EXPLORING GENE REGULATORY DIFFERENCES BETWEEN MONOGYNE AND POLYGYNE ALATE GYNES OF THE RED IMPORTED FIRE ANT *SOLENOPSIS INVICTA*

by

JOAN THERESA KING

(Under the Direction of Brendan G. Hunt)

ABSTRACT

Identifying the molecular basis of adaptation is a fundamental, but often challenging goal of evolutionary biology. Recently, a growing number of cases have been discovered where complex adaptive phenotypes are mediated by groups of tightly linked genes (supergenes). *Solenopsis invicta* possesses a supergene and is used as a model for the evolution of social behavior due to its polymorphism in colony social structure (two social forms: monogyne and polygyne). Previous studies have shown that many of the differentially expressed genes between the two social forms are within the supergene; however, many of these studies have been performed with whole-body samples and microarray analysis— which have limitations. By using specific tissues and RNA-seq, it was found that genotype (supergene presence or absence) has a larger effect than colony social environment on gene expression in ovaries, and that colony social environment affects brain gene expression more than genotype.

INDEX WORDS: Supergenes, *Solenopsis invicta*, Complex phenotype, Eusociality, Social behavior, RNA-seq, Differential gene expression

EXPLORING GENE REGULATORY DIFFERENCES BETWEEN MONOGYNE AND POLYGYNE ALATE GYNES OF THE RED IMPORTED FIRE ANT *SOLENOPSIS INVICTA*

By

JOAN THERESA KING

B.S., The University of Central Florida, 2013

A Thesis Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

MASTER OF SCIENCE

ATHENS, GEORGIA

2017

EXPLORING GENE REGULATORY DIFFERENCES BETWEEN MONOGYNE AND POLYGYNE ALATE GYNES OF THE RED IMPORTED FIRE ANT *SOLENOPSIS INVICTA*

by

JOAN THERESA KING

Major Professor: Brendan Hunt

Committee: Kenneth Ross
Patricia Moore

Electronic Version Approved:

Suzanne Barbour
Dean of the Graduate School
The University of Georgia
August 2017

DEDICATION

I would like to dedicate this thesis to three people: Dr. Linda Walters, Hester Dingle, and Jarrad Pond. Dr. Walters provided many amazing opportunities and guidance that were integral to my graduate career. She not only taught me about science, but about the importance of service and outreach. I cannot thank her enough for all of her hard work and for believing in me— she truly is an inspiring professor. As for Hester, I would not be studying what I am if it were not for her. Hester spent time with me, introducing the wonder of ants and their biology. I could not be happier with my field and research direction. Finally, I would like to thank Jarrad Pond for his patience, support, and friendship. Jarrad is an amazing human.

ACKNOWLEDGEMENTS

I would like to acknowledge several people that contributed to this body of scientific work (who later will be contributing authors to a future manuscript): Samuel Arsenault for his amazing assistance with bioinformatics and data analysis. Sasha Kay for teaching me molecular lab techniques, as well as performing RNA extractions and dissections. Kip Lacy for helping with *Solenopsis invicta* collecting and for preforming microsatellite genotype analysis. Dr. Kenneth Ross for providing guidance, resources, and support throughout my Master's work. Dr. Brendan Hunt for research and academic guidance, resources, and encouragement not only in research, but other academic pursuits, such as teaching and study abroad. I would further like to thank other major contributors to this work: Dr. Laurent Keller and Oksana Riba for providing a forthcoming *S. invicta* genome build Si_gnH_C3. Dr. Bob Schmitz and Nick Rohr for technical assistance and library preparation. Dr. Dietrich Gotzek for contributing to the experimental design and assistance with sampling. Finally, I would like to acknowledge Dr. Jarrad Pond for statistical guidance and proof-reading drafts, and Dr. Patricia Moore for helping me be less esoteric with ant natural history and bringing me back to the "full picture."

This study was supported in part by resources and technical expertise from the Georgia Advanced Computing Resource Center, a partnership between the University of Georgia's Office of the Vice President for Research and Office of the Vice President for Information Technology. Thank you, everyone!

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

I. Introduction

Understanding the molecular basis and evolutionary maintenance of complex suites of adaptive traits are challenging problems in evolutionary biology. In some cases, the co-inheritance of a suite of genes together shape variation in such traits (Tuttle et al., 2016; Wang et al., 2013). In particular, variation in clusters of two or more genetically-linked loci, or "supergenes," can collectively shape complex traits because these supergenes are comprised of individual genes that each affect different characteristics. Supergenes are increasingly being discovered to play a role in the mediation of complex adaptive phenotypes (Schwander, Libbrecht, & Keller, 2014).

The red imported fire ant, *Solenopsis invicta*, is a model for the evolution of social behavior because these ants exhibit a social polymorphism: they have a monogyne form and a polygyne form. The monogyne form contains one reproductive queen per colony while the polygyne form contains multiple reproductive queens per colony. Individuals differ in many morphological, behavioral, and physiological traits in accordance with their social form. This is due to a single Mendelian element, which comprises a large part of a chromosome (Ross & Keller, 1998). This "social chromosome" bears several inversions in the '*Sb*', versus '*SB*', variant. The *SB*/*SB* genotype in queens results in monogyne colony structure, *SB*/*Sb* results in polygyne colony structure, and *Sb*/*Sb* is effectively lethal (Wang et al. 2013, Huang & Wang

2014). This inversion region of the *Sb* chromosome is described as a supergene because it does not recombine with the *SB* variant and is inherited as a single unit (Huang & Wang, 2014). Recent differential gene expression studies have led to insight into how the supergene mediates the differences between social forms (Bourke & Mank, 2013; Wang et al., 2013). Further investigating differences in transcriptomes of *S. invicta* from different castes and developmental stages will pave the way to a better understanding of how supergenes mediate variation in complex phenotypes.

II. Complex Phenotypes and Supergenes

Understanding how complex phenotypes arise is a challenging topic in evolutionary biology because complex phenotypes involve multiple allelic combinations that can be rearranged by recombination (Thompson & Jiggins, 2014). Difficulties arise in the maintenance of complex phenotypes because individuals with divergent traits can interbreed (Thompson & Jiggins, 2014). Reducing recombination between favorable combinations of alleles is one way to ensure that alternative allelic combinations are maintained, as is the case with supergenes (Schwander et al., 2014; Thompson & Jiggins, 2014). Normally, genetic recombination increases the rate of adaptation by breaking down linkage disequilibrium generated by random genetic drift and selection, a phenomenon known as the Hill-Robertson effect (Comeron et al. 2008). However, the Hill-Robertson effect can also reduce the frequency of combinations of alleles that produce beneficial traits (Comeron, Williford, & Kliman, 2008). Recently, supergenes have been found to regulate a broad range of complex alternative phenotypes in a diverse group of organisms (Schwander et al., 2014) (Table 1). These complex alternative phenotypes can be behavioral, morphological, and/or physiological.

Table 1. Organisms currently known to have a supergene maintaining and regulating complex phenotypes**.**

Organism	Complex alternative phenotypes	References
Atlantic cod	Different migratory and stationary ecotypes	(Berg et al., 2016; Kirubakaran et al., 2016)
Cichlids	Differences in camouflage (orange-blotch color pattern)	(Roberts, Ser, & Kocher, 2009)
Turnera flowers	Floral heteromorphy: short-style phenotype	(Gilmartin & Li, 2010)
Heliconius butterflies	Multiple sympatric mimicry morphs	(Jones, Salazar, Jiggins, & Joron, 2012; Joron et al., 2011; Joron et al., 2006; Kunte et al., 2014; Le Poul et al., 2014)
Mice	t-haplotype (male sterilization through impaired sperm)	(Lyon, 2003)
Red imported fire ant	Polymorphic colony social structure	(Wang et al., 2013)
Ruff	Reproductive strategies	(Küpper et al., 2016; Lamichhaney et al., 2015)
White-throated sparrows	Morphs that widely differ in reproduction and behavior	(Knapton & Falls, 1983; Tuttle, 2003; Tuttle et al., 2016)

Solenopsis invicta is ideal for studying complex phenotypes regulated by a supergene and is increasingly becoming a model organism (Gotzek & Ross, 2007; Wang et al., 2013). This is because these ants are easy to collect, rear, and have a rich background on many aspects of their behavior, genetics, physiology, morphology, and natural history (Tschinkel, 2006).

III. Social Polymorphism in *S. invicta*— Natural History & Genetics

Since its introduction from South America into the United States and other parts of the world, *S. invicta* has been extensively studied (Callcott & Collins, 1996). Due to its nature as an invasive pest, much of the biology of this ant has been investigated (Tschinkel, 2006). Since variation of social form regarding the number of queens in a colony was discovered, there have been many studies on the morphology, physiology, and behavior of queens, workers, and males in both social forms of *S. invicta*.

A. Queen differences

Many phenotypic differences exist between queens of the three different social chromosome genotypes (Table 2). All queens with a copy of the *Sb* allele are polygyne and queens with the genotype *SB/SB* are monogyne. Workers from polygyne colonies will tolerate multiple reproductive females with a copy of *Sb*, and workers from monogyne colonies will tolerate only a single *SB/SB* queen (Ross & Keller, 1998). The most notable difference between the social forms is that in a monogyne colony, there is only one reproductive queen, while polygyne colonies have a few up to hundreds of reproductive queens (Ross, Vargo, & Keller, 1996; Tschinkel, 2006).

In social Hymenoptera, gynes are the female reproductives. Alate is a term that refers to a winged insect; therefore, an alate gyne is a winged female reproductive. It is important to differentiate between gyne and alate, because not all reproductive hymenopterans have alate forms; some species have ergatoid female reproductives that look morphologically similar to workers (Peeters, 1991) or gamergates (mated worker ants able to reproduce sexually) (Hölldobler & Wilson, 1990). In *S. invi*cta, female reproductives are alate gynes that participate in mating-flights, mate, and then dealate upon starting a colony. *SB/SB* alate gynes disperse far from the nest during their mating flights while typical *SB/Sb* polygyne alate gynes fly short distances at low levels (Goodisman, DeHeer, & Ross, 2000) and try to gain reentry into the parent or nearby polygyne nests (Gotzek & Ross, 2007). Very few *Sb/Sb* alate gynes participate in mating flights, if at all (Fritz, Vander Meer, & Preston, 2006). It has also been shown that the dispersal strategies of *SB/SB* alate gynes from polygyne nests are similar to those of *SB/SB* gynes from monogyne nests, even though they have slightly lower

weights than their monogyne counterparts (DeHeer, Goodisman, & Ross, 1999). Thus, it appears that genotype directly affects the dispersal behavior of alate gynes.

Table 2. *S. invicta* queen differences between genotypes.

Queen social chromosome genotype	*SB/SB*	*SB/Sb*	*Sb/Sb*	References
Social form	Monogyne	Polygyne	Polygyne, but is inviable	(Ross & Keller, 1998)
Number(s) of queens per colony	One	Multiple	N/A	(Gotzek & Ross, 2007; Tschinkel, 2006)
Mass prior to reproduction	High	Medium	Low	(DeHeer, 2002; DeHeer et al., 1999)
Dispersal and mating behavior	Alate gynes accumulate ample fat reserves, mate and disperse widely, independently and claustrally attempt to start colonies	Fly at low altitudes and try to gain reentry into the natal nest; Colony budding	Rarely participate in mating flights	(Goodisman et al., 2000; Gotzek & Ross, 2007; Ross & Keller, 1995)
Oviposition behavior and physiology	High oviposition rates	Low oviposition rates	N/A	(Vander Meer, Morel, & Lofgren, 1992)
Life span	After successful colony founding, are long-lived (6-8 years)	Short-lived (maximum of 2.5 years)	Most die shortly after eclosion	(Tschinkel, 2006; Vargo & Porter, 1989)

There are also differences in reproductive capacity between the two social forms. Once monogyne queens disperse and start a colony, they have higher oviposition rates and lose less weight per egg than individual polygyne queens (Vander Meer et al., 1992). *SB/SB* queens from polygyne colonies are assumed to perform like those from monogyne colonies since they have been shown to have similar morphologies and elicit

the same mating behaviors (DeHeer et al., 1999; Goodisman et al., 2000; Gotzek & Ross, 2007; Ross & Keller, 1995).

B. Worker differences

Like the female reproductives, workers differ by social form (Table 3). Worker sizes vary between the social forms, where polygyne workers are smaller and less variable in size than monogyne workers (Tschinkel, 2006). Polygyne workers are less aggressive towards conspecifics than monogyne workers, and polygyne mounds are closer together when compared to monogyne colonies (Greenberg, Fletcher, & Vinson, 1985; Tschinkel, 2006). Workers from polygyne colonies will only tolerate female reproductives with a copy of the *Sb* allele, whereas *SB/SB* reproductive females are executed (Huang & Wang, 2014; Ross & Keller, 1998).

C. Male differences

Male sexuals of *S. invicta* exhibit behavioral, morphological, and physiological differences due to the supergene genotype (Table 4). Males of hymenopteran taxa, including *S. invicta*, are generally haploid; however, *S. invicta* diploid males are frequently produced by polygyne colonies and are largely sterile (Hung, Vinson, & Summerlin, 1974; Ross & Fletcher, 1985). Diploid males will not be discussed in this review. *Sb* males have significantly less mass than their *SB* counterparts (Goodisman, Mack, Pearse, & Ross, 1999). Male reproduction is also affected by the supergene. Sperm counts differ according to genotype: *Sb* males do not have adequate sperm stores compared to *SB* males. Females will attempt to remate after copulation with an *Sb* male, while queens that mate with *SB* males rarely remate (Lawson, Vander Meer, &

Shoemaker, 2012). Like alate gynes and queens, workers in polygyne colonies only tolerate males that have a copy of the *Sb* allele. Most *SB* males are executed by polygyne workers (Fritz et al., 2006).

Table 3. *S. invicta* worker differences between genotypes.

Worker social form	*Monogyne*	*Polygyne*	References
Worker chromosome genotype	*SB/SB* only	*SB/SB* and *SB/Sb*. *Sb/Sb* is effectively lethal	(Ross, 1997)
Queen acceptance	*SB/SB* alate gynes and queens only	Alate gynes and queens must have a copy of *Sb*	(Ross & Keller, 1998)
Intercolony aggression and nest distribution	Intercolony Aggression	Less Intercolony aggression	(Morel, Vander Meer, & Lofgren, 1990)
Worker size and mass	Larger size on average	Smaller size on average	(Goodisman et al., 1999; Gotzek & Ross, 2007; Tschinkel, 2006)

IV. Differential Gene Expression Studies in *S. invicta*

The morphological, behavioral, and physiological nature of *S. invicta* is quite complex: there are many differences between these aspects within and between queens, workers, and males (Tables 2-4). Thus, there are vast amounts of natural history data available (Tschinkel, 2006). Moreover, there have been numerous experiments investigating the molecular biology of these fire ants— especially since the first *S. invicta* genome was published in 2011 (Wurm et al., 2011). I will discuss some studies on the genetic effects on reproduction and development, as well as effects of social environment, in *S. invicta*.

Table 4. *S. invicta* male differences between genotypes.

Male social chromosome genotype	*SB*	*Sb*	References
Male production by social form	Produced by both monogyne and polygyne colonies; however, many *SB* males are executed in monogyne colonies.	Polygyne colonies only	(Fritz et al., 2006)
Male mass	High	Low	(Goodisman et al., 1999)
Dispersal and mating behavior	Fly far from the natal nest. Alate gynes mate with *SB* males once	Alate gynes mated to *Sb* males remate	(Fritz et al., 2006; Goodisman et al., 2000; Hölldobler & Wilson, 1990; Lawson et al., 2012)
Sperm count	High	Low	(Lawson et al., 2012)

A. Genetic effects on reproduction and development

A single Mendelian factor, marked by alleles of the gene *Gp-9* (codes for an odorant-binding protein), comprises a large part of the social chromosome and is responsible for many of the differences found between monogyne and polygyne colonies (Gotzek & Ross, 2007; Keller & Ross, 1998). In the study that determined *Gp-9* is part of a non-recombining region of the *Sb* chromosome, Wang et al. (2013) found that 19 genes differentially expressed between *SB/SB* and *SB/Sb* workers were in the nonrecombining region (*Sb*) and that 38 genes were differentially expressed between *SB/SB* and *SB/Sb* virgin queens; 19 of which were found to be in *Sb*. Wang et al. (2013) not only characterized this genomic region, but also made differential gene expression comparisons of whole-body males and queens. RAD sequencing data from whole-body *SB* and *Sb* males, as well as RNA-seq analysis using whole-body *SB/Sb* queens,

showed that a putative acyl-CoA desaturase gene was expressed at significantly lower levels in *Sb* males — this is compelling because genes in this family are known to be involved in pheromone and cuticular hydrocarbon synthesis, which may implicate *Sb* involvement in odor differences between the alternative genotypes (Wang et al., 2013).

Gene expression profiles also differ between monogyne and polygyne workers. Thirty-nine genes were differentially expressed between whole-body workers of the two genotypes, *SB/SB* and *SB/Sb*, according to microarray analysis (Wang, Ross, & Keller, 2008). Of these, around two-thirds were more highly expressed in *SB/Sb* than *SB/SB* workers from polygyne colonies (Wang et al., 2008). However, Wang et al. (2008) also discovered 91 differentially expressed genes between *SB/SB* workers of the different social forms, of which over 75% were more highly expressed in polygyne workers compared to monogyne workers. At least for workers, social form appears to have a greater effect than supergene genotype on gene expression.

Since the discovery of the supergene in *S. invicta* in 2013, differential gene expression studies relating to reproduction and development have been conducted. Nipitwattanaphon et al. (2013) showed that many genes within the supergene are involved in regulating social organization in fire ants. Differences in physiology and behavior of the two social forms of queen fire ants can be attributed to widespread differences in gene expression. Microarray anaylsis was performed on whole-body unmated *SB/SB* and *SB/Sb* polygyne queens. Reproductive queens were also collected: *SB/SB* queens were collected from monogyne colonies (*SB/SB* don't reproduce within polygyne colonies because they will be executed. See Table 3), while *SB/Sb* queens were collected from polygyne colonies. RNA was extracted from one

SB/SB and one *Sb/Sb* queen for both a one day old and 11 day old from 16 colonies (later had 15 *SB/SB* 11-day old queens and 15 *SB/Sb* 11-day old queens due to poor hybridization against a standard reference). Field collected reprodutive queens from 8 colonies for *SB*/*SB* and *SB*/*Sb* were also used. Therefore, 1-day old queens, 11-day old queens, and reproductive queens were used for microarray anaylsis. RNA was extracted from the whole-body of each indiviual queen sample and their *S. invicta* microarray contained 11,024 cDNAs with a single PCR product estimated at 5956 unique genes (previously found by (Wang et al., 2007)) that matched transcripts from RNA-seq data of *S. invicta*. Six desaturase and three elongase genes differential expressed between *SB/SB* and *SB/Sb* queens were confirmed using real-time quantitative reverse transcription PCR (Qrt-PCR). Of the 5956 genes, the expression of 4386 were found to be attibuted to age class, while only 1028 for genotype (*SB/SB* vs *SB/Sb*). Interestingly, the up- or downregulation of 550 of these 1028 genes became faster or slower for older *SB/SB* queens compared to *SB/Sb* queens. The fewest differences in gene expression between *SB/SB* and *SB/Sb* queens were for 1-day old queens (38 genes), which increased in 11-day old queens (689 genes), and then decreased again for fully reproductive queens (295 genes). Even more compelling, 616 of the 689 genes that were differentially expressed between 11-day old queens were also differentially expressed between *SB/SB* and *SB/Sb* queens — *SB/SB* queens gain more fat and have a quicker maturation time than *SB/Sb* (Keller & Ross, 1999). Nipitwattanaphon et al. (2013) also found that genes inside the supergene were overrepresented among the differentially expressed genes in the various age classes of queens with the *SB/Sb* genotype. These data suggest that social chromosome

genotype (‘*SB*’ vs. ‘*Sb*’) affect phenotypic differences between monogyne and polygyne queens.

Nipitwattanaphon et al. (2014) explored the gene expression levels of diploid males, haploid males, and queens of *S. invicta* using microarray analysis in order to distinguish what the influence of ploidy level and sex are in determining transcription profiles. They found that ploidy level had a much greater effect on gene expression level than sex in pupae; however, after eclosion, sex strongly affects gene expression levels. Interestingly, there was also a difference between sperm and pheromone production: the level of expression of genes involved in sperm production were much higher in haploid males than in queens or diploid males. Therefore, their under-expression during the pupal stage may be the reason why diploid males are aspermatic (M. Nipitwattanaphon et al., 2014).

B. Social environment

Changes in the social environment can lead to changes in gene expression. Wurm, Wang, and Keller (2010) identified molecular changes in virgin queens that had begun competing for reproductive dominance in the nest, after they perceived the loss of their mother queen. In the monogyne form of *S. invicta*, colonies can produce hundreds to thousands of alate gynes. These alate gynes are fed by workers to build up fat reserves (Wurm, Wang, & Keller, 2010). Once these queens reach sexual maturity, it is in their best interest to not become reproductive within the colony because they will be executed by workers (Vargo & Laurel, 1994; Vargo & Porter, 1993). The normal mode of action is for these virgin alate gynes to participate in mating flights and found their own colony; however, if the mother queen dies, many of these alates will shed their

wings and begin emitting pheromonal signals. Therefore, if the mother queen is not present and emitting pheromones, virgin nestmate alate gynes will no longer perceive signals that a reproductive queen is present, and they will attempt to replace the mother queen (Fletcher, Cherix, & Blum, 1983; Vargo, 1999). To identify the molecular changes that follow after the loss of the mother queen is perceived, Wurm et al. (2010) collected virgin queens at different times before and after orphaning. RNA was extracted and through microarray analysis, they examined the expression of 10,000 genes and found that 297 of these genes were consistently differentially expressed after orphaning. These genes are purported to be involved in the signaling and onset of reproductive development due to annotations of these differentially expressed genes. They also found that three putative olfactory genes, two chemo-sensory proteins (CSPs), and one odorant binding protein (OBP) were upregulated in virgin queens. This study provided insight to reproduction and reproductive status in *S. invicta*, as well as observing that *Gp-9* had the highest sequence similarity to OBP — which it is now know that *Gp-9* is part of a social chromosome.

In 2013, Manfredini et al. performed a differential gene expression study to investigate whether social environment or reproductive state had a greater effect on gene expression. As discussed earlier, monogyne colonies are founded by a single reproductive queen, while polygyne queens try to gain reentry into the parent nest, join a nearby nest, or reproduce by budding (Goodisman et al., 2000; Gotzek & Ross, 2007; Ross & Keller, 1995). However, there are many instances where newly mated queens from monogyne colonies will find a colony in pleometrotic groups—this is in contrast to haplometrosis, were one queen founds a colony individually (Tschinkel & Howard,

1983). In *S. invicta*, when monogyne newly mated queens found a colony via pleometrosis, ultimately only one egg-laying queen remains the sole reproductive of the colony ("winner") because monogyne colonies only tolerate one reproducing queen, and all other cofoundresses must compete to the death ("losers") for the position (Balas & Adams, 1996; Bernascon, Krieger, & Keller, 1997; Keller & Ross, 1993). The transition from cooperation to rival fights occurs once the first cohort of workers emerges (Manfredini et al., 2013). Whether a newly mated monogyne queen starts a colony independently or with other cofoundresses, they have to face a claustral period in which they have enough reserves left in order to survive until her workers can care for her (Tschinkel, 2006). It has been shown that there are behavioral and physiological differences between haplometrotic queens and the "winner" and "loser" pleometrotic queens — haplometrotic queens lose more weight in the claustral period and produce more brood than pleometrotic queens (Tschinkel, 1995). Using microarray analysis with a set of 16,569 genes, gene expression patterns were investigated in foundress queens, and it was found that social environment has a great effect on the determination of the patterns of gene expression. In addition, Manfredini et al. found gene expression differences in genes involved in metabolism, stress response, aging, reproduction, and immunity — many of which are part of essential biological processes of *S. invicta* queens.

C. Communication in social environment

Since the discovery of the single mendelian factor that separates single versus multiple reproductive queens in *S. invicta* (Ross, 1992), there has been an interest in how genetic traits influence colony social structure. Before *Gp-9* was found to be part of

the supergene in *S. invicta*, it was known that social organization was under genetic control (Dietrich Gotzek & Kenneth G. Ross, 2007).

In *S. invicta*, OBPs are chemical carriers that are involved in many physiological processes, including chemical communication. *Gp-9* is an odorant binding protein (SiOBP3) and has been shown to be overexpressed in workers of polygyne colonies (Gotzek, Robertson, Wurm, & Shoemaker, 2011). In 2016, Zhang et al. investigated expression of odorant binding proteins (OBPs) of different tissues and castes of *S. invicta*. It was found that *SiOBPs* are differentially expressed in the various castes and tissues: workers had expression of a distinct set of *SiOBPs* highest in heads and antennae, lower levels of expression were found in the antennae of males, and the expression of *SiOBPs* in the antennae of female alates were found to be similar to that of the workers (Zhang, Wanchoo, Ortiz-Urquiza, Xia, & Keyhani, 2016).

The genes *Gp-9* and foraging (*for*) have been shown to be associated with behavioral polymorphisms in *S. invicta*. Lucas, Nicolas, and Keller (2015) investigated the expression of these two genes by analyzing brain quantitative PCR data of worker and nonreproductive alate gynes of both social forms. The expression of these two genes was investigated because 1) *for* has been shown to regulate behavioral polytheism in other social insects and 2) *Gp-9* (part of the *Sb* social chromosome) has been shown to be involved in social polymorphic colony structure in *S. invicta*. By exploring differences in gene expression between *SB/SB* and *SB/Sb* workers in addition to *SB/SB* and *SB/Sb* queens, a better understanding of the role of *Gp-9* in social organization could be attained. They found that depending on the task of the worker, the expression of *for* was higher in monogyne workers verses polygyne workers, and *SB/Sb*

workers have higher levels of *Gp-9* expression than *SB/SB* workers. *Gp-9^{b}* was also found to be highly expressed compared to *Gp-9^{B}* in heterozygote workers and alate gynes. In polygyne colonies, *for* was not significantly associated with *Gp-9;* therefore, they have independent effects on behavior (Lucas, Nicolas, & Keller, 2015).

V. Discussion

A growing body of literature where complex adaptive phenotypes are mediated by supergenes is emerging (Schwander et al., 2014; Thompson & Jiggins, 2014). Investigating *S. invicta* will help better understand how supergenes affect complex phenotypes. With the extensive natural history information, advancing genomic resources, and the availability and ease of rearing, *S. invicta* is a fantastic species to investigate this phenomenon.

Although there is a plethora of natural history data on *S. invicta*, there are still aspects of their biology that remain unknown. Sampling sexuals while in mating flights could elucidate the differences between genotypes because different behaviors are elicited depending on genotype; *SB/SB* fly high, far, and claustrally find a new colony, whereas *SB/Sb* do not fly as high or far and often try to gain reentry into the parent colony. New sampling techniques such as the high-altitude trapping system developed by Fritz, Fritz, and Vander Meer (2011) seem promising to collect sexuals previously thought to be too difficult to obtain. Continuing to link natural history studies with genetic expression analyses will led to a greater understanding of these ants' evolution and genetics. Transcriptomics are helping to unravel genes involved in physiology, morphology, and behavior. Gaining a better understanding of the genes involved in chemical signaling may lead to a better understanding of social organization. Further

investigating OBPs and CSPs, candidate genes that are involved the signaling between individuals of a colony, would be insightful, as they are important for odor and pheromone detection (Nipitwattanaphon, Wang, Dijkstra, & Keller, 2013). Nipitwattanaphon et al. (2013) found that odorant-binding proteins and a chemical signaling protein lie within the supergene and hypothesized that the odorant-binding protein *OBP2* may regulate social form by allowing *SB/Sb* workers to recognize and favor *SB/Sb* queens. Therefore, it would be interesting to further investigate how these proteins regulate social form via discrimination. Pathways involved in reproduction, for example, pheromone production, lipid storage and oogenesis, may also lead to a better understanding of how environment, genomic state, behavior, and physiology influence social evolution.

By using RNA-seq analysis with specific tissues, many of the limitations of microarray using full body can be overcome. Although microarray analysis can generate a vast amount of data, acquisition and data mining is difficult (Abdullah-Sayani, Bueno-de-Mesquita, & Van De Vijver, 2006). In addition, microarrays provide an indirect measure of relative concentration, have laborious set up in which arrays that have multiply related DNA or RNA sequences will not bind to the same probe on the array, and the DNA array is only capable of detecting sequences that it was specified to detect (Bumgarner, 2013). RNA-seq analysis will hopefully lead us to a better understanding of how supergenes contribute to complex phenotypes. Most of the present differential gene expression studies in *S. invicta* have used full body samples and have been limited to certain genotypes, especially of queens (*SB/SB* and *SB/Sb*). In the future, it would be interesting to see molecular analysis including *Sb/Sb*, as Nipitwattanaphon et

al. 2013 were not able to obtain enough individuals of these queens (rare and typically die soon after eclosion) (Tschinkel, 2006; Vargo & Porter, 1989).

CHAPTER 2

RESEARCH PROJECT

I. Introduction

Identifying the molecular basis of adaptation is a fundamental, but often challenging goal of evolutionary biology. Recently, a growing number of cases have been discovered where complex adaptive phenotypes are mediated by groups of tightly linked genes (Schwander et al., 2014; Thompson & Jiggins, 2014). These “supergenes” are clusters of two or more linked loci that each affect different characteristics. It is hypothesized that supergenes may help to alleviate the selective cost of recombination between co-adapted alleles (Thompson & Jiggins, 2014). Supergenes thus allow complex phenotypes to exist within an interbreeding population without maladaptive intermediate phenotypes. Recombination is reduced or inhibited by physical linkage of multiple loci within a supergene (e.g., brought about by a pericentric inversion (Schwander et al., 2014)) (Thompson & Jiggins, 2014). Examples of organisms with variation in complex phenotypes attributed to supergenes include white-throated sparrows, which exhibit morphs that differ greatly in reproductive behavior (Knapton & Falls, 1983; Tuttle, 2003; Tuttle et al., 2016), Atlantic cod, which maintain migratory and stationary ecotypes (Kirubakaran et al., 2016), and *Heliconius* butterflies, which possess multiple sympatric mimicry morphs (Jones et al., 2012; Joron et al., 2011; Joron et al., 2006; Kunte et al., 2014; Le Poul et al., 2014).

Solenopsis invicta, the red imported fire ant, is an important model for understanding the evolution of social behavior because this species exhibits a social polymorphism that is subject to the control of a supergene (Ross & Keller, 1998; Wang et al., 2013). Overall, the status of this supergene among a colony of ants results in two different social forms of *S. invicta*: the monogyne form and polygyne form. If a colony has the supergene present among its inhabitants, it is polygyne; if it is not present, a colony is monogyne. Monogyne *S. invicta* colonies contain one reproductive queen, while polygyne colonies contain many queens. These social forms also differ in many other life-history traits (Gotzek and Ross 2007, Huang and Wang 2014).

The supergene represents a single Mendelian element, marked by alleles of the gene *Gp-9*, that comprises a large part of chromosome 16 and is responsible for the differences observed between monogyne and polygyne colonies (DeHeer et al., 1999; Gotzek & Ross, 2007; Keller & Ross, 1998). This “social chromosome” bears several inversions in the ‘*Sb*’, versus ‘*SB*’, variant that spans approximately 13 megabases and contains more than 600 genes (Pracana, Priyam, Levantis, Nichols, & Wurm, 2017; Wang et al., 2013). Queens of the *SB*/*SB* genotype found colonies that exhibit a monogyne structure, *SB*/*Sb* queens found or join polygyne colonies, which exclusively accept *SB/Sb* queens, and *Sb*/*Sb* queens are typically incapable of reproduction (Keller & Ross, 1998; Ross & Keller, 1998). The inverted region of the *Sb* chromosome is described as a supergene because it does not exhibit substantial recombination with either the *SB* or *Sb* variants, and is inherited as a single unit (Wang et al., 2013).

Fire ant queens fly only during a mating flight and dealate (shed their wings) after mating, triggering the breakdown of flight muscle (Wheeler, 1996). The nutritional

constraints on queens starting new colonies are extreme — she must lay her eggs and rear her first set of workers completely on body reserves (Wheeler, 1996). *S. invicta* alate pre-reproductive queens (gynes) with an *SB/SB* genotype disperse widely and independently establish colonies, whereas *SB/Sb* gynes mate and attempt to remain in the natal nest, or partake in localized mating flights and enter other nearby polygyne nests (Goodisman et al., 2000; Gotzek & Ross, 2007; Ross & Keller, 1995). In conjunction with this variation in colony founding strategies, *SB/SB* gynes are observed to have higher nutrient reserves than *SB/Sb* gynes, which in turn possess higher nutrient reserves than *Sb/Sb* gynes, when embarking on mating flights (DeHeer et al., 1999; Ross & Shoemaker, 1997). Notably, *SB/SB* gynes produced by polygyne colonies display physiological traits and behaviors similar to the *SB/SB* gynes produced by monogyne colonies (DeHeer, 2002; DeHeer et al., 1999).

The physiological and behavioral differences between monogyne and polygyne gynes are the result of widespread differences in gene expression (Nipitwattanaphon et al., 2013). Previous studies have shown that many of the differentially expressed genes between the two social forms are within the supergene on the social chromosome (Nipitwattanaphon et al., 2013; Wang et al., 2008; Wang et al., 2013). Wang et al. 2013 showed using RAD sequencing analysis results from whole-body *SB* and *Sb* males, as well as RNA-seq analysis using whole-body *SB/Sb* queens, that out of 38 differentially expressed genes between 1-day-old virgin queens of genotypes $Gp\text{-}9^{BB}$ and $Gp\text{-}9^{Bb}$, that 15 could be mapped to known linkage groups and of those, four were located in the non-recombining region. Using microarray assays of whole-body queens, Nipitwattanaphon et al. (2013) further demonstrated expression differences between

queens of both social forms. Genes inside the supergene were overrepresented among differentially expressed genes in various age classes of queens with different social chromosome genotypes. These data suggest that phenotypic differences between monogyne and polygyne queens are due to their social chromosome genotype ('*SB*' vs. '*Sb*'), but how we perceive these gene expression differences is incomplete because whole-body samples cannot give specific tissue comparisons. Not only are these conclusions from microarray data incomplete, but microarrays in general have shortcomings — they are capable only of detecting expression differences for sequences that are used as probes and are poor at quantifying lowly and highly expressed genes (Kukurba & Montgomery, 2015).

We performed a differential gene expression analysis using tissue-specific RNA-seq in order to test effects of social environment and genotype on the phenotype of *S. invicta* gynes. To test for an effect of social chromosome genotype on gene expression, we compared the overall gene expression profiles of pre-mating flight alate gynes of polygyne colony origin (^{P}) bearing *Sb*/*Sb*, *SB*/*Sb*, or *SB*/*SB* genotypes to those of alate gynes of monogyne colony origin (^{M}), all of which have the *SB*/*SB* genotype. We sequenced transcripts from specific tissues, brain and ovary, that are known to be involved in the behavior and physiology of reproduction. Monogyne gynes begin oogenesis early, have high fecundity and high metabolic efficiency compared to polygyne gynes (DeHeer, 2002; Tschinkel, 2006). Brains secrete hormones involved in reproductive processes that may lead to differential fecundity between polygyne and monogyne alate gynes (Bendena, Garside, Yu, & Tobe, 1997; Boulay, Hooper-Bui, & Woodring, 2001). By disentangling gene expression profiles in relevant tissues, we

hope to gain a better understanding of how supergenes contribute to complex phenotypes.

II. Materials and methods

A. Alate gyne collection

Solenopsis invicta alate gynes were collected and snap frozen on dry ice in the field in Athens, Georgia, USA on days of mating flights, in April 2015. Alate gynes were collected from colonies of both the polygyne and monogyne social form. Colonies of both social forms in Athens, GA do not produce sexuals during the winter months and produce their first and largest pulse of sexuals in early- to mid-spring (Vargo & Fletcher, 1987). The first major mating-flights occur the day after a major rainfall, at the warmest part of the day (Collins & Scheffrahn, 2001). Any overwintering queens would have flown by late March (Fletcher & Blum, 1983). Samples of each social form were collected in areas where the *S. invicta* populations are known to be dominated by either the polygyne form or monogyne form. Workers were collected from monogyne colonies in addition to alate gynes. All samples were stored at -80°C.

B. Sample processing

RNA*later*®-ICE Frozen Tissue Transition Solution (Ambion™) was used to preserve RNA integrity. The manufactures' protocol for transitioning tissue from -70°C or colder to -20°C was modified as follows to maximize the extracted RNA quality. Alate gynes were stored in 15ml tubes at -80°C until sample processing. Following the addition of RNA*later*®-ICE, alate gynes were soaked for at least 24 hours at -20°C. After 24 hours, more RNA*later*®-ICE was added so that all gynes were submerged in the solution (approximately 3 times the volume of the gynes) and were then stored at -

20°C for another 24 hours. If there was not enough room for the gynes to be fully submerged, they were moved to a larger tube.

Alate gynes from polygyne colonies were then sorted by size class — small, medium, and large — in a Petri dish on top of dry ice. These size classes were chosen based on the degree of physogastry (enlargement due to growth of fat bodies) of the gaster. The smallest were pooled as prospective *Sb/Sb*^{P}, medium as *SB/Sb*^{P}, and largest as *SB/SB*^{P}. After the smallest and largest alate gynes were sampled, medium-sized gynes were randomly picked from the sample. Three legs were pulled off of each gyne in a pre-chilled weighing dish and forceps were cleaned with ethanol between each alate gyne dissection. After legs were pulled from each gyne for DNA extraction, the rest of the body was placed into a 1.5 ml microcentrifuge tube on dry ice and stored at -80°C until dissections and RNA extraction took place.

C. DNA extraction and genotyping

1. DNA extraction

DNA was extracted from three legs of each alate gyne using mill grinding and a modified Puregene DNA isolation kit (Gentra Systems, Qiagen, Valencia, California) protocol — legs from each individual sample were placed in sterilized (with UV crosslinker) safe-lock 1.5 mL tubes (Eppendorf) with 4 to 6 zirconia 2.0 mm diameter beads and 200 μL of PBS (1X concentration), and then fully ground with a Mixer Mill MM 301 instrument (set to 30 Hz). For DNA extraction, a cocktail of 200 μL of Cell Lysis Solution and 2 μL of Proteinase K (if more or less of the solution was needed, the amount of each reagent was scaled linearly, preserving proportions) was gently inverted 30 times and then added to each sample. After tubes were briefly vortexed, they were

incubated 4-8 hours in a 55°C water bath. Tubes were next cooled on ice for one minute and then centrifuged briefly (a few seconds). 140μL of Protein Precipitation Solution was added to each tube and then tubes were vortexed for ~20 seconds. After five minutes in a centrifuge, 500 μL of 100% isopropanol was then added to separate 1.5 mL centrifuge tubes, and the aqueous phase of the solution was added to the tubes containing 500 μL of 100% isopropanol and inverted. The mixture was centrifuged for another five minutes. Supernatant was then poured into a waste beaker and 500 μL of 70% ethanol were added to the tubes. Tubes were then inverted several times and then placed in the centrifuge for five minutes. After pouring out supernatant and allowing the tubes to dry while inverted for four to six hours, 35 μL of DNA Hydration Solution (1X TE) was added. For whole-body monogyne workers and alate gynes, 300 μL was added. For monogyne samples, in addition to alate gynes, DNA was extracted from the pooled whole-bodies of females to confirm monogyne colony identity — a monogyne colony should only contain *SB/SB*^{M} workers.

2. Genotyping

To genotype each alate gyne collected from presumed polygyne nests, a *Gp-9* PCR assay was performed (Valles & Porter, 2003). To confirm social form of alate gynes from monogyne colonies, 15-20 individuals (mix of workers and alate gynes) were pulled from each sample colony and the same *Gp-9* PCR assay was performed. Genotype was then independently confirmed with another round of PCR using a modified version of the protocol by Valles and Porter (2003), as in Goodisman et al. (2007). After genotyping, it was confirmed that sorting yielded eight alate gynes from monogyne colonies and eight alate gyne of each of the three genotypes (*SB/SB*^{P},

SB/Sb^{P}, and *Sb/Sb*^{P}) from polygyne nests. Eight replicates for RNA-seq analysis is considered sufficient for reliable results (Conesa et al., 2016).

3. Microsatellite data

All gynes sampled from polygyne nests were genotyped at nine previously described polymorphic microsatellite loci (Ascunce et al., 2011) to check for possible triploidy (Krieger et al., 1999) and to determine whether any of the gynes sampled from the same colony were siblings. These microsatellite data revealed that no sampled individuals were triploid and no individuals from six of the eight nests were siblings (Table 5 and Table 6). For two nests, it was impossible to determine sibling status using the initial 10 markers (*Gp-9* and nine microsatellites), so the three alate gynes from these nests (six gynes total) were genotyped at four additional microsatellites (Ascunce et al*.* 2011) (Table 6). Microsatellites were amplified using previously described methods (Ascunce et al*.* 2011) from the same stock DNA used for genotyping at *Gp-9*. PCR amplicons were diluted (34:1 or 45:1) and pooled into a single semi-skirted 96-well plate. 1.5 µL of this dilution was added to a plate, and both Liz 600 size standard (0.1 µL) and formamide were added to all dilutions before being run on an ABI-3730XL-96 capillary sequencer (Applied Biosystems) at the Georgia Genomics Facility at the University of Georgia. Microsatellite genotypes were scored using the GeneMarker software (SoftGenetics).

D. Tissue dissections and RNA extractions

1. Tissue dissections

Alate gynes from polygyne and monogyne colonies were dissected under an Olympus SZ61 stereomicroscope. Gynes were decapitated prior to brain extractions.

The head was held upside down in a droplet of RNAlater-ICE. A minutin probe was first used to make a linear series of perforations in the cuticle, extending from the postgena on either side, and along the occipital carina. Then, the probe was used to gently tear through the perforations to make a U-shaped opening in the cuticle. The second pair of forceps was used to tear away the cuticle from the back of the head, by grasping the trough of the “U” and pulling toward the mouthparts, revealing fibrous mandibular musculature that usually had to be removed to expose the brain and its associated glands. The minutin probe was used to disconnect the optic nerves of the compound eyes and ocelli from the cuticle inside the head, freeing the brain from the head capsule; the probe was also used to help clear the glands surrounding the brain. The majority of tracheae and glands were removed; however, there were some traces of tracheae and possibly residual gland tissue attached to the brains.

The ovaries were extricated from the hindgut, and any Malpighian tubules present were removed, along with some excess fat body. Due to the nature of tracheae entrenchment in the ovaries, removing all traces of tracheae from the ovary tissue is not possible. Tissues and body parts were placed in individual 1.5 ml microcentrifuge tubes, and stored at -20°C (short-term) or -80°C (long-term) until RNA and/or DNA extraction.

2. RNA extraction

RNA from brains was extracted using the RNeasy Micro Kit (Qiagen 74004). RNA extraction from ovaries was performed using the RNeasy Mini Kit (Qiagen, 7410). Extracted RNA integrity and concentration was evaluated on an Agilent 2100 bioanalyzer using RNA 6000 pico kit at the Georgia Genomics Facility. RNA integrity numbers (RIN) were calculated from the electrophoretic traces (Schroeder et al., 2006).

Samples with RIN scores lower than 6 were not used in this study. However, it is important to note that insect RNA may appear degraded when assessed using this method because rRNA profiles can differ significantly from the standard benchmark due to the 28S rRNA of most insects containing an endogenous “hidden break” — this occurs because during denaturation, the masking hydrogen bonds are disrupted, which releases two similar sized fragments that both migrate close to the 18S rRNA (Winnebeck, Millar, & Warman, 2010).

E. Library preparation and Illumina sequencing

Libraries were prepared by the Schmitz Lab in the Genetics Department at the University of Georgia following the Smart-seq2 protocol (Picelli et al., 2014). Samples were barcoded and pooled and then run on an Illumina NextSeq sequencer for 75 cycles to produce 75 bp single-end reads. Illumina sequencing was performed at the Georgia Genomics Facility at the University of Georgia. The numbers of reads produced for each sample are given in Table 7. 1.2 ng of RNA was used for brain libraries, 3.6 ng of RNA was used for typical ovary libraries. Ovary samples 107GO, 240AO and 30AO were used to test different concentrations of RNA; 1.2 ng of RNA was used for a low input ovary test, and 10 ng of RNA was used for a high input ovary test. The 1.2 ng samples are labeled with an ‘_L’ and the 10ng samples are labeled with an ‘_H.’

F. Data Analysis

1. Quality Control

Quality control (QC) is an important step in RNA-seq analysis. It ensures that RNA-seq results are reliable and reproducible (Conesa et al., 2016). Assessing the sequence quality of raw reads and discarding low-quality reads, percentage of mapped

reads (to reference genome), and quantification are all important steps of QC (Conesa et al., 2016). QC is a critical part of the RNA-seq pipeline because it helps remove biases in the data (Li, Nair, Wang, & Wang, 2014). Reads were removed that aligned to the *Myrmecia croslandi* (an Australian bull ant) genes for 18S rRNA, 5.8S rRNA, 28S rRNA using BLAT (v 3.5). BLAT is a pairwise sequence alignment algorithm that indexes the genome database, retains an index in memory, and scans the query sequence for matches (Kent, 2002). rRNA contaminants were removed through BLAT because a library should not be made of large amounts of rRNA (Adiconis et al., 2013; Conesa et al., 2016). Typically, rRNA constitutes over 90% of total RNA of the transcriptome; however, mRNA are the main focus of RNA-seq analysis (Conesa et al., 2016). Most of the rRNA should be removed during RNA-seq library preparation by poly-A mRNA isolation in the Smart-seq2 protocol. The Smart-seq2 protocol has limitations of a lack of strand specificity and inability to detect non-polyadenylated (polyA) RNA (Picelli et al., 2014). On the other hand, Smart-seq2 transcriptome libraries have improved accuracy, bias, coverage, and detection compared to other library preparation methods (Picelli et al., 2013). The remaining reads were trimmed with trimmomatic (v 0.32). Trimmomatic is a software tool that trims and filters reads for obtaining high quality data (Bolger, Lohse, & Usadel, 2014). Trimmomatic was used with the following parameters: ILLUMINACLIP:nextera_adapter.fa:2:30:10 to remove adapters,TRAILING:3 to remove trailing low quality or N bases (below quality 3), LEADING:3 to remove leading low quality or N bases (below quality 3), SLIDINGWINDOW:4:15 to scan the read with a 4-base wide sliding window, cutting

Table 5. Microsatellite data of gynes sampled from polygyne nests. Nests were genotyped at nine previously described polymorphic microsatellite loci: *Sol-42f, Sol-49*, *Sdag C27*, *Sdag C294*, *Sdag C536*, *Sol i129*, *Sol i120, Bertha*, and *Cassidy* (Ascunce et al., 2011). Each pair of columns represents a microsatellite locus. There are two for each sample alate gyne because each one is diploid. None were found to be triploid. The number in the 'Sample ID' column represents the colony of origin. Loci by colony are shaded in yellow when indicating that the *Gp-9* homozygote individuals are not siblings of the *Gp-9* heterozygotes. ** denotes no amplification.

Sample I. D.	Gp-9 genotype	C294		Sol-49		Bertha		Sol-42f		cassidy_PigTail		C536		I129		C27		I-120_PigTail	
104A	Bb	92	112	156	156	200	212	116	116	239	269	103	107	154	154	211	211	318	326
104D	BB	100	106	137	156	206	206	114	114	239	275	93	99	154	154	211	211	314	318
104G	bb	92	92	156	158	206	212	128	128	239	275	87	93	154	154	211	211	314	330
107A	Bb	92	106	144	158	206	206	116	140	239	275	87	93	146	154	207	211	314	326
107E	BB	106	112	144	158	206	206	116	140	239	275	101	107	146	154	211	211	314	318
107G	bb	92	92	137	156	206	206	116	128	239	269	93	93	150	154	207	213	314	320
107H	bb	92	92	156	162	206	212	114	118	275	275	87	107	154	154	211	211	318	330
15A	Bb	92	100	144	156	206	206	116	140	239	239	87	107	146	146	211	211	318	318
15D	BB	100	106	156	162	206	206	114	116	239	239	99	103	146	154	211	213	318	328
15G	bb	92	92	**	**	206	212	**	**	**	**	87	113	146	154	**	**	314	326
16C	Bb	92	112	156	162	200	206	114	140	269	275	99	103	154	154	211	211	318	330
16D	BB	106	106	162	162	206	206	114	128	239	275	99	103	146	154	211	213	318	318
16G	bb	92	92	144	156	206	206	114	128	239	269	93	93	154	154	211	211	314	320
19A	Bb	92	106	156	162	200	212	116	128	239	263	93	103	154	154	211	211	314	314
19D	BB	100	106	137	156	206	206	114	128	269	275	107	107	154	154	211	213	**	**
19G	bb	92	92	156	158	200	212	118	118	239	269	93	103	154	154	211	213	314	314
20A	Bb	92	106	137	158	206	212	116	128	269	275	101	107	154	154	211	211	314	314
20D	BB	100	112	158	158	200	200	116	118	239	269	93	103	154	154	207	211	314	318
20I	bb	92	92	137	156	200	200	114	116	239	269	91	107	154	154	211	213	318	320
30A	Bb	92	100	162	162	200	206	114	114	239	275	93	99	154	154	211	211	314	320
30E	BB	137	144	**	**	200	206	114	118	239	275	99	107	154	154	211	211	314	314
30G	bb	92	144	156	156	200	206	114	128	239	269	95	101	154	154	211	213	314	330
5A	Bb	92	106	156	162	200	200	114	128	239	275	93	101	154	154	211	211	314	318
5G	bb	92	144	156	156	200	206	114	128	269	275	93	107	154	154	211	211	314	318
1E	BB	100	112	156	158	200	200	114	114	269	269	87	99	146	154	211	211	314	320

Table 6. Additional microsatellite data of four additional microsatellites. *Sol i109*, *Sol i114*, *Sol i126*, and *Sunrise* (Ascunce et al., 2011) of alate gynes sampled from polygyne nests. Loci by colony are shaded in yellow to show that the *Gp-9* homozygote individuals are not siblings of the *Gp-9* heterozygotes.

Sample I.D.	*Gp-9* genotype	sunrise		i-*109*		i-*126*		i-*114*	
30A	*Bb*	79	85	147	147	209	221	303	317
30E	*BB*	79	85	147	147	209	218	303	317
30G	*bb*	79	79	147	149	218	218	307	307
5A	*Bb*	79	85	147	147	221	230	301	301
5G	*bb*	79	85	149	154	209	218	307	307
1E	*BB*	79	79	147	149	209	209	303	309

when the average quality per base drops below 15, and MINLEN:36 to drop reads below 36 bases long.

The percentage of reads removed due to rRNA contamination, percentage of reads dropped due to trimming, and new uniquely mapped reads (both percentage and number) are shown in Table 8. New mapped reads are the result of reads left after QC.

2. Differential gene expression analysis

Reads were aligned using STAR (v 2.5.3a) to the forthcoming *S. invicta* genome build Si_gnH_C3 using the parameters suggested for outputting to Cufflinks. Out of our 64 samples, one *SB/SB*^{P} ovary sample (140DO) was removed after STAR (alignment stage) because of low alignment to the *S. invicta* genome and due to poor QC measures. Normalization for comparing gene coverage values can be done by Reads

Table 7. Number of reads produced by Illumina NextSeq for each sample. Note that 1EB reads from two different sequencing runs were merged after alignment and 239AB is pending additional sequencing.

Library name	Original number of reads	Library name	Original number of reads
104AB	33111435	207AO	15811927
104AO	29722254	209AB	36482675
104DB	32732215	209AO	22261232
104DO	20904025	20AB	30318463
104GB	33878676	20AO	24884803
104GO	34550562	20DB	29627400
107AB	27502788	20DO	24697986
107AO	20771964	20IB	35051438
107EB	30656024	20IO	24176264
107EO	24775519	222AB	36416671
107GB	38562956	222AO	22014602
107GO	19406487	232AB	21200880
107GO_H	18351490	232AO	20869872
107GO_L	26786026	233AB	25074371
15AB	22800896	233AO	19701356
15AO	22005592	235AB	25610402
15DB	23733303	235AO	26981356
15DO	26072332	239AB	7619112
15GB	21425162	239AO	23420855
15GO	25317089	240AB	25357802
16CB	30045480	240AO	21752724
16CO	30355932	240AO_H	25256457
16DB	36780861	240AO_L	18595526
16DO	23067438	30AB	29616337
16GB	22058284	30AO	18501809
16GO	15906633	30AO_H	23468269
19AB	30207949	30AO_L	20902620
19AO	17961435	30EB	24667565
19DB	38456350	30EO	22998399
19DO	15728174	30GB	22736069
19GB	32027959	30GO	13054033
19GO	24219758	5AB	26072453
1EB	9518477	5AO	18256635
1EB_2	27920647	5GB	24418678
1EO	27346732	5GO	16532622
207AB	35901443	1EB *	37439124

Per Kilobase of transcript per Million mapped reads (RPKM) calculation — RPKM is a within-sample normalization method which corrects for differences in sequencing depth and gene length (Conesa et al., 2016; Mortazavi, Williams, McCue, Schaeffer, & Wold, 2008). The package edgeR for the open source software R Statistics with default parameters was used to calculate RPKM values and to perform differential gene expression analysis (Chen & McCarthy, 2017). The package pheatmap in R Statistics was used to both cluster the RPKM data as well as visualize the clustering as a heatmap (Kolde & Kolde, 2015). An agglomerative hierarchical clustering with complete-linkage was used to group the RPKM values, and the Pearson correlation coefficient was used as a similarity measure for the clustering (Gan, Ma, & Wu, 2007). Thus, with these parameters, the larger the linear correlation between the RPKM values (both within and across genes and within and across samples), the "closer" these samples are considered to be in the clustering. Using edgeR, variation in gene counts across samples was estimated by fitting generalized linear models (GLMs) to the counts data (number of reads) for each gene (attained as a result of the RNA-seq analysis). This variation is also known as the dispersion of the gene or the biological coefficient of variation (BCV) and is a measure of the level at which a gene differs across the individual replicate samples. With the dispersion of the genes estimated, the process of determining the number of differentially expressed genes (DEGs) can begin. To evaluate the number of DEGs between different test conditions, the read counts, dispersion estimates, and the appropriate test conditions for each replicate (i.e., *SB/SB*^{M}, *SB/SB*^{P}, *SB/Sb*^{P}, etc.) were passed to edgeR, and GLMs were fit for each gene within each test condition. Quasi-likelihood F-tests were then used to make

Table 8. Quality control summary. Percentage of reads removed due to rRNA contamination, percentage of reads dropped due to trimming, and new uniquely mapped reads are shown. Quality is colored coded: green is better than red.

Library name	% of reads removed due to rRNA contamination	% of reads dropped due to trimming	% of new uniquely mapped reads	Number of uniquely mapped reads
104AB	5.29%	1.23%	87.41%	28084517
104AO	4.17%	2.28%	89.82%	25825075
104DB	7.09%	1.24%	88.45%	27531766
104DO	0.62%	6.01%	0.68%	136376
104GB	5.36%	1.21%	87.62%	28789552
104GO	5.16%	2.31%	87.78%	28784770
107AB	8.43%	1.47%	90.21%	23249106
107AO	1.41%	2.13%	91.62%	19100827
107EB	7.84%	1.42%	89.22%	25786514
107EO	1.46%	4.23%	88.74%	21598033
107GB	16.14%	1.52%	89.86%	29651144
107GO	1.38%	2.07%	92.33%	17982845
107GO_H	1.82%	2.23%	92.37%	16916270
107GO_L	1.83%	4.37%	89.03%	23263453
15AB	7.13%	1.02%	89.10%	19212978
15AO	2.2%	4.46%	88.38%	18897077
15DB	21.51%	1.20%	88.60%	16820280
15DO	2.02%	4.06%	88.86%	22586328
15GB	6.52%	0.96%	87.78%	18055764
15GO	4.08%	4.21%	85.23%	20396157
16CB	23.02%	1.35%	82.47%	19203296
16CO	3.79%	4.20%	85.20%	24488418
16DB	19.84%	1.86%	86.69%	25808445
16DO	3.72%	4.07%	86.66%	19042982
16GB	12.26%	1.03%	84.18%	16706035
16GO	2.15%	4.13%	85.98%	13335789
19AB	8.57%	1.37%	90.58%	25549776
19AO	1.44%	4.27%	89.01%	15715749
19DB	3.64%	1.27%	83.91%	31514317
19DO	1.53%	4.44%	88.89%	13689671
19GB	6.39%	1.28%	88.81%	27187541
19GO	1.86%	4.31%	88.88%	20968237
1EB	42.13%	1.68%	80.75%	4453280
1EB_2	42.92%	1.40%	81.07%	12974638
1EO	7.16%	4.32%	84.68%	21136540
207AB	6.76%	1.35%	90.78%	30859433
207AO	1.41%	2.16%	92.57%	14680935
209AB	12.26%	1.18%	90.60%	29667681
209AO	1.56%	2.35%	92.72%	20673175
20AB	9%	1.34%	91.16%	25639343
20AO	1.41%	4.45%	88.25%	21522261
20DB	5.5%	1.28%	90.14%	25795429
20DO	1.84%	2.29%	91.37%	22529638
20IB	10.55%	1.89%	88.12%	28131878
20IO	1.81%	2.31%	91.51%	22070902
222AB	4.27%	1.22%	88.31%	31503576
222AO	1.46%	2.24%	92.89%	20482676
232AB	5.32%	1.00%	89.40%	18367287
232AO	0.68%	2.23%	92.61%	19503841
233AB	5.03%	1.05%	89.67%	21829573
233AO	1.15%	2.07%	92.68%	18387902
235AB	14.8%	1.15%	90.41%	20202222
235AO	1.28%	2.41%	92.77%	25107029
239AB	3.72%	1.22%	89.85%	6749504
239AO	1.13%	2.14%	92.57%	21825490
240AB	3.01%	1.01%	89.43%	22541563
240AO	1.72%	2.34%	92.98%	20212560
240AO_H	1.4%	2.36%	92.88%	23487033
240AO_L	1.44%	2.35%	93.02%	17328866
30AB	24.41%	1.41%	91.00%	20959780
30AO	2.81%	4.20%	85.51%	15355233
30AO_H	8.08%	4.71%	81.76%	17413907
30AO_L	2.15%	4.72%	85.12%	17258113
30EB	30.69%	1.38%	91.18%	15947226
30EO	1.54%	4.25%	89.13%	20134883
30GB	5.67%	0.92%	90.17%	19821727
30GO	1.39%	4.29%	88.74%	11373667
5AB	7.92%	1.28%	73.91%	18020584
5AO	2.33%	4.52%	77.24%	13658019
5GB	13.69%	1.13%	91.25%	19663510
5GO	2.58%	3.99%	87.47%	14032838

appropriative pair-wise comparisons between the models for different test conditions within each gene. For each pair-wise comparison, this results in a set of False Discovery Rate (FDR) corrected *p*-values for each gene. The Benjamini-Hochberg method of calculating FDR was used. Genes whose comparison results in a FDR-corrected *p*-value of less than 0.05 are considered to be differentially expressed between the two test conditions (Chen & McCarthy, 2017).

3. Gene Ontology (GO) term enrichment analysis

The newest release (6.16) of the *Drosophila melanogaster* genome from FlyBase (Gramates et al., 2017) was downloaded and gene names were converted to GenBank equivalents based on the included homology records and removed duplicated sequences using SeqKit (Shen, Le, Li, & Hu, 2016). Next, a local BLAST database of the *D. melanogaster* amino acid sequences was created using BLAST2GO (Conesa et al., 2005). The longest transcript for each gene in the *S. invicta* genome build Si_gnH_C3 annotation were gathered and BLASTp (as part of BLAST2GO) was used to query against the local *D. melanogaster* database. The top five BLAST hits with e-values less than 1×10^{-5} were recorded. All other parameters were BLASTp defaults. BLAST2GO mapped the BLASTp hits to corresponding Gene Ontology (GO) terms and a GO term annotation was generated. BLAST2GO's GO-Slim function was used to eliminate redundant GO terms. GO slims give a broad overview of the ontology content and are useful for giving a summary of the results of the GO annotation of a genome (Ashburner et al., 2000).

A list of significant genes was gathered for each comparison using edgeR and a background list of genes was generated based on the genes that passed coverage

filtration (genes with at least 1 count-per-million in 25 or more samples) (Robinson, McCarthy, & Smyth, 2010). The list of these genes was passed to topGO (Alexa & Rahnenfuhrer, 2010) where GO term enrichment was computed using the "weight01" algorithm and the "fisher" statistic. *p*-value less than 0.05 were deemed significant. AMIGO2 was used to get more details of the GO term descriptions (Balsa-Canto, Henriques, Gábor, & Banga, 2016; Consortium, 2015).

4. Overview of *S. invicta* Samples

In total, brain tissue and ovaries were dissected from eight single alate gynes of each genotype from eight nests of each social form (Table 9); importantly, all three *Gp-9* genotypes were represented for each polygyne nest. There is one less replicate (seven) for ovary samples because of one sample's low alignment to the *S. invicta* genome, as mentioned above. This led to a total of 63 samples to be sequenced.

Table 9. Number of replicates by genotype and tissue type. The superscripts M and P denote monogyne social form and polygyne social form, respectively. A total of eight individuals from eight different colonies comprise the *SB/SB* monogyne samples. A total of eight individuals of each genotype (*SB/SB*, *SB/Sb*, and *Sb/Sb*) from each of eight different nests, none of which are siblings (Tables 5 & 6), comprise the polygyne samples.

	Brains	Ovaries	Totals
***SB/SB*^{M}**	8	8	16
***SB/SB*^{P}**	8	7	15
***SB/Sb*^{P}**	8	8	16
***Sb/Sb*^{P}**	8	8	16
Totals	32	31	63

III. Results

Gene expression values from brain and ovary tissue samples form two distinct clusters (Figure 1). Certain clusters of genes have either higher or lower expression depending on the tissue type. It can be seen in the ovary samples that the top clusters of genes are upregulated more than the same set of genes for the brain samples. Furthermore, it can be seen in the brain samples that the bottom clusters of genes are upregulated more than the same set of genes for the ovary samples. This indicates a distinct difference in the gene expression profiles of the two tissue types, implying differences in regulation of expression between brains and ovaries.

The numbers of significant differentially expressed genes (DEGs) within most comparisons largely depends on tissue type (Figures 2 & 3). For clarity, figures 2 and 3 contain the “P” super script denoted for genotypes that would only be polygyne (since a copy of “*Sb*” makes an individual polygyne). There is a large difference in brain DEGs when comparing *SB/SB*^{M} versus *SB/SB*^{P} (476 DEGs; social environment effect), *SB/SB*^{P} versus *Sb/Sb*^{P} (446 DEGs; genotype effect), and *SB/SB*^{M} verses *Sb/Sb*^{P} (169 DEGs; social environment + genotype effect), respectively (Figure 2). Comparing *SB/Sb*^{P} to *Sb/Sb*^{P} has the fewest differentially expressed genes (21 DEGs) within the brain.

Ovaries appear to show a generally different trend in DEGs across comparisons. In Ovaries, the *SB/SB*^{M} versus *Sb/Sb*^{P} comparison possesses the most DEGs (2227; social environment + genotype effect); the comparisons of *SB/SB*^{P} versus *Sb/Sb*^{P} (503 DEGs; genotype effect) and *SB/SB*^{M} versus *SB/Sb*^{P} (684 DEGs) have similar and the next highest levels of differentially expressed genes; and the *SB/SB*^{M} versus *SB/SB*^{P}

(social environment effect), *SB/SB*^{P} versus *SB/Sb*^{P}, and *SB/Sb*^{P} versus *Sb/Sb*^{P} comparisons have the fewest DEGs in the ovaries.

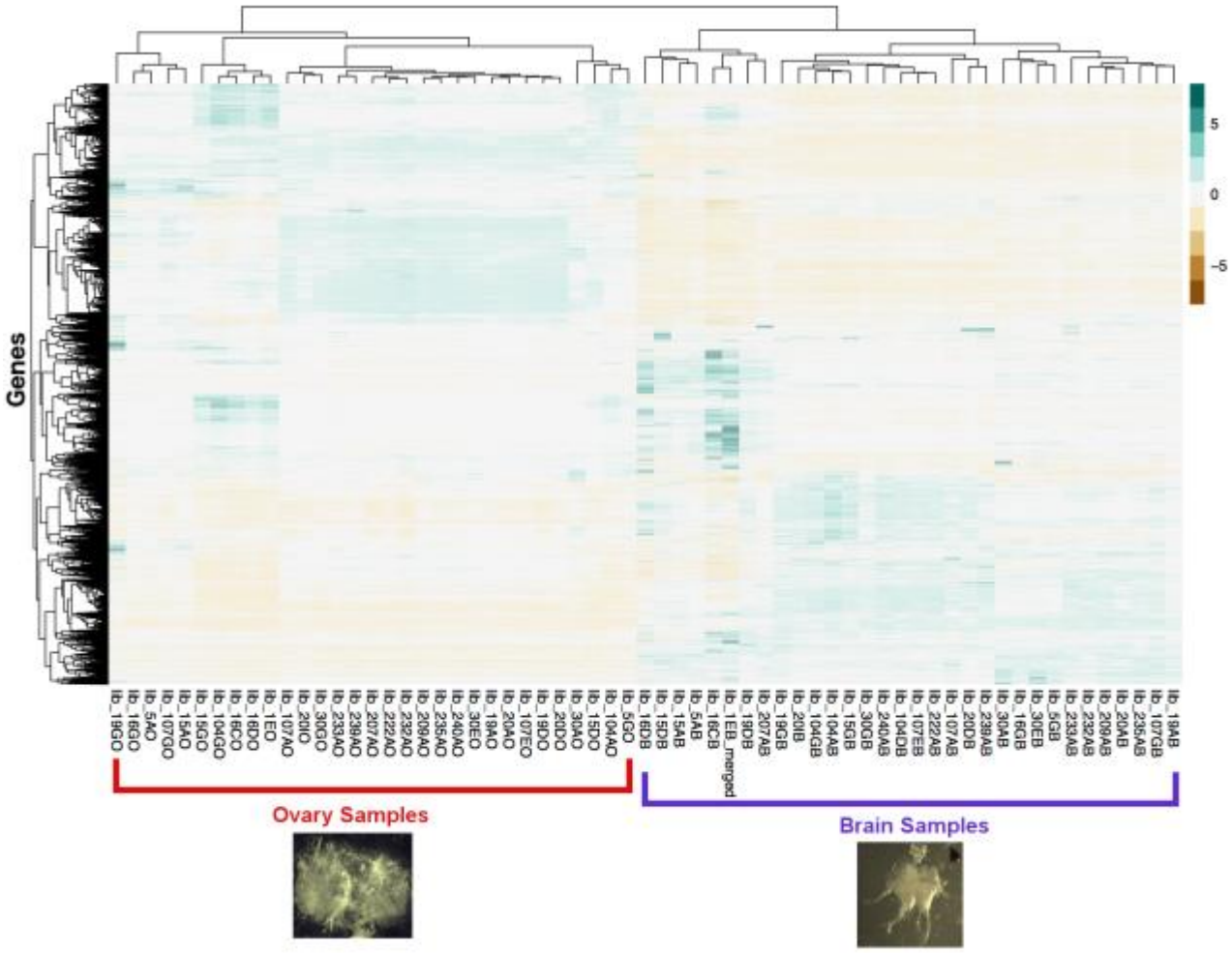

Figure 1. Heatmap of normalized RPKM data from libraries of brain and ovary (from alate gynes) samples show two distinct clusters. Samples are on the x-axis and genes are on the y-axis. Clustering method is Pearson correlation between RPKM values. Photographs of each tissue type are given at the bottom of the figure. Level of expression is given by the color bar on the right: dark green shows higher expression, while dark orange shows lower expression.

To focus in on the most significant DEGs, we report the number of genes that possess a magnitude fold change of greater than two (Figure 3 for brain tissue and

ovary tissue). The trends found when considering only DEGs with a magnitude fold change of greater than two are the same as when considering all the DEGs.

Comparing tissue types, we see the number of DEGs in the *SB/SB*^{M} versus *SB/SB*^{P} social environment comparison is larger in the brains (476 DEGs) compared to the ovaries (98 DEGs). Conversely, for the *SB/SB*^{M} versus *SB/Sb*^{P} and *SB/SB*^{M} versus *Sb/Sb*^{P} comparisons, we see a larger number of DEGs in the ovaries (684 and 2227 DEGs, respectively) compared to brains (81 and 169 DEGs, respectively). For the *SB/SB*^{P} vs *SB/Sb*^{P} comparison, the brain tissue shows 54 DEGs and the ovary tissue show 64 DEGs; for the *SB/Sb*^{P} vs *Sb/Sb*^{P} comparison, 21 DEGs are in the brain tissue and 11 DEGs in the ovaries; for the *SB/SB*^{P} vs *Sb/Sb*^{P} comparison, there are 46 DEGs within the brain tissue and 503 DEGs within the ovary tissue.

Three-way Venn diagrams show how many DEGs are shared and how many are unique for several different comparisons between social form and genotype (Figure 4). These diagrams reveal social environment effects on both brain tissue and ovary tissue (Figure 4. A & C) and genotypic effect for both tissues (Figure 4. B & D). This can help isolate what overall numbers of DEGs are shared and which are unique — further parsing effects of social environment versus genotype. The largest number of unique DEGs when comparing social form within brain tissue is between *SB/SB*^{M} and *SB/SB*^{P}, with 430 unique DEGs, followed by *SB/SB*^{M} and *Sb/Sb*^{P}, with 111 unique DEGs (Figure 4. A). Genotypic comparisons among brain tissue in polygynes show that the most unique DEGs are between *SB/SB*^{P} and *Sb/Sb*^{P}, with 397 unique DEGs (Figure 4. B). The most unique DEGs between social forms for the ovary tissue occur between *SB/SB*^{M} and *Sb/Sb*^{P}, with 1557 unique DEGs (Figure 4. C). Genotypic comparisons

among polygynes shows that the most unique DEGs are between *SB/SB*^{P} and *Sb/Sb*^{P} (431 unique genes) while there are no unique DEGs within ovary tissue between *SB/SB*^{P} and *SB/Sb*^{P}, and only 2 unique DEGs between *SB/Sb*^{P} and *Sb/Sb*^{P} within ovary tissue (Figure 4. D).

We conducted gene ontology (GO) biological process term enrichment analyses to gain insight into the functions of differentially expressed genes in our study. Differentially expressed genes for each tissue type, brain and ovary, for *SB/SB*^{M} vs *SB/SB*^{P} (social environment effect) and *SB/SB*^{P} vs *Sb/Sb*^{P} (genotype effect) and their significant GO terms and their descriptions are given (Table 10). Only two GO terms were shown to be significant for *SB/SB*^{P} vs *Sb/Sb*^{P}, GO:0021700 (developmental maturation) for brain tissue and GO:0040011 (locomotion) for ovary tissue. *SB/SB*^{M} vs *SB/SB*^{P}, for both tissues, yielded several terms: ten terms for brain and six terms for ovary (Table 10).

IV. Discussion

The overall clustering of our data set (Figure 1) reveals that much of the correlation between upregulation and downregulation of genes across samples can be explained by tissue origin of the sample. Brains and ovaries are very different organs that are located in different body regions (head and gaster) and have very different functions; thus, one would expect distinct patterns of gene expression within each tissue that outweigh any observed differences associated with social form or supergene genotype. Observing such a clear division in gene expression between tissue types helps solidify the foundation and integrity of the data as it is further interpreted for differential gene expression within tissue types.

Looking within ovary samples, as expected, genotype has a large effect on gene expression in ovaries. Comparisons involving *SB/SB* alate gynes (from both monogyne and polygyne colonies) and either *SB/Sb*^{P} or *Sb/Sb*^{P} samples generally result in a relatively large number of DEGs, whereas *SB/Sb*^{P} compared to *Sb/Sb*^{P} shows the smallest number (11 DEGs). Evidently, social chromosome has a large effect on overall gene expression in ovary tissue, whether in the heterozygous or homozygous state, suggesting almost complete dominance of the *Sb* element with respect to these expression profiles. This makes sense because queens with a copy of the "*Sb*" chromosome do not mature as quickly as *SB/SB* queens from monogyne colonies (DeHeer, 2002; Tschinkel, 2006). Additionally, *SB/SB*^{P} alate gynes from polygyne nests have been shown to have more mass and higher fat reserves than *SB/Sb*^{P} and *Sb/Sb*^{P} alate gynes (DeHeer et al*.* 1999; Ross & Shoemaker 1997). Therefore, it can be assumed that *SB/SB*^{P} alate gynes have morphological and physiological differences (compared to *SB/Sb*^{P} and *Sb/Sb*^{P} alate gynes, which are more genetically similar to one another) that are attributed to their genotype. We note that the *SB/SB*^{P} versus *SB/Sb*^{P} comparisons results in a small number of DEGs (64), but we hypothesize that these genes, though few in number, may result in the differences in *SB/SB*^{P} and *SB/Sb*^{P} gynes mentioned above. The detailed gene ontology studies conducted in future work will hopefully elucidate what exactly these genes do in *S. invicta*.

Furthermore, there also appears to be an additive effect of genotype and colony social structure in ovary tissue, as comparisons between *SB/SB*^{M} samples and either *SB/Sb*^{P} or *Sb/Sb*^{P} samples result in a larger number of DEGs compared to those made between *SB/SB*^{P} samples and either *SB/Sb*^{P} or *Sb/Sb*^{P} samples. This seems to

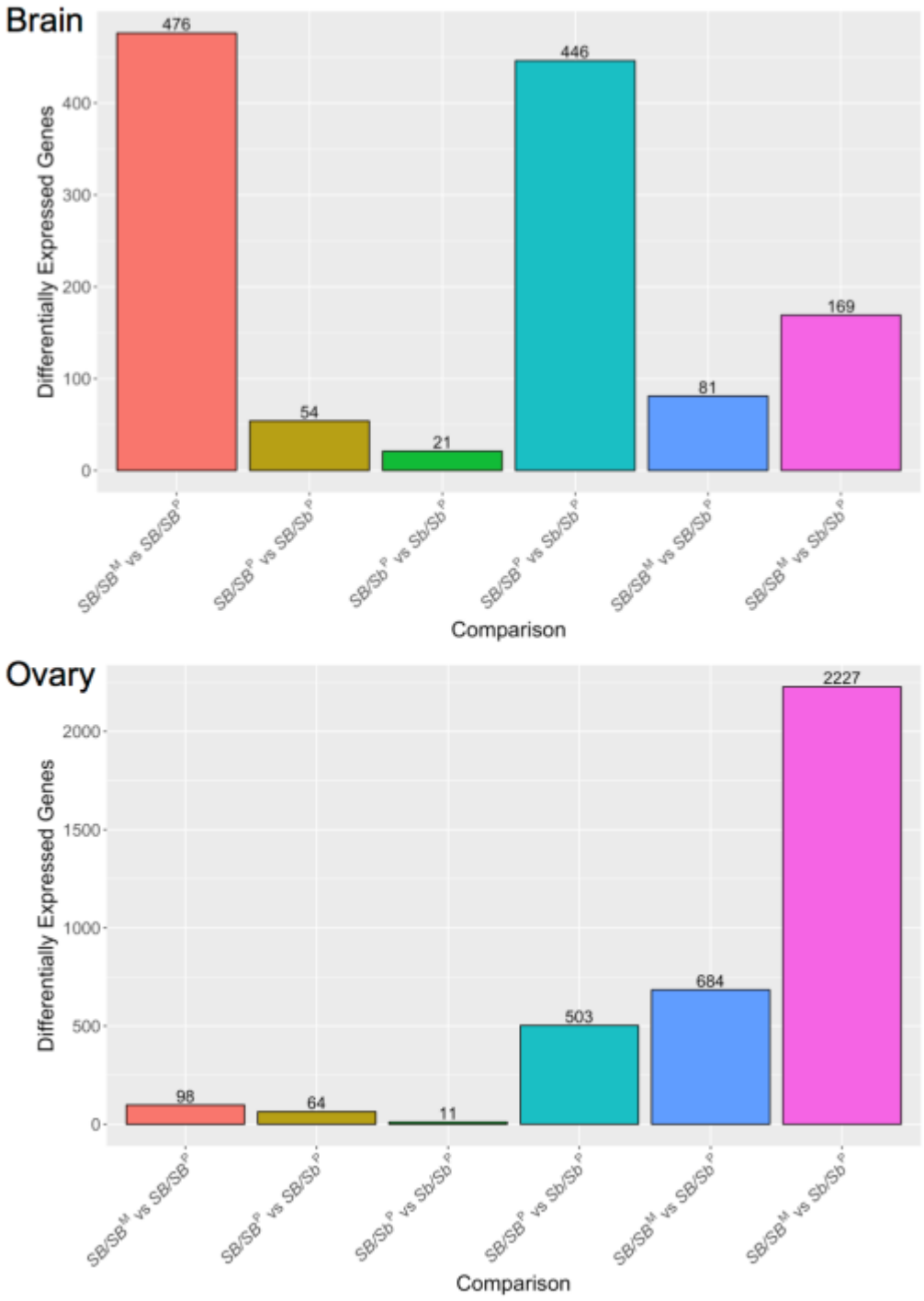

Figure 2. Number of differentially expressed genes (DEGs) per pairwise comparison of social form and genotype. Brain (top) and Ovary (bottom) show the DEGs significant at the false discovery rate (FDR) corrected value of $p < 0.05$. Each of the three supergene genotypes is listed, with the subscripts “M” representing samples from monogyne and “P” from polygyne nests, respectively. Note that y-axis scaling is different for each plot.

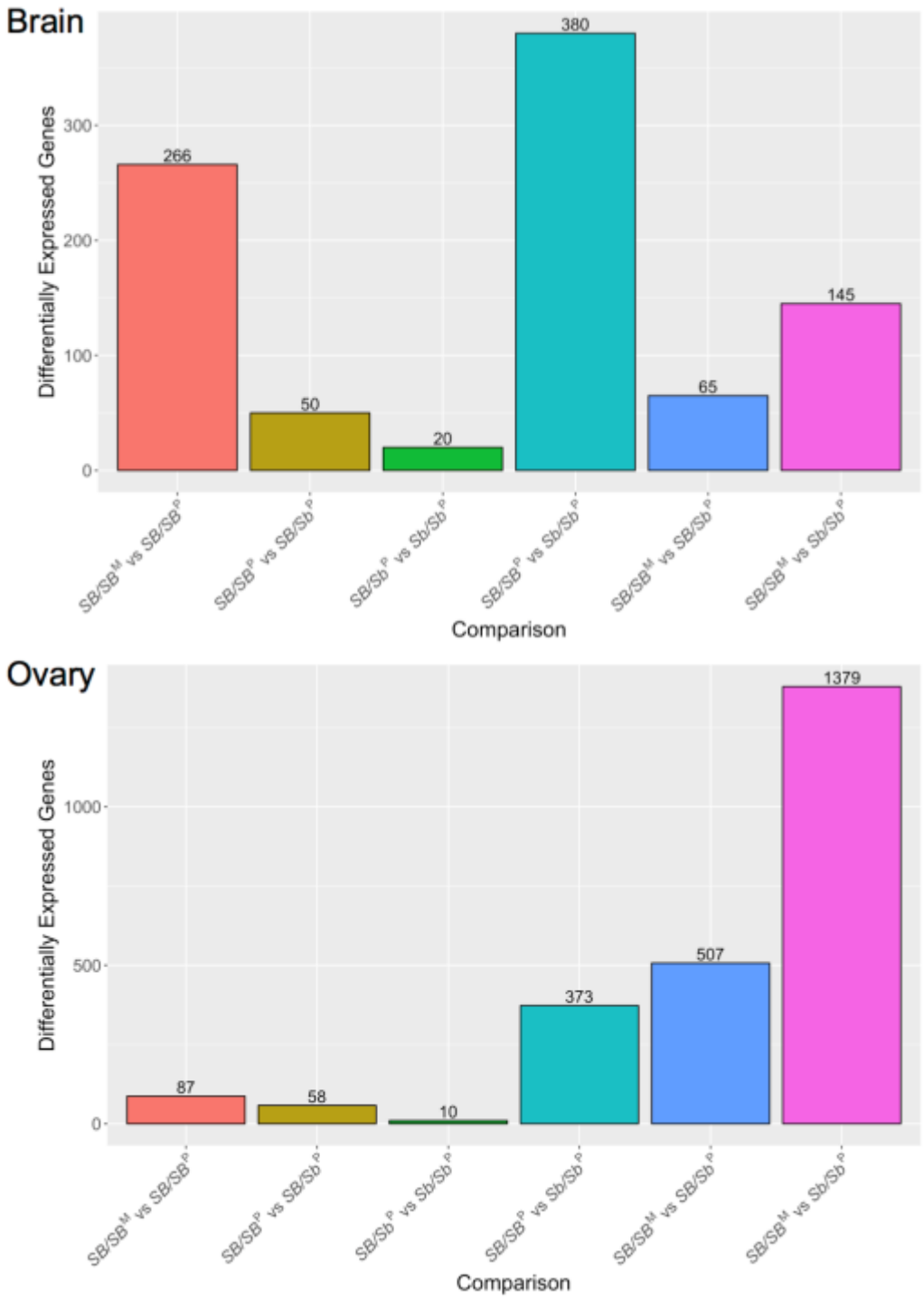

Figure 3. Brain (top) and Ovary (bottom) show DEGs with magnitude fold changes greater than 2 per pairwise comparison of social form and genotype. Each of the three genotypes are present with the subscripts “M” representing from monogyne and “P” representing polygyne, respectively. Note that y-axis scaling is different for each plot.

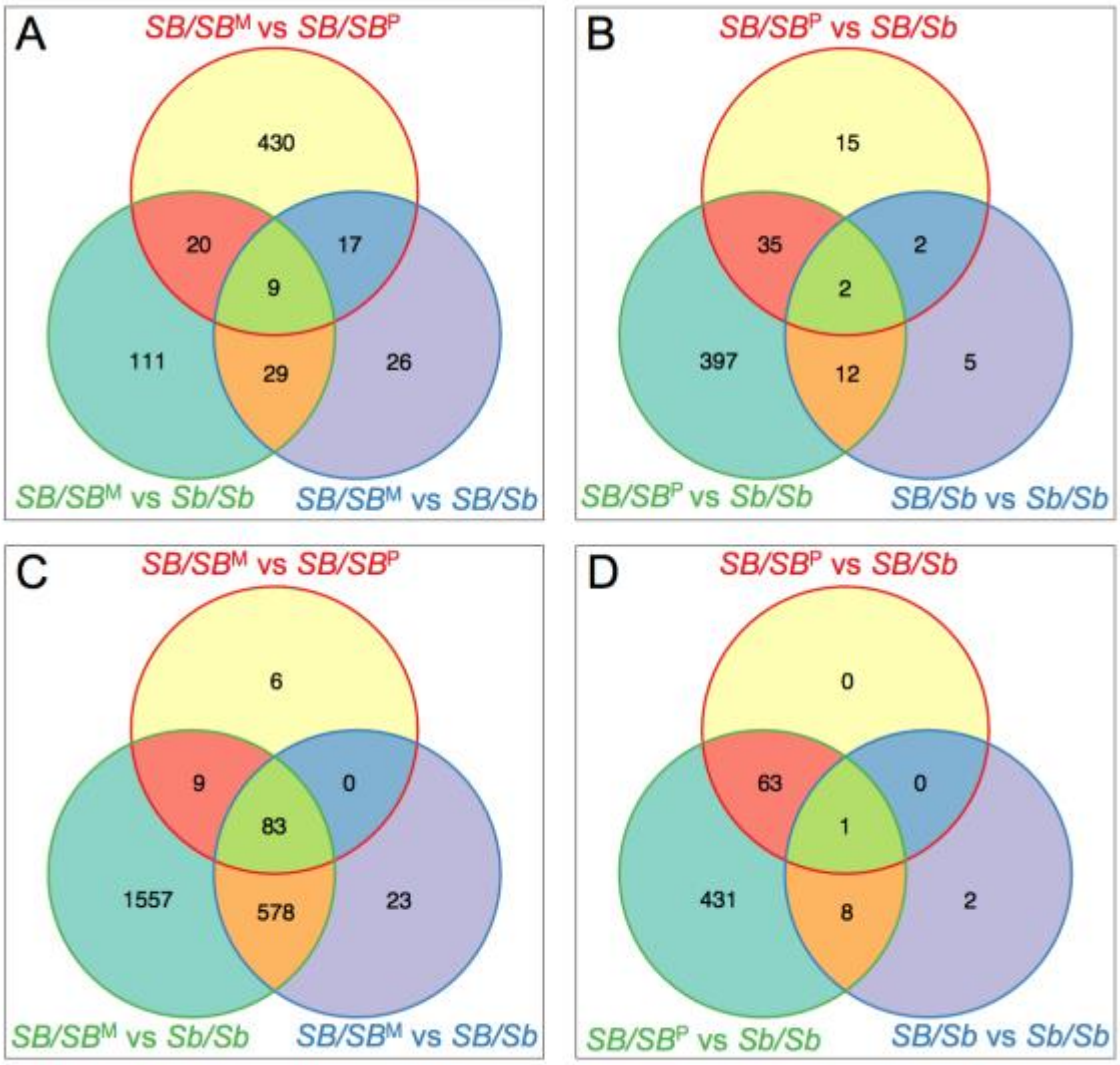

Figure 4. Venn diagrams showing the unique and shared DEGs within tissue types for comparisons between social form and genotype. Panels A and B show brain tissue DEGs for monogyne forms compared to polygyne forms (A) and genotypic comparisons among polygynes (B). Panels C and D show ovary tissue DEGs for monogyne forms compared to polygyne forms (C) and genotypic comparisons among polygynes (D).

Table 10. Biological process Gene Ontology (GO term) enrichment analyses using the GO-Slim function. Comparisons shown are for each tissue type for *SB/SB*^{M} vs *SB/SB*^{P} & *SB/SB*^{P} vs *Sb/Sb*^{P}. No FDR-correction step.

GO term	GO term description	Annotated	Significant	Expected	topGO Fisher *p*- value	Comparison	Tissue
GO:0006412	translation	283	29	14.99	0.00039	*SB/SB*^{M} vs *SB/SB*^{P}	Brain
GO:0006913	nucleocytoplasmic transport	67	10	3.55	0.00249	*SB/SB*^{M} vs *SB/SB*^{P}	Brain
GO:0006605	protein targeting	87	11	4.61	0.00578	*SB/SB*^{M} vs *SB/SB*^{P}	Brain
GO:0051186	cofactor metabolic process	73	9	3.87	0.01414	*SB/SB*^{M} vs *SB/SB*^{P}	Brain
GO:0043473	pigmentation	106	11	5.61	0.02368	*SB/SB*^{M} vs *SB/SB*^{P}	Brain
GO:0006397	mRNA processing	193	17	10.22	0.02584	*SB/SB*^{M} vs *SB/SB*^{P}	Brain
GO:0000902	cell morphogenesis	783	53	41.46	0.03047	*SB/SB*^{M} vs *SB/SB*^{P}	Brain
GO:0006810	transport	830	53	43.95	0.04209	*SB/SB*^{M} vs *SB/SB*^{P}	Brain
GO:0044281	small molecule metabolic process	563	39	29.81	0.04368	*SB/SB*^{M} vs *SB/SB*^{P}	Brain
GO:0006629	lipid metabolic process	162	14	8.58	0.04687	*SB/SB*^{M} vs *SB/SB*^{P}	Brain
GO:0021700	developmental maturation	236	17	10.07	0.023	*SB/SB*^{P} vs *Sb/Sb*^{P}	Brain
GO:0048856	anatomical structure development	2827	37	29.05	0.0066	*SB/SB*^{M} vs *SB/SB*^{P}	Ovary
GO:0051604	protein maturation	41	3	0.42	0.0083	*SB/SB*^{M} vs *SB/SB*^{P}	Ovary
GO:0034655	nucleobase-containing compound catabolic process	46	3	0.47	0.0114	*SB/SB*^{M} vs *SB/SB*^{P}	Ovary
GO:0008219	cell death	345	8	3.54	0.0231	*SB/SB*^{M} vs *SB/SB*^{P}	Ovary
GO:0061024	membrane organization	154	5	1.58	0.0234	*SB/SB*^{M} vs *SB/SB*^{P}	Ovary
GO:0006412	translation	283	7	2.91	0.0244	*SB/SB*^{M} vs *SB/SB*^{P}	Ovary
GO:0040011	locomotion	656	35	27.74	0.028	*SB/SB*^{P} vs *Sb/Sb*^{P}	Ovary

indicate a possible link with the weight differences that are observed between the different genotypes and their social colony of origin. For example, *SB/SB*^{P} alates have slightly less weight reserves (which are shown to have a correlation with reproduction) than *SB/SB*^{M} alates (DeHeer, 2002; DeHeer et al., 1999).

In brain tissue, as in the ovaries, it appears that social chromosome also has a large effect on gene expression. Comparisons involving *SB/SB* alate gynes (from both monogyne and polygyne colonies) and either *SB/Sb*^{P} or *Sb/Sb*^{P} samples generally result in a relatively large number of DEGs, whereas *SB/Sb*^{P} compared to *Sb/Sb*^{P} shows the smallest number (21 DEGs). Interestingly, comparing *SB/SB*^{M} vs *SB/SB*^{P} to *SB/SB*^{P} vs. *Sb/Sb*^{P} does not show much difference in the amount of DEGs (476 vs 446), while *SB/SB*^{P} vs *SB/Sb*^{P} does not yield a lot of DEGs (54). This is interesting because *SB/SB*^{P} alate gynes face challenges of survival within a polygyne colony: workers will execute them (Tschinkel, 2006; Vargo & Porter, 1989). In addition, in the brain, one of the largest differences in gene expression is seen when comparing the same genotype from different colony social environments (*SB/SB*^{M} to *SB/SB*^{P}; 476 DEGs). This suggests that social environment has an effect on gene expression within the brain tissue. It is known that polygyne workers will only tolerate reproductives with a copy of *Sb* and will execute reproductives without it (Huang & Wang, 2014; Ross & Keller, 1998). Perhaps the stress of workers executing *SB/SB*^{P} alate gynes affects genes involved in the expression of stress response; however, it will require further investigation to recognize whether or not this idea holds any validity.

GO term enrichment analysis revealed a broad overview of genes involved in biological processes. Two GO terms were shown to be significant for *SB/SB*^{P} vs

Sb/Sb^{P}, a term involved in developmental maturation present in brain tissue and a term involved in locomotion present in ovary tissue. It is important to keep in mind that GO term enrichment analysis used the GO-Slim function and topGO — which give a broad look at the ontology and may result in a relatively low number of returned GO terms; however, this method does eliminate redundant GO terms and it useful for initial GO analysis (Consortium, 2004). Interestingly, in brain tissue, 17 genes that are differentially expressed between *SB/SB*^{P} and *Sb/Sb*^{P} are involved in a developmental process, independent of morphogenetic change, that is required for an anatomical structure, cell or cellular component in order to work correctly as expected (Consortium, 2015). In ovary tissue, 35 genes that are differentially expressed between *SB/SB*^{P} and *Sb/Sb*^{P} are involved in self-propelled movement of a cell or organism from one location to another (Consortium, 2015). Further exploring gene ontology (GO) will help create a more delated list of genes and gene products, allowing for a better understanding of what genes are performing what roles (Harris et al., 2008). Only a subset of biological processes that were slimmed down using the GO-Slim function was investigated; cellular component and molecular function were not explored.

Cross comparing data to another computational pipeline will also reaffirm the methods used to obtain differential gene expression results. Using the Cufflinks pipeline to assemble transcripts and quantify transcripts to reanalyze the results will provide more support to our results. Cufflinks is a widely accepted and used tool for RNA-seq analysis (Trapnell et al., 2012).

Understanding how transcriptomes of organisms relate to their phenotypic differences, and ultimately how supergenes contribute to complex phenotypes requires

genomic tools and resources. Through RNA-seq analysis, identifying differences in gene expression in two specific tissues, brains and ovaries, of *S. invicta* may provide insight into how transcriptomes of *S. invicta* alate gynes relate to their social environment and/or genotype. Our study shows that both genotype and social environment have an effect on gene expression, however, it varies according to tissue type. For example, genotype may be more important for overall gene expression patterns within the ovaries. While in brain tissue, the effect is varied. Therefore, it appears that genotype and social environment both affect gene expression in brains and ovaries, but the specific ways in which they do varies across the tissue type. Furthermore, analysis of these RNA-seq data investigating alternative splicing social chromosome variants may lead to a better understanding of how genetic regulators affect complex phenotypes, because particular exons may be included in mRNAs in some tissues (brains vs. ovaries), while omitted in others (Black, 2003).

REFERENCES

Abdullah-Sayani, A., Bueno-de-Mesquita, J. M., & Van De Vijver, M. J. (2006). Technology Insight: tuning into the genetic orchestra using microarrays—limitations of DNA microarrays in clinical practice. *Nature Clinical Practice Oncology, 3*(9), 501-516.

Adiconis, X., Borges-Rivera, D., Satija, R., DeLuca, D. S., Busby, M. A., Berlin, A. M., . . . Fennell, T. (2013). Comparative analysis of RNA sequencing methods for degraded or low-input samples. *Nature Methods, 10*(7), 623-629.

Alexa, A., & Rahnenfuhrer, J. (2010). topGO: enrichment analysis for gene ontology. *R package version, 2*(0).

Ascunce, M. S., Yang, C.-C., Oakey, J., Calcaterra, L., Wu, W.-J., Shih, C.-J., . . . Shoemaker, D. (2011). Global invasion history of the fire ant Solenopsis invicta. *science, 331*(6020), 1066-1068.

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., . . . Eppig, J. T. (2000). Gene Ontology: tool for the unification of biology. *Nature Genetics, 25*(1), 25-29.

Balas, M. T., & Adams, E. S. (1996). The dissolution of cooperative groups: mechanisms of queen mortality in incipient fire ant colonies. *Behavioral Ecology and Sociobiology, 38*(6), 391-399.

Balsa-Canto, E., Henriques, D., Gábor, A., & Banga, J. R. (2016). AMIGO2, a toolbox for dynamic modeling, optimization and control in systems biology. *Bioinformatics, 32*(21), 3357-3359.

Bendena, W., Garside, C., Yu, C., & Tobe, S. (1997). Allatostatins: diversity in structure and function of an insect neuropeptide family. *Annals of the New York Academy of Sciences, 814*(1), 53-66.

Berg, P. R., Star, B., Pampoulie, C., Sodeland, M., Barth, J. M., Knutsen, H., . . . Jentoft, S. (2016). Three chromosomal rearrangements promote genomic divergence between migratory and stationary ecotypes of Atlantic cod. *Scientific reports, 6*.

Bernascon, G., Krieger, M. J., & Keller, L. (1997). Unequal partitioning of reproduction and investment between cooperating queens in the fire ant, Solenopsis invicta, as revealed by microsatellites. *Proceedings of the Royal Society of London B: Biological Sciences, 264*(1386), 1331-1336.

Black, D. L. (2003). Mechanisms of alternative pre-messenger RNA splicing. *Annual Review of Biochemistry, 72*(1), 291-336.

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics, 30*(15), 2114-2120.

Boulay, R., Hooper-Bui, L. M., & Woodring, J. (2001). Oviposition and oogenesis in virgin fire ant females Solenopsis invicta are associated with a high level of dopamine in the brain. *Physiological Entomology, 26*(4), 294-299.

Bourke, A. F., & Mank, J. E. (2013). Genetics: A social rearrangement. *Nature, 493*(7434), 612-613.

Bumgarner, R. (2013). Overview of DNA microarrays: types, applications, and their future. *Current protocols in molecular biology*, 22.21. 21-22.21. 11.

Callcott, A.-M. A., & Collins, H. L. (1996). Invasion and range expansion of imported fire ants (Hymenoptera: Formicidae) in North America from 1918-1995. *Florida Entomologist*, 240-251.

Chen, Y., & McCarthy, D. (2017). edgeR: differential expression analysis of digital gene expression data User's Guide.

Collins, L., & Scheffrahn, R. H. (2001). *Red Imported Fire Ant, Solenopsis invicta Buren (Insecta: Hymenoptera: Formicidae: Myrmicinae)*: University of Florida Cooperative Extension Service, Institute of Food and Agricultural Sciences, EDIS.

Comeron, J. M., Williford, A., & Kliman, R. (2008). The Hill–Robertson effect: evolutionary consequences of weak selection and linkage in finite populations. *Heredity, 100*(1), 19-31.

Conesa, A., Götz, S., García-Gómez, J. M., Terol, J., Talón, M., & Robles, M. (2005). Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics, 21*(18), 3674-3676.

Conesa, A., Madrigal, P., Tarazona, S., Gomez-Cabrero, D., Cervera, A., McPherson, A., . . . Zhang, X. (2016). A survey of best practices for RNA-seq data analysis. *Genome biology, 17*(1), 13.

Consortium, G. O. (2004). The Gene Ontology (GO) database and informatics resource. *Nucleic Acids Research, 32*(suppl 1), D258-D261.

Consortium, G. O. (2015). Gene ontology consortium: going forward. *Nucleic Acids Research, 43*(D1), D1049-D1056.

DeHeer, C. J. (2002). A comparison of the colony-founding potential of queens from single-and multiple-queen colonies of the fire ant *Solenopsis invicta*. *Animal Behaviour, 64*(4), 655-661.

DeHeer, C. J., Goodisman, M. A., & Ross, K. G. (1999). Queen Dispersal Strategies in the Multiple-Queen Form of the Fire Ant *Solenopsis invicta*. *The American Naturalist, 153*(6), 660-675.

Dietrich Gotzek, & Kenneth G. Ross. (2007). Genetic regulation of colony social organization in fire ants: an integrative overview. *The Quarterly Review of Biology, 82*(3), 201-226. doi:10.1086/519965

Fletcher, D. J., & Blum, M. S. (1983). The inhibitory pheromone of queen fire ants: effects of disinhibition on dealation and oviposition by virgin queens. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology, 153*(4), 467-475.

Fletcher, D. J., Cherix, D., & Blum, M. (1983). Some factors influencing dealation by virgin queen fire ants. *Insectes Sociaux, 30*(4), 443-454.

Fritz, G. N., Fritz, A. H., & Vander Meer, R. K. (2011). Sampling high-altitude and stratified mating flights of red imported fire ant. *Journal of Medical Entomology, 48*(3), 508-512.

Fritz, G. N., Vander Meer, R. K., & Preston, C. A. (2006). Selective male mortality in the red imported fire ant, Solenopsis invicta. *Genetics, 173*(1), 207-213.

Gan, G., Ma, C., & Wu, J. (2007). *Data clustering: theory, algorithms, and applications*: SIAM.

Gilmartin, P., & Li, J. (2010). Homing in on heterostyly. *Heredity, 105*(2), 161-162.

Goodisman, M. A., DeHeer, C. J., & Ross, K. G. (2000). Unusual Behavior of Polygyne Fire Ant Queens on Nuptial Flights. *Journal of Insect Behavior, 13*(3), 455-468.

Goodisman, M. A., Mack, P. D., Pearse, D. E., & Ross, K. G. (1999). Effects of a single gene on worker and male body mass in the fire ant Solenopsis invicta (Hymenoptera: Formicidae). *Annals of the Entomological Society of America, 92*(4), 563-570.

Gotzek, D., Robertson, H. M., Wurm, Y., & Shoemaker, D. (2011). Odorant binding proteins of the red imported fire ant, Solenopsis invicta: an example of the problems facing the analysis of widely divergent proteins. *PLoS One, 6*(1), e16289.

Gotzek, D., & Ross, K. G. (2007). Genetic regulation of colony social organization in fire ants: an integrative overview. *The Quarterly Review of Biology, 82*(3), 201-226.

Gramates, L. S., Marygold, S. J., Santos, G. d., Urbano, J.-M., Antonazzo, G., Matthews, B. B., . . . Emmert, D. B. (2017). FlyBase at 25: looking to the future. *Nucleic Acids Research, 45*(D1), D663-D671.

Greenberg, L., Fletcher, D. J., & Vinson, S. B. (1985). Differences in Worker Size and Mound Distribution in Monogynous and Polygynous Colonies of the Fire Ant *Solenopsis invicta* Buren. *Journal of the Kansas Entomological Society*, 9-18.

Harris, M. A., Deegan, J. I., Ireland, A., Lomax, J., Ashburner, M., Tweedie, S., . . . Day-Richter, J. (2008). The gene ontology project in 2008. *Nucleic Acids Research, 36*(SUPPL. 1).

Hölldobler, B., & Wilson, E. O. (1990). *The ants*: Harvard University Press.

Huang, Y. C., & Wang, J. (2014). Did the fire ant supergene evolve selfishly or socially? *Bioessays, 36*(2), 200-208. doi:10.1002/bies.201300103

Hung, A. C., Vinson, S. B., & Summerlin, J. W. (1974). Male sterility in the red imported fire ant, Solenopsis invicta. *Annals of the Entomological Society of America, 67*(6), 909-912.

Jones, R. T., Salazar, P. A., Jiggins, C. D., & Joron, M. (2012). *Evolution of a mimicry supergene from a multilocus architecture.* Paper presented at the Proc. R. Soc. B.

Joron, M., Frezal, L., Jones, R. T., Chamberlain, N. L., Lee, S. F., Haag, C. R., . . . Ferguson, L. (2011). Chromosomal rearrangements maintain a polymorphic supergene controlling butterfly mimicry. *Nature, 477*(7363), 203-206.

Joron, M., Papa, R., Beltrán, M., Chamberlain, N., Mavárez, J., Baxter, S., . . . Rogers, J. (2006). A conserved supergene locus controls colour pattern diversity in Heliconius butterflies. *PLoS Biology, 4*(10), e303.

Keller, L., & Ross, K. G. (1993). Phenotypic basis of reproductive success in a social insect: genetic and social determinants. *SCIENCE-NEW YORK THEN WASHINGTON-, 260*, 1107-1107.

Keller, L., & Ross, K. G. (1998). Selfish genes: a green beard in the red fire ant. *Nature, 394*(6693), 573-575.

Keller, L., & Ross, K. G. (1999). Major gene effects on phenotype and fitness: the relative roles of *Pgm-3* and *Gp-9* in introduced populations of the fire ant *Solenopsis invicta*. *Journal of Evolutionary Biology, 12*(4), 672-680.

Kent, W. J. (2002). BLAT—the BLAST-like alignment tool. *Genome Research, 12*(4), 656-664.

Kirubakaran, T. G., Grove, H., Kent, M. P., Sandve, S. R., Baranski, M., Nome, T., . . . Otterå, H. (2016). Two adjacent inversions maintain genomic differentiation between migratory and stationary ecotypes of Atlantic cod. *Molecular Ecology*.

Knapton, R., & Falls, J. (1983). Differences in parental contribution among pair types in the polymorphic white-throated sparrow. *Canadian Journal of Zoology, 61*(6), 1288-1292.

Kolde, R., & Kolde, M. R. (2015). Package 'pheatmap'.

Kukurba, K. R., & Montgomery, S. B. (2015). RNA sequencing and analysis. *Cold Spring Harbor Protocols, 2015*(11), pdb. top084970.

Kunte, K., Zhang, W., Tenger-Trolander, A., Palmer, D., Martin, A., Reed, R., . . . Kronforst, M. (2014). Doublesex is a mimicry supergene. *Nature, 507*(7491), 229-232.

Küpper, C., Stocks, M., Risse, J. E., dos Remedios, N., Farrell, L. L., McRae, S. B., . . . Verkuil, Y. I. (2016). A supergene determines highly divergent male reproductive morphs in the ruff. *Nature Genetics, 48*(1), 79-83.

Lamichhaney, S., Fan, G., Widemo, F., Gunnarsson, U., Thalmann, D. S., Hoeppner, M. P., . . . Zhang, H. (2015). Structural genomic changes underlie alternative reproductive strategies in the ruff (Philomachus pugnax). *Nature Genetics*.

Lawson, L. P., Vander Meer, R. K., & Shoemaker, D. (2012). Male reproductive fitness and queen polyandry are linked to variation in the supergene *Gp-9* in the fire ant *Solenopsis invicta*. *Proc Biol Sci, 279*(1741), 3217-3222. doi:10.1098/rspb.2012.0315

Le Poul, Y., Whibley, A., Chouteau, M., Prunier, F., Llaurens, V., & Joron, M. (2014). Evolution of dominance mechanisms at a butterfly mimicry supergene. *Nature communications, 5*.

Li, X., Nair, A., Wang, S., & Wang, L. (2014). Quality control of RNA-seq experiments. *Methods in molecular biology (Clifton, NJ), 1269*, 137-146.

Lucas, C., Nicolas, M., & Keller, L. (2015). Expression of *foraging* and *Gp-9* are associated with social organization in the fire ant *Solenopsis invicta*. *Insect Molecular Biology, 24*(1), 93-104.

Lyon, M. F. (2003). Transmission ratio distortion in mice. *Annual Review of Genetics, 37*(1), 393-408.

Manfredini, F., Riba-Grognuz, O., Wurm, Y., Keller, L., Shoemaker, D., & Grozinger, C. M. (2013). Sociogenomics of Cooperation and Conflict During Colony Founding in the Fire Ant *Solenopsis invicta*. *PLoS Genetics, 9*(8), e1003633. doi:10.1371/journal.pgen.1003633

Morel, L., Vander Meer, R. K., & Lofgren, C. S. (1990). Comparison of nestmate recognition between monogyne and polygyne populations of Solenopsis invicta (Hymenoptera: Formicidae). *Annals of the Entomological Society of America, 83*(3), 642-647.

Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L., & Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature Methods, 5*(7), 621-628.

Nipitwattanaphon, Wang, J., Dijkstra, M. B., & Keller, a. L. (2013). A simple genetic basis for complex social behaviour mediates widespread gene expression differences *Molecular Ecology, 22*, 3797-3813.

Nipitwattanaphon, M., Wang, J., Ross, K. G., Riba-Grognuz, O., Wurm, Y., Khurewathanakul, C., & Keller, L. (2014). Effects of ploidy and sex-locus genotype on gene expression patterns in the fire ant Solenopsis invicta. *Proceedings of the Royal Society B-Biological Sciences, 281*(1797), 8. doi:10.1098/rspb.2014.1776

Peeters, C. (1991). Ergatoid queens and intercastes in ants: two distinct adult forms which look morphologically intermediate between workers and winged queens. *Insectes Sociaux, 38*(1), 1-15.

Picelli, S., Björklund, Å. K., Faridani, O. R., Sagasser, S., Winberg, G., & Sandberg, R. (2013). Smart-seq2 for sensitive full-length transcriptome profiling in single cells. *Nature Methods, 10*(11), 1096-1098.

Picelli, S., Faridani, O. R., Björklund, Å. K., Winberg, G., Sagasser, S., & Sandberg, R. (2014). Full-length RNA-seq from single cells using Smart-seq2. *Nature protocols, 9*(1), 171-181.

Pracana, R., Priyam, A., Levantis, I., Nichols, R. A., & Wurm, Y. (2017). The fire ant social chromosome supergene variant Sb shows low diversity but high divergence from SB. *Molecular Ecology*.

Roberts, R. B., Ser, J. R., & Kocher, T. D. (2009). Sexual conflict resolved by invasion of a novel sex determiner in Lake Malawi cichlid fishes. *science, 326*(5955), 998-1001.

Robinson, M. D., McCarthy, D. J., & Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics, 26*(1), 139-140.

Ross, K. G. (1992). Strong selection on a gene that influences reproductive competition in a social insect. *Nature, 355*(6358), 347.

Ross, K. G. (1997). Multilocus evolution in fire ants: effects of selection, gene flow and recombination. *Genetics, 145*(4), 961-974.

Ross, K. G., & Fletcher, D. J. (1985). Genetic origin of male diploidy in the fire ant, Solenopsis invicta (Hymenoptera: Formicidae), and its evolutionary significance. *Evolution*, 888-903.

Ross, K. G., & Keller, L. (1995). Ecology and evolution of social organization: insights from fire ants and other highly eusocial insects. *Annual Review of Ecology and Systematics, 26*(1), 631-656.

Ross, K. G., & Keller, L. (1998). Genetic control of social organization in an ant. *Proceedings of the National Academy of Sciences, 95*(24), 14232-14237. doi:10.1073/pnas.95.24.14232

Ross, K. G., & Shoemaker, D. D. (1997). Nuclear and mitochondrial genetic structure in two social forms of the fire ant *Solenopsis invicta*: insights into transitions to an alternate social organization. *Heredity, 78*, 590-602.

Ross, K. G., Vargo, E. L., & Keller, L. (1996). Social evolution in a new environment: the case of introduced fire ants. *Proceedings of the National Academy of Sciences, 93*(7), 3021-3025.

Schroeder, A., Mueller, O., Stocker, S., Salowsky, R., Leiber, M., Gassmann, M., . . . Ragg, T. (2006). The RIN: an RNA integrity number for assigning integrity values to RNA measurements. *BMC Molecular Biology, 7*(1), 3.

Schwander, T., Libbrecht, R., & Keller, L. (2014). Supergenes and complex phenotypes. *Current Biology, 24*(7), R288-R294.

Shen, W., Le, S., Li, Y., & Hu, F. (2016). SeqKit: A Cross-Platform and Ultrafast Toolkit for FASTA/Q File Manipulation. *PLoS One, 11*(10), e0163962.

Thompson, M., & Jiggins, C. (2014). Supergenes and their role in evolution. *Heredity, 113*(1), 1-8.

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D. R., . . . Pachter, L. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature protocols, 7*(3), 562-578.

Tschinkel, W. R. (1995). Stimulation of fire ant queen fecundity by a highly specific brood stage. *Annals of the Entomological Society of America, 88*(6), 876-882.

Tschinkel, W. R. (2006). *The fire ants*: Harvard University Press.

Tschinkel, W. R., & Howard, D. F. (1983). Colony founding by pleometrosis in the fire ant, Solenopsis invicta. *Behavioral Ecology and Sociobiology, 12*(2), 103-113.

Tuttle, E. M. (2003). Alternative reproductive strategies in the white-throated sparrow: behavioral and genetic evidence. *Behavioral Ecology, 14*(3), 425-432.

Tuttle, E. M., Bergland, A. O., Korody, M. L., Brewer, M. S., Newhouse, D. J., Minx, P., . . . Warren, W. C. (2016). Divergence and functional degradation of a sex chromosome-like supergene. *Current Biology, 26*(3), 344-350.

Valles, S., & Porter, S. D. (2003). Identification of polygyne and monogyne fire ant colonies (Solenopsis invicta) by multiplex PCR of Gp-9 alleles. *Insectes Sociaux, 50*(2), 199-200.

Vander Meer, R. K. V., Morel, L., & Lofgren, C. S. (1992). A comparison of queen oviposition rates from monogyne and polygyne fire ant, *Solenopsis invicta*, colonies. *Physiological Entomology, 17*(4), 384-390.

Vargo, E. L. (1999). Reproductive development and ontogeny of queen pheromone production in the fire ant Solenopsis invicta. *Physiological Entomology, 24*(4), 370-376.

Vargo, E. L., & Fletcher, D. J. (1987). Effect of queen number on the production of sexuals in natural populations of the fire ant, Solenopsis invicta. *Physiological Entomology, 12*(1), 109-116.

Vargo, E. L., & Laurel, M. (1994). Studies on the mode of action of a queen primer pheromone of the fire ant Solenopsis invicta. *Journal of Insect Physiology, 40*(7), 601-610.

Vargo, E. L., & Porter, S. D. (1989). Colony reproduction by budding in the polygyne form of Solenopsis invicta (Hymenoptera: Formicidae). *Annals of the Entomological Society of America, 82*(3), 307-313.

Vargo, E. L., & Porter, S. D. (1993). Reproduction by virgin queen fire ants in queenless colonies: Comparative study of three taxa (Solenopsis richteri, hybridS.

invicta/richteri, S. geminata)(Hymenoptera: Formicidae). *Insectes Sociaux, 40*(3), 283-293.

Wang, J., Jemielity, S., Uva, P., Wurm, Y., Gräff, J., & Keller, L. (2007). An annotated cDNA library and microarray for large-scale gene-expression studies in the ant Solenopsis invicta. *Genome biology, 8*(1), R9.

Wang, J., Ross, K. G., & Keller, L. (2008). Genome-wide expression patterns and the genetic architecture of a fundamental social trait. *PLoS Genetics, 4*(7), e1000127. doi:10.1371/journal.pgen.1000127

Wang, J., Wurm, Y., Nipitwattanaphon, M., Riba-Grognuz, O., Huang, Y.-C., Shoemaker, D., & Keller, L. (2013). A Y-like social chromosome causes alternative colony organization in fire ants. *Nature, 493*(7434), 664-668. doi:http://www.nature.com/nature/journal/v493/n7434/abs/nature11832.html - supplementary-information

Wheeler, D. (1996). The role of nourishment in oogenesis. *Annual Review of Entomology, 41*(1), 407-431.

Winnebeck, E. C., Millar, C. D., & Warman, G. R. (2010). Why does insect RNA look degraded? *Journal of Insect Science, 10*(159), 1-7.

Wurm, Y., Wang, J., & Keller, L. (2010). Changes in reproductive roles are associated with changes in gene expression in fire ant queens. *Molecular Ecology, 19*(6), 1200-1211. doi:10.1111/j.1365-294X.2010.04561.x

Wurm, Y., Wang, J., Riba-Grognuz, O., Corona, M., Nygaard, S., Hunt, B. G., . . . Gotzek, D. (2011). The genome of the fire ant Solenopsis invicta. *Proceedings of the National Academy of Sciences, 108*(14), 5679-5684.

Zhang, W., Wanchoo, A., Ortiz-Urquiza, A., Xia, Y., & Keyhani, N. O. (2016). Tissue, developmental, and caste-specific expression of odorant binding proteins in a eusocial insect, the red imported fire ant, Solenopsis invicta. *Scientific reports, 6*.