ANALYSIS OF GENE EXPRESSION CHANGES IN RESPONSE TO FIELD-TO-LAB TRANSITION IN THE ARGENTINE ANT, *LINEPITHEMA HUMILE*

by

Mathew Dittmann

A Thesis

Submitted to the Faculty of Purdue University

In Partial Fulfillment of the Requirements for the degree of

Master of Science

Department of Entomology

West Lafayette, Indiana

December 2018

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF COMMITTEE APPROVAL

Dr. Gary Bennett, Co-Chair
School of Agriculture
Dr. Michael Scharf, Co-Chair
School of Agriculture
Dr. Grzegorz Buczkowski
School of Agriculture

Approved by:

Dr. Stephen Cameron
Head of the Graduate Program

ACKNOWLEDGMENTS

To my father, who convinced me to pursue Entomology as a career path.

To Dr. Susan McDowell, who introduced to the fascinating world of molecular biology.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Author: Dittmann, Mathew A. MS
Institution: Purdue University
Degree Received: December 2018
Title: Analysis of Gene Expression Changes in Response to Field-to-Lab Transition in the Argentine Ant, *Linepithema humile*.
Committee Chair: Gary Bennett, Michael Scharf

Gene expression research is a valuable tool for investigating how gene regulation and expression control the underlying behaviors that structure a eusocial insect colony. However, labs that focus on ant research frequently keep ant colonies in the lab for ease of sampling. These laboratories typically do not attempt to completely emulate the ant's natural environment, and thus can expose the colonies to drastically different environmental conditions and food sources than they are used to in the wild. These shifts in diet and environment can cause changes in the gene expression of the ants, affecting downstream behavioral and physiological systems. To examine the nature of these changes, colonies of the Argentine ant, *Linepithema humile* (Mayr, 1868), were excavated from North Carolina and transferred to the lab, where they were sampled monthly. Illumina and qPCR analyses were conducted on forager samples to detect any changes in gene expression. Approximately six percent of the Argentine ant genome showed changes in gene regulation after six months in the laboratory environment. The subset of these genes examined via qPCR show that the expression of many genes are correlated with each other, indicating that these genes might be a part of a regulatory network. These findings showed that ant colonies kept in the lab experience changes in gene expression, resulting in downstream effects. Therefore, lab ant colonies are not necessarily representative of wild colonies when conducting experiments on the gene expression, behavior, and physiology of these colonies.

CHAPTER 1. LITERATURE REVIEW

1.1 The Argentine Ant

The Argentine ant, *Linepithema humile* (Mayr, 1868), is one of the most widespread invasive species in the world, having become established on six continents and many smaller islands across the planet (Buczkowski et al. 2004). In the United States of America, the Argentine ant was first discovered in New Orleans (Newell 1908). From there, it has invaded much of the western and southern coasts, and has made inroads into the eastern coast as well Suarez et al. 2001). This invasion has resulted in a number of problems, such as the loss of native ant species and the resulting downstream effects, causing a nuisance by invading homes and foraging on human food, and agricultural damage through aphid farming (Daane et al. 2007; Rust et al. 1996; Suarez et al. 2000). In its native habitat, the Argentine ant is able to coexist with a number of other native species, and exhibits intraspecific aggression as well comparatively small territories (Tsutsui and Case 2007). However, introduced populations behave much differently, creating large, cooperative supercolonies that span hundreds of miles (Tsutsui et al. 2000). These supercolonies show much lower genetic diversity than native populations, indicating that the Argentine ant went through a genetic bottleneck when it was initially introduced. This reduction in genetic diversity translates into a loss of aggression towards conspecifics, resulting in the formation of these supercolonies. As a consequence of this unicoloniality, the Argentine ant is able to outcompete native ant species through sheer volume of workers (Holway 1998; Holway 1999; Human and Gordon 1996). However, native ant species aren't the only victim of this invasion, some vertebrate populations also show losses in response to Argentine ant invasion (Laakkonen et al. 2001; Suarez et al. 2000). These large supercolonies also become incredibly common nuisance pests for residents in areas that the

Argentine ant is present (Field et al. 2007). Finally, since Argentine ants show a preference for sugary liquids such as honeydew, they are frequently found tending aphid populations (Baker et al. 1985). Their protection of the aphids results in the failure of biological control methods and causes economic damage to crops (Phillips and Sherk 1991). Argentine ants are also capable of acting as a vector for plant diseases, such as their vectoring of avocado stem canker (El Hamalawi and Menge 1996). It is clear that the Argentine ants act as important pest species in environments that they invade, and more research needs to be done to develop better methods of controlling them.

1.2 Gene Expression in Eusocial Insects

Until recent years, the sociogenomic mechanisms that regulate eusocial behavior in insects have largely remained unknown due to the lack of practical tools for the analysis of large quantities of genetic data. Without tools capable of generating and analyzing data for thousands of genes at a time, research into genetic regulation of eusocial behaviors had been limited to small scale studies examining well-known genes and their downstream protein products (Engels 1974; Huarong Lin et al. 1999; Wheeler and Nijhout 1983). However, with the development of tools capable of assaying thousands of genes at a time, such as microarrays and later RNAseq, research into the complex web of genes that govern the caste systems and cooperative lifestyles associated with eusociality became practical (Chen and Chen 2003; Jongeneel et al. 2001). Using these tools, gene expression analysis has become a dominant influence in the way that eusociality is researched by looking for gene expression patterns that change due to influences from caste, age, sex, or a variety of other factors (Friedman and Gordon 2016).

1.3 Sex-Biased Gene Expression

The most obvious example of phenotypic plasticity in eusocial insects is the morphological and behavioral dimorphism that differentiates male and female colony members (Friedman and Gordon 2016). While most insects have ZW or X0 sex determination systems, Hymenopterans instead follow a haplodiploidy sex determination system where females have two sets of chromosomes and males only have one set (Mueller 1991). Differentiation in gene expression between males and females is apparent at all life stages, but the expression profiles of the two sexes grow more distinct as the organism moves through the life cycle (Harrison et al. 2015; Hoffman and Goodisman 2007). Analyses of gene expression patterns between males and females within ant colonies s how that while males do show distinct differences compared to females, the magnitude of difference depends on the caste of the female (Ometto et al. 2011). Comparisons in gene expression patterns show more similarity between the male and female reproductive castes than comparisons between males and sterile female castes, likely because fertile males and females share the same reproductive status, while males and sterile females do not share sex, behavioral roles, or reproductive status. In keeping with their different roles within the colony, expression of odorant-binding proteins and chemosensory proteins found in the antennae also show differentiation between males and workers, although the exact expression patterns and directions of caste bias vary between species (Mckenzie et al. 2014). While the differences between sexes in eusocial insects are the most obvious within a colony, males typically have very little involvement in the function of a colony, so examining differences in gene expression between sexes is of limited value. Instead, the most pertinent area to examine gene expression is the differences between the female castes, which do make up the bulk of a colony and are the most involved with its function.

1.4 Caste-Biased Gene Expression

Despite the single genome present in all the members of a eusocial insect colony, the members of the colony can display substantial phenotypic plasticity, resulting in distinct reproductive and nonreproductive castes (Wheeler 1986). Some of the more evolutionary derived ant species exhibit even greater plasticity through the presence of polymorphic subcastes of workers that have specialized functions within the colony. However, some eusocial species don't exhibit morphological differentiation between reproductive and sterile castes either because their ancestors never evolved the trait, such as in *Polistes* wasps, or because the morphological separation between reproductive and sterile castes was later lost, as in some ponerine ant species (Peeters and Crozier 1988; Sumner et al. 2006). Regardless of the amount of morphological variation present in the colony, the differences between the castes of eusocial taxa are regulated by the expression of certain genes in the genome (Aubin-Horth and Renn 2009; Linksvayer et al. 2012). Differential gene expression is involved with the regulation of caste behavior and physiology, and task polyethism (Friedman and Gordon 2016). The vast majority of recent gene expression research in ants focuses on discovering these underlying transient and persistent gene expression patterns that differentiate the castes present in colonies. These expression differences become apparent even before the adult stage is reached. For instance, juvenile hormone is a driver of caste and subcaste differentiation in many eusocial insect groups (Cornette et al. 2008; Rachinsky et al. 1990; Rajakumar et al. 2012; Wheeler and Nijhout 1981; Scharf et al. 2007). Juvenile hormone titers are affected by food intake, and by extension genes associated with nutrient sensing (Mutti et al. 2011; Tu et al. 2005). Genes such as *Am-ILP*, *Am-IR*, and *Am-InR* show distinct expression profiles in late-stage larvae prior for worker- and queen-destined larvae (de Azevedo and Hartfelder 2008; Wheeler et al. 2006). Manipulation of

expression patterns of nutrient sensing genes like these via RNAi-mediated knockdown can prevent queen formation even in queen-destined larvae (Mutti et al. 2011; Patel et al. 2007; Wolschin et al. 2011). While genes associated with juvenile hormone levels have been investigated extensively and determined to control caste determination in honeybees, there are other genes which show differential expression between caste-destined larvae but have not been confirmed to actually influence the transition, so there are plenty of other targets to investigate (Evans et al. 1999; Pereboom et al. 2005). However, most gene expression research focuses on expression of genes within the adult castes of eusocial insects. Whole transcriptome studies in eusocial insects often find hundreds of genes that show differential expression between queens and workers (Feldmeyer et al. 2014; Ferreira et al. 2013; Hoffman and Goodisman 2007). However, in depth analysis of the effects that individual genes have on caste determination is rare. One particular gene family known has vitellogenins has garnered a large amount of scrutiny as a driver of caste differentiation. Vitellogenin is a protein that is used as an egg-yolk precursor protein in many different taxa (Hagedorn and Kunkel 1979). However, in eusocial insects, the functions of vitellogenin have diversified (Corona et al. 2013; Guidugli et al. 2005; Morandin et al. 2014). In addition to its role in egg yolk formation, the functions of the vitellogenin genes have become associated with regulation of worker task transitions, oxidative stress resistance and immune strength (Amdam, Simões, et al. 2004; Guidugli et al. 2005; Nelson et al. 2007; Seehuus et al. 2006). Since the queen's reproductive function to the colony is critical regardless of taxa, eusocial species typically have queens with an overexpression of at least one vitellogenin gene when compared to workers (Feldmeyer et al. 2014). Queen overexpression of vitellogenin is not universal however, and some eusocial species exhibit the reverse trend, with workers having an overexpression of vitellogenin as compared to queens (Harrison et al. 2015).

However, vitellogenin expression and function becomes more complicated with the expansion of vitellogenin and vitellogenin-like genes that has occurred in ant species (Morandin et al. 2014). When more than one vitellogenin gene is found in the genome, some genes show queen-biased expression and some show worker-biased expression, likely due to sub- and neofunctionalization that occurs when vitellogenin copies become present in the genome (Corona et al. 2013; Feldmeyer et al. 2014). While queens tend to show upregulation in some form of vitellogenin, there doesn't appear to be consistently queen-biased vitellogenin genes between ant species (Berens et al. 2015). The inconsistency in vitellogenin expression patterns between reproductive and sterile castes is likely due to the inconsistent amount of vitellogenin genes present in a given ant species genome, which was caused by multiple adaptive radiations that have happened to vitellogenin over ant evolutionary history (Corona et al. 2013). These biases in expression between reproductive and non-reproductive groups can remain even when caste polymorphism is lost in eusocial species. For example, the clonal raider ant *Ooceraea biroi* (formerly *Cerapachys*) has lost the queen phenotype, and instead the workers all produce clonal daughters through parthenogenesis (Tsuji and Yamauchi 1995). *O. biroi* has two ancestral vitellogenin genes, and despite the lack of a queen the reproductive division in expression between the two genes can still be seen. The *O. biroi* life cycle is larval-driven biphasic cycle, consisting of a reproductive phase in which eggs are laid and a brood care phase when the emerging larvae are tended and foraging is done (Ravary et al. 2006). During the reproductive phase, the "queenlike" vitellogenin is upregulated, and during the foraging/brood care phase the "workerlike" vitellogenin is upregulated (Oxley et al. 2014). Even species with a more typical caste-driven reproductive lifestyle can show reproductive division in the worker caste. When some species of bees and ants lose their queen, some of the workers in the newly queenless colonies will begin

taking on reproductive duties through parthenogenetic production of males or females; or by mating with a brother and becoming a gamergate (Bourke 1988; Peeters and Hölldobler 1995). When workers make these transitions, their gene expression patterns will shift to approach a more queen-like pattern (Feldmeyer et al. 2014; Harrison et al. 2015; Huarong Lin et al. 1999; Winston et al. 2007).

1.5 Subcaste-Biased Gene Expression

While comparing gene expression differences between the sterile and reproductive castes is the most popular way to study the underlying mechanisms of eusociality, studying the differences between worker subcastes is another important avenue of research to examine the complexities of eusocial insect societies. While majors and minors are both functionally part of the ant worker caste, they have morphological differences and focus on different tasks in the colony (Wheeler 1991). Some of these morphological and behavioral variations have their roots in gene expression. For example, the foraging gene is an ancestral gene regulating foraging behavior, energy storage, and metabolism, but in eusocial insects the foraging gene has taken on additional functions (Kent et al. 2009). It can control defensive responses, such as in major workers in *Pheidole pallidula*, where the foraging protein induces more aggressive defense responses in majors exposed to conspecifics, but the same effect does not occur in other minor workers within the species (Lucas and Sokolowski 2009). Tachykinin, a neuropeptide, shows similar characteristics in the worker castes of *Acromyrmex echinator* (Howe et al. 2016). Other studies comparing worker subcastes have shown differences in genes associated with chemical communication, muscle development, and neuron activity (Bonasio et al. 2010; Simola et al. 2013). While work has been done examining the differences in the patterns of gene expression

exhibited by adult majors and minors, there hasn't been any research to look into the genes involved in major determination during the larval and pupal stages (Anderson et al. 2008; Tian and Zhou 2014). The only research conducted thus far shows that soldier development is associated with social, environmental, and genetic factors, but so far no gene expression work has been done (Anderson et al. 2008). The only definitive research done shows that juvenile hormone regulates the development of soldiers, which is a similar regulatory mechanism found in termites (Tian and Zhou 2014; Wheeler and Nijhout 1981, 1983). JH has even been implicated in the formation of "supersoldiers" in certain species of *Pheidole* (Rajakumar et al. 2012). Research into termite soldier determination is more in-depth, and shows a number of genes, such as *hexamerins*, regulate soldier morphogenesis (Tian and Zhou 2014). Subcastes are not only limited to ants and termites, although when present in other eusocial species they usually only show behavioral differentiation, with no morphological differences between subcastes (Herb et al. 2012). Changes in methylation patterns are shown associated with these behavioral subcastes, as shown in honeybees. Since workers can transition from one behavioral subcaste to another and in some cases back again, methylation patterns associated with those subcastes are similarly able to change over time.

1.6 Subcaste-Biased Gene Expression

While morphological variation is the clearest example of differential gene expression driving phenotypic diversity in eusocial insects, it is by no means the only example. Differences in gene expression can also drive behavioral differences as well (Friedman and Gordon 2016; Mikheyev and Linksvayer 2015). In addition to the caste-associated behavioral differences outlined above, gene expression also shows pattern variation based on worker task polyethism,

frequently having to do with foraging behavior. The most well-studied instance of differential gene expression driving differences in behavior is in the *foraging* gene. In ants, foragers tend to show lower levels of expression of the *foraging* gene than nurse workers (Ingram et al. 2011; Ingram et al. 2005). In bees, the trend tends to be reversed, with forager bees showing higher levels expression of *foraging* than nurse bees (Ben-Shahar et al. 2002). Interestingly, *foraging* also shows differential expression when comparing polygyne and monogyne workers in *Solenopsis invicta* (Lucas et al. 2015). Workers in multi-queen colonies show lower levels of *foraging* than workers from single queen colonies. Many other genes also show differential expression based on behavior. Vitellogenin has been shown to be involved in the foraging/nursing dichotomy. In bees, higher titers of vitellogenin are associated with increased pollen storage and more frequent foraging (Amdam et al. 2004; Toth and Robinson 2007). In ants, some species show upregulation of specific vitellogenin genes when showing forager behavior (Corona et al. 2013; Oxley et al. 2014; Ravary et al. 2006). In *A. mellifera*, the gene *malvolio* is a manganese transport protein that controls sucrose responsiveness (Toth and Robinson 2007). Manipulation of *malvolio* through manganese treatment induces precocious foraging in worker bees. Interestingly, *malvolio* also shows increased expression in the head tissues of *O. biroi*, compared to body tissues (Oxley et al. 2014). Considering that sensitivity to sugars is important to foraging efficiency, the differential expression of *malvolio* shown in *O. biroi* might indicate similar foraging regulation to that found in *A. mellifera* (Judd and Fasnacht 2007). Circadian genes show similar behavior-biased expression as well (Ingram et al. 2009). *Pogonomyrmex occidentalis* foragers who work outside the nest show pronounced differences in levels of the *period* gene over the length of the day, while nurse ants that stay inside the nest, and thus are not exposed to the night-day cycle, exhibit a much lower variation in the *period* gene

(Ingram et al. 2009). The *period* gene seems to have a complex relationship with foraging behavior, based on research done on honeybees (Bloch et al. 2001). Older foragers show higher variation in expression, but precocious foragers that develop in colonies artificially deprived of foragers exhibit the same variation in expression. It's clear that *period* expression is regulated by multiple social, environmental, and age factors, so it's relationship with regulation of foraging behavior is complex. While the genes behind foraging behaviors have received attention, other ant behaviors such as nursing are comparatively less well-characterized, so research into genes regulating ant behavior remains a significant field.

1.7 Regulation of Gene Expression

Differential gene expression patterns are partially due to methylation patterns of the genome associated with caste and age (Alvarado et al. 2015). Methylation in the genome of eusocial insects has been shown to exhibit caste-specific patterns that influence the expression of genes that are methylated, although how closely methylation patterns follow transcriptional differences varies between species and specific genes (Bonasio et al. 2012; Lyko et al. 2010). Aside from influencing transcriptional regulation, genomic methylation has also been shown to be a regulator of caste differentiation in ants and honeybees (Li-Byarlay et al. 2013; Simola et al. 2013). Regulation of gene expression can also be influenced through post-transcriptional editing of the pre-mRNA molecule to create alternative splice sites, which results in the production of alternative transcripts and proteins (Laurencikiene et al. 2006). Manipulating miRNA or its target sites is also used to regulate gene expression (Borchert et al. 2009; Kawahara et al. 2007). MicroRNAs, also called miRNAs, are endogenous RNA molecules similar to synthetic siRNA molecules (Jackson and Standart 2007). miRNAs are used by the organism to regulate gene

expression, although these molecules work by binding to specific sites on the mRNA and impairing translation rather than acting like siRNAs, which induce mRNA degradation. The targets of miRNAs can be modified by adenosine deaminases that act on RNA (ADAR) through either modification of the miRNA during processing, or modification of mRNA to create new binding sites in the 3' UTR region of the RNA molecule (Borchert et al. 2009; Kawahara et al. 2007). ADAR has been found to be differentially expressed in the castes of *Acromyrmex echinatior*, and miRNA expression has also been shown to vary between major and minor workers in *Camponotus pennsylvanicus*, so miRNA is another method of gene regulation used by ants to differentiate caste gene expression patterns (Bonasio et al. 2010; Li et al. 2014). Both methylation and miRNAs are likely regulators of gene expression. The next step is to examine how gene expression is used by eusocial insects.

1.8 Caveats

Despite the large amount of research generated in recent years, a few caveats remain. First, examining differences in gene expression can be difficult because any bias in expression can be localized to specific tissues, and whole body preparations of samples can lose these differences due to the "noise" generated by the rest of the body (Lucas et al. 2017; McKenzie et al. 2014; Simola et al. 2013). Second, cross species analyses searching for consistent trends in gene expression patterns have borne very little fruit, finding few genes that show consistent differences between castes (Berens et al. 2015; Morandin et al. 2016; Sumner 2014). These findings run in contrast to the "genetic toolkit" hypothesis, which is the idea that the mechanisms of eusocial behavior were built from a conserved set of genes that are involved with basic physiological processes (Mikheyev and Linksvayer 2015). However, alternative hypotheses

have developed to explain this lack of consistency, suggesting instead that changes in molecular pathway regulation are the driver of these caste differences (Berens et al. 2015; Morandin et al. 2016; Sumner 2014). Third, while there is usually always a difference in gene expression patterns between castes or sexes, any differences in gene expression between subsets of a colony are dependent on the life stage being examined. For example, queen-destined larvae upregulate genes associated with metabolism and respiration compared to worker-destined larvae, but this pattern is not necessarily present in adults (Evans and Wheeler 2001). Many studies show that developmental stage is the largest driver of gene expression patterns (Harrison et al. 2015; Hoffman and Goodisman 2007; Morandin et al. 2015; Ometto et al. 2011). Therefore, when studying gene expression patterns it is important to keep in mind that the results of an experiment can only be applied to the developmental stage that was used for the experiment. Finally, studies that rely on colonies that have been reared in a laboratory environment, which removes them from most, if not all, of the natural influences that they would normally receive from the surrounding ecosystem, are not necessarily representative of wild colonies. This loss of information could potentially impact gene expression patterns, and thus should be examined. However, there are very few studies attempting to examine the affects that laboratory environments have on insects (Hoffmann and Ross 2018; Jandt et al. 2015).

1.9 Objectives and Hypotheses

The objectives of this research were to collect data and determine if changes in gene expression in Argentine ants occur in response the transition from a field environment to a lab environment. The central hypothesis of this work is that genes associated with elimination of toxic chemicals and cuticular hydrocarbon synthesis will shift in response to this change in environment.

CHAPTER 2. COMPARISON OF GENE EXPRESSION BETWEEN LABORATORY AND FIELD COLONIES

2.1 Introduction

Eusocial insect societies are characterized by three things: cooperative brood care, overlapping generations of individuals, and a reproductive division of labor (Wilson 1971). Much research has been conducted in order to understand the behavioral and genetic systems underlying eusocial behavior (Toth and Robinson 2007; Wheeler 1986). Investigations into how eusocial structures are maintained by insects have looked at the effects of hormones such as juvenile hormone, nutritional thresholds causing larvae to go down different developmental pathways, behavior characteristics such as policing of worker egg laying, and genetic drivers of caste differentiation such as yellow and royal jelly genes (Anderson et al. 2008; Drapeau et al. 2006; Gamboa and Breed 1977; Patel et al. 2007; Ratnieks 1988; Robinson 1987). However, until recently large scale studies into the gene networks that regulate these mechanisms has been impractical due to having no feasible way to generate and analyze large quantities of genomic and transcriptomic data. However, with the advent of genomic and transcriptomic analysis, analyzing the expression profiles of an entire genome becomes practical, allowing for comparisons between different eusocial insect castes, subcastes, and sexes (Friedman and Gordon 2016).

When doing long-term studies on insects, many researchers find it beneficial to establish laboratory populations for ease of collecting samples and running experiments. This is especially useful for the study of ant colonies, as it provides easy access to reproductive and immature life stages, which are normally difficult to find in the field. However, prior research

has shown that keeping insects in the lab over multiple generations can cause shifts in their behavior and physiology resulting in loss of genetic variation or changes to tolerance of heat, desiccation, or UV exposure (Hoffmann and Ross 2018; Jandt et al. 2015). In particular, laboratory adaptation has been shown to decrease amylase cytochrome P450 expression and increase hexamerin expression in housefly populations (Højland et al. 2014). However, very little research has been done on the effect of laboratory rearing on the expression of genes, and the research that has been done is typically conducted on short lived species like mosquitos and houseflies (Aguilar et al. 2011; Højland et al. 2014; Højland et al. 2014). Since gene expression is primarily used for protein production, calculating gene expression levels and how they change can serve as a proxy for determining how modifications in protein production are affecting the physiology and behavior of lab specimens. This is especially important information when considering the plethora of gene expression research being conducted on ants and the fact that the lack of generational changes leaves gene expression the only avenue through which laboratory adaptation can occur.

The purpose of this research was to examine how the Argentine ant transcriptome changes when colonies were taken from the field and raised in a laboratory environment. Two methods were used to discover any potential changes. First, the entire transcriptome was compared between field and lab ants to determine if there were any particular gene families or functional categories that changes are focused in. Second, a subset of the genes found to be differentially expressed in the first objective were analyzed on a month-by-month basis to see at what point gene expression starts to change.

2.2 Materials and Methods

2.2.1 Illumina Transcriptome Analysis

2.2.1.1 Colony Retrieval and Sampling

Four nests of the Argentine ant, *Linepithema humile* (Mayr, 1868), were dug up from soil around Forsyth Technical Community College in Winston-Salem, North Carolina. The colonies were transported from the site and placed in Fluon-coated plastic trays in Smith Hall at Purdue University. Foraging worker samples were taken from each colony immediately upon arrival at Purdue. The colonies were extracted from the soil and moved to test tubes wrapped in aluminum foil in the same plastic trays. Colonies were kept in the lab for a period of six months, provided water and sugar water *ad libitum*, and fed on a diet of cockroaches and modified Bhatkar-Whitcomb diet (Bhatkar and Whitcomb 1970). Every month, a sample of foragers were taken from each colony and kept in RNAlater at -80^{o}C until RNA extraction was done.

2.2.1.2 RNA Extraction and Quantification

Ant samples were homogenized using a motorized micropestle apparatus. The homogenized ant samples had their RNA extracted using an SV Total RNA Isolation System Kit (Z3101, ProMega Corporation, 2800 Woods Hollow Road, Madison, WI 53711). RNA was quantified using a NanoDrop 2000 spectrophotometer (ND-2000, Thermo Fisher Scientific, 168 Third Avenue, Waltham, MA 02451). 0-month samples and all samples from colony 1 were inadequate for analysis and discarded. Aliquots of RNA suspension for all 1-month and 6-month samples containing 200 ng of RNA were submitted to the Genomics Core for Illumina HiSeq analysis. Further aliquots of 200 ng RNA were converted to cDNA using a SensiFAST cDNA Synthesis Kit (BIO-65053, Bioline USA Inc, 305 Constitution Drive, Taunton, MA 02780).

2.2.1.3 Illumina HiSeq Analysis

RNA aliquots were analyzed for RNA integrity via RIN scoring (Mueller and Schroeder 2004). To purify the RNA content, the RNA was subjected to ribodepletion to remove unwanted rRNA from the samples using a TruSeq Stranded Total RNA Library Prep Gold kit (20020598, Illumina, 5200 Illumina Way, San Diego, CA 92122). Ribodepleted RNA samples were processed using a Library Preparation Kit (KR1139, Kapa Biosystems, 200 Ballardvale St, Suite 350, Wilmington, MA 01887) and analyzed using an Illumina HiSeq 2500. Data from the Illumina analysis were submitted to Bioinformatics Core for analysis.

2.2.1.4 Bioinformatic Analysis of Illumina Data

Raw Illumina reads were obtained from the Purdue Genomics Core facility. Sequence data quality was determined using FastQC software (version 0.11.2). Quality trimming was performed using FASTX-Toolkit (version 0.0.14) to remove the bases with less than Phred33 score of 30, and resulting reads of at least 50 bp were retained (which comprised >99% of total reads for most samples). Reads were sorted into rRNA and non-rRNA fractions using sortMERNA tool and non-rRNA read fraction was utilized for all downstream analyses.

Non-rRNA reads were mapped against the indexed *L. humile* reference genome using STAR aligner (version 2.5.2b) with default parameters. STAR derived mapping results and annotation (GFF) file for reference genome were fed to HTSeq package (version 0.7.0) to obtain read counts for each gene feature for each replicate. Counts from all replicates were merged together using custom Perl script to generate a gene counts matrix for both samples (1-month and 6-month). Genes with 0 counts across all replicates were discarded from the counts matrix. When genes have 0 counts in one sample but not in others, the counts were converted from 0 to 1 to avoid having infinite values being calculated for fold change. Final combined counts matrix

was utilized for further differential gene expression (DGE) analysis by DESeq2 and edgeR packages. Additionally, DGE was calculated using the tuxedo protocol which directly used STAR mapping files (bam) instead of count matrix. The tuxedo protocol uses Fragments per kilobase of exon per million reads mapped (FPKM) which is corrected (normalized) for the length of the gene and the library size to represent the gene-expression as compared to raw counts in edgeR and DESeq2.

DGE analysis between 1-month (field) and 6-month (lab) samples was carried out using 'R-Bioconductor' package (version 3.3.2) and two different methods (**DESeq2** and **edgeR**). Both edgeR (version 3.16.5) and DESeq2 (version 1.14.1) use the negative binomial distribution based data model and perform specific estimate variance-mean tests. Both methods determine differentially expressed genes with *P*-value and adjusted *P*-values of false discovery rate (FDR) to correct for multiple tests. The quality of counts matrix was verified by determining basic statistics such as data range and matrix size prior to statistical tests. The DEseq2 package provides methods to test for DGE by use of negative binomial generalized linear models, the estimates of dispersion (measure for sample variance) and logarithmic fold changes. DESeq2 applied Empirical Bayes shrinkage for dispersion estimation and Wald test was used for significance testing and DGE. In the **edgeR** package, an edgeR object was created using the counts matrix, and providing library sizes and experimental design. Normalization factors were calculated for the counts matrix, followed by estimation of common dispersion of counts. EdgeR package performed an 'exact' test to calculate DGE. The tuxedo protocol starts with combined mapping files for each sample which are then processed through the **Cufflinks8** (version 2.2.1) suite of programs (Cufflinks, Cuffmerge, Cuffquant and Cuffdiff) to determine DGE. Briefly, cufflinks performs the transcript assembly for each sample each replicate, cuffmerge combines

the assemblies into a master transcriptome, cuffquant calculates the genes and transcript expression profiles, and cuffdiff compares these expression profiles to determine DGE. A pairwise comparison of Control and Treatment samples was performed using cuffdiff with default parameters. Each replicate was used to build a model, then these models are averaged to provide a single global model representing all conditions in the experiment and used for dispersion estimate. A t-test was performed to measure the DGE with *P*-values and also calculated adjusted *P*-values of false discovery rate (FDR) to correct for multiple tests. A gene was considered as exhibiting differential gene expression when two or more methods detected differential gene expression in that gene.

2.2.1.5 Blast2GO Analysis

Genes showing a significant change in expression were separated into upregulated and downregulated categories, and each group was further broken down into high, medium, and low categories to denote the magnitude of the change. For each list of genes, protein sequence data was obtained from GenBank and loaded into BLAST2GO (Version 4.1). Sequence data was blasted on GenBank, to obtain description information for the genes. After sequences were identified, genes were mapped and annotated using default settings. InterPro was also searched for sequence hits and the results merged into annotation. Gene Ontology (GO) terms were then graphed for molecular function, and common GO terms were obtained for category of genes.

2.2.2 Quantitative Real Time PCR (qPCR) Analysis of Selected Transcripts

2.2.2.1 Colony Retrieval and Sampling

Four nests of the Argentine ant, *Linepithema humile* (Mayr, 1868), were dug up from soil around Forsyth Technical Community College in Winston-Salem, North Carolina. The colonies were transported from the site and placed in Fluon-coated plastic trays in Smith Hall at Purdue

University. Foraging worker samples were taken from each colony immediately upon arrival at Purdue. The colonies were extracted from the soil and moved to test tubes wrapped in aluminum foil in the same plastic trays. Colonies were kept in the lab for a period of six months, provided water and sugar water *ad libitum*, and fed on a diet of cockroaches and modified Bhatkar-Whitcomb diet (Bhatkar and Whitcomb 1970). Every month, a sample of foragers were taken from each colony and kept in RNAlater at -80^{o}C until RNA extraction was done.

2.2.2.2 RNA Extraction and Quantification

Ant sample replicates from Months 0-6 were homogenized using a motorized micropestle apparatus. The homogenized ant samples had their RNA extracted using an SV Total RNA Isolation System Kit (Z3101, ProMega Corporation, 2800 Woods Hollow Road, Madison, WI 53711). RNA was quantified using a NanoDrop 2000 spectrophotometer (ND-2000, Thermo Fisher Scientific, 168 Third Avenue, Waltham, MA 02451). Aliquots of 200 ng RNA were converted to cDNA using a SensiFAST cDNA Synthesis Kit (BIO-65053, Bioline USA Inc, 305 Constitution Drive, Taunton, MA 02780). cDNA was aliquoted into 10µL aliquots and stored at -80^{o}C.

2.2.2.3 qPCR Target Selection

Illumina data were searched for genes from specific physiological and functional categories that show significant changes in gene expression. Thirteen genes showing diversity in direction and magnitude of expression change were chosen for qPCR analysis. Five genes that showed no changes in expression were included for comparison.

2.2.2.4 qPCR Analysis

The cDNA aliquots synthesized from 200 ng RNA were diluted 1/10 to ensure enough samples was present to conduct all analyses. Sequence data for each of the chosen targets was obtained from the Illumina analysis, and primers were generated from these sequences using NCBI Primer-BLAST. qPCR reactions were done using a SensiFAST SYBR No-ROX Kit (BIO-98005, Bioline USA Inc, 305 Constitution Drive, Taunton, MA 02780).

2.2.2.5 Statistical Analysis

For each gene, ΔCt values for Month 1 through Month 6 samples (lab samples) were compared to Month 0 samples (field samples) on a month to month basis via Mann-Whitney U-Test. Fold change was calculated through the ΔΔCt method. In addition, the ΔC_t values for each of the genes were compared to each other in a regression analysis.

2.3 Results

2.3.1 Illumina Transcriptome Analysis

Overall, the colony Illumina samples clustered more closely together based on time point (Month 1 vs. Month 6) rather than by colony identity (Figure 1). Illumina analysis discovered a total of 765 (6%) transcripts that were differentially expressed between lab and field samples (Figure 2). Of those 765 transcripts, 735 were genes while the remaining 30 transcripts were non-coding sequences (See Table 1). Most of the differentially expressed genes were downregulated (79%), although most of them only showed an average log2 fold change between 0 and -1 (77%) (Table 2). There were a small amount of genes that were upregulated (21%), and a narrow majority of them(58%) had an average log2 fold change between 0 and 1 (Table 3).

Gene Ontology (GO) analysis indicated that the most common GO terms for both upregulated and downregulated genes were for binding and catalytic activity (Tables 2 and 3). The binding category was often further broken down into ion binding, protein binding, heterocyclic and organic compound binding. The catalytic activity category was further broken down into primarily hydrolase activity, although oxidoreductase and structural molecule activity was also represented.

The volcano plot (Figure 3) shows that while many more genes showed upregulation in the transcriptome, a large portion of them were not significantly different than the field samples. Although the number of downregulated genes was more limited, a larger portion of them were found to be significant, there were more genes that exhibited both a large downregulation and a larger amount of transcriptional activity. Overall, the genes that are responsive to the transition tend towards reducing expression, likely to acclimate to changes in diet and environmental conditions.

2.3.2 qPCR Analysis

Results for the qPCR analysis were broken up into five graphs for legibility, based on general functional categories that describe the genes. The first gene category encompasses genes involved with immune defense and detoxification of allelochemicals (Figure 4). In this category, genes from the cytochrome p450 6A and 6B families show changes in gene expression over months 1-6 of lab rearing. The 6A-like gene shows upregulation from month 4 onwards, while the 6B-like gene shows sudden downregulation at month 2 before returning to prior levels.

The second gene category encompasses genes involved with storage of amino acids and sugars (Figure 6). In this category, hexamerin shows strong and consistent upregulation after

month 0, while apolipophorins and arylphorin show weaker and less consistent upregulation. Fatty acid synthase shows no change in gene expression at any time.

The third category encompasses genes that are associated with digestion (Figure 7). The only gene to show any kind of changes in this group is lysosomal aspartic protease, which shows a sudden and strong upregulation from month 1 onwards.

The fourth category encompasses genes associated with oogenesis (Figure 8). Vitellogenins 1 and 3 both showed downregulation in month 2, but showed no further changes. The other two genes showed no changes in gene expression during the time points sampled.

The final gene category encompasses genes that didn't fit into any specific category (see Figure 5). In this category, the genes sialin and clavesin both exhibited consistent upregulation from month four onwards. Aside from brief upregulation of adenosylhomocysteinase at month 5, there were no other changes in gene expression.

Finally, many of the qPCR target genes show high intercorrelation with other genes when comparing plots of their monthly qPCR ΔC_t values to one another (Figure 9). For example, some instances of correlation are to be expected, such as vitellogenins exhibiting high intercorrelation due to their sequence homology and function within the organism. But these same vitellogenins also exhibit high correlations with a diverse array of genes such as cytochrome p450s, chymotrypsins, and desaturases. There are a number of these correlations present in the qPCR dataset, such as cytochrome p450s and chymotrypsins or adenosylhomocysteinase and digestive enzymes. These correlations suggest the presence of a network or networks of interconnected genes present within the Argentine ant genome that are involved in adaptive responses to the laboratory environment.

2.4 Discussion

While the transition from field to laboratory does cause some genes to change their level of expression, it is also clear that the vast majority of the transcripts expressed by an organism remain at similar levels regardless of its environment. Based on how the transcriptome shifted over time, worker ants clearly experienced a change in how their physiology is being regulated in response to the laboratory environment. The question then becomes: which factors are driving which of the changes within the group of genes that are responding to this environmental transition? The transition from field to laboratory living is not a singular, easily isolated phenomenon, but a change in a number of different systems, each with their own downstream effects. First, there is the drastic environmental change from the field to the lab. In the field, colonies are experiencing temperature and humidity variability, weather extremes, and changes in photoperiod. Prior research shows that Argentine ants vary their foraging activity by changes in temperature and weather, and colonies have seasonal cycles driven by changes in temperature and photoperiod (Rust et al. 2000; Krusheknycky et al. 2005). The laboratory environment represents liberation from environmental variation, and provides ants with an ideal environment that enables them to continuously forage at maximum efficiency. Second, the microenvironment of the nesting and foraging area is significantly different than would be found in the field. Nests are made in tough, impermeable glass tubes wrapped in foil, rather than grainy soil or layers of wood or leaf detritus. Foraging is conducted on bare plastic, metal, and glass, rather than on plants, stone, and soil. The glass nests can hold water much more effectively than leaves and soil, and thus stable humid conditions ideal for brood growth are much easier to maintain.

Additionally, lab foragers will not encounter many of the bacterial, fungal, or plant based challenges that their field counterparts will experience over their lifetime, thanks to sanitized lab

conditions and ample supplies of clean nutrient sources provided to the lab colony. This diet itself represents another major change in the transition from lab to field. While foragers in the wild will travel far from the nest in search of nectar sources, aphids for honeydew, and protein sources both hunted and scavenged, lab colonies are provided with ample supplies of sugar water, pre-killed insects, and diet mixed from artificial materials. Not only do these ready sources of nutrients reduce the need for foraging, they likely change the biochemistry of the colony itself. Diet is one of the factors affecting the cuticular hydrocarbon profiles that colonies use to differentiate nestmate from conspecific, and is also one of the major determinants of the fauna present within the ant gut microbiome (Hu et al. 2014; Liang et al. 2000). Gut microbiomes have been shown to have important symbioses with their hosts in other insect systems, such as lignocellulose digestion in termites or pheromone production in cockroaches (Scharf et al. 2011; Wada-Katsumata et al. 2015). Such a drastic change in diet could easily cause changes in the gut microbiota of these lab colonies, which can in turn lead to downstream effects in other physiological and biochemical systems in the organism. For example, as a result of hydrocarbon and microbiome changes, the supercolonies that these lab colonies are taken from might no longer recognize them as nestmates.

The qPCR analysis was restricted to a small set of genes, but revealed detailed changes in gene expression/transcript abundance from month 0 to month 6. Comparing each of the genes to one another via pairwise regression scatterplots revealed that many of these genes shows similar co-expression, indicating that the genes showing differential regulation as a result of the lab-to-field transition might be part of a gene network or networks. While the exact relationships between some of these seemingly co-expressed genes are difficult to ascertain, some of them show comparatively obvious relationships. The first relationship within the data shows that

nutrient storage and transport proteins such as arylphorin and hexamerin are upregulated in the later months following laboratory introduction and show significant correlation in their expression with one another (Figure 8) (Burmester 1999; Telfer et al. 1983; Weers and Ryan 2006),. Previously, research seemed to indicate that these storage proteins were typically only present in immature stages, and disappear from the adult insect proteome (Burmester 1999). However, adult ants have been shown to express storage hexamerins into their adult stages (Martinez and Wheeler 1993). Ant colonies kept in the lab are typically given constant access to nutrients, whereas field colonies need to forage and thus are subject to resource scarcity, competition, and worker loss. Therefore, since nutrients are much more easily accessible in a laboratory environment (due to constant access to sugar water and regular additions of dead insect matter and artificial diet), workers are able to gather more nutrients with less energy expenditure. Lab workers appear to take advantage of this nutrient abundance by producing more proteins to adequately store and transport them within the body. However, worker ants tend to accumulate storage proteins in the hemolymph and fat body as brood production falls, typically in the fall (Wheeler and Martinez 1995). Therefore, an alternative explanation is that the ant colonies are still able to determine seasonal changes despite being kept under stable light conditions. Alternatively, while research does indicate that circadian rhythm and clock genes are involved in seasonal shifts in gene expression, it requires a change in photoperiod as well (Koštál 2011; Saunders 2010). There does not appear to be any research in the literature indicating that a "circadian calendar" gene is solely responsible for seasonal behavioral or genetic shifts in Argentine ants, so the only potential evidence for this phenomena is the storage protein gene expression patterns described in this work.

The second major relationship involves a trio of genes: *clavesin*, *sialin*, and *lysosomal aspartic protease* (LAP). Aspartic proteases are primarily involved in either digestion of food in the digestive tract or degradation of cellular debris within lysosomes (Cho et al. 1991; Dittmer and Raikhel 1997; Terra et al. 1994). Additionally, *lysosomal aspartic protease* has been specifically implicated in the breakdown of excess vitellogenin in the fat body after oogenesis takes place (Dittmer and Raikhel 1997; Raikhel 1986). Clavesins are a family of genes that are found in neuron cells, where they are involved in lysosome formation and morphology (Katoh et al. 2009). *Sialin* is a transport protein that transports aspartate and glutamate into synaptic vesicles, and transports sialic acid across lysosome membranes (Miyaji, Omote, and Moriyama 2011). Sialic acids are a versatile group of molecules, with a variety of functions (Schauer 2000). In particular, *sialin* and sialic acid have been found within neurons of *D. melanogaster* (Laridon et al. 2008; Roth et al. 2018). Together, the three genes LAP, *clavesin*, and *sialin* are all involved with lysosomes, and *clavesin* and *sialin* are both associated with lysosomes in neuron cells. These genes also show correlation in their expression with one another (Figure 8). The function of LAP within the Argentine ant remains unclear, however. Considering the upregulation in *clavesin* expression, it could be produced for neuronal lysosomes. However, research has shown that Argentine ant workers are functionally sterile, producing no eggs even when colonies lack queens (Passera et al. 1988). It is also possible that LAP is one of the factors keeping Argentine ants from producing eggs by degrading any vitellogenin that is produced for oogenesis in the fat body, and thus LAP is potentially a key gene in the maintenance of the reproductive division of labor present in the colony. Since increased diet in Argentine ant queens is correlated with increased fecundity, it is possible that the abundant diet available to the

colony is causing increased production of oogenesis-related vitellogenins, and a commensurate increase in LAP to prevent worker ovary development (Keller et al. 1989).

2.5 Conclusions

In conclusion, the results of this study have two main implications. First, ant colonies collected from the field and raised in a laboratory environment do experience changes in how their genes are expressed. These genes showing differential regulation seem to be involved in interconnected gene networks that have wide ranging effects on physiological systems such as nutrient storage and neuron function. To further elucidate what behavioral and physiological systems are changing as a result of this field-to-lab transition, experiments can be conducted in the lab by manipulating environmental variables. For instance, Argentine ant colonies that have been in the lab for long periods of time will no longer exhibit the seasonal queen execution conducted by workers to kickstart production of reproductive-destined brood (Keller et al. 1989). Creating artificial temperature and/or photoperiod variability could be done to see what cues are necessary for the performance of this behavior.

Furthermore, while research has shown that separating out worker groups and feeding them different diets can trigger aggressive behaviors between workers from the same colony (via changes in hydrocarbon decomposition), we don't know the underlying mechanism causing it (Liang et al. 2000). Considering gut microbiota assist in pheromone production in other insect groups like cockroaches, it is plausible that ant gut microbiota are one of the factors involved the creation of different cuticular hydrocarbon blends due to different dietary inputs. Testing this idea would be as simple as feeding worker groups antibiotics for a period of time and reintroducing them to their home colony and seeing if their nestmates will act aggressively towards them. Many other experiments can be done in this way, like testing the effects of plastic

vs soil substrates, or nesting substrates and water requirements, in order to properly establish how the laboratory environment can modify the behavior and physiology of the ant colonies kept there.

The second major implication of this work comes from the finding that lab and field colonies are not functionally identical in regards to patterns of gene expression. Colonies taken from the field do experience changes in how their genes are expressed as they acclimate to the laboratory environment. Since gene expression has become a popular avenue of research in examining eusocial colony structures in ants, it is important to understand that results derived from gene expression studies conducted on lab colonies are not necessarily going to be applicable to colonies in the field, and so field studies need to be conducted as well (Friedman and Gordon 2016) whenever feasible.

2.6 Figures and Tables

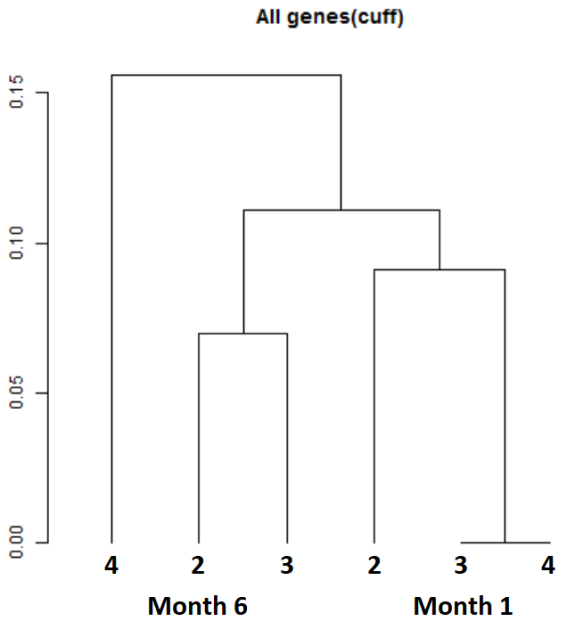


Figure 1: Dendrogram of Month-6 and Month-1 colony samples submitted for Illumina analysis. Samples cluster together by time point rather than colony number. The dendrogram was created using CummeRbund package and DGE data from cuffdiff, which provides insight into the relationships between conditions for various gene sets from sample replicates.

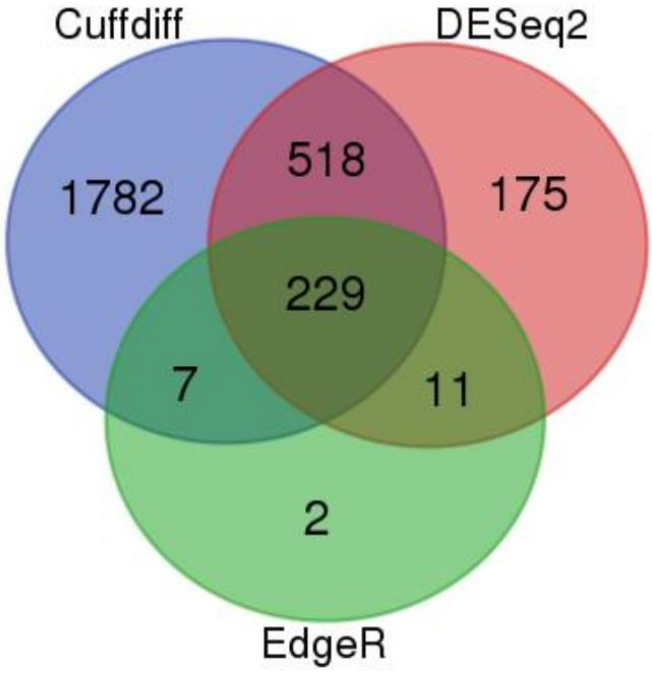


Figure 2: Venn diagram of the number of differentially expressed genes discovered by each analytical method, Cuffdiff, DESeq2, and EdgeR.

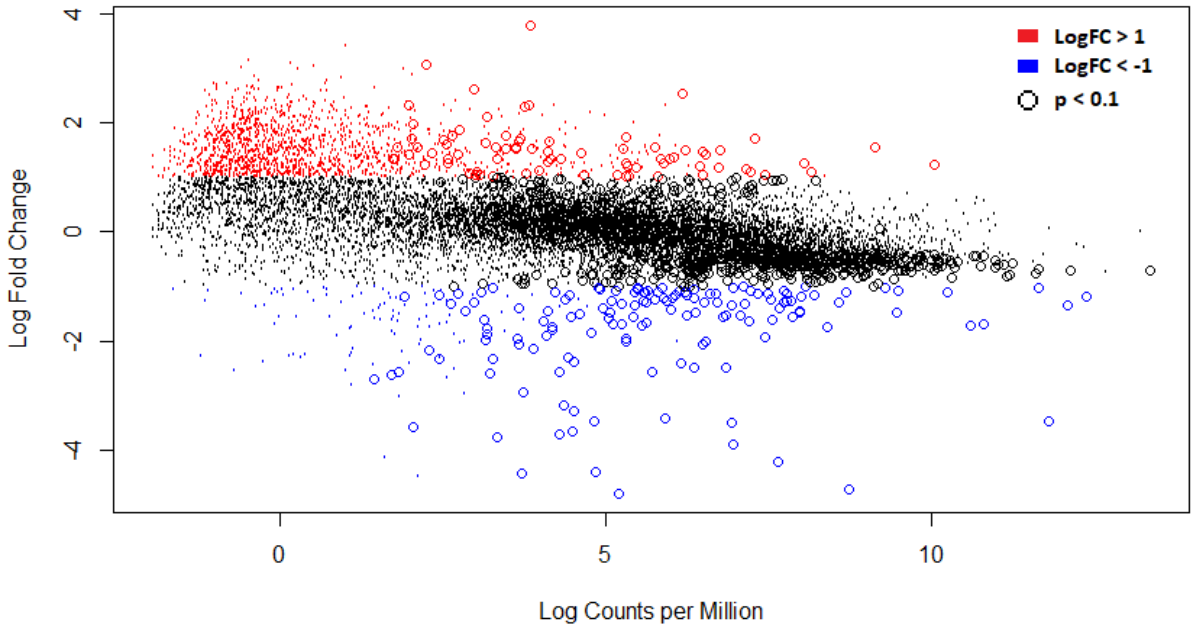


Figure 3: Volcano plot of Log Counts/Million vs average Log2 Fold Change data calculated from the three analytical methods used in the Illumina analysis, Cuffdiff, DESeq2, and EdgeR. Genes were considered significantly differently expressed if 2+ analysis methods designated them such.

Table 1: Direction and magnitude of change in expression for differentially expressed genes. Genes primarily show downregulation than upregulation.

Category	# of Genes
Upregulation	156
High Upregulation (AvgLog2FC > 2)	5
Medium Upregulation (AvgLog2FC > 1)	61
Low Upregulation (AvgLof2FC > 0)	90
Downregulation	579
Low Downregulation (AvgLof2FC < 0)	446
Medium Downregulation (AvgLof2FC < -1)	100
High Downregulation (AvgLof2FC < -2)	33
Total	735

Table 2: Frequent GO terms for downregulated genes, organized based on magnitude of downregulation. Binding GO terms are commonly downregulated.

GO #	GO Term	# of Sequences
High Downregulation (-2 > Avg Log2 FC)		33
5488	Binding	12
5515	Protein Binding	5
16787	Hydrolase Activity	4
22892	Substrate-Specific Transporter Activity	4
97367	Carbohydrate Derivative Binding	4
43167	Ion Binding	4
Unannotated		16
Medium Downregulation (-2 > Avg Log2 FC > -1)		100
5488	Binding	43
3824	Catalytic Activity	34
1901363	Heterocyclic Compound Binding	28
28	Organic Cyclic Compound Binding	28
20	Ion Binding	20
16787	Hydrolase Activity	16
15	Protein Binding	15
Unannotated		33
Low Downregulation (-1 > Avg Log2 FC)		446
5488	Binding	242
5515	Protein Binding	119
97159	Organic Cyclic Compound Binding	112
1901363	Heterocyclic Compound Binding	112
3824	Catalytic Activity	96
43167	Ion Binding	88
3676	Nucleic Acid Binding	69
Unannotated		135

Table 3: Frequent GO terms for upregulated genes, organized based on magnitude of upregulation. Binding and catalytic activity GO terms are commonly upregulated.

GO #	GO Term	# of Sequences
	High Upregulation (Avg Log2 FC > 2)	5
3824	Catalytic Activity	2
16787	Hydrolase Activity	2
Unannotated		3
	Medium Upregulation (2 > Avg Log2 FC > 1)	61
3824	Catalytic Activity	25
5488	Binding	21
16787	Hydrolase Activity	11
5515	Protein Binding	10
16491	Oxireductase Activity	8
43167	Ion Binding	8
Unannotated		18
	Low Upregulation (1 > Avg Log2 FC)	90
5488	Binding	27
3824	Catalytic Activity	19
5515	Protein Binding	16
5198	Structural Molecule Activity	11
97159	Organic Cyclic Compound Binding	11
1901363	Heterocyclic Compound Binding	11
Unannotated		35

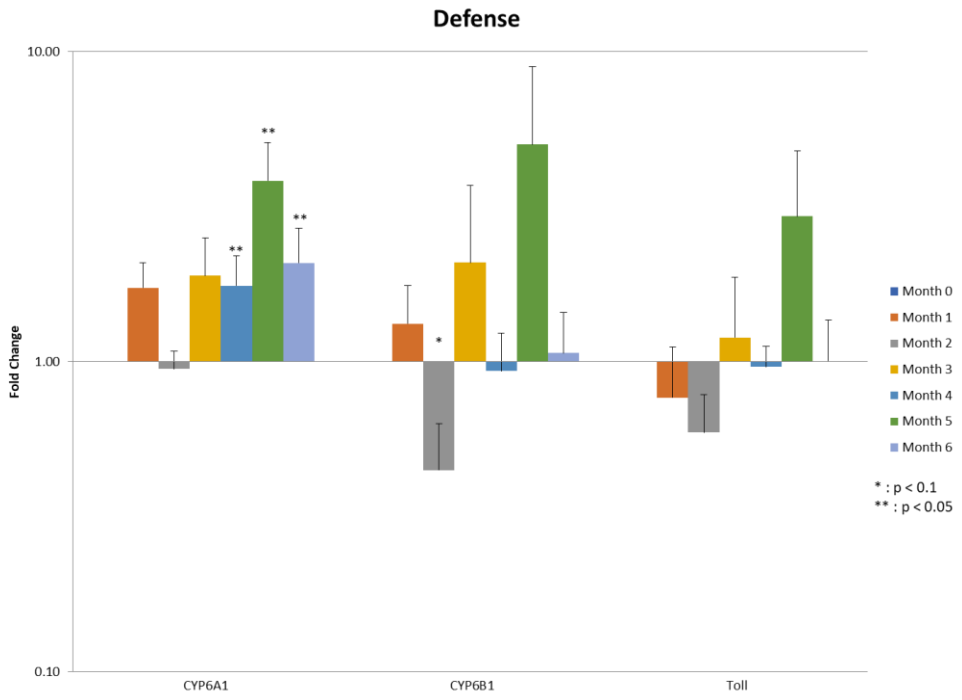


Figure 4: Monthly fold change data for genes involved in defense. CYP6A1 shows upregulation from Month 3 onwards. Month 0 results are not visible due to their normalization to 1.0.

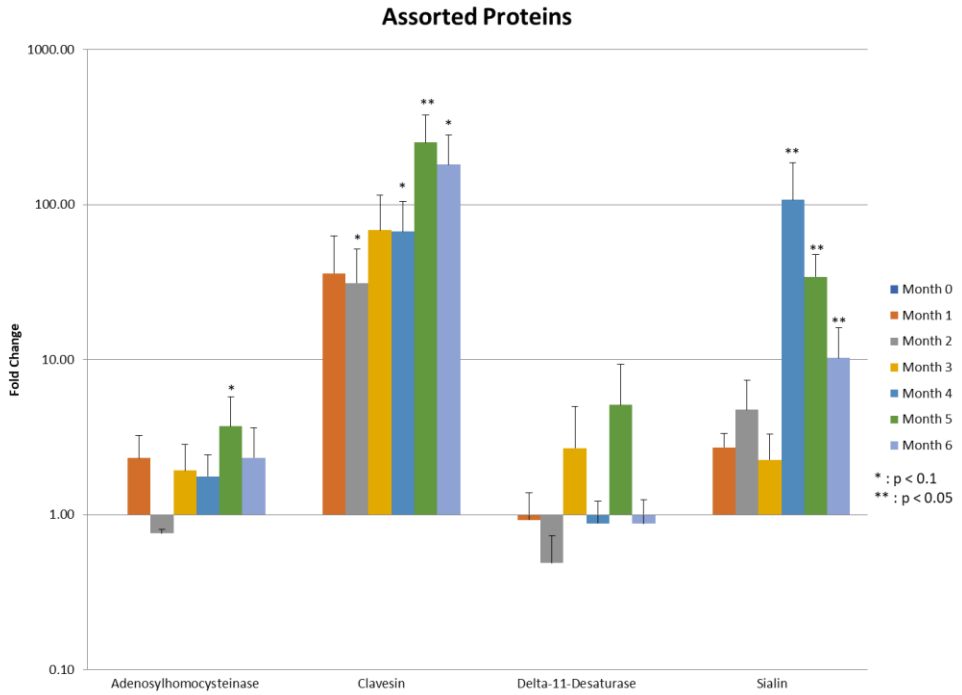


Figure 5: Monthly fold change data for genes with no particular relationships. Clavesin and sialin both show upregulation in later months. Month 0 results are not visible due to their normalization to 1.0.

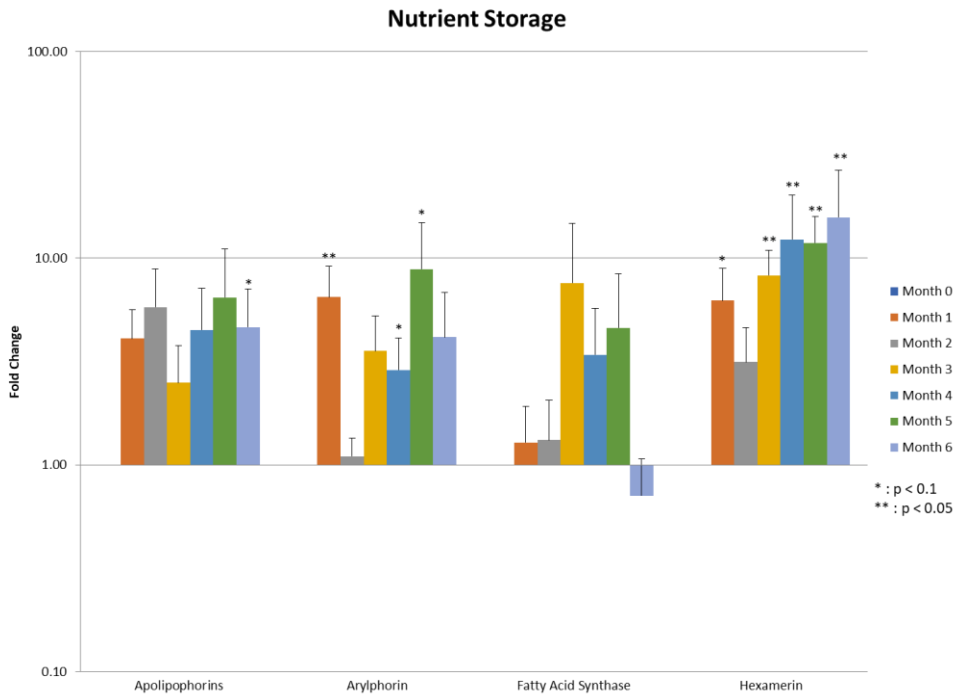


Figure 6: Monthly fold change data for genes associated with nutrient storage. Hexamerin shows consistent upregulation after lab introduction, while apolipophorins and arylphorin show less consistent upregulation in later months. Month 0 results are not visible due to their normalization to 1.0.

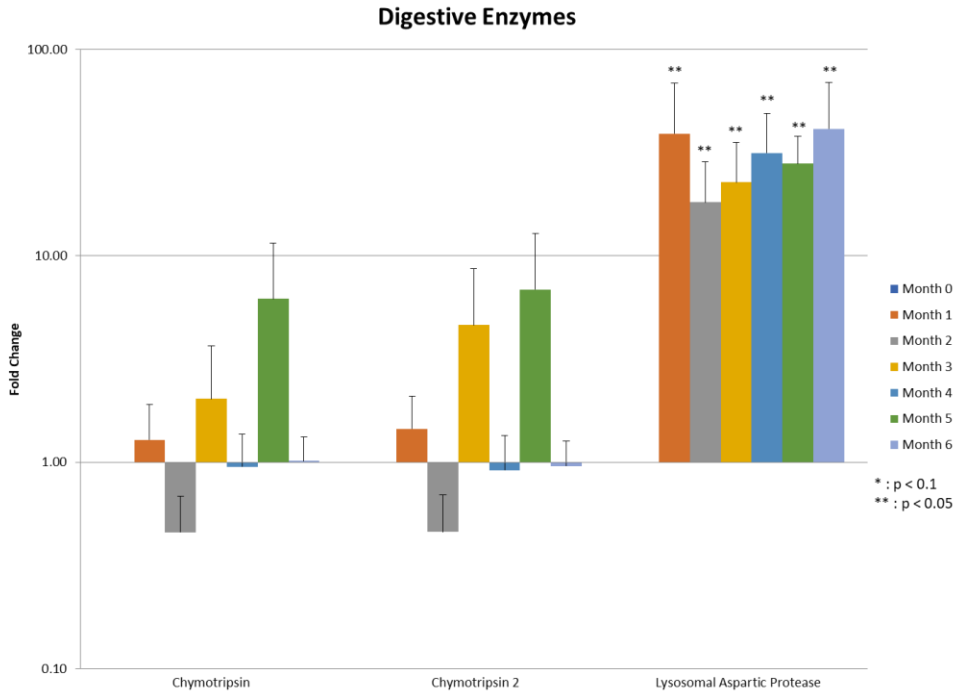


Figure 7: Monthly fold change data for genes associated with digestive enzymes. Lysosomal aspartic protease shows strong upregulation after laboratory introduction. Month 0 results are not visible due to their normalization to 1.0.

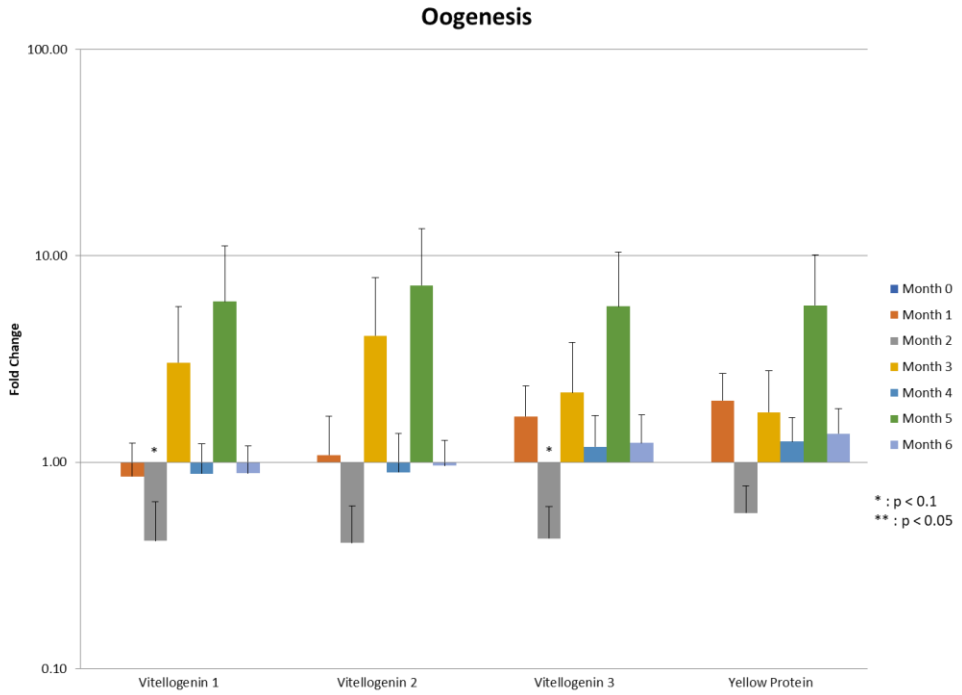


Figure 8: Monthly fold changes for genes associated with oogenesis. No consistent change in regulation was found for any of these genes. Month 0 results are not visible due to their normalization to 1.0.

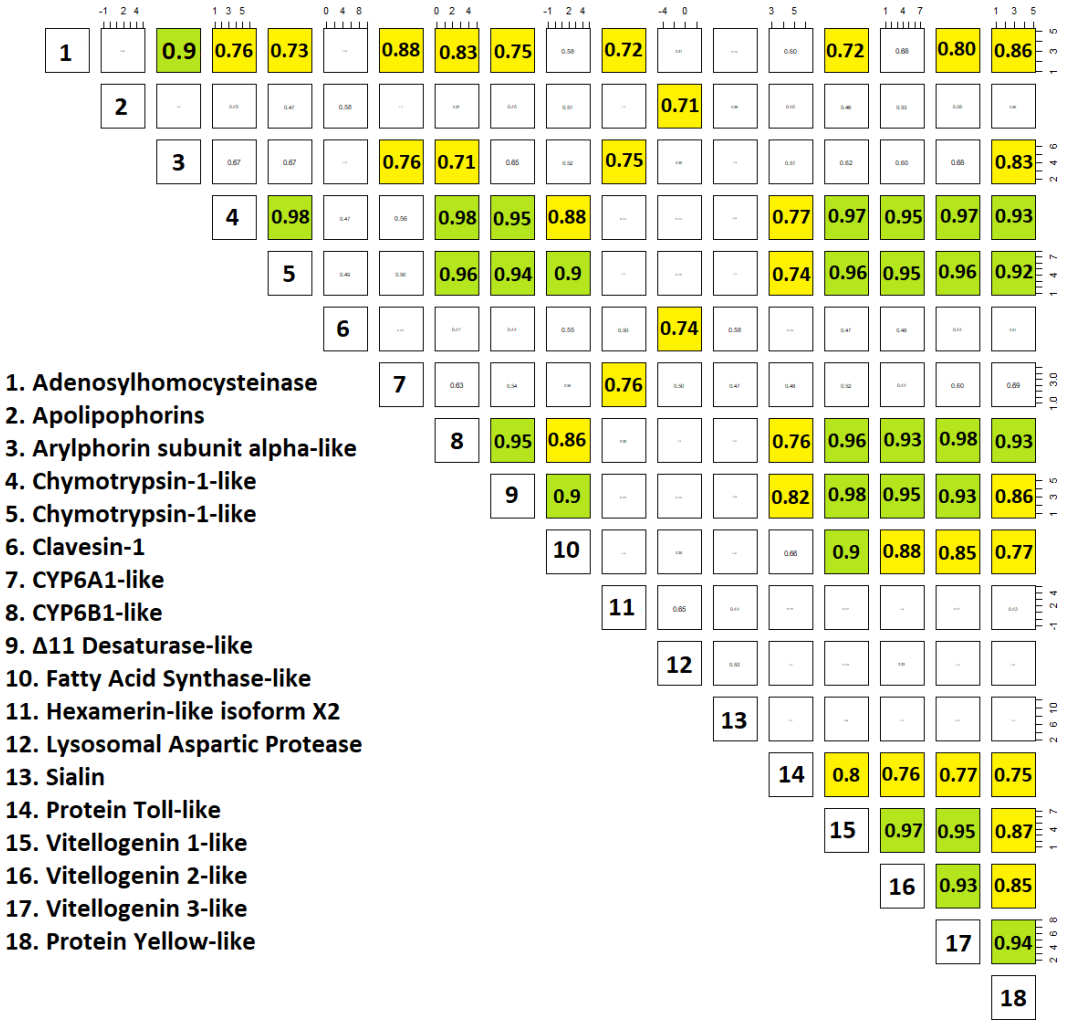


Figure 9: LOWESS plots and r^2 comparing ΔCt values for each gene with each other. The numbers 1-18 represent the gene identities as shown. Comparisons having r^2 value >0.9 are highlighted in green and those between 0.7-0.9 are highlighted yellow.

APPENDIX A. BIOINFORMATICS CORE ANALYSIS OF ILLUMINA DATA

PURDUE UNIVERSITY Discovery Park **Bioinformatics Core** ARP AGRICULTURAL RESEARCH AT PURDUE

March 29, 2017

Project Information:

PI: Dr. Gary Bennett
Analysis done by: Sagar Utturkar

Project: Differential Gene Expression analysis of laboratory-grown and wild-type ants

(1) Treatment – 6 month old ants
(2) Control – 1 month old ants

Input files: 100 base-pair Paired-end (PE) Illumina read files obtained from the Purdue Genomics Core facility.

Data Statistics:

Samples:

Treatment		Control	
2-6_1.fastq	2-6_2.fastq	2-1_1.fastq	2-1_2.fastq
3-6_1.fastq	3-6_2.fastq	3-1_1.fastq	3-2_2.fastq
4-6_1.fastq	4-6_2.fastq	4-1_1.fastq	4-1_2.fastq

Genomic resources: *Linepithema humile* genomic resources at NCBI are available at here.
The assembly file: GCF_000217595.1_Lhum_UMD_V04_genomic.fna
Annotation file: GCF_000217595.1_Lhum_UMD_V04_genomic.gff
Coding sequences: GCF_000217595.1_Lhum_UMD_V04_cds_from_genomic.fna

Materials and Methods:

Quality control: Raw Illumina reads were obtained from the Purdue Genomics Core facility. Sequence data quality was determined using **FastQC**[1] software (version 0.11.2). Quality trimming was performed using **FASTX-Toolkit**[2] (version 0.0.14) to remove the bases with less than Phred33 score of 30, and resulting reads of at least 50 base pair were retained (which comprised >99% of total reads for most samples). Reads were sorted into rRNA and non-rRNA fractions using **sortMERNA**[3] tool and non-rRNA fraction was utilized for all downstream analyses.

Read-mapping: non-rRNA reads were mapped against the indexed *Linepithema humile* reference genome using **STAR**[4] aligner (version 2.5.2b) with default parameters.

Counts: STAR derived mapping results and annotation (GFF) file for reference genome were fed to **HTSeq**[5] package (version 0.7.0) to obtain the read counts for each gene feature for each replicate. Counts from all replicates were merged together using custom Perl script to generate a gene counts matrix for both samples (Control and Treatment). Genes with 0 counts across all replicates were discarded from the counts matrix. When genes have 0 counts in one sample but not in others, the counts were converted from 0 to 1 to avoid having infinite values being calculated for fold change. Final combined counts matrix was utilized for further differential gene expression (DGE) analysis by **DESeq2**[6] and **edgeR**[7] packages. Additionally, DGE was calculated using the **tuxedo protocol**[8] which directly utilizes the STAR mapping files (bam) instead of count matrix. The tuxedo protocol uses Fragments Per Kilobase of exon per Million reads

mapped (FPKM) which is corrected (normalized) for the length of the gene and the library size to represent the gene-expression as compared to raw counts in edgeR and DESeq2.

Differential Gene Expression (DGE) analysis:
DGE analysis between samples - Control and Treatment was carried out using 'R-Bioconductor' package (version 3.3.2) and two different methods (**DESeq2** and **edgeR**). Both edgeR (version 3.16.5) and DESeq2 (version 1.14.1) use the negative binomial distribution based data model and perform specific estimate variance-mean tests. Both methods determine differentially expressed genes with *P*-value and adjusted *P*-values of false discovery rate (FDR) to correct for multiple tests. The quality of counts matrix was verified by determining basic statistics such as data range and matrix size prior to statistical tests. The DEseq2 package provides methods to test for DGE by use of negative binomial generalized linear models, the estimates of dispersion (measure for sample variance) and logarithmic fold changes. DESeq2 applied Empirical Bayes shrinkage for dispersion estimation and Wald test was used for significance testing and DGE. In **edgeR** package, an edgeR object was created using the counts matrix, and providing library sizes and experimental design. Normalization factors were calculated for the counts matrix, followed by estimation of common dispersion of counts. EdgeR package performed an 'exact' test to calculate DGE.

The tuxedo protocol starts with combined mapping files for each sample which are then processed through **Cufflinks**[8] (version 2.2.1) suite of programs (Cufflinks, Cuffmerge, Cuffquant and Cuffdiff) to determine DGE. Briefly, cufflinks perform the transcript assembly for each sample each replicate, cuffmerge combines the assemblies into a master transcriptome, cuffquant calculates the genes and transcript expression profiles and finally, cuffdiff compares these expression profiles to determine DGE. A pairwise comparison of Control and Treatment samples was performed using cuffdiff with default parameters. Each replicate was used to build a model, then these models are averaged to provide a single global model representing all conditions in the experiment and used for dispersion estimate. A t-test was performed to measure the DGE with *P*-value and also calculated adjusted *P*-values of false discovery rate (FDR) to correct for multiple tests.

Result files:

1. **Data Statistics:**
   Approximately (62-89%) of the raw reads from each sample were corresponding to non-rRNA fractions and same were used for the mapping against the reference genome. The mapping rate was in the range of (32-40%) for control samples while for treatment sample it was in the range of (69-94%). From sample 2-1, 2 million unmapped reads were randomly selected and BLASTN against non-redundant (NR) database was performed. However, 98% of the reads did not have any hits in NR database and origin of the unmapped reads could not be confirmed. "Mapping_statistics.xlsx" file describes the sorting and mapping statistics in two separate sheets. Descriptions for the "sortMERNA statistics" are self-explanatory. Descriptions of each field in "mapping statistics" are as below:
   a. Group – Group Name
   b. Sample Name – Sample Name
   c. File Name – Filename for raw sequence data file
   d. Total Reads - Total number of reads
   e. Quality Control Reads – Number of reads remained after quality trimming and filtering
   f. % Reads Passing QC – Percentage of reads remained after quality trimming and filtering

g. Reads went into mapping – Total number of reads used for mapping
h. Total Reads Mapped – Number of reads mapped to reference genome
i. Overall Mapping Rate – combined mapping rate from read pairs
j. % Mapped Reads in Gene Counts – Percentage of reads mapped only to gene features (this excludes any reads mapped within intergenic regions).

2. **Master Annotation:**

The annotations for the ant genes were fetched from the CDS (coding sequences) and GTF files and provided in the file "**Master_Annotation.xlsx**". Please note that each gene/locus correspond to multiple CDS which represent transcript isoforms of the same gene. In differential gene expression file, annotation for the one corresponding CDS have been linked. It is recommended that the user should check the master file and the annotation of all associated CDS sequences.

3. **Count Matrix:**

The "**Counts.xslx**" file describes the raw read counts for all the genes in each sample/replicate. Description of each field is as below:

a. Locus ID – Gene ID obtained from the GTF files
b. Read counts for the total 6 replicates corresponding to Control and Treatment
c. Rest of the columns correspond to annotation information

4. **Differential Gene Expression:**

Differential gene expression results are provided in the file "**Differential_Expression.xlsx**" in five different sheets. Description for each sheet is as follows:

The "**Trt_vs_Ctrl_EdgeR**" sheet described the DGE results obtained by edgeR method. Description of each field is as below:

a. Locus ID – Gene ID obtained from the annotation files
b. logFC – Describes log to base 2 value of fold change
c. Foldchange – Describes fold change within Treatment as compared to Control
d. logCPM – Log-counts-per-million values
e. P-Value – Describes probability of obtaining a test statistic at least as extreme as the one that was actually observed, assuming that the null hypothesis is true.
f. FDR- Describes false discovery rate values derived from multiple testing with Benjamini-Hochberg procedure
g. Rest of the columns correspond to annotation information

The "**Trt_vs_Ctrl_DESeq2**" sheet described the DGE results obtained by DESeq2 method. Description of each field is as below:

a. Locus ID –Gene ID obtained from annotation files
b. baseMean - Mean normalized counts, averaged over all samples from both conditions
c. log2FoldChange- Logarithm (to base 2) of the fold change
d. Foldchange - Fold change from condition Treatment to Control
e. pvalue – *P*-value for the statistical significance of this change

f. padj (FDR) – *P*-value adjusted for multiple testing with the Benjamini-Hochberg procedure
g. Rest of the columns correspond to annotation information

The "**Trt_vs_Ctrl_Cuffdiff**" sheet described the DGE results obtained by Cufflinks (Tuxedo) protocol. Description of each field is as below:

Note: Sample 1 correspond to Control and Sample 2 correspond to Treatment.

a. Locus ID – Gene ID obtained from annotation file
b. Locus – Genomic coordinates
c. Sample_1 (Control) – the first sample being compared
d. Sample_2 (Treatment) – the second sample being compared
e. Status - If there were enough samples to test then status will be "OK", if there were not enough samples then status will be "NOTEST"
f. Value_1 – FPKM of the gene in sample 1 (Control)
g. Value_2 – FPKM of the gene in sample 2 (Treatment)
h. log2foldChange - Describes log to base 2 value of fold change (Value_1/Value_2)
i. FoldChange - The fold change in Treatment samples as compared to Control
j. Test_stat – Test statistic value used to compute significance of the observed change in FPKM
k. p_value – describes the probability of obtaining a test statistic at least as extreme as the one that was actually observed, assuming that the null hypothesis is true.
l. q_value (FDR) – The FDR-adjusted p-value of the test statistic
m. Significant – whether or not the gene is differentially expressed (based on a q-score of ≤ 0.1)
n. Rest of the columns correspond to annotation information

Note: The genes described in the below sheets are after filtering at **FDR cutoff of ≤ 0.1**. **However, the user can select their own cutoff and filter genes from the original files.**

The "**Trt_vs_Ctrl_All3_Methods**" lists differentially expressed genes (FDR ≤ 0.1) which are common across all three (DESeq2/EdgeR/Cuffdiff) methods. Description of each field is as below:

a. Locus ID - Gene ID
b. DESeq2_log2FC – log2fold change from DESeq2 method
c. EdgeR_log2FC – log2fold change from edgeR method
d. cuffdiff_log2FC – log2fold change from cuffdiff method
e. Average_log2FC – average log2FC value
f. Rest of the columns correspond to annotation information

The "**Trt_vs_Ctrl_Atleast2_Methods**" sheet lists differentially expressed genes (FDR ≤ 0.1) from at least two methods. Description of each field is as below:

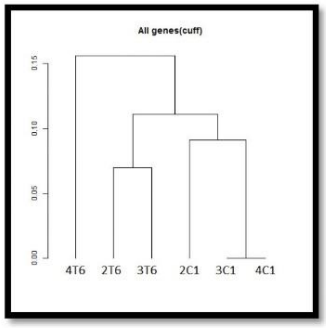
a. ID - Gene ID.
b. DESeq2_log2FC – log2fold change from DESeq2 method
c. EdgeR_log2FC – log2fold change from edgeR method
d. cuffdiff_log2FC – log2fold change from cuffdiff method
e. Average_log2FC – average log2FC value
f. Method – list of methods by which gene detected as differentially expressed

g. Rest of the columns correspond to annotation information

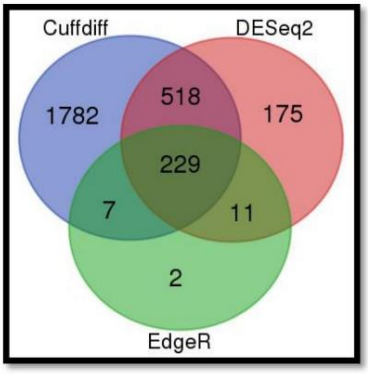
Additional figures:

Additional figures created to represent various aspects of differentially expressed genes.

1) The dendrogram was created using **CummeRbund**[9] package and DGE data from cuffdiff which provides insight into the relationships between conditions for various gene sets from sample/replicates. Please note that designations "C" and "T" for control and treatment were added to the sample names for convenience.



2) Differentially expressed genes from each method were filtered using cutoffs (FDR ≤ 0.1 for edgeR and cuffdiff data, and adjusted p-value ≤ 0.1 for DESeq2 data). Filtered results are as below.
   i) No. of significant DGE in edgeR (FDR ≤ 0.1) = 249
   ii) No. of significant DGE in DESeq2 (padj ≤ 0.1) = 933
   iii) No. of significant DGE in CUFFDIFF (FDR ≤ 0.1) = 2536



Venn diagrams were created using a tool (http://bioinformatics.psb.ugent.be/webtools/Venn/) which suggest 229 genes are differentially expressed across all three methods while total 765 genes are

differentially expressed by at least two methods. We suggest to use the set of 765 differentially expressed genes detected by at least two methods for downstream analysis.

References:

1 Babraham-Bioinformatics. *FastQC: a quality control tool for high throughput sequence data*, <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/> (2010).

2 Hannon-Lab. *FASTX-Toolkit - FASTQ/A short-reads pre-processing tools*, <http://hannonlab.cshl.edu/fastx_toolkit/index.html> (2010).

3 Kopylova, E., Noe, L. & Touzet, H. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics* **28**, 3211-3217, doi:10.1093/bioinformatics/bts611 (2012).

4 Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21, doi:10.1093/bioinformatics/bts635 (2013).

5 Anders, S., Pyl, P. T. & Huber, W. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166-169, doi:10.1093/bioinformatics/btu638 (2015).

6 Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550, doi:10.1186/s13059-014-0550-8 (2014).

7 Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139-140, doi:10.1093/bioinformatics/btp616 (2010).

8 Trapnell, C. *et al.* Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol* **28**, 511-515, doi:10.1038/nbt.1621 (2010).

9 Kellis, M. *CummeRbund: R package designed to aid and simplify the task of analyzing Cufflinks RNA-Seq output*, <http://compbio.mit.edu/cummeRbund/> (2011).

APPENDIX B. GENE ONTOLOGY ANALYSIS PLOTS

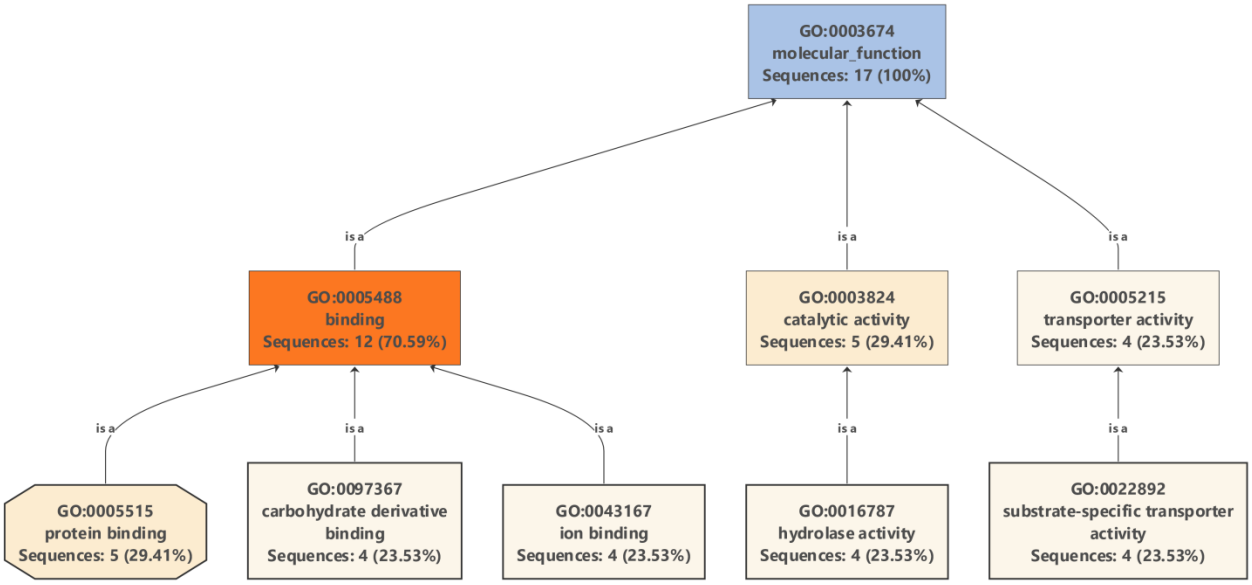


Figure 10: GO web for highly downregulated genes.

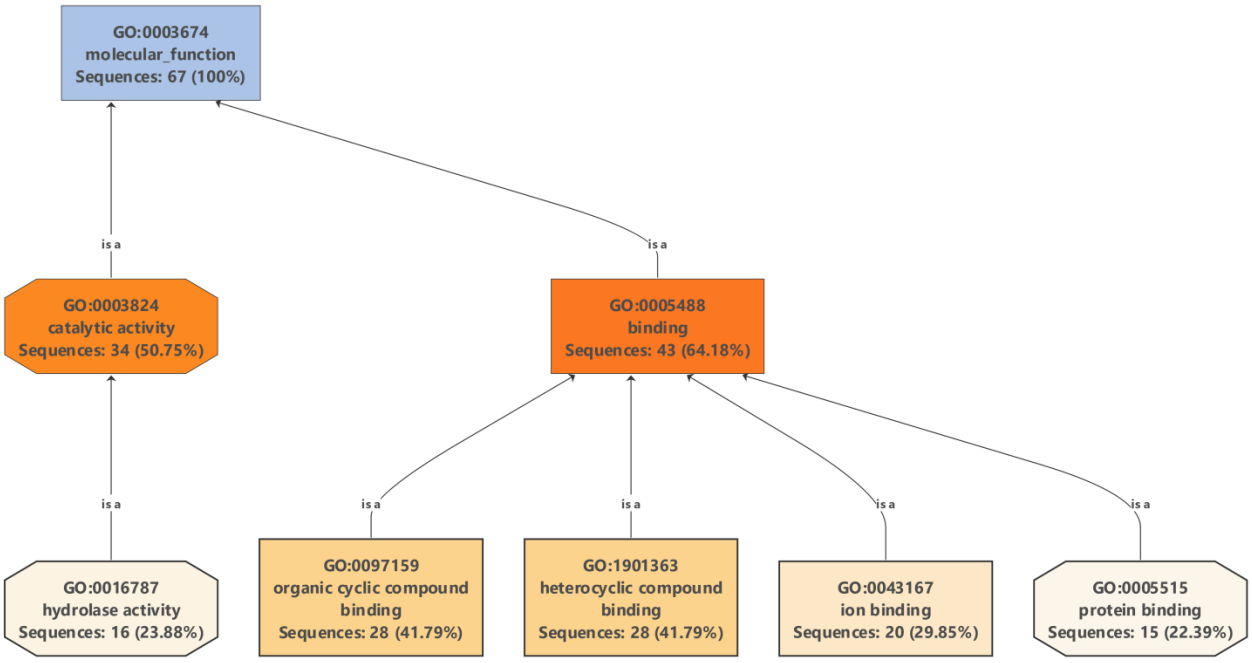


Figure 11: GO web for medium downregulated genes.

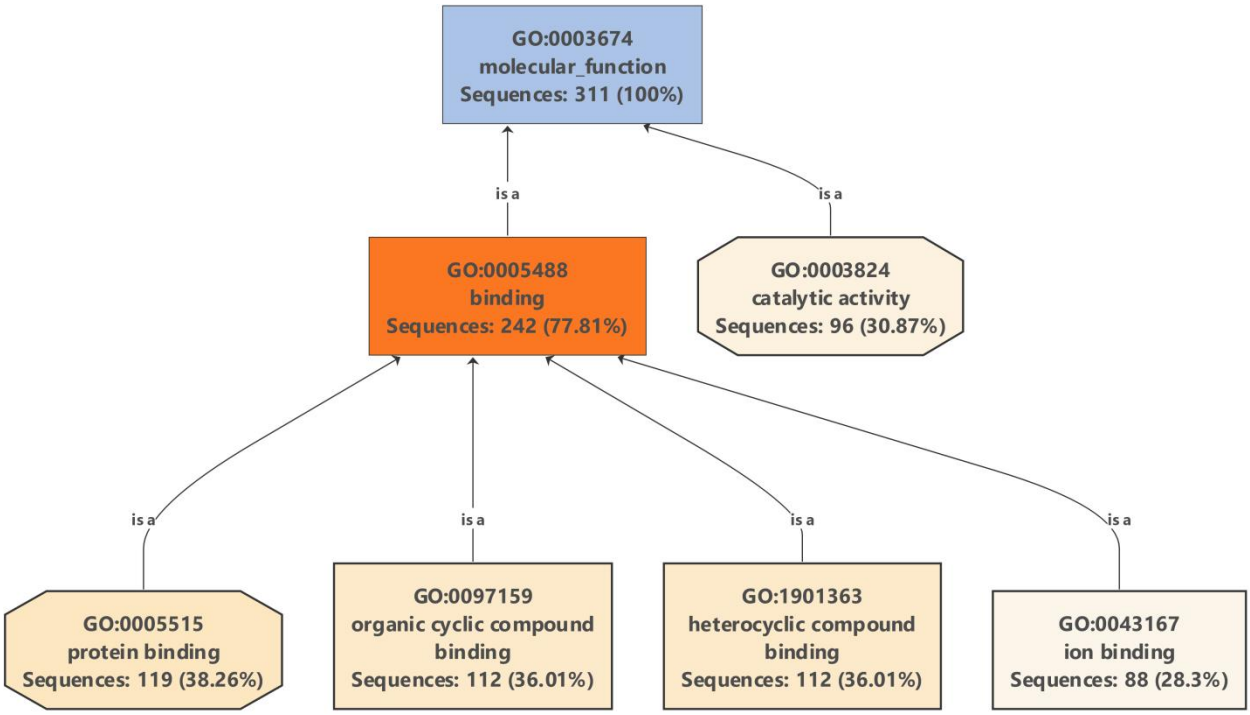


Figure 12: GO web for low downregulated genes.

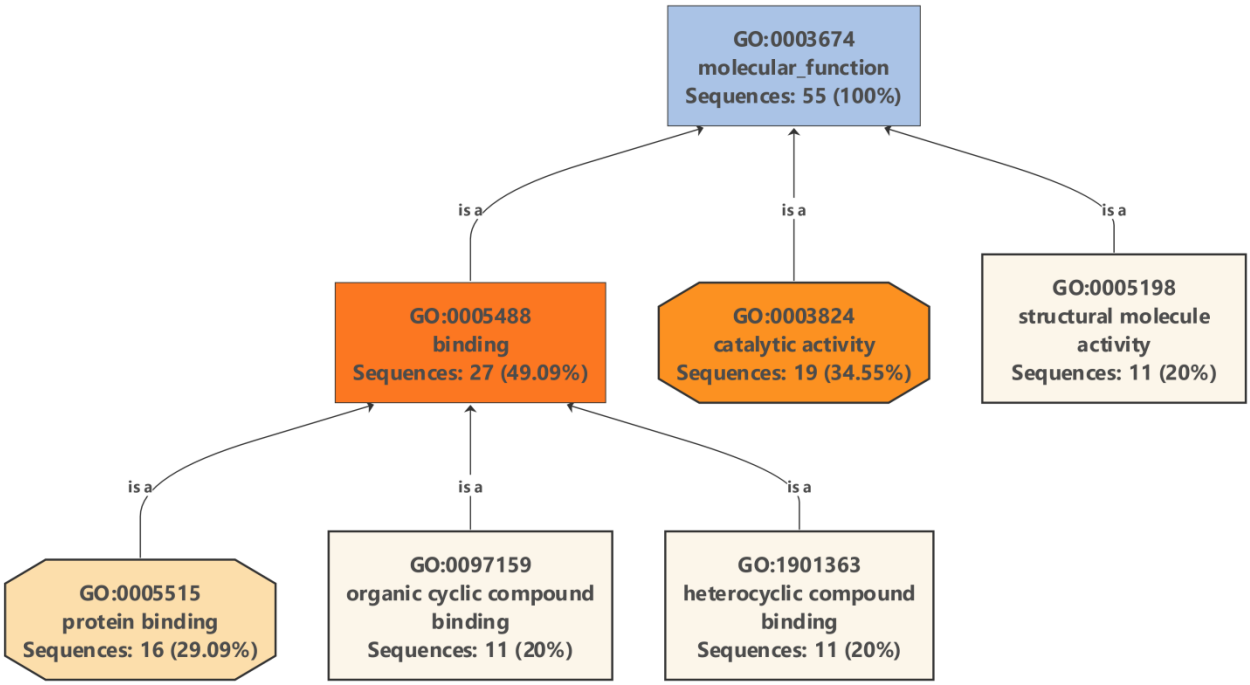


Figure 13: GO web for low upregulated genes.

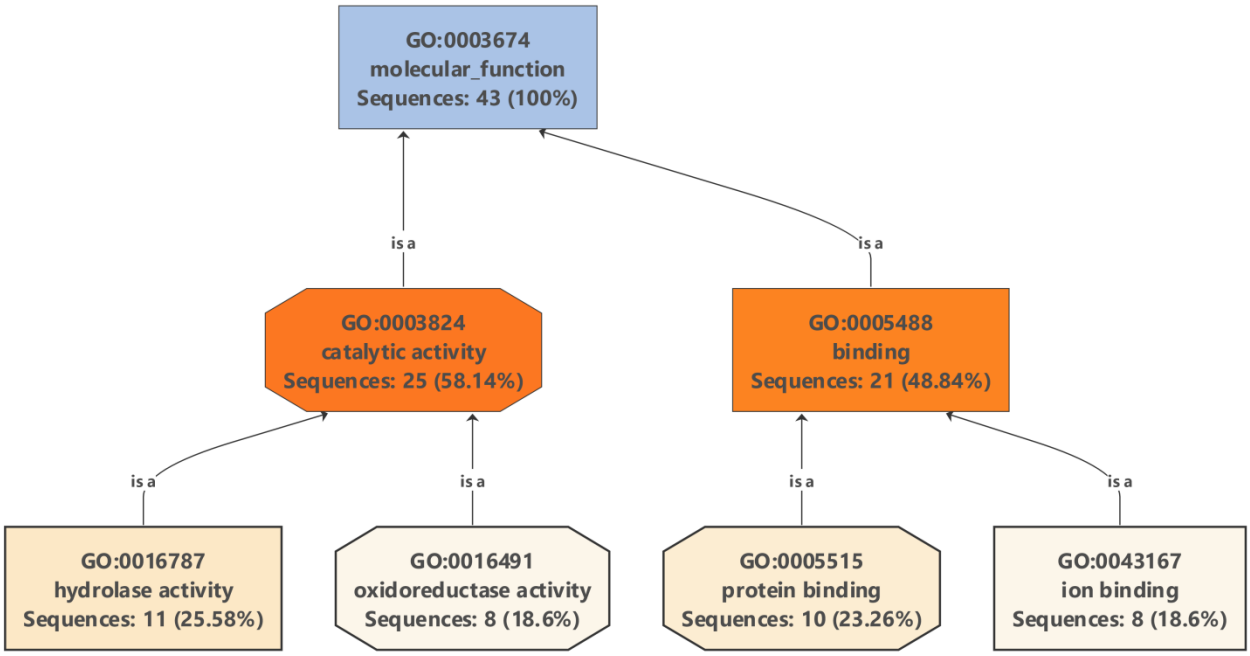


Figure 14: GO web for medium upregulated genes.

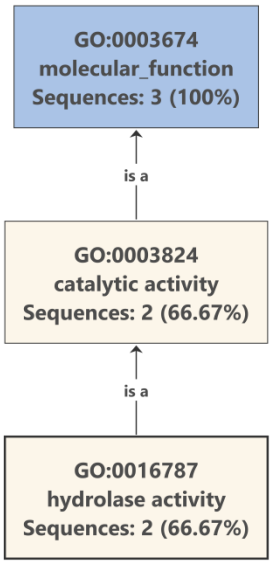


Figure 15: GO web for high upregulated genes.

APPENDIX C. QC INFORMATION FOR RNA SAMPLES

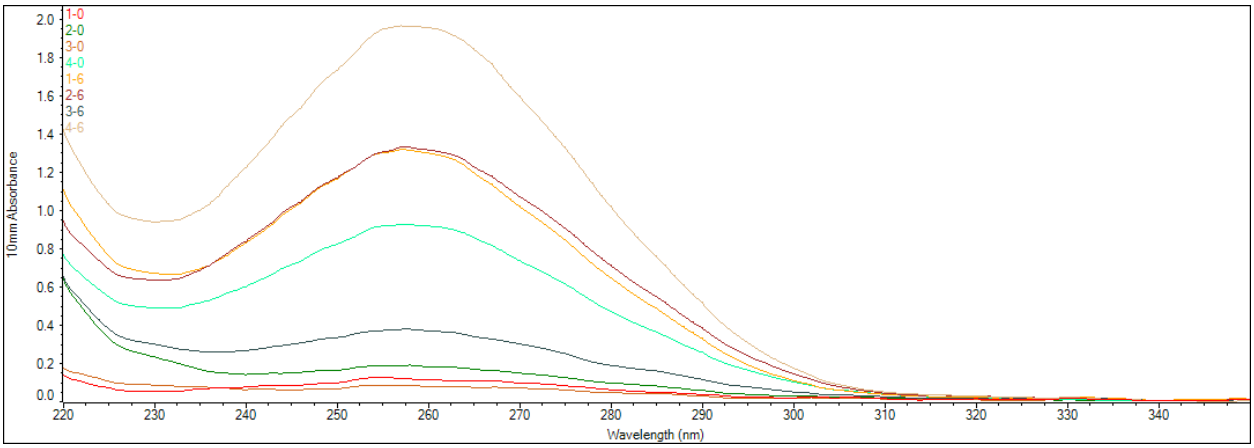


Figure 16: Nanodrop Absorbance spectra for Illumina samples.

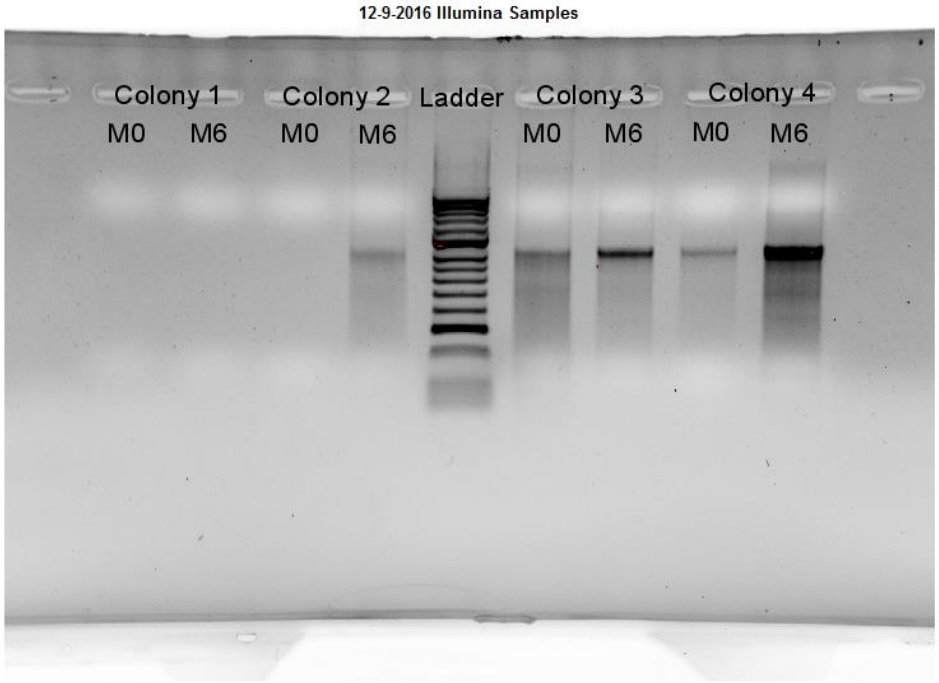


Figure 17: Agarose gel of 0 and 6 month RNA samples for Illumina analysis.

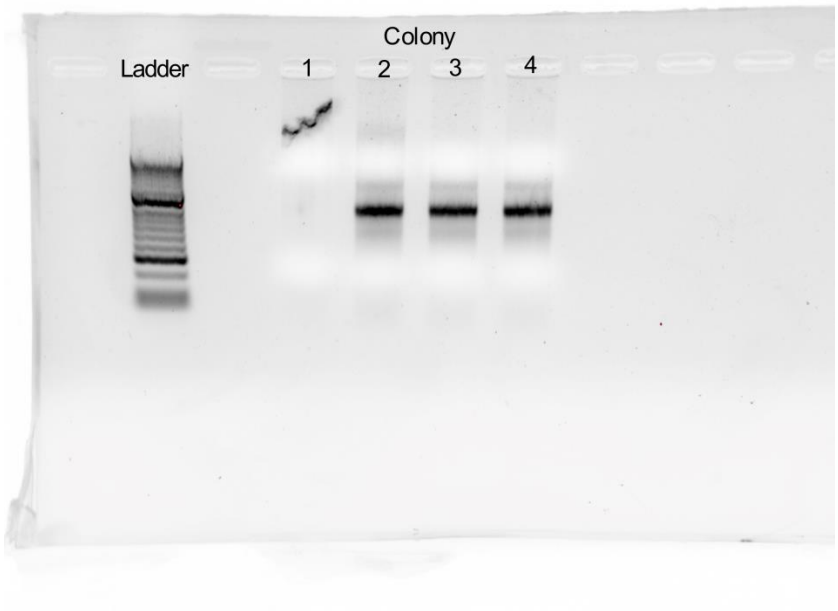


Figure 18: Agarose gel of 1 month RNA samples for Illumina analysis.

Table 4: Table showing RNA yields in (ng/µL) for each colony at each time point.

	Month 0	Month 1	Month 2	Month 3	Month 4	Month 5	Month 6
Colony 1	52.2	49	70.3	39	41.3	15	16.8
Colony 2	40.9	30.9	42	27.7	46.1	18.2	33.1
Colony 3	46.7	57.2	63.3	58.8	57	49.3	68.3
Colony 4	121.4	46.2	64.6	72.8	77.5	38.8	51.1

APPENDIX D. PRIMER INFORMATION

Table 5: Table containing gene identity and primer sequence information for each qPCR target.

Target	Accession #	Sequence	Tm (^{oC})	GC %	Product Size
actin-5, muscle-specific	XP_012231852.1	TTGGGTATGGAATCCTGCGG	68.9	55	161 bp
		AGGGCGGTGATTTCCTTCTG	67.4	55	
acyl-CoA Delta(11) desaturase-like	XP_012234970.1	GGCATCCAGTGTCTTGAGCA	66.6	50	105 bp
		ATGCCATTCTTGTCGCCTGT	66.8	55	
adenosylhomocysteinase	XP_012235416.1	CTGGACTTAGCCGAATGGGG	67.7	60	184 bp
		GAGCACCAAGTTCCACCAGT	64.2	55	
apolipophorins	XP_012225513.1	TACTGCGAGCAAAGCAAGGA	66	50	182 bp
		GCGGAGTGATCGAATTTGGC	68.9	55	
arylphorin subunit alpha-like	XP_012230053.1	GGTCTCAACATCCTCGGCAA	67.5	55	178 bp
		CGGGATCTCTCATGGACGTG	68.1	60	
chymotrypsin-1-like (1)	XP_012215860.1	CTCCTATCTGCTTGGCCGAC	66	60	207 bp
		CGCCGATTCCACGCTTATTG	69.4	55	
chymotrypsin-1-like (2)	XP_012215998.1	ATTCCCGTGGCAATGTTCCT	67.5	50	234 bp
		TTCACCCAAGATGCCTCTCG	67.6	55	
clavesin-1	XP_012233421.1	AGAGAGACCGAGGAGAATGTGA	64.2	50	221 bp
		TTCTCATCGCTCGGCATCAA	69.2	50	
cytochrome P450 6A1-like	XP_012216733.1	GGCAAAAGATCCCTGGGTGA	68.3	55	161 bp
		CGTTCGGGAAAGACGGAGAA	68.4	55	
cytochrome P450 6B1-like	XP_012234552.1	ACACCAAGCAAACTTAAGAGTATGG	63.2	40	146 bp
		GTAGCACTCAAGTCAGCGGT	62.1	55	

fatty acid synthase-like	XP_012233894.1	CTGCCGCAAGTGGTTATTCG	67.6	55	156 bp
		GCATGTGCGACGATGGAAAT	68.2	50	
hexamerin-like isoform X2	XP_012214584.1	ATCAACAGCAGCAGGTCCAA	65.8	50	187 bp
		ATGTTCTGCTCGCTCTGGAC	64.6	55	
lysosomal aspartic protease	XP_012236064.1	AGATGGGCACGACGATTTGT	66.7	50	140 bp
		TGCCGGAGCAAATCCTACTC	66.3	55	
protein toll-like	XP_012226636.1	CGCATTGATAGAGATCCACTGC	65.7	50	115 bp
		ACTCGGTAAATTGCAGTTGTCG	65	45.4	
protein yellow-like	XP_012227993.1	ACAGGACGGCGACAAAATCT	65.8	50	249 bp
		GGATGCGTGAGACTGTTCCA	66.4	55	
sialin	XP_012216319.1	ACCGAGCAACAGTAGCAGTC	62.1	55	193 bp
		TTCTCGTTCCTCCAAGTGCC	66.5	55	
vitellogenin-1-like	XP_012221881.1	ACAGCCCCGTGAAATAGTCG	66.1	55	143 bp
		TTAGCACGCCCTGTTCAACT	64.6	50	
vitellogenin-2-like	XP_012221806.1	AGGCCGAAAACCAGTGACAA	66.7	50	196 bp
		ACGTGTCAGTGCAGTCAGAG	61.9	55	
vitellogenin-3-like	XP_012221453.1	GATGTACGGTCCGGAATGCA	68	55	177 bp
		TGCAGATCCGGCCGTTTAAT	68.4	50	

APPENDIX E. A SUMMARY OF ILLUMINA ANALYSIS (USING CUFFDIFF, EDGE-R, AND DESEQ2) SHOWING GENES PASSING IN ALL THREE METHODS

Table 6: Table containing Illumina data for qPCR targets that do not pass all three Illumina analysis methods.

qPCR Targets Chosen From Illumina Analysis Showing Genes Passing in Two Methods									
Locus_ID	**DeSeq2_log2FC**	**EdgeR_log2FC**	**cuffdiff_log2FC**	**Average_log2FC**	**Method**	**CDS_ID**	**Contig_ID**	**Protein_ID**	**Desription**
LOC105674706	1.284008945	1.714645638	2.93932	1.979324861	Cuffdiff DESeq2	lcl|NW_012160746.1_cds_XP_012226636.1_14435	NW_012160746.1	protein_id=XP_012226636.1	protein toll-like
LOC105668478	-0.702527986	-0.790845068	-0.90264	-0.798671018	Cuffdiff DESeq2	lcl|NW_012160474.1_cds_XP_012216319.1_4884	NW_012160474.1	protein_id=XP_012216319.1	sialin
LOC105675420	0.619486182	0.691454878	0.862012	0.724317687	Cuffdiff DESeq2	lcl|NW_012160764.1_cds_XP_012227993.1_15937	NW_012160764.1	protein_id=XP_012227993.1	protein yellow-like
qPCR Targets Chosen From Illumina Analysis Showing Genes Exhibiting No Differential Expression									
LOC105674040	-0.275107863	-0.315291394	-0.242484			lcl|NW_012160731.1_cds_XP_012225513.1_13454	NW_012160731.1	protein_id=XP_012225513.1	apolipophorins
LOC105668771	-0.310016497	-0.352483085	-0.164229			lcl|NW_012160514.1_cds_XP_012216733.1_5276	NW_012160514.1	protein_id=XP_012216733.1	cytochrome P450 6A1-like
LOC105679217	0.011874723	0.008073747	0.146802			lcl|NW_012160246.1_cds_XP_012234552.1_1916	NW_012160246.1	protein_id=XP_012234552.1	cytochrome P450 6B1-like
LOC105671912	-0.520813912	-1.626357836	-2.50284			lcl|NW_012160688.1_cds_XP_012221881.1_10115	NW_012160688.1	protein_id=XP_012221881.1	vitellogenin-1-like
LOC105671676	-0.55232368	-0.671696384	-0.680662			lcl|NW_012160685.1_cds_XP_012221453.1_9856	NW_012160685.1	protein_id=XP_012221453.1	vitellogenin-3-like

Table 7: Table containing Illumina data for all transcripts that passed all three Illumina analysis methods (Cuffdiff, DESeq2, and EdgeR)

Locus_ID	**DeSeq2_log2FC**	**EdgeR_log2FC**	**cuffdiff_log2FC**	**Average_log2FC**	**CDS_ID**	**Contig_ID**	**Protein_ID**	**Desription**
LOC105675317	-0.685621478	-0.724505094	-1.09311	-0.834412191	lcl|NW_012160764.1_cds_XP_012227766.1_15726	NW_012160764.1	protein_id=XP_012227766.1	LOW QUALITY PROTEIN: RING finger protein 44
LOC105677569	-0.771354937	-0.823687464	-0.77817	-0.7910708	lcl|NW_012160815.1_cds_XP_012231709.1_18954	NW_012160815.1	protein_id=XP_012231709.1	protein charybde-like
LOC105671992	-0.666905606	-0.686118007	-0.73331	-0.695444538	lcl|NW_012160689.1_cds_XP_012222038.1_10244	NW_012160689.1	protein_id=XP_012222038.1	glutamate-gated chloride channel isoform X6
LOC105667639	-0.802161083	-0.847958261	-0.920689	-0.856936115	lcl|NW_012160412.1_cds_XP_012214981.1_3670	NW_012160412.1	protein_id=XP_012214981.1	slit homolog 3 protein-like isoform X2
LOC105670012	-0.913864323	-1.023926358	-1.18691	-1.041566894	lcl|NW_012160620.1_cds_XP_012218710.1_7195	NW_012160620.1	protein_id=XP_012218710.1	LOW QUALITY PROTEIN: putative ATP-dependent RNA helicase me31b
LOC105678264	-0.95030706	-1.052725311	-0.907185	-0.970072457	lcl|NW_012160832.1_cds_XP_012232877.1_20095	NW_012160832.1	protein_id=XP_012232877.1	LOW QUALITY PROTEIN: uncharacterized protein LOC105678264
LOC105679444	-1.180057822	-1.379647831	-1.48687	-1.348858551	lcl|NW_012160300.1_cds_XP_012234900.1_2164	NW_012160300.1	protein_id=XP_012234900.1	farnesol dehydrogenase-like
LOC105668134	1.030384425	1.230449922	2.27925	1.513361449		NW_012160461.1		lncRNA
LOC105670425	-0.707972969	-0.747474442	-0.571871	-0.675772804	lcl|NW_012160657.1_cds_XP_012219367.1_7816	NW_012160657.1	protein_id=XP_012219367.1	uncharacterized protein LOC105670425 isoform X1

LOC105671886	-0.681360824	-0.702130463	-0.711422	-0.698304429	lcl|NW_012160688.1_cds_XP_012221828.1_10126	NW_012160688.1	protein_id=XP_012221828.1	serine/threonine-protein phosphatase 2B catalytic subunit 2-like isoform X5
LOC105671872	-1.921620397	-2.576076164	-2.74118	-2.412958854	lcl|NW_012160688.1_cds_XP_012221806.1_10114	NW_012160688.1	protein_id=XP_012221806.1	vitellogenin-2-like
LOC105678792	-0.701068795	-0.725461484	-0.777604	-0.734711426	lcl|NW_012160840.1_cds_XP_012233859.1_21093	NW_012160840.1	protein_id=XP_012233859.1	zinc finger protein 395 isoform X2
LOC105670289	-0.96534783	-1.03176562	-0.976947	-0.991353483	lcl|NW_012160649.1_cds_XP_012219149.1_7570	NW_012160649.1	protein_id=XP_012219149.1	troponin T, skeletal muscle isoform X3
LOC105672663	0.873146264	0.957139	1.03079	0.953691755	lcl|NW_012160700.1_cds_XP_012223168.1_11293	NW_012160700.1	protein_id=XP_012223168.1	40S ribosomal protein S16
LOC105669680	-0.649057269	-0.680438481	-0.639077	-0.656190917	lcl|NW_012160613.1_cds_XP_012218168.1_6728	NW_012160613.1	protein_id=XP_012218168.1	glucose transporter type 1 isoform X10
LOC105668345	-0.951311014	-1.025746973	-1.05636	-1.011139329	lcl|NW_012160474.1_cds_XP_012216103.1_4592	NW_012160474.1	protein_id=XP_012216103.1	mediator of DNA damage checkpoint protein 1-like
LOC105673048	1.328986576	1.511698023	1.76704	1.5359082	lcl|NW_012160715.1_cds_XP_012223802.1_12157	NW_012160715.1	protein_id=XP_012223802.1	uncharacterized protein LOC105673048
LOC105679107	-0.697637211	-0.743729331	-0.78341	-0.741592181	lcl|NW_012160844.1_cds_XP_012234343.1_21434	NW_012160844.1	protein_id=XP_012234343.1	serine/threonine-protein kinase PLK1-like isoform X1
LOC105670150	-0.875362147	-0.937357753	-1.58277	-1.131829967	lcl|NW_012160636.1_cds_XP_012218921.1_7389	NW_012160636.1	protein_id=XP_012218921.1	zinc finger protein 362
LOC105679506	-0.800128744	-0.895639062	-0.789337	-0.828368269	lcl|NW_012160307.1_cds_XP_012234995.1_2233	NW_012160307.1	protein_id=XP_012234995.1	LOW QUALITY PROTEIN: fatty acid synthase-like
LOC105678989	1.047022644	1.317742124	1.38833	1.251031589	lcl|NW_012160841.1_cds_XP_012234177.1_21223	NW_012160841.1	protein_id=XP_012234177.1	carbonic anhydrase 2-like isoform X1
LOC105676549	-0.995500571	-1.140027101	-1.3275	-1.154342557	lcl|NW_012160211.1_cds_XP_012229974.1_1507	NW_012160211.1	protein_id=XP_012229974.1	fatty acyl-CoA reductase 1
LOC105673922	0.997026329	1.117261552	1.30471	1.13966596	lcl|NW_012160180.1_cds_XP_012225341.1_1041	NW_012160180.1	protein_id=XP_012225341.1	iron-sulfur cluster assembly enzyme ISCU, mitochondrial
LOC105667982	-1.135761951	-1.421462629	-3.19948	-1.918901527	lcl|NW_012160432.1_cds_XP_012215555.1_4337	NW_012160432.1	protein_id=XP_012215555.1	LOW QUALITY PROTEIN: uncharacterized protein LOC105667982
LOC105669832	1.069353406	1.45544787	1.93228	1.485693759	lcl|NW_012160615.1_cds_XP_012218397.1_6795	NW_012160615.1	protein_id=XP_012218397.1	phospholipase A2-like
LOC105677373	-1.596924735	-2.284639912	-2.54303	-2.141531549	lcl|NW_012160815.1_cds_XP_012231372.1_18716	NW_012160815.1	protein_id=XP_012231372.1	LOW QUALITY PROTEIN: probable WRKY transcription factor protein 1
LOC105676065	-2.646110295	-3.413688921	-4.19505	-3.418283072	lcl|NW_012160776.1_cds_XP_012229079.1_16747	NW_012160776.1	protein_id=XP_012229079.1	uncharacterized protein LOC105676065
LOC105674129	-1.274516524	-1.450162236	-1.54457	-1.42308292	lcl|NW_012160732.1_cds_XP_012225684.1_13579	NW_012160732.1	protein_id=XP_012225684.1	poly(A) RNA polymerase gld-2 homolog A-like
LOC105673828	-0.645660035	-0.670145087	-0.651218	-0.655674374	lcl|NW_012160728.1_cds_XP_012225189.1_13074	NW_012160728.1	protein_id=XP_012225189.1	DEAD-box ATP-dependent RNA helicase 20-like
LOC105669975	-1.100874711	-1.200824318	-1.41277	-1.238156343	lcl|NW_012160620.1_cds_XP_012218633.1_7163	NW_012160620.1	protein_id=XP_012218633.1	UDP-glucuronosyltransferase-like isoform X3
LOC105669816	-1.022512412	-1.133209415	-1.2921	-1.149273942	lcl|NW_012160615.1_cds_XP_012218375.1_6869	NW_012160615.1	protein_id=XP_012218375.1	ubiquitin carboxyl-terminal hydrolase 36
LOC105678225	-0.634297608	-0.661562609	-0.591601	-0.629153739	lcl|NW_012160832.1_cds_XP_012232797.1_20168	NW_012160832.1	protein_id=XP_012232797.1	protein unc-13 homolog B isoform X9
LOC105674948	-0.798370096	-0.841880464	-0.867988	-0.83607952	lcl|NW_012160750.1_cds_XP_012227064.1_14835	NW_012160750.1	protein_id=XP_012227064.1	putative uncharacterized protein DDB_G0271606
LOC105675705	-0.90455984	-0.982291876	-1.44639	-1.111080572	lcl|NW_012160768.1_cds_XP_012228460.1_16260	NW_012160768.1	protein_id=XP_012228460.1	uncharacterized protein LOC105675705 isoform X1
LOC105672377	0.727458541	0.76378484	0.972448	0.82123046	lcl|NW_012160698.1_cds_XP_012222695.1_10901	NW_012160698.1	protein_id=XP_012222695.1	uncharacterized protein LOC105672377
LOC105679585	-0.733927216	-0.761002296	-0.695683	-0.730204171	lcl|NW_012160332.1_cds_XP_012235113.1_2369	NW_012160332.1	protein_id=XP_012235113.1	sodium channel protein para isoform X13
LOC105678133	-1.346455336	-1.514664257	-1.59184	-1.484319864	lcl|NW_012160830.1_cds_XP_012232641.1_19826	NW_012160830.1	protein_id=XP_012232641.1	muscle LIM protein Mlp84B-like isoform X5
LOC105670100	1.182800465	1.322123085	1.50654	1.337154517	lcl|NW_012160627.1_cds_XP_012218853.1_7352	NW_012160627.1	protein_id=XP_012218853.1	cubilin-like
LOC105668197	-1.253122802	-1.326344144	-1.33007	-1.303178982	lcl|NW_012160474.1_cds_XP_012215850.1_4681	NW_012160474.1	protein_id=XP_012215850.1	uncharacterized protein LOC105668197

LOC105675857	-0.761187008	-0.788201881	-0.689855	-0.74641463	lcl|NW_012160769.1_cds_XP_012228748.1_16323	NW_012160769.1	protein_id=XP_012228748.1	potassium voltage-gated channel protein Shaw-like isoform X2
LOC105668293	1.557243304	1.821101587	1.70215	1.693498297	lcl|NW_012160474.1_cds_XP_012215998.1_5039	NW_012160474.1	protein_id=XP_012215998.1	chymotrypsin-1-like
LOC105679844	-0.771141842	-0.848069834	-0.807949	-0.809053559	lcl|NW_012160333.1_cds_XP_012235551.1_2706	NW_012160333.1	protein_id=XP_012235551.1	uncharacterized protein LOC105679844 isoform X1
LOC105669845	-0.963368495	-1.090594949	-1.02336	-1.025774481	lcl|NW_012160615.1_cds_XP_012218432.1_6805	NW_012160615.1	protein_id=XP_012218432.1	3-ketoacyl-CoA thiolase, mitochondrial isoform X1
LOC105672768	-0.76488386	-0.829685398	-1.08417	-0.892913086	lcl|NW_012160707.1_cds_XP_012223342.1_11468	NW_012160707.1	protein_id=XP_012223342.1	solute carrier family 2, facilitated glucose transporter member 1-like isoform X3
LOC105670685	-0.981684153	-1.413301053	-1.57269	-1.322558402	lcl|NW_012160665.1_cds_XP_012219762.1_8201	NW_012160665.1	protein_id=XP_012219762.1	glycine-rich RNA-binding protein 1-like
LOC105668327	-1.022707172	-1.113946737	-1.08308	-1.073244636	lcl|NW_012160474.1_cds_XP_012216077.1_4663	NW_012160474.1	protein_id=XP_012216077.1	uncharacterized protein LOC105668327 isoform X8
LOC105671505	0.814527694	0.901657243	1.07717	0.931118312	lcl|NW_012160683.1_cds_XP_012221159.1_8937	NW_012160683.1	protein_id=XP_012221159.1	60S ribosomal protein L14
LOC105670392	-1.636020008	-1.983765124	-2.43339	-2.017725044	lcl|NW_012160651.1_cds_XP_012219322.1_7730	NW_012160651.1	protein_id=XP_012219322.1	maternal protein exuperantia
LOC105675449	-1.485896434	-5.682870558	-7.1965	-4.788422331	lcl|NW_012160764.1_cds_XP_012228047.1_15969	NW_012160764.1	protein_id=XP_012228047.1	uncharacterized protein LOC105675449 isoform X1
LOC105673337	-0.691375777	-0.721203139	-0.698123	-0.703567305	lcl|NW_012160717.1_cds_XP_012224321.1_12308	NW_012160717.1	protein_id=XP_012224321.1	uncharacterized protein LOC105673337 isoform X10
LOC105676035	-0.74970006	-0.835580675	-0.808793	-0.798024578	lcl|NW_012160774.1_cds_XP_012229021.1_16584	NW_012160774.1	protein_id=XP_012229021.1	paramyosin, long form-like
LOC105673415	1.386317455	1.506762328	1.63087	1.507983261	lcl|NW_012160717.1_cds_XP_012224453.1_12446	NW_012160717.1	protein_id=XP_012224453.1	FMRFamide receptor-like isoform X1
LOC105677668	-1.030994196	-1.189435548	-1.24481	-1.155079915	lcl|NW_012160816.1_cds_XP_012231851.1_19159	NW_012160816.1	protein_id=XP_012231851.1	actin-5, muscle-specific
LOC105670431	-1.025316874	-1.138017034	-1.17535	-1.112894636	lcl|NW_012160659.1_cds_XP_012219375.1_7821	NW_012160659.1	protein_id=XP_012219375.1	glucosylceramidase-like
LOC105671968	-0.749844613	-0.789916206	-0.532095	-0.690618606	lcl|NW_012160689.1_cds_XP_012221978.1_10258	NW_012160689.1	protein_id=XP_012221978.1	LOW QUALITY PROTEIN: uncharacterized protein LOC105671968
LOC105669230	-1.095902969	-1.205160003	-1.24237	-1.181144324	lcl|NW_012158052.1_cds_XP_012219470.1_91	NW_012158052.1	protein_id=XP_012219470.1	twitchin isoform X26
LOC105673345	-1.401822618	-1.635620855	-2.10411	-1.713851158	lcl|NW_012160717.1_cds_XP_012224336.1_12476	NW_012160717.1	protein_id=XP_012224336.1	LOW QUALITY PROTEIN: uncharacterized protein LOC105673345
LOC105680109	-1.418923694	-1.603289779	-1.57815	-1.533454491		NW_012158668.1		Pseudogene
LOC105673753	1.481017502	1.552903574	0.96975	1.334557025	lcl|NW_012160728.1_cds_XP_012225027.1_12947	NW_012160728.1	protein_id=XP_012225027.1	uncharacterized protein LOC105673753 isoform X2
LOC105679121	-0.706075502	-0.737911704	-0.81442	-0.752802402	lcl|NW_012160844.1_cds_XP_012234377.1_21409	NW_012160844.1	protein_id=XP_012234377.1	uncharacterized protein LOC105679121 isoform X6
LOC105668241	-1.149773084	-2.160487166	-2.43408	-1.914780083	lcl|NW_012160474.1_cds_XP_012215909.1_5001	NW_012160474.1	protein_id=XP_012215909.1	endochitinase
LOC105676735	-0.746004751	-0.794996416	-0.742069	-0.761023389	lcl|NW_012160785.1_cds_XP_012230271.1_17757	NW_012160785.1	protein_id=XP_012230271.1	homeobox protein onecut-like
LOC105678543	-0.947188228	-1.115087612	-1.06743	-1.04323528	lcl|NW_012160836.1_cds_XP_012233421.1_20461	NW_012160836.1	protein_id=XP_012233421.1	clavesin-1
LOC105667815	-0.722169869	-0.78183103	-0.887797	-0.797265966	lcl|NW_012160418.1_cds_XP_012215290.1_4121	NW_012160418.1	protein_id=XP_012215290.1	rho-associated protein kinase 1 isoform X2
LOC105673538	-0.820408798	-0.861959975	-0.924793	-0.869053924	lcl|NW_012160723.1_cds_XP_012224679.1_12664	NW_012160723.1	protein_id=XP_012224679.1	LOW QUALITY PROTEIN: actin cytoskeleton-regulatory complex protein PAN1
LOC105678518	-0.980236237	-1.072425458	-1.59436	-1.215673898	lcl|NW_012160836.1_cds_XP_012233363.1_20283	NW_012160836.1	protein_id=XP_012233363.1	uncharacterized protein LOC105678518
LOC105673930	-0.723894382	-0.781216009	-1.12138	-0.875496797	lcl|NW_012160729.1_cds_XP_012225349.1_13195	NW_012160729.1	protein_id=XP_012225349.1	MLX-interacting protein isoform X2
LOC105668957	-1.202358756	-1.34383871	-1.39398	-1.313392489	lcl|NW_012159323.1_cds_XP_012217130.1_502	NW_012159323.1	protein_id=XP_012217130.1	probable cytochrome P450 6a14
LOC105677167	-0.614179783	-0.632730762	-0.576904	-0.607938182	lcl|NW_012160220.1_cds_XP_012231378.1_16651	NW_012160220.1	protein_id=XP_012231378.1	calcium-activated potassium channel slowpoke isoform X47

LOC105678319	-1.378461033	-1.561802805	-1.84145	-1.593904613	lcl|NW_012160223.1_cds_XP_012232972.1_1792	NW_012160223.1	protein_id=XP_012232972.1	transcription initiation factor TFIID subunit 12b-like isoform X1
LOC105679378	-0.715159104	-0.749607275	-0.784776	-0.74984746	lcl|NW_012160278.1_cds_XP_012234800.1_2086	NW_012160278.1	protein_id=XP_012234800.1	microtubule-associated serine/threonine-protein kinase 3 isoform X1
LOC105677053	-2.657798522	-3.198745218	-4.02888	-3.295141247	lcl|NW_012160220.1_cds_XP_012230816.1_1666	NW_012160220.1	protein_id=XP_012230816.1	uncharacterized protein LOC105677053
LOC105670996	-1.758674255	-2.725785898	-2.62853	-2.370996718	lcl|NW_012160677.1_cds_XP_012220218.1_8578	NW_012160677.1	protein_id=XP_012220218.1	uncharacterized protein LOC105670996
LOC105676445	-1.103165495	-1.298412833	-2.06517	-1.488916109	lcl|NW_012160777.1_cds_XP_012229762.1_1749 8	NW_012160777.1	protein_id=XP_012229762.1	B-cell lymphoma/leukemia 11A
LOC105671398	-1.362627976	-1.673990513	-2.48123	-1.83928283	lcl|NW_012160683.1_cds_XP_012220951.1_9271	NW_012160683.1	protein_id=XP_012220951.1	myb-like protein I isoform X2
LOC105676627	-1.108826846	-1.194509112	-1.00888	-1.104071986	lcl|NW_012160212.1_cds_XP_012230098.1_1543	NW_012160212.1	protein_id=XP_012230098.1	uncharacterized protein LOC105676627
LOC105680137	0.867947712	0.945219053	1.28228	1.031815588	lcl|NW_012160366.1_cds_XP_012236013.1_3193	NW_012160366.1	protein_id=XP_012236013.1	tryptophan 2,3-dioxygenase
LOC105678496	2.159534464	3.28027657	3.74918	3.062997011		NW_012160223.1		lncRNA
LOC105675138	-0.682522668	-0.721122797	-0.861007	-0.754884155	lcl|NW_012160754.1_cds_XP_012227447.1_1511 2	NW_012160754.1	protein_id=XP_012227447.1	ELAV-like protein 3 isoform X1
LOC105675477	-1.013696621	-1.287392607	-1.38736	-1.229483076	lcl|NW_012160764.1_cds_XP_012228085.1_1559 2	NW_012160764.1	protein_id=XP_012228085.1	astakine-like
LOC105677470	-1.390007345	-1.677270183	-2.03231	-1.699862509	lcl|NW_012160221.1_cds_XP_012231530.1_1688	NW_012160221.1	protein_id=XP_012231530.1	diphosphoinositol polyphosphate phosphohydrolase 1 isoform X1
LOC105668043	-1.387944019	-1.791593472	-2.06391	-1.74781583	lcl|NW_012160436.1_cds_XP_012215639.1_4393	NW_012160436.1	protein_id=XP_012215639.1	uncharacterized protein LOC105668043
LOC105676478	-2.798559942	-3.623668631	-4.07713	-3.499786191	lcl|NW_012160211.1_cds_XP_012229836.1_1513	NW_012160211.1	protein_id=XP_012229836.1	uncharacterized protein LOC105676478
LOC105678650	-1.466663839	-2.146075987	-2.03837	-1.883703275	lcl|NW_012160840.1_cds_XP_012233574.1_2115 2	NW_012160840.1	protein_id=XP_012233574.1	elongation of very long chain fatty acids protein AAEL008004-like
LOC105673603	-1.026546902	-1.154882735	-1.33952	-1.173649879	lcl|NW_012160723.1_cds_XP_012224788.1_1268 8	NW_012160723.1	protein_id=XP_012224788.1	uncharacterized protein LOC105673603
LOC105674060	-1.068056893	-1.216642603	-1.5275	-1.270733165	lcl|NW_012160731.1_cds_XP_012225555.1_1345 1	NW_012160731.1	protein_id=XP_012225555.1	uncharacterized protein LOC105674060 isoform X2
LOC105678899	0.730790036	0.77786395	0.965804	0.824819329	lcl|NW_012160840.1_cds_XP_012234029.1_2090 2	NW_012160840.1	protein_id=XP_012234029.1	40S ribosomal protein S12
LOC105667777	-0.650025899	-0.671383886	-0.739307	-0.686905595	lcl|NW_012160418.1_cds_XP_012215205.1_3895	NW_012160418.1	protein_id=XP_012215205.1	uncharacterized protein LOC105667777 isoform X1
LOC105679298	-0.914202786	-0.97060646	-1.36631	-1.083706415	lcl|NW_012160266.1_cds_XP_012234648.1_1940	NW_012160266.1	protein_id=XP_012234648.1	patched domain-containing protein 3-like
LOC105670906	1.145971815	1.51139375	1.19718	1.284848522	lcl|NW_012160673.1_cds_XP_012220076.1_8457	NW_012160673.1	protein_id=XP_012220076.1	leucine-rich repeat transmembrane neuronal protein 2-like
LOC105671418	-1.144795641	-1.235738676	-1.40729	-1.262608106	lcl|NW_012160683.1_cds_XP_012220993.1_9385	NW_012160683.1	protein_id=XP_012220993.1	LOW QUALITY PROTEIN: R3H domain-containing protein 1
LOC105672735	-1.00393014	-1.072505851	-1.19661	-1.09101533		NW_012160704.1		Pseudogene
LOC105667544	-1.647212479	-1.809701047	-1.64558	-1.700831175	lcl|NW_012158778.1_cds_XP_012214806.1_321	NW_012158778.1	protein_id=XP_012214806.1	fatty acid synthase-like
LOC105676070	-1.701804605	-1.971377684	-2.50632	-2.059834096	lcl|NW_012160776.1_cds_XP_012229085.1_1690 0	NW_012160776.1	protein_id=XP_012229085.1	uncharacterized protein LOC105676070
LOC105672009	-0.74909241	-0.832091445	-0.968422	-0.849868618	lcl|NW_012160689.1_cds_XP_012222087.1_1033 7	NW_012160689.1	protein_id=XP_012222087.1	microtubule-associated protein RP/EB family member 1-like isoform X2
LOC105673098	-0.983762282	-1.069197427	-1.06312	-1.038693236	lcl|NW_012160715.1_cds_XP_012223881.1_1215 6	NW_012160715.1	protein_id=XP_012223881.1	myosin heavy chain, muscle
LOC105677525	-1.193569001	-1.391588651	-1.74666	-1.443939217	lcl|NW_012160221.1_cds_XP_012231637.1_1669	NW_012160221.1	protein_id=XP_012231637.1	protein LSM14 homolog A
LOC105670995	0.935065534	1.051016714	1.3041	1.096727416	lcl|NW_012160677.1_cds_XP_012220216.1_8576	NW_012160677.1	protein_id=XP_012220216.1	uncharacterized protein LOC105670995 isoform X2
LOC105672916	-1.042110161	-1.118031062	-0.893339	-1.017826741	lcl|NW_012160711.1_cds_XP_012223610.1_1164 9	NW_012160711.1	protein_id=XP_012223610.1	E3 ubiquitin-protein ligase SIAH1-like

LOC105668386	0.904192241	0.982539779	1.18333	1.023354007	lcl|NW_012160474.1_cds_XP_012216171.1_4646	NW_012160474.1	protein_id=XP_012216171.1	trifunctional purine biosynthetic protein adenosine-3
LOC105673187	1.053472138	1.538851287	1.74614	1.446154475	lcl|NW_012160715.1_cds_XP_012224065.1_1210 1	NW_012160715.1	protein_id=XP_012224065.1	4-hydroxyphenylpyruvate dioxygenase
LOC105668873	-1.341117017	-1.499059688	-1.86924	-1.569805568	lcl|NW_012160516.1_cds_XP_012216941.1_5593	NW_012160516.1	protein_id=XP_012216941.1	bestrophin-4-like
LOC105671203	1.056164958	1.434416229	1.78689	1.425823729	lcl|NW_012160683.1_cds_XP_012220583.1_8952	NW_012160683.1	protein_id=XP_012220583.1	uncharacterized protein LOC105671203
LOC105674350	-1.028483398	-1.150998352	-1.62441	-1.267963917	lcl|NW_012160744.1_cds_XP_012226025.1_1430 4	NW_012160744.1	protein_id=XP_012226025.1	uncharacterized protein LOC105674350 isoform X4
LOC105677667	-1.078645553	-1.261946871	-1.38675	-1.242447475	lcl|NW_012160816.1_cds_XP_012231850.1_1915 7	NW_012160816.1	protein_id=XP_012231850.1	actin, muscle
LOC105673752	1.115711452	1.182135533	1.31293	1.203592328	lcl|NW_012160728.1_cds_XP_012225025.1_1294 9	NW_012160728.1	protein_id=XP_012225025.1	endothelin-converting enzyme 1-like isoform X2
LOC105668059	-1.004127244	-1.123004843	-1.18254	-1.103224029	lcl|NW_012160440.1_cds_XP_012215656.1_4410	NW_012160440.1	protein_id=XP_012215656.1	putative fatty acyl-CoA reductase CG5065
LOC105667316	0.808970593	0.864877618	1.26013	0.977992737	lcl|NW_012160385.1_cds_XP_012214458.1_3302	NW_012160385.1	protein_id=XP_012214458.1	uncharacterized protein LOC105667316
LOC105676346	-1.422431406	-3.689994081	-6.13951	-3.750645162	lcl|NW_012160777.1_cds_XP_012229567.1_1719 5	NW_012160777.1	protein_id=XP_012229567.1	uncharacterized protein LOC105676346 isoform X2
LOC105673581	-0.774295306	-0.832971436	-0.683602	-0.763622914	lcl|NW_012160723.1_cds_XP_012224751.1_1255 5	NW_012160723.1	protein_id=XP_012224751.1	longitudinals lacking protein, isoforms H/M/V-like isoform X2
LOC105667548	-1.596710236	-1.76182901	-1.76327	-1.707269749	lcl|NW_012160404.1_cds_XP_012214810.1_3633	NW_012160404.1	protein_id=XP_012214810.1	LOW QUALITY PROTEIN: fatty acid synthase-like
LOC105674475	-1.304521931	-1.514409918	-1.58784	-1.46892395	lcl|NW_012160188.1_cds_XP_012226229.1_1233	NW_012160188.1	protein_id=XP_012226229.1	silk gland factor 1
LOC105672698	1.304194714	1.419772894	1.65479	1.459585869	lcl|NW_012160701.1_cds_XP_012223219.1_1131 9	NW_012160701.1	protein_id=XP_012223219.1	uncharacterized protein LOC105672698
LOC105676503	-0.685989238	-0.721614408	-0.960206	-0.789269882	lcl|NW_012160777.1_cds_XP_012229879.1_1754 4	NW_012160777.1	protein_id=XP_012229879.1	arf-GAP with SH3 domain, ANK repeat and PH domain-containing protein 2 isoform X1
LOC105674360	-0.916753962	-1.021016985	-1.30678	-1.081516982	lcl|NW_012160185.1_cds_XP_012226051.1_1183	NW_012160185.1	protein_id=XP_012226051.1	high-affinity choline transporter 1-like
LOC105677670	-1.222610398	-1.298154154	-1.35421	-1.291658184	lcl|NW_012160816.1_cds_XP_012231852.1_1915 8	NW_012160816.1	protein_id=XP_012231852.1	actin-5, muscle-specific
LOC105677043	-1.02611698	-1.174260775	-1.22228	-1.140885918	lcl|NW_012160796.1_cds_XP_012230779.1_1800 9	NW_012160796.1	protein_id=XP_012230779.1	cytochrome P450 4C1-like
LOC105667536	-2.317367642	-2.522253791	-2.66062	-2.500080478		NW_012158778.1		Pseudogene
LOC105677636	-1.105942306	-1.191062408	-1.2845	-1.193834905	lcl|NW_012160816.1_cds_XP_012231808.1_1909 6	NW_012160816.1	protein_id=XP_012231808.1	FK506-binding protein 5 isoform X10
LOC105670452	-1.187245057	-1.367438897	-1.48749	-1.347391318	lcl|NW_012160660.1_cds_XP_012219404.1_7838	NW_012160660.1	protein_id=XP_012219404.1	circadian clock-controlled protein
LOC105671099	-3.040045738	-3.891408349	-4.71298	-3.881478029	lcl|NW_012160681.1_cds_XP_012220422.1_8784	NW_012160681.1	protein_id=XP_012220422.1	cytoplasmic polyadenylation element-binding protein 1-B isoform X4
LOC105667323	-1.057739625	-1.136858974	-1.13233	-1.1089762	lcl|NW_012160385.1_cds_XP_012214474.1_3356	NW_012160385.1	protein_id=XP_012214474.1	uncharacterized protein LOC105667323 isoform X1
LOC105676185	-1.071319477	-1.155446827	-1.13657	-1.121112101	lcl|NW_012158190.1_cds_XP_012229366.1_152	NW_012158190.1	protein_id=XP_012229366.1	LOW QUALITY PROTEIN: fatty acid synthase-like
LOC105674466	-1.01505663	-1.078460332	-1.17547	-1.089662321	lcl|NW_012160744.1_cds_XP_012226209.1_1384 4	NW_012160744.1	protein_id=XP_012226209.1	uncharacterized protein LOC105674466 isoform X2
LOC105677852	2.380216002	4.386104382	4.54985	3.772056795		NW_012160223.1		lncRNA
LOC105678573	-1.456799825	-2.234249197	-3.29912	-2.330056341	lcl|NW_012160836.1_cds_XP_012233461.1_2044 7	NW_012160836.1	protein_id=XP_012233461.1	uncharacterized protein LOC105678573
LOC105672983	-0.785840585	-0.887670613	-0.985246	-0.886252399	lcl|NW_012160712.1_cds_XP_012223716.1_1180 7	NW_012160712.1	protein_id=XP_012223716.1	monocyte to macrophage differentiation factor 2
LOC105678719	-0.681299847	-0.725790151	-0.861355	-0.756148333	lcl|NW_012160840.1_cds_XP_012233687.1_2105 8	NW_012160840.1	protein_id=XP_012233687.1	uncharacterized protein LOC105678719 isoform X2
LOC105675057	-2.059204125	-3.009104939	-3.75203	-2.940113021	lcl|NW_012160752.1_cds_XP_012227276.1_1501 7	NW_012160752.1	protein_id=XP_012227276.1	uncharacterized protein LOC105675057 isoform X1

LOC105670897	-0.989699177	-4.420861397	inf	-2.705280287		NW_012160673.1		lncRNA
LOC105668703	1.198098243	1.350560159	1.55917	1.369276134	lcl|NW_012160514.1_cds_XP_012216620.1_5496	NW_012160514.1	protein_id=XP_012216620.1	membrane metallo-endopeptidase-like 1
LOC105677855	-0.81045578	-0.877050644	-0.820066	-0.835857475	lcl|NW_012160818.1_cds_XP_012232156.1_1939 2	NW_012160818.1	protein_id=XP_012232156.1	putative uncharacterized protein DDB_G0286901 isoform X2
LOC105669367	-1.270381132	-1.496619089	-1.8261	-1.531033407	lcl|NW_012160586.1_cds_XP_012217695.1_6273	NW_012160586.1	protein_id=XP_012217695.1	glycogen
LOC105676170	-1.07999227	-1.279842004	-1.48584	-1.281891425	lcl|NW_012160776.1_cds_XP_012229298.1_1690 2	NW_012160776.1	protein_id=XP_012229298.1	acyl-CoA-binding protein homolog isoform X2
LOC105670866	-1.445146648	-3.778085097	inf	-2.611615873		NW_012160673.1		lncRNA
LOC105671854	-1.000495416	-1.395633208	-1.42526	-1.273796208	lcl|NW_012160079.1_cds_XP_012221781.1_820	NW_012160079.1	protein_id=XP_012221781.1	uncharacterized protein LOC105671854
LOC105677822	-0.833924653	-0.885799896	-1.07396	-0.931228183	lcl|NW_012160818.1_cds_XP_012232117.1_1941 6	NW_012160818.1	protein_id=XP_012232117.1	LOW QUALITY PROTEIN: lymphoid-restricted membrane protein-like
LOC105670858	1.181546557	1.454712553	1.78025	1.472169703	lcl|NW_012160672.1_cds_XP_012220026.1_8400	NW_012160672.1	protein_id=XP_012220026.1	leucine-rich repeat-containing protein egg-6-like
LOC105677705	-0.683307642	-0.710462689	-0.709736	-0.701168777	lcl|NW_012160816.1_cds_XP_012231927.1_1920 1	NW_012160816.1	protein_id=XP_012231927.1	guanine nucleotide-binding protein G(s) subunit alpha
LOC105667827	-0.771915485	-0.823990791	-0.942701	-0.846202425	lcl|NW_012160418.1_cds_XP_012215306.1_4128	NW_012160418.1	protein_id=XP_012215306.1	LOW QUALITY PROTEIN: filamin-A-like
LOC105675169	-0.710272115	-0.754542172	-0.707686	-0.724166762	lcl|NW_012160754.1_cds_XP_012227543.1_1531 3	NW_012160754.1	protein_id=XP_012227543.1	muscle M-line assembly protein unc-89 isoform X4
LOC105672116	-1.12814795	-1.25455047	-1.53519	-1.305962807	lcl|NW_012160693.1_cds_XP_012222251.1_1056 5	NW_012160693.1	protein_id=XP_012222251.1	angiotensin-converting enzyme-like isoform X2
LOC105669910	-0.7559647	-0.788498485	-1.04432	-0.862927728	lcl|NW_012160617.1_cds_XP_012218547.1_7075	NW_012160617.1	protein_id=XP_012218547.1	echinoderm microtubule-associated protein-like 2 isoform X5
LOC105672585	-0.892868282	-0.985845049	-1.48684	-1.12185111	lcl|NW_012160699.1_cds_XP_012223046.1_1098 2	NW_012160699.1	protein_id=XP_012223046.1	protein tincar isoform X4
LOC105677683	-0.938336168	-1.093539992	-1.01708	-1.01631872	lcl|NW_012158220.1_cds_XP_012231916.1_174	NW_012158220.1	protein_id=XP_012231916.1	uncharacterized protein LOC105677683
LOC105678601	-0.745924057	-0.799236481	-0.868638	-0.804599513	lcl|NW_012160837.1_cds_XP_012233501.1_2065 6	NW_012160837.1	protein_id=XP_012233501.1	uncharacterized protein LOC105678601
LOC105670034	-1.059489554	-1.250337604	-1.35894	-1.222922386	lcl|NW_012160620.1_cds_XP_012218749.1_7120	NW_012160620.1	protein_id=XP_012218749.1	glucosylceramidase-like
LOC105676608	0.827475064	0.911578516	1.09887	0.945974527		NW_012160777.1		lncRNA
LOC105667386	-1.794880392	-2.80117103	-2.88321	-2.493087141	lcl|NW_012160390.1_cds_XP_012214584.1_3451	NW_012160390.1	protein_id=XP_012214584.1	hexamerin-like isoform X2
LOC105677927	-2.094685668	-2.433544637	-3.15565	-2.561293435	lcl|NW_012160824.1_cds_XP_012232256.1_1949 7	NW_012160824.1	protein_id=XP_012232256.1	Niemann-Pick C1 protein-like isoform X2
LOC105671653	-0.993518497	-1.040191187	-1.31401	-1.115906561	lcl|NW_012160685.1_cds_XP_012221407.1_9676	NW_012160685.1	protein_id=XP_012221407.1	plasma membrane calcium-transporting ATPase 3 isoform X8
LOC105679066	-1.365705512	-1.545083731	-1.77573	-1.562173081	lcl|NW_012160844.1_cds_XP_012234267.1_2141 1	NW_012160844.1	protein_id=XP_012234267.1	myb-like protein Q isoform X1
LOC105669486	-1.002679188	-1.178704723	-1.20993	-1.13043797	lcl|NW_012160596.1_cds_XP_012217901.1_6465	NW_012160596.1	protein_id=XP_012217901.1	LOW QUALITY PROTEIN: cytochrome P450 4g15
LOC105677680	-0.677245269	-0.712643797	-0.770846	-0.720245022	lcl|NW_012160816.1_cds_XP_012231868.1_1924 8	NW_012160816.1	protein_id=XP_012231868.1	ankyrin repeat domain-containing protein 17 isoform X2
LOC105670612	0.800536872	0.938653437	1.12127	0.95348677	lcl|NW_012160664.1_cds_XP_012219645.1_8162	NW_012160664.1	protein_id=XP_012219645.1	40S ribosomal protein S17
LOC105678045	-0.652435062	-0.683202376	-0.797595	-0.711077479	lcl|NW_012160828.1_cds_XP_012232465.1_1970 4	NW_012160828.1	protein_id=XP_012232465.1	gamma-aminobutyric acid receptor subunit beta isoform X13
LOC105671873	-1.897643777	-4.158390654	-4.34275	-3.466261477	lcl|NW_012160688.1_cds_XP_012221807.1_1011 6	NW_012160688.1	protein_id=XP_012221807.1	uncharacterized protein LOC105671873
LOC105673689	0.956771941	1.202534245	1.35156	1.170288729	lcl|NW_012160727.1_cds_XP_012224918.1_1288 3	NW_012160727.1	protein_id=XP_012224918.1	60S acidic ribosomal protein P2
LOC105670541	-3.78994262	-4.976552348	-5.41855	-4.728348323	lcl|NW_012160663.1_cds_XP_012219518.1_7954	NW_012160663.1	protein_id=XP_012219518.1	mediator of RNA polymerase II transcription subunit 15-like isoform X2

LOC105670297	0.97047471	1.154754373	1.02925	1.051493028	lcl|NW_012160649.1_cds_XP_012219166.1_7613	NW_012160649.1	protein_id=XP_012219166.1	LOW QUALITY PROTEIN: hemocytin
LOC105676017	1.240777479	1.534995257	1.89909	1.558287579	lcl|NW_012160773.1_cds_XP_012228972.1_1652 4	NW_012160773.1	protein_id=XP_012228972.1	fatty acid synthase-like
LOC105670072	-1.403183967	-2.308490811	-4.05914	-2.590271593	lcl|NW_012160620.1_cds_XP_012218821.1_7119	NW_012160620.1	protein_id=XP_012218821.1	ribonucleoside-diphosphate reductase large subunit-like
LOC105675442	1.027730485	1.366212952	1.15216	1.182034479	lcl|NW_012160764.1_cds_XP_012228027.1_1543 5	NW_012160764.1	protein_id=XP_012228027.1	uncharacterized protein LOC105675442 isoform X2
LOC105671210	0.744044102	0.777662643	0.909051	0.810252582	lcl|NW_012160683.1_cds_XP_012220595.1_8928	NW_012160683.1	protein_id=XP_012220595.1	60S ribosomal protein L31
LOC105667809	1.121712662	1.202322207	1.40331	1.24244829	lcl|NW_012160418.1_cds_XP_012215274.1_4065	NW_012160418.1	protein_id=XP_012215274.1	protein lethal(2)essential for life-like
LOC105673090	-1.539553958	-1.648658432	-1.75036	-1.646190797	lcl|NW_012160715.1_cds_XP_012223869.1_1204 6	NW_012160715.1	protein_id=XP_012223869.1	sarcalumenin isoform X2
LOC105672643	-1.578984935	-2.217369187	-3.09025	-2.295534707	lcl|NW_012160699.1_cds_XP_012223141.1_1104 1	NW_012160699.1	protein_id=XP_012223141.1	fibulin-2-like isoform X2
LOC105672088	-1.197404898	-1.460476595	-1.75136	-1.469747164	lcl|NW_012160692.1_cds_XP_012222219.1_1041 7	NW_012160692.1	protein_id=XP_012222219.1	lipid storage droplets surface-binding protein 2-like
LOC105674935	-0.643498402	-0.66958524	-0.620379	-0.644487547	lcl|NW_012160190.1_cds_XP_012227063.1_1275	NW_012160190.1	protein_id=XP_012227063.1	calmodulin-binding transcription activator 2 isoform X2
LOC105668115	-1.457591183	-1.700803281	-1.77233	-1.643574821	lcl|NW_012160461.1_cds_XP_012215721.1_4534	NW_012160461.1	protein_id=XP_012215721.1	alpha-methylacyl-CoA racemase-like
LOC105677647	-0.877183047	-0.96508062	-1.22725	-1.023171222	lcl|NW_012160816.1_cds_XP_012231825.1_1908 5	NW_012160816.1	protein_id=XP_012231825.1	diacylglycerol O-acyltransferase 2-like
LOC105675202	1.28466368	1.68320396	2.28391	1.750592547		NW_012160754.1		lncRNA
LOC105673413	1.792932009	2.127326222	2.39493	2.105062744	lcl|NW_012160717.1_cds_XP_012224450.1_1244 3	NW_012160717.1	protein_id=XP_012224450.1	thyrotropin-releasing hormone receptor-like
LOC105674248	1.222286065	1.404777391	1.82714	1.484734485	lcl|NW_012160739.1_cds_XP_012225866.1_1374 3	NW_012160739.1	protein_id=XP_012225866.1	LOW QUALITY PROTEIN: branched-chain-amino-acid aminotransferase, cytosolic
LOC105672171	-0.784264203	-0.844168772	-0.877228	-0.835220325	lcl|NW_012160080.1_cds_XP_012222341.1_865	NW_012160080.1	protein_id=XP_012222341.1	uncharacterized protein LOC105672171
LOC105679756	1.073387745	1.197639099	1.33812	1.203048948	lcl|NW_012160333.1_cds_XP_012235416.1_2793	NW_012160333.1	protein_id=XP_012235416.1	adenosylhomocysteinase
LOC105671881	-0.700644584	-0.760107791	-0.798923	-0.753225125	lcl|NW_012160688.1_cds_XP_012221814.1_1016 7	NW_012160688.1	protein_id=XP_012221814.1	LOW QUALITY PROTEIN: microtubule-actin cross-linking factor 1
LOC105670177	-1.402595934	-1.614002067	-2.21398	-1.743526	lcl|NW_012160639.1_cds_XP_012218975.1_7468	NW_012160639.1	protein_id=XP_012218975.1	LOW QUALITY PROTEIN: protein bicaudal C
LOC105668207	1.973902409	2.820918209	3.05616	2.616993539	lcl|NW_012160474.1_cds_XP_012215860.1_5038	NW_012160474.1	protein_id=XP_012215860.1	chymotrypsin-1-like
LOC105674950	-0.760787661	-0.810417049	-0.83322	-0.801474903	lcl|NW_012160750.1_cds_XP_012227071.1_1484 1	NW_012160750.1	protein_id=XP_012227071.1	chromodomain-helicase-DNA-binding protein 7 isoform X6
LOC105678491	0.819322046	0.963912978	1.30752	1.030251675		NW_012160223.1		lncRNA
LOC105675233	-0.693190777	-0.722712423	-0.6895	-0.701801067	lcl|NW_012160763.1_cds_XP_012227651.1_1536 9	NW_012160763.1	protein_id=XP_012227651.1	LOW QUALITY PROTEIN: nesprin-1-like
LOC105671534	-0.770141104	-0.846945364	-1.19142	-0.936168823	lcl|NW_012160684.1_cds_XP_012221197.1_9559	NW_012160684.1	protein_id=XP_012221197.1	moesin/ezrin/radixin homolog 1 isoform X1
LOC105670011	-3.389365751	-4.261957603	-4.98859	-4.213304451	lcl|NW_012160620.1_cds_XP_012218709.1_7197	NW_012160620.1	protein_id=XP_012218709.1	vitellogenin receptor-like
LOC105676988	-1.19818156	-1.519864021	-2.64747	-1.788505194	lcl|NW_012160796.1_cds_XP_012230665.1_1821 7	NW_012160796.1	protein_id=XP_012230665.1	zinc finger protein 764
LOC105678787	1.280619024	1.560488683	1.70187	1.514325902	lcl|NW_012158257.1_cds_XP_012233894.1_188	NW_012158257.1	protein_id=XP_012233894.1	fatty acid synthase-like
LOC105669318	1.231021741	1.590851258	2.32022	1.714031	lcl|NW_012160578.1_cds_XP_012217627.1_6189	NW_012160578.1	protein_id=XP_012217627.1	kynurenine/alpha-aminoadipate aminotransferase, mitochondrial-like
LOC105676600	-3.103687939	-3.998208109	-4.04679	-3.716228683	lcl|NW_012160777.1_cds_XP_012230053.1_1727 1	NW_012160777.1	protein_id=XP_012230053.1	arylphorin subunit alpha-like
LOC105670722	-0.716861718	-0.762384582	-0.773913	-0.7510531	lcl|NW_012160668.1_cds_XP_012219829.1_8246	NW_012160668.1	protein_id=XP_012219829.1	F-box only protein 11

LOC105670444	-1.688247847	-2.117096496	-2.39646	-2.067268114	lcl|NW_012160660.1_cds_XP_012219388.1_7835	NW_012160660.1	protein_id=XP_012219388.1	uncharacterized protein LOC105670444
LOC105675123	-1.253729106	-1.37903141	-1.42754	-1.353433505	lcl|NW_012160193.1_cds_XP_012227415.1_1340	NW_012160193.1	protein_id=XP_012227415.1	calcium-transporting ATPase sarcoplasmic/endoplasmic reticulum type isoform X1
LOC105667374	-1.588727056	-1.808299238	-2.62407	-2.007032098	lcl|NW_012160390.1_cds_XP_012214561.1_3442	NW_012160390.1	protein_id=XP_012214561.1	cationic amino acid transporter 3 isoform X3
LOC105670286	-0.644943698	-0.673323248	-0.690644	-0.669636982	lcl|NW_012160649.1_cds_XP_012219135.1_7601	NW_012160649.1	protein_id=XP_012219135.1	trithorax group protein osa isoform X5
LOC105679486	1.028741884	1.280240187	1.67556	1.32818069	lcl|NW_012160307.1_cds_XP_012234970.1_2236	NW_012160307.1	protein_id=XP_012234970.1	acyl-CoA Delta(11) desaturase-like
LOC105677931	-1.911787117	-2.653071025	-3.11127	-2.558709381	lcl|NW_012160824.1_cds_XP_012232265.1_1951 8	NW_012160824.1	protein_id=XP_012232265.1	ankyrin repeat and SAM domain-containing protein 3-like isoform X1
LOC105677026	0.752124259	0.836216678	1.17903	0.922456979	lcl|NW_012160796.1_cds_XP_012230747.1_1821 2	NW_012160796.1	protein_id=XP_012230747.1	spondin-1-like isoform X4
LOC105673069	-0.726368249	-0.759115353	-0.606128	-0.697203867	lcl|NW_012160178.1_cds_XP_012223846.1_1005	NW_012160178.1	protein_id=XP_012223846.1	uncharacterized protein LOC105673069 isoform X2
LOC105672664	-0.942100782	-1.04206777	-1.02062	-1.001596184	lcl|NW_012160700.1_cds_XP_012223169.1_1130 5	NW_012160700.1	protein_id=XP_012223169.1	UDP-glucuronosyltransferase 2C1-like
LOC105674873	-1.018336311	-1.130289528	-1.49362	-1.214081946	lcl|NW_012160748.1_cds_XP_012226925.1_1479 8	NW_012160748.1	protein_id=XP_012226925.1	glucose dehydrogenase
LOC105668189	-1.473825015	-2.52968349	-2.991	-2.331502835	lcl|NW_012160472.1_cds_XP_012215844.1_4579	NW_012160472.1	protein_id=XP_012215844.1	uncharacterized protein LOC105668189
LOC105667359	-1.059708905	-1.204309797	-1.54133	-1.268449567	lcl|NW_012160390.1_cds_XP_012214530.1_3422	NW_012160390.1	protein_id=XP_012214530.1	rho guanine nucleotide exchange factor 10-like protein
LOC105678421	-1.153510922	-1.313042653	-2.39867	-1.621741192	lcl|NW_012160836.1_cds_XP_012233179.1_2058 9	NW_012160836.1	protein_id=XP_012233179.1	uncharacterized protein LOC105678421 isoform X2
LOC105668361	-1.217212306	-1.325328315	-1.62608	-1.389540207	lcl|NW_012160474.1_cds_XP_012216127.1_5027	NW_012160474.1	protein_id=XP_012216127.1	sushi, von Willebrand factor type A, EGF and pentraxin domain-containing protein 1-like
LOC105675938	-0.749707646	-0.804246739	-1.21711	-0.923688128	lcl|NW_012160771.1_cds_XP_012228868.1_1645 6	NW_012160771.1	protein_id=XP_012228868.1	protein sickie-like
LOC105670518	-0.864513558	-0.939719615	-1.61646	-1.140231058	lcl|NW_012160663.1_cds_XP_012219487.1_8032	NW_012160663.1	protein_id=XP_012219487.1	LOW QUALITY PROTEIN: sodium/calcium exchanger 2
LOC105673571	1.566748699	1.672013914	1.95065	1.729804204	lcl|NW_012160723.1_cds_XP_012224725.1_1278 7	NW_012160723.1	protein_id=XP_012224725.1	phosphoserine phosphatase isoform X4
LOC105676633	1.32055027	1.496135899	1.79245	1.536378723	lcl|NW_012160777.1_cds_XP_012230101.1_1751 7	NW_012160777.1	protein_id=XP_012230101.1	MAM and LDL-receptor class A domain-containing protein 2-like
LOC105676765	-1.166980884	-1.491066898	-1.99291	-1.550319261	lcl|NW_012160785.1_cds_XP_012230332.1_1770 9	NW_012160785.1	protein_id=XP_012230332.1	serine/threonine-protein kinase polo
LOC105668448	1.087314202	1.568554715	1.36768	1.341182972	lcl|NW_012160474.1_cds_XP_012216263.1_4881	NW_012160474.1	protein_id=XP_012216263.1	homogentisate 1,2-dioxygenase
LOC105667677	-1.182061439	-1.561763149	-1.4534	-1.399074863	lcl|NW_012160413.1_cds_XP_012215052.1_3856	NW_012160413.1	protein_id=XP_012215052.1	zinc carboxypeptidase-like
LOC105675105	1.926661657	2.297195052	2.6078	2.277218903	lcl|NW_012160193.1_cds_XP_012227353.1_1299	NW_012160193.1	protein_id=XP_012227353.1	LOW QUALITY PROTEIN: fatty acid synthase-like
LOC105672596	-1.249269459	-1.552189434	-1.88604	-1.562499631	lcl|NW_012160699.1_cds_XP_012223059.1_1094 3	NW_012160699.1	protein_id=XP_012223059.1	exonuclease 3'-5' domain-containing protein 2
LOC105671774	-0.637239561	-0.665868578	-0.798356	-0.700488046	lcl|NW_012160686.1_cds_XP_012221642.1_9902	NW_012160686.1	protein_id=XP_012221642.1	LOW QUALITY PROTEIN: basement membrane-specific heparan sulfate proteoglycan core protein
LOC105677022	-1.10023115	-1.244228688	-1.54101	-1.295156613	lcl|NW_012160796.1_cds_XP_012230736.1_1811 4	NW_012160796.1	protein_id=XP_012230736.1	trithorax group protein osa-like isoform X1
LOC105668221	-0.710992237	-0.731253912	-0.569137	-0.67046105	lcl|NW_012160474.1_cds_XP_012215884.1_4620	NW_012160474.1	protein_id=XP_012215884.1	poly(rC)-binding protein 4 isoform X4
LOC105668635	-1.008054084	-1.268201689	-2.2472	-1.507818591	lcl|NW_012160509.1_cds_XP_012216516.1_5187	NW_012160509.1	protein_id=XP_012216516.1	LOW QUALITY PROTEIN: putative epidermal cell surface receptor
LOC105675259	-0.837550561	-0.904615714	-1.00666	-0.916275425	lcl|NW_012160763.1_cds_XP_012227677.1_1538 3	NW_012160763.1	protein_id=XP_012227677.1	probable fatty acid-binding protein isoform X1
LOC105680170	-1.383035407	-1.888052572	-1.82708	-1.699389326	lcl|NW_012158698.1_cds_XP_012236064.1_288	NW_012158698.1	protein_id=XP_012236064.1	lysosomal aspartic protease
LOC105675679	-2.281530826	-4.608238237	-6.29278	-4.394183021	lcl|NW_012160768.1_cds_XP_012228395.1_1617 5	NW_012160768.1	protein_id=XP_012228395.1	aminopeptidase N-like isoform X2

LOC105668206	-1.190625281	-1.845343765	-1.92972	-1.655229682	lcl|NW_012160474.1_cds_XP_012215858.1_5036	NW_012160474.1	protein_id=XP_012215858.1	transmembrane protease serine 9-like
LOC105673741	1.070074598	1.174040283	1.43563	1.226581627	lcl|NW_012160728.1_cds_XP_012225005.1_1309 2	NW_012160728.1	protein_id=XP_012225005.1	uncharacterized protein LOC105673741 isoform X1
LOC105671506	1.674118336	1.840430852	1.44225	1.652266396	lcl|NW_012160683.1_cds_XP_012221160.1_8953	NW_012160683.1	protein_id=XP_012221160.1	uncharacterized protein LOC105671506
LOC105675201	1.02633169	1.169699856	1.54762	1.247883849	lcl|NW_012160754.1_cds_XP_012227604.1_1518 9	NW_012160754.1	protein_id=XP_012227604.1	glycine N-methyltransferase isoform X2
LOC105674720	1.459537878	1.642749887	1.4422	1.514829255		NW_012160746.1		Pseudogene
LOC105673449	-0.801493864	-0.868270264	-1.03696	-0.902241376	lcl|NW_012160180.1_cds_XP_012224509.1_1052	NW_012160180.1	protein_id=XP_012224509.1	E3 ubiquitin-protein ligase Bre1 isoform X1
LOC105670403	0.963868652	1.04225526	1.25906	1.088394637	lcl|NW_012160651.1_cds_XP_012219341.1_7685	NW_012160651.1	protein_id=XP_012219341.1	S-adenosylmethionine synthase isoform X1
LOC105673193	-0.722156362	-0.782499826	-0.907172	-0.803942729	lcl|NW_012160715.1_cds_XP_012224084.1_1199 9	NW_012160715.1	protein_id=XP_012224084.1	transmembrane protein 131
LOC105673269	-0.79152446	-0.851898542	-1.2352	-0.959541001	lcl|NW_012160179.1_cds_XP_012224224.1_1037	NW_012160179.1	protein_id=XP_012224224.1	uncharacterized protein LOC105673269
LOC105677240	-1.352535778	-1.603429126	-1.61831	-1.524758301	lcl|NW_012160802.1_cds_XP_012231124.1_1849 1	NW_012160802.1	protein_id=XP_012231124.1	general odorant-binding protein 56d-like

REFERENCES

Aguilar, Ruth et al. “Genome-Wide Analysis of Transcriptomic Divergence between Laboratory Colony and Field Anopheles Gambiae Mosquitoes of the M and S Molecular Forms.” *Insect Molecular Biology* 19.5 (2011): 695–705. Web.

Alvarado, Sebastian et al. “Epigenetic Variation in the Egfr Gene Generates Quantitative Variation in a Complex Trait in Ants.” *Nature Communications* 6 (2015): n. pag. Web.

Amdam, Gro V., Zilá L P Simões, et al. “Hormonal Control of the Yolk Precursor Vitellogenin Regulates Immune Function and Longevity in Honeybees.” *Experimental Gerontology* 39.5 (2004): 767–773. Web. 24 Jan. 2018.

Amdam, Gro V., K. Norberg, et al. “Reproductive Ground Plan May Mediate Colony-Level Selection Effects on Individual Foraging Behavior in Honey Bees.” *Proceedings of the National Academy of Sciences* 101.31 (2004): 11350–11355. Web.

Anderson, Kirk E, Timothy A Linksvayer, and Chris R Smith. “The Causes and Consequences of Genetic Caste Determination in Ants (Hymenoptera : Formicidae).” *Ecology* 11.August (2008): 119–132. Web. 13 Mar. 2018.

Aubin-Horth, Nadia, and Susan C.P. Renn. “Genomic Reaction Norms: Using Integrative Biology to Understand Molecular Mechanisms of Phenotypic Plasticity.” *Molecular Ecology* 18.18 (2009): 3763–3780. Web.

Baker, T C, S E Van Vorhis Key, and Lyle K. Gaston. “Bait-Preference Tests for the Argentine Ant (Hymenoptera: Formicidae).” *Journal of Economic Entomology* 78.5 (1985): 1083–1088. Web. 23 Jan. 2018.

Ben-Shahar, Y. et al. “Influence of Gene Action across Different Time Scales on Behavior.” *Science* 296.5568 (2002): 741–744. Web.

Berens, Ali J., James H. Hunt, and Amy L. Toth. “Comparative Transcriptomics of Convergent Evolution: Different Genes but Conserved Pathways Underlie Caste Phenotypes across Lineages of Eusocial Insects.” *Molecular Biology and Evolution* 32.3 (2015): 690–703. Web.

Bhatkar, A., and W. H. Whitcomb. “Artificial Diet for Rearing Various Species of Ants.” *The Florida Entomologist* 53.4 (1970): 229. Web.

Bloch, Guy, Dan P. Toma, and Gene E. Robinson. “Behavioral Rhythmicity, Age, Division of Labor and Period Expression in the Honey Bee Brain.” *Journal of Biological Rhythms* 16.5 (2001): 444–456. Web. 5 Mar. 2018.

Bonasio, Roberto, Qiye Li, et al. “Genome-Wide and Caste-Specific DNA Methylomes of the Ants Camponotus Floridanus and Harpegnathos Saltator.” *Current Biology* 22.19 (2012): 1755–1764. Web.

Bonasio, Roberto, Guojie Zhang, et al. “Genomic Comparison of the Ants Camponotus Floridanus and Harpegnathos Saltator.” *Science* 329.5995 (2010): 1068–1071. Web.

Borchert, Glen M et al. “Adenosine Deamination in Human Transcripts Generates Novel MicroRNA Binding Sites.” *Human Molecular Genetics* 18.24 (2009): 4801–4807. Web. 2 Mar. 2018.

Bourke, Andrew F. G. “Worker Reproduction in the Higher Eusocial Hymenoptera.” *The Quarterly Review of Biology* 63.3 (1988): 291–311. Web. 30 Jan. 2018.

Buczkowski, Grzegorz, Edward L. Vargo, and Jules Silverman. “The Diminutive Supercolony: The Argentine Ants of the Southeastern United States.” *Molecular Ecology* 13.8 (2004): 2235–2242. Web.

Burmester, Thorsten. “Evolution and Function of the Insect Hexamerins.” *European Journal of Entomology* 96.3 (1999): 213–225. Web.

Chen, J J, and C H Chen. “Microarray Gene Expression.” *Encyclopedia of Biopharmaceutical ...* (2003): 599–613. Web.

Cho, Wen Long, Tarlochan S. Dhadialla, and Alexander S. Raikhel. “Purification and Characterization of a Lysosomal Aspartic Protease with Cathepsin D Activity from the Mosquito.” *Insect Biochemistry* 21.2 (1991): 165–176. Web.

Cornette, R., S. Koshikawa, and T. Miura. “Histology of the Hormone-Producing Glands in the Damp-Wood Termite Hodotermopsis Sjostedti (Isoptera, Termopsidae): A Focus on Soldier Differentiation.” *Insectes Sociaux* 55.4 (2008): 407–416. Web. 16 Feb. 2018.

Corona, Miguel et al. “Vitellogenin Underwent Subfunctionalization to Acquire Caste and Behavioral Specific Expression in the Harvester Ant Pogonomyrmex Barbatus.” *PLoS Genetics* 9.8 (2013): n. pag. Web.

Daane, Kent M. et al. “Impacts of Argentine Ants on Mealybugs and Their Natural Enemies in California’s Coastal Vineyards.” *Ecological Entomology* 32.6 (2007): 583–596. Web. 11 Nov. 2015.

de Azevedo, Sergio Vicente, and Klaus Hartfelder. “The Insulin Signaling Pathway in Honey Bee (Apis Mellifera) Caste Development - Differential Expression of Insulin-like Peptides and Insulin Receptors in Queen and Worker Larvae.” *Journal of Insect Physiology* 54.6 (2008): 1064–1071. Web. 18 Feb. 2018.

Dittmer, Neal T., and Alexander S. Raikhel. “Analysis of the Mosquito Lysosomal Aspartic Protease Gene: An Insect Housekeeping Gene with Fat Body-Enhanced Expression.” *Insect Biochemistry and Molecular Biology* 27.4 (1997): 323–335. Web.

Drapeau, Mark David et al. “Evolution of the Yellow/Major Royal Jelly Protein Family and the Emergence of Social Behavior in Honey Bees.” *Genome Research* 16.11 (2006): 1385–1394. Web. 9 Apr. 2018.

El Hamalawi, Z A, and J A Menge. “The Role of Snails and Ants in Transmitting the Avocago Stem Canker Pathogen, Phytophthora Citricola.” *J. Am. Soc. Hort. Sci.* N.p., 1996. Web. 30 Nov. 2015.

Engels, Wolf. “Occurrence and Significance of Vitellogenins in Female Castes of Social Hymenoptera.” *Integrative and Comparative Biology* 14.4 (1974): 1229–1237. Web.

Evans, Jay D et al. “Differential Gene Expression between Developing Queens and Workers in the Honey Bee, Apis Mellifera.” *Proceedings of the National Academy of Sciences of the United States of America* 96.10 (1999): 5575–80. Web. 19 Feb. 2018.

Evans, Jay D, and Diana E Wheeler. “Gene Expression and the Evolution of Insect Polyphenisms.” *BioEssays* 23.1 (2001): 62–68. Web.

Feldmeyer, B., D. Elsner, and S. Foitzik. “Gene Expression Patterns Associated with Caste and Reproductive Status in Ants: Worker-Specific Genes Are More Derived than Queen-Specific Ones.” *Molecular Ecology* 23.1 (2014): 151–161. Web. 19 Feb. 2016.

Ferreira, Pedro G. et al. “Transcriptome Analyses of Primitively Eusocial Wasps Reveal Novel Insights into the Evolution of Sociality and the Origin of Alternative Phenotypes.” *Genome Biology* 14.2 (2013): R20. Web.

Field, Herb C et al. “A Survey of Structural Ant Pests in the Southwestern U.S.A. (Hymenoptera: Formicidae).” *Sociobiology* 2007: 151–164. Web. 23 Jan. 2018.

Friedman, D.A., and D.M. Gordon. “Ant Genetics: Reproductive Physiology, Worker Morphology, and Behavior.” *Annual Review of Neuroscience* 39.1 (2016): 41–56. Web.

Gamboa, George T., and Micheal D. Breed. “Behavioral Control of Workers by Queens in Primitively Eusocial Bees.” *Science* 195.4279 (1977): 694–696. Print.

Guidugli, Karina R. et al. “Vitellogenin Regulates Hormonal Dynamics in the Worker Caste of a Eusocial Insect.” *FEBS Letters* 579.22 (2005): 4961–4965. Web. 24 Jan. 2018.

Hagedorn, H H, and J G Kunkel. “Vitellogenin and Vitellin in Insects.” *Annual Review of Entomology* 24.1 (1979): 475–505. Web. 23 Jan. 2018.

Harrison, Mark C., Robert L. Hammond, and Eamonn B. Mallon. “Reproductive Workers Show Queenlike Gene Expression in an Intermediately Eusocial Insect, the Buff-Tailed Bumble Bee Bombus Terrestris.” *Molecular Ecology* 24.12 (2015): 3043–3063. Web.

Herb, Brian R et al. “Reversible Switching between Epigenetic States in Honeybee Behavioral Subcastes.” *Nature Neuroscience* 15.10 (2012): 1371–1373. Web. 26 Feb. 2018.

Hoffman, Eric A., and Michael A D Goodisman. “Gene Expression and the Evolution of Phenotypic Diversity in Social Wasps.” *BMC Biology* 5 (2007): n. pag. Web.

Hoffmann, Ary A, and Perran A Ross. “Rates and Patterns of Laboratory Adaptation in (Mostly) Insects.” *Journal of Economic Entomology* March (2018): 1–9. Web. 14 Mar. 2018.

Hoffmann, Ary A, and Perran A Ross. “Rates and Patterns of Laboratory Adaptation in (Mostly) Insects.” *Journal of Economic Entomology* March (2018): 1–9. Web. 14 Mar. 2018.

Højland, Dorte H., Karl Martin Vagn Jensen, and Michael Kristensen. “Adaptation of Musca Domestica L. Field Population to Laboratory Breeding Causes Transcriptional Alterations.” *PLoS ONE* 9.1 (2014): n. pag. Web.

Højland, Dorte H., Karl Martin Vagn Jensen, and Michael Kristensen. “A Comparative Study of P450 Gene Expression in Field and Laboratory Musca Domestica L. Strains.” *Pest Management Science* 70.8 (2014): 1237–1242. Web.

Holway, David a. “Effect of Argentine Ant Invasion on Ground Dwelling Arthropods in Northern California Riparian Woodlands.” N.p., 1998. Web. 29 Nov. 2015.

Holway, David A. “Competitive Mechanisms Underlying the Displacement of Native Ants By the Invasive Argentine Ant.” *Ecology*. N.p., 1999. Web. 8 Nov. 2015.

Howe, Jack, Morten Schiøtt, and Jacobus J. Boomsma. “Tachykinin Expression Levels Correlate with Caste-Specific Aggression in Workers of the Leaf-Cutting Ant Acromyrmex Echinatior.” *Frontiers in Ecology and Evolution* 4 (2016): 55. Web. 31 Jan. 2018.

Hu, Yi, Piotr Łukasik, Corrie S. Moreau, and Jacob A. Russell. 2014. “Correlates of Gut Community Composition across an Ant Species (Cephalotes Varians) Elucidate Causes and Consequences of Symbiotic Variability.” *Molecular Ecology* 23(6): 1284–1300.

Huarong Lin, Winston ML, Haunerland NH. “Influence of Age and Population Size on Ovarian Development , and of Trophallaxis on Ovarian Development and Vitellogenin Titres of Queenless Worker Honey Bee (Hymenoptera: Apidae).” *The Canadian Entomologist* 131 (1999): 695–706. Print.

Human, Kathleen G., and Deborah M. Gordon. “Exploitation and Interference Competition between the Invasive Argentine Ant, Linepithema Humile, and Native Ant Species.” *Oecologia*. N.p., 1996. Web. 8 Nov. 2015.

Ingram, Krista K., Scott Krummey, and Michelle LeRoux. “Expression Patterns of a Circadian Clock Gene Are Associated with Age-Related Polyethism in Harvester Ants, Pogonomyrmex Occidentalis.” *BMC Ecology* 9 (2009): 1–9. Web.

Ingram, Krista K., Peter Oefner, and Deborah M. Gordon. “Task-Specific Expression of the Foraging Gene in Harvester Ants.” *Molecular Ecology* 14.3 (2005): 813–818. Web.

Ingram, Krista K, Lindsay Kleeman, and Swetha Peteru. “Differential Regulation of the Foraging Gene Associated with Task Behaviors in Harvester Ants.” *BMC Ecology* 11.1 (2011): 19. Web.

Jackson, R. J., and N. Standart. “How Do MicroRNAs Regulate Gene Expression?” *Science’s STKE* 2007.367 (2007): re1-re1. Web. 2 Mar. 2018.

Jandt, Jennifer M. et al. “Lab Rearing Environment Perturbs Social Traits: A Case Study with Polistes Wasps.” *Behavioral Ecology* 26.5 (2015): 1274–1284. Web. 13 Mar. 2018.

Jongeneel, Victor et al. “RNA-Seq: A Revolutionary Tool for Transcriptomics.” *Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins* 10.1 (2001): 57–63. Web.

Judd, Timothy M, and Matthew P Fasnacht. “Distribution of Micronutrients in Social Insects: A Test in the Termite Reticulitermes Flavipes (Isoptera: Rhinotermitidae) and the Ant Myrmica Punctiventris (Hymenoptera: Formicidae).” *Ann. Entomol. Soc. Am.* 100.6 (2007): 893–899. Web. 31 Jan. 2018.

Katoh, Yohei et al. “The Clavesin Family, Neuron-Specific Lipid-and Clathrin-Binding Sec14 Proteins Regulating Lysosomal Morphology.” *Journal of Biological Chemistry* 284.40 (2009): 27646–27654. Web.

Kawahara, Yukio et al. “Redirection of Silencing Targets by Adenosine-to-Inosine Editing of MiRNAs.” *Science* 315.5815 (2007): 1137–1140. Web. 2 Mar. 2018.

Keller, Laurent, Passera, Luc, and Suzzoni, Jean-Pierre. 1989. “Queen Execution in the Argentine Ant, Iridomyrmex Humilis.” *Physiological Entomology* 14(2): 157–63.

Kent, Clement F. et al. “The Drosophila Foraging Gene Mediates Adult Plasticity and Gene-Environment Interactions in Behaviour, Metabolites, and Gene Expression in Response to Food Deprivation.” Ed. Greg Gibson. *PLoS Genetics* 5.8 (2009): e1000609. Web. 26 Jan. 2018.

Koštál, Vladimír. “Insect Photoperiodic Calendar and Circadian Clock: Independence, Cooperation, or Unity?” *Journal of Insect Physiology* 57.5 (2011): 538–556. Web.

Krushelnycky, Paul D., Stephanie M. Joe, Arthur C. Medeiros, and Curtis C. Daehler. 2005. “The Role of Abiotic Conditions in Shaping the Long-Term Patterns of a High-Elevation Argentine Ant Invasion.” *Diversity and Distributions* 11(4): 319–31.

Laakkonen, Juha, Robert N. Fisher, and Ted J Case. “Effect of Land Cover, Habitat Fragmentation and Ant Colonies on the Distribution and Abundance of Shrews in Southern California.” *Journal of Animal Ecology* 70.5 (2001): 776–788. Web. 23 Jan. 2018.

Laridon, Bram, Patrick Callaerts, and Koen Norga. “Embryonic Expression Patterns of Drosophila ACS Family Genes Related to the Human Sialin Gene.” *Gene Expression Patterns* 8.4 (2008): 275–283. Web.

Laurencikiene, Jurga et al. “RNA Editing and Alternative Splicing: The Importance of Co-Transcriptional Coordination.” *EMBO Reports* 7.3 (2006): 303–307. Web. 2 Mar. 2018.

Li-Byarlay, Hongmei et al. “RNA Interference Knockdown of DNA Methyl-Transferase 3 Affects Gene Alternative Splicing in the Honey Bee.” *Proceedings of the National Academy of Sciences* 110.31 (2013): 12750–12755. Web. 26 Feb. 2018.

Li, Qiye et al. “Caste-Specific RNA Editomes in the Leaf-Cutting Ant Acromyrmex Echinatior.” *Nature communications* 5.11 (2014): 4943. Web.

Liang, D., and Jules Silverman. 2000. “‘You Are What You Eat’: Diet Modifies Cuticular Hydrocarbons and Nestmate Recognition in the Argentine Ant, Linepithema Humile.” *Naturwissenschaften* 87(9): 412–16.

Linksvayer, Timothy A. et al. “Developmental Evolution in Social Insects: Regulatory Networks from Genes to Societies.” *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution* 318.3 (2012): 159–169. Web.

Lucas, C., M. Nicolas, and L. Keller. “Expression of Foraging and Gp-9 Are Associated with Social Organization in the Fire Ant Solenopsis Invicta.” *Insect Molecular Biology* 24.1 (2015): 93–104. Web.

Lucas, Christophe, and Marla B Sokolowski. “Molecular Basis for Changes in Behavioral State in Ant Social Behaviors.” *Proceedings of the National Academy of Sciences of the United States of America* 106.15 (2009): 6351635–6. Web.

Lucas, Eric R., Jonathan Romiguier, and Laurent Keller. “Gene Expression Is More Strongly Influenced by Age than Caste in the Ant Lasius Niger.” *Molecular Ecology* 26.19 (2017): 5058–5073. Web.

Lyko, Frank et al. “The Honey Bee Epigenomes: Differential Methylation of Brain DNA in Queens and Workers.” *PLoS Biology* 8.11 (2010): e1000506. Web. 26 Feb. 2018.

Martinez, Teresa, and Diana E Wheeler. “Identification of Two Storage Hexamers in the Ant, Camponotus Festinatus: Accumulation in Adult Queenless Workers.” *Insect Biochemistry and Molecular Biology* 23.2 (1993): 309–317. Web.

Mckenzie, Sean K, Peter R Oxley, and Daniel J C Kronauer. “Comparative Genomics and Transcriptomics in Ants Provide New Insights into the Evolution and Funcion of OBP and CSP BMC Genomics 2014.Pdf.” (2014): n. pag. Print.

McKenzie, Sean K, Peter R Oxley, and Daniel J C Kronauer. “Comparative Genomics and Transcriptomics in Ants Provide New Insights into the Evolution and Function of Odorant Binding and Chemosensory Proteins.” *BMC genomics* 15.1 (2014): 718. Web.

Mikheyev, Alexander S., and Timothy A. Linksvayer. “Genes Associated with Ant Social Behavior Show Distinct Transcriptional and Evolutionary Patterns.” *eLife* 2015.4 (2015): 1–17. Web.

Miyaji, Takaaki, Hiroshi Omote, and Yoshinori Moriyama. “Functional Characterization of Vesicular Excitatory Amino Acid Transport by Human Sialin.” *Journal of Neurochemistry* 119.1 (2011): 1–5. Web.

Morandin, C. et al. “Caste-Biases in Gene Expression Are Specific to Developmental Stage in the Ant Formica Exsecta.” *Journal of Evolutionary Biology* 28.9 (2015): 1705–1718. Web.

Morandin, Claire, Mandy M.Y. Tin, et al. “Erratum: Comparative Transcriptomics Reveals the Conserved Building Blocks Involved in Parallel Evolution of Diverse Phenotypic Traits in Ants [Genome Biol. 17, (2016), (43)].” *Genome Biology* 17.1 (2016): 1–19. Web.

Morandin, Claire, Heli Havukainen, et al. “Not Only for Egg Yolk-Functional and Evolutionary Insights from Expression, Selection, and Structural Analyses of Formica Ant Vitellogenins.” *Molecular Biology and Evolution* 31.8 (2014): 2181–2193. Web.

Mueller, Odilo, and Andreas Schroeder. “RNA Integrity Number (RIN) – Standardization of RNA Quality Control Application.” *Nano* (2004): 1–8. Web. 15 May 2018.

Mueller, U. G. “Haplodiploidy and the Evolution of Facultative Sex Ratios in a Primitively Eusocial Bee.” *Science* 254.5030 (1991): 442–444. Web. 10 Mar. 2018.

Mutti, N. S et al. “IRS and TOR Nutrient-Signaling Pathways Act via Juvenile Hormone to Influence Honey Bee Caste Fate.” *Journal of Experimental Biology* 214.23 (2011): 3977–3984. Web. 14 Feb. 2018.

Nelson, C Mindy et al. “The Gene Vitellogenin Has Multiple Coordinating Effects on Social Organization.” *PLoS Biology* 5.3 (2007): 0673–0677. Web. 24 Jan. 2018.

Newell, Wilmon. “Notes on the Habits of the Argentine or ‘New Orleans’ Ant, Iridomyrmex Humilis Mayr.” *Journal of Economic Entomology* 1.1 (1908): 21–34. Web. 23 Jan. 2018.

Ometto, Lino et al. “Evolution of Gene Expression in Fire Ants: The Effects of Developmental Stage, Caste, and Species.” *Molecular Biology and Evolution* 28.4 (2011): 1381–1392. Web. 19 Feb. 2016.

Oxley, Peter R. et al. “The Genome of the Clonal Raider Ant Cerapachys Biroi.” *Current Biology* 24.4 (2014): 451–458. Web.

Passera, L., L. Keller, and J. P. Suzzoni. 1988. “Control of Brood Male Production in the Argentine Ant Iridomyrmex Humilis (Mayr).” *Insectes Sociaux* 35(1): 19–33.

Patel, Avani et al. “The Making of a Queen: TOR Pathway Is a Key Player in Diphenic Caste Development.” *PLoS ONE* 2.6 (2007): n. pag. Web. 19 Feb. 2018.

Peeters, Christian, and Ross H. Crozier. “Caste and Reproduction in Ants: Not All Mated Egg-Layers Are ‘Queens.’” *Psyche (New York)* 95.3–4 (1988): 283–288. Web. 25 Feb. 2018.

Peeters, Christian, and B. Hölldobler. “Reproductive Cooperation between Queens and Their Mated Workers: The Complex Life History of an Ant with a Valuable Nest.” *Proceedings of the National Academy of Sciences of the United States of America* 92.24 (1995): 10977–10979. Web. 30 Jan. 2018.

Pereboom, Jeffrey J M et al. “Differential Gene Expression in Queen-Worker Caste Determination in Bumble-Bees.” *Proceedings of the Royal Society B: Biological Sciences* 272.1568 (2005): 1145–1152. Web. 19 Feb. 2018.

Phillips, P, and C Sherk. “To Control Mealybugs, Stop Honeydew-Seeking Ants.” *California Agriculture* 45.2 (1991): 26–28. Web. 11 Nov. 2015.

Rachinsky, Anna et al. “Caste and Metamorphosis: Hemolymph Titers of Juvenile Hormone and Ecdysteroids in Last Instar Honeybee Larvae.” *General and Comparative Endocrinology* 79.1 (1990): 31–38. Web. 15 Feb. 2018.

Raikhel, Alexander S. “Role of Lysosomes in Regulating of Vitellogenin Secretion in the Mosquito Fat Body.” *Journal of Insect Physiology* 32.7 (1986): 597–604. Web.

Rajakumar, Rajendhran et al. “Ancestral Developmental Potential Facilitates Parallel Evolution in Ants.” *Science* 335.6064 (2012): 79–82. Web.

Ratnieks, Francis L. W. “Reproductive Harmony via Mutual Policing by Workers in Eusocial Hymenoptera.” *The American Naturalist* 132.2 (1988): 217–236. Web.

Ravary, F., B. Jahyny, and P. Jaisson. “Brood Stimulation Controls the Phasic Reproductive Cycle of the Parthenogenetic Ant Cerapachys Biroi.” *Insectes Sociaux* 53.1 (2006): 20–26. Web. 25 Jan. 2018.

Robinson, Gene E. “Regulation of Honey Bee Age Polyethism by Juvenile Hormone.” *Behavioral Ecology and Sociobiology* 20.5 (1987): 329–338. Web. 14 Feb. 2018.

Roth, Jürgen et al. 2000. “Occurrence of Sialic Acids in Drosophila Melanogaster.” 256(5057): 673–75.

Rust, Michael K, Donald A Reierson, Eileen Paine, and Lisa J Blum. 2000. “Seasonal Activity and Bait Preferences of the Argentine Ant (Hymenoptera : Formicidae).” *Journal of Agricultural and Urban Entomology* 17(4): 201–12.

Rust, M K, K Haagsma, and D A Reierson. “Barrier Sprays to Control Argentine Ants (Hymenoptera: Formicidae).” *J. Econ. Entomol.* N.p., 1996. Web. 17 Dec. 2015.

Saunders, D. S. “Controversial Aspects of Photoperiodism in Insects and Mites.” *Journal of Insect Physiology* 56.11 (2010): 1491–1502. Web.

Scharf, M. E., Tarver, Mathew, and X. Zhou. 2006. "Hexamerin-Based Regulation of Juvenile Hormone-Dependent Gene Expression Underlies Phenotypic Plasticity in a Social Insect." *Development* 134(3): 601–10. Web.

Scharf, Michael E., Zachary J. Karl, Amit Sethi, and Drion G. Boucias. 2011. "Multiple Levels of Synergistic Collaboration in Termite Lignocellulose Digestion." *PLoS ONE* 6(7): 1–7.

Schauer, Roland. "Achievements and Challenges of Sialic Acid Research." *Glycoconjugate Journal* 17.7–9 (2000): 485–499. Web.

Seehuus, S.-C. et al. "Reproductive Protein Protects Functionally Sterile Honey Bee Workers from Oxidative Stress." *Proceedings of the National Academy of Sciences* 103.4 (2006): 962–967. Web. 24 Jan. 2018.

Simola, Daniel F. et al. "A Chromatin Link to Caste Identity in the Carpenter Ant Camponotus Floridanus." *Genome Research* 23.3 (2013): 486–496. Web.

Suarez, A V, J Q Richmond, and T J Case. "Prey Selection in Horned Lizards Following the Invasion of Argentine Ants in Southern California." *Ecological Applications*. N.p., 2000. Web. 29 Nov. 2015.

Suarez, Andrew V, David A Holway, and Ted J Case. "Patterns of Spread in Biological Invasions Dominated by Long-Distance Jump Dispersal: Insights from Argentine Ants." *Proceedings of the National Academy of Sciences* 98.3 (2001): 1095–1100. Web. 26 Oct. 2015.

Sumner, Seirian. "The Importance of Genomic Novelty in Social Evolution." *Molecular Ecology* 23.1 (2014): 26–28. Web.

Sumner, Seirian, J. J.M Pereboom, and William C Jordan. "Differential Gene Expression and Phenotypic Plasticity in Behavioural Castes of the Primitively Eusocial Wasp, Polistes Canadensis." *Proceedings of the Royal Society B: Biological Sciences* 273.1582 (2006): 19–26. Web. 25 Feb. 2018.

Telfer, William H, Pamela S Keim, and John H Law. “Arylphorin , a New Protein From Hyalophora Cecropia : Comparisons With Calliphorin and Manducin.” *Insect Biochemistry* 13.6 (1983): 601–613. Print.

Terra, Walter R et al. “Insect Digestive Enzymes : Properties , Compartmentalization and Function.” *Science* January (1994): n. pag. Print.

Tian, Li, and Xuguo Zhou. “The Soldiers in Societies: Defense, Regulation, and Evolution.” *International Journal of Biological Sciences* 2014: 296–308. Web. 13 Mar. 2018.

Toth, Amy L., and Gene E. Robinson. “Evo-Devo and the Evolution of Social Behavior.” *Trends in Genetics* 23.7 (2007): 334–341. Web.

Tsuji, K., and K. Yamauchi. “Production of Females by Parthenogenesis in the Ant, Cerapachys Biroi.” *Insectes Sociaux* 42.3 (1995): 333–336. Web. 25 Jan. 2018.

Tsutsui, Neil D et al. “Reduced Genetic Variation and the Success of an Invasive Species.” *Proceedings of the National Academy of Sciences of the United States of America* 97.11 (2000): 5948–53. Web. 7 Oct. 2015.

Tsutsui, Neil D, and Ted J Case. “Population Genetics and Colony Structure of the Argentine Ant (Linepithema Humile) in Its Native and Introduced Ranges.” *Evolution* 55.5 (2007): 976–985. Web. 11 Nov. 2015.

Tu, Meng-Ping, Chih-Ming Yin, and Marc Tatar. “Mutations in Insulin Signaling Pathway Alter Juvenile Hormone Synthesis in Drosophila Melanogaster.” *General and Comparative Endocrinology* 142.3 (2005): 347–356. Web.

Wada-Katsumata, Ayako et al. 2015. “Gut Bacteria Mediate Aggregation in the German Cockroach.” *Proceedings of the National Academy of Sciences* 112(51): 201504031.

Weers, P. M M, and Robert O. Ryan. “Apolipophorin III: Role Model Apolipoprotein.” *Insect Biochemistry and Molecular Biology* 36.4 SPEC. ISS. (2006): 231–240. Web.

Wheeler, Diana E. “Developmental and Physiological Determinants of Caste in Social Hymenoptera: Evolutionary Implications.” *The American Naturalist* 128.1 (1986): 13–34. Web. 23 Feb. 2018.

---. “The Developmental Basis of Worker Caste Polymorphisms in Ants.” *The American Naturalist* 138.5 (1991): 1218–1238. Web. 26 Jan. 2018.

Wheeler, Diana E, N. Buck, and Jay D Evans. “Expression of Insulin Pathway Genes during the Period of Caste Determination in the Honey Bee, Apis Mellifera.” *Insect Molecular Biology* 15.5 (2006): 597–602. Web. 18 Feb. 2018.

Wheeler, Diana E, and H. Frederik Nijhout. “Soldier Determination in Ants: New Role for Juvenile Hormone.” *Science* 213.4505 (1981): 361–363. Web. 14 Feb. 2018.

---. “Soldier Determination in Pheidole Bicarinata: Effect of Methoprene on Caste and Size within Castes.” *Journal of Insect Physiology* 29.11 (1983): 847–854. Web.

Wheeler, Diana E, and Teresa Martinez. “Storage Proteins in Ants (Hymenoptera : Formicidae).” *Comparative Biochemistry and Physiology* 112B.1 (1995): 15–19. Print.

Wilson, Edward O. *The Insect Societies*. Cambridge, Massachusetts, USA, Harvard University Press [Distributed by Oxford University Press, 1971. Web. 30 Oct. 2018.

Winston, Mark L. et al. “Genome-Wide Analysis Reveals Differences in Brain Gene Expression Patterns Associated with Caste and Reproductive Status in Honey Bees (Apis Mellifera).” N.p., 2007. Web. 19 Feb. 2016.

Wolschin, Florian, Navdeep S Mutti, and Gro V Amdam. “Insulin Receptor Substrate Influences Female Caste Development in Honeybees.” *Biology letters* 7.1 (2011): 112–5. Web. 19 Feb. 2018.