

© 2010 by Ying Wang. All rights reserved.

MOLECULAR MECHANISMS OF PHENOTYPIC PLASTICITY: EXPLORATIONS  
OF DNA METHYLATION AND TRANSCRIPTIONAL REGULATION USING THE  
HONEY BEE

BY

YING WANG

DISSERTATION

Submitted in partial fulfillment of the requirements  
for the degree of Doctor of Philosophy in Cell and Developmental Biology  
in the Graduate College of the  
University of Illinois at Urbana-Champaign, 2010

Urbana, Illinois

Doctoral Committee:

Professor David Clayton, Chair  
Professor Gene Robinson, Director of Research  
Assistant Professor Craig Mizzen, Co-Director of Research  
Assistant Professor Peter Jones  
Professor Hugh Robertson  
Assistant Professor Christopher Schoenherr

## **Abstract**

Honey bee behavioral maturation exemplifies phenotypic plasticity, the ability of a single genotype to produce multiple phenotypes in response to environmental conditions (Scheiner, 1993; West-Eberhard, 1989). Behavioral maturation in honey bees is characterized by a socially regulated transition of adult worker bees from brood caring (“nursing”) to foraging outside hives (Winston, 1991). It involves a suite of behavioral and physiological changes regulated by multiple environmental and internal factors. With the sequenced honey bee genome (Honey Bee Genome Sequencing Consortium, 2006), and development of genomic techniques, honey bee behavioral maturation provides a platform to understand molecular mechanisms underlying phenotypic plasticity.

It has been demonstrated that phenotypic plasticity is associated with massive gene expression changes (Aubin-Horth & Renn, 2009). Therefore, it is essential to understand mechanisms of transcriptional regulation related to alternative phenotypes. This thesis focuses on two important aspects of transcriptional regulation: DNA methylation and transcription regulation by a nuclear receptor, Ultraspiracle (USP).

Chapter 2 presents results that demonstrate that there is a functional CpG methylation system in honey bees, the first well-characterized functional DNA methylation system in insects. Catalytically active orthologs of vertebrate DNA methyltransferase (Dnmts) were identified along with methylated genes. Compared to vertebrate genomes, DNA methylations in honey bees are sparse, and all of the identified methylations are intragenic, similar to results shown in other invertebrates. Differences in DNA methylation density and positions suggest possibly different functions of DNA methylation in vertebrates and invertebrates. The characterization of a DNA methylation

system in honey bees paves the way to study its function in honey bee behavioral maturation. This paper was published in 2006 in *Science* 314: 645-647.

Chapter 3 presents a review of nutritional regulation of honey bee behavioral maturation and propose systems biology approaches to reveal the gene regulatory networks (GRNs) underlying behavioral maturation. It was published in 2010 in *Wiley Interdisciplinary Reviews: Systems Biology and Medicine* (published online in advance of print).

Chapter 4 presents results on transcriptional regulation related to behavioral maturation in the context of fat bodies gene expression change. As a first step, we focus on transcription regulation by USP, a nuclear receptor implicated in JH signaling. USP was first demonstrated to regulate behavioral maturation. *usp* knockdown in honey bee workers' fat bodies significantly delayed onset of foraging. Genome-wide USP binding sites were then mapped in the fat bodies of nurses and foragers respectively. No difference of USP binding was detected, suggesting USP regulates gene expression depending on the availability of co-factors or ligands and chromatin environments. Functional analysis of USP putative genes reveals several genes involved in multiple signaling pathways in behavioral maturation, raising the possibility that USP works as a master regulator to integrate different signals during honey bee behavioral maturation.

## Acknowledgments

Graduate school has been a wonderful and transforming experience for me thanks to many great people I met at the University of Illinois, Urbana-Champaign. I would like to express my gratitude to my advisers, Gene Robinson and Craig Mizzen, for being always enthusiastic, patient and encouraging. I am especially thankful for their great advice and solid support to my career choice. I would like to thank other members of my committee, David Clayton, Hugh Roberson, Peter Jones and Christopher Schoenherr, for their special interest, excellent advice, and effort to make time for me. Peter Jones deserves special thanks. He taught me essential lab skills, and his optimism always cheers me up. I also want to thank other faculty members in the department of Cell and Developmental Biology especially Brian Freeman, and Lisa Stubbs for their interest and support. Thomas Newman and Karen Pruiett offered many forms of assistance and advice in Robinson lab and bee lab, without which my research projects could not go smoothly. My colleagues in Robinson lab are fun, hard-working and intelligent, who stimulated, helped, and entertained me. I would like to thank Miguel Corona and Zhengzheng Liang for their company in late night experiments, and Seth Ament for his brilliance and non-stop challenges. My friends and housemates in Brooks co-op offered delicious food, interesting conversation and valuable friendship, which were essential for my sanity in graduate school. Last but not least, I would like to thank my mother and grandparents for their loving and enthusiastic support through all these years.

# Table of Contents

|                                                                                                                                                          |            |
|----------------------------------------------------------------------------------------------------------------------------------------------------------|------------|
| <b>Chapter 1: General Introduction.....</b>                                                                                                              | <b>1</b>   |
| <b>Chapter 2: Functional CpG Methylation System in a Social Insect.....</b>                                                                              | <b>10</b>  |
| <b>Chapter 3: Nutritional Regulation of Division of Labor in Honey Bees:<br/>Toward a Systems Biology Perspective.....</b>                               | <b>32</b>  |
| <b>Chapter 4: The Genomic Distribution of USP Suggests Possible Molecular<br/>Mechanisms for its Involvement in Honey Bee Behavioral Maturation.....</b> | <b>55</b>  |
| <b>Chapter 5: Conclusions.....</b>                                                                                                                       | <b>109</b> |
| <b>References.....</b>                                                                                                                                   | <b>111</b> |
| <b>Appendix A: ChIP-PCR Confirmation and Additional Antibody Testing<br/>Using ChIP-PCR.....</b>                                                         | <b>123</b> |
| <b>Appendix B: ChIP (ChIP-QPCR) Protocol.....</b>                                                                                                        | <b>129</b> |

# **Chapter1**

## **General Introduction**

Phenotypic plasticity is the ability of a single genotype to produce multiple phenotypes in response to environmental conditions (Scheiner, 1993; West-Eberhard, 1989). Plastic responses range from alternative morphologies to different physiological and behavioral states, and often involve changes in a combination of related traits (West-Eberhard, 1989). These responses can be stable for extended periods. At times, the changes may persist even after the environmental stimuli are removed. For instance, rats that received good maternal care (frequently licked and groomed by the mother) as pups are less responsive to stress and perform better in nursing their own offspring as adults (Weaver et al., 2004). These long-lasting physiological and behavioral changes suggest persistent alterations on the molecular level.

The molecular mechanisms of phenotypic plasticity are under active investigation using new tools in genomics (Aubin-Horth & Renn, 2009). This has revealed that complex changes in gene expression are associated with phenotypic plasticity. However, little is known about the regulatory mechanisms by which changes in gene expression are orchestrated by the external and internal factors. In particular, how does the coordinated gene expression pattern shift between different physiological and behavioral states? How is this transition initiated by external stimuli and then mediated by internal factors? And after the transition, how are the gene expression patterns and their associated phenotypes maintained? Better understanding of the roles of two interconnected regulatory systems

discussed in this thesis: DNA methylation (Chapter 2), and transcription factors (Chapter 4) are likely to shed light on the above questions.

The honey bee (*Apis mellifera*) provides a system to study phenotypic plasticity at the molecular level. Adult honey bee workers change their behavior dramatically with age, as part of the division of labor in the colony. For the first 2-3 weeks of their adult life, worker bees perform tasks inside the hive, as exemplified by brood and queen care (“nursing”). At around 3 weeks of age, worker bees transition to foraging outside the hive for the remainder of their life (Winston, 1991). This process of worker age-related division of labor is also known as behavior maturation. Besides the shift of behavioral repertoire, behavioral maturation is also characterized by dramatic changes in physiological states. For example, nurses and foragers differ significantly in their stores of abdominal lipids (Toth & Robinson, 2005) and proteins (Amdam, Norberg, Fondrk, & Page, 2004; Rutz & Lüscher, 1974), basal metabolic rates (Harrison, 1986), and brain metabolic capacity (Alaux et al., 2009; Ament, Corona, Pollock, & Robinson, 2008). So like other cases of phenotypic plasticity, behavioral maturation is a suite of coordinated changes in behavioral and physiological states.

Three features of honey bee behavioral maturation are of particular interest in relation to understanding phenotypic plasticity.

First, individual worker bees are able to respond to both internal factors and environmental cues by adjusting their timing of the onset of foraging. Among the many internal factors, Juvenile Hormone (JH) and nutritional status are best understood so far. JH titers increase dramatically before the onset of the foraging, and stay at a high level in foragers (Robinson, Page, Strambi, & Strambi, 1989; Rutz, Gerig, Wille, & Lüscher,

1976). Treating young bees with JH or methoprene a JH analog, speeds up the process (Jaycox, 1976; Robinson, 1985, 1987). Similarly, a dramatic change in nutritional status, a steep loss in abdominal lipid and protein stores, precedes the onset of foraging (Amdam et al., 2004; Rutz & Lüscher, 1974; Toth & Robinson, 2005). In addition, reducing the amount of abdominal lipid or a storage protein Vitellogenin (Vg) causes precocious foraging (Marco, Guidugli-Lazzarini, do Nascimento, Simões, & Hartfelder, 2008; Nelson, Ihle, Fondrk, Page, & Amdam, 2007; Toth, Kantarovich, Meisel, & Robinson, 2005).

Environmental cues also pace the behavioral maturation of individual bees, in part mediated by these internal factors. For instance, food deprivation or the loss of a colony's foraging force causes young workers to forage precociously (Schulz, Huang, & Robinson, 1998). A lack of foragers causes the JH titers of young bees to rise prematurely, and this results in precocious foraging (Robinson, Page, C. Strambi, & A. Strambi, 1989). Similarly, it is the decrease of individual food intake rather than the direct assessment of decreased food stores that causes precocious foraging in food-deprived colonies (Schulz et al., 1998).

Second, the transition from nurses to foragers involves massive changes in gene expression. For instance, up to 40% of honey bee genes are expressed differentially in the brains of nurses compared to foragers (Alaux et al., 2009; Whitfield, Cziko, & Robinson, 2003). These distinct gene expression patterns suggest that transcriptional control is important for regulating behavioral maturation.

Transcription factors are essential in transcriptional regulation. Transcription factors bind to regulatory DNA sequences and participate in interactions that ultimately

activate or repress the expression of downstream genes (Lee & Young, 2000), with the expressions of individual transcription factors being subject to similar regulation as well. Together, these hierarchical interactions comprise gene regulatory networks (GRNs) (Davidson, 2006). Internal factors such as hormones and nutritional status regulate gene expressions through GRNs. A first step in understanding the role of transcriptional regulation in behavioral maturation is to describe the GRNs in nurses and foragers. Towards this goal, the distribution of binding sites of a single transcription factor were identified across the genome in honey bee fat cells (Chapter 4).

Third, the behavioral and physiological states of nursing and foraging are highly stable. It is possible to reverse foragers to nurses but the process is difficult and the reversed foragers are never as efficient in nursing (Robinson et al., 1989). The stability may be maintained across several different levels, from gene expression to hormone titers. Little is known about how the distinct gene expression patterns of nurses and foragers are maintained, particularly on the molecular level, which led to our effort to characterize the honey bee DNA methylation system (Chapter 2). DNA methylation is implicated in maintaining stable gene expression patterns in other organisms (Bird, 2002), and thus represents a promising candidate. Indeed, in the aforementioned example of the adult rats affected by maternal care, DNA methylation is shown to be involved in maintaining the effect (Weaver et al., 2004). It is also possible that the intrinsic properties of the GRNs, such as hierarchical control and feedback loops stabilize the expression patterns (Davidson, 2006). The answer to this question lies in a deeper understanding of the topological properties of GRNs, which is a second impetus for the studies described in Chapter 4.

This thesis is comprised of 3 manuscripts. Chapter 2 was published in 2006 in *Science* 314: 645-647, Chapter 3 was published in 2010 in *Wiley Interdisciplinary Reviews: Systems Biology and Medicine* (published online in advance of print), and Chapter 4 will be submitted for publication. I would like to acknowledge that the work reported in Chapter 2 is the result of a collaboration involving 5 labs; Chapter 3 and behavioral experiments in Chapter 4 are the products of a collaboration between fellow graduate student Seth Ament and myself; and that the bioinformatic analyses of Chapter 4 have been greatly facilitated by collaboration with the labs of Sheng Zhong and Saurabh Sinha. Below I provide brief introductions for the remaining chapters of this thesis.

Chapter 2 describes the functional characterization of the honey bee DNA methylation system. DNA methylation is widespread among eukaryotic species, and its function has been studied intensively in mammals. In mammalian cells, it is implicated in many important cellular functions, including repression of the inactivated X chromosome in females, silencing transposons, and maintaining the allele-specific expression of imprinted genes (Chen & Li, 2004; Dodge et al., 2005; Karpf & Matsui, 2005). DNA methylation is considered to be one of the mechanisms to establish and maintain gene expression patterns (Bird, 2002), which makes it highly attractive in understanding phenotypic plasticity on the molecular level. However, DNA methylation in insects is not well-understood, partially due to the limited extent of DNA methylation in the popular insect model organism, *Drosophila melanogaster* (Marhold, Kramer, Kremmer, & Lyko, 2004).

The sequencing of the honey bee genome allowed us to use a bioinformatic approach together with molecular and biochemical analyses to characterize the honey bee

DNA methylation system. The system is truly functional with catalytically active orthologs of vertebrate DNA methyltransferase (Dnmts), and methylated genes. However, compared to vertebrates the global levels of DNA methylation are much lower in the honey bee. Moreover, all of the methylation we detected was intragenic, in contrast to the classical model of methylated promoter regions observed in vertebrates. It was confirmed that lower global levels of DNA methylation is a feature shared by some other invertebrate genomes (Bird, Taggart, & Smith, 1979; Suzuki, Kerr, De Sousa, & Bird, 2007; Tweedie, Charlton, Clark, & Bird, 1997). It has also been shown that DNA methylation in the intragenic regions is pervasive in many species, including humans (Eckhardt et al., 2006; Rakyan et al., 2004; Suzuki et al., 2007). Nonetheless, the function of intragenic DNA methylation is still not well-understood (Suzuki & Bird, 2008).

The characterization of the DNA methylation system paves the way to study its role in phenotypic plasticity in the honey bee. It has been demonstrated that a DNA methyltransferase (Dnmt3) is important in queen vs. worker determination (Kucharski, Maleszka, Foret, & Maleszka, 2008), another major form of phenotypic plasticity exhibited by the honey bee. It will be highly interesting in the future to determine if DNA methylation is also involved in the age-related division of labor. Previously, we faced the difficulty of efficiently locating methylated genes for functional analysis, which was the main hurdle preventing further investigations. The development of genome-wide “bisulfite sequencing” (BS-Seq) (Cokus et al., 2008) and direct sequencing techniques effectively solved this problem (Flusberg et al., 2010). It is an exciting time for others to continue the study of DNA methylation in the context of age-related division of labor.

Chapter 3 reviews current knowledge of the mechanisms by which the one of the internal factors, nutritional status, influences honey bee behavior maturation. It sets the stage for the studies described in Chapter 4. Chapter 3t also provides a framework for using systems biology approaches to gain a deeper insight into nutritional regulation of honey bee behavior, and phenotypic plasticity in general.

Chapter 4 describes experiments providing information crucial to constructing a predictive GRN model for the fat bodies of nurse and foragers. Fat bodies are the insect counterpart of vertebrate liver and adipose tissues combined (Arrese & Soulages, 2010; Leopold & Perrimon, 2007). Fat bodies undergo dramatic changes during behavioral maturation. Lipid stores drop steeply (Toth & Robinson, 2005), and about 20% of the genes are differentially expressed (Ament, Robinson et al., unpublished). In GRN models, gene expression is determined by sets of functionally active transcription factors, denoted as regulatory states (Davidson, 2006). We hypothesize that different regulatory states underlie the two behavioral and physiological stages.

To test this hypothesis, we focused on one transcription factor, Ultraspiracle (USP). USP was chosen for the following reasons. First, USP is implicated in JH signaling (Barchuk, Maleszka, & Simoes, 2004; Jones & Sharp, 1997; Velarde, Robinson, & Fahrbach, 2009). Second, the vertebrate ortholog of USP is Retinoid X Receptor (RXR), a master regulator of lipid metabolism (Chawla, Repa, Evans, & Mangelsdorf, 2001; Mangelsdorf, Evans et al., 1995; Sonoda, Pei, & Evans, 2008). Finally, since USP associates with both of the two internal factors (JH and nutrition), we hypothesize that USP also plays a regulatory role in behavioral maturation. Indeed, *usp* knockdown in the fat bodies of young bees delays the onset of foraging.

A first step to construct a USP-centered GRN is to find its downstream genes. We hypothesized that USP regulates different genes in nurses and foragers by binding to different genomic regions. Alternatively, differential regulation could be mediated in the absence of differential USP binding by the presence or absence of other ligands or cofactors involved in USP function. We performed chromatin immunoprecipitation coupled to genomic tiling arrays (ChIP-chip) using fat bodies of nurses and foragers. The ChIP-chip results were highly reproducible, and revealed around 1300 USP binding regions in the genome. No binding difference was detected between nurses and foragers, suggesting USP regulates gene expression depending on the availability of additional cofactors or ligands, presumably in different chromatin contexts.

Motif analysis of USP binding regions showed that the USP-binding motif in *Drosophila*, CF1, is significantly enriched in USP-binding regions, indicating that the basic regulatory mechanisms of honey bee USP are conserved to its orthologs in *Drosophila* and vertebrates: it binds to DNA directly through a highly conserved DNA motif. Putative USP target genes were identified within 10 kb of USP binding regions. Among these putative USP target genes, transcription factors are highly enriched. This suggests that USP is likely to occupy a high place in the fat body behavioral maturation GRN hierarchy, i.e. that it acts to regulate other regulators, instead of directly regulating effectors. Functional analysis of putative USP target genes also identified several genes involved in other signaling pathways that are known to be involved in honey bee behavioral maturation: *inR1* in insulin signaling (Ament et al., 2008; Wu & Brown, 2005), *jhamt* in JH synthesis (Bellés, Martín, & Piulachs, 2004), and *forage (for)*, a gene encoding a cGMP-dependent protein kinase (Ben-Shahar, Robichon, Sokolowski, &

Robinson, 2002; Osborne et al., 1997). Taken together, our observations raise intriguing possibility that USP functions as an integrator of different signaling pathways.

## Chapter 2

# Functional CpG Methylation System in a Social Insect

Previously published work<sup>1</sup>

### Abstract

DNA methylation systems are well characterized in vertebrates, but methylation in *Drosophila melanogaster* and other invertebrates remains controversial. Using the recently sequenced honey bee genome, we present a bioinformatic, molecular, and biochemical characterization of a functional DNA methylation system in an insect. We report on catalytically active orthologs of the vertebrate DNA methyltransferases Dnmt1 and Dnmt3a and b, two isoforms that contain a methyl-DNA binding domain, genomic 5-methyl-deoxycytosine, and CpG-methylated genes. The honey bee provides an opportunity to study the roles of methylation in social contexts.

### Introduction

Among the many important functions of CpG DNA methylation, sex-specific regulation of gene expression (imprinting) in vertebrates stands out because it provides insight into intragenomic conflict (Goll & Bestor, 2005; Queller, 2003). Provided social insects have CpG methylation, they would be ideal models to further explore the kin-

---

<sup>1</sup> Y. Wang et al., "Functional CpG methylation system in a social insect," *Science* 314, no. 5799 (2006): 645.

conflict theory of imprinting, because insect societies are composed of many different types of relatives and they interact with each other in many evolutionarily important contexts (Haig, 2000; Queller, 2003).

However, although widely conserved from yeast and fungi to plants to vertebrates, DNA methylation in insects is enigmatic. Evidence of CpG-methylated sequences exists for several insect species (Field, 2000; Field, Lyko, Mandrioli, & Pranter, 2004; Tweedie, Charlton, Clark, & Bird, 1997), but no bona fide invertebrate deoxycytosine methyltransferases (DNMTs) have been described. Conversely, the model insect *Drosophila melanogaster* shows limited DNA methylation, predominantly in asymmetric CpT and CpA dinucleotides (Marhold, Kramer, Kremmer, & Lyko, 2004), and this is attributed to the only DNMT family member encoded in its genome, dDNMT2, a tRNA<sup>Asp</sup> methyltransferase (Goll et al., 2006; Kunert, Marhold, Stanke, Stach, & Lyko, 2003). The *Drosophila* situation renders the fragmentary findings from the other insect species uninterpretable.

Here, we report that a social insect, the honey bee, *Apis mellifera*, has a fully functional CpG methylation system. We present biochemical, molecular, and genomic analyses, made possible by the recently sequenced honey bee genome (Honey Bee Genome Sequencing Consortium, 2006).

## Results

In addition to *Dnmt2*, we identified one ortholog for de novo methylation (*AmDnmt3*) and two orthologs for maintenance methylation (*AmDnmt1a* and *AmDnmt1b*) (Figure 2.1A and Figure 2.2A). *AmDnmt1a* and *AmDnmt1b* encode 70% identical ~1400–

amino acid proteins, which have 55% identity and 70% similarity to human DNMT1 over almost their entire length. Likewise, *AmDNMT3* shows strong sequence similarity to *hDNMT3A* and *hDNMT3B*, with 33 and 32% similarity over the whole gene and 61 and 66% identity in the catalytic domains, respectively.

The predicted AmDNMT-encoding genes are expressed in tissue-specific and developmentally regulated patterns (Figure 2.1A). Protein extracts from bees (Figure 2.1B) demonstrate the presence of catalytically active DNMT enzymes. In vitro biochemical analyses of the recombinant AmDNMT proteins confirmed that AmDNMT1A and AmDNMT3 are catalytically active DNMTs (Figure 2.1C) with similar characteristics to their vertebrate orthologs (Goll et al., 2006).

The honey bee CpG methylation system is functional in vivo. We isolated 5-methyl deoxycytosine (dCM) (Ramsahoye, 2002) from bee genomic DNA (Figure 2.1D) and using the methods of (Frigola, Ribas, Risques, & Peinado, 2002) have so far identified six genes methylated in vivo (Figures. 2.3 to 2.8). These analyses also revealed that non-CpG methylation is either extremely rare or nonexistent in honey bee. In each instance, the methylation was found exclusively in transcribed regions and predominantly in predicted exons with low G+C content and few CpGs overall. One methylated gene (*GB16767*) encodes an ortholog of mSin3A-associated protein 130 kD (SAP-130); Sin3 complexes regulate gene expression from yeast through humans (Silverstein & Ekwall, 2005), providing the potential for potent downstream genome-wide effects in honey bees.

Further analysis of the honey bee genome identified a gene homologous to the family of methyl-DNA binding domain (MBD) proteins (Figure 2.2B), some of which are effectors of DNA methylation (Klose & Bird, 2006). At least two expressed splicing

variants exist, both of which contain the most highly conserved amino acids in the MBD (Figure 2.2B). AmMBD-1 preferentially, but not exclusively, interacted with a methylated DNA probe in vitro (Figure 2.9, B and C), similar to MBDs in other species (Klose & Bird, 2006; Marhold, Brehm, & Kramer, 2004). The protein sequence, size, and in vitro binding characteristics indicate that AmMBD-1 is most similar to the vertebrate MBD3 subfamily. MBD3 proteins across species vary in their methyl-DNA binding specificities in vitro (Klose & Bird, 2006) and function in vivo as integral components of the Mi-2 chromatin remodeling complexes in both vertebrates and *Drosophila* (Marhold et al., 2004; Wade et al., 1999). Alternatively, the translation of DNA methylation marks in the honey bee may be achieved through other mechanisms (Hark et al., 2000).

## Discussion

On the basis of our findings, it is now possible to reflect on the earlier insect work (Field, 2000; Field et al., 2004; Tweedie et al., 1997) and predict that vertebrate-like systems of methylation are widespread in insects. If so, then *Drosophila* is of interest, not as a general model of insect methylation, but for unexplored evolutionary aspects of genome regulation, including the lack of a canonical telomeric (TTAGG) repeat (Duncan, 2002), the pairing of homologous chromosomes throughout interphase (Sahara, Marec, & Traut, 1999), and the lack of symmetrical or CpG methylation (Marhold et al., 2004).

Our findings also mean that the honey bee has the mechanisms that underlie imprinting, so tests of the kin-conflict theory in social insects are likely to be fruitful. Toward this end, we note several differences in honey bee methylation relative to vertebrate systems. First, methylation in vertebrates represses expression of repetitive

DNAs and retrotransposons to maintain genome integrity, but in honey bees (and other insects) intermediate and high repetitive DNA elements and transposons are not methylated (Field et al., 2004). This suggests that the role in mobile element repression has been lost or evolved after insects and vertebrates diverged (Field et al., 2004; Goll et al., 2006). Second, overall levels of methylation appear to be lower in the honey bee than in vertebrates, arguing against DNA methylation as a global mediator of bee heterochromatin. Third, honey bees possess two paralogs for methylation maintenance, making them the first animal discovered to express multiple somatic DNMT1 proteins. Fourth, all detected methylation was limited predominantly to the coding regions of genes. It remains to be seen how any of these differences relate to the functions of methylation in social contexts.

## **Materials and Methods**

### *AIMS and Bisulfite Genomic Sequencing*

AIMS was carried out essentially as described using different combinations of sensitive/insensitive isoschizomers to digest genomic DNA (Frigola et al., 2002). The isoschizomers used were EcoRII/MvaI that recognize CCWGG; SmaI/XmaI which recognize CCCGGG, and HpaII/MspI which recognize CCGG. Bisulfite sequencing was performed as described (Stirzaker, Song, Davidson, & Clark, 2004). Quantification was validated by direct bisulfite sequencing of 10 independent clones from adult drone testis, adult queen ovary, and five clones from adult worker brain.

Samples for AIMS were obtained as follows. Bee eggs (embryos) were collected by allowing a queen to lay in a small comb located in a hive designed for easy egg

retrieval (Karl Jenter, Nürtingen, Germany); following queen laying the egg-containing comb was transferred to an incubator (35°C, 80% humidity) for later collection. Larvae were collected from brood frames taken from the hive, incubated and snap frozen. Adult drones were collected from typical field colonies and queens were purchased from local beekeepers in Australia. Dissections of queen ovaries and drone genitals were carried out in bee Ringer solution. Samples were as follows. Eggs, 0-22h old (N = 2113 eggs, yielding 50 ug DNA); Eggs, 48-72 h old (2860, 150 ug DNA); Worker larvae, 2-5 days old (~25, ~ 10 ug DNA); Drone larvae, 2-5 days old (~15 mixed; ~ 12 ug); Adult workers, 1-day-old, thorax muscle (50, ~75 ug); Adult workers, brain, foragers (50, ~50 ug); Adult workers, brain, 1-day-old (50, ~75 ug), Adult workers, hypopharyngeal glands, 1-day-old (25, ~100 ug); Mated queens, ovary (3, ~ 250 ug of DNA); Young adult drones, testes (4, ~ 40 ug).

Preparation of high molecular DNA for AIMS was as follows. Frozen tissues were crushed in liquid nitrogen (either in a 1.5 ml Eppendorf tube or in a small mortar) and transferred to a tube containing NTE (100 mM NaCl, 50 mM Tris pH 8.0, 10 mM EDTA, 1% SDS and 0.01% Triton X-100) plus proteinase K (500 ug per ml, freshly dissolved in NTE). For tissues such as embryos and brains the homogenization step was not necessary as these tissues dissolve well in NTE with proteinase K. The volume of NTE was adjusted to ensure that the suspended homogenate was not too thick. Samples were incubated at 50°C for 2-3 h, extracted with 1 volume of phenol:chloroform and spun for 10 min at 10,000g. The upper phase was collected and treated with 1 ul of RNase A (10 mg/ml) for 5-10 min. After another phenol extraction DNA was precipitated with 2 vol of absolute ethanol (or 1 vol of isopropanol). For tissues with high content of DNA such as

ovaries and testes DNA was spooled using a small glass rod. For other tissues the precipitate was gently spun for 2 min at 2,000g and immediately dissolved in TE buffer by heating to 65 C. Final preps were stored at 4 C.

#### *Expression RT-PCR*

Total RNA was extracted from indicated source by a combined Trizol/Qiagen method (Stirzaker et al., 2004). RNA expression was evaluated by RT-PCR of ~0.1ug total RNA and 35 cycles (15'' at 94°C, 25'' at 56°C, and 30'' at 72°C) using primers listed in Table 2.1. Expected products for *AmDnmt1a* were 179 bp from mRNA compared to a 441 bp band from genomic DNA. Expected products for *AmDnmt1b* were 295 bp from mRNA compared to 387 bp from genomic DNA. Expected products for *AmDnmt3* were 338 bp from mRNA compared to 540 bp from genomic DNA.

#### *HPLC Analysis*

Genomic DNA was prepared from adult honey bees, 3-4 day larvae, or rat liver, digested and analyzed as described (Ramsahoye, 2002). HPLC was carried out at room temperature essentially as described using a Beckman HPLC with an absorbance detector. Deoxyribonucleotide 5'-monophosphates in DNA digests were resolved using an Alltech Adsorbosphere octadecylsilane-bonded silica column (250mmx4.6mm) eluted at room temperature with 50mM ammonium orthophosphate (pH4.1) at 0.8 ml per minute. The system was calibrated using samples of the common deoxyribonucleotide 5'-monophosphates (Sigma) and 5mC monophosphate (Reliable Biopharmaceutical Corp).

### *Cloning cDNAs and Expression of Recombinant AmDnmt and AmMBD-L Proteins*

Total RNA was extracted from brains of adult honey bees using Trizol (Invitrogen Inc). Putative full-length *AmMBD-l*, *AmMBD-s*, *AmDnmt1a*, *AmDnmt2*, and *AmDnmt3* cDNAs were generated by RT-PCR (SuperScript™ III One-step RT-PCT, Invitrogen Inc) using oligonucleotide primers designed according to sequences predicted bioinformatically to be the respective gene sequences (Table 2.1). PCR products were cloned into the pGEM-T vector System I (Promega, Inc) and sequenced. The catalytic domains of *AmDnmt1a* and *AmDnmt3* as well as the full-length *AmMBD-l*, *AmDnmt2* and *AmDnmt3* were further subcloned into pET23b (NdeI and XhoI) (Novagen). pET23b-*AmMBD-l* was transformed into BL21-AI (Invitrogen, Inc) for recombinant protein expression. Overnight cultures were diluted 1:100 in 500 ml LB, grown for 3 hr at 37° C, and induced with 1mM IPTG and 0.2% L-arabinose for 5 hr at 30° C. The 6X-His *AmMBD-l* protein was purified using 2ml Talon resin as per instructions (Novagen). The recombinant protein was diluted to 100mM NaCl, incubated with 2mM ATP at room temperature for 10 min to dissociate the DnaK chaperone, applied to a 1ml Mono-S column (GE Healthcare) and eluted over a 10ml linear gradient to 1M NaCl.

### *Protein Extracts*

Honey bee eggs (40 per preparation) or larvae (20 mg per preparation) were homogenized in Buffer A (350mM NaCl, 20mM HEPES pH 7.5, 10mM  $\beta$ -glycerophosphate, 1.5mM MgCl<sub>2</sub>, 0.5mM DTT, 10% glycerol, 1mM PMSF, 1 $\mu$ g/ml leupeptin, 1 $\mu$ g/ml aprotinin and 1 $\mu$ g/ml pepstatin) and centrifuged at 55,000 x g for 45 minutes at 4°C to obtain soluble protein.

### *DNA Methyltransferase Assay*

DNMT activity was assayed as described using soluble protein derived from either 40 eggs or 20 mg of larvae per reaction and 1  $\mu$ g poly(dI-dC) for a substrate (Issa et al., 1993). Recombinant DNMT activity was assayed as described using 17  $\mu$ l of a 50  $\mu$ l Quick-coupled TnT (Promega, Inc) reaction, from either AmDnmt1a full-length, AmDnmt3 catalytic domain (amino acid 401 – COOH), or pET23b, incubated with 3 $\mu$ g double-stranded linear plasmid DNA substrates and 3 $\mu$ Ci S-adenosyl-l-[methyl-3H] methionine in a total volume of 50  $\mu$ l for 2h at 37°C. Rnase A and T1 were added for 10 min and reactions were terminated in stop solution (1.0% SDS, 2.0mM EDTA, 3% 4-aminosalicylate, 5% butanol, 0.25 mg/ml calf thymus DNA, 1.0 mg/ml proteinase K) phenol extracted, ethanol precipitated and the DNA resuspended in 0.3 N NaOH and incubated at 37°C for 45 min. This solution was then spotted onto a Whatman GF/C filter, dried, washed with 10% trichloroacetic acid and 70% ethanol Incorporated [3H]-CH<sub>3</sub> was determined by liquid scintillation counting.

### *EMSA and Filter Binding Assay*

Binding reactions were performed as described with minor modifications (Wade et al., 1999). Briefly, 150ng of double-stranded oligonucleotide probes (Table 2.1) were radiolabeled with  $\gamma$ -32P-ATP, and gel purified. FPLC purified recombinant AmMBD-1 was incubated with 50,000 cpm of dsDNA probe alone or with unmethylated specific competitor DNAs (0.01 ng, 1ng, or 5ng) in binding buffer (12mM HEPES pH 7.9, 5mM MgCl<sub>2</sub>, 50mM NaCl, 0.12mM EDTA, 0.1% NP-40, 2mM DTT, 5% glycerol and 0.4 mg/ml BSA) for 45 min at 4°C. Gel mobility shifts were performed in 6%

polyacrylamide gels run in 0.5X TBE buffer. Filter binding reactions were as for EMSA with the exception that reactions were spotted onto P81 cellulose filters (Whatman), and washed 4X with wash buffer (20mM Tris pH 8.0, 50mM NaCl). Bound [<sup>32</sup>P] labeled DNA probe was determined by scintillation counting. Control reactions did not contain AmMBD-l protein and competitions were with unmethylated competitor (0.1ng, 1ng, or 5ng). All reactions were carried out in triplicate.

## **Acknowledgments**

We thank J. Maleszka, T. Newman, K. Pruiett, and C. J. Schoenherr for technical assistance and D. C. Queller, C. J. Schoenherr, and members of the Jones and Robinson labs for reviewing the manuscript. Supported by the Illinois Sociogenomics Initiative (G.E.R.) and the Roy J. Carver Charitable Trust (C.A.M.), and a Consolider Ingenio 2010 program (CSD2006-49 from the Spanish Ministry of Education and Science (M.J. and M.A.P.). Sequences have been deposited into GenBank: AmDnmt1a (GB15130), AmDnmt1b (GB19865), AmDnmt2 (GB10767), AmDnmt3 (HM067836), and AmMBD (REFSEQ XP\_392422).

## Tables

**Table 2.1:** Sequences of oligonucleotide primers used in chapter 2.

### Cloning cDNAs

#### **AmDNMT1A**

|               |                                                          |
|---------------|----------------------------------------------------------|
| AmDnmt1a (F1) | 5'-CACACACATATG AAG TCT AAT ACA TCA AAA TTAA CAA ACA TCT |
| AmDnmt1a (R1) | 5'-CACACAGTCGAC ATC ATT TTG TAT TTT TAT ATCAAT TTC AGT   |

#### **AmDNMT2**

|              |                                                                           |
|--------------|---------------------------------------------------------------------------|
| AmDnmt2 (F1) | 5'-CACACACATATG AGA GTG TTG GAA TTA TAC AGT GGTATT GGT                    |
| AmDnmt2 (R1) | 5'-CACACAGTCGAC TGT AAT CTT TCT TTC AGT ATA TAATAA AAA TAT<br>TAA TCT GCT |

#### **AmDNMT3**

|              |                                                                  |
|--------------|------------------------------------------------------------------|
| AmDnmt3 (F1) | 5'-CACACACATATG TTG TCC GAG GAA GGG AAA TTA TGGTGT               |
| AmDnmt3 (R1) | 5'-CACACAGTCGAC GAA GAT TGA GAA ACA GTT TGT AAATGT AAA TGA<br>AG |

#### **AmMBD**

|            |                                                             |
|------------|-------------------------------------------------------------|
| AmMBD (F1) | 5'-CACACACATATG, AAT ATG TCA GTG GAA AAA AAG AAATAT CCA TCG |
| AmMBD (R1) | 5'-CACACACTCGAG, TAT AGG AAT AGA TCG TAA AGC GTCTTG         |

### RT-PCR analysis

|               |                             |
|---------------|-----------------------------|
| AmDnmt1a (F1) | 5'-CAATGGGACCTAAGAGACAA     |
| AmDnmt1a (R1) | 5'-GGATCAGAATCATCTGCTTCTTGA |

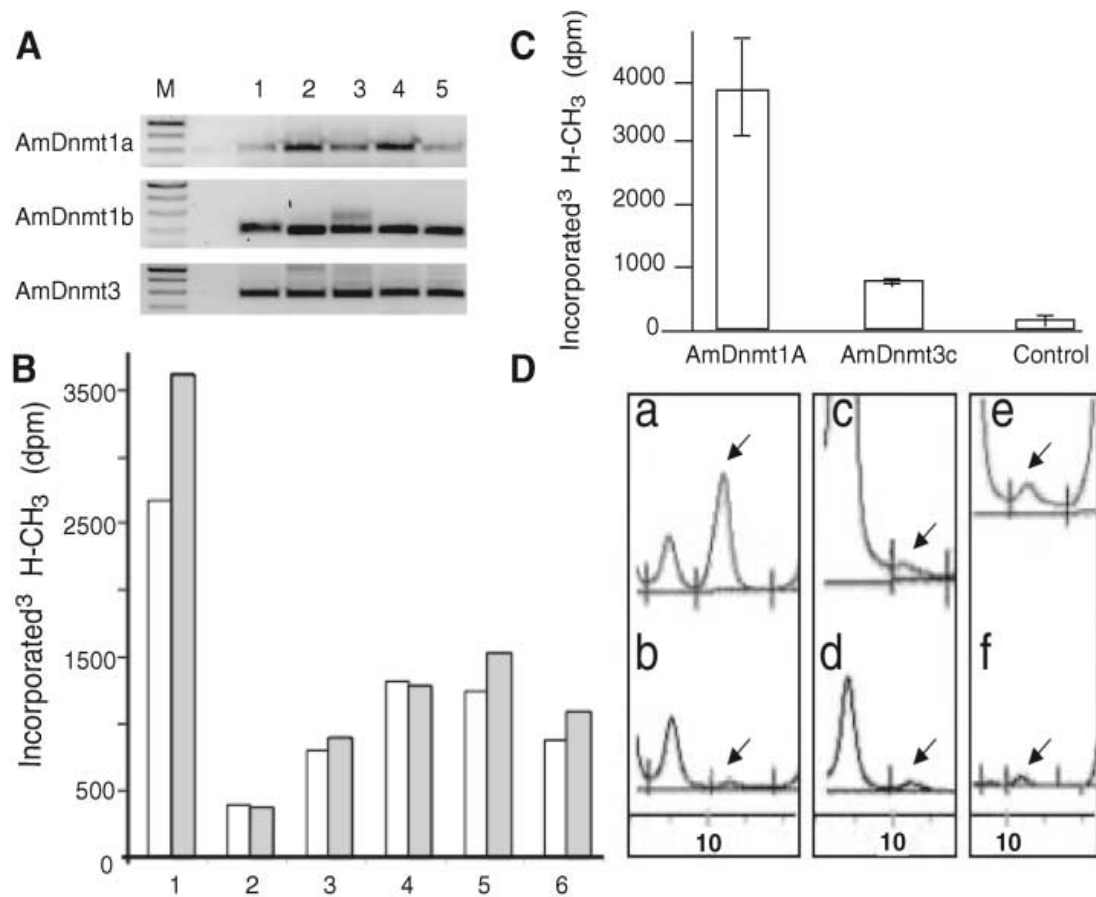
**Table 2.1 (cont.):**

|               |                              |
|---------------|------------------------------|
| AmDnmt1b (F1) | 5'-GCATCAGCAGGACCAGAAGATCC   |
| AmDnmt1b (R1) | 5'-CTCCACAATCTGGCTGTTGACAAG  |
| AmDnmt3 (F1)  | 5'-CAGCGATGACCTGCGATCGGCGATA |
| AmDnmt3 (R1)  | 5'-CATGCAGGCGGTGCAGTA        |

### EMSA and Filter binding

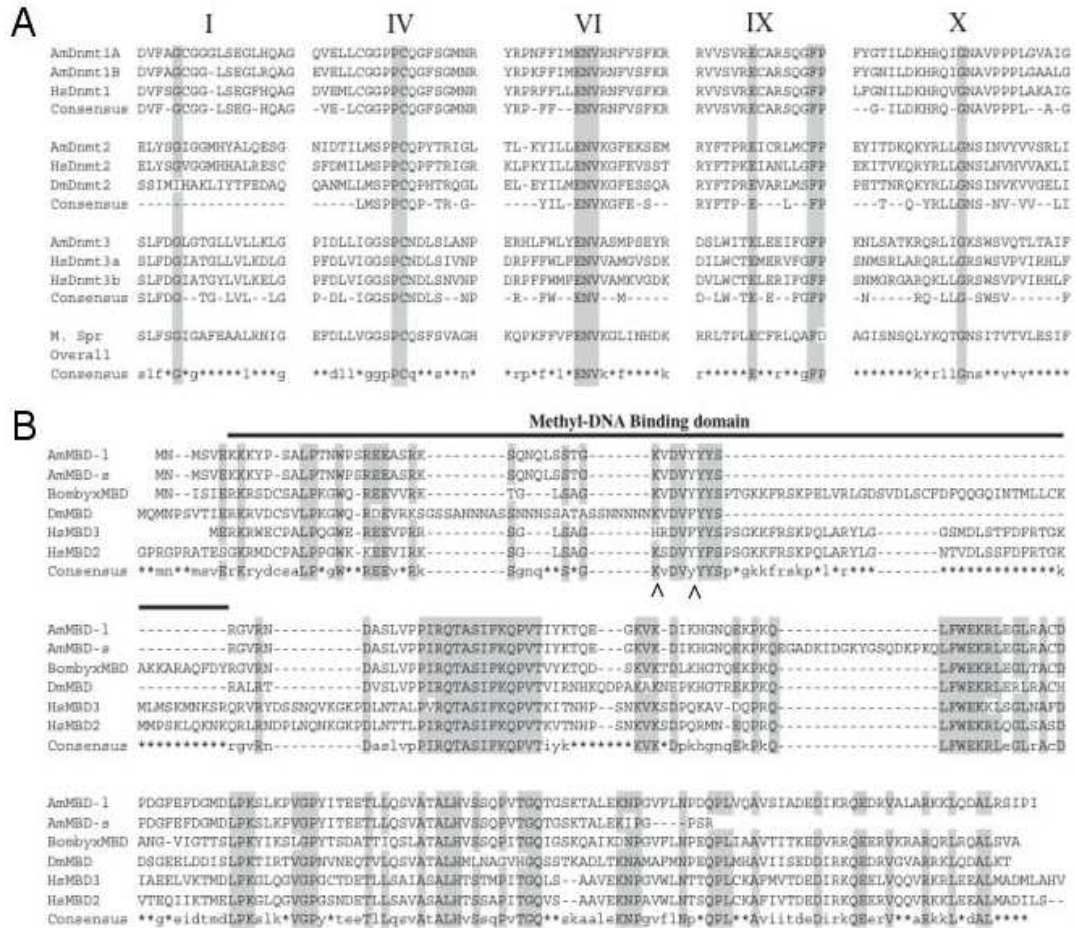
|       |                                               |
|-------|-----------------------------------------------|
| MS30F | 5'-GGTTTATACGCGGGAGTTGCCGCGTGGTCC             |
| MS30R | 5'-GGACCACGCGGCAACTCCCGCGTATAAACC             |
| MSCGF | 5'-GGTTTATAMGMGGGAGTTGCMGMGTGGTCC             |
| MSCGR | 5'-GGACCAMGMGGCAACTCCMGMGTATAAACC             |
| GACF  | 5'-GATCCGACGACGACGACGACGACGACGACGACGACGACGATC |
| GACR  | 5'-GATCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGGATC |
| GAMF  | 5'-GATCMGAMGAMGAMGAMGAMGAMGAMGAMGAMGAMGATC    |
| GAMR  | 5'-GATMGTMGTMGTMGTMGTMGTMGTMGTMGTMGTMGTMGGATC |

## Figures

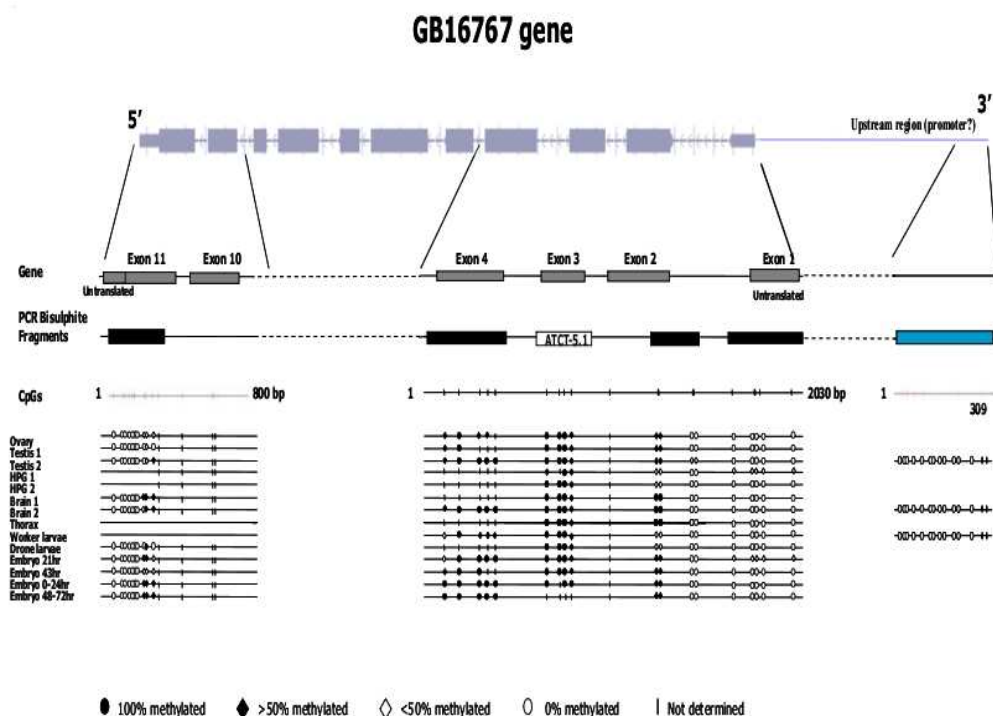


**Figure 2.1.** DNA methylation in the honey bee. (A) Expression of *AmDNMT1* and *AmDNMT3* genes. Reverse transcription polymerase chain reaction (RT-PCR) of total RNA from multiple tissues and stages of development (24-hour-old embryo, adult worker ovary, queen larva, drone genitals and sperm, and adult worker brain; lanes 1 to 5, respectively) indicates expression of AmDnmt1a (top), AmDnmt1b (middle), and AmDnmt3 (bottom). Data are from a representative experiment repeated twice with similar results. For each experiment, total RNA was extracted from 50 eggs (24 hours old), one adult queen ovary, one queen larva, four adult drone genitals, and 10 brains from 8-day-old worker bees. M = 100–base pair ladder. (B) AmDNMT1A and AmDNMT3 proteins are catalytically active DNA methyltransferases in vivo. DNMT assays were performed on honey bee protein extracts prepared from either 40 embryos (lane 3), or 20 mg of tissue from 2-day-old, 3-day-old, or 5-day-old larvae (lanes 4 to 6, respectively). *Xenopus* egg extract (lane 1) was used as a positive control for DNMT activity, and bovine serum albumin (BSA) (lane 2) was used as a negative control. Experiments were carried out four times, each in duplicate, resulting in a highly reproducible pattern of activity. Results from a representative DNMT activity assay are shown. The two bars represent duplicates from one experiment. dpm, disintegrations per minute. (C) AmDNMT1 and AmDNMT3 proteins are catalytically active DNA methyltransferases in vitro. In vitro transcribed and translated full-length AmDNMT1A, AmDNMT3 catalytic domain (amino acid 401 to the carboxyl terminus) (AmDnmt3c), and empty vector (control) were analyzed for DNA methyltransferase activity using an unmethylated doublestranded DNA template. The average of three independent experiments

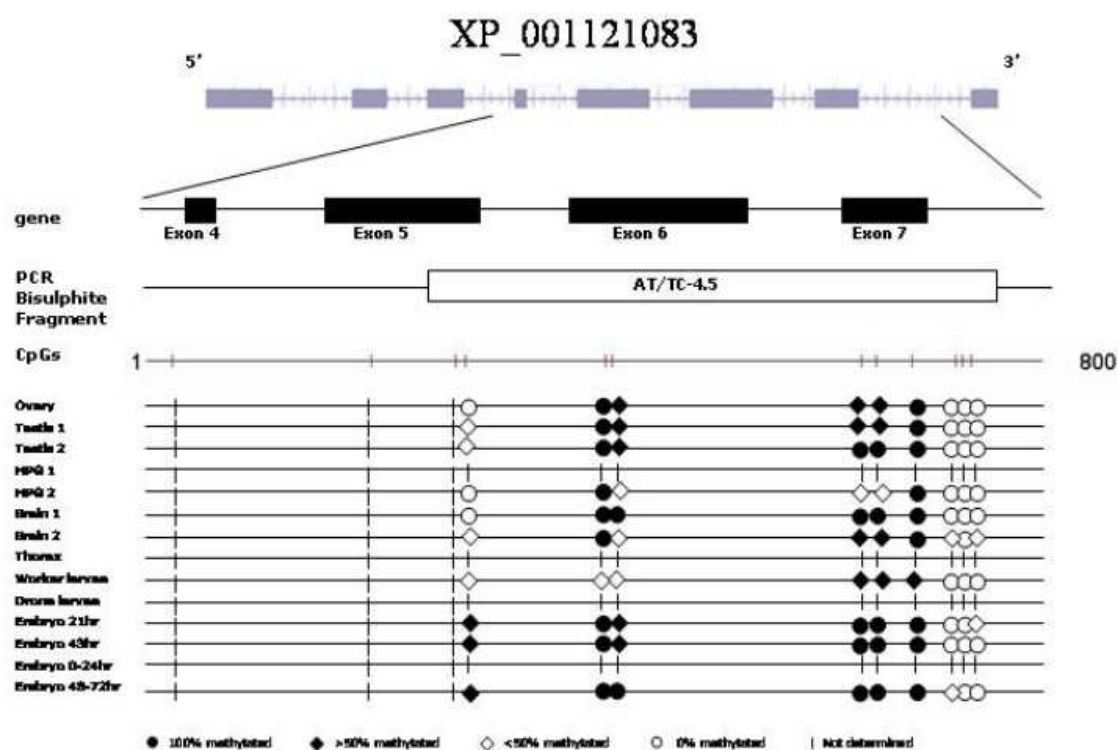
**Figure 2.1 (cont.)** ( $\pm$ SD) is shown. (D) Honey bee genome contains 5-methyl deoxycytosine. Genomic DNA from adult honey bees (a, b, d, and e), larvae (c), or rat liver (f), was digested to nucleotide monophosphates and analyzed by reverse phase high-performance liquid chromatography for dCM. Digested bee DNA was spiked (a) with exogenous dCM and analyzed in parallel with the digested DNA (b), clearly identifying a peak in adult honey bee coincident with dCM enrichment (indicated by arrows). DNA from 3- to 4-day-old larvae (c) and adults (d) both contain dCM. The dCM peak from the larger sample (73.8  $\mu$ g) of adult bee DNA (e) co-elutes with that from a lesser amount (17.8  $\mu$ g) of rat liver DNA (f). The x axis is time (min) and the y axis is absorbance at 260 nm.



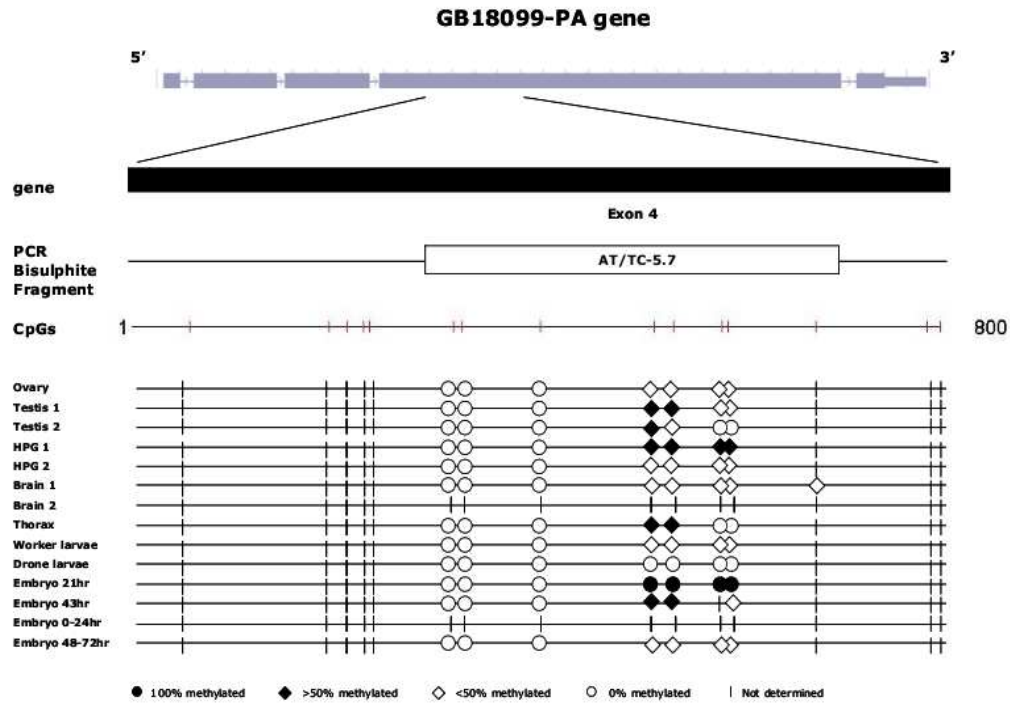
**Figure 2.2.** The honey bee genome encodes orthologs of vertebrate DNA methylation proteins. Predicted amino acid sequences for AmDNMT proteins and the AmMBDs were aligned with known orthologs using the CLUSTAL W program. Conserved amino acids are identified in the consensus sequences with residues conserved across all family members in shaded uppercase. A) Sequence alignment of honey bee, human, *Drosophila*, and prokaryotic (*M. spr.*) (Buhk et al., 1984) methyltransferases across the five methyltransferase motifs highly conserved in the catalytic domains of all DNMT family members. B) Sequence alignments of the deduced translation of the two honey bee MBD isoforms compared with MBD proteins from human (hMBD2 aa 137-393 and hMBD3), *Bombyx mori*, and *D. melanogaster* (dMBD). AmMBD-1 encodes 205 aa while AmMBD-s encodes 183 aa. Key residues (Y41 and K37, denoted by ^) required for recognition of the methyl group and methyl DNA binding activity are conserved (Hendrich & Tweedie, 2003).



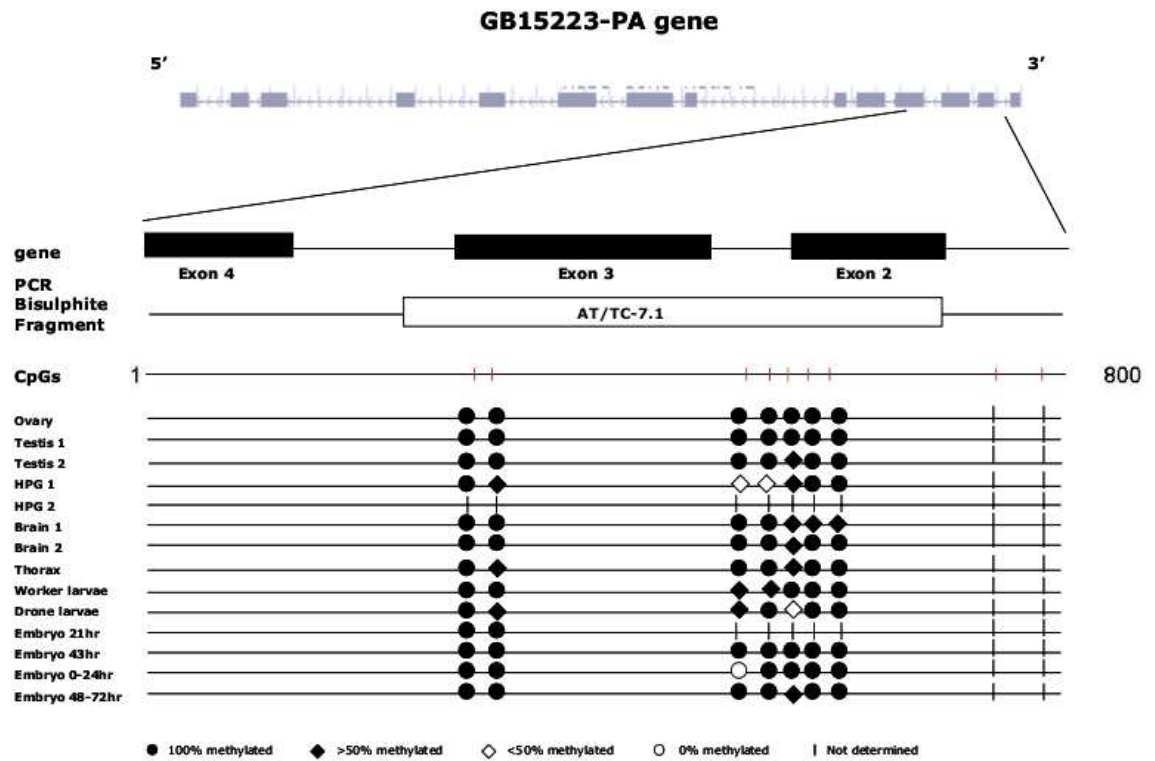
**Figure 2.3.** Predicted gene GB16767 is methylated in the honey bee genome. Analysis of the ATCT-5.1 (methylated) sequence placed it in exon 3 of the annotated gene GB16767. Bisulfite sequencing of the flanking regions identified a total of 17 CpGs, 12 of which were methylated to varying degrees in at least 1 tissue analyzed. GB16767 exons 3 and 4 are methylated exclusively at CpGs in all tissues examined. Each data set represents bisulfite sequencing of a pool of bisulfite converted bee genomic DNA. Each data point is derived from the percent of C converted to T at a particular site, indicative of dCM.



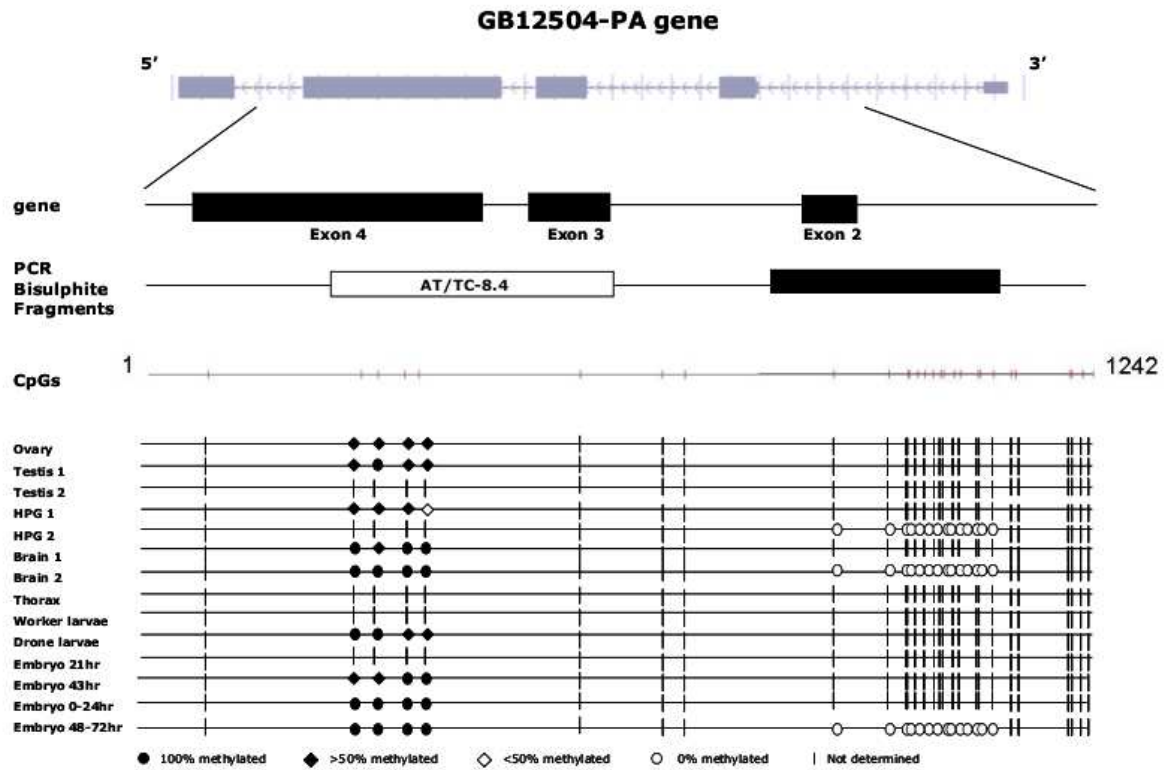
**Figure 2.4.** Predicted gene XP\_001121083 is methylated in the honey bee genome. Analysis of the AT/TC-4.5 (methylated) sequence placed it in exons 5, 6 and 7 of the gene XP\_001121083 predicted by Gnomon. It is similar to *Drosophila* CG12042-PA. Bisulfite sequencing of this region identified a total of 10 CpGs, 8 of which were methylated to varying degrees in at least 1 tissue analyzed. Methylation is exclusively at CpGs in all tissues examined. Data sets and data points as described in Fig 2.3 legend.



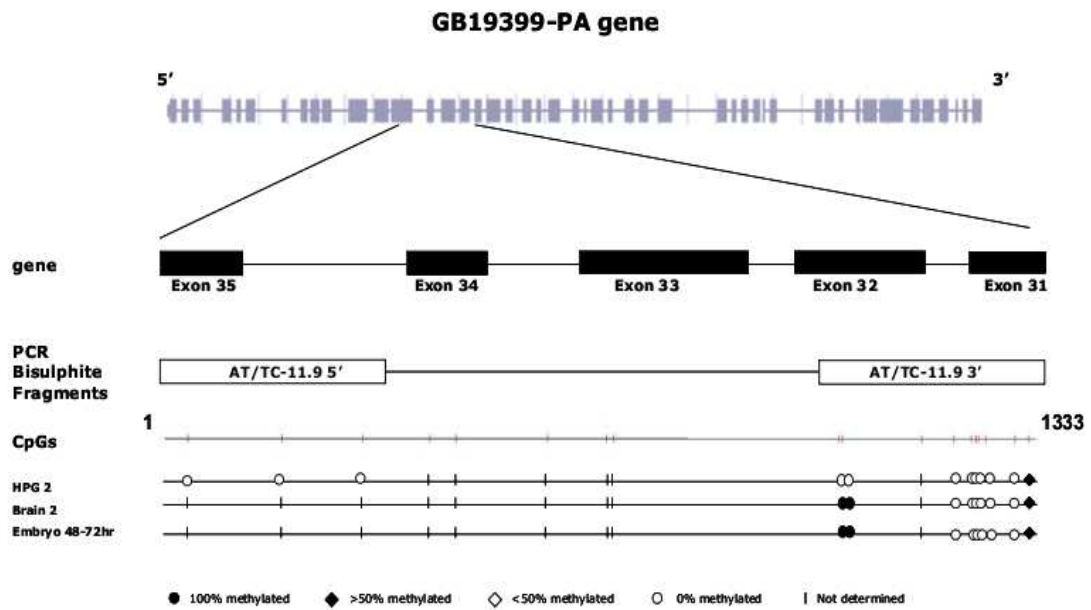
**Figure 2.5.** Predicted gene GB18099-PA is methylated in the honey bee genome. Analysis of the AT/TC-5.7 (methylated) sequence placed it in exons 4 of the predicted gene by Gnomon. Bisulfite sequencing of this region identified a total of 8 CpGs, 5 of which were methylated to varying degrees in at least 1 tissue analyzed. Methylation is exclusively at CpGs in all tissues examined. Data sets and data points as described in Fig 2.3 legend



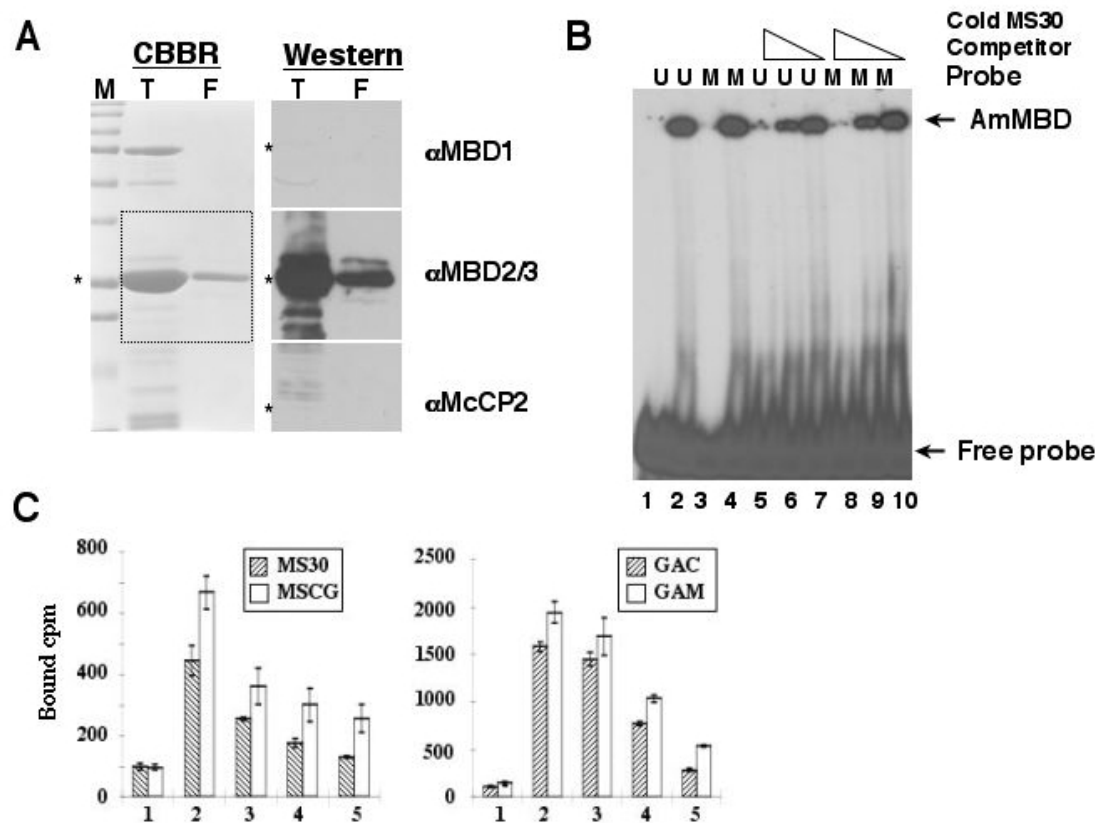
**Figure 2.6.** Predicted gene GB15223-PA is methylated in the honey bee genome. Analysis of the AT/TC-7.1 (methylated) sequence placed it in exons 2 and 3 of the predicted gene by Gnomon. Bisulfite sequencing of this region identified a total of 7 CpGs, and all of them were methylated to varying degrees in at least 1 tissue analyzed. Methylation is exclusively at CpGs in all tissues examined. Data sets and data points as described in Fig 2.3 legend



**Figure 2.7.** Predicted gene GB12504-PA is methylated in the honey bee genome. Analysis of the AT/TC-8.4 (methylated) sequence placed it in exons 3 and 4 of the predicted gene by Gnomon. Bisulfite sequencing of this sequence and a flanking region identified a total of 19 CpGs, 4 of which were methylated to varying degrees in at least 1 tissue analyzed. Methylation is exclusively at CpGs in all tissues examined. Data sets and data points as described in Fig 2.3 legend.



**Figure 2.8.** Predicted gene GB19399-PA is methylated in the honey bee genome. Analysis of the AT/TC-11.9 5' and AT/TC-11.9 3' (methylated) sequences placed them in exon 35 and exons 31 and 32, respectively, of the predicted gene by Gnomon. Bisulfite sequencing of these regions identified a total of 13 CpGs, 3 of which were methylated to varying degrees in at least 1 tissue analyzed. Methylation is exclusively at CpGs in all tissues examined. Data sets and data points as described in Fig 2.3 legend



**Figure 2.9** AmMBD-1 interacts preferentially with methylated dsDNA in vitro. A) Recombinant AmMBD-1 (indicated by \*) was purified by cobalt affinity chromatography and MonoS ion exchange to a single detectable polypeptide (left panel, Coomassie Brilliant Blue R250). Western blotting (showing the boxed region) using polyclonal antibodies reactive to vertebrate MBD1 (upper), MeCP2 (lower), or MBD2/3 (middle) showed AmMBD-1 is cross-reactive with MBD2/3 antiserum (right) (Wade et al., 1999). B) EMSA identified AmMBD-1 as a DNA binding protein by its ability to bind both methylated (lane 4, MSCG) and unmethylated (lane 2, MS30) DNA in vitro (lanes 1 and 3 are no protein controls). Competition for DNA binding was carried out using 0.1 ng, 1 ng or 5 ng of cold unmethylated MS30 DNA. C) Filter-binding assays to quantify binding of recombinant AmMBD-1 to an unmethylated or methylated DNA probe (lane 2) and uncovered a consistent binding preference for methylated DNA. Competition for AmMBD-1 DNA binding (0.1 ng, 1 ng or 5 ng cold unmethylated DNA, lanes 3, 4, 5 respectively) indicated unmethylated DNA was less efficient at competing the AmMBD-1 interaction with a methylated DNA probe than an unmethylated DNA probe, supporting AmMBD-1 binds preferentially to methylated DNA. Lane 1 is free probe control.

## Chapter 3

# Nutritional Regulation of Division of Labor in Honey Bees: Toward a Systems Biology Perspective

Previous published work<sup>2</sup>

### Abstract

Organisms adapt their behavior and physiology to environmental conditions through processes of phenotypic plasticity. In one well-studied example, the division of labor among worker honey bees involves a stereotyped yet plastic pattern of behavioral and physiological maturation. Early in life, workers perform brood care and other in-hive tasks and have large internal nutrient stores; later in life, they forage for nectar and pollen outside the hive and have small nutrient stores. The pace of maturation depends on colony conditions, and the environmental, physiological, and genomic mechanisms by which this occurs are being actively investigated. Here we review current knowledge of the mechanisms by which a key environmental variable, nutritional status, influences worker honey bee division of labor. These studies demonstrate that changes in individual nutritional status and conserved food-related molecular and hormonal pathways regulate the age at which individual bees begin to forage. We then outline ways in which systems biology approaches, enabled by the sequencing of the honey bee genome, will allow

---

<sup>2</sup> Seth A. Ament, Ying Wang, and Gene E. Robinson, "Nutritional regulation of division of labor in honey bees: toward a systems biology perspective," *Wiley Interdisciplinary Reviews: Systems Biology and Medicine* (2010): published on line.

researchers to gain deeper insight into nutritional regulation of honey bee behavior, and phenotypic plasticity in general.

## Introduction

Many animals are able to alter their behavior and physiology in response to changes in the environment. At times, these changes in behavior and physiology are stable for long periods, a phenomenon known as phenotypic plasticity (West-Eberhard, 2003). For instance, short periods of food deprivation stimulate feeding and the mobilization of stored nutrients meet an individual's immediate energetic needs. But prolonged food deprivation can also lead to much longer-term effects, causing individuals to enter extended periods of inactivity (Fielenbach & Antebi, 2008), alter their reproductive strategy (Wade & Schoneider, 1992), or lose their position in a dominance hierarchy (Markiewicz & O'Donnell, 2001). In humans, chronic food deprivation early in life may lead to a propensity toward obesity and diabetes in later life (Kahn, 2001; Ravelli, Stein, & Susser, 1976). The mechanisms that enable and constrain plasticity in behavior and physiology are not well understood, but it is clear that they often involve coordinated and long-lasting changes in gene expression, brain circuitry, and physiology (Whitfield et al., 2006) (reviewed in Robinson, Fernald, & Clayton, 2008). We are developing a systems biology approach to studying the mechanisms that underlie phenotypic plasticity in the honey bee, *Apis mellifera*.

The adult worker honey bee exhibits phenotypic plasticity as part of the division of labor that defines the roles of each individual in the colony (see a later session: honey bee primer). For the first 2-3 weeks of their adult life, bees perform brood care (nursing)

and other tasks inside the hive. They then transition to foraging for nectar and pollen outside the hive for the final 1-2 weeks of their life (Winston, 1991).

Nursing and foraging define the two most distinct and highly stable states in the life of a bee. Under ordinary circumstances, once a bee transitions into the foraging state she ceases to perform other tasks; although some foragers can be forced to perform brood care by removing all of a colony's younger bees, these reverted nurses are inefficient at rearing brood, suggesting that some aspects of the transition are irreversible (Robinson, 1992). Microarray experiments show that the brains of nurses and foragers differ in the expression of hundreds to thousands of genes, up to 40% of the genes expressed in the brain (Alaux et al., 2009; Whitfield, Cziko, & Robinson, 2003). This finding, together with results from neuroanatomical and neurochemical studies (Robinson, 2002), demonstrates that honey bee behavioral states are defined by major changes in the brain.

The behavioral states of nursing and foraging are also associated with a number of physiological changes outside the brain. A honey bee loses half its abdominal lipid stores during the transition from working in the hive to foraging. The lipid loss is stable; a forager's lipid stores subsequently remain low even when food is plentiful (Toth & Robinson, 2005). The coordination of these changes in the brain and periphery and the importance of nutritional changes to division of labor will be described in detail below.

Although the stages of behavioral maturation are stereotyped and stable in honey bees, their timing is plastic. When older bees are removed from the hive, young bees rapidly establish a division of labor; some individuals begin to forage as young as 5 days of (adult) age (Robinson, 1992), over 2 weeks earlier than normal. The pace of behavioral maturation also depends on the needs of the colony. For instance, when a colony lacks

stored pollen and honey, bees also become precocious foragers, helping to replenish these essential provisions (Schulz, Huang, & Robinson, 1998). Plasticity in the age at onset of foraging makes it possible with the honey bee to dissociate age and behavior, something usually impossible to do when studying maturation. Most of the maturational changes in physiology (Toth & Robinson, 2005) and brain gene expression (Alaux et al., 2009; Whitfield et al., 2003) discovered in honey bees are associated with a bee's behavior, and not her age.

The stability of nursing and foraging, along with the ability to influence the transition between these states through a variety of techniques, has made honey bee division of labor an important model for understanding phenotypic plasticity (Robinson, 2002; Whitfield et al., 2006). Here we review current knowledge of the mechanisms by which a key environmental variable, nutritional status, influences division of labor. We then outline ways in which systems biology approaches, enabled by the sequencing of the honey bee genome, will allow researchers to gain deeper insight into nutritional regulation of honey bee behavior, and phenotypic plasticity in general.

### *Honey Bee Primer*

Honey bees are social insects, living together in colonies containing tens of thousands of individuals (for an overview of honey bee biology, see Winston, 1991). Colony life is organized by a complex and sophisticated division of labor. Each colony contains a single queen, who is specialized for reproduction and spends most of her time laying eggs. Males, called drones, are relatively rare, and their sole role is to mate. The vast majority of the individuals in the hive are sterile worker bees that are responsible for

all of the other tasks performed by the colony. The tasks performed by worker bees are further divided up among individuals via the process of behavioral maturation that is the subject of this review. For the first 2-3 weeks of adult life, worker bees specialize on brood care (nursing). They then switch for a few days to any of a number of more specialized tasks such as building honeycomb cells, storing food in honeycomb cells, or guarding the hive entrance against intruders. Finally, for the remaining 1-2 weeks of their life, worker bees forage outside the hive for nectar and pollen, the colony's sole sources of food.

Honey bees have been the subjects of scientific study for hundreds of years (reviewed in Hölldobler & Wilson, 2008). Bees are noted models for behavioral plasticity, communication (Seeley, 1996), and learning and memory (Hammer & Menzel, 1995) among other subjects. Many of these studies have relied on three key attributes of the honey bee system. First, the millennia-old tradition of beekeeping allows for high-throughput experiments under natural and seminatural colony conditions in the field, and large numbers of bees of all life stages are readily available for experiments (Winston, 1991). Second, causal experiments are possible using a variety of environmental, pharmacological, and molecular manipulations of whole colonies or of individuals within colonies (Ament, Corona, Pollock, & Robinson, 2008; Ben-Shahar, Robichon, Sokolowski, & Robinson, 2002; Jaycox, 1976; Nelson, Ihle, Fondrk, Page, & Amdam, 2007; Schulz & Robinson, 2001; Schulz et al., 1998; Toth, Kantarovich, Meisel, & Robinson, 2005). For studies of division of labor, these experiments typically involve marking bees of known age early in life, manipulating them, and tracking the age at which they first begin to forage. Often these studies are done using single-cohort colonies

that are made from a single cohort of 1000-2000 one-day-old bees at the start of the experiment; because bees begin foraging at a much younger age in the absence of a preexisting foraging force, the length of the experiment is reduced and certain treatment designs become more feasible. Finally, the sequencing of the honey bee genome (Honey Bee Genome Sequencing Consortium, 2006) has enabled efficient discovery of candidate genes using microarrays (Grozingier, Sharabash, Whitfield, & Robinson, 2003; Whitfield et al., 2006; Whitfield et al., 2003), informatics (Sinha, Ling, Whitfield, Zhai, & Robinson, 2006), and other functional genomics technologies.

An important attribute of honey bee societies is that the colony is an integrated level of biological organization, not merely an aggregation of individuals (Hölldobler & Wilson, 2008). Colonies have specific attributes or traits, but they can be traced to the traits expressed in individuals. A prime example of this comes from artificial selection of colonies for high and low pollen hoarding. In these experiments, selection for a colony-level trait, the amount of stored pollen in the hive, revealed genetic variation for a variety of nutritionally related traits in individuals, including the age at first foraging, foraging preference for nectar or pollen, and gustatory sensitivity (Page, Erber, & Fondrk, 1998; Pankiw & Page, 2001). Both the proximate and ultimate mechanisms by which social phenotypes arise from individual phenotypes are active areas of research (Smith, Toth, Suarez, & Robinson, 2008).

## **Nutritional Regulation of Division of Labor in Honey Bee Colonies**

Given that many of the tasks performed by honey bees relate to collecting, handling, and distributing food, it is perhaps not surprising that their division of labor is regulated by nutritional stimuli. A number of studies have now elucidated mechanisms by which changes in colony nutrition lead to changes in individual behavior, including roles for well-conserved signaling pathways that regulate hunger and the more straightforward food-related behaviors of solitary species.

Honey bees differ from nonsocial species in that food-related tasks are performed in an intrinsically social context. Because a colony of social insects is widely thought to be the unit upon which natural selection acts, tasks performed by workers are typically interpreted in terms of how they benefit the whole colony, rather than the individual worker. For instance, foragers directly consume very little of the food they collect (Winston, 1991). Instead, they initially store the nectar that they have collected in their foregut, sometimes called the honey stomach, and then regurgitate it to food storing bees, who are responsible for converting it into honey and storing it. Much of the actual food consumption occurs via sharing among adult bees of processed nectar and jelly, a proteinaceous glandular secretion produced primarily by nurses and subsequently passed to foragers (Crailsheim, 1992; Crailsheim et al., 1992). Moreover, a forager consumes honey inside the hive before leaving the hive for a foraging trip (Winston, 1991), so a bee is not food-deprived when she leaves the hive to forage. For these reasons, while division of labor revolves around the performance of food-related tasks, foraging itself is not an appetitive or consummatory behavior in the traditional sense. It has therefore been all the

more exciting to learn over the past decade that nutritionally related mechanisms within individuals play an important role in the regulation of honey bee social behavior. Some of the key findings are described in the following paragraphs.

In the first paper describing an effect of nutrition on honey bee division of labor, Schulz et al. (Schulz et al., 1998) monitored the age at first foraging for bees in matched colonies that were either well-fed with excess honey and pollen or food-deprived with no pollen and only enough honey to last 2-3 days. Bees from food-deprived colonies began foraging earlier than bees from well-fed colonies, establishing that food availability inside the hive was involved in behavioral maturation. The important factor affecting behavioral maturation appears to be intake of nutrients by individuals. To establish this, the authors uncoupled the usually linked factors of food availability and food storage. They created well-fed bees in starved colonies by providing a feeder inside the hive to nurture individual bees, but bees were prevented from storing food in honeycomb cells by use of a vacuum pump mounted on the back of the honeycomb frame. Under these conditions, bees foraged as though they were well-fed, indicating that individual food intake rather than assessment of stored food caused precocious foraging in food-deprived colonies.

The most obvious means by which individual nutrition could lead to long-term changes in behavior is through changes in physiology. Nurses and foragers differ dramatically in several aspects of nutritional physiology. Toth and Robinson (Toth & Robinson, 2005) studied the relationship between a bee's lipid stores and her behavioral state. Bees reach adulthood with low lipid stores and gain lipid mass during the first few days of adult life. They maintain large lipid stores for as long as they nurse brood, but

during the transition to foraging they lose half of their abdominal lipid. Interestingly, this lipid loss does not depend on the performance of energy-intensive flights outside the hive; experimentally hive-restricted bees lose almost as much lipid mass as controls during the days leading up to foraging. In a follow-up experiment, Toth et al. (Toth et al., 2005) treated bees with a fatty acid synthesis inhibitor to artificially reduce lipid stores of bees in an otherwise well-fed colony. Individuals with reduced lipid stores initiated foraging earlier than controls, suggesting that changes in lipid stores can accelerate behavioral maturation.

Independent evidence for the causal effects of stored nutrients on honey bee behavioral maturation comes from the work on the storage protein Vitellogenin (Vg). Like lipid stores, Vg titers and *vg* transcription are high in nurses and decline prior to the onset of foraging (Amdam, Norberg, Fondrk, & Page, 2004; Rutz & Lüscher, 1974). Inhibition of *vg* synthesis using RNAi causes precocious foraging (Antonio, Guidugli-Lazzarini, do Nascimento, Simões, & Hartfelder, 2008; Nelson et al., 2007). This result indicates that storage proteins, like stored lipids, are causal for the timing of behavioral maturation. In the colony environment, the availability of food to each individual is influenced both by the availability of forage outside the hive and of stored nectar and pollen, and by explicitly social factors such as the sharing of food among individuals (Crailsheim, 1992). We can summarize that the availability of food to each individual within the colony leads to changes in her ability to store nutrients inside her body, and that these changes in physiology cause changes in behavior.

Bees at different stages of behavioral maturation also differ in their basal metabolism. Harrison (Harrison, 1986) demonstrated that foragers have greater capacity

for flight than nurses due to a higher capacity for energy metabolism in flight muscle. Interestingly, although energy metabolism in flight muscle and in the body as a whole increases during behavioral maturation, the metabolic capacity of the brain is actually lower in foragers than nurses. This was initially discovered through pathway analysis of microarray experiments (Ament et al., 2008) (Figure 3.1) and later confirmed by assays of mitochondrial function (Alaux et al., 2009). These results suggest that the bee brain has specific metabolic needs that are distinct from those of the rest of the animal, as is well established in other species (Bough et al., 2006; Mootha et al., 2003). Nonetheless, behavior is responsive to the nutritional and energetic state of the organism as a whole. This is accomplished in both vertebrates and invertebrates by changes in the activity of brain circuits that are both uniquely exposed to the conditions of the circulatory system and signaled to change their activity via endocrine signals (Leopold & Perrimon, 2007; Morton, Cummings, Baskin, Barsh, & Schwartz, 2006; Wu & Brown, 2005).

## **Roles for Conserved Food-related Pathways in Division of Labor in Honey Bee Colonies**

The recent sequencing of the honey bee genome (Honey Bee Genome Sequencing Consortium, 2006) has made it possible to explore the roles of nutritionally related molecular pathways in the regulation of honey bee division of labor. The mechanisms that regulate food intake and nutritional physiology are under active investigation in many laboratories, in part because of the clinical need to develop treatments for obesity (Haslam & James, 2005). In insects, the synthesis and metabolism of lipids, proteins, and carbohydrates are accomplished by the fat bodies, which are most prevalent lining the walls of the abdomen and have functions comparable to both the liver and adipose tissues

of vertebrates (Chapman, 1998; Leopold & Perrimon, 2007). As in vertebrates, there is active communication between these peripheral tissues and the nutrient-sensing brain circuits that regulate both behavior and physiology (Leopold & Perrimon, 2007), and conserved signaling pathways have roles in this process across vertebrate and invertebrate taxa (Leopold & Perrimon, 2007; Wu & Brown, 2005).

Many insights on nutritional regulation have come from traditional genetic model systems such as the fruit fly *Drosophila melanogaster* and the mouse *Mus musculus*. Conserved food-related pathways have generally been studied in the context of homeostasis, and changes in nutrient stores (either wasting or obesity) are viewed as dysfunctions. In the honey bee the situation is different - the extreme changes in behavior that occur during behavioral maturation are associated with striking changes in nutritional physiology and food-related behaviors, namely, the loss of abdominal lipid that precedes the onset of foraging. It might thus be possible to use the honey bee to identify mechanisms that regulate stable lipid loss. Studying these mechanisms in the bee also provides an opportunity to explore the adaptation of conserved, nutritionally related signaling pathways to the derived, social context, perhaps providing insights into the evolution of behavior.

Insulin signaling and the related target of rapamycin (TOR) pathway are the best understood, most well conserved, pathways linking nutrition to changes in physiology and behavior (Morton, Cummings, Baskin, Barsh, & Schwartz, 2006). In *Drosophila* and rodents, insulin and TOR act as satiety signals in the brain, decreasing food intake in response to acute or chronic increases in circulating nutrients, such as after a meal or in the context of elevated adiposity (Cota et al., 2006; Morton et al., 2006; Wu, Zhao, &

Shen, 2005). Recently, it has been shown in the honey bee that insulin-related transcripts are upregulated in the head and fat bodies of foragers compared to nurses (Ament, Corona, Pollock, & Robinson, 2008; Corona et al., 2007). Moreover, pharmacological manipulations of the TOR pathway led to a delay in the initiation of foraging in a seasonally dependent fashion (Ament et al., 2008). These results suggest a causal role for conserved nutrient-sensing pathways in the regulation of honey bee division of labor.

In both vertebrates and invertebrates, insulin-like peptides are typically secreted in response to high levels of circulating nutrients and synthesized at greater levels in response to high levels of stored nutrients, acting in both cases as negative feedback signals that inhibit further food consumption (Morton et al., 2006). Interestingly, while stored nutrients influence insulin signaling gene expression in the honey bee, they do so in the opposite direction: bees with larger lipid stores have lower expression of insulin-like peptides, whether in the context of natural maturation (nurses vs. foragers), manipulations of colonies (well-fed vs. food-deprived colonies), or manipulations of individuals (caged bees fed a rich or poor diet) (Ament et al., 2008).

It is possible that this reversed polarity in the relationship between insulin signaling and nutrition reflects a change in the adipostatic set point of foragers relative to nurses, rather than the traditional homeostatic mechanism associated with insulin signaling (Ament et al., 2008). In this view, the combination of high insulin synthesis and high insulin sensitivity reflects, or perhaps causes, a shift from high to low adiposity during behavioral maturation (and in response to experimental nutritional manipulations). Similar reasoning has been used to explain relationships between nutrient-sensing pathways and variation in nutrient stores in the contexts of mammalian torpor

(Klingenspor, Niggemann, & Heldmaier, 2000; Tups et al., 2006) and insect diapause (Williams et al., 2006). Reversed IIS gene expression and the suggested set point regulation do not occur in all contexts in the life of a bee. Injections of insulin peptides acutely reduce blood carbohydrates in the honey bee, as in other species (Bounias & Pacheco, 1979). And a more traditional (homeostatic) relationship between nutrition and insulin signaling is seen during larval development (Wheeler, Buck, & Evans, 2006). These results suggest that the cooption of insulin signaling for the regulation of adult worker division of labor entailed changes in some aspect of the signaling system, but that the system is intact in its fundamentals within the species.

The discovery that nutrient-sensing pathways are involved in the regulation of worker division of labor helps to understand earlier results implicating feeding-related genes in this process. The gene known as *foraging* (*for*) encodes a cGMP-dependent protein kinase (PKG) (Osborne et al., 1997). Variant alleles of *for* were originally shown in *Drosophila* to underlie naturally occurring variation in larval foraging strategies, and brain expression differences in its bee homolog were subsequently shown to regulate behavioral maturation (Ben-Shahar, Robichon, Sokolowski, & Robinson, 2002). The effect of *for* on behavioral maturation in the bee was initially linked to age-related changes in phototaxis (both foragers and bees treated with a membrane-permeable form of cGMP have increased attraction to light) (Ben-Shahar, Leung, Pak, Sokolowski, & Robinson, 2003). However, since these initial studies, *for* has been shown in *Drosophila* to regulate nutrient storage and nutrient utilization (Kaun, Chakaborty-Chatterjee, & Sokolowski, 2008; Kaun et al., 2007). Perhaps *for* also has a role in regulating similar physiological processes in the context of worker honey bee division of labor.

*Malvolio* (*mvl*) is another gene involved in regulating feeding-related behavior in *Drosophila* shown to also be involved in regulating honey bee division of labor. This gene was initially identified in *Drosophila* in a screen for mutants with defects in taste behavior (Rodrigues, Cheah, Ray, & Chia, 1995) and is involved in the transport of manganese ( $Mn^{2+}$ ) ions (Orgad, Nelson, Segal, & Nelson, 1998). The bee homolog of *mvl* is highly expressed in the brains of foragers compared to nurses, and  $Mn^{2+}$  supplementation causes both precocious foraging and increased responsiveness to sugar (Ben-Shahar, Dudek, & Robinson, 2004). Together, the involvement of insulin signaling, *for*, and *mvl* suggest that there may be a general role for feeding-related pathways in the regulation of worker division of labor.

The role of nutritionally related pathways in honey bee division of labor also provides an important new perspective on the role that the multifunctional insect hormone juvenile hormone (JH) plays in this process. A role for JH in division of labor was first shown more than 30 years ago (Jaycox, 1976), and a variety of subsequent studies established that JH paces behavioral maturation (Robinson, Page, Strambi, & Strambi, 1989; Sullivan, Jassim, Fahrbach, & Robinson, 2000) and underlies age-related changes in energy metabolism (Sullivan et al., 2003), brain neurochemical levels (Schulz, Sullivan, & Robinson, 2002), and, perhaps, brain structure (Velarde, Robinson, & Fahrbach, 2009). In adult worker bees, JH and the key storage protein vitellogenin are mutually inhibitory, and it has been proposed that this feedback loop serves as a timing mechanism in behavioral maturation (Amdam & Omholt, 2003). Recent studies in a variety of insect species have shown that JH and insulin signaling often have overlapping functions, and there appears to be extensive crosstalk between the two signaling

pathways (Corona et al., 2007; Tu, Yin, & Tatar, 2005; Wu & Brown, 2005). Consequently, it has been possible to link nutrient-sensing mechanisms to JH signaling (Corona et al., 2007; Page & Amdam, 2007). Elucidating the precise interactions between these signaling pathways and their effects on target genes would be greatly facilitated by systems biology approaches.

The results reviewed here on conserved feeding-related pathways provide context for the role of the neuromodulator octopamine (OA) in both worker division of labor (Schulz & Robinson, 2001) and in the assessment of food rewards during honey bee foraging (Barron, Maleszka, Meer, & Robinson, 2007; Scheiner, Plückhahn, Öney, Blenau, & Erber, 2002). OA has well-established roles in the behavioral and physiological responses to nutritional cues in a variety of insect species (Adamo, Linn, & Hoy, 1995; Roeder, 1999), and OA is known to interact with JH in honey bees (Kaatz, Eichmüller, & Kreissl, 1994; Schulz et al., 2002). Thus, the regulation of division of labor involves a web of interconnected feeding and nutritionally related pathways.

## **Division of Labor as Nutritionally Mediated Plasticity**

Based on the studies reviewed above, it is now clear that nutritionally related pathways regulate division of labor in honey bee colonies. We stress three characteristics of this system that are likely to inform future studies. First, the regulation of behavioral maturation involves the coordination of changes occurring in multiple tissues, in part via hormonal signaling pathways. Insulin and JH are synthesized in glandular cells, in or associated with the brain, and target both brain circuits and peripheral tissues. Vitellogenin and lipids are released from the fat bodies and are thought to directly or

indirectly modulate the activity of brain circuits. Some octopaminergic neurons are also involved in transmitting peripheral nutritional signals into the brain (Roeder, 1999). The importance of these cross-tissue communication signals, as well as the strong coordination between nutrient stores and behavioral state, suggests that any model for the regulation of division of labor should explicate the changes induced by these signals in both the periphery and brain, with an emphasis on the fat bodies and on specific nutrient-sensing circuitry.

Second, behavioral maturation involves massive changes in gene expression. As described above, as many as 40% of genes are differentially expressed in the brains of nurses and foragers (Whitfield, Cziko, & Robinson, 2003), and a similarly large number of genes change expression in the fat cells (Ament and Robinson, unpublished data). These results suggest that transcriptional control is an important part of the regulation of division of labor.

Finally, the behavioral and physiological states of nursing and foraging are highly stable, and these two groups of bees typically have different responses to similar conditions. When a colony's nutrient stores run low, existing foragers immediately leave the hive in search of nectar and pollen, but it takes at least a day for bees that had not previously been foraging to enter the foraging force (Schulz, Huang, & Robinson, 1998). This delay suggests that the transition into foraging requires a maturational process. The regulation of physiology seemingly also involves a stable state change since a forager's lipid stores remain low in well-fed colonies, even in bees that are forced to revert from foraging back to nursing (Toth & Robinson, 2005). Together, these results suggest that the hormones and related signals that regulate behavioral maturation are involved in

setting target cells into distinct regulatory states in nurses and foragers. An important goal for future research is to elucidate how signaling pathways generate these state transitions at the molecular level, in both fat cells and in the brain.

## **Toward a Systems Biology of Honey Bee Division of Labor**

Division of labor, like other examples of phenotypic plasticity, involves massive and stable changes in gene expression and regulatory roles for hormones and signal transduction mechanisms (Smith, Toth, Suarez, & Robinson, 2008). While signal transduction mechanisms are likely the switches by which environmental cues regulate phenotypes, the broader changes in gene expression have important roles as the effectors of each behavioral or physiological state. According to this line of reasoning, various forms of phenotypic plasticity such as division of labor are regulated through hormonally related transcriptional networks that link environmental changes to alternative gene expression states, ultimately generating differences in phenotype. Systems biology provides the framework to test the generality of this idea.

The regulation of cellular regulatory states is best understood in the context of development, in which different regulatory states specify the different body parts that make up an animal. During development, the progression of states is controlled by hierarchical gene regulatory networks (GRNs) (reviewed in Davidson, 2006). In these networks, signaling pathways specify when and where transcription factors (TFs) activate batteries of target genes that determine each cell's phenotype. We hypothesize that in the context of honey bee division of labor, the hormonal signaling pathways that influence behavioral and physiological maturation do so by affecting the activity of GRNs in much

the same way (Figure 3.2). The difference is that these changes occur in cells that have already differentiated so that the regulatory networks specify the physiological, rather than developmental states of cells.

A challenge in modeling the GRNs underlying division of labor will be to characterize the network properties that both enable and constrain phenotypic change. Specifically, how do nutritionally related hormones determine network regulatory states, and how do these regulatory states in turn determine the behavioral and physiological characteristics of nurses and foragers? These questions can be addressed by characterizing the GRNs that underlie each behavioral state, focusing on the parts of the network that are regulated by hormones. A specific hypothesis, derived from studies of transcriptional networks in other contexts (Workman et al., 2006), is that the increase in insulin and JH during maturation induces forager-like traits by changing the locations that hormonally responsive TFs bind in the genome, leading to activation of foraging-related genes. Alternatively, as has been revealed in other contexts for some nuclear hormone receptors (Perissi & Rosenfeld, 2005), hormonally related TFs may bind at the same locations throughout maturation but have different effects on transcription depending on the presence or absence of ligands and cofactors. These changes in the activity of sequence-specific TFs are likely to occur both through changes in the expression of the TFs themselves and through changes in the open or closed states of chromatin structures at target genes. Although understanding the dynamic properties of transcriptional networks is an active area of research (Bonneau, 2008), it is known that network connections can be turned on or off rapidly in some contexts. For instance, the response of single-celled organisms to DNA damage involves changes in active network

connections that occur within minutes of chemical perturbations (Workman et al., 2006), so it is reasonable to think that changes in transcription factor activity within GRNs occur in real time during plasticity. In honey bees, relevant GRNs are located in the fat cells and brain circuits, which are among the tissues that mediate physiological and behavioral differences between nurses and foragers.

The GRN framework suggests an established methodology through which to elucidate the mechanisms that specify the states of fat and brain cells in nurses and foragers. Currently available technologies in honey bees now allow elucidation of the active GRNs in nurse and forager tissues. Functional interactions between genes can be inferred from large-scale microarray experiments by generating gene co-expression networks. Causal relationships within these correlational networks can then be established in a variety of ways. For instance, direct targets of TFs can be identified by characterizing the genomic binding sites for TFs using chromatin immunoprecipitation coupled to genomic tiling microarrays (ChIP-chip) or high-throughput sequencing (ChIP-seq). Targets of TFs can also be elucidated by gene expression profiling after genetic manipulations of TFs; genetic perturbation using RNA interference is now routine in the bee, especially when targeting peripheral tissues (Amdam, Simoes, Guidugli, Norberg, & Omholt, 2003; Barchuk, Figueiredo, & Simoes, 2008; Patel et al., 2007). Finally, the functional relationships among genes discovered using any of these techniques can be inputted into computational models to construct a quantitative model of a GRN (Workman et al., 2006) (reviewed in Bonneau, 2008).

Between brain circuits and fat cells, it seems more prudent to begin by characterizing the GRNs in fat cells. This is because it is straightforward to harvest large

numbers of relatively homogeneous cells from the fat bodies and because the roles of insulin and JH are well known in this tissue (Leopold & Perrimon, 2007). In order to identify dynamic changes in the network, GRNs will need to be deduced in both nurses and foragers. Since key hormones are already known, characterizing the targets of only a small handful of TFs will likely be sufficient to gain significant insight into this system. In time, this approach could be expanded to elucidate GRNs in brain circuits and for additional TFs. In addition to advancing our understanding of plasticity, this approach also has the potential to identify genes that regulate the stable weight loss that occurs during the maturation process.

## **Conclusion**

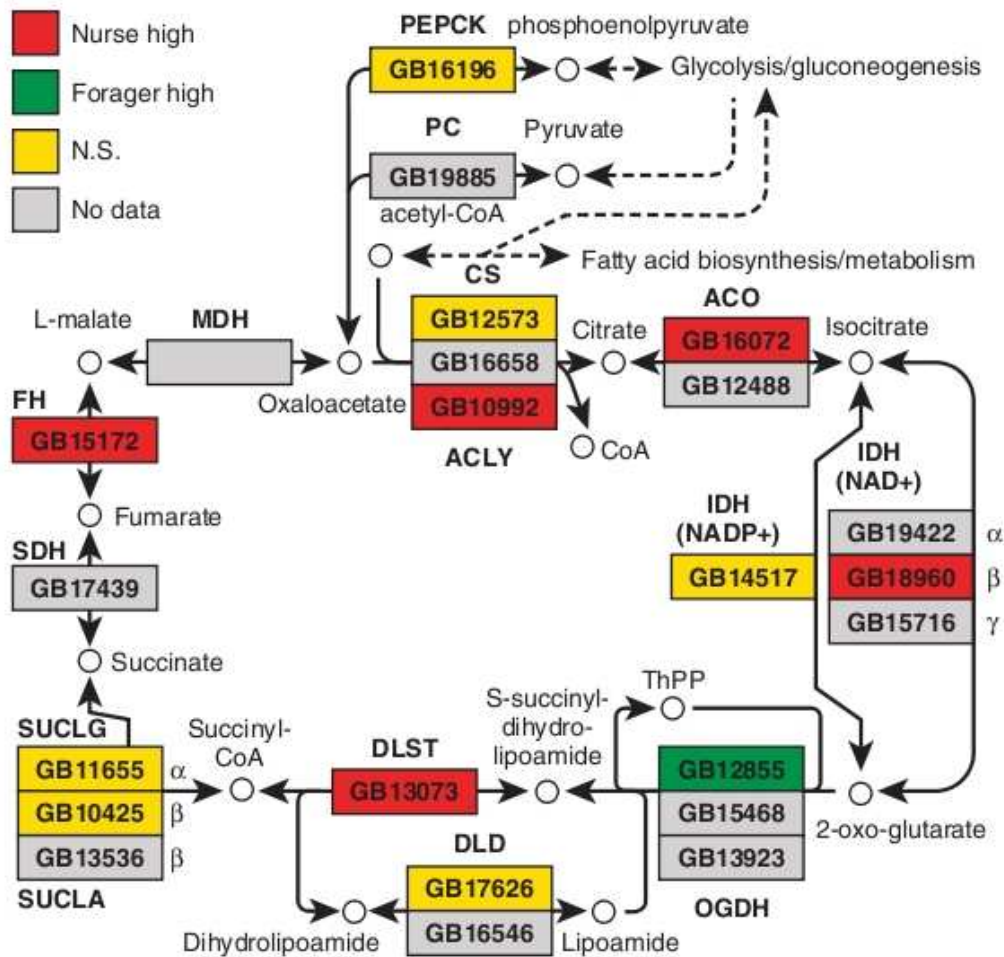
In the future, it will be interesting to compare the structures of GRNs underlying development to GRNs underlying behavioral plasticity. One might speculate that GRNs underlying plasticity are different in structure, owing to the need for greater flexibility. Body parts (such as an eye or a heart) are constructed once during development and must form in the same way regardless of conditions; perhaps for this reason, the GRNs that specify these body parts are extremely dense - i.e., containing many interconnected TFs - a wiring pattern that is thought to be very stable (Davidson, 2006). Perhaps there are design principles for behavioral GRNs that allow animals to rapidly transition back and forth between regulatory states during processes of phenotypic plasticity. Comparing behaviorally related GRNs in the honey bee and in solitary species could also shed light on how the social environment regulates the behaviors of individuals. The answers will

require many experiments over many years, but now is the time to begin framing the questions.

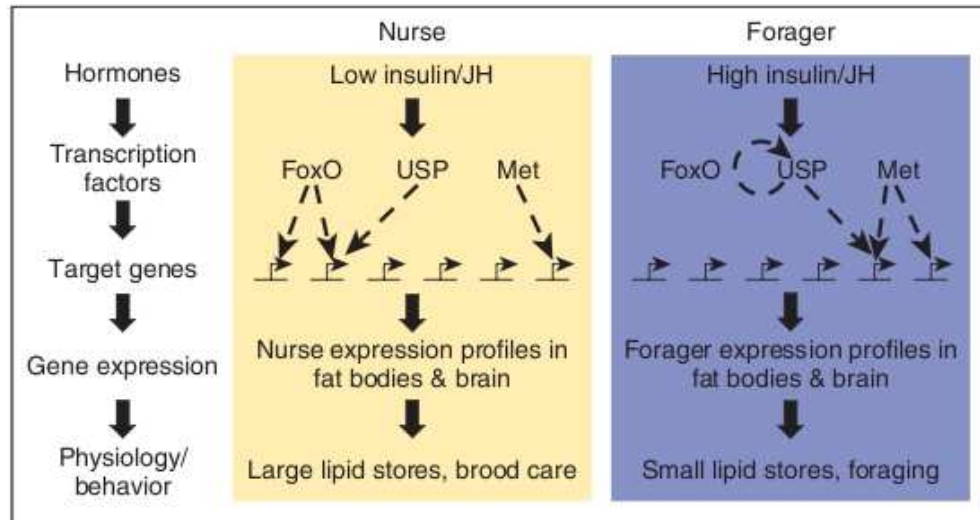
## **Acknowledgments**

We acknowledge grant support from NIH DK082605 (GER, PI), and Craig Mizzen, Lisa Stubbs, Sheng Zhong, members of GER's laboratory, and two anonymous reviewers for helpful suggestions that improved this article.

## Figures



**Figure 3.1.** Changes in brain energy metabolism during worker maturation elucidated by pathway analysis of data from brain microarray experiments. Gene expression data from nurse and forager brains were mapped to energy metabolism pathways compiled by the Kyoto Encyclopedia of Genes and Genomes (Kanehisa & Goto, 2000). Genes in the tricarboxylic acid pathway (shown) and other energy metabolism pathways were predominantly upregulated in nurse brains relative to foragers (Ament, Corona, Pollock, & Robinson, 2008). Solid lines indicate the predicted enzymatic reactions catalyzed by the product of each gene, listed by its identifier in the Official Gene Set 2 from the honey bee genome sequencing project (Honey Bee Genome Sequencing Consortium, 2006). Dotted lines indicate indirect links to other metabolic pathways. ACLY: ATP citrate (pro-S)-lyase; ACO: aconitate hydratase; CS: citrate synthase; DLD: dihydrolipoamide dehydrogenase; DLST: dihydrolipoamide succinyltransferase; FH: fumarate hydratase; IDH: isocitrate dehydrogenase; MDH: malate dehydrogenase; OGDH: 2-oxoglutarate dehydrogenase E1 component; PC: pyruvate carboxylase; PEPCCK: phosphoenolpyruvate carboxykinase; SDH: succinate dehydrogenase; SUCLA: succinate-CoA ligase (ADP-forming); SUCLG: succinate-CoA ligase (GDP-forming).



**Figure 3.2.** Theoretical model for the role of gene regulatory networks (GRNs) in the regulation of worker division of labor. Signaling through insulin-like peptides and juvenile hormone (JH) is low in nurses and higher in foragers (Ament et al., 2008; G. E Robinson, 1987). These hormones regulate gene expression through interactions with transcription factors (TFs), some of which have already been identified in other insect species. Known transcriptional regulators include *FoxO*, which is involved in insulin action (Puig, Marr, Ruhf, & Tjian, 2003), as well as *ultraspiracle* (*usp*) and *methoprene-tolerant* (*met*), both of which are associated with JH (Jones & Sharp, 1997; Konopova & Jindra, 2007; Li, Zhang, Robinson, & Palli, 2007). Increased insulin signaling in foragers is likely to repress *FoxO* target genes by preventing FoxO protein from binding to their promoters (Puig et al., 2003). Increased JH signaling causes increased *usp* expression in honey bees (Jones & Sharp, 1997), as well as other hypothetical changes in target gene activation by USP and MET. According to this framework, interactions among these and other TFs lead to the distinct gene expression profiles of nurses and foragers in brain (Whitfield et al., 2006; Whitfield et al., 2003) and fat bodies (Ament and Robinson, unpublished data). These hormonally controlled GRNs are hypothesized to be causal for behavioral maturation and stable lipid loss.

## Chapter 4

# The Genomic Distribution of USP Suggests Possible Molecular Mechanisms for its Involvement in Honey Bee Behavioral Maturation

### Abstract

Honey bee behavioral maturation provides a system to understand phenotypic plasticity on the molecular level. In this chapter, regulatory mechanisms of Ultraspiracle (USP), a nuclear receptor implicated in JH signaling (Barchuk, Maleszka, & Simoes, 2004; Jones & Sharp, 1997; Velarde, Robinson & Fahrbach, 2009), are explored in the context of the Gene Regulatory Networks (GRNs) involved in behavioral maturation. We first demonstrated that *usp* knockdown in worker bees' fat bodies delays the onset of foraging, suggesting that USP plays an important role in fat bodies during behavioral maturation. Genome-wide USP binding sites were then mapped using fat body cells from nurses and foragers respectively. Few binding differences were detected between nurses and foragers, suggesting UPS regulates its targets by alternating its regulatory activities depending on the availability of cofactors and/or ligands, rather than by differential binding to its target sites in DNA. Enrichment of *Drosophila* USP binding motifs in *Apis* USP binding regions, and significant overlaps of putative target genes between *Drosophila* and *Apis*, indicate a conserved mechanism of USP in *Apis* and *Drosophila*. Moreover, analyses of putative USP target genes raised intriguing hypotheses of how USP integrates different signaling pathways in honey bee behavioral maturation.

## Introduction

Phenotypic plasticity has fascinated generations of biologists. Phenotypic plasticity is the ability of a single genotype to produce multiple stable phenotypes, combinations of coordinated traits in morphology, physiology and behavior (Scheiner, 1993; West-Eberhard, 1989). Even though the significance of phenotypic plasticity is readily apparent, the molecular mechanisms involved are still largely unknown. With well-characterized phenotypes and a sequenced genome (Honey Bee Genome Sequencing Consortium, 2006), behavioral maturation in adult worker honey bees provides a useful system to understand molecular mechanisms of phenotypic plasticity.

Adult honey bee workers undergo an age-related transition from working inside hives to foraging outside, a process known as behavioral maturation. For the first 2 weeks in their adult life, honey bee workers mainly perform inside-hive tasks such as brood feeding (“nursing”); at about 3 weeks of age, workers shift to foraging for the remainder of their lives (Winston, 1991).

Like other examples of phenotypic plasticity, behavioral maturation involves a suite of coordinated changes in behavior and physiology. One of the most prominent physiological differences between nurses and foragers are their content of abdominal lipids, stored primarily in the fat body, a homologue of vertebrate liver and adipose tissues combined (Arrese & Soulages, 2010; Chapman, 1998). Fat bodies primarily consist of two cell types: trophocytes and oenocytes. It has been demonstrated that in *Drosophila* larvae, oenocytes take up lipid during starvation, and oenocytes are required for depleting lipid from fat bodies. The functional relationship between trophocytes and oenocytes in adult honey bees are not still well-understood. (Gutierrez, Wiggins,

Fielding, & Gould, 2007). Prior to the onset of foraging, worker bees lose as much as half of their abdominal lipid stores (Toth & Robinson, 2005). Histological analysis revealed that lipid granules are remarkably less in forager fat bodies, but no dramatic changes of cell morphology and cell numbers were reported (Toth & Robinson, 2005).

The decrease of lipid stores does not only accommodate metabolic requirements for different tasks. Young bees treated with lipid synthesis inhibitor have smaller lipid stores and forage precociously (Toth, Kantarovich, Meisel, & Robinson, 2005), suggesting abdominal lipid stores reflect additional changes in fat bodies that play an active role in regulating behavioral maturation. It is possible that fat bodies send regulatory signals to the brain through neural and hormonal pathways, as has been demonstrated in vertebrates (Leopold & Perrimon, 2007). The apparent active role of fat bodies in behavioral maturation makes it an attractive candidate for further investigation. Moreover, unlike brains, which contain an array of diverse cell types assembled into intricate structures, fat bodies, as mentioned above, consist mainly of just 2 cell types: trophocytes and oenocytes, providing an easier characterization on the molecular level. For these reasons, in this chapter, we focus on fat bodies in investigating behavioral maturation.

To elucidate molecular mechanisms of behavioral maturation, we use systems biology approaches to study transcriptional regulation in fat bodies during behavioral maturation. The following 3 features of behavioral maturation justified our focus on transcription level and the use of systems biology methodology:

First, behavioral maturation is a regulated process. The timing of behavioral maturation is socially regulated: bees can speed up, slow down or even reverse the process of maturation in response to colony needs (Leoncini et al., 2004). For instance, food deprivation or the lack of a foraging force can induce precocious foraging in workers as young as 5 days old (Schulz, Huang, & Robinson, 1998). The regulatory effects of colony needs are mediated by direct communication with pheromones and several internal factors such as hormone levels and individual nutritional status.

One of the most intensively studied regulators is juvenile hormone (JH). Foragers have significantly increased JH titers compared to nurses ( Robinson, Page, Strambi, & Strambi, 1989; Rutz, Gerig, Wille, & Lüscher, 1976). Additionally, worker bees treated with JH or methoprene, a JH analog, forage precociously (Jaycox, 1976; Robinson, 1985, 1987). Another factor, Vitellogenin (Vg), is closely related to JH functions. Vg is a precursor to yolk protein, and is secreted by fat bodies into hemolymph. Besides functioning as a storage protein, Vg has widespread endocrine effects, and thus is also considered to be a hormone-like factor. Vg titers decrease sharply before the onset of foraging (Amdam, Norberg, Fondrk, & Page, 2004; Rutz & Lüscher, 1974), and decreased Vg titers trigger early onset foraging (Marco, Guidugli-Lazzarini, do Nascimento, Simões, & Hartfelder, 2008; Nelson, Ihle, Fondrk, Page, & Amdam, 2007). Topical application of methoprene decreases Vitellogenin while vg knockdown increases JH. It has been proposed that their reciprocal inhibition underlies the clear-cut transition to foraging, and maintains foraging status ( Amdam & Omholt, 2003; Page & Amdam, 2007).

Besides hormonal factors, colony food deprivation causes precocious foraging through limiting individual food intakes, suggesting nutrition itself is also an important regulator in behavioral maturation. As mentioned previously, abdominal lipid levels tightly correlate with behavioral states. Furthermore, several evolutionarily conserved, nutritionally related pathways are implicated in regulating the process, for example, insulin pathways (Ament, Corona, Pollock, & Robinson, 2008) and cGMP dependent kinase pathways (Ben-Shahar, Robichon, Sokolowski, & Robinson, 2002). It has been demonstrated that manipulations of both pathways respectively can causally affect the onset age of foraging.

Second, behavioral maturation is regulated on the transcriptional level. Results from DNA microarray analysis showed that manifold changes in gene expression are associated with behavioral maturation. Up to 40% of the genes expressed in honey bee brains are differentially expressed between nurses and foragers (Alaux et al., 2009; Whitfield, Cziko, & Robinson, 2003). Furthermore, manipulations of signals involved in behavioral maturation affect brain gene expression patterns in a way that reflects their corresponding behavioral effect. For example, as mentioned above, methoprene treatment causes precocious foraging. Meanwhile the treatment also shifts brain gene expression so that it becomes more forager-like (Whitfield et al., 2006). This result indicates that JH modulates behavioral maturation at the level of transcription in honey bee brains.

Like brain, differences in fat body gene expression are associated with behavioral maturation. 2641 transcripts, representing 21% of the genes expressed in fat bodies, are differentially expressed between nurses and foragers (Ament, Robinson et al., unpublished). Notably, many genes involved in lipid metabolism were up-regulated in

nurses. In contrast, many genes related to carbohydrate metabolism were up-regulated in foragers. This result makes sense since it reflects different metabolic requirements of brood-care and foraging. More importantly, the results indicate that the metabolic shift between nurses and foragers is achieved and/or maintained at least in part by transcriptional regulation of important metabolic enzymes.

Besides being correlated with different behavioral status, fat body gene expression is also regulated by the above-mentioned regulators of behavioral maturation (Ament, Robinson et al., unpublished), for instance, nutrition. Bees fed with a rich diet have increased expression of genes involved in lipid metabolism, which is similar to nurses. Bees fed with a poor diet have a higher gene expression of glycolytic genes, as foragers do (Ament, Robinson et al., unpublished).

However, nutrition alone cannot explain the gene expression differences between nurses and foragers, suggesting transcriptional regulation by other factors. Indeed, manipulations of other regulators, such as methoprene treatment and *vg* knockdown, also affect fat body gene expression (Ament, Robinson et al., unpublished). Comparisons of these gene expression patterns suggest overlapping yet distinct mechanisms by which different factors affect behavioral maturation.

Third, as mentioned above, multiple regulators influence behavioral maturation on the transcription level, however, it is still unclear how these signals integrate. Pharmacological treatment and gene expression data revealed interactions among the 4 regulators mentioned above: JH, Vg, nutrition and insulin. From what we have learned so far, it seems that JH is at a central position, interacting with all other regulators: First, as mentioned previously, methoprene treatment decreases Vg titers, while *vg* knockdown

increases JH (Amdam & Omholt, 2003; Page & Amdam, 2007). Second, JH signaling is implicated in nutrition, specifically lipid metabolism. Gene expression data have shown that JH epoxide hydrolase (Jhe), a JH degrading enzyme, is up-regulated across several lipogenic conditions, suggesting a strong association between low JH signaling and large lipid stores (Ament, Robinson et al, unpublished). Third, JH has been shown to be regulated by insulin signaling in fruit flies (Wu & Brown, 2005), and is implicated in regulating insulin production in honey bees (Corona et al., 2007). Even though much is known about interactions between individual pathways, the essential question still remains: how do the aforementioned signaling pathways work together as a coherent system to regulate behavioral maturation? Specifically, if JH is a central regulator, what are the molecular mechanisms for JH to integrate other signals on the transcription level?

To answer this question, a gene regulatory network (GRN) framework is proposed (Chapter 3). GRN models were originally used in the context of developmental biology (Davidson, 2006). Developmental processes consist of a progression of regulatory states, each of which are defined by a distinct set of active transcription factors. By analogy, we hypothesized that different behavioral and physiological statuses in behavioral maturation are associated with distinct regulatory states. Distinct regulatory states can be embodied by different sets of active transcription factors, or by transcription factors with alternating activities. In this chapter, we focus on the latter situation. We explored how a transcription factor changes its activity by differentially regulating its downstream targets during behavioral maturation. As a first step towards describing the GRN underlying behavioral maturation, the function of a key transcription factor Ultraspiracle (USP) in behavioral maturation was investigated and its genomic binding sites were mapped using

chromatin immunoprecipitation followed by whole genome tiling array analysis (ChIP-chip).

USP was chosen because of its potential role as a signaling integrator. USP forms heterodimers with ecdysone receptor (EcR). Together, USP and EcR work as a functional ecdysone receptor (Yao et al., 1993). USP is also implicated in JH signaling (Barchuk et al., 2004; Jones & Sharp, 1997; Velarde et al., 2009). Although it is still controversial whether USP functions as a JH receptor, USP seems to play a role in transmitting JH signals: treating honey bee pupa with methoprene, a JH homologue, induced USP expression (Barchuk et al., 2004). In addition, the vertebrate USP ortholog is Retinoic X Receptor (RXR) (King-Jones & Thummel, 2005; Velarde, Robinson, & Fahrbach, 2006), which has been shown to be a master regulator in lipid metabolism (Chawla, Repa, Evans, & Mangelsdorf, 2001). RXR is able to form heterodimers with several nuclear receptors, that are involved in different aspects of lipid metabolism (Mangelsdorf, Evans et al., 1995). Notably, heterodimer of PPAR and RXR influence the storage of fatty acids in adipose tissues (Chawla et al., 2001; Sonoda, Pei, & Evans, 2008). A PPAR homologue has not been identified in honey bees (King-Jones & Thummel, 2005; Velarde et al., 2006), so it is not straightforward to deduce if this function of RXR is conserved with USP. It is possible that USP performs a similar function with a different heterodimer partner. It has been shown that USP forms heterodimers with nuclear receptors other than EcR (Zhu, Miura, Chen, & Raikhel, 2000; Zhu, Miura, Chen, & Raikhel, 2003), suggesting multiple heterodimer partners of USP in honey bees, presumably regulating different pathways. However, the function of USP is not well understood beyond its role in ecdysone signaling.

In this chapter, USP is demonstrated to be another regulator in honey bee behavioral maturation. To elucidate its regulatory mechanisms under the GRN framework, genome-wide binding sites of USP and putative USP target genes were identified, which serves as a first step to understand the complexity of molecular mechanisms regulating honey bee behavioral maturation.

## **Materials and Methods**

### *Honey Bee Collections*

Honey bees (*Apis mellifera*) used in the experiments were reared in the Bee Research Facility at the University of Illinois at Urbana-Champaign, and maintained according to standard apicultural procedures. Honey bees were collected from colonies headed by queens artificially inseminated with sperm from a single drone, which reduces genetic variation within the colony.

For behavioral collections, nurses and foragers were identified and collected according to specific standards (Huang, Robinson, & Borst, 1994). Nurses were collected after they repeatedly lowered their heads into honeycomb cells containing larvae, and foragers were collected as they returned to the hive with visible loads of pollen on their legs. Foragers and nurses from 3 different colonies were collected for ChIP-chip.

To obtain one-day-old bees, sealed combs of brood-containing pupae that would enclose within 1 day were removed from colonies and placed in an incubator at 34° C. Newly emerged bees were collected from the incubator within 24 hrs.

### *RNA Interference*

A protocol of RNA interference was adapted from Amdam et al., (Amdam et al., 2003). dsRNAs were provided by Beeologics LLC (Miami, FL) (Maori et al., 2009). Briefly, DNA fragments were cloned into plasmid pDrive with 5' T7-promoters on both DNA strands using sequence specific primers (see Table 4.5) (Barchuk, Figueiredo, & Simoes, 2008). Then the sequences of interest were excised from the plasmids, gel-purified and used as templates for in vitro T7 transcription. The resulting RNA was treated with DNase followed by phenol extraction and ethanol precipitation. The dsRNAs were purified desalted using RNA purification columns to remove residual chemicals, and dissolved in nuclease-free water. dsRNAs were derived from the sequences of *usp* and *gfp* as a negative control.

dsRNA injections were performed as Amdam et al. and Velarde et al. (Velarde et al., 2009). 1 ul of 20ug/ul dsRNA was injected between the second and third abdominal segments of one-day-old bees. The needle of the injector punctured abdomens in a very shallow angle so that no harm would be done to the digestive tracks. After injection, the needle was not immediately pulled out to prevent hemolymph leaking. Bees with serious bleeding after injection were discarded immediately after injection. The treated bees were checked everyday to keep a record of mortality.

The bees used were marked with a dot of paint (Testor's PLA) on their thorax to differentiate different treatment groups. A group of around 50 bees (25 injected with *dsusp* and 25 injected with *dsgfp*) were placed into 10 x 10 x 7 cm Plexiglas cages, and kept in a 34 ° C incubator in constant dark for 3 days. The cages were provided with *ad*

*libitum* sugar syrup (50% sucrose/water w/v) and pollen/honey paste (45% pollen, 45% honey, and 10% water, w/w/v), which were supplied by two feeding tubes.

#### *Tissue Preparation for Gene Expression Analysis*

To evaluate *usp* knockdown in honey bee heads and abdominal fat bodies, bees were flash frozen immediately after collection. Frozen heads were homogenized directly in tissue lysis buffer. Whole abdomens were thawed in RNAlater-ICE. Digestive tracks and other abdominal tissues were removed, and fat bodies were left attached to abdominal cuticle. The prepared tissues were used directly for total RNA extraction or stored in -80 ° C.

For total RNA extraction, a commercial kit (Rneasy, Qiagen, Valencia, CA) was used. Fat bodies with cuticle were homogenized in the tissue lysis buffer from the kit. During homogenization, fat cells were dislodged in the lysis buffer, and pieces of cuticle were removed by centrifugation.

#### *Quantitative, real-time RT-PCR*

cDNA was synthesized from 200ng of total RNA with ArrayScript reverse transcriptase (Ambion, Applied Biosystems, Carlsbad, CA) and random hexamers as primers. Quantitative PCR was performed in triplicates on 384-well plates with Sybr Green master mix (Roche) and specific primers. ABI Prism7900 was used to conduct the PCR cycles and record the fluorescence. The results of experimental genes were log transformed and normalized to spike-in control gene *rcp* using a standard curve method.

### *Behavioral Analyses of Age at First Foraging*

Methods are essentially the same as Ament et al. (Ament et al., 2008), with slight modifications. On the 4th day of treatment in cages, all the treated bees were introduced into single-cohort host colonies. The treated bees were marked on the thorax with a small dot of enamel paint with color codes corresponding to different treatment groups. The single-cohort colonies were prepared the day before the introduction, and were made with around 1000 one-day-old bees and a queen. In each trial, 4 cages with about 200 treated bees were placed into the host colony.

After introduction, the colonies were observed for 5-7 days. Returning foragers were identified by their extended abdomens or the presence of pollen loads on their legs. All of the foragers observed for the first time were captured briefly with a soft forceps, and were marked by an additional dot of paint on the abdomen so that they would not be counted a second time. A record was kept on when and which group of bees was first observed foraging. At the end of the observation, colonies were killed in very early mornings by flash-freezing with liquid nitrogen. Marked bees were counted.

### *Antibodies*

The antibody for honey bee USP used in ChIP-chip was obtained from K. White at the University of Chicago. The antibody was produced in rabbits against a 109 amino acid peptide. The peptide matches the region amino acid 75 to 183 of honey bee USP, including a DNA binding domain. Antisera were purified using protein A affinity purification. The antibody was tested using western blots. A single band at the predicted molecular weight was observed (Figure 4.1).

### *Fat Body Dissection and Preparation for Chip*

Fat bodies were dissected following the method described by Kaatz et al. (Kaatz, Hagedorn, & Engels, 1985). Briefly, the abdomen was separated from thorax. And the last two segments of the abdomen were removed. The sting apparatus and digestive tracks were pulled away. The remaining abdominal cuticle was cut laterally into ventral and dorsal halves, and the former was discarded. Since fat body cells cannot be dissected without being destroyed, the whole dorsal abdominal cuticle was used in the following steps.

Fat bodies, still attached to the dorsal half of cuticle, were fixed with 1% formaldehyde in PBS for 10 minutes at room temperature. The reaction was stopped by adding 100ul 1.25M Glycine stock and incubated at room temperature for 10 minutes. Then the cuticle with fat bodies was rinsed twice in cold PBS. Fat bodies were collected from the cuticle with fine forceps and were placed in 1ml cell lysis buffer ( 50mM Tris Hcl pH 8.1, 2mM EDTA pH 8.0, 0.5% NP-40, 10% glycerol) on ice. 4 individuals were pooled for each ChIP reaction. Fat bodies were homogenized on ice using a Downs homogenizer until no big clusters of cells were visible. The samples were then centrifuged at 2000rpm for 10 minutes at 4 ° Celsius, and supernatant was removed. The collected nuclei were flash-frozen in liquid nitrogen and stored in -80 ° C.

### *ChIP-Chip*

Standard ChIP reactions were performed using buffers from a commercial kit EZ-ChIP (Millipore, Billerica, MA) following the protocol with slight modifications according to other protocols (Lee, Johnstone, & Young, 2006; Weinmann, Bartley, Zhang,

Zhang, & Farnham, 2001). (See Appendix B for a detailed protocol.) Fresh or stored nuclei (from the previous session) were suspended in SDS lysis buffer. The pellet was fully disrupted by pipetting. The suspended nuclei were then sonicated in a 4 ° C water bath for 5 minutes; with a 30 second break for every 30 second sonication. We used Bioruptor (Diagenode, Liège - Belgium) on a high power setting. The chromatin solution was centrifuged and diluted with ChIP dilution buffer. 10ug of USP antibody was added to the solution and incubated at 4° Celsius overnight. The next morning, 100ul magnetic beads Protein A/G pre-blocked with BSA and sheared salmon sperm DNA were added to each reaction. It was followed by incubation and washes with low salt/high salt/LiCl buffer once and twice with TE. After DNA-protein complexes were eluded from the beads, the DNA was released by incubated at 65° C for 4-5 hours in the presence of NaCl. Precipitated DNA was treated with RNaseA and protease K followed by phenol/chloroform exaction and ethanol precipitation, and dissolved in nuclease-free water. Amplifications were performed using a Genome Amplification kit following a protocol from P. Farnham lab.

Custom whole genome tiling arrays were produced by NimbleGen Systems, Inc. (Madison, WI). The probe sequences were derived from the honey bee genome Assembly 4.0 masked from repetitive sequences. The probes were 50 base pairs long with 100 base pair resolution. All NimbleGen arrays were hybridized, and the data were extracted according to standard operating procedures by NimbleGen Systems Inc. Signal Map Software provided by NimbleGen Systems, Inc. was used to visualize the results.

### *ChIP-chip Data Analysis*

To identify USP binding sites we used two peak-calling algorithms Mpeak described in Kim et al. (Kim et al., 2005) and Tamalpais (Bieda, Xu, Singer, Green, & Farnham, 2006). We included the union of peaks found in both methods, since each utilized different statistical models that seemed to identify peaks with different shapes. The Mpeak algorithm assumes DNA fragments ends are exponentially distributed relative to true binding sites. Thus, the signal at a certain position that is  $c$  base pairs away from the true binding sites can be approximated by the probability that cuts occur farther than  $c$  base pairs. Since we monitor the logarithm of the hybridization signals, the signal intensity for a certain site should decrease linearly with distance from the true binding sites. Using this rationale, a double regression model was adopted in Mpeak to fit neighboring log ratio signals to asymmetric triangles centered on the candidate binding sites. Tamalpais uses a different statistical model. It does not depend on any assumptions about shape and amplitude of peaks representing true binding sites. Instead it uses a threshold for the log<sub>2</sub> ratio of individual probes, and a cut-off for the “run length”, or the width for a valid peak. We set the threshold at the 98th percentile of the signal intensity observed for all of the probes on one array. The cut-off value for peak width was determined indirectly by a p-value calculated using Erdos-Renyi Law. We used  $p < 0.05$ .

To compare the peak region binding strength between nurses and foragers, the raw log<sub>2</sub>(ChIP/ genomic DAN) value of the probes within the peak regions from the 6 arrays (N1, N2, N3, F1, F2, F3) were extracted. To avoid bias in the future comparison of nurse and forager arrays, these values were normalized across the 6 arrays using quantile normalization. Each peak region contains multiple probes with 50 base pairs in-between.

The ratio values of these probes within the peak regions are dependent. We use the pairwise t-test to test whether nurse and forager have the same mean binding intensity. Specifically, for each probe the difference  $D = (N1 + N2 + N3)/3 - (F1 + F2 + 3)/3$  was calculated, and pairwise t-test was used to test the non-hypothesis  $D=0$ . The t-tests were conducted for each of all the peak regions. To adjust for the multiple tests, FDR for each test was calculated using pvlaue method. The FDR cutoff of 0.05 was used to select peak regions with different binding strength.

Additional ChIP-chip analyses were coded in Perl. Motif analyses were performed with the software SWAN developed by Saurabh Sinha et al. (Kim et al., 2010).

## Results

### *USP Knockdown in Fat Bodies Affects Honey Bee Behavioral Maturation*

USP has been implicated in JH signaling in honey bees by several lines of evidence (Barchuk et al., 2004; G. Jones & Sharp, 1997; Velarde et al., 2009). The importance of JH in behavioral maturation raises the possibility that USP is also involved in regulating the timing of onset of foraging. We tested whether USP influences honey bee behavioral maturation by determining the effect of *usp* knockdown in fat bodies on the onset age of foraging. Since JH positively regulates USP, and USP has been suggested to be downstream in JH signaling, we hypothesized that knockdown of USP would delays the onset of foraging.

We first tested the effect of *dsusp* injection on *usp* expression in abdominal fat bodies and heads. It was confirmed that *usp* mRNA was significantly decreased in abdominal fat bodies on the third day after *dsusp* injection, while its expression in heads

(including brain, fat bodies in head capsule, and brain associated glands) was not affected (Figure 4.2). Since USP is not highly expressed in brain associated glands, the expression levels of heads roughly reflect USP expression in brains. Therefore, the behavioral effects of *usp* knockdown should be caused mainly by its reduced expression in peripheral tissues.

Knockdown of *usp* in fat bodies delayed the onset age of foraging with some colony variation. In a combined statistical analysis of 9 trials with 5 source colonies, bees injected with *dsusp* delayed their onset of foraging significantly compared to bees injected with *dsgfp* control (Cox regression,  $p < 0.05$ , Figure 4.3). There was also a significant interaction between the effects of dsRNA knockdown and source colonies, caused mainly by one source colony that showed a reversed behavioral effect of *dsusp* knockdown (Figure 4.3).

The reversed behavioral effect was not caused by weather or the host colony. In the final behavioral trial, dsRNA treated bees from 3 different source colonies were introduced into a common host colony at the same time and observed for the same period. The trend of reversed behavioral effects persisted only for bees from the above-mentioned source colony, even though the result was not statistically significant. It is possible that genetic differences as well as the nutritional status of young bees from different source colonies caused the differential behavioral response to *usp* knockdown.

### *Genome-wide Comparisons of USP Binding Sites in Nurses and Foragers Showed Few Differential Bindings*

As a first step to understand how USP paces honey bee behavioral maturation, genomic binding locations of USP were mapped in fat bodies of nurses and foragers using ChIP-chip. *usp* is not differentially expressed in fat bodies of nurses and foragers (Ament et al. Mechanisms of stable lipid loss in a social insect 2010, submitted to *PLoS Biology*), however, it can still differentially regulate its target genes in nurses and foragers. We hypothesized that differential USP binding to its binding sites may directly determine the differences in gene expression between the two behavioral states. If this hypothesis is true, then we should detect two overlapping but distinct sets of USP binding regions from nurses and foragers. Alternatively, USP may regulate gene expression by repressing or activating its target genes depending on the availability of ligands or cofactors and chromatin environments in different contexts without alternating binding sites occupancy. If the latter hypothesis is true, then there will be no difference in USP binding regions between nurses and foragers.

ChIP-chip results of 3 biological replicates of nurses and foragers revealed few difference in USP binding regions in the fat bodies of nurses and foragers, respectively. USP binding regions were identified using peak extraction algorithms Tamalpais (Bieda et al., 2006) with results of each individual trial. Pairwise comparisons of the peak regions showed high correlations in all possible comparisons from the 6 trials (Figure 4.4). The correlation coefficients range from 0.73 to 0.92, with the highest correlation coefficients between the second nurse trial and second forager trial.

To further detect whether there are peak regions with different binding strength between nurses and foragers, pairwise t-tests were performed to compare the signal intensity of peak regions. With a FDR cutoff of 0.05, only 56 regions were found to have different binding intensities between nurses and foragers. A total of 38 putative USP target genes (see Genomic Locations of USP Binding Sites for identification of putative USP target genes) are associated with these regions (Table 4.6). Among these putative USP targets, 18 have *Drosophila* orthologs, of which 6 are development related genes.

High correlations of peak regions between different trials indicate the ChIP-chip results are reproducible. The rarity of distinctive binding regions between nurse and forager trials suggests USP regulates gene expression primarily by alternating its regulatory mode rather than by differentially binding to its cognate sites, which is consistent with the regulating mechanisms of several nuclear receptors described in vertebrates (Mangelsdorf et al., 1995).

#### *Drosophila USP Motifs are Significantly Enriched in Apis USP Binding Sites*

1360 genomic regions were identified as USP binding regions using the union of two peak extraction results from Tamalpais (Bieda et al., 2006) and MPeak (Kim et al., 2005) across both nurse and forager samples. The number of USP binding sites is comparable to the number of *Drosophila* USP binding sites in Kc167 cells (1450 USP binding sites). Since it has been demonstrated that nuclear receptors have conserve core binding motifs from fruit flies to vertebrates, it is highly likely that honey bee USP recognizes a motif similar to the *Drosophila* USP (CF1) motif (Table 4.1). A scan using the motif searching software SWAN (Kim et al., 2010) revealed that 53% (716 out of

1360) of the USP binding regions contain at least one CF1 motif. A comparable proportion was reported for *Drosophila* USP binding regions identified by a similar methodology. USP binding regions without CF1 motifs might contain degenerate USP binding motifs that are elusive to motif searching algorithms, or USP might be able to bind DNA in an indirect manner, or they might be false positive results.

Even though a CF1 motif was not detected in all USP binding regions, it is highly enriched in the USP binding regions (hypergeometric test,  $p < 10E-6$ ). In fact, it is the most highly enriched *Drosophila* motif in all of the tested motifs from TRANSFAC (Matys et al., 2003). The significant enrichment of CF1 motif provides partial validation of the quality of the USP ChIP-chip results, and suggests honey bee USP recognizes an evolutionarily conserved nuclear receptor binding site.

The distribution of CF1 motifs associated with USP putative target genes (see Genomic Locations of USP Binding Sites for the identification of USP putative target genes) is shown in Figure 4.7. Of the 356 genes with at least one CF1 motif in the associated USP binding region(s), most of the genes have less than 3 CF1 motifs. There are 14 genes with 10 or more CF1 motifs. Gene *GB17945*, an ortholog of *Drosophila* transcription factor *cut*, has 29 CF motifs. A closer inspection revealed that exons of this gene spreads over a huge genomic area which includes 13 USP binding regions. One of these binding regions contain as many as 9 CF1 motifs.

### *Other Transcription Factor Binding Motifs Co-localize With CF1 in USP Binding Regions*

Besides CF1 motifs, there are 4 other *Drosophila* transcription factor binding motifs enriched in honey bee UPS binding regions (Table 4.1), including binding motifs of ADF1 (SWAN, p-value=2E-6) and Ftz-f1 (SWAN, p-value=2E-5). Ftz-f1 is an important regulator in development, closely related to the ecdysone signaling pathway (Sullivan & Thummel, 2003) and suggested to be related to JH signaling (Zhu, Chen, & Raikhel, 2003). Enrichment of Ftz-f1 binding motifs in honey bee USP binding regions might reflect these connections.

ADF1 is a member of myb-related family of transcription factors. *Drosophila* ADF1 mutants have defective long term memory, consistent with suggestions that it regulates synaptic growth and is involved in neuroplasticity (DeZazzo et al., 2000). Interestingly, the ADF1 binding motif was shown to be enriched in another behaviorally related gene set in honey bees. The ADF1 binding motif was enriched in genes up-regulated by methoprene in honey bee brains (Sinha, Ling, Whitfield, Zhai, & Robinson, 2006), suggesting that the ADF1 motif is involved in JH signaling. Given that USP is also implicated in JH signaling, enrichment of ADF1 motifs in USP binding regions suggest the possibility of combinatorial regulation by USP and ADF1 in JH signaling.

In *Drosophila*, large-scale profiling of the binding sites of a set of transcription factors has revealed clusters of transcription factor binding sites, named transcription factor binding hotspots (Moorman et al., 2006). We tested if the 4 enriched binding motifs form similar clusters. Only 8% of all the UPS binding regions that have at least one of the 4 motifs have 2 or more motifs out of 4, indicating the 4 motifs tend to be

detected in different USP binding regions instead of forming clusters, even though they are all enriched in USP binding regions. The result suggests that these 4 motifs might work with USP in different modes of regulation, similar to the mechanisms of RXR regulation of lipid metabolism (Chawla, Repa, Evans, & Mangelsdorf, 2001; Sonoda, Pei, & Evans, 2008).

### *PPAR Motifs are Enriched in USP Binding Regions, Suggesting New Heterodimer*

#### *Partners of USP*

USP has heterodimer partners besides EcR in other insects, but it is not clear whether this is also true in the honey bee. Interestingly, vertebrate PPAR binding motifs are significantly enriched in honey bee USP binding regions (SWAN,  $p=7E-6$ ) (Table 4.1). PPAR motifs consist of a direct repeat of two half site separated by one nucleotide (DR1) (Mangelsdorf et al., 1995). The PPAR half site is sufficiently similar to CF1 motif that it was recognized as CF1 motif in most cases by SWAN. The structure of the PPAR motif suggests that like its vertebrate homologue, honey bee USP might also form heterodimers with another nuclear receptor.

Since EcR is a known USP partner, it is possible PPAR motifs represent the binding sites of USP/EcR heterodimers. In *Drosophila*, most of the EcR/USP binding motifs are palindromic sequences, but it has been shown that EcR/USP can also function through DR1 motifs in vivo as well (Fang, Xu, Jones, & Jones, 2005). To test whether the identified PPAR motifs represent EcR/USP binding motifs in honey bee, we compared the honey bee genes having PPAR motifs with those known to be bound by EcR/USP in *Drosophila* (Gauhar et al., 2009). Out of 278 genes with PPAR motifs, only

46 genes are bound by EcR/USP in *Drosophila*. The EcR targets are not enriched in genes with PPAR motifs, suggesting there is at least one nuclear receptor other than EcR that forms heterodimers with USP in honey bees.

### *Genomic Locations of USP Binding Sites*

Using the official honey bee gene set (Honey Bee Genome Sequencing Consortium, 2006), we identified genes within 10 kb of USP binding regions as possible targets regulated by USP (Gauhar et al., 2009). Instead of assigning the binding regions to their closest neighboring genes, we included all the genes within the range, reasoning that distal regulatory sites are common and that the genes nearest to binding sites are not necessarily the relevant USP targets. Among the identified USP putative target genes approximately 23% have USP binding regions within 5kb upstream of their translation start sites. About 44% have USP binding sites within the gene body, which is similar to the percentage of PPAR binding sites in vertebrate 3T3-L1 cells (Lefterova et al., 2008; Nielsen et al., 2008). Since the identified peak regions are approximately 500bp to 1000bp in length, it is difficult to pin down the exact locations of peaks relative to their target genes. Using the middle points of the highest scoring probes, we tested whether the peak regions were preferentially located in introns compared to exons. No such location bias was detected. The widespread distribution of USP binding regions in distal extragenic regions and intragenic regions is consistent with the genomic binding patterns described for transcription factors, such as FOXA2 and TCF4 in recent publications (Hatzis et al., 2008; Wederell et al., 2008).

USP binding regions occur upstream, downstream, or within a gene with similar frequencies except for one hot spot. There are more USP binding sites spanning or within 1kb downstream of translational start sites compared to the average occurrence frequencies in other regions (Figure 4.5). The frequencies that USP binding sites are around and downstream of gene translational start sites are 2.5 and 3 times, respectively, greater than the average frequency in other regions. The genomic distribution of USP binding regions relative to its putative target genes resembles that of RXR and PPAR in murine 3T3-L1 cells (Hamza et al., 2009).

#### *Transcription Factors are Overrepresented in USP Putative Target Genes*

One notable feature of the set of putative USP target genes we identified is that it contains many genes encoding transcription factors. In this case, a honey bee gene is considered to be a transcription factor if its *Drosophila* ortholog is a transcription factor. Using this stringent criteria, 64 transcription factors were identified out of 521 genes that have a *Drosophila* ortholog (Table 4.2). Among these transcription factors are *hairy* and *tai*, which are known USP/EcR targets in *Drosophila* (Gauhar et al., 2009). Transcription factors are highly overrepresented in the genes near USP binding sites (hypergeometric test,  $p=2E-15$ ). This result suggests USP might regulate most of its effector genes indirectly through regulating their corresponding transcription factors. It also partially explains how USP functions as a master regulator and how the knockdown of a single gene can have a significant behavioral effect.

### *Functional Analysis of Putative USP Target Genes*

Functional analysis of putative USP target genes was performed using their *Drosophila* orthologs, since the functions of honey bee genes are not well-characterized. Gene orthologs were identified based on similarities of translated protein sequences. 512 genes out of 820 near USP binding sites which have a *Drosophila* ortholog were used in Gene Ontology (GO) analyses. Consistent with the enrichment of transcription factors, the GO term “transcription factor activity” is highly significant (DAVID, p-value=4.8E-21, benjamini correction) (Dennis et al., 2003; Huang, Sherman, & Lempicki, 2009). Another notable result of the GO analyses is that many development-related GO terms are significantly enriched in this gene set, including “instar larval or pupal development” (DAVID, p-value=3.8E-10, benjamini correction), “metamorphosis” (DAVID, p-value=6.2E-10, benjamini correction), and “imaginal disc development” (DAVID, p-value=5.2E-10, benjamini correction). Ecdysone is an important regulator in developmental processes, and JH works as a modulator, most notably in the process of metamorphosis. Ecdysone receptor EcR/USP binds to genes involved in the developmental process, the enriched development-related GO terms may reflect that USP functions as an ecdysone receptor or USP is involved in JH signaling. There might not be a clear distinction, since EcR and USP have been suggested to be the key players through which ecdysone and JH interact (see “USP, a master regulator” below for further discussion).

The function of JH is well-characterized in behavioral maturation, but it is still unclear whether ecdysone signaling is involved. It has been proposed that ecdysone signaling is related to behavioral maturation. It is possible that developmental genes are

're-used' in the process of behavioral maturation, probably through signaling pathways other than ecdysone (see “USP, a master regulator” below for further discussion).

Among putative USP targets, several groups of genes with shared characters and genes in important signaling pathways are of special interest. The first group consists of 7 nuclear receptors: E75B, E78C, Hr4, Hr38, SVP, Hr46 and USP (Table 4.3), out of a total of 19 nuclear receptors that have been discovered in the honey bee genome (Velarde et al., 2006). Inclusion of USP itself as a putative USP target suggests the possibility for feedforward/feedback regulation. Supporting this idea, it has been shown that USP expression was up-regulated in Kenyon cells coincident with a moderate endogenous JH titer increase (Barchuk et al., 2004); in contrast, high dose JH treatment induced a transient down-regulation of USP transcripts (Velarde et al., 2009).

E75B, E78C, Hr4 and Hr46 have been demonstrated to be direct targets of the 20E-EcR-USP complex in *Drosophila* (Beckstead, Lam, & Thummel, 2005; Sullivan & Thummel, 2003; White, Hurban, Watanabe, & Hogness, 1997), and may be similarly regulated in honey bees. Alternatively, like other USP putative targets involved in ecdysone signaling, these nuclear receptors may transmit signals from USP independently from ecdysone signaling. It has been suggested that the other signal is JH (see “ USP, a master regulator” for further discussion).

Both Hr38 and SVP can form heterodimers with USP (Zhu et al., 2000; Zhu et al., 2003), indicating that like its vertebrate homologue RXR, insect USP has different heterodimer partners, even though it is unclear which signaling pathways they regulate. Regulation of putative heterodimer partners may provide an extra layer of regulation that renders a more precise and/or rapid control in signaling cascades.

It is intriguing to see a big group of nuclear receptors among putative USP target genes in honey bees. Similarly, in *Drosophila*, many nuclear receptors are direct targets of EcR/USP in ecdysone signaling (Beckstead, Lam, & Thummel, 2005; Sullivan & Thummel, 2003; White, Hurban, Watanabe, & Hogness, 1997). Although these two sets of nuclear receptors overlap in honey bees and *Drosophila*, there are divergences, which may reflect species differences. It is also possible the different downstream nuclear receptors indicate other functions of USP beyond ecdysone signaling, possibly mediating JH signaling. Like ecdysone signaling in developmental processes, during which many regulated pathways are coordinated, USP might coordinate many pathways during behavioral maturation through direct regulation of nuclear receptors. In any case, the result reveals a general mechanism of nuclear receptor signaling in insects. Instead of regulating parallel pathways, nuclear receptors seem to form regulatory hierarchies among themselves. USP occupies a high position in the nuclear receptor hierarchy, regulating many nuclear receptors.

The second group of genes is involved in lipid and energy metabolism, including lipid biosynthetic and catabolic processes, and an enzyme on the mitochondrial respiratory chain (Table 4.4). Among USP putative targets, lipid metabolism is not significantly enriched as a GO term, suggesting that unlike its vertebrate homologue RXR, USP does not directly regulate most metabolic genes. Instead, it seems that USP directly regulates transcription factors which in turn regulate effector genes. But it is difficult to infer those secondary targets using only the presence of USP binding sites. Interestingly, all 4 of the genes involved in lipid metabolism are down-regulated in fat bodies of bees depleted of Vg using RNAi (Ament, Robinson et al., unpublished),

suggesting a correlation between Vg signaling and these genes. This connection might provide a channel to link USP and Vg signaling.

There are also several important signaling genes among putative USP targets, which generate intriguing hypotheses about how different signaling pathways are coordinated through USP regulation. 3 USP putative targets are of special interest:

First, *insulin receptor 1 (inR1)* has a USP binding site across its translation start site, and the USP binding site contains a CF1 motif around 300bp upstream of the translation start site. Second, one of the putative USP targets is *JH acid methyltransferase (jhamt)*, encoding an enzyme involved in the final steps of JH synthesis. It transfers a methyl group from S-adenosyl-L-methionine (SAM) to farnesoic acid or epoxyfarnesoic acid. There are 3 USP binding sites upstream of the *jhamt* translation start site, each of which contains at least one CF1 motif (Figure 4.6). Third, *foraging (for)*, encoding a gene involved in PKG pathway is also a putative USP target. *For* has been shown to be involved in behavioral maturation (Ben-Shahar et al., 2002). A distal USP binding site was discovered 7kb upstream of this gene, suggesting that it may be regulated by USP. These 3 putative USP target genes raised the possibility that JH signaling integrates or modulates insulin and *for* pathways, and regulates its own synthesis through USP.

#### *Conservation of Putative USP Target Genes*

Conservation of USP target genes was assessed by comparing the putative honey bee USP target genes with the genes in *Drosophila* identified using similar techniques and criteria in a recent publication (Gauhar et al., 2009). Interestingly, there is a large difference in the numbers of genes found near USP binding regions between fruit flies

and honey bees, even though the numbers of USP binding regions are similar in these two species. 243 out of 521 honey bee genes that have a *Drosophila* ortholog are also found near to USP binding sites in *Drosophila*. This substantial overlap is unlikely due to random chance (hypergeometric test,  $p=1.3E-5$ ). This overlap of putative USP target genes between honey bees and fruit flies suggests a functional conservation of USP between the two species.

In *Drosophila*, USP has been demonstrated to heterodimerize with EcR to function as an ecdysone receptor. It is possible that the demonstrated functional conservation between honey bee and fruit fly USP is due solely to their conserved roles in ecdysone signaling. Putative USP target genes in honey bees were compared to putative EcR/USP target genes (bound by both USP and EcR) in *Drosophila*. 91 overlapping genes were identified including well-characterized ecdysone responsive genes *eip74EF*, *eip75B*, *eip78C*, *inR1*, etc. Gene Ontology (GO) analyses (DAVID) showed these 91 genes were significantly enriched in several GO terms related with developmental processes, which is consistent with the role of ecdysone in development. However, the putative EcR/USP target genes are not significantly enriched in the conserved USP putative target genes between honey bees and fruit flies (hypergeometric test,  $p=0.47$ ), indicating the functional conservation between honey bee and fruit fly USP goes beyond USP's function in ecdysone signaling.

Comparison between genes near USP binding sites in honey bees and vertebrates is challenging because of major differences in experimental platforms, the difficulty of identification and confirmation of orthologs, and the different criteria to identify putative target genes. A direct comparison of genes near USP binding sites in honey bees to genes

near RXR/PPAR sites in mouse 3T3-L1 cells from the recent two publications revealed several overlapping genes including the well-characterized RXR targets UPP2 and PCK2 (Wakabayashi et al., 2009). The overlap is not statistically significant, so it is difficult to conclude whether the overlapped genes are due to functional conservation of USP and RXR or caused by random chance.

## Discussion

Honey bee behavioral maturation represents an intriguing case of phenotypic plasticity. The recent sequencing of the honey bee genome and development of genomic tools make it a tractable model to study the molecular mechanisms underlying phenotypic plasticity using systems biology approaches. This chapter presents a primary exploration towards understanding behavioral maturation in terms of describing transcriptional regulation under the framework of gene regulatory networks (GRNs). One piece of information important in constructing GRNs, the genomic binding sites of one nuclear receptor, USP, was obtained. Analysis of putative UPS targets provides possible directional edges in the regulatory networks underlying behavioral maturation.

We first showed that USP is involved in regulating behavioral maturation. *usp* knockdown in worker bee fat bodies delays the onset of foraging. The regulatory mechanisms of USP were then explored. USP might differentially regulate gene expression through two possible mechanisms. By binding to different genomic regions, USP regulates different sets of genes in nurses and foragers. Alternatively, the regulatory activities depend on the availability of its ligands and cofactors.

Few differences in the genomic distribution of USP were observed between nurses and foragers. Among around 1300 USP binding regions, only 56 regions were detected

with significant binding difference between nurses and foragers (FDR =0.05). To determine if these differences in binding strength are biologically meaningful, 38 genes near to these USP binding sites were identified. According to a recent gene expression comparison in fat bodies of nurses and foragers (Ament et al. Mechanisms of stable lipid loss in a social insect 2010, submitted to *PLoS Biology*), only 5 transcripts are differentially expressed between nurse and foragers, suggesting that most binding differences are not associated with differential regulation. Among these 5 transcripts, one transcript (*GB11000*) does not have *Drosophila* ortholog; one transcript (*GB14436*) has a *Drosophila* ortholog (*FBgn0033906*) with unknown functions; the rest 3 genes (*GB13199*, *GB30199*, and *GB11410*) have *Drosophila* orthologs that encode proteins involved in uridine phosphorylation, cell adhesion, and protein kinase A binding respectively. On one hand, the rarity of differential bindings, and the lack of differentially expressed genes suggest that the differences in regulation by USP involves alternate ligands or co-factors, similar to what has been described for several vertebrate nuclear receptors. On the other hand, ChIP-chip performed with whole genome titling arrays may not have enough resolution to detect subtle binding differences. The question should be revisited using a more powerful technique such as ChIP-seq in the future.

Identified USP binding regions were further characterized. USP binding sites are significantly enriched with CF1 motifs, a *Drosophila* USP consensus binding site, indicating honey bee USP binds to DNA through a conserved DNA motif. Functional analysis revealed enrichment of transcription factors among putative USP targets, suggesting USP is high in regulatory hierarchies.

Besides insights on basic molecular mechanisms of USP regulation, it may be more important to use the information provided by USP binding sites to understand a bigger biological question: what can we learn about the regulations of behavioral maturation from these results? To answer this question, I would like to first analyze the role of fat bodies in behavioral maturation, then explore possible mechanisms of USP coordinating different signaling pathways, and finally discuss the limitations of and perspectives provided by our systems biology approach.

### *The Fat Body, a Control Center*

It is intriguing that *usp* knockdown in fat bodies affects honey bee behavioral maturation. Previous studies have mostly focused on gene expression changes in brains (Alaux et al., 2009; Whitfield, Cziko, & Robinson, 2003), which are considered to be a more direct control of behavior. Recently, behavior and physiology studies have demonstrated that the nutritional status of individual bees is important in initiating or maintaining changes during behavioral maturation (Ament et al., 2008; Nelson et al., 2007; Toth et al., 2005), presumably through fat bodies. Fat bodies are functionally equivalent to liver and adipose tissues in vertebrates (Leopold & Perrimon, 2007). Fat bodies have been proposed to be an integrator of signals from other organs so that it can perform multiple metabolic functions to fulfill the changing physiological requirements of the whole body (Arrese & Soulages, 2010). And their roles are not merely passive. Indeed, it has been demonstrated that in *Drosophila* fat bodies can sense nutrients and control neural secretion cells in the brain through humoral signals, even though the signals have not been identified (Géminard, Rulifson, & Léopold, 2009). By analogy,

there might be a similar signal between honey bee fat bodies and brains to coordinate physiological and behavioral changes. This hypothesis leads to two related questions: what is the signal? And how does USP knockdown make fat bodies change the production of the signal?

One possible candidate is Vitellogenin (Vg), a major reproductive protein in insects in general, which is also considered to be an endocrine factor in honey bees. In adult worker bees, Vg is secreted mainly by abdominal fat bodies into hemolymph, making it a perfect candidate to transmit signals from fat bodies to brain. Even though a putative Vg receptor does not appear to be expressed in honey bee brains (Corona, Robinson, unpublished), the case remains unclear, and an indirect effect on brain is still a possibility. It has been proposed that high Vg titers in young workers inhibit foraging (Amdam et al., 2004; Rutz & Lüscher, 1974), and reduced Vg titers are known to promote precocious foraging (Marco et al., 2008; Nelson et al., 2007). Moreover, Vg and JH reciprocally inhibit each other, and control the timing of behavioral maturation (Amdam & Omholt, 2003; Guidugli et al., 2005). Given that USP is implicated in JH signaling, it is possible to hypothesize that USP knockdown impedes JH signaling, which releases the inhibition of Vg production and thus delays the onset of foraging. However, there are no USP binding sites near Vg, which might be a false negative result. It is possible that USP indirectly regulates Vg through regulating other transcription factors or regulators. To test this hypothesis, it is essential to determine whether USP knockdown could de-repress Vg expression in JH treated bees.

Besides Vg, there are many other possible signals. To explore these possibilities, it is essential to understand how *usp* knockdown changes fat bodies. No obvious

morphological difference was observed during fat body dissection. It will be interesting to determine if *usp* knockdown changes lipid level stored in fat bodies. Furthermore, it will be informative to measure the expression of important metabolic genes. Admittedly, it might be difficult to identify direct effects of *usp* knockdown on fat bodies, apart from indirect effects caused by signals from other tissues such as brain in response to changes in fat bodies. The direct and indirect effects might be separated by culturing different combinations of treated fat bodies and brains ex vivo, and profile gene expression in both tissues. Gene expression patterns could then be analyzed using GNR model to reveal the signaling pathways in both tissues.

#### *USP, a Master Regulator*

RXR, USP's vertebrate homologue forms heterodimers with several different nuclear receptors to regulate related yet distinct pathways respectively (Chawla et al., 2001; Mangelsdorf et al., 1995; Sonoda et al., 2008). It has been suggested that like its vertebrate homologue, USP can form heterodimers with several nuclear receptors, and function as a master regulator (Yao et al., 1993; Zhu et al., 2000; Zhu et al., 2003). One of the USP heterodimer partners, EcR is well-characterized in *Drosophila*. Together with USP, EcR functions as an ecdysone receptor (Yao et al., 1993). Other heterodimer partners such as Hr38 and SVP have been suggested, but their mechanisms and signaling pathways are still unclear (J Zhu et al., 2000; Jinsong Zhu et al., 2003).

The USP ChIP-chip results presented here provide some evidence supporting the aforementioned picture of USP regulation. Putative USP targets are enriched in developmental genes and genes involved in edysone signaling, supporting USP's role as

a functional ecdysone receptor. Hr38 and SVP are among the putative USP targets, supporting their involvement in some aspects of USP signaling. In addition, binding motifs of vertebrate PPAR/RXR heterodimers are significantly enriched in USP binding regions, and these motifs are unlikely to be EcR/USP binding sites, suggesting a different heterodimer partner other than EcR.

Besides ecdysone signaling, USP is also implicated in JH signaling. In insects, JH and ecdysone signaling are interconnected, and together, they regulate several biological processes, such as metamorphosis and vitellogenesis (Hagedorn, 1983; Riddiford, 1985). It has been proposed that JH and ecdysone pathways are connected through USP (Fang et al., 2005; Lezzi et al., 2009). We have shown that many ecdysone responsive genes in developmental processes are among USP putative targets: E75B, E78C, Hr4 and Hr46. It will be interesting to determine whether these genes are involved in behavioral maturation and if so how. Genes that are up-regulated in forager fat bodies are also enriched in development-related GO terms (Ament et al. Mechanisms of stable lipid loss in a social insect 2010, submitted to *PLoS Biology*), supporting that many developmental genes are re-used in behavioral maturation. There are two non-exclusive hypotheses concerning how these genes are involved in behavioral maturation.

One hypothesis is that these genes function in response to ecdysone. It has been demonstrated that ecdysone signaling can regulate different biological processes through conserved downstream genes (Li et al., 2000; Raikhel et al., 2002; Velarde et al., 2009). So it is possible that the canonical ecdysteroid transcriptional cascade previously defined for metamorphosis and vitellogenesis is used in behavioral maturation. However, one difficulty associated with this hypothesis is that ecdysone titers remains low in adult

worker bees, except for a small peak around 3 days after eclosion (Hartfelder et al., 2002). The function of this early ecdysone peak is unknown. It has been proposed that this ecdysone peak triggers cell differentiation of wax glands of young workers (Hartfelder et al., 2002). However, the transcriptional response is not restricted to wax glands. It has been shown that in honey bee Kenyon cells, EcR, USP and several other ecdysone responsive genes are up-regulated, coincident with the ecdysone peak (Velarde et al., 2009). It is possible that this early ecdysone peak triggers developmental process in multiple tissues, setting the stage for the dramatic JH titer increase later in behavioral maturation.

Alternatively, independent of ecdysone signaling, USP and its downstream genes function in response to JH signaling. In young bees, high dose JH treatment affects gene expression of USP and some other ecdysone responsive genes, but not EcR (Velarde et al., 2009), suggesting those ecdysone responsive genes are also part of JH signaling through USP regulation. The suggested overlapping between ecdysone responsive genes and JH responsive genes provides the potential for interactions between these two important hormonal signals. This hypothesis also fits well with the aforementioned gene expression data (Ament, Robinson et al., unpublished), explaining why developmental genes are enriched in genes up-regulated in forager fat bodies.

USP putative targets go beyond ecdysone responsive genes. Functional analyses revealed that there are several genes involved in other signaling pathways of behavioral maturation. To be specific, there are *InR1* in insulin signaling, *JHamt* in JH synthesis, and *for*, encoding a cGMP-dependent protein kinase.

First, the suggested regulation of *inRI* by USP provides a possible channel for insulin-JH interactions during honey bee behavioral maturation. Interactions between JH and insulin signaling have been demonstrated in insects in contexts other than behavioral maturation. In *Drosophila*, mutation in insulin signaling impedes JH synthesis (Tu, Yin, & Tatar, 2005). In honey bees, JH treatment significantly increases the expression of insulin-like peptide 1 (ILP-1) in the heads of young workers (Corona et al., 2007). Moreover, like JH signaling, insulin signaling is also involved in behavioral maturation. Foragers have a higher *inRI* expression level in abdominal fat bodies compared to that of nurses, coincident with foragers' higher JH titers (Ament et al., 2008). However, no direct regulation of *inRI* by JH have been established in behavioral maturation. The USP binding site near *inRI* suggests a possible mechanism that JH signaling might crosstalk to insulin signal through USP regulation of insulin genes, and this connection might be utilized in behavioral maturation.

Second, JHamt as a putative USP target gene implies feedback regulation of JH production through USP. JHamt is an enzyme involved in the final steps of JH synthesis (Bellés, Martín, & Piulachs, 2004). It is not clear which enzymes are rate-limiting in JH synthesis in honey bees. In adult female mosquitoes (*Aedes aegypti*), it has been suggested the rate-limiting steps in JH synthesis occur prior to JHamt (Mayoral et al., 2009). In silkworm larvae (*Bombyx mori*), however, JHamt is a rate-limiting enzyme in JH synthesis, and its expression level is tightly correlated with JH titers (Sheng, Ma, Cao, Jiang, & Li, 2008). Similarly, during the period of adult emergence of the Eri silkworm (*Samia cynthia ricini*), the mRNA level and enzyme activity were closely correlated with the patterns of JH synthesis, suggesting JHamt is a key regulatory enzyme in JH synthesis

(Sheng et al., 2008). JH synthesis is generally known to occur in *corpora allata* (CA). In young adult Eri silkworm (*Samia cynthia ricini*), *jhamt* mRNAs were detected in CA exclusively. However, honey bee *jhamt* seems to be expressed in many tissues in adult bees. It will be interesting to investigate the role of JHamt in JH synthesis in honey bee CA. Moreover, it is still unknown what the function of JHamt is in fat bodies. Exploring whether and how *jhamt* expression is regulated by USP in CA and fat bodies might shed light on an important feedback mechanism through which JH synthesis is regulated.

Third, *foraging* (*for*), a gene involved in behavioral maturation, is also a putative USP target. *for* encodes a cGMP-dependent protein kinase (PKG) (Osborne et al., 1997). Naturally-occurring *for* alleles were originally shown in *Drosophila* to be associated with variation in larval foraging strategies. In honey bees, differences of brain expression of *for* were shown to regulate behavioral maturation. There is no gene expression difference of *for* in fat bodies of nurses and foragers, suggesting its function in behavioral maturation is mainly mediated through brain gene expression. It seems that USP does not display tissue-specific binding to its targets, but nonetheless regulates certain targets only in certain tissues (see below for further discussion).

In summary, putative USP targets suggest potential mechanisms to integrate different regulatory signaling pathways in behavioral maturation, suggesting USP acts as a master regulator. However, USP binding near the gene does not necessarily mean the genes are regulated by USP in that tissue. These intriguing regulatory relationships need to be directly tested, and supported by independent evidence.

### *Systems Biology, Limitations and Hopes*

Analyses of USP binding sites in honey bee fat bodies generated intriguing hypotheses about the regulatory mechanisms of behavioral maturation. These hypotheses should be evaluated with caution and directly tested using independent assays due to the limitations of the approach. One major limitation is that the relationship between transcription factor binding and transcriptional regulation is still unclear (Farnham, 2009). There are 3 important remaining problems that await clarification:

First, the assignment of a specific binding site to a target gene is not accurate. With the accumulation of genomic maps of transcription factor binding sites, it is clear that transcription factor binding is not restricted to proximal promoter regions. Instead, transcriptional binding sites can occur anywhere along genes, and binding at distal regulatory elements is common. Comparisons of the genome-wide profiles for a large number of factors in the ENCODE pilot project (Birney et al., 2007) showed that less than 10% of the factors tested had greater than 50% of their binding sites within 2.5 kb of a transcription start site (Rozowsky et al., 2009). In mammalian genomes, distal regulatory regions can be as far as 100kb away from the regulated gene (Carroll et al., 2005). Given the difficulties associated with assigning binding sites to nearby genes, additional information such as the sequences of insulators DNA elements was used to improve predictions (Xi et al., 2007). With the development of RNA deep sequencing and ChIP-seq, better characterizations of the honey bee genome such as more accurate transcription start sites and histone modifications are on the horizon (Johnson, Mortazavi, Myers, & Wold, 2007). This new knowledge will tremendously facilitate our understanding of transcription factor binding data.

Second, the relationship between tissue-specific binding and tissue-specific regulation of nuclear receptors is still unknown, so it is difficult to infer which genes with USP binding sites are regulated in fat bodies, or how reliable it is to predict regulatory relationships in other tissues using USP binding information obtained from fat bodies. It seems that at least in some cases, transcription factor binding sites in one tissue/cell type can predict regulatory relationships in other tissues. In a recent publication, 42% of the EcR/USP binding sites obtained from Kc167 cells are nearby 20HE-regulated genes in these cells. Another 44% of the binding sites are close to 20HE-regulated genes identified in other tissues, indicating that while EcR/USP binding regions appear to be non-tissue-specific, the activation by the receptor complex occurs in a tissue-specific manner (Gauhar et al., 2009). If this conclusion holds true for honey bee USP, regulation of USP should be tested in multiple tissues besides fat bodies. For example, *jhamt*, a JH synthesis enzyme, has 3 USP binding sites identified in fat bodies. While it may be interesting to explore *jhamt*'s function in fat bodies to test if USP regulates JH synthesis through regulating the expression of *jhamt*, it makes more sense to use corpora allata, the paired glands where JH is synthesized.

Third, mechanisms of transcriptional regulation are under re-examination (Farnham, 2009). A general assumption of many previous analyses of transcriptional regulation is that transcription factors act individually and have a specific role in regulating a specific set of genes. It has been recently suggested that transcription factors might have 'community' functions at some binding sites. For instance, a cluster of transcription factors might regulate gene expression by maintaining open chromatin structures, or transcription factors can function as “bumper proteins” to restrict scanning

regions for other transcription functions. These possible alternative regulatory mechanisms make us re-consider concepts used under general assumptions. For example, with alternative mechanisms, transcription factors might not be so much regulators rather than conditions, the presentation of which enables regulation. As genome-wide binding information of more transcription factors are obtained, a clear picture of regulatory mechanisms should be revealed.

Despite the limitations of genome-wide transcription factor binding data, they have greatly advanced our understanding of transcription regulation. In this case, genomic binding sites of honey bee USP provides insights on the regulatory mechanisms of this nuclear receptor, and generated a list of putative USP target genes. Few USP binding differences were detected between nurses and foragers. On one hand, this result suggests possible regulatory mechanisms of USP. On the other hand, the result might due to a lack of sensitivity of ChIP-chip methods. New techniques such as ChIP-seq might provide a better resolution of differential USP binding. Moreover, the results suggested several hypotheses that will promote the understanding of behavioral maturation on the molecular level. These hypotheses can be tested individually by a combination of in vitro and behavioral assays. Alternatively, or in addition, the genome-wide USP binding information can be used to construct gene regulatory networks. As a framework, the GRN model combines information from transcription factor binding data, gene expression data, and gene expression coupled with RNAi or pharmacological treatments, which provides a powerful tool to systematically analyze regulatory relationships on the genomic level. This approach would be particularly useful to understand the complexity of honey bee behavioral maturation.

## **Acknowledgments**

I would like to thank Thomas Newman, Yupeng Zheng, Lisa Stubbs, Brain Freeman for their technique advice; Feng Hong and Sheng Zhong for help with ChIP-chip peak finding software; Saurabh Sinha for motif analysis; Satwik Rajaram for help with Perl programming; Seth Ament for assistance in collection of bees used in ChIP-chip; Karen Pruiett for beekeeping; and members of Robinson lab for helpful discussions and moral support.

## Tables

**Table 4.1:** Significant enriched motifs in USP binding regions found using SWAN. Columns from left to right are: motif names, associated transcription factors, fly or vertebrate source, p-values, and graphic representations of position weight matrix.

|                      |        |      |              |  |
|----------------------|--------|------|--------------|--|
| GGGTCA               | CF1    | FLY  | 1.166802e-06 |  |
| BMGYBGYYGY<br>NGMVBV | ADF1   | FLY  | 2.420274e-06 |  |
| TTCCSGGAA            | Dstat  | FLY  | 1.428478e-05 |  |
| YSAAGGWCRC<br>HRM    | ftz-fl | FLY  | 2.415777e-05 |  |
| BYGTGRGAAM<br>CBNDVD | SUH    | FLY  | 6.928148e-05 |  |
| V_PPARG_02           | PPAR   | VERT | 6.981086e-06 |  |

**Table 4.2:** Transcription factors among putative USP target genes.

|         |         |             |                                         |
|---------|---------|-------------|-----------------------------------------|
| GB10341 | CG11648 | FBgn0000015 | Abd-B / Abdominal B                     |
| GB10650 | CG33183 | FBgn0000448 | Hr46 / Hormone receptor-like in 46      |
| GB10759 | CG32180 | FBgn0000567 | Eip74EF / Ecdysone-induced protein 74EF |
| GB11229 | CG5580  | FBgn0010575 | sbb / scribbler                         |
| GB11420 | CG32611 | FBgn0052611 | CG32611 / CG32611                       |
| GB11441 | CG10601 | FBgn0014343 | mirr / mirror                           |
| GB11566 | CG12154 | FBgn0004102 | oc / ocelliless                         |
| GB11582 | CG12399 | FBgn0011648 | Mad / Mothers against dpp               |
| GB11654 | CG15455 | FBgn0031121 | CG15455 / CG15455                       |
| GB11761 | CG9656  | FBgn0001138 | grn / grain                             |
| GB11857 | CG11966 | FBgn0037645 | CG11966 / CG11966                       |
| GB11901 | CG6993  | FBgn0003513 | ss / spineless                          |
| GB12313 | CG2052  | FBgn0039905 | CG2052 / CG2052                         |
| GB12355 | CG1464  | FBgn0005558 | ey / eyeless                            |
| GB12480 | CG1133  | FBgn0003002 | opa / odd paired                        |
| GB12566 | CG8817  | FBgn0041111 | lilli / lilliputian                     |
| GB12972 | CG2939  | FBgn0004567 | slp2 / sloppy paired 2                  |
| GB13030 | CG5058  | FBgn0004586 | grh / grainy head                       |
| GB13693 | CG7368  | FBgn0036179 | CG7368 / CG7368                         |
| GB14416 | CG10002 | FBgn0000659 | fkh / fork head                         |
| GB14651 | CG9908  | FBgn0000459 | disco / disconnected                    |
| GB14672 | CG16779 | FBgn0037698 | CG16779 / CG16779                       |
| GB14841 | CG31612 | FBgn0051612 | CG31612 / CG31612                       |
| GB14857 | CG6494  | FBgn0001168 | h / hairy                               |
| GB14998 | CG4354  | FBgn0005638 | slbo / slow border cells                |
| GB15081 | CG3411  | FBgn0004101 | bs / blistered                          |
| GB15295 | CG1447  | FBgn0020912 | Ptx1 /                                  |
| GB15421 | CG7847  | FBgn0003499 | sr / stripe                             |
| GB15632 | CG2692  | FBgn0001147 | gsb-n / gooseberry-neuro                |
| GB15643 | CG11641 | FBgn0033288 | CG11641 / CG11641                       |
| GB15698 | CG10488 | FBgn0000625 | eyg / eyegone                           |
| GB16085 | CG10037 | FBgn0003995 | wl / ventral veins lacking              |
| GB16145 | CG3242  | FBgn0004892 | sob / sister of odd and bowl            |
| GB16177 | CG11152 | FBgn0039937 | CG11152 / CG11152                       |
| GB16191 | CG5893  | FBgn0000411 | D / Dichaete                            |
| GB16259 | CG6534  | FBgn0002941 | slou / slouch                           |
| GB16457 | CG2762  | FBgn0003963 | ush / u-shaped                          |
| GB16648 | CG4380  | FBgn0003964 | usp / ultraspiracle                     |
| GB16784 | CG6172  | FBgn0003986 | vnd / ventral nervous system defective  |
| GB16786 | CG7951  | FBgn0015542 | sima / similar                          |
| GB16863 | CG16902 | FBgn0023546 | Hr4 / Hr4                               |
| GB17100 | CG11502 | FBgn0003651 | svp / seven up                          |
| GB17213 | CG11798 | FBgn0015371 | chn / charlatan                         |
| GB17617 | CG14307 | FBgn0004652 | fru / fruitless                         |
| GB18348 | CG17117 | FBgn0001235 | hth / homothorax                        |
| GB18397 | CG11049 | FBgn0005561 | sv / shaven                             |
| GB18541 | CG5403  | FBgn0004795 | retn / retained                         |

**Table 4.2 (cont.):** Transcription factors among putative USP target genes.

|         |         |             |                                                                           |
|---------|---------|-------------|---------------------------------------------------------------------------|
| GB18560 | CG5575  | FBgn0011236 | ken / ken and barbie                                                      |
| GB18613 | CG17077 | FBgn0003118 | pnt / pointed                                                             |
| GB18685 | CG6964  | FBgn0010825 | Gug / Grunge                                                              |
| GB18923 | CG4257  | FBgn0016917 | Stat92E / Signal-transducer and activator of transcription protein at 92E |
| GB19470 | CG12296 | FBgn0013469 | klu / klumpfuss                                                           |
| GB19738 | CG10325 | FBgn0000014 | abd-A / abdominal A                                                       |
| GB19838 | CG9403  | FBgn0041145 | jing /                                                                    |
| GB19879 | CG1832  | FBgn0032979 | CG1832 / CG1832                                                           |
| GB19937 | CG18024 | FBgn0029123 | SoxN / SoxNeuro                                                           |
| GB19951 | CG15552 | FBgn0024288 | Sox100B / Sox100B                                                         |
| GB19998 | CG5799  | FBgn0020307 | dve / defective proventriculus                                            |
| GB30126 | CG2380  | FBgn0042696 | Nfl / Nuclear factor I                                                    |
| GB30139 | CG6883  | FBgn0003749 | trh / trachealess                                                         |
| GB30205 | CG17964 | FBgn0019664 | pan / pangolin                                                            |
| GB30210 | CG4952  | FBgn0005677 | dac / dachshund                                                           |
| GB30226 | CG18023 | FBgn0004865 | Eip78C / Ecdysone-induced protein 78C                                     |
| GB30234 | CG7018  | FBgn0005658 | Ets65A / Ets at 65A                                                       |

**Table 4.3:** Nuclear receptors among putative USP target genes.

|               | Vertebrate homologue | CF1 motifs | Overlap with other results |
|---------------|----------------------|------------|----------------------------|
| USP           | RXR                  | yes        | fruit fly                  |
| SVP           | COUP-TF              | yes        | fruit fly, mouse           |
| Eip75B        | REV-ERB              | yes        | fruit fly                  |
| Eip78C        | REV-ERB              | no         | fruit fly                  |
| Hr4           | GCNF                 | yes        | .                          |
| Hr38(GB15642) | NURR1/NGFIB          | yes        | .                          |
| Hr38(GB17814) | NURR1/NGFIB          | no         | .                          |
| Hr46          | ROR                  | no         | .                          |

**Table 4.4:** Putative USP target genes involved in lipid and energy metabolism.

|         | Gene                                 | CF1 motifs | Overlap with other results |
|---------|--------------------------------------|------------|----------------------------|
| GB12778 | acyl-CoA synthetase family member 2  | yes        | mouse, PPAR; fruit fly     |
| GB13289 | acyl-CoA synthetase family member 3  | no         | fruit fly                  |
| GB13198 | Acyl-CoA binding domain containing 3 | yes        | .                          |
| GB30513 | PLCB4                                | no         | .                          |
| GB10253 | levy                                 | yes        | fruit fly                  |

**Table 4.5:** Primers used in dsRNA synthesis and RT-PCR.

|                    |                                                       |
|--------------------|-------------------------------------------------------|
| T7-USP-apis-241F   | TAATACGACTCACTATAGAGTCCTTCCGGCCCCCA                   |
| T7-USP-apis-756R   | TAATACGACTCACTATAGGCAAATGTGCGACACTGCA                 |
| T7-GFP-F           | TAATACGACTCACTATAGGGCGAGC<br>CAACACTTGTCACTACTTTCTCTT |
| T7-GFP-R           | TAATACGACTCACTATAGGGCGAAGGTAATGGTTGTCTGGTAAAAGGAC     |
| GB16648-USP F1100  | AGGGTTATTGCCGCGTAGCT                                  |
| GB16648-USP R-1168 | CGGGCAGGCGTAGAAGTAATT                                 |
| RCPI-211F          | TCAATTAACTCGGAATCGGA                                  |
| RCPI-275R          | CCTGGATTTCCTGCTGAT                                    |

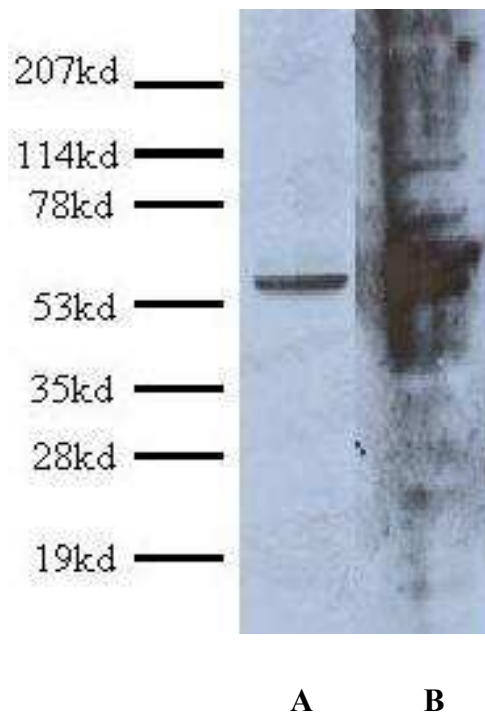
**Table 4.6:** Peak regions and associated USP putative genes with different USP binding intensity between nurses and foragers.

| <b>Genomic Regions</b>   | <b>FDR</b> | <b>Associated Genes</b> | <b><i>Drosophila</i> ortholog</b> |
|--------------------------|------------|-------------------------|-----------------------------------|
| Group4.27_487725_488259  | 0.04       | GB30199-RA              | FBgn0000497                       |
| Group6.12_220225_220671  | 0.03       | GB17560-RA              | FBgn0035538                       |
| Group7.13_204424_205853  | 0.03       | GB12861-RA              | .                                 |
| Group5.21_195023_195657  | 0.04       | GB13750-RA              | FBgn0039286                       |
| Group5.21_195023_195657  | 0.04       | GB19081-RA              | FBgn0083946                       |
| Group1.1_1064915_1065762 | 0.02       | GB14436-RA              | FBgn0033906                       |
| Group1.1_1064915_1065762 | 0.02       | GB12075-RA              | .                                 |
| GroupUn.4564_506_1061    | 0.04       | GB15366-RA              | .                                 |
| Group16.19_354406_355566 | 0.04       | GB18833-RA              | FBgn0000028                       |
| GroupUn.4525_426_1071    | 0.05       | GB14768-RA              | .                                 |
| GroupUn.8489_410_1769    | 0.02       | GB17314-RA              | .                                 |
| Group16.19_49422_49972   | 0.03       | GB11316-RA              | FBgn0051646                       |
| Group3.26_51023_51558    | 0.03       | GB15270-RA              | FBgn0039260                       |
| GroupUn.3127_525_985     | 0.05       | GB18343-RA              | .                                 |
| Group8.24_519823_520152  | 0.01       | GB18150-RA              | FBgn0001402                       |
| Group5.11_446805_448767  | 0.05       | GB18027-RA              | FBgn0034412                       |
| Group5.11_446805_448767  | 0.05       | GB13431-RA              | .                                 |
| Group6.34_484023_484955  | 0.02       | GB11946-RA              | FBgn0002643                       |
| Group5.33_213902_214251  | 0.03       | GB13477-RA              | FBgn0005511                       |
| Group5.33_213902_214251  | 0.03       | GB20112-RA              | FBgn0010341                       |
| Group13.6_297222_297552  | 0.04       | GB20131-RA              | FBgn0031854                       |
| Group3.31_293702_294158  | 0.05       | GB16863-RA              | FBgn0023546                       |
| Group4.27_185122_185755  | 0.01       | GB11000-RA              | .                                 |
| Group4.27_185122_185755  | 0.01       | GB11410-RA              | FBgn0027932                       |
| Group3.36_8116_10667     | 0.03       | GB12386-RA              | .                                 |
| Group3.36_8116_10667     | 0.03       | GB13359-RA              | FBgn0033017                       |
| Group3.36_8116_10667     | 0.03       | GB19473-RA              | FBgn0033021                       |

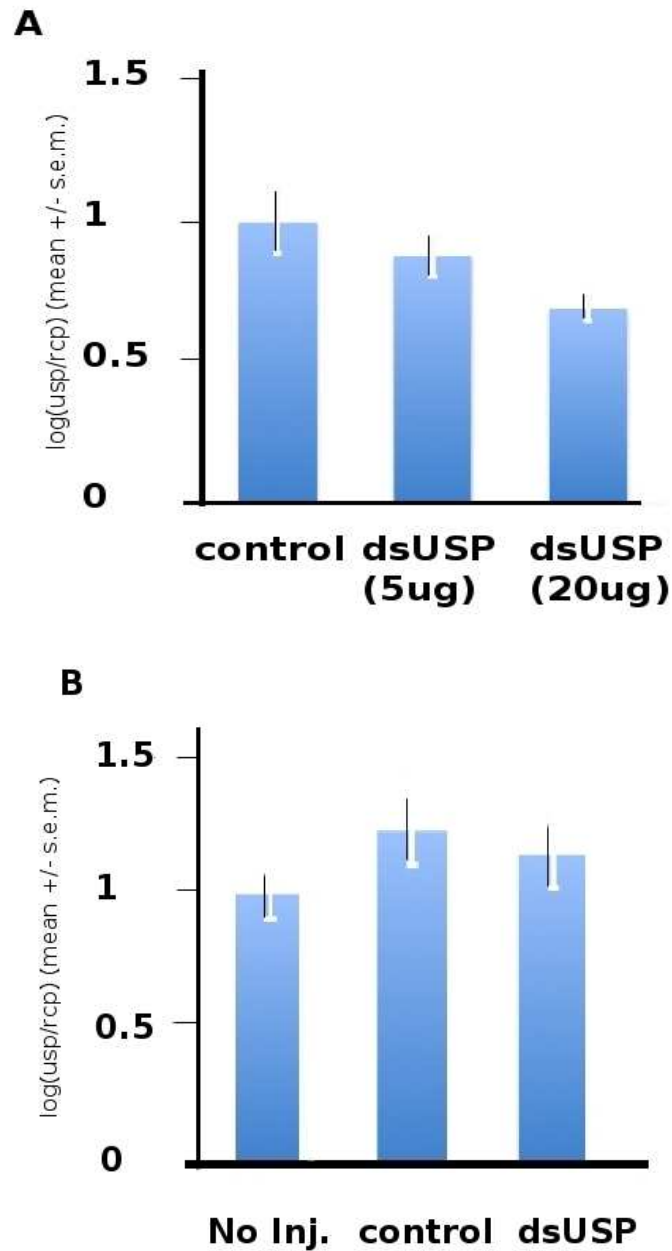
**Table 4.6 (cont.):** Peak regions and associated USP putative genes with different USP binding intensity between nurses and foragers.

| <b>Genomic Regions</b>    | <b>FDR</b> | <b>Associated Genes</b> | <b><i>Drosophila</i> ortholog</b> |
|---------------------------|------------|-------------------------|-----------------------------------|
| Group1.66_1563820_1564365 | 0.04       | GB19115-RA              | .                                 |
| GroupUn.5624_126_867      | 0.05       | GB17340-RA              | .                                 |
| Group1.62_313502_314072   | 0.04       | GB19438-RA              | .                                 |
| GroupUn.7787_6_1050       | 0.05       | GB10479-RA              | .                                 |
| Group7.3_157026_157360    | 0.01       | GB17667-RA              | FBgn0037213                       |
| GroupUn.7346_619_1057     | 0          | GB19641-RA              | FBgn0036784                       |
| Group5.22_591910_594450   | 0.02       | GB18706-RA              | FBgn0034408                       |
| Group13.6_230315_230854   | 0.05       | GB20131-RA              | FBgn0031854                       |
| Group3.21_272723_273467   | 0.05       | GB12947-RA              | FBgn0045852                       |
| Group7.28_138223_139476   | 0.03       | GB13199-RA              | FBgn0039464                       |
| Group1.1_648803_649472    | 0.02       | .                       | .                                 |
| Group1.24_312411_313260   | 0.05       | .                       | .                                 |
| Group10.17_172609_173253  | 0.05       | .                       | .                                 |
| Group10.45_319706_320355  | 0.05       | .                       | .                                 |
| Group11.19_84226_84869    | 0.02       | .                       | .                                 |
| Group11.21_399616_400150  | 0.05       | .                       | .                                 |
| Group11.8_206713_207163   | 0.03       | .                       | .                                 |
| Group16.19_456924_457473  | 0.05       | .                       | .                                 |
| Group16.9_1292124_1292458 | 0.04       | .                       | .                                 |
| Group4.22_66322_66861     | 0.01       | .                       | .                                 |
| Group5.22_658111_658570   | 0.04       | .                       | .                                 |
| Group8.16_373415_373862   | 0.05       | .                       | .                                 |
| Group9.24_666526_667450   | 0.03       | .                       | .                                 |
| GroupUn.1647_326_871      | 0.04       | .                       | .                                 |
| GroupUn.4191_915_1482     | 0.03       | .                       | .                                 |
| GroupUn.4671_709_1166     | 0.03       | .                       | .                                 |
| GroupUn.4709_1_457        | 0.03       | .                       | .                                 |
| GroupUn.500_19326_19763   | 0          | .                       | .                                 |
| GroupUn.5535_301_1258     | 0.01       | .                       | .                                 |
| GroupUn.5916_11_880       | 0.05       | .                       | .                                 |
| GroupUn.6004_106_1670     | 0.05       | .                       | .                                 |
| GroupUn.7340_26_575       | 0.05       | .                       | .                                 |
| GroupUn.7681_11_1064      | 0.05       | .                       | .                                 |
| GroupUn.7919_405_1260     | 0.02       | .                       | .                                 |
| GroupUn.7977_211_867      | 0.03       | .                       | .                                 |
| GroupUn.854_24406_24872   | 0          | .                       | .                                 |

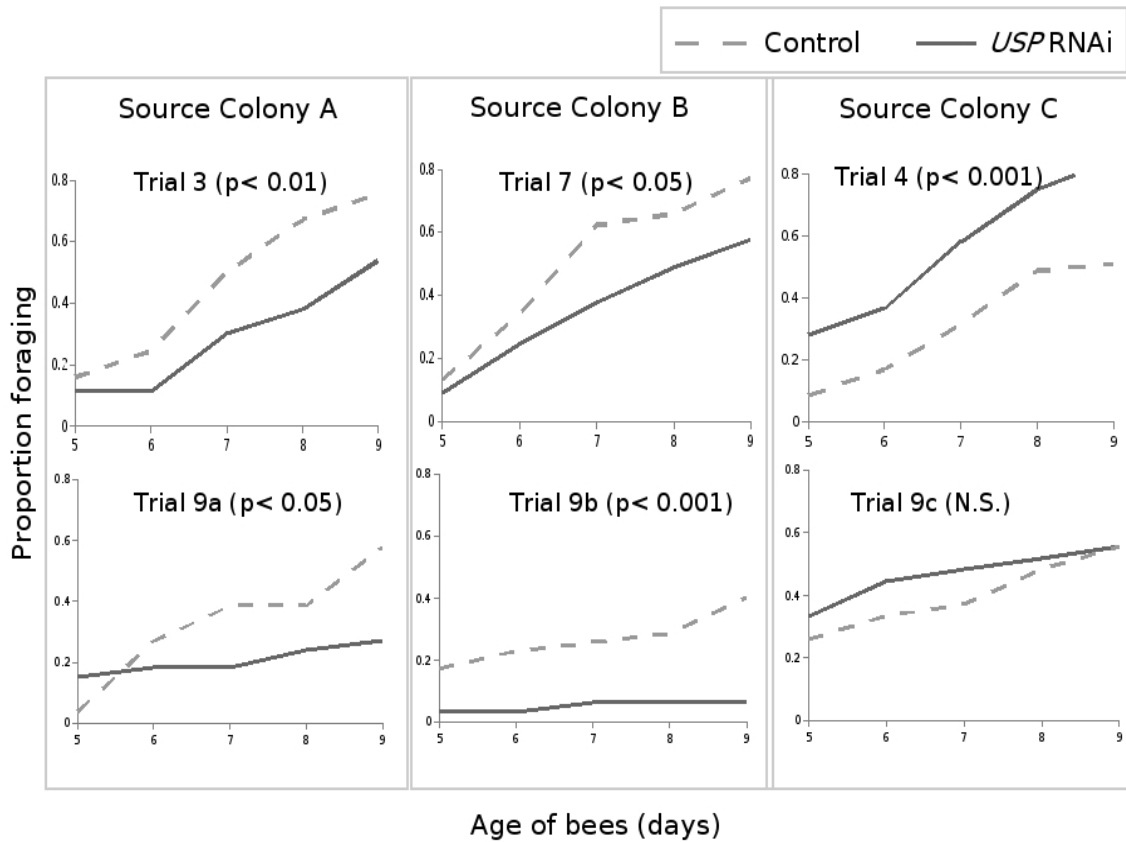
## Figures



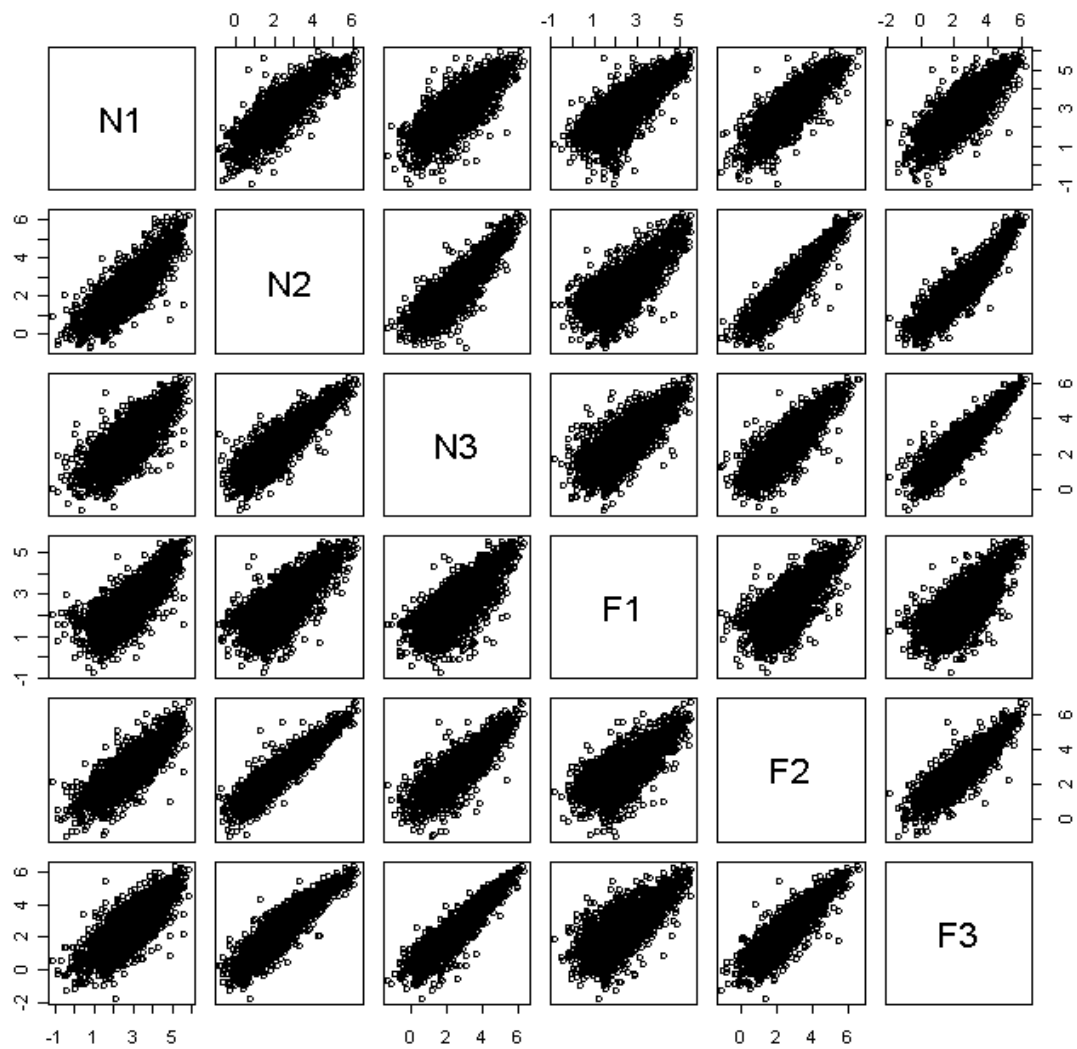
**Figure 4.1.** Western blots of antibodies that recognize honey bee USP. Fat body protein extracts were used in both assays. Lane (A) shows the results of the antibody used in ChIP-chip; lane (B) shows a different USP antibody generated using a different peptide as antigen. Both antibodies recognized a band at the predicted molecular weight, but the latter has more background signals. Both western blots were performed using same washing conditions and with the same secondary antibodies.



**Figure 4.2.** USP knockdown in honey bee fat bodies (**A**) but not heads (**B**). On the 4<sup>th</sup> day after dsRNA injection, bees were sacrificed and USP levels were determined by Q-PCR in fat bodies and heads respectively. USP mRNA levels were log transformed normalized to a spike-in gene rcp. In fat bodies (**A**), USP mRNA level was significantly decreased in bees injected with 20ug of dsUSP, compared to bees injected with control dsRNA ( $p < 0.05$ ,  $n=8$ ). Bees injected with 5ug of dsUSP showed intermediate expression. In honey bee heads (**B**), no effect of dsUSP was detected among bees without injection, bees injected with control dsRNA and bees injected with dsUSP.

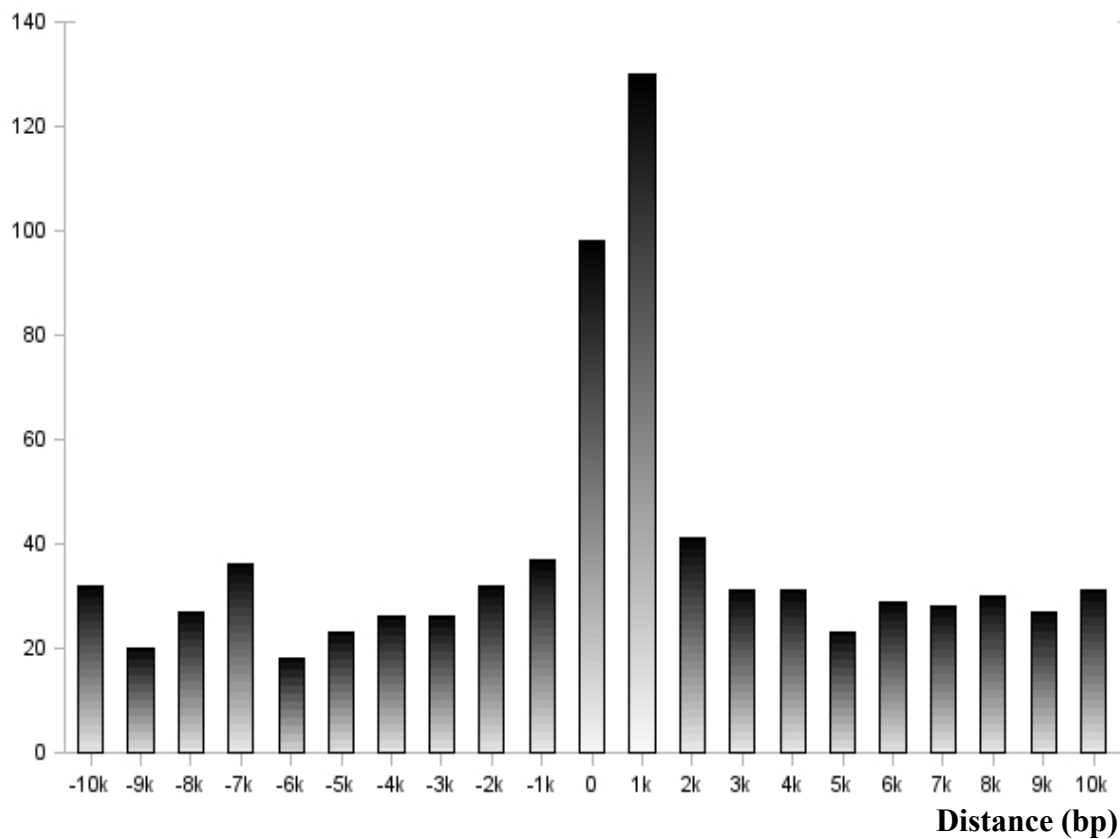


**Figure 4.3.** Delayed behavioral maturation caused by *usp* knockdown. Proportions of bees that initiated foraging at 5-9 days of age after injected with *dsusp* or control dsRNA. Data from 3 source colonies were shown. All trials were performed in summer 2009. In the first 8 trials, bees from a single source colony were put into each host colonies. In trial 9, bees from the 3 source colonies were co-hosted in a single colony during the same period to eliminate weather and host colony effects. p-values were derived from Cox regression.

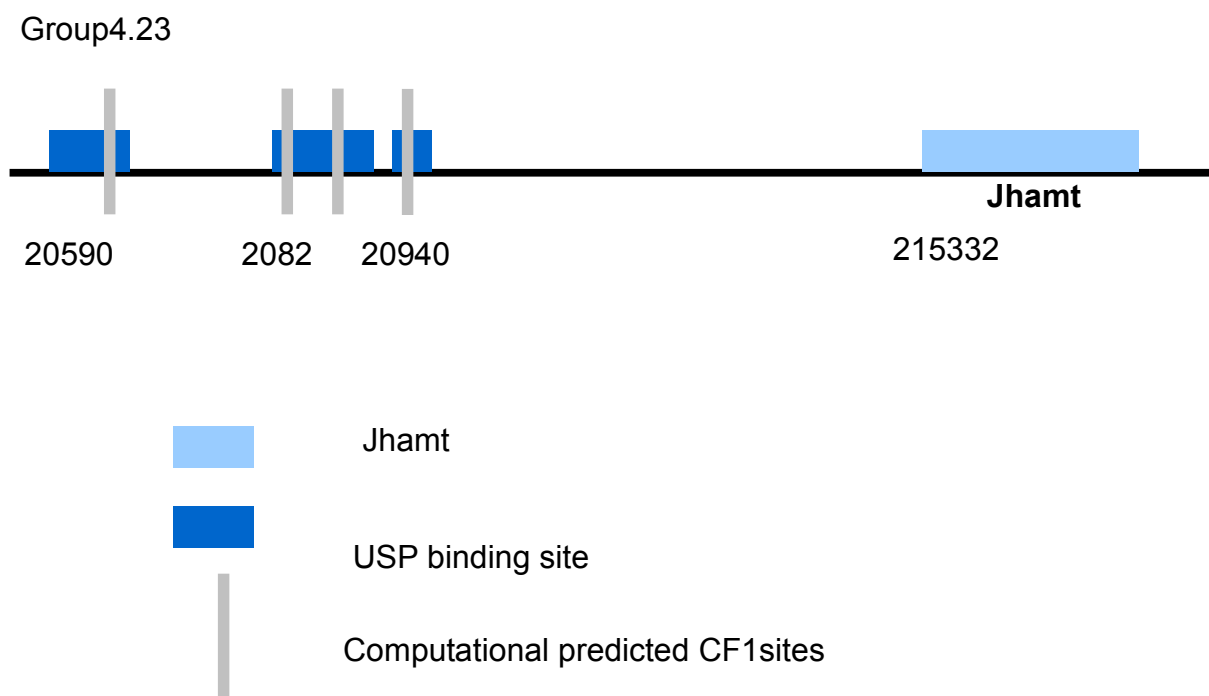


**Figure 4.4.** Scatter plots of comparisons between all 6 ChIP-chip assays showing high correlation of peak regions. Log2 ratios of immunoprecipitated DNA / input genomic DNA from each array were used in the comparison. In every small graph, the x-axis and y-axis each represent data from one array. For example, the y-axes of graphs in the first row, and x-axes of graphs in the first column represent data from nurse trial 1.

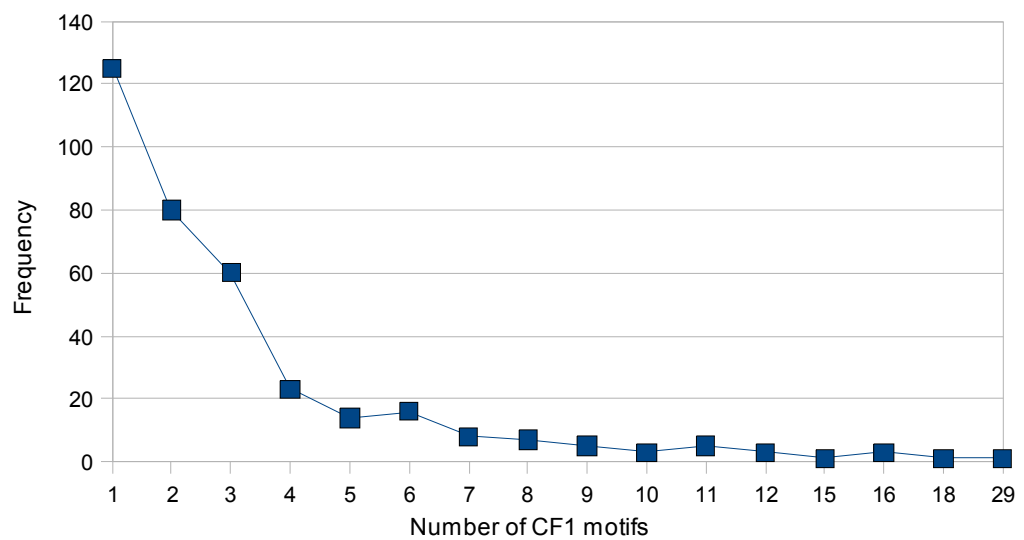
## Frequency



**Figure 4.5.** Histogram of peak regions at different distance to their nearby genes. Frequencies of peak regions in different distance with the nearby genes were shown at an interval of 1kb. The distance between peak regions and genes was measured by the minimal distance between a peak region and the predicted translation start sites of the genes. If a translation start site occurs inside of a peak region, the distance is defined as 0.



**Figure 4.6.** An example of USP binding regions and a USP putative target gene. Relative positions of USP binding regions near *jhamt* are shown together with CF1 motifs detected using SWAN.



**Figure 4.7.** The distribution of the numbers of CF1 motifs associated with USP putative target genes. CF1 motifs were counted inside of USP binding regions within 10kb of USP putative target genes. Frequencies of each count were represented on y-axis.

## Chapter 5

### Conclusions

Adult worker honey bees (*Apis mellifera*) undergo an age-related task transition, called behavioral maturation (Winston, 1991), which provides a system to study molecular mechanisms of phenotypic plasticity. Two possible mechanisms, DNA methylation and transcription factor regulation, were explored.

Characterization of the functional DNA methylation system in honey bees revealed catalytically active enzymes similar to those in vertebrates. However, the identified DNA methylations all occur in intragenic regions rather than promoter regions, suggesting different functions of DNA methylation in invertebrates compared to vertebrates. Several years ago when the project was first underway, functional analysis of how DNA methylation involved in honey bee behavioral maturation was hindered by difficulties identifying methylated loci in a genome-wide manner, partially due to relatively low DNA methylation densities in the honey bee genome. Development of new technologies offered new opportunities for understanding functions of DNA methylation in honey bees. With much cheaper DNA sequencing techniques, “bisulfite sequencing” (BS-Seq) (Cokus et al., 2008) can be used for genome-wide detection of DNA methylation. A new DNA sequencing technique is under improvement, which promises to detect methylated nucleotides directly (Flusberg et al., 2010). Meanwhile, several bioinformatic analyses have been done to predict DNA methylation patterns in the honey bee genome (Elango, Hunt, Goodisman, & Yi, 2009; Foret, Kucharski, Pittelkow,

Lockett, & Maleszka, 2009). All these new developments in this young field make it possible to perform an in-depth analysis of honey bee DNA methylation and its function in behavioral maturation.

A second exploration focused on the regulatory functions of a nuclear receptor, USP. It was demonstrated that USP in fat bodies is important in pacing behavioral maturation. The mechanisms by which USP regulates differential gene expressions underlying behavioral maturation were framed in a Gene Regulatory Network (GRN) model. As a first step towards describing USP-centered GRNs, genome-wide USP binding sites were mapped in honey bee fat bodies and putative target genes were identified near those binding sites. Transcription factors were significantly enriched in USP putative target genes, suggesting USP regulates effector genes through regulating their corresponding transcription factors. Functional analysis of putative USP targets also suggests USP might coordinate multiple signaling pathways involved in behavioral maturation.

The hypotheses generated by USP binding sites need to be tested in further investigations. The suggested regulatory relationships can be individually tested by experiments such as luciferase assays (Brasier, Tate, & Habener, 1989) or chromatin conformation capture (Miele, Gheldof, Tabuchi, Dostie, & Dekker, 2006). Additionally, in the GRN models, genome-wide transcription data can be combined with gene expression data, RNAi coupled with gene expression data, and bioinformatic analyses. With the accumulation of genomic data and the development of more sophisticated modeling methods, a more complete regulatory picture underlying honey bee behavioral maturation will eventually emerge.

## References

- Adamo, S. A., Linn, C. E., & Hoy, R. R. (1995). The role of neurohormonal octopamine during 'fight or flight' behaviour in the field cricket *Gryllus bimaculatus*. *Journal of experimental Biology*, 198(8), 1691.
- Alaux, C., Le Conte, Y., Adams, H. A., Rodriguez-Zas, S., Grozinger, C. M., Sinha, S., & Robinson, G. E. (2009). Regulation of brain gene expression in honey bees by brood pheromone. *Genes, Brain, and Behavior*, 8(3), 309-319.
- Alaux, C., Sinha, S., Hasadsri, L., Hunt, G. J., Guzmán-Novoa, E., DeGrandi-Hoffman, G., Uribe-Rubio, J. L., et al. (2009). Honey bee aggression supports a link between gene regulation and behavioral evolution. *Proceedings of the National Academy of Sciences*, 106(36), 15400.
- Amdam, G. V., Norberg, K., Fondrk, M. K., & Page, R. E. (2004). Reproductive ground plan may mediate colony-level selection effects on individual foraging behavior in honey bees. *Proceedings of the National Academy of Sciences of the United States of America*, 101(31), 11350-11355.
- Amdam, G. V., & Omholt, S. W. (2003). The hive bee to forager transition in honeybee colonies: the double repressor hypothesis. *Journal of theoretical biology*, 223(4), 451-464.
- Amdam, G. V., Simoes, Z. L., Guidugli, K. R., Norberg, K., & Omholt, S. W. (2003). Disruption of *vitellogenin* gene function in adult honeybees by intra-abdominal injection of double-stranded RNA. *BMC biotechnology*, 3(1), 1.
- Ament, S. A., Corona, M., Pollock, H. S., & Robinson, G. E. (2008). Insulin signaling is involved in the regulation of worker division of labor in honey bee colonies. *Proceedings of the National Academy of Sciences of the United States of America*, 105(11), 4226-4231.
- Arrese, E. L., & Soulages, J. L. (2010). Insect Fat Body: Energy, Metabolism, and Regulation. *Annual Review of Entomology*, 55(1), 207-225.
- Aubin-Horth, N., & Renn, S. C. (2009). Genomic reaction norms: using integrative biology to understand molecular mechanisms of phenotypic plasticity. *Molecular Ecology*, 18(18), 3763-3780.
- Barchuk, A. R., Figueiredo, V. L., & Simoes, Z. L. (2008). Downregulation of *ultraspiracle* gene expression delays pupal development in honeybees. *Journal of insect physiology*, 54(6), 1035-1040.
- Barchuk, A. R., Maleszka, R., & Simoes, Z. L. (2004). *Apis mellifera* *ultraspiracle*: cDNA sequence and rapid up-regulation by juvenile hormone. *Insect Molecular Biology*, 13(5), 459-467.
- Barron, A. B., Maleszka, R., Vander Meer, R. K., & Robinson, G. E. (2007). Octopamine modulates honey bee dance behavior. *Proceedings of the National Academy of Sciences*, 104(5), 1703.
- Beckstead, R., Lam, G., & Thummel, C. (2005). The genomic response to 20-hydroxyecdysone at the onset of *Drosophila* metamorphosis. *Genome biology*, 6(12), R99.
- Bellés, X., Martín, D., & Piulachs, M. D. (2004). The mevalonate pathway and the synthesis of juvenile hormone in insects.

- Ben-Shahar, Y., Robichon, A., Sokolowski, M. B., & Robinson, G. E. (2002). Influence of gene action across different time scales on behavior. *Science*, 296(5568), 741-744.
- Ben-Shahar, Y., Dudek, N. L., & Robinson, G. E. (2004). Phenotypic deconstruction reveals involvement of manganese transporter malvolio in honey bee division of labor. *Journal of Experimental Biology*, 207(19), 3281.
- Ben-Shahar, Y., Leung, H., Pak, W. L., Sokolowski, M. B., & Robinson, G. E. (2003). cGMP-dependent changes in phototaxis: a possible role for the foraging gene in honey bee division of labor. *J Exp Biol*, 206(14), 2507-2515.
- Bieda, M., Xu, X., Singer, M. A., Green, R., & Farnham, P. J. (2006). Unbiased location analysis of E2F1-binding sites suggests a widespread role for E2F1 in the human genome. *Genome Research*, 16(5), 595.
- Bird, A. (2002). DNA methylation patterns and epigenetic memory. *Genes & development*, 16(1), 6.
- Bird, A. P., Taggart, M. H., & Smith, B. A. (1979). Methylated and unmethylated DNA compartments in the sea urchin genome. *Cell*, 17(4), 889-901.
- Birney, E., Stamatoyannopoulos, J. A., Dutta, A., Guigó, R., Gingeras, T. R., Margulies, E. H., Weng, Z., et al. (2007). Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature*, 447(7146), 799–816.
- Bonneau, R. (2008). Learning biological networks: from modules to dynamics. *Nature chemical biology*, 4(11), 658–664.
- Bough, K. J., Wetherington, J., Hassel, B., Pare, J. F., Gawryluk, J. W., Greene, J. G., Shaw, R., et al. (2006). Mitochondrial biogenesis in the anticonvulsant mechanism of the ketogenic diet. *Annals of neurology*, 60(2), 223–235.
- Bounias, M., & Pacheco, H. (1979). Susceptibility of the honeybee glycemia to injection of insulin and glucagon 'in vivo'. *Comptes Rendus Hebdomadaires des Seances de l'Academie des Sciences. Serie D (France)*.
- Brasier, A. R., Tate, J. E., & Habener, J. F. (1989). Optimized use of the firefly luciferase assay as a reporter gene in mammalian cell lines. *Biotechniques*, 7(10), 1116–1122.
- Buhk, H. J., Behrens, B., Tailor, R., Wilke, K., Prada, J. J., Günthert, U., Noyer-Weidner, M., et al. (1984). Restriction and modification in *Bacillus subtilis*: nucleotide sequence, functional organization and product of the DNA methyltransferase gene of bacteriophage SPR. *Gene*, 29(1-2), 51-61.
- Carroll, J. S., Liu, X. S., Brodsky, A. S., Li, W., Meyer, C. A., Szary, A. J., Eeckhoute, J., et al. (2005). Chromosome-wide mapping of estrogen receptor binding reveals long-range regulation requiring the forkhead protein FoxA1. *Cell*, 122(1), 33-43.
- Chapman, R. F. (1998). *The Insects: Structure and Function* (4th ed.). Cambridge University Press.
- Chawla, A., Repa, J. J., Evans, R. M., & Mangelsdorf, D. J. (2001). Nuclear receptors and lipid physiology: opening the X-files. *Science*, 294(5548), 1866.
- Chen, T., & Li, E. (2004). Structure and function of eukaryotic DNA methyltransferases. *Current Topics in Developmental Biology*, 60, 55-89.
- Cokus, S. J., Feng, S., Zhang, X., Chen, Z., Merriman, B., Haudenschild, C. D., Pradhan, S., et al. (2008). Shotgun bisulphite sequencing of the *Arabidopsis* genome reveals DNA methylation patterning. *Nature*, 452(7184), 215-219.

- Corona, M., Velarde, R. A., Remolina, S., Moran-Lauter, A., Wang, Y., Hughes, K. A., & Robinson, G. E. (2007). Vitellogenin, juvenile hormone, insulin signaling, and queen honey bee longevity. *Proceedings of the National Academy of Sciences*, 104(17), 7128-7133.
- Cota, D., Proulx, K., Smith, K. A., Kozma, S. C., Thomas, G., Woods, S. C., & Seeley, R. J. (2006). Hypothalamic mTOR signaling regulates food intake. *Science*, 312(5775), 927.
- Crailsheim, K. (1992). The flow of jelly within a honeybee colony. *Journal of Comparative Physiology B: Biochemical, Systemic, and Environmental Physiology*, 162(8), 681-689.
- Crailsheim, K., Schneider, L. H. W., 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. *Journal of Insect Physiology*, 38(6), 409-419.
- Davidson, E. H. (2006). *The regulatory genome: gene regulatory networks in development and evolution*. Academic Press.
- Dennis, G., Sherman, B. T., Hosack, D. A., Yang, J., Gao, W., Lane, H. C., & Lempicki, R. A. (2003). DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biology*, 4(5), P3.
- DeZazzo, J., Sandstrom, D., de Belle, S., Velinzon, K., Smith, P., Grady, L., DelVecchio, M., et al. (2000). nalyot, a mutation of the *Drosophila* myb-related Adf1 transcription factor, disrupts synapse formation and olfactory memory. *Neuron*, 27(1), 145-158.
- Dodge, J. E., Okano, M., Dick, F., Tsujimoto, N., Chen, T., Wang, S., Ueda, Y., et al. (2005). Inactivation of Dnmt3b in Mouse Embryonic Fibroblasts Results in DNA Hypomethylation, Chromosomal Instability, and Spontaneous Immortalization. *Journal of Biological Chemistry*, 280(18), 17986-17991.
- Duncan, I. W. (2002). Transvection effects in *Drosophila*. *Annual Review of Genetics*, 36, 521-556.
- Eckhardt, F., Lewin, J., Cortese, R., Rakyan, V. K., Attwood, J., Burger, M., Burton, J., et al. (2006). DNA methylation profiling of human chromosomes 6, 20 and 22. *Nat Genet*, 38(12), 1378-1385.
- Elango, N., Hunt, B. G., Goodisman, M. A., & Yi, S. V. (2009). DNA methylation is widespread and associated with differential gene expression in castes of the honeybee, *Apis mellifera*. *Proceedings of the National Academy of Sciences*, 106(27), 11206.
- Fang, F., Xu, Y., Jones, D., & Jones, G. (2005). Interactions of ultraspiracle with ecdysone receptor in the transduction of ecdysone-and juvenile hormone-signaling. *FEBS Journal*, 272(7), 1577-1589.
- Farnham, P. J. (2009). Insights from genomic profiling of transcription factors. *Nature Reviews Genetics*, 10(9), 605-616.
- Field, L. M. (2000). Methylation and expression of amplified esterase genes in the aphid *Myzus persicae* (Sulzer). *Biochemical Journal*, 349(Pt 3), 863-868.
- Field, L. M., Lyko, F., Mandrioli, M., & Pranter, G. (2004). DNA methylation in insects. *Insect Molecular Biology*, 13(2), 109-115.

- Fielenbach, N., & Antebi, A. (2008). *C. elegans* dauer formation and the molecular basis of plasticity. *Genes & Development*, 22(16), 2149-2165.
- Flusberg, B. A., Webster, D. R., Lee, J. H., Travers, K. J., Olivares, E. C., Clark, T. A., Korlach, J., et al. (2010). Direct detection of DNA methylation during single-molecule, real-time sequencing. *Nat Meth*, 7(6), 461-465.
- Foret, S., Kucharski, R., Pittelkow, Y., Lockett, G., & Maleszka, R. (2009). Epigenetic regulation of the honey bee transcriptome: unravelling the nature of methylated genes. *BMC Genomics*, 10(1), 472.
- Frigola, J., Ribas, M., Risques, R. A., & Peinado, M. A. (2002). Methylome profiling of cancer cells by amplification of inter-methylated sites (AIMS). *Nucleic acids research*, 30(7), e28.
- Gauhar, Z., Sun, L. V., Hua, S., Mason, C. E., Fuchs, F., Li, T. R., Boutros, M., et al. (2009). Genomic mapping of binding regions for the Ecdysone receptor protein complex. *Genome research*, 19(6), 1006.
- Géminard, C., Rulifson, E. J., & Léopold, P. (2009). Remote control of insulin secretion by fat cells in *Drosophila*. *Cell Metabolism*, 10(3), 199–207.
- Goll, M. G., & Bestor, T. H. (2005). Eukaryotic cytosine methyltransferases. *Annual Review of Biochemistry*, 74, 481-514.
- Goll, M. G., Kirpekar, F., Maggert, K. A., Yoder, J. A., Hsieh, C. L., Zhang, X., Golic, K. G., et al. (2006). Methylation of tRNA<sup>Asp</sup> by the DNA methyltransferase homolog Dnmt2. *Science*, 311(5759), 395.
- Grozinger, C. M., Sharabash, N. M., Whitfield, C. W., & Robinson, G. E. (2003). Pheromone-mediated gene expression in the honey bee brain. *Proceedings of the National Academy of Sciences of the United States of America*, 100 Suppl 2, 14519-14525.
- Gutierrez, E., Wiggins, D., Fielding, B., & Gould, A. P. (2007). Specialized hepatocyte-like cells regulate *Drosophila* lipid metabolism. *Nature*, 445(7125), 275-280.
- Hagedorn, H. H. (1983). The role of ecdysteroids in the adult insect. *Endocrinology of insects*, 1.
- Haig, D. (2000). The kinship theory of imprinting. *Annu. Rev. Ecol. Syst*, 31, 9–32.
- Hammer, M., & Menzel, R. (1995). Learning and memory in the honeybee. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 15(3 Pt 1), 1617-1630.
- Hamza, M. S., Pott, S., Vega, V. B., Thomsen, J. S., Kandhadayar, G. S., Ng, P. W., Chiu, K. P., et al. (2009). De-novo identification of PPARgamma/RXR binding sites and direct targets during adipogenesis. *PLOS one*, 4(3).
- Hark, A. T., Schoenherr, C. J., Katz, D. J., Ingram, R. S., Levorse, J. M., & Tilghman, S. M. (2000). CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. *Nature*, 405(6785), 486-489.
- Harrison, J. M. (1986). Caste-specific changes in honeybee flight capacity. *Physiological zoology*, 175–187.
- Hartfelder, K., Bitondi, M. M. G., Santana, W. C., & Simoes, Z. L. P. (2002). Ecdysteroid titer and reproduction in queens and workers of the honey bee and of a stingless bee: loss of ecdysteroid function at increasing levels of sociality? *Insect biochemistry and molecular biology*, 32(2), 211–216.

- Hartfelder, K. (2005). Vitellogenin regulates hormonal dynamics in the worker caste of a eusocial insect. *FEBS letters*, 579(22), 4961–4965.
- Haslam, D. W., & James, W. P. T. (2005). Obesity. *Lancet*, 366(9492), 1197–1209.
- Hatzis, P., Van Der Flier, L. G., Van Driel, M. A., Guryev, V., Nielsen, F., Denissov, S., Nijman, I. J., et al. (2008). Genome-wide pattern of TCF7L2/TCF4 chromatin occupancy in colorectal cancer cells. *Molecular and cellular biology*, 28(8), 2732.
- Hendrich, B., & Tweedie, S. (2003). The methyl-CpG binding domain and the evolving role of DNA methylation in animals. *Trends in Genetics: TIG*, 19(5), 269–277.
- Hölldobler, B., & Wilson, E. O. (2008). *The Superorganism: The Beauty, Elegance, and Strangeness of Insect Societies* (1st ed.). W. W. Norton & Company.
- Honey Bee Genome Sequencing Consortium. (2006). Insights into social insects from the genome of the honeybee *Apis mellifera*. *Nature*, 443(7114), 931–949.
- Huang, Z. Y., Robinson, G. E., & Borst, D. W. (1994). Physiological correlates of division of labor among similarly aged honey bees. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, 174(6), 731–739.
- Huang, D. W., Sherman, B. T., & Lempicki, R. A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols*, 4(1), 44–57.
- Issa, J. P., Vertino, P. M., Wu, J., Sazawal, S., Celano, P., Nelkin, B. D., Hamilton, S. R., et al. (1993). Increased cytosine DNA-methyltransferase activity during colon cancer progression. *Journal of the National Cancer Institute*, 85(15), 1235–1240.
- Jaycox, E. R. (1976). Behavioral Changes in Worker Honey Bees (*Apis mellifera* L.) after Injection with Synthetic Juvenile Hormone (Hymenoptera: Apidae). *Journal of the Kansas Entomological Society*, 49(2), 165–170.
- Johnson, D. S., Mortazavi, A., Myers, R. M., & Wold, B. (2007). Genome-wide mapping of in vivo protein-DNA interactions. *Science*, 316(5830), 1497.
- Jones, G., & Sharp, P. A. (1997). Ultraspiracle: an invertebrate nuclear receptor for juvenile hormones. *Proceedings of the National Academy of Sciences of the United States of America*, 94(25), 13499.
- Kaatz, H., Eichmüller, S., & Kreissl, S. (1994). Stimulatory effect of octopamine on juvenile hormone biosynthesis in honey bees (*Apis mellifera*): physiological and immunocytochemical evidence. *Journal of insect physiology*, 40(10), 865–872.
- Kaatz, H. H., Hagedorn, H. H., & Engels, W. (1985). Culture of honeybee organs: development of a new medium and the importance of tracheation. *In Vitro Cellular & Developmental Biology-Plant*, 21(6), 347–352.
- Kahn, H. S. (2001). Glucose tolerance in adults after prenatal exposure to famine. *Lancet*, 357(9270), 1798–1799.
- Kanehisa, M., & Goto, S. (2000). KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic acids research*, 28(1), 27.
- Karpf, A. R., & Matsui, S. (2005). Genetic Disruption of Cytosine DNA Methyltransferase Enzymes Induces Chromosomal Instability in Human Cancer Cells. *Cancer Research*, 65(19), 8635–8639.
- Kaun, K. R., Chakaborty-Chatterjee, M., & Sokolowski, M. B. (2008). Natural variation in plasticity of glucose homeostasis and food intake. *Journal of Experimental Biology*, 211(19), 3160.

- Kaun, K. R., Riedl, C. A., Chakaborty-Chatterjee, M., Belay, A. T., Douglas, S. J., Gibbs, A. G., & Sokolowski, M. B. (2007). Natural variation in food acquisition mediated via a *Drosophila* cGMP-dependent protein kinase. *Journal of Experimental Biology*, 210(20), 3547.
- Kim, J., Cunningham, R., James, B., Wyder, S., Gibson, J. D., Niehuis, O., Zdobnov, E. M., et al. (2010). Functional Characterization of Transcription Factor Motifs Using Cross-species Comparison across Large Evolutionary Distances. *PLoS Comput Biol*, 6(1).
- Kim, T. H., Barrera, L. O., Zheng, M., Qu, C., Singer, M. A., Richmond, T. A., Wu, Y., et al. (2005). A high-resolution map of active promoters in the human genome. *Nature*, 436(7052), 876–880.
- King-Jones, K., & Thummel, C. S. (2005). Nuclear receptors—a perspective from *Drosophila*. *Nature Reviews Genetics*, 6(4), 311–323.
- Klingenspor, M., Niggemann, H., & Heldmaier, G. (2000). Modulation of leptin sensitivity by short photoperiod acclimation in the Djungarian hamster, *Phodopus sungorus*. *Journal of Comparative Physiology B: Biochemical, Systemic, and Environmental Physiology*, 170(1), 37–43.
- Klose, R. J., & Bird, A. P. (2006). Genomic DNA methylation: the mark and its mediators. *Trends in Biochemical Sciences*, 31(2), 89–97.
- Konopova, B., & Jindra, M. (2007). Juvenile hormone resistance gene Methoprene-tolerant controls entry into metamorphosis in the beetle *Tribolium castaneum*. *Proceedings of the National Academy of Sciences*, 104(25), 10488.
- Kucharski, R., Maleszka, J., Foret, S., & Maleszka, R. (2008). Nutritional Control of Reproductive Status in Honeybees via DNA Methylation. *Science*, 319(5871), 1827–1830.
- Kunert, N., Marhold, J., Stanke, J., Stach, D., & Lyko, F. (2003). A Dnmt2-like protein mediates DNA methylation in *Drosophila*. *Development*, 130(21), 5083.
- Lee, T. I., Johnstone, S. E., & Young, R. A. (2006). Chromatin immunoprecipitation and microarray-based analysis of protein location. *Nature Protocols*, 1(2), 729–748.
- Lee, T. I., & Young, R. A. (2000). Transcription of eukaryotic protein-coding genes. *Annual review of genetics*, 34(1), 77–137.
- Lefterova, M. I., Zhang, Y., Steger, D. J., Schupp, M., Schug, J., Cristancho, A., Feng, D., et al. (2008). PPARgamma and C/EBP factors orchestrate adipocyte biology via adjacent binding on a genome-wide scale. *Genes & development*, 22(21), 2941.
- Leoncini, I., Le Conte, Y., Costagliola, G., Plettner, E., Toth, A. L., Wang, M., Huang, Z., et al. (2004). Regulation of behavioral maturation by a primer pheromone produced by adult worker honey bees. *Proceedings of the National Academy of Sciences*, 101(50), 17559.
- Leopold, P., & Perrimon, N. (2007). *Drosophila* and the genetics of the internal milieu. *Nature*, 450(7167), 186–188.
- Lezzi, M., Bergman, T., Henrich, V. C., Vöggtli, M., Frömel, C., Grebe, M., Przibilla, S., et al. (2009). Ligand-induced heterodimerization between the ligand binding domains of the *Drosophila* ecdysteroid receptor and ultraspiracle. *European Journal of Biochemistry*, 269(13), 3237–3245.

- Li, C., Kapitskaya, M. Z., Zhu, J., Miura, K., Segraves, W., & Raikhel, A. S. (2000). Conserved molecular mechanism for the stage specificity of the mosquito vitellogenic response to ecdysone. *Developmental Biology*, 224(1), 96–110.
- Li, Y., Zhang, Z., Robinson, G. E., & Palli, S. R. (2007). Identification and characterization of a juvenile hormone response element and its binding proteins. *Journal of Biological Chemistry*, 282(52), 37605.
- Mangelsdorf, D. J., Evans, R. M., & others. (1995). The RXR heterodimers and orphan receptors. *Cell*, 83(6), 841–850.
- Maori, E., Paldi, N., Shafir, S., Kalev, H., Tsur, E., Glick, E., & Sela, I. (2009). IAPV, a bee-affecting virus associated with colony collapse disorder can be silenced by dsRNA ingestion. *Insect molecular biology*, 18(1), 55–60.
- Marco Antonio, D. S., Guidugli-Lazzarini, K. R., do Nascimento, A. M., Simões, Z. L. P., & Hartfelder, K. (2008). RNAi-mediated silencing of *vitellogenin* gene function turns honeybee (*Apis mellifera*) workers into extremely precocious foragers. *Die Naturwissenschaften*, 95(10), 953–961.
- Marhold, J., Kramer, K., Kremmer, E., & Lyko, F. (2004). The *Drosophila* MBD2/3 protein mediates interactions between the MI-2 chromatin complex and CpT/A-methylated DNA. *Development*, 131(24), 6033.
- Marhold, J., Brehm, A., & Kramer, K. (2004). The *Drosophila* methyl-DNA binding protein MBD2/3 interacts with the NuRD complex via p55 and MI-2. *BMC Molecular Biology*, 5(1), 20.
- Markiewicz, D. A., & O'Donnell, S. (2001). Social dominance, task performance and nutrition: implications for reproduction in eusocial wasps. *Journal of Comparative Physiology. A, Sensory, Neural, and Behavioral Physiology*, 187(5), 327–333.
- Matys, V., Fricke, E., Geffers, R., Gössling, E., Haubrock, M., Hehl, R., Hornischer, K., et al. (2003). TRANSFAC: transcriptional regulation, from patterns to profiles. *Nucleic Acids Research*, 31(1), 374–378.
- Mayoral, J. G., Nouzova, M., Yoshiyama, M., Shinoda, T., Hernandez-Martinez, S., Dolgih, E., Turjanski, A. G., et al. (2009). Molecular and functional characterization of a juvenile hormone acid methyltransferase expressed in the corpora allata of mosquitoes. *Insect biochemistry and molecular biology*, 39(1), 31–37.
- Miele, A., Gheldof, N., Tabuchi, T. M., Dostie, J., & Dekker, J. (2006). Mapping chromatin interactions by chromosome conformation capture. *Current protocols in molecular biology*, 21(11).
- Moorman, C., Sun, L. V., Wang, J., De Wit, E., Talhout, W., Ward, L. D., Greil, F., et al. (2006). Hotspots of transcription factor colocalization in the genome of *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences*, 103(32), 12027.
- Mootha, V. K., Bunkenborg, J., Olsen, J. V., Hjerrild, M., Wisniewski, J. R., Stahl, E., Bolouri, M. S., et al. (2003). Integrated analysis of protein composition, tissue diversity, and gene regulation in mouse mitochondria. *Cell*, 115(5), 629–640.
- Morton, G. J., Cummings, D. E., Baskin, D. G., Barsh, G. S., & Schwartz, M. W. (2006). Central nervous system control of food intake and body weight. *Nature*, 443(7109), 289–295.

- Nelson, C. M., Ihle, K. E., Fondrk, M. K., Page, R. E., & Amdam, G. V. (2007). The gene vitellogenin has multiple coordinating effects on social organization. *PLoS Biology*, 5(3), e62.
- Nielsen, R., Pedersen, T., Hagenbeek, D., Moulos, P., Siersbaek, R., Megens, E., Denissov, S., et al. (2008). Genome-wide profiling of PPARgamma: RXR and RNA polymerase II occupancy reveals temporal activation of distinct metabolic pathways and changes in RXR dimer composition during adipogenesis. *Genes & development*, 22(21), 2953.
- Orgad, S., Nelson, H., Segal, D., & Nelson, N. (1998). Metal ions suppress the abnormal taste behavior of the *Drosophila* mutant malvolio. *Journal of Experimental Biology*, 201(1), 115.
- Osborne, K. A., Robichon, A., Burgess, E., Butland, S., Shaw, R. A., Coulthard, A., Pereira, H. S., et al. (1997). Natural behavior polymorphism due to a cGMP-dependent protein kinase of *Drosophila*. *Science*, 277(5327), 834.
- Page Jr, R. E., & Amdam, G. V. (2007). The making of a social insect: developmental architectures of social design. *BioEssays: news and reviews in molecular, cellular and developmental biology*, 29(4), 334.
- Page, R. E., Erber, J., & Fondrk, M. K. (1998). The effect of genotype on response thresholds to sucrose and foraging behavior of honey bees (*Apis mellifera* L.). *Journal of Comparative Physiology. A, Sensory, Neural, and Behavioral Physiology*, 182(4), 489-500.
- Pankiw, T., & Page, R. E. (2001). Genotype and colony environment affect honeybee (*Apis mellifera* L.) development and foraging behavior. *Behavioral Ecology and Sociobiology*, 51(1), 87-94.
- Patel, A., Fondrk, M. K., Kaftanoglu, O., Emore, C., Hunt, G., Frederick, K., & Amdam, G. V. (2007). The making of a queen: TOR pathway is a key player in diphenic caste development. *PLoS One*, 2(6).
- Perissi, V., & Rosenfeld, M. G. (2005). Controlling nuclear receptors: the circular logic of cofactor cycles. *Nature Reviews Molecular Cell Biology*, 6(7), 542-554.
- Puig, O., Marr, M. T., Ruhf, M. L., & Tjian, R. (2003). Control of cell number by *Drosophila* FOXO: downstream and feedback regulation of the insulin receptor pathway. *Genes & development*, 17(16), 2006.
- Queller, D. C. (2003). Theory of genomic imprinting conflict in social insects. *BMC Evolutionary Biology*, 3, 15. doi:10.1186/1471-2148-3-15
- Raikhel, A. S., Kokoza, V. A., Zhu, J., Martin, D., Wang, S. F., Li, C., Sun, G., et al. (2002). Molecular biology of mosquito vitellogenesis: from basic studies to genetic engineering of antipathogen immunity. *Insect biochemistry and molecular biology*, 32(10), 1275-1286.
- Rakyan, V. K., Hildmann, T., Novik, K. L., Lewin, J., Tost, J., Cox, A. V., Andrews, T. D., et al. (2004). DNA methylation profiling of the human major histocompatibility complex: a pilot study for the human epigenome project. *PLoS Biology*, 2(12), e405.
- Ramsahoye, B. H. (2002). Measurement of genome-wide DNA cytosine-5 methylation by reversed-phase high-pressure liquid chromatography. *Methods in Molecular Biology*, 200, 17-27.

- Ravelli, G. P., Stein, Z. A., & Susser, M. W. (1976). Obesity in young men after famine exposure in utero and early infancy. *The New England Journal of Medicine*, 295(7), 349-353.
- Renn, S. C. P., Aubin-Horth, N., & Hofmann, H. A. (2008). Fish and chips: functional genomics of social plasticity in an African cichlid fish. *The Journal of Experimental Biology*, 211(Pt 18), 3041-3056.
- Riddiford, L. M. (2008). Juvenile hormone action: a 2007 perspective. *Journal of insect physiology*, 54(6), 895-901.
- Riddiford, L. M. (1985). Hormone action at the cellular level. *Comprehensive insect physiology, biochemistry and pharmacology*, 8, 37-84.
- Robinson, G. E. (1992). Regulation of division of labor in insect societies. *Annual Review of Entomology*, 37, 637-665.
- Robinson, G. E. (1985). Effects of a juvenile hormone analogue on honey bee foraging behaviour and alarm pheromone production. *Journal of Insect Physiology*, 31(4), 277-282.
- Robinson, G. E. (1987). Regulation of honey bee age polyethism by juvenile hormone. *Behavioral Ecology and Sociobiology*, 20(5), 329-338.
- Robinson, G. E., Page Jr, R. E., Strambi, C., & Strambi, A. (1989). Hormonal and genetic control of behavioral integration in honey bee colonies. *Science*, 246(4926), 109.
- Robinson, G. E. (2002). Genomics and integrative analyses of division of labor in honeybee colonies. *The American Naturalist*, 160 Suppl 6, S160-172.
- Robinson, G. E., Fernald, R. D., & Clayton, D. F. (2008). Genes and social behavior. *Science (New York, N.Y.)*, 322(5903), 896-900.
- Rodrigues, V., Cheah, P. Y., Ray, K., & Chia, W. (1995). malvolio, the *Drosophila* homologue of mouse NRAMP-1 (Bcg), is expressed in macrophages and in the nervous system and is required for normal taste behaviour. *The EMBO Journal*, 14(13), 3007.
- Roeder, T. (1999). Octopamine in invertebrates. *Progress in neurobiology*, 59(5), 533-561.
- Rozowsky, J., Euskirchen, G., Auerbach, R. K., Zhang, Z. D., Gibson, T., Bjornson, R., Carriero, N., et al. (2009). PeakSeq enables systematic scoring of ChIP-seq experiments relative to controls. *Nature biotechnology*, 27(1), 66-75.
- Rutz, W., Gerig, L., Wille, H., & Lüscher, M. (1976). The function of juvenile hormone in adult worker honeybees, *Apis mellifera*. *Journal of Insect Physiology*, 22(11), 1485-1491.
- Rutz, W., & Lüscher, M. (1974). The occurrence of vitellogenin in workers and queens of *Apis mellifica* and the possibility of its transmission to the queen. *Journal of Insect Physiology*, 20(5), 897-909.
- Sahara, K., Marec, F., & Traut, W. (1999). TTAGG telomeric repeats in chromosomes of some insects and other arthropods. *Chromosome Research: An International Journal on the Molecular, Supramolecular and Evolutionary Aspects of Chromosome Biology*, 7(6), 449-460.
- Scheiner, R., Plückhahn, S., Öney, B., Blenau, W., & Erber, J. (2002). Behavioural pharmacology of octopamine, tyramine and dopamine in honey bees. *Behavioural brain research*, 136(2), 545-553.

- Scheiner, S. M. (1993). Genetics and evolution of phenotypic plasticity. *Annual Review of Ecology and Systematics*, 24(1), 35–68.
- Schulz, D. J., & Robinson, G. E. (2001). Octopamine influences division of labor in honey bee colonies. *Journal of Comparative Physiology. A, Sensory, Neural, and Behavioral Physiology*, 187(1), 53–61.
- Schulz, D. J., Huang, Z. Y., & Robinson, G. E. (1998). Effects of colony food shortage on behavioral development in honey bees. *Behavioral Ecology and Sociobiology*, 42(5), 295–303.
- Schulz, D. J., Sullivan, J. P., & Robinson, G. E. (2002). Juvenile hormone and octopamine in the regulation of division of labor in honey bee colonies. *Hormones and behavior*, 42(2), 222–231.
- Seeley, T. D. (1996). *The Wisdom of the Hive: The Social Physiology of Honey Bee Colonies*. Harvard University Press.
- Sheng, Z., Ma, L., Cao, M. X., Jiang, R. J., & Li, S. (2008). Juvenile hormone acid methyl transferase is a key regulatory enzyme for juvenile hormone synthesis in the Eri silkworm, *Samia cynthia ricini*. *Archives of Insect Biochemistry and Physiology*, 69(3), 143–154.
- Silverstein, R. A., & Ekwall, K. (2005). Sin3: a flexible regulator of global gene expression and genome stability. *Current Genetics*, 47(1), 1–17.
- Sinha, S., Ling, X., Whitfield, C. W., Zhai, C., & Robinson, G. E. (2006). Genome scan for cis-regulatory DNA motifs associated with social behavior in honey bees. *Proceedings of the National Academy of Sciences of the United States of America*, 103(44), 16352–16357.
- Smith, C. R., Toth, A. L., Suarez, A. V., & Robinson, G. E. (2008). Genetic and genomic analyses of the division of labour in insect societies. *Nature Reviews. Genetics*, 9(10), 735–748.
- Sonoda, J., Pei, L., & Evans, R. M. (2008). Nuclear receptors: decoding metabolic disease. *FEBS letters*, 582(1), 2–9.
- Stirzaker, C., Song, J. Z., Davidson, B., & Clark, S. J. (2004). Transcriptional gene silencing promotes DNA hypermethylation through a sequential change in chromatin modifications in cancer cells. *Cancer Research*, 64(11), 3871–3877.
- Sullivan, A. A., & Thummel, C. S. (2003). Temporal profiles of nuclear receptor gene expression reveal coordinate transcriptional responses during *Drosophila* development. *Molecular Endocrinology*, 17(11), 2125.
- Sullivan, J. P., Fahrbach, S. E., Harrison, J. F., Capaldi, E. A., Fewell, J. H., & Robinson, G. E. (2003). Juvenile hormone and division of labor in honey bee colonies: effects of allatectomy on flight behavior and metabolism. *Journal of Experimental Biology*, 206(13), 2287.
- Sullivan, J. P., Jassim, O., Fahrbach, S. E., & Robinson, G. E. (2000). Juvenile hormone paces behavioral development in the adult worker honey bee. *Hormones and behavior*, 37(1), 1–14.
- Suzuki, M. M., Kerr, A. R. W., De Sousa, D., & Bird, A. (2007). CpG methylation is targeted to transcription units in an invertebrate genome. *Genome Research*, 17(5), 625–631.
- Suzuki, M. M., & Bird, A. (2008). DNA methylation landscapes: provocative insights from epigenomics. *Nature Reviews Genetics*, 9(6), 465–476.

- Toth, A. L., & Robinson, G. E. (2005). Worker nutrition and division of labour in honeybees. *Animal Behaviour*, 69(2), 427–435.
- Toth, A. L., Kantarovich, S., Meisel, A. F., & Robinson, G. E. (2005). Nutritional status influences socially regulated foraging ontogeny in honey bees. *The Journal of Experimental Biology*, 208(Pt 24), 4641–4649.
- Tu, M. P., Yin, C. M., & Tatar, M. (2005). Mutations in insulin signaling pathway alter juvenile hormone synthesis in *Drosophila melanogaster*. *General and comparative endocrinology*, 142(3), 347–356.
- Tups, A., Helwig, M., Stohr, S., Barrett, P., Mercer, J. G., & Klingenspor, M. (2006). Photoperiodic regulation of insulin receptor mRNA and intracellular insulin signaling in the arcuate nucleus of the Siberian hamster, *Phodopus sungorus*. *American Journal of Physiology- Regulatory, Integrative and Comparative Physiology*, 291(3), R643.
- Tweedie, S., Charlton, J., Clark, V., & Bird, A. (1997). Methylation of genomes and genes at the invertebrate-vertebrate boundary. *Molecular and Cellular Biology*, 17(3), 1469–1475.
- Tweedie, S., Charlton, J., Clark, V., & Bird, A. (1997). Methylation of genomes and genes at the invertebrate-vertebrate boundary. *Molecular and cellular biology*, 17(3), 1469.
- Ueda, H., Shinoda, T., & Hiruma, K. (2009). Spatial expression of the mevalonate enzymes involved in juvenile hormone biosynthesis in the corpora allata in *Bombyx mori*. *Journal of insect physiology*.
- Velarde, R. A., Robinson, G. E., & Fahrbach, S. E. (2006). Nuclear receptors of the honey bee: annotation and expression in the adult brain. *Insect Molecular Biology*, 15(5), 583.
- Velarde, R. A., Robinson, G. E., & Fahrbach, S. E. (2009). Coordinated responses to developmental hormones in the Kenyon cells of the adult worker honey bee brain (*Apis mellifera* L.). *Journal of insect physiology*, 55(1), 59–69.
- Wade, G., & Schoneider, J. (1992). Metabolic fuels and reproduction in female mammals. *Neurosci Biobehav Rev*, 16, 235–272.
- Wade, P. A., Geggion, A., Jones, P. L., Ballestar, E., Aubry, F., & Wolffe, A. P. (1999). Mi-2 complex couples DNA methylation to chromatin remodelling and histone deacetylation. *Nature Genetics*, 23(1), 62–66.
- Wakabayashi, K., Okamura, M., Tsutsumi, S., Nishikawa, N., Tanaka, T., Sakakibara, I., Kitakami, J., et al. (2009). PPAR gamma/RXR alpha heterodimer targets gene of histone modification enzyme PR-Set7/Setd8 and regulates adipogenesis through a positive feedback loop. *Molecular and Cellular Biology*.
- Weaver, I. C. G., Cervoni, N., Champagne, F. A., D'Alessio, A. C., Sharma, S., Seckl, J. R., Dymov, S., et al. (2004). Epigenetic programming by maternal behavior. *Nature Neuroscience*, 7(8), 847–854.
- Wederell, E. D., Bilenky, M., Cullum, R., Thiessen, N., Dagpinar, M., Delaney, A., Varhol, R., et al. (2008). Global analysis of in vivo Foxa2-binding sites in mouse adult liver using massively parallel sequencing. *Nucleic Acids Research*.
- Weinmann, A. S., Bartley, S. M., Zhang, T., Zhang, M. Q., & Farnham, P. J. (2001). Use of chromatin immunoprecipitation to clone novel E2F target promoters. *Molecular and Cellular Biology*, 21(20), 6820.

- West-Eberhard, M. J. (2003). *Developmental Plasticity and Evolution* (1st ed.). Oxford University Press, USA.
- West-Eberhard, M. J. (1989). Phenotypic plasticity and the origins of diversity. *Annual Review of Ecology and Systematics*, 20(1), 249–278.
- Wheeler, D. E., Buck, N., & Evans, J. D. (2006). Expression of insulin pathway genes during the period of caste determination in the honey bee, *Apis mellifera*. *Insect Molecular Biology*, 15(5), 597.
- White, K. P., Hurban, P., Watanabe, T., & Hogness, D. S. (1997). Coordination of *Drosophila* metamorphosis by two ecdysone-induced nuclear receptors. *Science*, 276(5309), 114.
- Whitfield, C. W., Ben-Shahar, Y., Brillet, C., Leoncini, I., Crauser, D., Leconte, Y., Rodriguez-Zas, S., et al. (2006). Genomic dissection of behavioral maturation in the honey bee. *Proceedings of the National Academy of Sciences*, 103(44), 16068–16075.
- Whitfield, C. W., Cziko, A., & Robinson, G. E. (2003). Gene expression profiles in the brain predict behavior in individual honey bees. *Science*, 302(5643), 296–299.
- Williams, K. D., Busto, M., Suster, M. L., So, A. K., Ben-Shahar, Y., Leever, S. J., & Sokolowski, M. B. (2006). Natural variation in *Drosophila melanogaster* diapause due to the insulin-regulated PI3-kinase. *Proceedings of the National Academy of Sciences*, 103(43), 15911.
- Winston, M. L. (1991). *The Biology of the Honey Bee*. Harvard University Press.
- Workman, C. T., Mak, H. C., McCuine, S., Tagne, J. B., Agarwal, M., Ozier, O., Begley, T. J., et al. (2006). A systems approach to mapping DNA damage response pathways. *Science*, 312(5776), 1054.
- Wu, Q., & Brown, M. R. (2005). Signaling and function of insulin-like peptides in insects.
- Wu, Q., Zhao, Z., & Shen, P. (2005). Regulation of aversion to noxious food by *Drosophila* neuropeptide Y–and insulin-like systems. *Nature neuroscience*, 8(10), 1350–1355.
- Xi, H., Shulha, H. P., Lin, J. M., Vales, T. R., Fu, Y., Bodine, D. M., McKay, R. D., et al. (2007). Identification and characterization of cell type-specific and ubiquitous chromatin regulatory structures in the human genome. *PLoS Genet*, 3(8), 1377–1388.
- Yao, T. P., Forman, B. M., Jiang, Z., Cherbas, L., Chen, J. D., McKeown, M., Cherbas, P., et al. (1993). Functional ecdysone receptor is the product of EcR and Ultraspiracle genes. *Nature*, 366, 476–479.
- Zhu, J., Miura, K., Chen, L., & Raikhel, A. S. (2000). AHR38, a homolog of NGFI-B, inhibits formation of the functional ecdysteroid receptor in the mosquito *Aedes aegypti*. *The EMBO Journal*, 19(2), 253–262.
- Zhu, J., Chen, L., & Raikhel, A. S. (2003). Posttranscriptional control of the competence factor  $\beta$ FTZ-F1 by juvenile hormone in the mosquito *Aedes aegypti*. *Proceedings of the National Academy of Sciences*, 100(23), 13338–13343.
- Zhu, J., Miura, K., Chen, L., & Raikhel, A. S. (2003). Cyclicity of mosquito vitellogenic ecdysteroid-mediated signaling is modulated by alternative dimerization of the RXR homologue Ultraspiracle. *Proceedings of the National Academy of Sciences*, 100(2), 544–549.

## **Appendix A**

### **ChIP-PCR Confirmation and Additional Antibody Testing Using ChIP-PCR**

#### **ChIP-PCR confirmation of ChIP-chip results**

ChIP-PCR assays were performed to verify ChIP-chip results. A different honey bee USP antibody that recognizes a different region of USP from the ChIP-chip antibody was used in these assays. Foragers from 3 typical colonies were used in 3 independent ChIP assays. 2 peak regions and 2 negative control regions from ChIP-chip results were detected in QPCR (Figure A1). QPCR results of peak regions were normalized with the average of QPCR results of the two negative control regions. So the relative binding is a ratio of enrichment of peak region over negative control regions.

For both peak regions, the relative binding is consistently higher in assays using anti-serum compared to that in assays using pre-bleed, suggesting specific antibodies in the anti-serum precipitated the DNA fragments of the peak regions. Interestingly, the differences of relative binding between anti-serum and pre-bleed are consistently bigger for peak region 1 than peak region 2, which is similar to ChIP-chip results. This binding difference suggests that USP binds to different genomic regions with different strength.

Statistical analysis revealed that there is a significant difference in relative binding between anti-serum and pre-bleed in peak 1 region (paired t-test,  $p$ -value  $< 0.05$ ). The difference is not significant in peak 2 region at the  $p$ -value cut off at 0.05, even though

the trend of the difference is consistent among 3 trials. It is likely that there is a consistent small difference, which is not detected with small sample size.

## **Using ChIP-chip assays to screen other honey bee USP antibodies**

The aforementioned 2 peak regions and 2 negative control regions were used to screen another 3 antibodies we have to test their specificities in ChIP-chip assays (Figure A2). One trial using antibodies 102 and 103 , respectively, and 3 trials using antibody 105 (a typical result is shown in Figure A2) were performed. Antibodies 102 and 105 showed no or reversed difference of relative binding between the anti-serum and pre-bleed, suggesting these antibodies do not function well in ChIP assays. Another antibody 103 showed expected trend. However, most of its relative binding value is below 1, meaning there is more binding in negative control regions compared to peak regions on absolute measures, which is a problematic sign for this antibody to be used in ChIP assays.

## **Antibody Information**

The antibody used for ChIP-PCR confirmation and the additional antibodies were produced in rabbits using purified USP protein fragment. The protein fragment contains the first 100 amino acids from the N-terminus of USP. The USP protein fragment has a his-tag on its N-terminus for purification and was expressed in E.coli. The column purified protein were sent to Cocalico Biologicals (Reamstown, PA) for antibody production. 4 rabbits were injected. Pre-bleed and anti-serum from the 4 rabbits were

harvested (numbered as UIC102, UIC103, UIC104 and UIC105). The antibodies were tested with western blots. UIC104 was tested in brain immunostaining and was used in ChIP-PCR confirmation.

## Tables

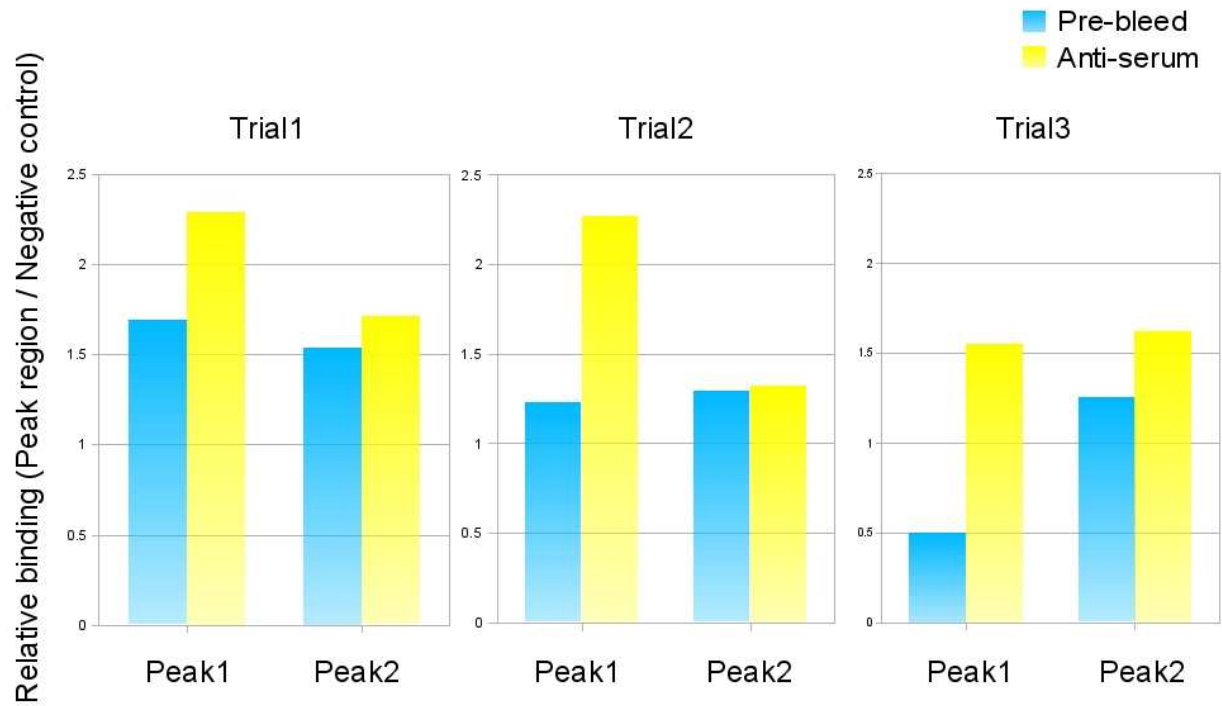
**TableA.1:** Genomic coordinates of peak regions and negative control regions used in ChIP-PCR.

| Region             | Group      | Start | End   |
|--------------------|------------|-------|-------|
| Peak1              | GROUP7.13  | 38312 | 38952 |
| Peak2              | GROUP14.21 | 26426 | 26967 |
| Negative Control 1 | GROUP1.17  | 1     | 4700  |
| Negative Control 2 | GROUP1.21  | 1000  | 5000  |

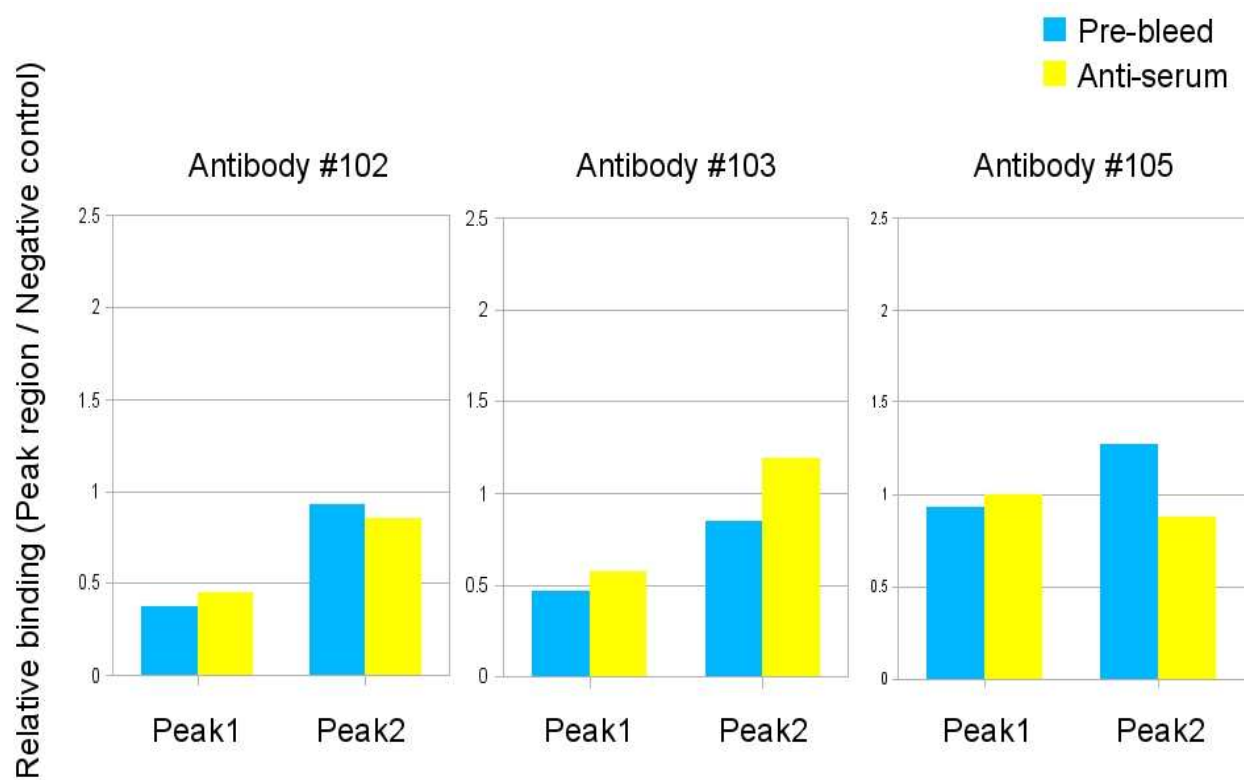
**TableA.2:** Primers used in ChIP-QPCR.

|                 |                            |
|-----------------|----------------------------|
| <b>Name</b>     | Group7.13_38313_38952-164F |
| <b>Sequence</b> | AAAGAGGCGAAGAAAGCAGGA      |
| <b>Name</b>     | Group7.13_38313_38952-233R |
| <b>Sequence</b> | TTTAGCCACACATCTCACCTT      |
| <b>Name</b>     | Group14.21_26626_26870-17F |
| <b>Sequence</b> | AGGAGATCCGGGATTCTAGAGAG    |
| <b>Name</b>     | Group14.21_26626_26870-86R |
| <b>Sequence</b> | TCACCGCCTTCCTTCGTAC        |
| <b>Name</b>     | ChIP_NC_1_1.17-3326F       |
| <b>Sequence</b> | TTCTGAAGCAGGAACTGGATTG     |
| <b>Name</b>     | ChIP_NC_1_1.17-3408R       |
| <b>Sequence</b> | CACTTCTGAGTCTGACCGGAATAA   |
| <b>Name</b>     | ChIP_NC_2_1.21-166F        |
| <b>Sequence</b> | GCACAAGGCTTCTGGAATTTG      |
| <b>Name</b>     | ChIP_NC_2_1.21-218R        |
| <b>Sequence</b> | TCGGAGGATCTTCTTAATTGCT     |

## Figures



**Figure A.1.** ChIP-PCR confirmation of two peak regions detected by ChIP-chip.



**Figure A.2.** ChIP-PCR results of the same peak regions using additional antibodies.

## Appendix B

### ChIP (ChIP-QPCR) Protocol

#### Day1

1. Turn on the water bath to let the water cooled to 4°C and make sure the water level is maintained at the indicated water level mark.

Supplement the Nuclear Lysis Buffer with protease inhibitors.

Make sure there is no SDS precipitates in the Nuclear Lysis Buffer.

Supplement the ice-cold dilute buffer with protease inhibitors.

Remove the stored sample from -80°C and put it on ice.

2. Re-suspend nuclei pellet in 0.31ml of Nuclear Lysis Buffer.

Sonicate: 10min, power High, 30 sec on, 30 sec off.

**NOTE:** it is always necessary to test the sonication conditions if the sample prepare method is changed.

3. Centrifuge samples for 10 mins at 13,000 rpm at 4C.

4. Put 140ul of the supernatant into each of 2 new silicon coated EP tubes.

Put 20ul of the supernatant into another EP tube as genomic DNA input, and store in -20C.

Add 1260ul of dilution buffer (supplemented with PI inhibitors) into the 2 tubes with supernatant in them.

5. Add 10ul of USP #104 anti-serum and 10ul #104 pre-bleed into the diluted cell lysis. Incubate with rotation at 4°C overnight.

---

Make Protein G magnetic particles slurry (500ml for 5 precipitations):

1. Take 0.5ml of Protein G magnetic particles from Dynal vial into EP tube.
2. Pellet beads by placing tube into magnetic particle separator stand for 30 sec, until solution is clear and brown particles stick to the wall.
3. Remove solution.
4. Remove tube from stand, re-suspend pellet in 0.5ml TE pH8.1. Rotate at RT for 15 min.
5. Pellet particles. Remove TE.
6. Repeat step 4, 5 for 2 more times.
7. Re-suspend protein G particles in 400ul TE 8.1 add 50ul BSA(10mg/ml) and 50ul of sonicated salmon sperm DNA (Stratagene 10mg/ml stock).
8. Rotate at 4°C for 1 hour before used in ChIP.

The prepared slurry can be stored for at least 2 month at 4°C. But I prefer to prepare them fresh very time.

## Day2

1. Add 100ul of salmon sperm DNA/BSA/Protein G magnetic particles and incubate 8-10 hours at 4°C with rotation.

2. Pellet the magnetic beads.

Wash beads in ice cold buffers 1ml each for 3-5 min at 4°C in the of order listed bellow:

- a. Low salt immune complex wash buffer
- b. High salt immune complex wash
- c. LiCl immune complex wash buffer
- d. TE buffer. Twice.

3. While washing the beads, make 10ml 1M NaHCO<sub>3</sub> . Vortex to make sure that here is no precipitates in the buffer. Set the heating block to 65°C. Put the input tube on ice to thaw.

4. Make Elution Buffer for all IP tubes as well as the input tube. For each tube prepare 400 ul of elution buffer: 20ul 20% SDS, 40ul 1M NaHCO<sub>3</sub> and 340ul sterile distilled water.

For 6 reactions (4 IP tubes and 2 inputs), prepare 140ul SDS, 280ul 1M NaHCO<sub>3</sub> and 2380ul sterile distilled water.

5. Add 380 ul of elution buffer into input (So 400ul in total) and set on ice.

6. Add 200ul of Elution buffer to each tube containing the magnetic beads, incubate at RT for 15 min.
7. Pellet, and collect the supernatant into new microfuge tubes.
8. Repeat step 6 and 7, this time heat the tubes on the heating block (65°C) to help elution. Now we have 400ul elution buffer in each tube.  
  
To all tubes, add 16ul 5M NaCl and incubate at 65C for 4-5 hour (Don't incubate it longer than 5 hr for ChIP-chip).
9. Wait until the temperature go back to RT, add to all tubes 1ul of Rnase A and incubate for 45min at 37°C.
10. Add 8ul 0.5M EDTA, 16ul 1M Tris-HCl and 1ul Proteinase K and incubate at 45°C for 1-2 hours.
11. Phenol/CCl4 extraction.
12. Dissolve in 50ul dd water.
13. Q-PCR

## Q-PCR

1. The input we get here is actually  $1/7$  of the should-be input. The aimed dilution of dilution serial is 0.1 input, 0.01 input, 0.001 input and 0.0001 input.
2. For 2 sets of standard curves, we need 11ul of 0.1 input.  $11 \times 0.1 \times 7 = 7.7\text{ul}$
3. So for 0.1 input: 7.7ul input + 3.4ul water.
4. Usual dilution of the inputs.