

DIFFERENTIAL EXPRESSION OF *FOR*, *FAX*, AND *U2AF* ORTHOLOGS AMONG THREE
TERMITE CASTES OF THE TERMITE, *RETICULITERMES FLAVIPES* (ISOPTERA:
RHINOTERMITIDAE)

by

JOSHUA RAYMOND URBAN

A.S., Cloud County Community College, 2005
B.S., Kansas Wesleyan University, 2007

A THESIS

submitted in partial fulfillment of the requirements for the degree

MASTER OF SCIENCE

Department of Entomology
College of Agriculture

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2010

Approved by:

Major Professor
Srini Kambhampati

Copyright

JOSHUA RAYMOND URBAN

2010

Abstract

Termites (Isoptera) are eusocial insects and exhibit highly complex eusocial behavior. Eusociality is characterized by the presence of castes (workers, soldiers, reproductives), polyphenisms (same genotype exhibiting multiple phenotypes), flexible developmental pathways, complex communication, cooperative brood care, construction and maintenance of complex nests, and division of labor. Previous studies on honey bees implicated several genes in caste-specific behavior; here, we investigate if orthologs of such genes are present in termites and if so, whether they are expressed differentially among the castes. A candidate gene approach using degenerate primers was used to amplify three candidate genes in the termite *Reticulitermes flavipes*. Quantitative real time PCR analysis revealed differential expression among termite workers, soldiers, and alates, with a general pattern of higher expression in alates. These results provide information on three novel genes in the termite *R. flavipes*.

Table of Contents

| | |
|--|------|
| List of Figures | vi |
| List of Tables | vii |
| Acknowledgements..... | viii |
| CHAPTER 1 - Expression of termite orthologs | 1 |
| Introduction..... | 1 |
| Sociogenomics | 1 |
| Termites | 2 |
| Materials and Methods..... | 5 |
| Termite Collection | 5 |
| Gene Selection and Primer Design | 5 |
| CODEHOP | 6 |
| RNA Isolation, cDNA Synthesis, PCR, and Sequencing | 6 |
| Sequence Analysis | 7 |
| Quantitative PCR | 7 |
| Caste-Specific Gene Expression..... | 7 |
| Individual Variation of Gene Expression | 9 |
| PCR Analysis of Head Specificity | 10 |
| Results..... | 12 |
| Amplification of Candidate Genes..... | 12 |
| Quantitative PCR for Caste-Specific Expression | 12 |
| cGMP-Dependent Protein Kinase:..... | 12 |
| Failed Axon Connections:..... | 13 |
| Nuclear Riboprotein Auxiliary Factor: | 13 |
| Expression of Candidate Genes lacks Head Specificity | 14 |
| Individual Variation in Gene Expression..... | 14 |
| Discussion..... | 16 |
| Individual Variation in Gene Expression..... | 16 |
| Caste-Specific Gene Expression | 18 |

| | |
|---|----|
| cGMP Dependent Protein Kinase: | 18 |
| Failed Axon Connections:..... | 22 |
| Nuclear Ribonucleoprotein Auxiliary Factor: | 24 |
| Figures and Tables | 27 |
| References..... | 34 |

List of Figures

- Figure 1.1 Quantitative real-time PCR analysis of *Rffor* (A and D), *Rffax* (B and E), and *RfU2AF* (C and F) expression in pooled heads of termite workers, soldiers, and alates from two unrelated colonies from Kansas (A, B, C) and Florida (D, E, F). Expression levels were calculated and normalized to the reference gene β -Actin. Values represent the means of three experimental replicates, with error bars representing variation in gene expression present between the three experimental replicates (10 heads per caste per pooled sample). Bars labeled with the same letter (a or b) are not significantly different from one another by the least squares difference test (LSD, $\alpha = 0.05$). 27
- Figure 1.2 Quantitative real-time PCR analysis of individual variation in gene expression of *Rffor* (A), *Rffax*(B), and *RfU2AF* (C). Expression was assessed in the heads of termite individuals from each of three termite castes of a Florida colony. Four biological replicates were performed, and the values indicated are the means of three technical replicates, with the error bars representing the standard experimental error present in the three replicates performed for each quantitative PCR sample. Data from individuals were combined to calculate the means of the four biological replicates to show caste-specific trends of *Rffor* (D), *Rffax* (E), and *RfU2AF* (F) expression. Error bars represent the natural variation in gene expression present in the four biological replicates (individuals). Bars labeled with the same letter (a or b) are not significantly different by the least squares difference test (LSD, $\alpha = 0.05$). 28
- Figure 1.3 Reverse transcriptase PCR analysis to determine whether the three candidate genes, *Rffor* (97 bp), *Rffax* (80 bp), and *RfU2AF* (82 bp) exhibit expression patterns localized to the tissues in the termite head. PCR was carried out using cDNA synthesized from total RNA from the bodies of those termites whose heads comprised the first pool of five in our Florida sample. PCR products were analyzed by electrophoresis on a 1% agarose gel in 1X TAE buffer and visualized under ultra-violet light using ethidium bromide staining. Three termite castes, workers, soldiers, and alates, were assayed, with all three showing successful amplification, indicating that the expression of these transcripts is not localized to the head, but rather that expression is present throughout the termite body, regardless of caste..... 29

List of Tables

| | |
|---|----|
| Table 1.1 Candidate genes were selected for amplification in the termite <i>R. flavipes</i> . Sequences for primer design were obtained from the NCBI database via honey bee accession numbers seen here. Honey bee expression levels for each gene are shown, as well as their putative function as defined by sequence similarity to known <i>D. melanogaster</i> genes. Relative honey bee gene expression levels are presented here as a ratio of forager levels to nurses. Levels for the foraging gene represent the average of three trials performed by Ben-Shahar et al. 2002..... | 30 |
| Table 1.2 Degenerate CODEHOP primers were designed for amplification of candidate genes in termites. Primers are listed in 5'-3' direction, with their expected fragment lengths provided. Two sets of primers are listed for those genes that required nested PCR for successful amplification. The first set represents the primers that were used for the outer PCR, and the second for the inner PCR. In addition, the primers for B-Actin, which served as our positive control in PCR reactions and as our reference gene for quantitative analysis, are listed..... | 31 |
| Table 1.3 Termite specific quantitative primers were designed from candidate gene DNA sequences gathered from degenerate PCR and sequencing experiments. Termite gene names are based on their similarity to known genes as determined through BLAST analysis (Table 1.4). Primers are listed in the 5'-3' direction and expected fragment lengths are indicated..... | 32 |
| Table 1.4 BLAST analysis of termite gene fragment sequences revealed their putative functions based on their similarity to homologous sequences from other insect species. Similarity is presented as the percent identity, with an e-value showing the significance of the match. The records indicated here represent the sequence hits with the highest values for each of the three candidate genes..... | 33 |

Acknowledgements

I would like to thank my major advisor, Dr. Srinivas Kambhampati for his aid in my project, as well as for his support throughout my program. I'd also like to thank my advisory committee members Dr. Jeremy Marshall and Dr. Ted Morgan for their insight and guidance throughout my research.

I would like to thank Dr. Michael Scharf (University of Florida) for supplying the *R. flavipes* for our Florida sampling in this study. I thank Xin Zhang, Dr. Yoonseong Park, and Juraj Koci for their help with quantitative PCR. I thank Paola Sotelo, Sharon Starkey, and Marimuthu Murugan for their help with SAS and the statistical analysis of my data.

I thank my friends Damian Tamayo, April Sutton, Diana Heustis, Matthew Steller, Michael Franklin, Blake Brewer, Briana Brewer, Lisa Brewer, Dave Brewer, Dane Baxa, Michael Tate, James Rivers, Dr. Art Neuberger, Dr. Weston Opitz, Leah Johnson, Gerald Turbiville, Jamie Pettus, Jason Harris, Kolin Morton, and Sam Danker for their presence in my life and their moral support throughout my degree.

Finally, I thank my family: my parents Joe and Carol Urban, my brother and sister-in-law Jeremy and Erin Urban, my sister and brother-in-law Audra and Jeff Cleveland, my brother and sister-in-law Michael and Katie Urban, my nieces and nephews, my mother-in-law Joni Wetzel, and my brother-in-law Jesse Wetzel, as well as the rest of my loved ones for their emotional support throughout my education. A special thanks to my brother Michael for his critical advice and constant voice of reason as I have been confronted with trial after trial throughout my graduate career. Michael, along with my uncle Richard Hubert, are silently responsible for my drive to better myself through education and to follow in their footsteps. Most importantly, I

thank my wife and love of my life, Leslie Urban. She has been with me through times of joy and sorrow and has always provided me with love and support. She is my strength in times and weakness and my friend in times of happiness. She has only shown support for my decisions during my graduate career and has been a blessing in my life. Last, I'd like to thank all of my loved ones who are no longer with me. They will always be missed, but have shaped me into the person I am today. I couldn't have done this without them. I love you all.

This work was supported by a grant from the Ecological Genomics Institute (ECOGEN) at Kansas State University, the Arthropod Genomics Center of Kansas State University, as well as by support from the National Science Foundation and Kansas State University's GK-12 program (EIDRoP).

CHAPTER 1 - Expression of termite orthologs

Introduction

Sociogenomics

Differential expression has become an important tool for discovering the underlying genetic aspects of social behavior. This is true of highly social insects, such as social bees and wasps, ants, and termites, which are divided into distinct morphological and behavioral phenotypes called castes that underlie division of labor (Evans and Wheeler, 1999). The castes found in honey bees and other social insect colonies exhibit polyphenic differences between individuals, which arise as a result of cues received during development and often involve many characters (Evans and Wheeler, 2001b). Much progress has been made to understand the genetic basis of caste formation in social insects in the area of gene expression, and studies on the honey bee have been important in elucidating aspects of neurology, behavior, learning, and memory (Whitfield et al., 2002; Denison and Raymond-Delpech, 2008). The roles of colony members are often determined by the events that occur during larval development, and studies on honey bees indicate that there are significant gene expression differences during caste differentiation (Evans and Wheeler, 1999; Evans and Wheeler, 2001a), behavioral development (Kucharski and Maleszka, 2002), and among individuals (Whitfield et al., 2003; Tsuchimoto et al., 2004).

In eusocial insects, social and environmental cues can alter gene expression in the brain to influence behavior, and genetic variation can influence brain function and social behavior (Robinson et al., 2008). For example, genes have previously been identified in honey bees that were characteristic of individuals performing certain caste-specific behaviors (Whitfield et al., 2003). The expression profiles of these genes in individual brains correctly predicted the

behavior (nurse or forager) in 95% of the bees, indicating a robust association between brain gene expression in the individual and caste-specific behavior (Whitfield et al. 2003). Social insects' genomes are highly sensitive to social influence (Robinson et al., 2005), and an individual's gene expression and/or behavior can change based on the needs of the colony. Whitfield et al. (2003) provide a great example of the relationship between social life and gene expression, in that the need for nurses in the honey bee colony leads to an alteration in gene expression, which, in turn, constitutes a change in the social behavior of each individual bee.

Social behaviors are likely very complex, involving many genes and gene networks, and are influenced by interactions among colony members, resulting in various behavioral and morphological changes. The studies in honey bees have clearly shown that there is a strong, tightly regulated control of gene expression influencing the caste system within a colony. However, an interesting, and as yet unanswered question is: do the same gene expression changes mediate social behavior across evolutionarily distant eusocial species?

Termites

Termites are eusocial insects in the Order Isoptera, of which there are about 2,600 species (Kambhampati and Eggleton, 2000). Termites exhibit highly complex eusocial behavior, including the presence of castes, polyphenisms, flexible developmental pathways, complex communication, symbiosis with protozoans and bacteria, construction and maintenance of complex nests, use of "agriculture," and other traits (Abe et al., 2000). They are hemimetabolous and, unlike eusocial hymenopterans (all ants, some bees and wasps), which exhibit haplodiploidy and whose young require continual care from adult workers, termite colonies consist of diploid males and females and they, along with the young, are recruited into the social hierarchy immediately (Miura et al., 2001; Zhou et al., 2006).

Termites and honey bees diverged nearly 300 million years ago, and the termites' complex social organization is believed to have arisen independently from that of the eusocial hymenopterans (Wilson, 1971; Thorne, 1997; Thorne and Traniello, 2003). Regardless of their novel beginnings, convergent evolution has resulted in very similar caste systems in the two groups (Wilson, 1971). However, comparative studies of eusocial termites to non-eusocial termites cannot be done as they can in Hymenoptera, as all extant species of termites are eusocial. Therefore, molecular and genomic studies comprise the most viable avenue by which to investigate the genetic basis of eusociality in termites.

Termite castes consist of workers, soldiers, reproductives, and various intermediate forms. Each of these castes has distinct behavior, physiology, and morphology (Vargo and Husseneder, 2009). In the termite *Reticulitermes flavipes*, workers are white, eyeless, and wingless individuals that are responsible for foraging, tunnel building, and brood care. Soldiers are similar in body size to workers, with the exception of the presence of a large, sclerotized head. They lack eyes and wings, but are fitted with a large pair of mandibles, which makes them well suited for their role in colony defense. The enlarged mandibles however, render them incapable of feeding and thus, they rely on the workers for nourishment. Reproductives have black, sclerotized bodies and possess eyes as well as wings. Alates, with their capacity to disperse, found new colonies and begin reproducing shortly thereafter (Scharf, 2003b).

As with honey bees and other social insects (Evans and Wheeler, 1999; Evans and Wheeler, 2001b), it is logical to assume that termite polyphenisms and associated developmental processes may also occur by means of differential gene expression (Miura et al., 1999; Miura, 2001). Although there are many resources from which one can gather information about honey bees and other model organisms, there is very little known about termites. An EST database of

genes from *R. flavipes* has been assembled (Steller et al., 2010); however, the absence of a fully sequenced genome greatly limits the progress into termite genomics. As such, whether orthologs of those genes associated with caste-specific behavior in honey bees and other social insects are also present in termites is not known. This lack of sequence information and functional characterization of genes for the termite *R. flavipes*, has led us to adopt the candidate gene approach, in order to learn more about these fascinating insects.

Here, our goal was to determine if homologs of selected genes, previously implicated in social behavior of honey bees (Ben-Shahar et al., 2002; Whitfield et al., 2003), are present and differentially expressed in the termite *R. flavipes*. We first set out to amplify selected honey bee candidate genes in termite workers, soldiers, and alates using degenerate PCR primers. Our second objective was to use quantitative real-time PCR to investigate whether the candidate genes exhibit caste-specific differential expression in the heads of termite workers, soldiers, and alates. Third, we sought to determine whether expression of the candidate genes was specific to the tissues in the head or if they were expressed elsewhere in the termite body. In these studies, we had three hypotheses: (i) that homologs of honey bee candidate genes would be present in termites, (ii) that those genes would exhibit caste-specific differential expression, and (iii) that the candidate genes would exhibit head specific expression.

Materials and Methods

Termite Collection

R. flavipes were collected from several different locations in Riley County Kansas. Some were collected on the Kansas State University campus, but most were collected in the natural wildlife area located at Pillsbury Crossing on Pillsbury Road near Manhattan, KS (39°07'47"N, 96°26'26"W). All termites were gathered from fallen, naturally infested wood and stored at -80°C. They were then placed on dry ice and separated from any wood and debris. They were identified and sorted by their castes (workers, soldiers, nymphs, larvae) and again stored at -80°C. Samples from a single Florida colony were provided by Michael E. Scharf of the University of Florida, Gainesville. Samples were kept at -80°C until needed.

Gene Selection and Primer Design

The selection of the candidate genes being investigated in this study was based upon those described by Whitfield et al. (2003). Whitfield et al. (2003) described a set of genes, whose combined expression predicted an individual honey bee's behavior (nursing or foraging). There were 17 genes identified in this study as being critical for caste-specific behavior in honey bees, and as having strong sequence similarity to functionally annotated *Drosophila melanogaster* genes. Of these, 12 were initially investigated for their presence in the termite *R. flavipes* due to their predictive nature of behavior in honey bees. In addition to these, this research also examines the presence and differential expression of *Amfor*, the foraging gene in the honey bee *Apis mellifera*, which, in a separate study, has been implicated in the behavioral shift from nursing to foraging in honey bees (Ben-Shahar et al., 2002).

CODEHOP

Consensus degenerate hybrid oligonucleotide primers (CODEHOPs) (Rose et al., 1998; Rose et al., 2003) were designed for the candidate genes to compensate for any base differences that may be present between *R. flavipes* sequence and that of other insects. This was done by first obtaining the candidate honey bee gene sequences from the National Center for Biotechnology Information's (NCBI) database via the accession numbers provided by Whitfield et al. (2003) (Table 1.1). Sequences were then run through NCBI's BLASTx to obtain homologous amino acid sequences from other insect species represented in the database. Multiple homologous insect sequences were then aligned using the online ClustalW2 (ClustalW2) sequence alignment program under its default settings. The multiple sequence alignment (MSA) files were then carved into blocks using the blocks multiple alignment processor (Multiple Alignment processor), and then imported into the CODEHOP program, adopting the codon usage table of *Blattella germanica*, to generate primers. Primers for the candidate genes were selected based on certain criteria; a G, C content of no more than 50-60%, a melting temp (T_m) of 50-65°C, low degeneracy, a high clamp score, and were selected so that they would amplify a region around 200 bases in length.

RNA Isolation, cDNA Synthesis, PCR, and Sequencing

For initial screening of candidate genes' presence in termites, PCR amplification was carried out using termites collected from Pillsbury Crossing, KS mentioned above. Total RNA, rather than DNA was targeted in order to ensure that amplification of candidate genes in termites was indicative of those genes actually being expressed in termites. Total RNA was isolated from termite workers' heads using Trizol® reagent and was followed by cDNA synthesis using the SuperScript™III First-Strand Synthesis System (Invitrogen). Polymerase chain reaction (PCR)

was carried out in 25µl volume reactions using CODEHOP primers (Table 1.2). Successful PCR programs were as follows: 95 °C for 3 min, then 35 cycles at 95°C for 30s, 50°C for 45s and 72°C for 30s, followed by 10 mins at 72°C. PCR products were then analyzed by electrophoresis on a 1% agarose gel in 1X TAE buffer and visualized under ultra-violet light using ethidium bromide staining. Bands of interest were excised from the gel and purified using a QIAquick® Gel Extraction Kit (Qiagen). Purified samples were then cloned into the pCR®2.1 vector, followed by transforming the construct into competent *Escherichia coli* using a TA Cloning kit (Invitrogen). Inserts from positive colonies (identified using x-gal blue-white selection) were amplified via PCR using M13 primers as follows: M13 forward (TCACACAGGAAACAGCT-ATGA) and M13 reverse (CGCCAGGGTTTTCCAGTC-ACGAC) shown 5' to 3'. Products were then visualized, purified, and sent to the Kansas State Genotyping and Sequencing Facility for sequencing in a 3730 DNA Analyzer (Applied Biosystems).

Sequence Analysis

The edges of vector sequence, obtained from the TA Cloning kit (Invitrogen) manual, were identified and trimmed from the sequence in a word processing program. Next, using the program Sequencher (v4.7, Gene Codes Corp.), sequences from multiple clones of a single DNA fragment were aligned to assure confidence in the sequencing results. This also allowed us to easily identify sequences of the complementary strands. Sequences were then put into NCBI's BLASTx to confirm identity to homologous genes in other insect species.

Quantitative PCR

Caste-Specific Gene Expression

Quantitative real-time PCR (qPCR) was performed on termite workers, soldiers, and alates from both a Kansas and a Florida colony to determine if the candidate genes exhibit caste-

specific differential gene expression. Termites from the Kansas colony were collected on the Kansas State University campus and those that made up the Florida colony were those provided by Michael E. Scharf of the University of Florida, Gainesville. Total RNA was isolated from pools of termite heads from each respective caste of each colony using Trizol® reagent. cDNA templates were then synthesized from these total RNA isolations using a SuperScript™ III First-Strand Synthesis System mentioned above. Termite specific primers for qPCR (Table 1.3) were designed using Beacon Designer (v7.5, Premier Biosoft International) and DNA fragment sequences obtained from CODEHOP amplification above. Specificity of quantitative primers was confirmed by PCR, followed by electrophoresis on a 1% agarose gel in 1X TAE buffer and visualized under ultra-violet light using ethidium bromide staining. PCR products were cloned and sequenced for confirmation of identity, using the same techniques described above.

qPCR was performed using iQ SYBR Green Supermix in a BIORAD iCycler real time PCR machine on 96-well plates with optical sealing tape. All reactions were carried out in triplicate, and in 25µl volumes using a PCR program that included an initial denaturation step of 95°C for 5 minutes, followed by 40 cycles of 95°C for 15 seconds, 55°C for 30 seconds, and 70°C for 30 seconds. Data collection and real-time analysis took place immediately following each cycle.

Template calibrations were performed using β-Actin expression levels and the $2^{-\Delta\Delta CT}$ method (Livak and Schmittgen, 2001). Individual primer efficiency tests were carried out using three 5-fold dilutions of an already 10-fold diluted worker cDNA template (.10x, .02x, .004x, .0008x), using three technical replicates for every sample. Primer efficiencies were determined by the iCycler software (v3.1.7050, Biorad Inc). For gene expression tests, three experimental replicates were performed (three pools of 10 individuals), each consisting of three technical

replicates per template per gene, and were run at the .02x template dilution. A Ct value (i.e., the PCR cycle number where amplification causes the amount of product to cross a set threshold) was calculated for each gene for each caste. Δ Ct values for candidate genes were calculated in relation to the most reliable reference gene β -Actin (Zhou et al., 2006), as $Ct_{\text{Target}} - Ct_{\beta\text{-Actin}}$, resulting in candidate gene expression values being corrected for by the amount of β -Actin in the sample (important note: a higher Ct value means a lower amount of gene product in the sample, as it takes more amplification cycles to cross the set threshold). The results are presented as the means of three experimental replicates, and the data for each sample (Kansas or Florida) were analyzed using PROC GLM in SAS (SAS Institute, 1985). Differences in the mean gene expression among castes within each population were evaluated for significance using the least squares difference tests (LSD) with $\alpha = 0.05$.

This protocol was followed for both the Kansas and Florida colonies, with one exception. cDNA for Kansas samples was synthesized from RNA isolated from pools of ten individuals, whereas the cDNA for Florida samples was synthesized from RNA isolated from pools containing the heads of five individuals. This difference was due simply to the limitation of resources.

Individual Variation of Gene Expression

Social information can have profound effects on the onset of behaviors with respect to the age, caste, and even sex of individuals (Robinson, 2008). As a result, genomes are highly sensitive to social influence (Robinson, 2005), and an individual's gene expression and or behavior can change based on the social environment. Consequently, individuals of the same caste may exhibit differential expression of many critical genes, including those of the candidate genes investigated here. For the purpose of this study, we examined the overall trends between

castes. For this reason, we opted to use pools of individuals, which provided us with an average level of gene expression across individuals of a given caste.

In addition to the pooled samples however, experiments were performed to assess the gene expression of individuals. This allowed for the investigation of the underlying cause of the error bars present in our pooled samples. Theoretically, the error bars represent the standard deviation between the three experimental replicates of pooled samples. In order to implicate individual variation in the cause of the error bars, rather than experimental error, this experiment identifies differential expression among termites of the same caste. For this, the experimental protocol was the same as described above with the exception that expression was analyzed in individual worker, soldier, and alate heads for each of the three candidate genes. Four biological replicates were performed, with termite samples from the same Florida colony described earlier. Again, Ct values for candidate genes were calculated in relation to the most reliable reference gene β -Actin.

PCR Analysis of Head Specificity

Upon completion of quantitative analysis of the candidate genes from the heads of the termites, additional PCR was performed to determine whether expression of the transcripts was localized to the head of the termites specifically. PCR was carried out using the bodies of those termites whose heads comprised the first pool of five in our Florida sample. Total RNA was isolated from those termite workers', soldiers', and alates' bodies (thorax + abdomen) using Trizol® reagent and was followed by cDNA synthesis as described above. Polymerase chain reaction (PCR) was carried out in 25 μ l volume reactions using the previously mentioned program with quantitative primers. PCR products were then analyzed by electrophoresis on a

1% agarose gel in 1X TAE buffer and visualized under ultra-violet light using ethidium bromide staining.

Results

Amplification of Candidate Genes

Of those genes whose presence was initially investigated in termites, 3 genes were successfully amplified. CODEHOPs were designed (Table 1.2) that successfully amplified a 220bp fragment of a cGMP-dependent protein kinase gene (*Rffor*), commonly referred to as the foraging gene, a 221 bp cDNA fragment of the *failed axon connections* gene (*Rffax*), and a 195bp fragment of a nuclear riboprotein auxiliary factor (*RfU2AF*) using termite worker cDNA synthesized from termite head total RNA. Sequencing, followed by BLAST analysis of these fragments revealed maximum identity values of 75% in *Acrythosiphon pisum*, 98% in *Culex quinquefasciatus*, and 83% in *Tribolium castaneum* respectively (Table 1.4). The three termite homologs were named *Rffor*, *Rffax*, and *RfU2AF* for their putative functions as revealed by their most prominent hits during BLAST analysis. Other insect species represented in the BLAST results include *Aedes aegypti*, *Bombyx mori*, *Anopheles gambiae*, *A. mellifera*, *Nasonia vitripennis*, and various *Drosophila* species. Initial amplification of candidate genes was in the worker caste only. Further reverse transcriptase PCR experiments confirmed the expression of these genes in soldiers and alates.

Quantitative PCR for Caste-Specific Expression

cGMP-Dependent Protein Kinase:

qPCR analysis of termites from a Kansas colony showed that the expression of *Rffor* varies among workers, soldiers, and alates. Relative to alates, *Rffor* levels were lower in both the worker and soldier castes (Fig. 1.1A). *Rffor* was expressed at higher levels in alates (7.74 ± 0.21 SEM) and workers (8.69 ± 0.19) compared to that in soldiers (9.54 ± 0.49). Statistical analysis

of the expression data revealed that *Rffor* levels in alates are significantly different from those in soldiers, whereas those levels present in workers are not significantly different from either soldiers or alates.

Rffor expression levels were also assessed in an unrelated Florida termite colony (Fig. 1.1D); however, no significant differences were observed between termite castes in the Florida samples.

Failed Axon Connections:

qPCR analysis of termites from a Kansas colony showed that the expression of *Rffax* is different among workers, soldiers, and alates (Fig. 1.1B). The highest reported level of expression (6.42 ± 0.13) was determined to be present in alates, and was significantly greater than those found in soldiers (7.79 ± 0.44). Those levels seen in workers (7.27 ± 0.16) showed no significant differences between either soldiers or alates.

Analysis of expression levels in the Florida population (Fig. 1.1E) revealed worker and alate levels (7.62 ± 0.28 and 7.64 ± 0.19 respectively) to be very similar, and consequently statistically insignificant. However, levels present in soldiers (9.03 ± 0.18) were significantly lower than both workers and alates.

Nuclear Riboprotein Auxiliary Factor:

Differential expression analysis of termites from a Kansas colony showed that the expression of *RfU2AF* is different among workers, soldiers, and alates (Fig. 1.1C). The highest reported level of expression (6.86 ± 0.31) was determined to be present in alates, and was significantly greater than those found in soldiers (8.17 ± 0.45). Those levels seen in workers (7.34 ± 0.10) showed no significant differences between either soldiers or alates.

Analysis of *RfU2AF* expression levels in the Florida population (Fig. 1.1F) revealed worker and alate levels (8.37 ± 0.14 and 8.37 ± 0.17 respectively) to be very similar, and consequently statistically insignificant. However, levels present in soldiers (9.82 ± 0.20) were significantly lower than those in both workers and alates.

Expression of Candidate Genes lacks Head Specificity

PCR was carried out to determine whether the three candidate genes exhibit expression patterns localized to the tissues in the termite head. When PCR was carried out using cDNA from the termite body, all three candidate genes were successfully amplified. Successful amplification was noted in workers, soldiers, and alates (Fig. 1.3), indicating that the expression of these transcripts is not localized to the head, but rather that expression is present throughout the termite body, regardless of caste.

Individual Variation in Gene Expression

Individual quantitative real-time PCR analysis was performed to identify whether variation in the pooled samples can be attributed to individual variation or experimental error. Analysis of the data generated by individual analysis revealed significant gene expression differences among individuals within the worker, soldier, and alate castes for all three of the candidate genes investigated here (Fig. 1.2 A-C). For example, looking at *Rffor* expression in workers, the first replicate exhibited expression of the transcript at levels nearly a quarter that in the other three. This was deemed significantly different from replicates two and three, as well as replicate four, which in turn was significantly different from replicates two and three. Similar degrees of significance can be seen in the soldiers and alates as well, as there are four levels of significance present in soldiers and three in alates.

Data from the expression analysis of individuals were also combined to generate caste-specific means, which were compared to those levels of expression achieved through caste-specific expression experiments. Analysis of the means for the individuals of each caste in the Florida population (Fig. 1.2 D-F) revealed worker and alate levels (2.59 ± 0.58 and 2.53 ± 0.68 respectively) to be very similar, and consequently statistically insignificant. However, levels present in soldiers (6.27 ± 1.14) were significantly lower than both workers and alates. We see that the trends of expression for *Rffor* are still in the same direction among castes as determined in the pooled individuals (Fig. 1.1 D-F). Soldiers still exhibit levels that are significantly lower than those present in either workers or alates, which are not significantly different from one another. Caste-specific levels of expression from the individual analysis exhibit similar trends to those in the pooled samples of the Florida colony for all three of the candidate genes.

Discussion

Sociogenomics finds its basis on two important insights. First is the idea that social life has a biological/genetic basis and is therefore influenced to some extent by genes and the forces of evolution (Wilson, 1975). Second is the realization that the molecular functions of many genes are highly conserved across species, even for complex traits (Robinson et al., 2005). Though much is known about the molecular aspects of differentiation in model organisms such as *D. melanogaster* and *A. mellifera*, there is little known about such processes in termites. Combining the candidate gene approach with degenerate PCR, as done here, molecular biologists are extending knowledge from model organisms to new study species around the world. Although this research has only focused on the presence and differential expression of a set of candidate genes, we can infer the function of these genes in termites based on molecular work in model organisms. However, it is important to note that the functions presented here represent hypotheses. No experiments were performed to identify gene functions, and it is possible that these genes may perform entirely different functions from those suggested here.

Individual Variation in Gene Expression

Results from the individual analysis provide a window into the individual variation in gene expression that is present within a population of individuals from the same colony. Those individuals investigated were members of the same caste and presumably exhibiting the same types of behavior. Yet, we see that there are significant differences among those individuals. Taking the means of those four individuals measured, we see that the caste-levels of expression exhibit trends that are similar to those found in our pooled individuals from the same Florida colony.

Though the trends between castes remain, it is important to note the difference witnessed in the gene expression data generated by the pooled Florida sample and the Florida individuals that were examined. In looking at the data generated by the quantitative analysis of individuals, we see that the levels of expression exhibit an increase compared to those in the pooled Florida samples for all castes and all three candidate genes. The larger Ct values (lower expression) seen in our pooled samples may be an artifact of the experimental design. β -Actin has been shown to be a reliable reference gene (Zhou et al., 2006), but we know that it is expressed at high levels. As a result, the pooled samples would contain a large volume of the transcript compared to other, less abundant transcripts in the cDNA. The higher levels of representation then would lead to a larger gap in the Ct values generated by the quantitative analysis, and thus lead to higher Δ Ct values. In the individual analysis however, the lack of pooling gives a more reliable comparison between β -Actin and target genes and thus, lower Δ Ct values. Our results from pooled samples and individuals suggest that, whereas using pooled individuals is sufficient for generating an overall picture of gene expression patterns, several individuals of the same caste must be assayed to obtain a full picture of inter-individual variation in gene expression levels. The source of variation, we suspect, arises due to social conditions within the colony, age of individuals, and sex of individuals.

While error bars in the analysis of individuals represent differences between technical replicates and thus, standard error, those present in our pooled results indicate the variation between our three experimental replicates. The individual quantitative analysis leads us to believe that the variation present between pools is a direct result of the variation among individuals from each of the three samplings. By using three replicates of ten, our sampling for

this study became 30 individuals. We believe this provides an adequate measure of the general trends present between the three castes investigated here.

Caste-Specific Gene Expression

cGMP Dependent Protein Kinase:

After amplifying *Rffor*, BLAST analysis indicated that *Rffor* likely encodes a cGMP-dependent protein kinase, making it likely to be an ortholog of the foraging gene. The foraging gene is most notably characterized by *A. mellifera*'s '*Amfor*' and *D. melanogaster*'s '*for*' as it has been shown to play an instrumental role in different foraging strategies of fruit flies (Osborne et al., 1997), as well as in the onset of the foraging behavior in honey bees (Ben-Shahar et al., 2002). In fruit flies, which are solitary insects, there exist two different alleles of the foraging gene; *for^R*, which is responsible for the 'rover' foraging behavior, and *for^S*, which is responsible for the 'sitter' behavior. Both naturally occurring alleles code for the same cGMP-dependent protein kinase (PKG), the difference being in the level of expression of the enzyme. Flies with the *for^R* allele produce higher levels of PKG and exhibit the 'rover' behavior, foraging more actively than flies with the *for^S* allele, which show lower production of PKG and are known as 'sitters' (Osborne et al., 1997; Sokolowski, 2001).

Ben-Shahar et al. (2002) have shown that *Amfor* is involved in developmentally regulated behavioral variation in honey bees, suggesting that the transition from nursing to foraging behavior is a result of a complex process of behavioral maturation, rather than a genetic dichotomy. Brain expression of *Amfor* was shown to be higher (by a factor of 2- to 8-fold) in foragers than in nurse bees. Bees exhibiting the nursing behavior, when pharmacologically treated with cGMP, show a precocious increase in *Amfor* expression in the brain (Ben-Shahar et

al., 2002) and transition to foraging behavior, demonstrating a strong link between *Amfor* expression and the behavioral switch.

In our sampling of Kansas termites, the alates exhibited the highest level of *Rffor* expression, nearly double that of workers and triple that of soldiers (Fig. 1.1A). Though not as extreme as the 8-fold difference seen between nursing and foraging bees, the 2- to 3-fold up-regulation of *Rffor* in alates is still revealing. Based on studies on other insects, we can infer more information regarding *Rffor*'s possible function and role in social behavior of termites.

In honey bees, *Amfor* levels are elevated in forager bees relative to nurses, and *Amfor* is highly expressed in the lamina of the optic lobes, as well as in the Kenyon cells within the mushroom bodies (Ben-Shahar et al., 2002), which specifically receive visual input (Gronenberg, 2001). Thus, *Amfor* may play a role in “higher order integration of visual information associated with orientation and foraging behavior” (Ben-Shahar et al., 2002). More specifically, the upregulation of *Amfor*, and thus PKG levels, is involved in regulating phototaxis, and is associated with working outside of the hive (Ben-Shahar et al., 2003).

Spatial expression analysis of *for* in *D. melanogaster* revealed high levels of *for* expression in several neuronal clusters. Of notable importance, it was found to have high expression in the optic lobes as in honey bees, and to be present in the antennal nerve of the fruit fly (Belay et al., 2007). The antennal nerve carries axons of the olfactory, auditory, and mechanosensory neurons (Stocker, 1994). The level of *for* expression, dependent upon the presence of either the *for^R* or *for^S* allele, has even been implicated in the degree of short vs. long-term associative olfactory learning in *D. melanogaster*, identifying the mushroom bodies as the “spatial focus of the action of PKG on learning performance (Mery et al., 2007).” These findings in *Drosophila* suggest the possible importance of *for* in receiving not only visual

stimuli, but also auditory and olfactory stimuli, and perhaps even the reception of pheromone signals.

In termites, alates, which possess eyes, leave the colony in nuptial flights (swarms), after which males and females drop to the ground, lose their wings, form monogamous pairs, and search for a suitable nesting site to form a new colony (Vargo and Husseneder, 2009). Pairing generally takes place on the ground, though it has been recorded in a few instances that females alight first, and then males fly low to the ground, seeking them out (Fuller, 1915). Fuller suggested that males locate females by visual cues, such as “the play of sunlight on the half-spread wings of *Macrotermes natalensis* poised on grass spears, and the fluttering wings of *Odontotermes badius* hanging from pendulous plants” (Nutting, 1969). Females of *Odontotermes assmuthi* perform a calling behavior to attract males, during which they remain stationary and move their abdomens violently from side to side. We hypothesize that *Rffor*, due to its relationship with cGMP, may be involved in the development of the visual system (Gibbs et al. 2001; Ben-Shahar et al., 2003) in termites and is critical for functions related to visually identifying a suitable partner as Fuller (1915) suggested.

In an analysis of the calling and tandem behavior in *Reticulitermes lucifugus* however, Buchli (1960) suggested that a chemotropic factor must be present. Females of this species also seek out elevated terrain and raise their abdomens as previously mentioned, with the exception of side-to-side movement. After observing the behavior of tandem running pairs, Buchli (1960) concluded that the male must be initially attracted, and his following response sustained, by a chemotropic factor produced by an active female. It has since been determined that termites rely heavily on chemical communication including trail, sex, and alarm pheromones, as well as hydrocarbons for nest mate recognition (Costa-Leonardo et al., 2009). It’s possible then, that

mate selection in termites may be a product of both visual and chemical stimuli, and a cGMP-dependent protein kinase may play a role in the processing of visual, chemical, and olfactory stimuli.

The foraging gene may also have an impact on termites' identification of potential nesting sites or even aid in orientation during flight. As *Amfor* does in honey bee foragers, *Rffor* could facilitate an increase in phototaxis in alates, aiding in their ability to locate tunnel exits when making their departure from the colony. It's importance in interpreting visual cues may also allow for the avoidance of predation, as termite alates have many natural predators that prey upon them during their swarming flights (Nutting, 1969).

No significant differences were recognized between workers and soldiers. This seems conducive to a life within the colony. Both termite workers and soldiers are blind and spend their lives within the dark, subterranean tunnels of the nest. *Rffor* levels were elevated somewhat in workers (though not significantly) compared to that in soldiers. We hypothesize that the subtle difference noted here may be attributed to the tunnel building and foraging responsibilities of the workers. Also, while workers and soldiers do lack the presence of eyes, there may still be a role for *Rffor* expression in positive, and even possibly negative phototaxis. The presence of ocelli (simple eyes), which allow insects to tell the difference between dark and light, may lead workers or soldiers to exhibit positive phototaxis, causing them move towards a lighted opening. This can be in an effort to forage, to build tunnels to the surface of the soil for alate dispersal as described in *R. lucifugus* (Grassé, 1942), or to defend the colony if the nest walls should be breached. *Rffor* could also influence termite workers to exhibit an increase in negative phototaxis, as they may shy away from openings in the tunnel walls to avoid predation. *Rffor*

could also be important to workers and soldiers in regards to maintaining their orientation within the complex networks of tunnels that make up termite colonies.

Quantitative results from our biological replicate in Florida however showed no significant differences among the three termite castes tested. Soldiers continued to show the lowest levels of *Rffor*, and workers had levels comparable to those in alates. If this colony was preparing for dispersal, workers and soldiers may all be engaging in more of those behaviors mentioned above including tunnel building, foraging, colony defense, positive and negative phototaxis, and orientation within the colony.

Failed Axon Connections:

In our Kansas sample, we found that alates had significantly higher levels of *Rffax* expression than soldiers, whereas worker levels showed no significance (Fig. 1.1B). *Rffax*'s strong sequence similarity to the functionally annotated *D. melanogaster* gene *fax* (failed axon connections), suggests that *Rffax* may play a role in the changes in brain structure and chemistry that precede shifts in complex behaviors in termites comparable to the transition from nursing to foraging behavior in honey bees (Robinson, 2002). The *fax* gene is shown to be up-regulated in nursing bees relative to foragers, as foragers only express the transcript at ~ 65% the level seen in nurses. This seems to be contradictory to the termite *Rffax* levels presented here (Fig. 1.1B, 1.1E). In our sampling, alates have the highest level of expression, nearly double that of workers and soldiers. However, while it would be easier to explain the role of *Rffax* in termites if we had seen elevated levels in workers as we do in nursing bees, this pattern of expression does still make sense. Unlike honey bees, which have two distinct castes in nurses and foragers, termites exhibit intermediary forms between workers and soldiers (pre-soldiers), as well as workers and alates (nymphs) (Zhou et al., 2006). Therefore, while nurses are the immediate precursors to

foragers in honey bees, termite workers are several stages removed from alates and therefore do not exhibit the characteristic increase in *fax* expression that may accompany the behavioral shift to the alate form. In order to see a similar trend as that found in honey bees, it may be useful to undertake quantitative PCR of *Rffax* expression levels in nymphs relative to alates, as they represent the most immediate form preceding the development into alates.

The similar expression levels between workers and alates in our Florida sample may be due to the time point at which the termites were collected. If the workers, soldiers, and alates were all collected at the same time, it's possible that the upregulation of *Rffax* is due to the colony's preparation for dispersal. Workers collected at this point may be preparing to develop into nymphs and then to alates, and therefore may have increased levels of *Rffax* due to the changes in brain structure that likely precede the shift from work inside the colony to life outside of the nest (Robinson, 2002). It's possible that the Kansas colony was not preparing for dispersal at the time the workers were collected, which could explain the lower levels in Kansas workers.

In *D. melanogaster*, *fax* expression is localized in the cellular membranes, suggesting a possible role in cell-cell interactions (Hill et al., 1995). It has also been found to be present in embryonic neurons, including the AVK interneurons, and it is known to regulate neural outgrowth and neural path finding (Much et al., 2000). Thus, *fax* may play a functional role in the structural and chemical changes in the brain that precede the onset of social behaviors (Robinson, 2002). Despite its importance in the onset of foraging behavior in honey bees, *fax* levels decrease in bees after prolonged foraging, showing a foraging-dependent decline in the learning performance of foraging bees (Wolschin et al., 2009).

Termite worker differentiation into the soldier caste then would likely constitute an increased level of *Rffax* expression, as again there is a shift in behavior associated with this

developmental progression. Therefore, changes in brain structure and chemistry would be expected (Robinson, 2002), which explains the presence of *Rffax* expression in soldiers. However, significant down regulation of *Rffax* was observed in soldiers relative to workers and alates in both the Kansas and Florida samples. The lowered expression we see in our sampling could be due to the fact that when a termite differentiates into a presoldier, and consequently a soldier, it has reached a developmental endpoint (Scharf et al., 2003a). We hypothesize then that, as in honey bees, there may be a mechanism of programmed senescence associated with the duration of time an individual has spent performing soldier functions. Individual analysis in soldiers shows variable differences of nearly 8-fold among individuals. Because our sampling was blind to age or sex, there may be older individuals represented in the sample, and thus, lower levels of expression as a result of senescence.

Nuclear Ribonucleoprotein Auxiliary Factor:

Pre-mRNA splicing has long been known to contribute to the production of a variety of transcripts and proteins by alternatively splicing exons into mRNA transcripts prior to their exit from the nucleus. Interactions such as these aren't always one to one, so that one pre-mRNA splicing factor is responsible for the splicing of a single RNA. Alternative splicing of different regions (exons) within a single mRNA may be regulated by a coordinated set of proteins (Park et al., 2004). The gene *RfU2AF*, identified here, is so named after the results of BLAST analysis revealed its putative identity. This 195 bp fragment, isolated from termite head total RNA, was determined through BLAST analysis to be most similar to a pre-mRNA splicing factor. More specifically, the most highly reported identity was that of a U2 nuclear ribonucleoprotein auxiliary factor large subunit (U2AF50), though it is also similar to the U2AF small subunit in some organisms.

Though we hypothesize its function to likely be the alternative splicing of different mRNAs, *RfU2AF*'s specific interactions are not known. In *Drosophila*, U2AF50 is a non-snRNP that has been shown to be essential for viability (Kanaar et al., 1993), to play a role in sexual differentiation in *Drosophila* (Valcárcel et al., 1993; Nagengast et al., 2003), as well as many other diverse functions, including an unexpected role in the active nuclear export of intron-less mRNAs (Blanchette et al., 2004). As a pre-mRNA splicing factor, *RfU2AF* is likely involved in a number of different biological and chemical pathways. Its expression pattern presented here can't bring to light any role in sex-differentiation, as all termites sampled here were already differentiated. Although all three termite castes studied here showed expression of the *RfU2AF* transcript, the highest levels were observed in alates. This was true for both colonies, with soldiers exhibiting a significant down-regulation of the transcript, and workers showing no significant differences. The higher levels of expression in workers and alates may indicate *RfU2AF*'s role in splicing and regulating genes that may be caste-specific, and possibly play a role in caste-specific behavior or physiology related to dispersal or foraging. This gene was shown to be upregulated in foraging honey bees (Whitfield et al., 2003), but no conclusions were reached as to its function.

In conclusion, fragments of three genes, namely, *foraging*, *fax*, and *U2AF* were successfully PCR-amplified and their expression levels quantified in workers, soldiers, and alates of the termite *R. flavipes*. Caste-specific differential expression of these three genes suggests differing roles for each in either the behavioral or physiological differences noted between these three very different termite castes. A preliminary look at the spatial expression of these transcripts suggests that they can be found in different regions throughout the body. Here, we presented hypotheses for the potential functions of these candidate genes in termites based upon

literature from model organisms. In order to confirm the function of these genes and to pinpoint areas of expression, it is necessary to extend our experiments to RNA interference studies, as well as quantitative PCR to investigate the spatial and temporal expression of these genes in specific tissues during all stages of termite development. To fully understand the division of labor in *R. flavipes*, it is important to investigate other genes and networks, as it is unlikely that any one can regulate the complexity of caste differentiation and caste differences on its own.

Figures and Tables

Figure 1.1 Quantitative real-time PCR analysis of *Rffor* (A and D), *Rffax* (B and E), and *RfU2Af* (C and F) expression in pooled heads of termite workers, soldiers, and alates from two unrelated colonies from Kansas (A, B, C) and Florida (D, E, F). Expression levels were calculated and normalized to the reference gene β -Actin. Values represent the means of three experimental replicates, with error bars representing variation in gene expression present between the three experimental replicates (10 heads per caste per pooled sample). Bars labeled with the same letter (a or b) are not significantly different from one another by the least squares difference test (LSD, $\alpha = 0.05$).

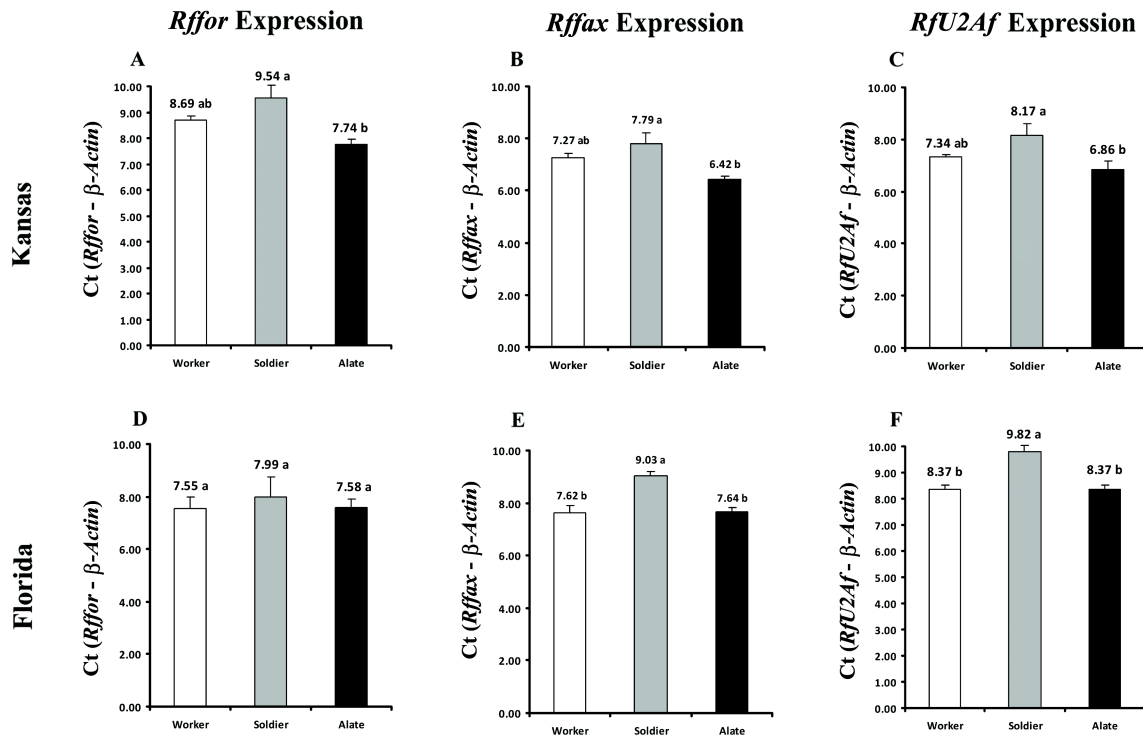


Figure 1.2 Quantitative real-time PCR analysis of individual variation in gene expression of *Rffor* (A) , *Rffax*(B) , and *RfU2AF* (C). Expression was assessed in the heads of termite individuals from each of three termite castes of a Florida colony. Four biological replicates were performed, and the values indicated are the means of three technical replicates, with the error bars representing the standard experimental error present in the three replicates performed for each quantitative PCR sample. Data from individuals were combined to calculate the means of the four biological replicates to show caste-specific trends of *Rffor* (D), *Rffax* (E), and *RfU2AF* (F) expression. Error bars represent the natural variation in gene expression present in the four biological replicates (individuals). Bars labeled with the same letter (a or b) are not significantly different by the least squares difference test (LSD, $\alpha = 0.05$).

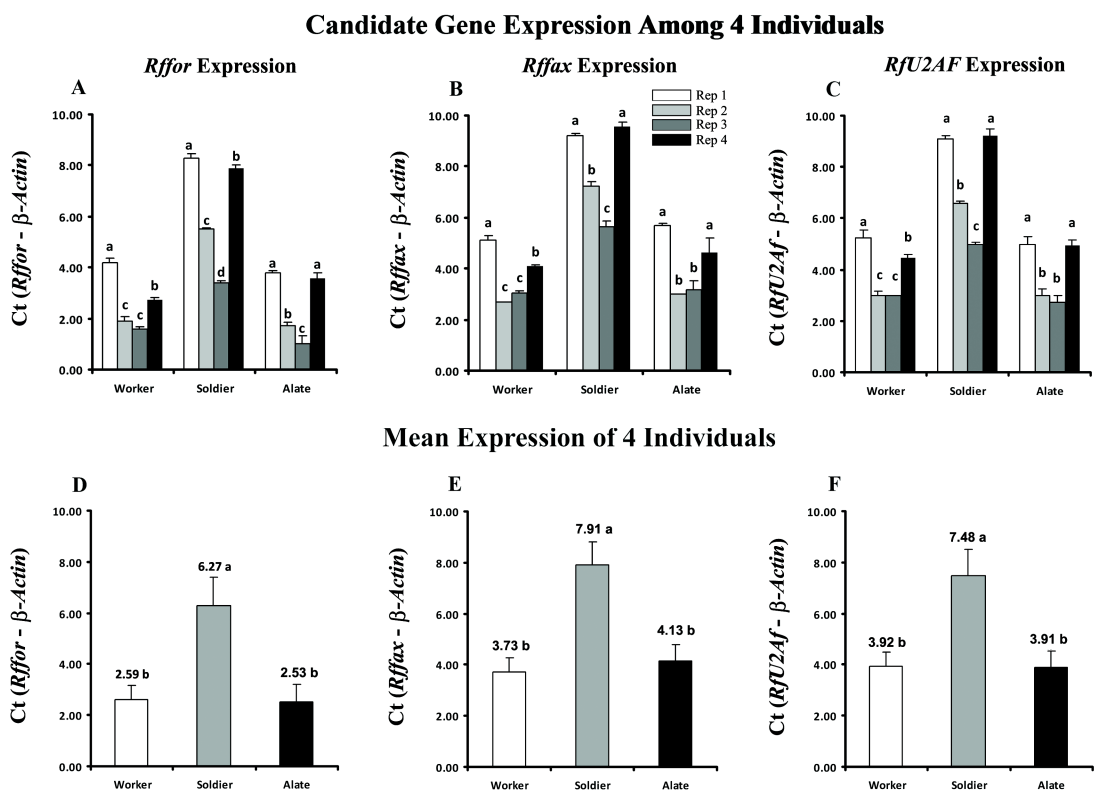


Figure 1.3 Reverse transcriptase PCR analysis to determine whether the three candidate genes, Rffor (97 bp), Rffax (80 bp), and RfU2AF (82 bp) exhibit expression patterns localized to the tissues in the termite head. PCR was carried out using cDNA synthesized from total RNA from the bodies of those termites whose heads comprised the first pool of five in our Florida sample. PCR products were analyzed by electrophoresis on a 1% agarose gel in 1X TAE buffer and visualized under ultra-violet light using ethidium bromide staining. Three termite castes, workers, soldiers, and alates, were assayed, with all three showing successful amplification, indicating that the expression of these transcripts is not localized to the head, but rather that expression is present throughout the termite body, regardless of caste.

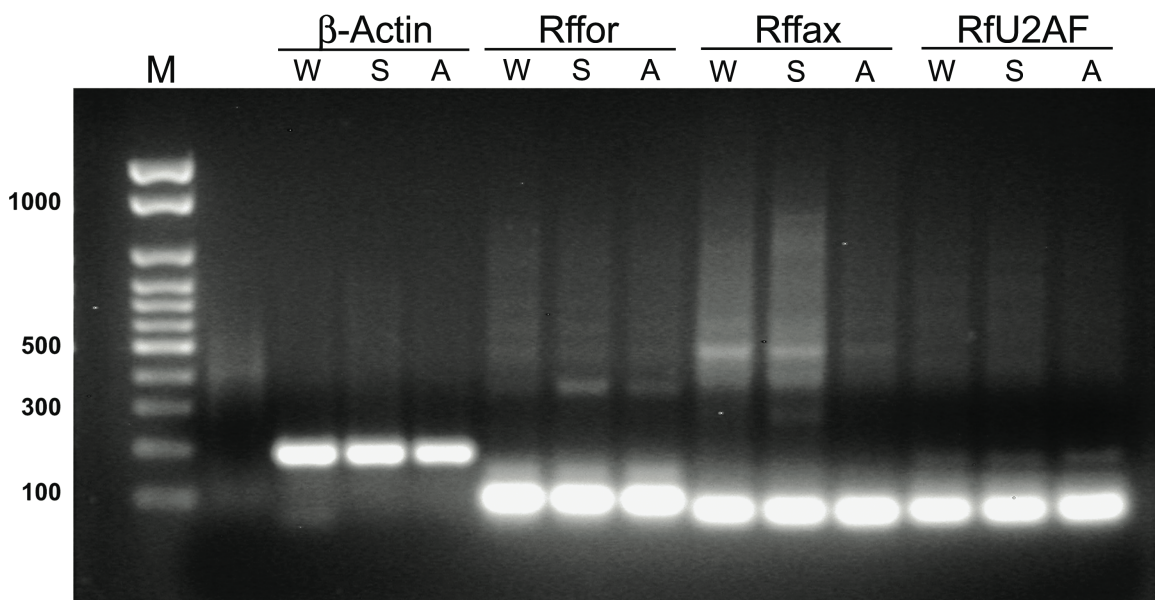


Table 1.1 Candidate genes were selected for amplification in the termite *R. flavipes*. Sequences for primer design were obtained from the NCBI database via honey bee accession numbers seen here. Honey bee expression levels for each gene are shown, as well as their putative function as defined by sequence similarity to known *D. melanogaster* genes. Relative honey bee gene expression levels are presented here as a ratio of forager levels to nurses. Levels for the foraging gene represent the average of three trials performed by Ben-Shahar et al. 2002.

| Candidate Gene / Putative Function | Honey Bee Accession | Honeybee F/N | Best <i>Drosophila</i> Match |
|---|----------------------------|---------------------|-------------------------------------|
| Foraging | AF469010 | ~ 4.00 | <i>for</i> |
| Axonogenesis | BI504916 | 0.63 | <i>fax</i> |
| Pre-mRNA Splicing Factor | BI508614 | 1.18 | <i>U2af50</i> |

Table 1.2 Degenerate CODEHOP primers were designed for amplification of candidate genes in termites. Primers are listed in 5'-3' direction, with their expected fragment lengths provided. Two sets of primers are listed for those genes that required nested PCR for successful amplification. The first set represents the primers that were used for the outer PCR, and the second for the inner PCR. In addition, the primers for B-Actin, which served as our positive control in PCR reactions and as our reference gene for quantitative analysis, are listed.

| Candidate Gene | Forward Primer | Reverse Primer | Bp |
|--------------------------|---------------------------------|---------------------------------|-----------|
| Foraging | GAACGACAGTGTTTCCAGACAathatgatmg | CCTGTCCCTTGGAAATAATGaaraangtrtc | 220 |
| Axonogenesis | TCGTGGAAGTGAACGGAgargarathgc | GCATGCTGCAGGTCACCTtrtancyyt | 221 |
| Pre-mRNA Splicing Factor | AGCCATCCCTGTACTgggaygtnc | CCAGTCCGGACAGATGCatytytgrtt | 357 |
| | TTACACCAATGCAGTACAAGgcnatgcargc | CAGATGCATCTGCTGGTTGaaraaytccat | 195 |
| B-Actin | AACTGGGACGACATGGAGAAGAT | GCCAAGTCCAGACGCAGGAT | 310 |

Table 1.3 Termite specific quantitative primers were designed from candidate gene DNA sequences gathered from degenerate PCR and sequencing experiments. Termite gene names are based on their similarity to known genes as determined through BLAST analysis (Table 1.4). Primers are listed in the 5'-3' direction and expected fragment lengths are indicated.

| Termite Gene | Forward Primer | Reverse Primer | Frag. Lgth. (bases) |
|---------------------|-----------------------|--------------------------|----------------------------|
| <i>Rffor</i> | CGGCTGGAGGACAATATCAAC | CCACTTCAAGAACATCAGCAATC | 97 |
| <i>Rffax</i> | AGTGAACGGAGAGGAGATTGC | CCGCATCCAAATCCTTCTCAAATC | 80 |
| <i>RfU2AF</i> | TGCTGTACCTGTTGTTGGC | CTCTGTTACACCGAAAGGAATG | 82 |
| <i>B-Actin</i> | TACAATGAACTGCGAGTGG | TGAGTCCAGCACGATACC | 195 |

Table 1.4 BLAST analysis of termite gene fragment sequences revealed their putative functions based on their similarity to homologous sequences from other insect species. Similarity is presented as the percent identity, with an e-value showing the significance of the match. The records indicated here represent the sequence hits with the highest values for each of the three candidate genes.

| Termite Gene | Honey bee putative function | Definition | E-Value | Identity | Organism |
|---------------------|------------------------------------|--|----------------|-----------------|-------------------------------|
| <i>Rffax</i> | Axonogenesis | Failed axon connections | 2^{e-27} | 75% | <i>Acyrtosiphon pisum</i> |
| <i>RfU2AF</i> | Pre-mRNA splicing factor | Splicing factor u2af large subunit | 5^{e-23} | 98% | <i>Culex quinquefasciatus</i> |
| <i>Rffor</i> | Foraging gene | Similar to cGMP-dependent protein kinase | 1^{e-26} | 83% | <i>Tribolium castaneum</i> |

References

- Abe, T., Bignell, D. E., and Higashi, M., Eds. 2000. *Termites: Evolution, Sociality, Symbioses, Ecology*. Kluwer Academic Publishers, Dordrecht, The Netherlands.
- Belay, A. T., Scheiner, R., So, A. K., Douglas, S. J., Chakaborty-Chatterjee, M., Levine, J. D., Sokolowski, M. B. 2007. The foraging gene of *Drosophila melanogaster*: spatial expression analysis and sucrose responsiveness. *J. Comp. Neurol.* 504: 570-582.
- Ben-Shahar Y, Leung HT, Pak WL, Sokolowski MB, Robinson GE. 2003. cGMP-dependent changes in phototaxis: a possible role for the *foraging* gene in honey bee division of labor. *Journ. Exp. Biol.* 206:2507–2515.
- Ben-Shahar, Y., Robichon, A., Sokolowski, M. B., Robinson, G. E. 2002. Influence of gene action across different time scales. *Science.* 296: 741-744.
- Blanchette, M., Labourier, E., Green, R. E., Brenner, S. E., Rio, D. C. 2004. Genome-wide analysis reveals an unexpected function for the *Drosophila* splicing factor U2AF50 in the nuclear export of intronless mRNAs. *Mol. Cell.* 14: 775-786.
- Buchli, H. 1960. Les tropismes lors de la parade des imagos de *Reticulitermes lucifugus*. *R. Vie Milieu* 6: 308–315.
- Clustal W2 – Multiple Sequence Alignment Tool. Available from:
<http://www.ebi.ac.uk/Tools/clustalw2/index.html>
- Costa-Leonardo, A. M., Casarin, F. E., Lima, J. T. 2009. Chemical communication in Isoptera. *Neotropical Entomology.* 38(1): 001-006.
- Denison, R., Raymond-Delpech, V. 2008. Insights into the molecular basis of social behavior from studies on the honeybee, *Apis mellifera*. *Invert Neurosci.* 8: 1-9.
- Evans J. D., Wheeler D. E. 1999. Differential gene expression between developing queens and workers in the honey bee, *Apis mellifera*. *Proc Natl Acad Sci USA.* 96: 5575-5580.
- Evans J. D., Wheeler D. E. 2001a. Expression profiles during honey bee caste determination. *Genome Biology.* 2(1), RESEARCH001.
- Evans J. D., Wheeler D. E. 2001b. Gene expression and the evolution of insect polyphenisms. *BioEssays.* 23:62-68.
- Fuller, C. 1915. Observations on some South African termites. *Annals of the Natal Museum.* 3:329–505.

- Gibbs, S. M., Becker, A., Hardy, R. W., Truman, J. W. 2001. Soluble guanylate cyclase is required during development for visual system function in *Drosophila*. *J. Neurosc.* 21:7705-7714.
- Grassé, P.-P. 1942. L'essaimage des termites. Essai d'analyse causale d'un complexe instinctif. *Bull. Biol. France Belg.* 76: 347-382.
- Gronenberg, W. 2001. Subdivisions of hymenopteran mushroom body calyces by their afferent supply. *J. Comp. Neurol.* 435: 474-489.
- Hill, K. K., Bedian, V., Juang, J. L., Hoffmann, F. M. 1995. Genetic interactions between the *Drosophila* Abelson (Abl) tyrosine kinase and failed axon connections (fax), a novel protein in axon bundles. *Genetics.* 141: 595-606.
- Kambhampati, S. and P. Eggleton. 2000. Taxonomy and phylogeny of termites. In *Termites: Evolution, Sociality, Symbioses, Ecology.* (Abe, t., Higashi, T., and Bignell, D., Eds.) Kluwer Academic Publishers, Dordrecht, The Netherlands.
- Kanaar, R., Roche, S. E., Beall, E. L., Green, M. R., Rio, D. C. 1993. The conserved pre-mRNA splicing factor U2AF from *Drosophila*: requirement for viability. *Science.* 262: 569-573.
- Kucharski, R. and Maleszka, R. 2002. Evaluation of differential gene expression during behavioral development in the honey bee using microarrays and northern blots. *Genome Biol.* 3(2), RESEARCH0007.
- Livak, K. J. & Schmittgen, T. D. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta C_T}$ method. *Methods.* 25: 402-408.
- Mery, F., Belay, A. T., So, A. K., Sokolowski, M. B., Kawecki, T. J. 2007. Natural polymorphism affecting learning and memory in *Drosophila*. *Proc. Natl. Acad. Sci. U.S.A.* 104: 13051-13055.
- Miura, T. 2001. Morphogenesis and gene expression in the soldier-caste differentiation of termites. *Insectes soc.* 48: 216-223.
- Miura, T., Kamikouchi A., Sawata M., Takeuchi H., Natori S., Kubo T., Matsumoto T. 1999. Soldier cast-specific gene expression in the mandibular glands of *Hodotermopsis japonica*. *Proc Natl Acad Sci USA.* 96: 13874-13879.
- Much, J. W., Slade, D. J., Klampert, K., Garriga, G., Wightman, B. 2000. The fax-1 nuclear hormone receptor regulates axon pathfinding and neurotransmitter expression. *Dev.* 127: 703-712.

Multiple Alignment processor [http://blocks.fhrc.org/blocks/process_blocks.html]

- Nagengast, A. A., Stitzinger, S. M., Tseng, C., Mount, S. M., Salz, H. K. 2003. Sex-lethal splicing autoregulation in vivo: interactions between SEX-LETHAL, the U1 snRNP and U2AF underlie male exon skipping. *Dev.* 130: 463-471.
- Nutting, W.L. 1969. Flight and colony foundation. In: *Biology of Termites*, Vol. 1(K. Krishna and F.M. Weesner, Eds.), Academic Press, New York. Pp. 233-282.
- Osborne, K.A., Robichon, A., Burgess, E., Butland, S., Shaw, R. A., Coulthard, A., Pereira, H. S., Greenspan, R. J., Sokolowski, M. B. 1997. Natural behaviour polymorphism due to a cGMP-dependent protein kinase of *Drosophila*. *Science*. 277: 834–836
- Park, J. W., Parisky, K., Celotto, A. M., Reenan, R. A., Graveley, B. R. 2004. Identification of alternative splicing regulators by RNA interference in *Drosophila*. *PNAS*. 101:15974-15979.
- Robinson, G. E. 2002. Genomics and integrative analyses of division of labor in honeybee colonies. *Am. Nat.* 160:S160-S172.
- Robinson, G. E., Fernald, R. D., Clayton, D. F. 2008. Genes and social behavior. *Science*. 302: 896-900.
- Robinson, G. E., Grozinger, C. M., and Whitfield, C. W. 2005. Sociogenomics: Social life in molecular terms. *Nat Rev Gen.* 6: 257-270.
- Rose, T. M., Henikoff, J. G., Henikoff, S. 2003. CODEHOP (Consensus-Degenerate Oligonucleotide Primer) PCR primer design. *Nucleic Acids Research*. 31: 3763-3766.
- Rose, T. M., Schultz, E. R., Henikoff, J. G., Pietrokovski, S., McCallum, C. M., Henikoff, S. 1998. Consensus-degenerate hybrid oligonucleotide primers for amplification of distantly related sequences. *Nucleic Acids Research*. 26: 1628-1635.
- SAS Institute. 1985. SAS user's guide: basics. SAS Institute. Cary, NC.
- Scharf, M. E., Ratliff, C. R., Hoteling, J. T., Pittendrigh, B. R. and Bennett, G. W. 2003a. Caste differentiation responses of two sympatric *Reticulitermes* termite species to juvenile hormone homologs and synthetic juvenoids in two laboratory assays. *Ins Soc.* 50: 346-354.
- Scharf, M. E., Wu-Scharf, D., Pittendrigh, B. R., Bennet, G. W. 2003b. Caste-and development-associated gene expression in a lower termite. *Genome Biol.* 4: R62.
- Sokolowski, M. B. 2001. *Drosophila*: Genetics meets behaviour. *Nat. Rev. Genet.* 2: 879–890.

- Steller, M. M., Kabmhampati, S., Caragea, D. Comparative analysis of expressed sequence tags from three castes and two life stages of the termite *Reticulitermes flavipes*. *BMC Genomics* 11: 463 doi:10.1186/1471-2164-11-463.
- Stocker, R. F. 1994. The organization of the chemosensory system in *Drosophila melanogaster*: a review. *Cell Tissue Res.* 275: 3-26.
- Thorne B. L. 1997. Evolution of eusociality in termites. *Ann. Rev. Ecol. Syst.* 28:27-54.
- Thorne, B. L., Traniello, J. F. 2003. Comparative social biology of basal taxa of ants and termites. *Annu. Rev. Entomol.* 48:283-306.
- Tsuchimoto, M., Aoki, M., Takada, M., Kanou, Y., Sasagawa, H., Kitagawa, Y., Kadowaki, T. 2004. The changes of gene expression in honey bee (*Apis mellifera*) brains associated with ages. *Zoological Science.* 21:23-28
- Valcárcel, J., Singh, R., Zamore, P. D., Green, M. R. 1993. The protein Sex-lethal antagonizes the splicing factor U2AF to regulate alternative splicing of *transformer* pre-mRNA. *Nature.* 362: 171-175.
- Vargo, E., Husseneder C. 2009. Biology of subterranean termites: insights from molecular studies of *Reticulitermes* and *Coptotermes*. *Ann Rev Entomol.* 54:379–403.
- Whitfield, C. W., Band, M. R., Bonaldo, M. F., Kumar, C. G., Liu, L., Pardina, J. R., Robertson, H. M., Soares, M. B., Robinson, G. E. 2002. Annotated expressed sequence tags and cDNA microarrays for studies of brain and behavior in the honey bee. *Genome Res.* 12:555
- Whitfield, C. W., Cziko, A. M., Robinson, G. E. 2003. Gene expression profiles in the brain Predict behavior in individual honey bees. *Science.* 302:296-299a
- Wilson, E. O. 1971. Caste: Termites. In: *The Insect Societies*. Belknap Press of Harvard University Press, Cambridge, Massachusetts. pp 183-210.
- Wilson, E. O. 1975. *Sociobiology: the New Synthesis*. Belknap Press of Harvard University Press, Cambridge, Massachusetts. 697pp.
- Wolschin, F., Munch, D., Amdam, G. V. 2009. Structural and proteomic analyses reveal regional brain differences during honeybee aging. *J. Exp. Biol.* 212: 4027-4032.
- Zhou, X., Oi, F. M., Scharf, M. E. 2006. Social exploitation of hexamerin: RNAi reveals a major caste-regulatory factor in termites. *PNAS.* 103: 4499- 4504.