Simple inheritance, complex regulation: supergene-mediated fire ant queen polymorphism

Samuel Arsenault¹, Joanie King², Sasha Kay¹, Kip Lacy³, Kenneth Ross⁴, and Brendan Hunt¹

¹University of Georgia ²Texas A and M University College Station ³Rockefeller University ⁴University of Georgia College of Agricultural and Environmental Sciences

May 8, 2020

Abstract

The fire ant Solenopsis invicta exists in two alternate social forms: monogyne nests contain a single reproductive queen and polygyne nests contain multiple reproductive queens. This colony-level social polymorphism corresponds with individual differences in queen physiology, queen dispersal patterns, and worker discrimination behaviors, all evidently regulated by an inversion-based supergene that spans more than 13Mb of a "social chromosome," contains over 400 protein-coding genes, and rarely undergoes recombination. The specific mechanisms by which this supergene influences expression of the many distinctive features that characterize the alternate forms remain almost wholly unknown. To advance our understanding of these mechanisms, we explore effects of social chromosome genotype and natal colony social form on gene expression in virgin queens sampled as they embarked on nuptial flights, using RNA-sequencing of two important tissues. We observe relatively minor effects of natal social form, that is, of the social/developmental environment, on gene expression profiles, but substantial effects of genotype, including i) supergene-associated gene upregulation, ii) allele-specific expression, and iii) pronounced extra-supergene trans-regulatory effects. These findings, along with observed spatial variation in differential and allele-specific expression within the supergene region, highlight the complex gene regulatory landscape that emerged following evolutionary divergence of the inversion-mediated Sb haplotype from its homolog that largely retained the ancestral gene order. The distinctive social chromosome-linked gene expression trajectories we document at the onset of a queen's reproductive life expand the known record of relevant molecular correlates of a complex social polymorphism and point to putative genetic underpinnings of the alternate social syndromes.

Introduction

Supergenes are regions of the genome that contain multiple genes held in strong linkage disequilibrium due to severely restricted recombination (Schwander, Libbrecht, & Keller, 2014). They are often formed by chromosomal rearrangements, such as inversions, which inhibit crossing over during meiosis, thereby facilitating the preservation of favorable allelic combinations across numerous loci (Faria, Johannesson, Butlin, & Westram, 2019; Schwander et al., 2014). Chromosomal rearrangements are fundamental to sex chromosome evolution (Hoffmann & Rieseberg, 2008; White, Kitano, & Peichel, 2015) and are now recognized as taxonomically widespread, evolutionarily significant mediators of complex trait variation (Faria et al., 2019; Schwander et al., 2014; Wellenreuther & Bernatchez, 2018). For example, supergenes have been discovered that determine variation in social behavior, sexual dimorphism, and dispersal strategies in animals as diverse as insects (Purcell, Brelsford, Wurm, Perrin, & Chapuisat, 2014; Wang et al., 2013), birds (Küpper et al., 2016; Lamichhaney et al., 2015; Tuttle et al., 2016), and fishes (Kirubakaran et al., 2016; Pearse et al., 2019). Despite their acknowledged role in facilitating the evolution of complex trait polymorphisms, supergenes are not without evolutionary costs. Diminished recombination and resulting linkage disequilibrium within supergenes are expected to impede the ability of selection to act on individual mutations, with the result that spread of beneficial mutations and elimination of deleterious ones are constrained by the genetic background on which they arise (Hill-Robertson interference) (Hill & Robertson, 1966; Wang et al., 2013). The subsequent reduction in the efficacy of natural selection typically culminates in some level of degeneration of the non-recombining region, usually manifested as accumulation of fixed deleterious mutations, loss of genetic diversity, suppression of gene expression, gene deletion, and/or proliferation of transposable elements (Bachtrog et al., 2011; Stolle et al., 2019; Tuttle et al., 2016). Thus, the complex makeup of supergenes as mosaics of tightly linked adaptive, neutral, and even deleterious genetic variants makes dissection and elucidation of their genotype to phenotype map especially challenging.

The fire ant Solenopsis invicta displays naturally occurring variation in fundamental, ecologically important traits comprising two social syndromes (monogyny and polygyny) that collectively are under the control of a supergene (Wang et al., 2013). The eponymous difference in the syndromes is the number of reproductive (egglaying) queens in a colony, with colonies of the monogyne form containing only a single reproductive queen and colonies of the polygyne form containing few to many queens (Gotzek and Ross 2007; Tschinkel 2006). The difference in queen number corresponds with differences in numerous individual-level attributes including: the extent of energy reserves, types of dispersal behaviors, and reproductive ontogenies of young queens (C. DeHeer, Goodisman, & Ross, 1999; C. J. DeHeer, 2002); the fecundity of mature reproductive queens (DeHeer 2002; Tschinkel 2006); and the form of worker discrimination directed toward queens attempting to join a colony (Keller & Ross, 1998). The S. invicta supergene regulating social form arose with the appearance of multiple adjacent chromosomal inversions that span more than 13Mb of DNA on the "social chromosome" and include over 400 mapped protein-coding genes (Y. Huang, Dang, Chang, Wang, & Wang, 2018; Wang et al., 2013; Yan et al., 2020). Recombination is dramatically reduced between the inverted region of the social chromosome (denoted as the Sb haplotype) and its homologs that show synteny with the presumed ancestral haplotype (denoted SB) (Stolle et al., 2019; Wang et al., 2013; Yan et al., 2020). Moreover, Sb/Sbhomokaryotypes (homozygotes) in the invasive U.S. range effectively do not survive to reproduce (Hallar, Krieger, & Ross, 2007); thus, like the mammalian Y-chromosome, recombination is strongly reduced between variants of the Sb haplotype as well as between the Sb and SB haplotypes. The entirety of the fire ant supergene can be characterized as an enormous, complex genomic module that collectively regulates colony social form and allied traits while propagating via simple Mendelian inheritance (Keller & Ross, 1998; Wang et al., 2013).

Despite a recent surge of interest in the importance of supergenes in regulating complex trait variation (Schwander et al., 2014: Wellenreuther & Bernatchez, 2018), progress in disentangling the functional mechanisms by which they exert their phenotypic effects has been slow. As suggested above, supergenes may harbor relatively few genes with direct functional effects on adaptive traits that reside in a sea of elements with neutral or even deleterious fitness consequences, and identification of such loci is complicated by the strong linkage disequilibrium observed across the supergene region. Moreover, supergene-associated traits may be regulated by a combination of protein structural variants and variant DNA regulatory elements, necessitating careful investigation and evaluation of the role of each type, as well as possible joint effects (e.g., allele-specific expression of protein-coding genes). In S. invicta, recent studies have defined the specific inversion breakpoints in the Sb haplotype (Y. Huang et al., 2018; Yan et al., 2020) and have begun to characterize protein sequence divergence between the Sb and SB haplotypes in both the invasive U.S. range (Pracana, Priyam, Levantis, Nichols, & Wurm, 2017) and native South American range (Yan et al., 2020). However, fundamental questions remain as to how differences between the supergene haplotypes, ranging in scope from single nucleotide polymorphisms, to copy number variants, to major structural rearrangements, influence gene regulation and induce alternate complex social phenotypes (Fontana et al., 2019; Gotzek & Ross, 2007; Y. C. Huang & Wang, 2014).

To explore in greater detail the genotype to phenotype relationship between the Sb supergene haplotype and the fire ant polygyne syndrome, we generated and analyzed RNA-seq data from brains and ovaries of individual, sexually mature, similarly-aged unmated queens (gynes) of both social forms of *S. invicta*. Organ-specific sampling was utilized to minimize the effects of variable tissue composition, with brains and ovaries targeted because they are likely the primary organs mediating the behavioral and reproductive differences between the social forms. Samples of gynes with all three social chromosome genotypes (*SB/SB*, *SB/Sb*, and *Sb/Sb*) were included in our experiment (Figure 1), which is noteworthy by virtue of inclusion of the rare gynes from polygyne nests bearing either homozygous genotype. We focus on gene expression effects of the two most conspicuous factors relevant to fire ant social organization: an individual's social chromosome (supergene) genotype and its natal colony social form (developmental environment).

Materials and Methods

Alate queen (gyne) collection. Young *Solenopsis invicta* gynes exhibit distinct rates of sexual and reproductive maturation depending on social chromosome genotype (Nipitwattanaphon, Wang, Dijkstra, & Keller, 2013); these gynes do not attempt to leave their natal nest on mating flights until they are fully sexually mature. We collected gynes embarking on mating flights in the field to achieve physiological matching across genotypes, colonies, and forms based on maturity rather than strict age-matching (Figure 1), although, based on the colony phenologies and weather conditions at the collection locale in spring of 2015 in Athens-Clarke Co., Georgia, USA, it is doubtful that any sampled gynes differed by more than a couple of weeks in age from one another. Alate (winged) gynes were aspirated from the tops of mounds and frozen on dry ice in the field in and around Athens-Clarke Co. on days of mating flights in April 2015. Alate gynes were collected from polygyne nests at three sites (four nests from 33°55'18"N 83°21'2"W, three nests from 33°54'48"N 83°20'46"W, and one nest from 33°54'36"N 83°20'6"W), and from eight monogyne nests at one site (33°42'17"N 83°16'53"W). All samples were stored at -80°C pending further processing. We processed a total of eight gynes from as many monogyne colonies and three gynes each (one of each genotype) from eight polygyne colonies.

Sample genotyping. RNA*later* -ICE (Thermo Fisher Scientific, MA) was added to tubes holding the frozen gynes, which were submerged for >24 hours at -20° C. Gynes from polygyne colonies then were sorted by size in a Petri dish on top of dry ice. The smallest gynes were selected as prospective Sb/Sb, medium as SB/Sb, and the largest as SB/SB, based on prior data on the association of gyne mass with Gp-9 genotype (C. DeHeer et al., 1999). Legs were removed for DNA extraction and genotype confirmation, then bodies were stored at -80°C until dissection and RNA extraction. DNA was extracted using a Puregene DNA isolation kit (Qiagen, Valencia, CA). To obtain the social chromosome genotype of each gyne from presumed polygyne colonies, a Gp-9 PCR assay was performed (Valles & Porter, 2003). The presence of gynes with SB/Sb and Sb/Sb genotypes also confirmed colony social form of polygyne colonies because only polygyne colonies contain individuals with the Sb haplotype (for which the $Gp-9^b$ allele is fully diagnostic). To confirm social form of the monogyne colonies from which gynes were sampled, 15-20 individuals (mix of workers and alate gynes) were collected from each colony and the same Gp-9 PCR assay was performed. All gynes sampled from polygyne colonies were genotyped at 9-13 previously described polymorphic microsatellite loci (Ascunce, Bouwma, & Shoemaker, 2009). These microsatellite data revealed that no sampled individuals were triploid (Krieger, Ross, Chang, & Keller, 1999) and that no sampled nestmates were siblings (Supplementary Data 1).

Organ dissection. Gynes were dissected under an Olympus SZ61 stereomicroscope. After making an opening in the cuticle on the back of the head, a minutin probe was used to free the brain from the head capsule. The majority of tracheae and glands were removed from each brain; however, some traces of tracheae and possibly residual gland material remained. The ovaries were extricated from the hindgut, and any Malpighian tubules present were removed, along with some excess fat body. Due to the extensive tracheal intrusion in the ovaries, removing all traces of tracheae and fat body from the ovaries was not feasible. Tissues were stored at -20°C in RNA*later* -ICE until RNA extraction.

RNA extraction, library preparation, and sequencing. RNA was extracted from brains using the RNeasy Micro Kit and from ovaries using the RNeasy Mini Kit, both with DNase treatment (Qiagen, Valencia, CA). Extracted total RNA integrity and concentration were evaluated for every sample on an Agilent 2100 bioanalyzer (Supplementary Data 1). Illumina sequencing libraries were prepared following the Smart-seq2 protocol

(Picelli et al., 2014), which was developed for low input RNA sequencing applications. Based on bioanalyzer readings, approximately 1.2 ng of total RNA was used to make each brain library and approximately 3.6 ng of total RNA was used to make each ovary library (Supplementary Data 1). Samples were barcoded and pooled prior to sequencing at the Georgia Genomics and Bioinformatics Core (Athens, GA) on an Illumina NextSeq instrument to produce 75 bp single-end reads. We sequenced all brain samples on one high-output flow cell and all ovaries on a second high-output flow cell.

Quality control and read alignment. Most rRNA presumably was removed during RNA-Seq library preparation by poly-A mRNA isolation in the Smart-seq2 protocol. Nonetheless, any reads in our data that aligned to Myrmecia croslandi (Australian bull ant) 18S rRNA, 5.8S rRNA, or 28S rRNA genes were removed using BLAT (version 3.5). The remaining reads were trimmed with Trimmomatic version 0.32 (TRAILING:3, LEADING:3, SLIDINGWINDOW:4:15, MINLEN:36) (Bolger, Lohse, & Usadel, 2014). Reads were aligned to S. invicta assembly gnG (NCBI accession GCF_000188075.1, generated from an SB male) using STAR version 2.5.3a (Dobin et al., 2013; Wurm et al., 2011). We utilized previously published linkage mapping information to determine which scaffolds mapped to each linkage group as well as to the Sb supergene (Pracana, Priyam, et al., 2017; Wang et al., 2013). One SB/SB polygyne ovary sample (104DO) was removed from further analysis due to excessively low mapping (~0.9% uniquely mapped reads) (Supplementary Data 1). In addition, we resequenced two libraries that had lower than expected numbers of sequenced reads (239AB and 1EB). After alignment, we had between 10 million and 31 million uniquely mapped reads available for each sample (Supplementary Data 1).

Differential Expression Analysis. Aligned reads were mapped to the NCBI annotation release 100 RefSeq gene models using the FeatureCounts function from the RSubread package (Liao, Smyth, & Shi, 2013). Genes were filtered from the analysis if fewer than 8 libraries had a Counts-per-million (CPM) > 10/9 at that locus (more than 10 reads aligned to a gene in the majority of libraries). This decreased the possibility of inferring differential expression from structured background noise. The data were then analyzed using two statistical approaches in edgeR (Robinson, McCarthy, & Smyth, 2009) – a 'multifactorial model' and a 'pairwise model'. The multifactorial-model parameters included colony identity, social form, organ type, social chromosome genotype, and a covariance term for organ and genotype. We used this model to ask broad questions about our data while making use of all of the information from all of our samples. The pairwise comparisons simply treated each sample type (e.g., polygyne SB/SB brain, polygyne SB/Sb brain, etc.) as an independent treatment for analyzing specific differences between two types. For each approach, genes were considered differentially expressed if they exhibited an FDR-corrected P-value < 0.01. Each comparison was visualized using volcano plots with a custom script using the ggplot2 R package (Villanueva & Chen, 2019).

Overlap of significant differences for various comparisons was visualized using UpSetR (Conway, Lex, & Gehlenborg, 2017). Next, candidate gene categories were further refined by classifying genes as 'ovary-specific,' 'brain-specific,' or 'Sb/Sb -specific' if they exhibited an FDR-corrected P-value < 0.01 in the relevant comparisons (e.g., ovary SB/SB versus SB/Sb and ovarySB/SB versus Sb/Sb in the ovary-specific category) while also exhibiting an FDR-corrected P-value > 0.1 in the other categories (e.g., brain SB/SB versus SB/Sb and brainSB/SB versus Sb/Sb in the ovary-specific category). We note that genes that exhibit brain- or ovary-specific expression differences in our study, as well as those that do not, may exhibit differential expression in unsampled tissues. Visualization of the genomic location of differentially expressed genes (DEGs) was performed using KaryoPloteR (Gel & Serra, 2017). Non-random spatial distribution of DEGs, with respect to higher or lower expression of Sb -bearing samples, was assessed using a two-sided Wald-Wolfowitz runs test (Wald & Wolfowitz, 1940) after significant DEGs were converted into binary representation based on the directionality of differential expression.

Allele-specific Expression Analysis. To perform allele-specific expression analysis, we generated a highconfidence set of haplotype-specific SNPs from our RNA-seq data following a previously described protocol (Auwera et al., 2013). In short, we performed a 2-pass alignment of the reads using STAR (Dobin et al., 2013), formatted the reads for GATK using Picard ("Picard toolkit," 2018), then identified variants using the GATK HaplotypeCaller (Auwera et al., 2013). We leveraged our homozygous samples (SB/SB andSb/Sb) to retain only those variants that consistently differentiated SB and Sb haplotypes. We removed Sbvariants that were found as polymorphisms in the SB population as well as those that were found in fewer than three Sb/Sb samples. A total of 3994 SNPs diagnostic for the Sb and SB haplotypes remained following this procedure (Supplementary Data 2). We then n-masked the genome for these variants using Bedtools (Quinlan & Hall, 2010) to eliminate any alignment bias and re-aligned the heterozygous SB/Sb samples to the masked genome using the 2-pass method in STAR (Dobin et al., 2013). The GATK ASEReadCounter function was used to generate a per-SNP count matrix for SB/Sb samples (Auwera et al., 2013), which were mapped to genes using Bedtools intersect. Both the SNP-level and gene-level allele-specific expression counts were next passed to edgeR and fit to a model with variance terms for colony identity, organ, allele, and a covariance term for organ and allele effects (McCarthy, Chen, & Smyth, 2012). Genes and SNPs were considered to display significant allele-specific expression if the FDR-corrected P < 0.01. Visualization of the genomic location of SNPs displaying significant allele-specific expression was performed using KaryoPloteR (Gel & Serra, 2017). The overlap between genes displaying differential expression and allele-specific expression results was visualized using UpSetR (Conway et al., 2017).

Results

Effects of social chromosome genotype on gene expression. To analyze the effect of the supergene-bearing social chromosome's genotype on gene expression patterns in *S. invicta*gynes, we first utilized a generalized linear model accommodating variance due to the experimental factors (colony identity, natal social form, tissue type, genotype, and covariation between tissue type and genotype) in our RNA-seq data (multifactorial-model). We found that the different social chromosome genotypes are associated with many highly significant differences in gene expression independent of tissue type, and that presence of the inversion-based *Sb* haplotype leads to a relative increase in gene expression at most of the significantly affected loci (FDR < 0.01; Figures 2A,B and S1). This pattern is particularly prominent in the *SB/SB* versus *SB/Sb* comparison, where 96% of the differentially expressed genes (DEGs) are upregulated in heterozygous gynes relative to homozygous *SB/SB*gynes (Figures 2A and S1). It is less pronounced in the *SB/SB* versus *Sb/Sb* comparison (72% of DEGs are upregulated in*Sb/Sb* individuals), although there are more than double the number of DEGs in this comparison (as expected given the greater difference in *Sb* copy number in the latter versus the former comparisor; Figures 2B and S1). These data are consistent with a scenario in which the *Sb* haplotype drives widespread upregulation of gene expression, but such a mechanism would appear to be moderated in the absence of a copy of the *SB* haplotype.

We next tested for organ- and genotype-specific effects of the supergene using pairwise comparisons exclusively between polygyne-reared individuals. We observed identical numbers of significant DEGs in the brains and ovaries in comparisons of SB/SB versus SB/Sbgynes (42 DEGs in each; FDR < 0.01; Figures 2C and S1). More than double this number was observed in both organs in the SB/SB versus Sb/Sb comparisons, again as expected given the greater difference in Sb copy number in the latter comparison, with the highest number of such DEGs occurring in the ovaries (Figures 2C and S1). Minimal functional enrichment for GO terms was observed in any of the comparisons, but this could be due in part to the lack of functional information for many genes in the *S. invicta* genome (Supplementary Data 3).

We observed a strong bias in the localization of DEGs to the supergene region in the comparisons of different social chromosome genotypes (Figure 2D). More than a quarter of these DEGs occur within the supergene region despite this region accounting for only 4% of mapped gene space in our assembly (X²-test; P < 0.01). Moreover, we observed distinct patterns of allelic dominance with respect to social chromosome genotype for DEGs within the supergene region as compared to the rest of the genome (Figure S3, Watson's Two-Sample Test of Homogeneity, P < 0.05). These results suggest that *cis* -regulatory divergence is an important outcome of sequence divergence between the Sb and SB supergene haplotypes. However, we also observed widespread *trans* -acting effects of Sb presence, with between 49% and 78% of DEGs attributable to social chromosome genotype located outside of the supergene region (Figures 2D and S2). This is note-worthy because, while the Sb and SB supergene regions rarely undergo recombination and consequently bear many fixed genetic differences, the remainder of the genome recombines freely and shows minimal sequence

differentiation between males bearing or lacking Sb (Yan et al., 2020).

Effects of natal social form on gene expression. Monogyne and polygyne colonies of S. invicta differ in many ways other than colony queen number, for example adult worker to brood ratios, dispersion and connectivity of nests, and population densities (Tschinkel 2006). The cuticular semiochemical profiles and reproductive capacity of queens differ between the two forms, presumably driven largely by their social chromosome genotypes (C. J. DeHeer, 2002; Eliyahu, Ross, Haight, Keller, & Liebig, 2011), yet indirect genetic effects stemming from the different supergene compositions of the colony worker force in the two forms also potentially affect the phenotypes of developing gyne offspring, independently of gyne genotype. For example, substantial effects of the social developmental environment on gene expression were previously observed in workers of S. invictacolonies (Wang, Ross, & Keller, 2008). Thus, we expected to see differences in gene expression between gynes reared in these very different types of societies even if they possess the same social chromosome genotype. Surprisingly, we found that no genes were significantly differentially expressed solely in response to social form of origin when assessed across tissues by a multifactorial-model (Figure 3A). Similarly, pairwise comparisons of specific tissues from SB/SB gynes from different natal colony types (monogyne or polygyne) revealed only 22 social form-dependent DEGs (FDR < 0.01), and these were confined to brains (Figure 3B). However, we did observe an apparent synergistic effect between natal colony social form and social chromosome genotype on differential expression in the ovaries. We observed 43 DEGs between ovaries of SB/SB and SB/Sb individuals when both were reared in polygyne colonies but this increased to 205 DEGs when polygyne-reared SB/Sb individuals were compared to monogyne-reared SB/SB individuals (Figure 3C).

Allele-specific expression in the supergene region. The high proportion of genes upregulated in SB/Sb heterozygotes relative to SB/SB homozygotes that occur within the supergene region can be explained by two potential mechanisms operating within SB/Sb individuals: 1) elevated expression of the Sballeles relative to alleles from the SB homologous region, or 2) a relatively balanced increase in the expression of both the Sband SB alleles. Elevated expression of Sb alleles would be consistent with differences in SB and Sb chromatin structure, mutations in supergene regulatory elements that disproportionately influence Sb -linked alleles in cis, or the presence of Sb-specific paralogs (Fontana et al., 2019). A balanced increase in SB and Sb alleles would be consistent with Sb -induced activation of regulatory elements that can interact with both alleles of a gene, or with Sb -induced chromatin remodeling that extends to both alleles. In order to distinguish between these mechanisms, we evaluated allele-specific expression (ASE) in heterozygotes at genes within the supergene region by means of diagnostic SNPs (SB versus Sb, derived from our RNA-seq data). Our analysis identified 28 genes (of the 222 with sufficient diagnostic SNP coverage) exhibiting significant ASE (FDR < 0.01, Figure 4A, Supplementary Data 2). Notably, we observed similar frequencies of significant ASE for the SB - and Sb -linked alleles, as reported previously in S. invictausing a more limited experimental design (Wang et al., 2013). However, of the nine genes that exhibit significantly elevated expression in SB/Sbrelative to SB/SB gynes, six also exhibit significant Sb ASE (Figure 4B), providing support for the importance of elevated expression of the Sb alleles relative to alleles from the SB homologous region. The other three of these genes exhibit differential expression without significant ASE (Figure 4B), serving as examples of relatively balanced upregulation of SB - and Sb -linked alleles in heterozygotes.

Our analysis of ASE within the supergene region of the genome allowed us to investigate putative SB dosage compensation in heterozygous individuals. Genes subject to such dosage compensation are expected to lack significant differential expression between SB/SB and SB/Sb gynes (FDR > 0.01) but to exhibit significant SBallele-specific overexpression (P < 0.01). We identified twelve such genes (Supplementary Data 2), raising the prospect that SB dosage compensation operates at these loci, possibly due to diminished functionality of the homologous Sbvariants.

Heterogeneity of differential expression in the supergene region. In order to gain insight into the regulatory topology of the supergene region, we assessed whether DEGs and genes with ASE are distributed uniformly throughout the supergene region or located in discernible clusters. Plots of gene expression pattern by social chromosome position qualitatively illustrate that genes exhibiting either type of differential expression are

widespread yet spatially clustered according to effect within the supergene region (Figure 5, S5, Supplementary Data 2). To statistically assess whether DEGs are randomly distributed throughout the region, we employed Wald-Wolfowitz runs tests (Wald & Wolfowitz, 1940) on binary representations of significant DEGs, specifying whether gene expression is elevated in the presence or absence of the *Sb*haplotype and ordering DEGs based on relative linear position in the supergene region (Supplementary Data 2). For DEGs between *SB/SB* and *SB/Sb* gynes, only one *SB/SB*- upregulated gene maps to the supergene and, unsurprisingly, the test is nonsignificant (P > 0.05). In contrast, among DEGs between *SB/SB* and *Sb/Sb* individuals, we found that genes with significantly elevated expression in the presence or absence of the *Sb*haplotype, respectively, are not distributed randomly throughout the supergene region (Wald-Wolfowitz runs test, P = 0.019). This heterogeneous pattern persists when the analyses are repeated using a recently published, improved genome assembly (Figure S5, Supplementary Data 4) (Yan et al., 2020).

In order to illustrate the contrasting patterns of gene regulation observed within the fire ant supergene, we assigned genes to two exemplary regions. The first contains loci that exhibit a high degree of SB allele-specific expression coupled with SB/SB -elevated expression in the SB/SB versus Sb/Sb comparisons (region 1); the second is characterized by Sb allele-specific expression coupled with elevated expression in heterozygotes and Sb homozygotes relative to SB homozygotes (region 2). The appearance of these distinct regions shows that supergene haplotype divergence has generated complex, non-uniform changes in the transcriptional landscape.

Candidate genes associated with polygyny. To glean functional insights into the molecular regulation of colony social form in S. invicta, we grouped strongly supported DEGs by patterns of differential expression into four categories: 1) those with expression differences between polygyne-reared SB/SB and SB/Sb gynes that were **brain-specific**, 2) those with differences between gynes of these same classes that were **ovary-specific** , 3) those with differences between SB/SB and Sb/Sb gynes but not SB/SB and SB/Sb gynes, independent of natal social form or organ type (Sb/Sb-specific), and 4) those with differences in all comparisons of Sb- supergene presence/absence, independent of natal social form or organ type (constitutive). We hypothesized that brain-specific DEGs would exist because of the dramatically different dispersal, mating, and colony-founding behaviors of gynes bearing or lacking the supergene. Similarly, we hypothesized that ovary-specific DEGs would exist because of the dramatic differences in rates of ovary development and associated physiological processes (e.g., fat and storage protein deposition) of young queens of different supergene status. Given the effective lethality of the Sb/Sb genotype in gynes (Hallar et al., 2007), we expected that genes differentially expressed solely in the SB/SB-Sb/Sb comparisons might give some indications as to the molecular causes of the recessive lethal effect. Finally, the genes always differentially expressed when comparing individuals with and without the Sb haplotype, given their ubiquitous differential expression, are prime candidates for drivers of the multifaceted Sb- mediated polygyne phenotype. A list of genes in each category is presented in Table 1 (a complete description of expression statistics for each gene and a table of genes that appeared as DEGs in multiple analyses but do not fit into one of our four candidate gene categories is provided in Supplementary Data 2).

Discussion

Our experimental design leverages samples from two organs and the full breadth of social chromosome genotypes and social forms to provide a high-resolution snapshot of the transcriptomic effects of these factors in fire ants, with the aim of exploring the direct and indirect genetic effects of the supergene. Our focal samples, young queens (gynes) embarking on nuptial flights, arguably are the most relevant caste/life stage on which to conduct such a study. These potential future reproductives were collected while engaged in a pivotal activity that immediately precedes nest founding (monogyne form) or joining (polygyne form), thus marking the initiation of a queen's reproductive life. Fire ant workers are obligately sterile, though essential to colony fitness, while males are short-lived and appear to play no meaningful role in colony social life. Inclusion of SB/SB gynes from both monogyne and polygyne colonies facilitates direct comparison of the gene regulatory consequences of the contrasting developmental environments of the two colony types, while sampling two organs of particular relevance to the alternate social syndromes (brains mediating queen

dispersal/nest-founding behaviors and ovaries enabling reproduction) from individual ants, with appropriate biological replication, allows us to begin linking individual facets of the complex phenotype of polygyny to putative genetic mechanisms. Expression profiling of two organs also provides the opportunity to screen for loci with putative constitutive regulatory effects, that is, genes differentially expressed in multiple cellular contexts and thus possibly mediating body-wide effects of Sb. Finally, inclusion of the rare homozygous Sb/Sb genotype allows exploration of the potential genetic causes of the documented, presumably endogenous effective lethality of this genotype.

The genome-wide perspective. Considering the striking number of individual- and colony-level differences between the monogyne and polygyne forms of S. invicta –e.g., in worker behavior towards queens, queen cuticular chemical profiles and reproductive trajectories, worker gene expression patterns, and many other biological features (Eliyahu et al., 2011; Keller & Ross, 1998; Trible & Ross, 2016; Wang et al., 2008)we predicted that differences in the social environment in which gynes are reared would be reflected in gene expression differences well beyond those that can be explained by social chromosome genotype alone. Although we observed no significantly differentially expressed genes (DEGs) directly attributable to colony social form in our multifactorial-model analysis (FDR < 0.01), polygyne-reared SB/Sb gynes did exhibit many more DEGs in ovaries when compared to monogyne-reared SB/SB gynes (205 DEGs) than when compared to polygyne-reared SB/SB gynes (43 DEGs; Figure 3). This result mirrors the previously observed small, but statistically significant, higher weight of monogyne-reared, as compared to polygyne-reared, SB/SB gynes (all SB/SB gynes, however, are far heavier than polygyne-reared SB/Sb or Sb/Sb gynes) (C. DeHeer et al., 1999). Gyne weight in fire ants reflects amount of storage reserves accumulated during sexual maturation and profoundly affects prospects for successful independent founding (e.g., the ability to produce workers solitarily), versus colony joining, as an individual reproductive strategy (C. J. DeHeer, 2002). Our gene expression results point to an interaction between the effects of natal colony type and supergene-mediated gene regulation that is superimposed on the small effect of natal colony type alone.

We predicted that gene expression differences associated with social chromosome genotype would preferentially involve genes located within the supergene region (Nipitwattanaphon et al., 2013; Wang et al., 2013). The reason is that extensive sequence divergence exists between the Sb and SB haplotypes of the supergene in U.S. populations of S. invicta (estimated at 1.4 SNPs per Kb; (Pracana, Priyam, et al., 2017; Wang et al., 2013)), contrasted with the rest of the genome, which recombines freely and is subject to constant high levels of admixture between the social forms (Ross, Krieger, Keller, & Shoemaker, 2007; Ross, Shoemaker, Krieger, DeHeer, & Keller, 1999; Shoemaker, Deheer, Krieger, & Ross, 2006), mainly via polygyne-derived SB/Sb gynes mating with monogyne-derived SB (haploid) males (Ross & Shoemaker, 2006). Consequently, linkage disequilibrium outside of the supergene region is minimal in both social forms (Ross & Shoemaker, 2018; Yan et al., 2020). Our data expand upon the previous finding that genes differentially expressed in association with social chromosome genotype are disproportionately localized within the supergene region (Nipitwattanaphon et al., 2013; Wang et al., 2013), illustrating the importance of supergene *cis* -regulatory evolution. Nonetheless, many of the genes differentially expressed in association with social chromosome genotype are located elsewhere in the genome, illustrating that *trans* -acting effects, or other epistatic interactions involving the supergene, are also widespread and potentially of major significance in the evolution of the supergene and the alternate social forms in fire ants. This is not surprising given the prevalence of trans-acting effects documented in supergene systems mediating sexual dimorphism (Parsch & Ellegren, 2013; Wijchers & Festenstein, 2011) and, strikingly, even in the relatively small (~100kb) supergene regulating mimicry in *Papilio* polytes butterflies (Kunte et al., 2014).

An important, related point is that the majority of supergene-driven DEGs in our multifactorial-model comparisons show higher expression in individuals carrying a copy of the Sb supergene haplotype, both within and outside of the supergene region (Figures 2, S2). Preliminary indications of preferentially elevated gene expression in the presence of the Sb haplotype have been reported (Wang et al., 2013), but the broad scope of this phenomenon was not established previously. Thus, the observed "degenerative expansion" (Stolle et al., 2019) of the Y-like Sb supergene haplotype is, somewhat paradoxically, accompanied by an increase in transcriptional activity for many Sb alleles at loci within the supergene. This Sb upregulation could

be driven by a combination of cis- regulatory evolution, increased chromatin accessibility, and/or recent Sbspecific gene duplication events that appear to have occurred in conjunction with the expansion of the Sb haplotype (Fontana et al., 2019; Stolle et al., 2019).

Although such gene duplication within the inversion-derived Sbhaplotype likely explains some of the widespread Sb -mediated upregulation observed in our analyses, we think it is unlikely to be the sole driver of this phenomenon for three reasons. First, while Fontana et al. (2019) show a general increase in transcript levels from genes with Sb- specific paralogs, they found no significant correlation between gene expression level and copy number, implying that regulatory element evolution may play a larger role than copy number in observed expression differences between SB and Sb haplotypes in S. *invicta*. Second, we detected several genes in our analysis that exhibited differential expression without significant allele-specific expression, indicating that their increased transcription is facilitated by both the Sb and SBhaplotypes rather than being a simple consequence of possession of extra copies of the Sb -linked variants. Finally, while a number of our candidate genes have undergone gene duplication events, the increase in gene expression we observe in Sb -carrying individuals is significantly higher than one would expect from a simple gene duplication event (or even several), with \log_2 -fold changes greater than 5 in many duplicated genes (Table 1). Ultimately, the Sb -upregulated DEGs we report reflect an increase in the number of identical or highly similar transcripts in Sb- carrying individuals that collectively influence downstream pathways and phenotypes, no matter the specific causal genetic architecture.

Our comparisons of supergene effects on gene expression in the brains and ovaries revealed some key differences in the profiles of the two organs. In particular, ovary DEGs are less concentrated in the supergene than brain DEGs (Figure 2D). This indicates that regulatory effects of the supergene may be more indirect in the ovaries than the brain, involving more diffuse sets of genetic pathways regulated largely by *trans* -acting effects. Moreover, many more DEGs are detected in ovaries than the brain in comparisons of the two homozygous classes (SB/SB versus Sb/Sb; Figure 2), suggesting that the transcriptional consequences of supergene homozygosity are particularly pronounced in the ovarian tissues of Sb/Sb gynes. Because these individuals do not survive to reproduce in U.S. populations, it was not possible to gauge the effects of these expression differences with respect to fecundity; however, maturing Sb/Sb gynes have been shown to suffer curtailed development in terms of other physiological markers associated with onset of oogenesis compared to gynes of the other two genotypes (Hallar et al., 2007).

The supergene perspective. Given that minimal levels of recombination are observed within the Sb haplotype (Ross & Shoemaker, 2018; Wang et al., 2013; Yan et al., 2020), its constituent genes are effectively inherited as a single Mendelian element and are in near-complete linkage disequilibrium. Nonetheless, expression profiles and associated selective pressures are expected to vary among supergene genes in relation to their specific functions. Indeed, we observed spatially heterogeneous gene expression patterns across the supergene region in S. invicta gynes, with genes expressed more highly inSb than SB, as well as genes showing the opposite pattern, clustered together more commonly than expected by chance.

We highlighted two regions of the supergene with these distinctive gene and allele-specific expression patterns (Figure 5,S5). The first, here denoted region 1, exhibits minimal differential expression in the SB/SB-SB/Sb comparison, while showing extensive SB/SB -biased expression in the SB/SB-Sb/Sb comparison. Concurrently, this region shows disproportionate SB -biased allele expression in heterozygous SB/Sb individuals. These results point to diminished expression of the alleles within the Sb haplotype, compensated by means of overexpression of the SB versions of those genes. We hypothesize that genes subject to this SB dosage compensation in region 1 (and elsewhere) contribute to the observed inviability of Sb/Sb queens (Hallar et al., 2007). Dosage compensation is a commonly observed phenomenon in sex chromosome systems (Disteche, 2012; White et al., 2015) and has been documented recently in the supergene of the white-throated sparrow (Sun, Huh, Zinzow-kramer, Maney, & Yi, 2018). A second region of the supergene, here denoted region 2, contains genes exhibiting strong Sb -biased gene expression and Sb allele-specific expression (Figure 5), indicating upregulation driven largely by expression of the Sb -linked alleles. We hypothesize that genes in region 2 (and other genes conforming to this regulatory pattern) act as functional drivers of the polygyne

syndrome and thus are important candidates of major effect on this social polymorphism.

Gene-specific, functional observations. We leveraged our sampling scheme to search for genes exhibiting differential expression patterns that parallel specific phenotypic differences observed between gynes of each social chromosome genotype. Specifically, we classified DEGs into four categories: brain-specific, ovary-specific, Sb/Sb-specific, and constitutive (Table 1). Among eight detected brain-specific DEGs, *Ribonuclease H-like* is notable for having a human ortholog linked to splicing and neurological disfunction (Bin et al., 2016). Additionally, among the twelve detected ovary-specific DEGs, *sphingosine-1-phosphate lyase* is notable for having a *Drosophila* ortholog involved in sphingolipid metabolism that is linked to reproductive deficiencies (Phan et al., 2007). This gene also stands out as the only characterized gene in the ovary-specific category with lower expression in heterozygotes than SB/SB homozygotes.

Among eight detected Sb/Sb- specific DEGs (those with differences in both organs but solely in the Sb/Sb-SB/SB comparison), pre-mRNA-splicing factor Slu7-like and peptide deformylase, mitochondrial-like are notable for their putative impacts on the production of properly spliced and processed proteins. Finally,

among the thirteen detected constitutive DEGs (those with differences in all comparisons of Sb- supergene presence/absence) we identified notable genes with orthologs involved in metabolism (NADH dehydrogenase), odorant perception (pheromone binding protein GP-9-like, also referred to as an odorant binding protein) (Leal, 2013), and brain development and sperm motility (growth arrest-specific protein-8) (zur Lage, Newton, & Jarman, 2019).Gas-8 may be implicated in differences in traits as diverse as gyne dispersal behavior (C. DeHeer et al., 1999) and male reproductive capacity (Lawson, Vander Meer, & Shoemaker, 2012). Interestingly, one of the constitutive supergene-effect loci we identify appears to be located outside of the supergene region: nose-resistant to fluoxetine protein 6-like is affiliated with defects in lipid transport to the ovaries and with lifespan in Caenorhabditis elegans and humans (Brejning et al., 2014; Choy & Thomas, 1999). This gene is a candidate for direct, trans -regulation by the Sb supergene and may contribute to the diminished early fecundity of polygyneSB/Sb queens relative to SB/SB queens (C. J. DeHeer, 2002; Keller & Ross, 1993). However, given recent evidence that duplicate genes may arise by translocation into the Sb supergene from other genomic locations (Fontana et al., 2019), improved Sb assemblies and annotation will be necessary to further explore this hypothesis.

Pheromone binding protein GP-9-like belongs to one of the best studied gene families in S. invicta: the insect odorant binding protein (OBP) family (Keller & Ross, 1998, 1999; Pracana, Levantis, et al., 2017). These genes are thought to produce proteins that function as molecular carriers, transporting odorant molecules to their receptors in some canonical study systems, though OBPs likely also function outside of the chemosensory system as general chaperone molecules (Leal, 2013). In S. invicta, OBPs are of particular interest, as it has long been known that the fire ant social polymorphism is marked by fixed amino acid differences at an OBP, General protein-9 (encoded by the gene Gp-9) (Ross, 1997). We found that the most strongly differentially expressed OBP-encoding gene in our data was not the widely studied Gp-9, but a paralog, OBP12, that exists as duplicated tandem genes on the Sb haplotype and as a single-copy gene on the SB haplotype (Fontana et al., 2019; Pracana, Levantis, et al., 2017). This Sb -specific duplication of OBP12 corresponds with a unique increase in gene expression and Sb allele-specific expression among S. invicta OBP genes (reads from both Sb paralogs evidently mapped to the single OBP12 encoded in the SB reference genome in our analyses; Figure S4). Our data are largely consistent with the recent findings by Dang et al. (2019) regarding increased expression of OBP12 in antennae of Sb -carrying workers, though we also observed increased OBP12 expression in both the brains and ovaries of Sb -carrying gynes, which may mitigate against the notion of antennal-specific neofunctionalization (Dang, Cohanim, Fontana, Privman, & Wang, 2019). Additionally, the >12-fold increase in expression we observed for OBP12 in Sb -carrying individuals compared to SB/SB individuals is much higher than the simple doubling one might expect if the duplicate was expressed similarly to its progenitor (Table 1, Figure S4). This points to effects of gene regulatory evolution. Our findings are consistent with the hypothesis that duplication of OBP12 in the Sb haplotype and subsequent Sb- upregulation of the derivative OBP12 may facilitate some differences in sensory functions and response behaviors between monogyne and polygyne fire ants (Keller & Ross, 1998;

Trible & Ross, 2016).

Conclusion. We find that the gene regulation underlying a major alternate form of social organization in fire ants is more complex than the simple inheritance patterns or broad structural differences in the implicated supergene may superficially suggest. As hypothesized (Y. C. Huang & Wang, 2014), gene expression evolution appears to play a large role in the evolution of the complex phenotype of polygyny in *S. invicta*. We show that possession of the *Sb* inversion-mediated haplotype by a queen leads to changes in gene expression both within and beyond the boundaries of the constitutive supergene inversions, that such supergene-mediated effects interact with those of the developmental environment, and that genes differentially expressed in queens bearing different supergene genotypes are heterogeneously distributed along the supergene. Further work exploring the interplay of protein sequence, gene expression, and regulatory element evolution is necessary to better understand the genetic causes of a fundamental shift in fire ant colony social organization with widespread behavioral, ecological, and economic consequences (Tschinkel 2006).

Acknowledgements

We thank Dietrich Gotzek for the idea to sample gynes embarking on mating flights, Bob Schmitz and Nick Rohr for library preparation, Yannick Wurm and Rodrigo Pracana for providing linkage mapping and OBP annotation information, and 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. This work was supported by U.S. NSF grants to B.H. and K.R. (1755130) and K.R. (1354479) and U.S. Federal Hatch funds to B.H. and K.R.

References

Ascunce, M. S., Bouwma, A. M., & Shoemaker, D. (2009). Characterization of 24 microsatellite markers in 11 species of fire ants in the genus Solenopsis (Hymenoptera: Formicidae). *Molecular Ecology Resources*, 9 (6), 1475–1479. doi:10.1111/j.1755-0998.2009.02688.x

Auwera, G. A., Carneiro, M. O., Hartl, C., Poplin, R., del Angel, G., Levy-Moonshine, A., ... DePristo, M. A. (2013). From FastQ Data to High-Confidence Variant Calls: The Genome Analysis Toolkit Best Practices Pipeline. *Current Protocols in Bioinformatics*, 43 (1), 11.10.1-11.10.33. doi:10.1002/0471250953.bi1110s43

Bachtrog, D., Kirkpatrick, M., Mank, J. E., McDaniel, S. F., Pires, J. C., Rice, W. R., & Valenzuela, N. (2011). Are all sex chromosomes created equal? *Trends in Genetics*, 27 (9), 350–357. doi:10.1016/j.tig.2011.05.005

Bin, D., Zhang, P., Zhang, Y., Pan, C., Meng, G., Xiao, X., ... Zhang, L. (2016). RNaseH2A is involved in human gliomagenesis through the regulation of cell proliferation and apoptosis. *Oncology Reports*, 36 (1), 173–180. doi:10.3892/or.2016.4802

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, 30 (15), 2114–2120. doi:10.1093/bioinformatics/btu170

Brejning, J., Norgaard, S., Scholer, L., Morthorst, T. H., Jakobsen, H., Lithgow, G. J., ... Olsen, A. (2014). Loss of NDG-4 extends lifespan and stress resistance in Caenorhabditis elegans. *Aging Cell*, 13 (1), 156–64. doi:10.1111/acel.12165

Choy, R. K., & Thomas, J. H. (1999). Fluoxetine-resistant mutants in C. elegans define a novel family of transmembrane proteins. *Molecular Cell*, 4 (2), 143–52. doi:10.1016/S1097-2765(00)80362-7

Conway, J. R., Lex, A., & Gehlenborg, N. (2017). UpSetR: an R package for the visualization of intersecting sets and their properties. *Bioinformatics (Oxford, England)*, 33 (18), 2938–2940. doi:10.1093/bioinformatics/btx364

Dang, V., Cohanim, A. B., Fontana, S., Privman, E., & Wang, J. (2019). Has gene expression neofunctionalization in the fire ant antennae contributed to queen discrimination behavior? *Ecology and Evolution*. doi:10.1002/ece3.5748 DeHeer, C., Goodisman, M., & Ross, K. (1999). Queen Dispersal Strategies in the Multiple-Queen Form of the Fire Ant Solenopsis invicta. *The American Naturalist*, 153 (6), 660–675.

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. doi:10.1006/anbe.2002.3095

Disteche, C. M. (2012). Dosage Compensation of the Sex Chromosomes. Annual Review of Genetics , 46 (1), 537–560. doi:10.1146/annurev-genet-110711-155454

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., ... Gingeras, T. R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, 29 (1), 15–21. doi:10.1093/bioinformatics/bts635

Eliyahu, D., Ross, K. G., Haight, K. L., Keller, L., & Liebig, J. (2011). Venom Alkaloid and Cuticular Hydrocarbon Profiles Are Associated with Social Organization, Queen Fertility Status, and Queen Genotype in the Fire Ant Solenopsis invicta. *Journal of Chemical Ecology*, 37 (11), 1242–1254. doi:10.1007/s10886-011-0037-y

Faria, R., Johannesson, K., Butlin, R. K., & Westram, A. M. (2019). Evolving Inversions. *Trends in Ecology* and Evolution . doi:10.1016/j.tree.2018.12.005

Fontana, S., Chang, N., Chang, T., Lee, C.-C., Dang, V.-D., & Wang, J. (2019). The fire ant social supergene is characterized by extensive gene and transposable element copy number variation. *Molecular Ecology*, mec.15308. doi:10.1111/mec.15308

Gel, B., & Serra, E. (2017). karyoploteR: an R/Bioconductor package to plot customizable genomes displaying arbitrary data. *Bioinformatics*, 33 (19), 3088–3090. doi:10.1093/bioinformatics/btx346

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

Hallar, B. L., Krieger, M. J. B., & Ross, K. G. (2007). Potential cause of lethality of an allele implicated in social evolution in fire ants. *Genetica*, 131 (1), 69–79. doi:10.1007/s10709-006-9114-5

Hill, W. G., & Robertson, A. (1966). The effect of linkage on limits to artificial selection. *Genetical Research*, 8 (3), 269–294. doi:10.1017/S0016672300010156

Hoffmann, A. A., & Rieseberg, L. H. (2008). Revisiting the Impact of Inversions in Evolution: From Population Genetic Markers to Drivers of Adaptive Shifts and Speciation? *Annual Review of Ecology, Evolution, and Systematics*, 39 (1), 21–42. doi:10.1146/annurev.ecolsys.39.110707.173532

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

Huang, Y., Dang, V. D., Chang, N., Wang, J., & Wang, J. (2018). Multiple large inversions and breakpoint rewiring of gene expression in the evolution of the fire ant social supergene. *Proceedings of the Royal Society B: Biological Sciences*, 285. doi:10.5061/dryad.2458p4r

Keller, L., & Ross, K. G. (1993). Phenotypic Basis of Reproductive Success in a Social Insect: Genetic and Social Determinants. *Science*, 260 (5111), 1107–1110. doi:10.1126/science.260.5111.1107

Keller, L., & Ross, K. G. (1998). Selfish genes: a green beard in the red fire ant. *Nature* , 394 (6693), 573–575. doi:10.1038/29064

Keller, & Ross. (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. doi:10.1046/j.1420-9101.1999.00064.x

Kirubakaran, T. G., Grove, H., Kent, M. P., Sandve, S. R., Baranski, M., Nome, T., ... Andersen, O. (2016). Two adjacent inversions maintain genomic differentiation between migratory and stationary ecotypes of Atlantic cod. *Molecular Ecology*, 25 (10), 2130–2143. doi:10.1111/mec.13592

Krieger, M. J. B., Ross, K. G., Chang, C. W. Y., & Keller, L. (1999). Frequency and origin of triploidy in the fire ant Solenopsis invicta. *Heredity*, 82 (2), 142–150. doi:10.1038/sj.hdy.6884600

Kunte, K., Zhang, W., Tenger-Trolander, A., Palmer, D. H., Martin, A., Reed, R. D., ... Kronforst, M. R. (2014). Doublesex is a mimicry supergene. *Nature*, 507 (7491), 229–232. doi:10.1038/nature13112

Kupper, C., Stocks, M., Risse, J. E., dos Remedios, N., Farrell, L. L., McRae, S. B., ... Burke, T. (2016). A supergene determines highly divergent male reproductive morphs in the ruff. *Nature Genetics*, 48 (1), 79–83. doi:10.1038/ng.3443

Lamichhaney, S., Fan, G., Widemo, F., Gunnarsson, U., Thalmann, D. S., Hoeppner, M. P., ... Andersson, L. (2015). Structural genomic changes underlie alternative reproductive strategies in the ruff (Philomachus pugnax). *Nature Genetics*, 48 (1), 84–88. doi:10.1038/ng.3430

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. *Proceedings of the Royal Society B: Biological Sciences*, 279 (1741), 3217–3222. doi:10.1098/rspb.2012.0315

Leal, W. S. (2013). Odorant Reception in Insects: Roles of Receptors, Binding Proteins, and Degrading Enzymes. Annual Review of Entomology, 58 (1), 373–391. doi:10.1146/annurev-ento-120811-153635

Liao, Y., Smyth, G. K., & Shi, W. (2013). The Subread aligner: Fast, accurate and scalable read mapping by seed-and-vote. *Nucleic Acids Research*, 41 (10), 1–17. doi:10.1093/nar/gkt214

McCarthy, D. J., Chen, Y., & Smyth, G. K. (2012). Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Research*, 40 (10), 4288–4297. doi:10.1093/nar/gks042

Nipitwattanaphon, M., Wang, J., Dijkstra, M. B., & Keller, L. (2013). A simple genetic basis for complex social behaviour mediates widespread gene expression differences. *Molecular Ecology*, 22 (14), 3797–3813. doi:10.1111/mec.12346

Parsch, J., & Ellegren, H. (2013). The evolutionary causes and consequences of sex-biased gene expression. Nature Reviews Genetics, 14 (2), 83–87. doi:10.1038/nrg3376

Pearse, D. E., Barson, N. J., Nome, T., Gao, G., Campbell, M. A., Abadia-Cardoso, A., ... Lien, S. (2019). Sex-dependent dominance maintains migration supergene in rainbow trout. *Nature Ecology & Evolution*, 1–12. doi:10.1038/s41559-019-1044-6

Phan, V. H., Herr, D. R., Panton, D., Fyrst, H., Saba, J. D., & Harris, G. L. (2007). Disruption of sphingolipid metabolism elicits apoptosis-associated reproductive defects in Drosophila. *Developmental Biology*, 309 (2), 329–41. doi:10.1016/j.ydbio.2007.07.021

Picard toolkit. (2018). Broad Institute, GitHub Repository. Broad Institute.

Picelli, S., Faridani, O. R., Bjorklund, A. K., Winberg, G., Sagasser, S., & Sandberg, R. (2014). Full-length RNA-seq from single cells using Smart-seq2. *Nat. Protocols*, 9 (1), 171–181.

Pracana, R., Levantis, I., Martinez-Ruiz, C., Stolle, E., Priyam, A., & Wurm, Y. (2017). Fire ant social chromosomes: Differences in number, sequence and expression of odorant binding proteins. *Evolution Letters*, 1 (4), 199–210. doi:10.1002/evl3.22

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* ,26 (11), 2864–2879. doi:10.1111/mec.14054

Purcell, J., Brelsford, A., Wurm, Y., Perrin, N., & Chapuisat, M. (2014). Convergent genetic architecture underlies social organization in ants. *Current Biology*, 24 (22), 2728–2732. doi:10.1016/j.cub.2014.09.071

Quinlan, A. R., & Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics (Oxford, England), 26 (6), 841–842. doi:10.1093/bioinformatics/btq033

Robinson, M. D., McCarthy, D. J., & Smyth, G. K. (2009). edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, 26 (1), 139–140. doi:10.1093/bioinformatics/btp616

Ross, K. G. (1997). Multilocus evolution in fire ants: effects of selection, gene flow and recombination. *Genetics*, 145 (4), 961–74.

Ross, K. G., Krieger, M. J. B., Keller, L., & Shoemaker, D. D. (2007). Genetic variation and structure in native populations of the fire ant Solenopsis invicta: evolutionary and demographic implications. *Biological Journal of the Linnean Society*, 92 (3), 541–560. doi:10.1111/j.1095-8312.2007.00853.x

Ross, K. G., & Shoemaker, D. (2018). Unexpected patterns of segregation distortion at a selfish supergene in the fire ant Solenopsis invicta. *BMC Genetics*, 19 (1), 1–22. doi:10.1186/s12863-018-0685-9

Ross, K. G., & Shoemaker, D. D. (2006). An Unusual Pattern of Gene Flow between the Two Social Forms of the Fire Ant Solenopsis invicta. *Evolution*, 47 (5), 1595. doi:10.2307/2410170

Ross, K. G., Shoemaker, D. D., Krieger, M. J. B., DeHeer, C. J., & Keller, L. (1999). Assessing genetic structure with multiple classes of molecular markers: a case study involving the introduced fire ant Solenopsis invicta. *Molecular Biology and Evolution*, *16* (4), 525–543. doi:10.1093/oxfordjournals.molbev.a026134

Schwander, T., Libbrecht, R., & Keller, L. (2014). Supergenes and complex phenotypes. *Current Biology*, 24 (7), R288–R294. doi:10.1016/j.cub.2014.01.056

Shoemaker, D. D., Deheer, C. J., Krieger, M. J. B., & Ross, K. G. (2006). Population Genetics of the Invasive Fire Ant Solenopsis invicta (Hymenoptera: Formicidae) in the United States. *Annals of the Entomological Society of America*, 99 (6), 1213–1233. doi:10.1603/0013-8746(2006)99[1213:pgotif]2.0.co;2

Stolle, E., Pracana, R., Howard, P., Paris, C. I., Brown, S. J., Castillo-Carrillo, C., ... Wurm, Y. (2019). Degenerative Expansion of a Young Supergene. *Molecular Biology and Evolution*, 36 (3), 553–561. doi:10.1093/molbev/msy236

Sun, D., Huh, I., Zinzow-kramer, W. M., Maney, D. L., & Yi, S. V. (2018). Rapid regulatory evolution of a nonrecombining autosome linked to divergent behavioral phenotypes. *Proc Natl Acad Sci U S A*, 115 (11), 1–6. doi:10.1073/pnas.1717721115

Trible, W., & Ross, K. G. (2016). Chemical communication of queen supergene status in an ant. *Journal of Evolutionary Biology*, 29 (3), 502–513. doi:10.1111/jeb.12799

Tschinkel, W. R. (2006). The Fire Ants . Belknap Press.

Tuttle, E. M., Bergland, A. O., Korody, M. L., Brewer, M. S., Newhouse, D. J., Minx, P., ... Balakrishnan, C. N. (2016). Divergence and Functional Degradation of a Sex Chromosome-like Supergene. *Current Biology*, 26 (3), 344–350. doi:10.1016/j.cub.2015.11.069

Valles, S. M., & 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. doi:10.1007/s00040-003-0662-8

Villanueva, R. A. M., & Chen, Z. J. (2019). ggplot2: Elegant Graphics for Data Analysis (2nd ed.). Measurement: Interdisciplinary Research and Perspectives, 17 (3), 160–167. doi:10.1080/15366367.2019.1565254

Wald, A., & Wolfowitz, J. (1940). On a Test Whether Two Samples are from the Same Population. *The* Annals of Mathematical Statistics ,11 (2), 147–162. doi:10.1214/aoms/1177731909

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:10.1038/nature11832

Wellenreuther, M., & Bernatchez, L. (2018). Eco-Evolutionary Genomics of Chromosomal Inversions. *Trends* in Ecology and Evolution ,33 (6), 427–440. doi:10.1016/j.tree.2018.04.002

White, M. A., Kitano, J., & Peichel, C. L. (2015). Purifying Selection Maintains Dosage-Sensitive Genes during Degeneration of the Threespine Stickleback Y Chromosome. *Molecular Biology and Evolution*, 32 (8), 1981–1995. doi:10.1093/molbev/msv078

Wijchers, P. J., & Festenstein, R. J. (2011). Epigenetic regulation of autosomal gene expression by sex chromosomes. *Trends in Genetics*, 27 (4), 132–140. doi:10.1016/j.tig.2011.01.004

Wurm, Y., Wang, J., Riba-Grognuz, O., Corona, M., Nygaard, S., Hunt, B. G., ... Keller, L. (2011). The genome of the fire ant Solenopsis invicta. *Proceedings of the National Academy of Sciences*, 108 (14), 5679–5684. doi:10.1073/pnas.1009690108

Yan, Z., Martin, S. H., Gotzek, D., Arsenault, S. V, Duchen, P., Helleu, Q., ... Keller, L. (2020). Evolution of a supergene that regulates a trans-species social polymorphism. *Nature Ecology & Evolution*, 4 (2), 240–249. doi:10.1038/s41559-019-1081-1

zur Lage, P., Newton, F. G., & Jarman, A. P. (2019). Survey of the Ciliary Motility Machinery of Drosophila Sperm and Ciliated Mechanosensory Neurons Reveals Unexpected Cell-Type Specific Variations: A Model for Motile Ciliopathies. *Frontiers in Genetics*, 10, 24. doi:10.3389/fgene.2019.00024

Data availability.

Raw reads and analyzed data for the RNA-sequencing libraries are available at the Gene Expression Omnibus – accession GSE149726. Other data supporting the findings of this study are available within the paper and its supplementary files. An R markdown document of differential expression analyses and an image of the workspace used for these analyses can be found at https://github.com/ArsenaultResearch/Arsenault-King-Simple-Inheritance.

Author Contributions

J.K., K.R., and B.H. conceived of the study. J.K. and K.R. collected samples. J.K. size-sorted gynes, performed DNA extractions, and genotyped samples. S.K. performed dissections and RNA extractions. K.L. performed microsatellite analysis. S.A., J.K., and B.H. designed bioinformatic analyses and S.A. implemented them. S.A., K.R., and B.H. wrote the manuscript with feedback from all authors.

Figures



Figure 1: Experimental design. Monogyne and polygyne Solenopsis invicta colonies differ in the number of reproductive queens. This difference is affiliated with an inversion-based supergene haplotype on the social chromosome (chromosome 16) denoted as Sb; the homologous region is designated SB. Monogyne colonies produce only SB/SB gynes (unmated winged queens), which are relatively high in weight upon sexual maturity. Polygyne colonies produce 1)SB/SB gynes, which are slightly lighter than monogyne SB/SB gynes at maturity, 2) SB/Sb gynes, which are of intermediate weight, and 3) Sb/Sb gynes, which are low in weight. Monogyne-derived SB/SB gynes fly to a high altitude to mate and independently (claustrally) found a new colony upon landing. The relatively rare polygyne-derived SB/SB gynes evidently do the same, although empirical evidence for this is lacking (denoted by a question mark). Polygyne SB/Sb gynes fly lower to the ground and attempt to enter and join an existing colony after flying. PolygyneSb/Sb gynes depart on mating flights but rarely survive them. Gynes of each type were collected as they left their nests on mating flights; their brains and ovaries subsequently were dissected for RNA-sequencing.



Figure 2: Supergene-mediated differential gene expression. (A) Volcano plot of gene expression differences between SB/SB and SB/Sb gynes, independent of organ and social-environment effects (multifactorial-model). Aqua dots represent statistically significant differences (FDR < 0.01). (B) Volcano plot of gene expression differences between SB/SB and Sb/Sb gynes. (C) UpSet plot showing organspecific patterns of differential gene expression associated with differences in social chromosome genotype in polygyne-reared gynes. The pattern of genes common to different comparisons is indicated by the dots below the vertical bar plot. Numbers in parentheses indicate the genes remaining in a category after elimination of cusp cases (genes with FDR = 0.01-0.1 in other contexts, described in Methods). (D) Heatmap of the distribution of differentially expressed genes across the linkage-mapped genome in polygyne-reared gynes. The region spanned by the lg16 Sb supergene is marked in green. Labels to the right of the heatmap refer to candidate genes in each of the categories described in panel C.



Figure 3: Effects of social form on gene expression. (A) Volcano plot of gene expression differences due to differences in social form of origin of each gyne across all three social chromosome genotypes and both organs (multifactorial-model). Euler diagrams of the number of DEGs common to different pairwise comparisons are shown for brains (B) and ovaries (C). The 'm' and 'p' prefixes denote the natal social form of origin of gynes (monogyne and polygyne, respectively).



Figure 4: Allele-specific expression within the supergene. (A) Volcano plot of allele-specific expression (ASE) for genes with informative SNPs. Aqua dots represent significant instances (edgeR glmQLFtest, FDR < 0.01), with the names of the genes showing the most extreme ASE ($|Log_2FC| > 3$) indicated. (B) UpSet plot showing overlap between genes with significant differential expression (DE, FDR < 0.01) in the *SB/SB* versus *SB/Sb* multifactorial-model comparison and those with significant ASE in *SB/Sb* heterozygotes (context for increased relative expression shown with arrows). Only genes subjected to both types of analysis are included in this plot.



Figure 5: Spatial patterns of differentially expressed genes (DEGs) and allele-specific expression (ASE) within the supergene. The social chromosome (lg16) physical map is shown at the bottom, with the supergene location marked in green above it. Panel A shows ASE of significant individual SNPs (FDR < 0.01), with olive dots representing elevated SB- linked expression and blue dots elevated Sb- linked expression. Heatmaps shown at the upper and lower margins of panel A represent densities of SNPs that display SB orSb ASE, respectively. Significant DEGs (FDR < 0.01) identified in two comparative contexts are represented in Panels B and C, with genes upregulated in the presence of Sb marked by aqua dots and genes upregulated in SB/SB marked by gold dots (multifactorial-model). Heatmaps of DEG density across 500kb windows are shown at the upper and lower margins of panels B and C. The Y-axes indicate log₂ folddifference in expression and the size of the dots indicates FDR-corrected P-value. The upper heatmaps in Panels B and C represent SB/SB -upregulated genes, while the lower heatmaps represent S_{-}/Sb -upregulated genes. Region 1, demarcated by yellow dotted lines, delineates a hotspot of SB- linked expression; Region 2, demarcated by blue dotted lines, delineates a hotspot of Sb -linked expression.

GeneID ¹	Description ¹	SB/SB vs. SB/S
		Brain log ₂ FC ²
Constitutive		
LOC105206625	putative nuclease HARBI1	-3.58**
LOC105194501	pheromone-binding protein Gp-9-like	-3.88**
LOC105194902	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 9, mitochondrial-like	-5.42**
LOC105196479	nose resistant to fluoxetine protein 6-like	-5.03**
LOC105198369	putative nuclease HARBI1	-4.7**
LOC105199315	growth arrest-specific protein 8	-4.38**
	7 uncharacterized genes not included ⁵	
Brain-specific		
LOC105205135	USP6 N-terminal-like protein	-3.27**
LOC105207200	probable serine hydrolase	-5.97**
LOC105193214	ATP-dependent RNA helicase vasa, isoform A	-0.99**
LOC105194434	limulus clotting factor C-like	-4.48**
LOC105198311	ribonuclease H-like, transcript variant X1	7.21**
LOC105199289	sperm flagellar protein 1-like	-3.2**
LOC105199833	low-density lipoprotein receptor-like	-2.47**
	1 uncharacterized gene not included ⁵	
Ovary-specific		
LOC105203027	sphingosine-1-phosphate lyase	0.05
LOC105203994	general odorant-binding protein 69a-like, transcript variant X2	0.05
LOC105206806	WAS/WASL-interacting protein family member 1-like	-1.7
LOC105207398	enhancer of split mbeta protein-like	-0.15
LOC105194452	JNK-interacting protein 1, transcript variant X1	0.23
LOC105198124	high affinity choline transporter 1-like	-0.08

Table 1: Candidate genes associated with social form based on gene expression patterns.

GeneID ¹	Description ¹	SB/SB vs. SB/S
LOC105200547	protein SCAI 5 uncharacterized genes not included 5	0.02
Sb/Sb-specific	nra mRNA enlicing factor Shu like	1 16
LOC105202812 LOC105203028	peptide deformylase, mitochondrial-like	0.49
LOC105206526	dopamine receptor 1	0.40
LOC105199521	4 uncharacterized genes not included 5	1.09

Gray shading indicates nonsignificant gene expression effects. Candidate genes discussed in text are highlighted with bold underlining.¹ GeneID and description taken from the NCBI RefSeq annotation release $100.^2$ Log2-Fold change (log₂FC) values are included with the corresponding statistical significance (* FDR p < 0.05, ** FDR p < 0.01). Positive values indicate elevated expression in *SB/SB* samples relative to samples with *Sb*-bearing genotypes, negative values indicate elevated expression for *Sb*-bearing genotypes.³ Presence in the supergene is based on the previously published linkage mapping data (Pracana, Priyam, et al., 2017; Wang et al., 2013).⁴ Additional support indicates whether each gene has been reported as differentially expressed in previous studies.**a** indicates (Wang et al., 2008), **b** indicates (Wang et al., 2013), **c** indicates (Nipitwattanaphon et al., 2013),**d** indicates (Pracana, Levantis, et al., 2017), and eindicates haplotype-specific gene duplications identified in (Fontana et al., 2019). Criteria for inclusion of genes in each category were specific, conservative P-value thresholds conceived in our study (see Methods for details).⁵ Details on uncharacterized candidate loci expression are reported in Supplementary Data 2

SUPPLEMENTARY DATA

Supplementary Data 1: RNA quality, microsatellite-based relatedness assay, sequence quality control and alignment statistics for each of our samples.

Supplementary Data 2: Compendium of differential expression and allele-specific expression for all genes.

Supplementary Data 3: GO term enrichment for differential expression analyses (P < 0.05).

Supplementary Data 4: Overlap and correlations for differentially expressed genes when using the Wurm et al. 2011 and Yan et al. 2020 genome assemblies.

SUPPLEMENTARY METHODS

Polar Dominance Analysis. To quantify the dominance patterns of the various differentially expressed genes, we needed to leverage information from both the SB/SB vs. SB/Sb comparison as well as the SB/SB vs. Sb/Sb comparison at the same time. Additionally, we needed to check all genes effected by the presence of the supergene, not simply the genes that were differentially expressed in either comparison. To this end we used the full glm approach querying for any genes whose expression was significantly affected by the presence of the supergene (FDR < 0.01). We then computed the angle between the log2-fold change in the SB/SB vs. SB/Sb comparison and the log2-fold change in the SB/SB vs. Sb/Sb comparison (Figure S3). This angle is descriptive of the dominance pattern of that particular gene.

Odorant Binding Protein Differential Expression. Leveraging the gene models from previously published work (Pracana, Levantis, et al., 2017), we tested for differential expression amongst the most complete set of odorant binding proteins (OBPs) available in the system. We re-mapped our aligned reads to the OBP gene models using RSubread's FeatureCounts function (Liao et al., 2013) and then computed differential expression using edgeR as described in the Methods (McCarthy et al., 2012). We then plotted the log2-fold change and FDR-corrected p-value using the pheatmap package in R. GP-9 is described here as SiOBP3.

GO term enrichment analysis. GO terms were determined for each gene using BLAST2GO to search for homologous sequences in the Drosophila melanogaster genome (version r6.20). We then used topGO with

default parameters to compute enriched GO terms amongst significantly differentially expressed genes (FDR < 0.01) compared against a background of genes that passed our coverage filtration. Terms with a P < 0.05 were deemed significant.

SUPPLEMENTARY FIGURES



Supplementary Figure S1: Volcano Plots of tissue-specific pairwise comparisons. (A) Volcano plot of gene expression differences between polygyne SB/Sb individuals and SB/SB individuals in the brain. (B) Volcano plot of gene expression differences between polygyne Sb/Sb and SB/SB individuals in the brain. (C) Volcano plot of gene expression differences between polygyne heterozygous and SB/SB individuals in the ovaries. (D) Volcano plot of gene expression differences between polygyne Sb/Sb and SB/SB individuals in the ovaries.



Supplementary Figure S2: Gene expression directionality outside and within the supergene region. Stacked bar plots showing the directional pattern of differential expression of genes outside and within the supergene. Numbers indicate the number of differentially expressed genes (FDR < 0.01) in a given category. "Multifactorial" indicates that the DEG's were inferred using the Multifactorial-model method. "Brain" and "Ovary" indicate that the DEGs were inferred using tissue-specific pairwise comparisons of polygyne individuals. * P < 0.05, ** P < 0.01, X²-test.



Supplementary Figure S3: Gene expression dominance with respect to a sample's social chromosome genotype. (A) Explanation of the conversion of log-fold change values for SB/SB vs. SB/Sb and SB/SB vs. Sb/Sb comparisons into a single angle value representing the dominance pattern of a given gene. For space, in this figure SB/SB is abbreviated BB, SB/Sb is abbreviated Bb, and Sb/Sb is abbreviated bb. (B) Polar plot of the dominance patterns of genes differentially expressed in response to changes in supergene genotype within the supergene itself. The angle represents the expression pattern while the radial distance is the negative log-transformed FDR-corrected p-value (further out implying more significant). (C) Polar plot of the dominance patterns of genes outside of the supergene and differentially expressed in response to changes in supergene genotype.



Supplementary Figure S4: Differential gene expression of Odorant Binding Proteins by supergene status and social form. (A) Gene expression \log_2 -Fold Change (logFC) values for our tissue-specific, multifactorial, and social form pairwise comparisons. Negative logFC values (red) indicate higher expression in*Sb* -carrying individuals or the polygyne social form. (B) FDR-corrected p-value heatmap for the differential expression analyses in Panel A. The lighter the tile, the more significant that particular gene is. OBP gene models were provided by Pracana et al. 2018.



Supplementary Figure S5: Spatial patterns of differentially expressed genes (DEGs) within the supergene in the Yan et al. 2020 SB assembly. The social chromosome (chr16) physical map is shown at the bottom, with the supergene location marked in green, blue, and red above it. Each color represents a region generated by a particular inversion. Red indicates the first inversion, blue the second, and green the third. Additionally, there is a small purple region on the right end of the supergene that was present in both the first and second inversion. Each track represents a differential expression analysis with GLM comparisons describing the comparisons using a multifactorial model. Dots indicate significant DEGs (FDR < 0.01) with genes upregulated in the presence of Sb marked by aqua dots and genes upregulated in SB/SB marked by yellow dots. The Y-axes indicate log₂ fold-difference in expression and the size of the dots indicates FDR-corrected P-value. Region 1, demarcated by the yellow line at the top of the plot, delineates a hotspot of SB- linked expression; Region 2, demarcated by a blue line, delineates a hotspot of Sb -linked expression.

Hosted file

MolEcol_Manuscript_Submission2_untracked.docx available at https://authorea.com/users/319470/ articles/449199-simple-inheritance-complex-regulation-supergene-mediated-fire-ant-queenpolymorphism