. 2013b. Functional analysis of domain of unknown function (DUF) 1943, DUF1944 and von Willebrand factor type D domain (VWD) in vitellogenin2 in zebrafish. *Dev Comp Immunol*. 41:469–476.

Sun C, Zhang S. 2015a. Immune-Relevant and Antioxidant Activities of Vitellogenin and Yolk Proteins in Fish. *Nutrients*. 7:8818–8829.

Sun C, Zhang S. 2015b. Immune-Relevant and Antioxidant Activities of Vitellogenin and Yolk Proteins in Fish. *Nutrients*. 7:8818–8829.

Teixeira A das D, Marques-Araújo S, Zanuncio JC, Serrão JE. 2015. Peritrophic membrane origin in adult bees (Hymenoptera): Immunolocalization. *Micron*. 68:91–97.

Thompson DM, Khalil SMS, Jeffers LA, Sonenshine DE, Mitchell RD, Osgood CJ, Michael Roe R. 2007. Sequence and the developmental and tissue-specific regulation of the first complete vitellogenin messenger RNA from ticks responsible for heme sequestration. *Insect Biochem Mol Biol*. 37:363–374.

Thompson GJ, Kucharski R, Maleszka R, Oldroyd BP. 2006. Towards a molecular definition of worker sterility: differential gene expression and reproductive plasticity in honey bees. *Insect Mol Biol*. 15:537–644.

Tidbury HJ, Pedersen AB, Boots M. 2011. Within and transgenerational immune priming in an insect to a DNA virus. *Proc R Soc Lond B Biol Sci*. 278:871–876.

Tjong H, Zhou H-X. 2007. DISPLAR: an accurate method for predicting DNA-binding sites on protein surfaces. *Nucleic Acids Res*. 35:1465–1477.

Tong Z, Li L, Pawar R, Zhang S. 2010. Vitellogenin is an acute phase protein with bacterial-binding and inhibiting activities. *Immunobiology*. 215:898–902.

Townsend GF, Lucas CC. 1940. The chemical nature of royal jelly. *Biochem J*. 34:1155–1162.

Traniello JFA, Rosengaus RB, Savoie K. 2002. The development of immunity in a social insect: Evidence for the group facilitation of disease resistance. *Proc Natl Acad Sci U S A*. 99:6838–6842.

Tufail M, Takeda M. 2008. Molecular characteristics of insect vitellogenins. *J Insect Physiol*. 54:1447–1458.

Uttenweiler-Joseph S, Moniatte M, Lagueux M, Dorsselaer AV, Hoffmann JA, Bulet P. 1998. Differential display of peptides induced during the immune response of Drosophila: A matrix-assisted laser desorption ionization time-of-flight mass spectrometry study. *Proc Natl Acad Sci*. 95:11342–11347.

APPENDIX A

CHAPTER 5

Supplementary Information Text

Extended technical description of vitellogenin cutting inhibition

Results. To determine how the N-sheet domain is cleaved from the Vg molecule, we searched for candidate enzymes by applying an array of protease inhibitor molecules on fat body tissue homogenate rich in Vg. Use of lysate instead of pure target protein is necessary, since pure sample has been cleared of proteases (1). Without any protease inhibitors, full length Vg gets fragmented at room temperature in the tissue homogenate in 2 h (Fig. S2; the "control" sample kept on ice versus the "no inhibitors" samples). We included broad-range inhibitors and the highly specific caspase 1 inhibitor Ac-Tyr-Val-Ala-Asp-acyloxymethyl ketone (YVAD-aomk), because there is a caspase 1 cut site in the Vg sequence that, hypothetically, would produce an N-terminal fragment of 40 kDa (2). Vg cutting was partially inhibited by YVAD-aomk, leupeptin (serine and cysteine protease inhibitor), EDTA (metalloprotease and phosphatase inhibitor) and phophatase inhibitor cocktail, but not by E64 (inhibitor of papain-like, but not caspase-like cysteine proteases (3)) or 3,4-dichloroisocoumarin (DCI) (serine protease inhibitor). These results show that Vg is cleaved by at least a caspase 1-like enzyme.

Discussion. This study expands our understanding of the enzymatic activity required for Vg cutting (Fig. S2). Honey bee Vg N-sheet cutting site is located in a polyserine linker, whose dephosphorylation appears play a role in the proteolytic cleavage (4). Our inhibition assay strengthens the significance of the removal of phosphate groups, since we show that phosphatase inhibitors inhibit Vg cutting. The metalloprotease inhibitor EDTA inhibited Vg cutting similarly to phosphatase inhibitors. EDTA is known to inhibit phosphatases (5), which may explain the result. Alternatively, EDTA might inhibit some

other enzymes possibly involved in the cascade that leads into Vg cleavage. Moreover, this assay predicts that the cutting of Vg N-sheet is regulated by at least caspase 1-like enzyme, because there is cutting inhibition by the highly specific caspase 1 inhibitor YVAD-aomk and also by leupeptin whose inhibition range covers caspases. We were led to these findings by an earlier realization that caspase 1 is the only relevant (non-gut) enzyme with a cut site in the polyserine linker (4). In mammals, it is well-established that caspase 1 activity is regulated by phosphorylation events (6). The mammalian caspase 1 is involved in many processes, most notably, inflammation and response to intracellular bacterial infection (reviewed by (7)). However, mammalian caspase activities do not directly translate to insects. Very little information is currently available about honey bee caspases in the literature.

Method. Three winter worker honey bee individuals were anesthetized in cold, and their guts and ovaries were removed like before (8). The abdomens were detached, immediately cooled in liquid nitrogen and stored in -80ºC. The abdomens were homogenized in 1.5 ml ice cold PBS and insoluble material was removed by centrifugation. The inhibitors tested were (final concentration according to manufacturer's instructions): E64 (10 µM), leupeptin (100 µM), DCI (100 µM), YVAD-aomk (100 µM), EDTA (5 mM) and Roche PhosSTOP inhibitor cocktail (x 2). The inhibitors were incubated with 9 µl of honey bee protein extract in PBS in total volume of 10 µl for 2 h in 28 ºC in triplicates and blotted. As controls, we had samples without inhibitors, and a sample that was kept on ice.

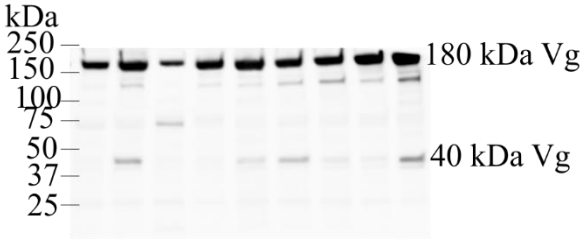

Fig. S1. Western blot of vitellogenin N-terminal antibody (targeting amino acids 24-360) in honey bee tissue lysate samples. Abdomen of nine winter worker honey bee individuals were homogenized and western blotted (20 µg total protein per lane). There is a strong full-length vitellogenin band (180 kDa) in each sample. In addition, there are shorter fragments, whose presence and strength varies in individual samples. The previously identified 40 kDa N-terminal vitellogenin fragment is indicated. The other fragments of unknown function detected by this antibody in some individuals are ~75 kDa and ~125 kDa.

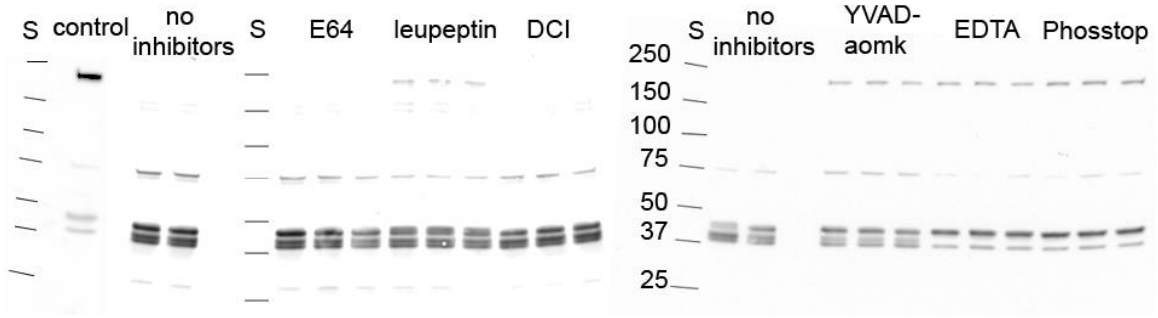

Fig. S2. Western blot -based honey bee vitellogenin cutting inhibition assay. Un-cut, full-length vitellogenin is marked with 180 and the N-terminal domain is marked with 40 according to their size. S = size standard. The control is honey bee fat body protein extract kept on ice. The full-length vitellogenin is fully cut in 2 h in 28 °C in the absence of inhibitors (no inhibitors, N=2). The following treatments were found to prevent the cutting of vitellogenin: leupeptin, YVAD-aomk, EDTA and PhosSTOP. The treatments were done in triplicates. The figure is a combination of two blots.

APPENDIX B

CHAPTER 6

Table S1: Proportion of each RNA-seq sample that aligned to 2 viral genomes, Varroa Destructor Virus-1 (VDV) and Deformed Wing Virus (DWV)

CASTE	COLONY ID	VDV	DWV
Nurse	1	0.25	0.00
Nurse	2	0.10	0.15
Nurse	3	0.08	0.13
Forager	1	0.37	0.01
Forager	2	0.15	0.59
Forager	3	0.29	0.26

APPENDIX C

SUPLLEMENTARY REFERENCES

1. Kondo T, Yokokura T, Nagata S (1997) Activation of distinct caspase-like proteases by Fas and reaper in Drosophila cells. *Proc Natl Acad Sci* 94(22):11951–11956.

2. Havukainen H, Halskau Ø, Skjaerven L, Smedal B, Amdam GV (2011) Deconstructing honeybee vitellogenin: novel 40 kDa fragment assigned to its N terminus. *J Exp Biol* 214(4):582–592.

3. Rozman-Pungerčar J, et al. (2003) Inhibition of papain-like cysteine proteases and legumain by caspase-specific inhibitors: when reaction mechanism is more important than specificity. *Cell Death Differ* 10(8):881–888.

4. Havukainen H, Underhaug J, Wolschin F, Amdam G, Halskau Ø (2012) A vitellogenin polyserine cleavage site: highly disordered conformation protected from proteolysis by phosphorylation. *J Exp Biol* 215(11):1837–1846.

5. Whisnant AR, Gilman SD (2002) Studies of reversible inhibition, irreversible inhibition, and activation of alkaline phosphatase by capillary electrophoresis. *Anal Biochem* 307(2):226–234.

6. Yang J, Liu Z, Xiao TS (2017) Post-translational regulation of inflammasomes. *Cell Mol Immunol* 14(1):65–79.

7. Denes A, Lopez-Castejon G, Brough D (2012) Caspase-1: is IL-1 just the tip of the *ICEberg*? *Cell Death Dis* 3(7):e338.

8. Havukainen H, Halskau Ø, Skjaerven L, Smedal B, Amdam GV (2011) Deconstructing honeybee vitellogenin: novel 40 kDa fragment assigned to its N terminus. *J Exp Biol* 214(4):582–592.