Genome-Wide Expression Patterns and the Genetic Architecture of a Fundamental Social Trait

John Wang1, Kenneth G. Ross2, Laurent Keller1

Abstract

Explaining how interactions between genes and the environment influence social behavior is a fundamental research goal, yet there is limited relevant information for species exhibiting natural variation in social organization. The fire ant *Solenopsis invicta* is characterized by a remarkable form of social polymorphism, with the presence of one or several queens per colony and the expression of other phenotypic and behavioral differences being completely associated with allelic variation at a single Mendelian factor marked by the gene Gp-9. Microarray analyses of adult workers revealed that differences in the Gp-9 genotype are associated with the differential expression of an unexpectedly small number of genes, many of which have predicted functions, implying a role in chemical communication relevant to the regulation of colony queen number. Even more surprisingly, worker gene expression profiles are more strongly influenced by indirect effects associated with the Gp-9 genotypic composition within their colony than by the direct effect of their own Gp-9 genotype. This constitutes an unusual example of an “extended phenotype” and suggests a complex genetic architecture with a single Mendelian factor, directly and indirectly influencing the individual behaviors that, in aggregate, produce an emergent colony-level phenotype.

Introduction

Considerable interest surrounds the genetic architectures underlying fundamental adaptive traits in wild populations [1–5]. In social organisms, such interest centers on the numbers and types of genes directly regulating expression of the individual behaviors that, in aggregate, create social organization, as well as genes in interactants that indirectly influence expression of socially relevant behaviors by altering the social environment [6–12]. This indirect influence is mediated by interactions of the genotype of a given individual with those of other group members who collectively comprise the social environment. Information on the genetic architecture of social organization is essential to constructing realistic models of social evolution that can answer questions about the numbers and types of genetic changes necessary to change a solitary to a social animal or to convert a simple society to a large and highly complex one [13].

A remarkable case of a fundamental social polymorphism that appears to be under simple genetic control (single Mendelian factor of large effect) is variation in colony social organization in the fire ant *Solenopsis invicta*. In this species a single genomic region marked by the protein-encoding gene Gp-9 is implicated in the production of two distinct types of queens that differ in physiology, fecundity and behavior [14–19]. This genetic factor also determines whether workers tolerate a single fertile queen (monogyne social form) or multiple queens (polygyne social form) in their colony. Colonies containing only homozygous Gp-9 BB workers accept only a single queen, whereas colonies containing both Gp-9 BB and Gp-9 Bb workers invariably accept multiple queens, but only those bearing a Gp-9 b haplotype [20–23]. The near complete absence of adult workers and queens with a bb genotype stems from the deleterious effects associated with the genomic region marked by the b allele, inducing homozygous females to die shortly after they eclose from the pupa [22,24]. The monogyne and polygyne social forms also differ in a number of important reproductive, behavioral, and life history traits besides colony queen number [25,26], differences that are also completely associated with differences at the genomic region marked by Gp-9. In contrast, there is a complete lack of differentiation at genes not tightly linked to Gp-9, presumably because frequent matings between sexuals from sympatric monogyne and polygyne colonies result in extensive gene flow between the forms [22,27–29].

Colony queen number in *S. invicta* is regulated by the workers, which collectively decide which and how many queens from within or outside the colony are recruited as new egg-layers [21], largely on the basis of chemical signals emanating from the queens [20]. Workers in monogyne colonies (all of which possess the BB genotype) accept only a single replacement queen that must also bear genotype BB, whereas workers in polygyne colonies accept multiple queens, each of which must possess the b haplotype. Significantly, the presence of as few as 5–10% of workers with the b haplotype induces the entire colony worker force, including BB workers, to become tolerant of multiple Bb queens and thus display the polygyne social phenotype [23]. Thus the genomic region marked by Gp-9 exerts indirect genetic effects [9], in that the presence of the b variant in a colony induces changes in the social behavior of all colony members, even those lacking the b haplotype. Although the identity of the product of Gp-9 as an
Author Summary

Fundamental research goals for scientists interested in social evolution are to determine the numbers and types of genes that directly regulate individual social behaviors as well as to understand how the social environment indirectly influences the expression of socially relevant traits. The fire ant Solenopsis invicta features a remarkable form of social variation in which the occurrence of two distinct social types that differ in colony queen number is associated with genetic differences at a genomic region marked by the gene Gp-9. Our analyses of gene expression profiles in fire ant workers revealed that differences in Gp-9 genotype are associated with the differential expression of an unexpectedly small number of genes, many of which are predicted to function in chemical communication relevant to the regulation of colony queen number. Surprisingly, worker gene expression profiles are more strongly influenced by indirect effects associated with the social environment within their colony than by the direct effect of their own Gp-9 genotype. These results suggest a complex genetic architecture underlying the control of colony queen number in fire ants, with a single Mendelian factor directly regulating, and the social environment indirectly influencing, the expression of the individual behaviors that, in aggregate, yield an emergent colony social organization.

odorant binding protein (OBP) and other lines of evidence suggest that the gene may play a direct part in regulating social organization via a role in chemical communication, it remains an open question whether other genes tightly linked to Gp-9 (possibly locked up in an inversion with it) are also involved [22].

The first aim of this study was to investigate whether variation in the genomic region marked by Gp-9 is associated with differences in patterns of expression of genes other than Gp-9 in workers. The second aim was to study how the social environment (i.e., presence or absence of nestmate workers with the b allele) can alter individual gene expression patterns. To answer these questions and begin to address the issue of how variation at a single Mendelian factor can directly and indirectly affect gene expression to produce a complex colony-level phenotype, we employed a fire ant microarray platform representing some 10,000 genes [30].

Results/Discussion

To determine the effect of Gp-9 genotype on gene expression in focal individuals, we compared expression profiles between BB and Bb adult workers from 20 polygynous S. invicta colonies. This comparison revealed 39 genes consistently differentially expressed between workers of the two genotypes, of which about two-thirds were more highly expressed in Bb than BB workers (Figure 1; see also confirmation of microarray data by quantitative RT-PCR [qRT-PCR] in Text S1 and Table S1). Sixteen of these genes did not significantly match any sequence in the public databases (BLASTX, threshold E-value = 1e-5), ten matched predicted proteins of unknown function, and the remaining 13 matched genes with a known or inferred function (Table 1).

Three gene categories were overrepresented among the genes differently expressed between workers of alternate genotypes (Table 1 and Table S2; all P<0.05). The first two are the allergen and odorant binding protein categories, which collectively include five genes likely to contribute to chemical signaling and response, the essential components of regulation of colony queen number and social organization. In ants, venom allergens are proteins released from the venom sac, an organ that, in queens, appears to store and release chemical signals allowing recognition by workers [25]. Similarly, the two odorant binding genes that encode members of the insect OBP protein family (as does Gp-9) and the antennal chemosensory protein may be involved in pheromone transduction, thereby potentially influencing the abilities of workers of the two Gp-9 genotypes to recognize and discriminate among queens. Experiments from other systems are suggestive that changes in the expression levels of OBP could influence discriminatory behavior by modulating the threshold for a particular response; differential regulation of OBPs has been observed in Drosophila following mating [31], exposure to starvation stress [32], and alcohol tolerance development after exposure to alcohol [33]. Additionally, genetic and biochemical evidence suggests that OBPs may interact combinatorially in odor discrimination [34,35].

The third overrepresented category comprises two transposons, which are of special interest with respect to properties that may be shared between the genomic region including Gp-9 and regions containing the sex-determining genes in species with sex chromosomes [36]. The b haplotype is found only in the polygenic social form, just as the Y chromosome is found only in males in species with male heterogamy. By analogy with the Y chromosome, theoretical predictions and empirical observations suggest that the Gp-9 b region should (i) accumulate genes beneficial in the polygenic social environment (as the Y chromosome accumulates genes beneficial to male function [37]), (ii) evolve reduced recombination to preserve associations of genes advantageous for polygyny (as occurs for genes advantageous to males on the Y chromosome [38,39], and (iii) accumulate deleterious alleles and transposable elements (because of reduced recombination [38,39]). Consistent with these expectations, the genomic region marked by Gp-9 is characterized by low recombination [40,41], the b haplotype is a homozygous lethal [15,24,40], and the piggyBac-like transposon, which is differentially expressed between workers of alternate genotypes, appears to occur almost exclusively in individuals possessing haplotype b (data not shown). Thus the strong expression of at least this transposon in b-bearing workers, which constitutes the most extreme expression difference among the 13 genes with annotated matches (Table 1), likely reflects its unique insertion in b haplotypes. While this distribution could signify that the piggyBac-like transposon directly affects the differential expression of other candidate genes in BB and Bb workers, we note that, consistent with earlier protein electrophoresis data [40], no significant difference in the expression levels of Gp-9 was detected between workers of the two genotypes; therefore, whatever elements control the differential expression of genes in parallel with Gp-9 genotype appear not to regulate the expression of Gp-9 itself.

To determine the indirect effects of colony Gp-9 genotype composition as well as other aspects of the social environment on worker gene expression, while controlling for individual Gp-9 genotype, we compared profiles of adult workers bearing genotype BB between 20 polygynous and 20 monogynous colonies. This comparison revealed 91 genes consistently differentially expressed between workers of the alternate forms, of which over three-quarters were more highly expressed in polygynous than monogynous workers (Figure 2; see also confirmation of microarray data by qRT-PCR in Text S1 and Table S1). Forty-five of these genes did not significantly match any sequence in the public databases (BLASTX, threshold E-value = 1e-5), 13 matched predicted proteins of unknown function, and the remaining 33 matched previously annotated genes (Table 2).

Three gene categories (mitochondrial, prefoldin complex, and viral genes) were overrepresented among the genes that were differentially expressed between BB workers from monogynous and
polygyne colonies (Table 2 and Table S2; all $P<0.05$). The 11 genes encoding mitochondrial and prefoldin complex (molecular chaperone) proteins were all up-regulated in polygyne compared to monogyne workers. Increased mitochondrial gene expression may reflect increased oxidative metabolism, while increased prefoldin expression may indicate higher protein synthesis rates, possibly in relation to the relatively smaller size and higher metabolic rates of polygyne workers [42].

The pattern of expression of the six genes in the viral gene category is consistent with the expectation that differences in social organization affect susceptibility to pathogens and parasites. In the monogyne form, there is intense selection against susceptible infected individuals during independent colony founding, a stage that colonies of the polygyne form never display [26]. Accordingly, we found that workers in the polygyne form express more sequences corresponding to viral genes than their counterparts in the monogyne form, presumably because of relaxed selection and generally greater susceptibility in the former [see also [43]]. Based on sequence similarity and correlated expression across our experiments, we identified six gene products that likely represent three different viruses, a ssRNA negative-strand (−) virus and two ssRNA positive-strand (+) viruses, one of which is the SINV-2 virus [44]. All 20 polygyne study colonies showed evidence of infection with at least one virus (mean number of viral types per colony, 2.5±0.67), whereas only three of the 20 monogyne colonies showed evidence of infection, in all cases by a single viral type. Finally, the pattern of expression of another socially-regulated gene, this one encoding a defensin (a class of small protein antibiotics active against viruses, bacteria, and fungi [45]), also is consistent with greater selection for resistance in the monogyne form, as this gene was more highly expressed in monogyne than polygyne workers.

The numbers of genes differentially expressed in the genotype (39) and social form (91) comparisons are relatively low compared to the numbers expected based on other published microarray experiments. There are several possible explanations for this. First, the use of whole worker bodies as the source of RNA may decrease the probability of detecting genes whose level of expression varies among cells or tissues. Second, our comparisons were performed on groups of workers originating from different colonies, thus adding a colony-level effect to, and thus increasing the total variance in, gene expression. Finally, workers of alternate genotype or social form apparently exhibit fewer phenotypic differences than queens [20,21,46,47], possibly reflecting the involvement of fewer differentially expressed genes in the former caste.

Remarkably, there was almost no overlap between genes whose level of expression was influenced by the focal workers’ Gp-9 genotypes and genes whose expression was influenced by the social environment, with only one of the 129 differentially expressed genes appearing in both categories. This demonstrates an almost complete decoupling of the direct effects of the genomic region marked by Gp-9 and the indirect effects mediated by social interactions within colonies. Moreover, there is little indication of an interaction between these direct and indirect effects; genes...
whose levels of expression depend on the Gp-9 genotype of focal individuals generally are expressed at similar levels in the two social forms when genotype is held constant (Figure 1), whereas genes whose levels of expression depend on the indirect influence of the social environment almost always are expressed at similar levels in polygyne workers of different Gp-9 genotypes (Figure 2; see also Figure S1).

This study reveals that variation at the S. invicta genomic region marked by Gp-9 is associated with the differential expression in workers of a relatively small number of genes, that with the exception of the piggyBac-like transposon, presumably are unlinked to Gp-9. A high proportion of these differentially expressed genes have putative functions implying a role in chemical signaling and behavior relevant to the regulation of colony queen number and, therefore, these genes may have a primary function in determining social organization. The number of such genes is unexpectedly low given the profound behavioral, physiological, and life-history differences between the two social forms and the fact that widespread changes in gene expression patterns can be observed after just a few generations of selection [48,49]. A perhaps more surprising finding is that worker gene expression profiles are significantly more strongly influenced by indirect effects associated with the Gp-9 genotypic composition within their colony than by the direct effect of their own Gp-9 genotype (chi-squared test, P<0.001), with the indirect-effect genes largely implicated in the secondary differences in colony social characteristics expected between the forms. While several studies have demonstrated that the social environment can modulate gene expression [50–53], and others have revealed indirect genetic effects on phenotypes or levels of gene expression [10,54–60], this is the first example of a naturally occurring polymorphic Mendelian element that affects gene expression in other group members. The finding of a complex genetic architecture directly and indirectly influencing the individual behaviors that, in aggregate, generate a fundamental colony-level social phenotype represents an unusual example of an “extended phenotype” [61].

### Materials and Methods

#### Sample Collection

Colonies of S. invicta were collected near Athens, Georgia (eight polygyne and eight monogyne colonies in 2004; eight polygyne and eight monogyne colonies in 2006) and near Hammond, Louisiana (four polygyne and four monogyne colonies in 2006), USA. All colonies were returned to the laboratory and reared for one month under standard conditions [62]. We determined the social form of each study colony using several lines of evidence. Nest density, worker size distribution, and nest brood composition were used to make initial identifications of social form in the field (see [63]). Subsequently, polygyny was confirmed for each suspected polygyne colony by discovering two or more wingless inseminated (reproductive) queens, while monogyny was confirmed in each suspected monogynous colony by discovering a single, highly physogastric, wingless inseminated queen. The social form of each colony was further substantiated by electrophoretically detecting the b allele of Gp-9 in pooled samples of 20 female inhabitants of each polygyne colony and failing to detect the allele in such samples from each monogynous colony (the b allele is completely diagnostic for polygyny in S. invicta in the USA [14,40,63]).

#### mRNA Isolation and Microarray Hybridization

From each polygyne colony, 24–40 medium-sized adult workers were haphazardly collected from the foraging area of each colony and individually flash-frozen with liquid nitrogen in tubes containing 1 g of 1.4 mm ceramic beads (Quackenbush). From each monogyne colony, 10 medium-sized adult workers were collected in an identical fashion. For the 2004 samples, individual
ants were homogenized in Trizol (Invitrogen) using a Fastprep bead shaker, and DNA and RNA were extracted using the manufacturer’s recommended protocol. For the 2006 samples, individual ants were homogenized in RLT + buffer, and DNA and RNA were extracted using the AllPrep RNA/DNA Mini Kit (Qiagen). An RFLP analysis was used to determine the Gp-9 genotype of each individual in polygyne colonies [14]. We pooled the RNA from 7–10 BB workers and 7–10 Bb workers from each polygyne colony. Although bb workers generally are rare due to deleterious effects associated with the genotype [24,40], we found 13 such workers from eight colonies. Two pooled bb samples were created, one for 2004 (nine workers from four colonies) and one for 2006 (four workers from four colonies). These samples were hybridized (after amplification) but not included in the statistical analysis due to the small number and pooling of individuals across colonies. RNA from ten workers from each monogyne colony (all with genotype BB) was pooled by colony.

For both the 2004 and 2006 samples, pooled total RNA was linearly amplified once (Ambion MessageAmp II kit), then labeled using a modified version of the aminoallyl-labeling method in which reverse transcription is performed in the presence of aminoallyl-dUTP and the resulting cDNA is coupled to Cy3 or Cy5 fluorescent monomers [64,65]. Briefly, amplified RNA (~5 μg) was mixed with random 9mers (2 μg/μl), 0.5 μl of Alien mRNA Spike mix (Stratagene), and water for a final volume of 17.5 μL. This RNA/primer mix was incubated for 10 min at 70°C, then held for 5 min on ice. Reverse transcription was performed for 2 h at 50°C after adding 6 μL of 5× first-strand buffer, 3 μL of 0.1 M DTT, 0.6 μL of 50× aminoallyl-dNTP mix (25 mM dATP, 25 mM dCTP, 25 mM dGTP, 15 mM dTTP, 10 mM aminoallyl-dUTP), 1 μL of RNase inhibitor (15 U/μL, Invitrogen), and 2 μL of SuperScript III reverse transcriptase (200 U/μL, Invitrogen). The RNA was then hydrolyzed by adding 15 μL of 0.1 M NaOH and incubating for 10 min at 70°C. The pH was neutralized by adding 15 μL of 0.1 M HCl.

The aminoallyl-labeled cDNA was purified with a modified Qiaquick PCR Purification Kit (QIAGEN) and coupled to Cy3 or Cy5 dyes [64]. The combined Cy3- and Cy5-labeled probes were purified using the Qiaquick PCR Purification Kit (QIAGEN) and eluted in 72 μL of elution buffer. After adding 13.5 μL of 20× SSC, 2.7 μL of yeast tRNA (2 μg/μL), 2.7 μL of polyA DNA (2μg/μL; Sigma), and 1.62 μL of 10% SDS, the probe was denatured at 100°C for 45 sec and hybridized to the microarray slides at 64°C overnight. Excess probe was removed by washing for 2×5 min in 2× SSC, 0.1% SDS; 2×1 min in 0.2× SSC; 1×1 min in 0.1× SSC; and 1×5 min in 0.1× SSC, 0.1% Triton at room temperature.

Experimental samples were labeled with Cy3 and were hybridized against Cy5-labeled “common reference” RNA on our custom-made spotted cDNA microarrays. We employed a common reference design because not all samples provided...
enough amplified RNA for multiple hybridizations (e.g., for loop foraging area of 30 colonies (eight and seven of each social form from isolated en masse from hundreds of adult workers collected from the experimental sample. For the 2006 samples, we amplified total RNA from 25% of the amplified RNA from each experimental sample. The signal intensities for each spot (representing 9,722 putative genes) were considered in the analyses.

For both batches, only the 18,438 spots yielding a single PCR product were used, one set printed in 2004 and the other in 2006. For both experimental samples, we amplified total RNA isolated en masse from hundreds of adult workers collected from the foraging area of 30 colonies (eight and seven of each social form from Georgia and Louisiana, respectively). The microarrays were made from 22,560 independent cDNAs generated from a fire ant expressed sequence tag project and are estimated to represent 11,864 different genes [30]. Two different batches of microarrays were used, one set printed in 2004 and the other in 2006. For both batches, only the 18,438 spots yielding a single PCR product (representing 9,722 putative genes) were considered in the analyses. Images of the competitive hybridization were obtained with an Agilent Technologies Scanner. The signal intensities for each spot

**Table 2.** Genes differentially expressed between *S. invicta* workers from monogyne and polygyne colonies bearing the BB genotype at Gp-9 that significantly match annotated genes in public databases.

| Fire ant gene | Putative gene product of best match | F-value | Gene category | Expression Ratio (P/M) | P-value |
|---------------|-----------------------------------|---------|---------------|-----------------------|---------|
| SJWH118C2Z.scf | Y-box protein (A2A246) | 1.00E-38 | regulation of transcription | 0.69 | 7.24E-07 |
| SL.C1.20.209S.Contig1 | defense-2 (Q5MQL3) | 3.00E-15 | immunity | 0.70 | 8.39E-05 |
| SL.C1.221.2132.Contig1 | prenylycysteine oxidase (EGNAP00000029873) | 1.00E-21 | oxidereductase | 0.77 | 3.98E-04 |
| SL.C1.11.1163.Contig1 | alpha-glucosidase (hbg3) (Q2S876) | 8.00E-55 | metabolism | 0.78 | 4.39E-04 |
| SL.C1.31.3193.Contig1 | S-methyl-5-thioadenosine phosphorylase (AGAP005129-PA) | 1.00E-44 | transferase | 0.79 | 4.72E-05 |
| SJWF02AX.scf | alanine-glyoxylate aminotransferase (Q6DW8) | 3.00E-17 | transferase | 0.84 | 7.20E-05 |
| SL.C5.16.571.Contig1 | ribosomal protein L3 (Q56F10) | 0 | translation | 0.84 | 3.32E-04 |
| SL.C6.61.615.Contig1 | probable allergen protein (Q18530) | 5.00E-27 | allerger | 0.89 | 2.34E-04 |
| SL.C3.30.3064.Contig1 | mitochondrial import inner membrane translocase subunit Tim9 (Q17HY2) | 1.00E-28 | mitochondrial | 1.14 | 7.25E-04 |
| SL.C1.11.1166.Contig1 | dyned light chain 2B (Q174A0) | 6.00E-43 | microtubule motor activity | 1.15 | 8.58E-05 |
| SL.C1.21.2117.Contig1 | prefoldin subunit 4 (Q17U1) | 1.00E-36 | prefoldin chaperone | 1.17 | 2.84E-04 |
| SL.C1.25.2556.Contig2 | prefoldin subunit 6 (A21449) | 2.00E-30 | prefoldin chaperone | 1.17 | 7.62E-04 |
| SL.C1.0.0.000.Contig1 | prefoldin subunit 2 (Q16LV2) | 7.00E-11 | prefoldin chaperone | 1.18 | 7.32E-04 |
| SL.C1.1.105.Contig3 | c-Myc-binding protein (Q8ROR8) | 6.00E-12 | regulation of MYC | 1.24 | 7.47E-04 |
| SJWF04AAB.scf | mitochondrial ribosomal protein L32 (Q2BC5) | 4.00E-29 | mitochondrial | 1.25 | 1.01E-04 |
| SL.C1.12.1258.Contig1 | succinate-ubiquinone reductase membrane anchor subunit, mitochondrial (Q9CV5) | 3.00E-28 | mitochondrial | 1.28 | 3.24E-04 |
| SL.C1.10.1.1128.Contig1 | small nuclear ribonucleoprotein G (snRNPG) (Q56F16) | 1.00E-32 | RNA processing | 1.29 | 4.57E-04 |
| SL.C1.31.341.Contig1 | mitochondrial import receptor subunit TOM7 (Q4TC18) | 1.00E-11 | mitochondrial | 1.29 | 9.18E-05 |
| SL.C1.77.737.Contig1 | putative Deoxyxystydylate deaminase (Q16G52) | 7.00E-72 | nucleic acid metabolism | 1.29 | 4.87E-04 |
| SL.C1.55.15.SJW108DQ.scf | ATP synthase-like protein, mitochondrial (Q0PXW9) | 5.00E-07 | mitochondrial | 1.41 | 2.12E-04 |
| SL.C4.51.6000.Contig1 | DNA polymerase v (Q17DX3) | 1.00E-09 | DNA replication | 1.51 | 8.54E-06 |
| SL.C9.19.997.Contig1 | septin-2 (Q29BR7) | 1.00E-13 | cell division | 1.55 | 2.00E-05 |
| SL.C1.9.9422.Contig1 | ribonuclease H (Q56A61) | 1.00E-11 | nucleic acid metabolism | 1.74 | 2.11E-04 |
| SL.C1.41.1644.SJW108AB.scf | cytochrome b (cytB) (Q6WTV3) | 2.00E-59 | mitochondrial | 2.14 | 5.13E-05 |
| SL.C1.0.041.SJWH03BOX.scf | cytochrome oxidadse subunit I (COI) (Q1PLX5) | 1.00E-73 | mitochondrial | 2.21 | 9.20E-05 |
| SL.C1.31.3197.Contig1 | mitochondrial protein L22, mitochondrial (Q29K4) | 2.00E-55 | mitochondrial | 2.81 | 1.06E-05 |
| SL.C1.41.1449.Contig1 | bone morphogenetic protein (Q16JR6) | 9.00E-51 | intercellular signaling | 2.92 | 2.27E-04 |
| SL.C1.41.1435.Contig1 | polypeptide (sRNA([-]) virus) (Q38048) | 4.00E-15 | virus | 3.87 | 4.71E-04 |
| SL.C1.18.1832.Contig1 | putative structural protein of *S. invicta* virus 2 (ASH89) | 2.00E-81 | virus | 2.88E-06 |
| SL.C1.28.2823.Contig1 | polypeptide (sRNA([-]) virus) (Q38048) | 1.00E-33 | virus | 12.57 | 1.12E-04 |
| SL.C1.42.4295.Contig1 | putative structural protein of *S. invicta* virus 2 (ASH889) | 1.00E-120 | virus | 14.30 | 3.75E-09 |
| SL.C2.52.5211.Contig1 | (pv4)Non-capsid protein (sRNA([-]) virus) (Q11437) | 9.00E-11 | virus | 51.98 | 1.14E-11 |
| SL.C1.66.6100.Contig1 | non-structural protein of *S. invicta* virus 2 (ASH892) | 1.00E-126 | virus | 87.39 | 6.73E-10 |

*Threshold, F<e-5. See Table S4 for all BLASTX matches with E<1 for all 91 differentially expressed genes.

*Excludes Ensembl Api gene predictions. Accession numbers of best matches (TrEMBL, Swiss-Prot, or Ensembl databases) are shown in parentheses.

*Expression ratios are based on averages for all monogyne (M) and polygyne (P) BB workers in the 40 study colonies. Elevated expression in P workers relative to M workers is highlighted with bold italics.

*P-values from ANOVA calculations are averages for genes represented by more than one significantly differentially expressed clone on the microarray (10% FDR).
were extracted from the images using GenePix software. After scanning, bad spots were flagged and the background-subtracted median foreground values were used as the intensity levels in the subsequent analysis. All spots with a positive intensity were considered for the subsequent analyses (i.e., no threshold filtering was used). Raw intensity data were converted to normalized log2 ratios using “print-tip specific” loess normalization (within arrays; marray Bioconductor package, R [66]).

Selected gene expression results were confirmed using qRT-PCR (see Text S1 and Table S1). Primers used for qRT-PCR are listed in Table S5.

Data Analysis

For the genotypic analysis, we tested for differential expression of each gene between samples of BB and Bb workers in the 20 polygyne colonies using a 2-factor mixed-model ANOVA of the form:

\[ Y = \mu + \sim \ \text{BATCH} + \ \text{GENOTYPE} + \epsilon, \]

where \( Y \), representing the reference/sample log-transformed ratio for a spot, is the sum of effects. The symbol \( \mu \) represents the overall average log-transformed ratio for a given spot over all experiments. BATCH is a random effect (denoted by \( \sim \)) with two levels, batch_2004 and batch_2006, that accounts for the variation between profiles that appeared similar to those of a small number of samples. However, these samples yielded expression ratios using “print-tip specific” loess normalization (within arrays; marray Bioconductor package, R [66]).

For the social form analysis, we tested for differential expression of each gene between BB and Bb workers in the 20 polygyne colonies by using the same 2-factor mixed-model ANOVA, but with the variable SOCIAL FORM (monogyn or polygyne) replacing the variable GENOTYPE.

Statistical significance of the expression differences detected by the ANOVA calculations was additionally evaluated by means of non-parametric Mann-Whitney tests conducted on the normalized, batch-adjusted data (Figure S1). Expression differences between polygyne workers of different Gp-9 genotype (BB, Bb) were confirmed to be highly significant for all 39 genes identified by the ANOVA (all \( P < 0.002 \)), as were expression differences between BB workers of different social form for all 91 genes identified by the ANOVA (all \( P < 0.002 \)). In contrast, among the 39 genes influenced by Gp-9 genotype, only seven (18%) showed significantly different expression between monogyne and polygyne workers with the BB genotype (0.001 \(< P < 0.041 \)), while among the 91 genes influenced by social form, only three (3.3%) showed significantly different expression between BB and Bb workers of the polygyne form (0.0001 \(< P < 0.047 \)) (see Figure S1). Given the large number of these tests performed, some 5% of the significant results are presumed to represent Type I errors.

Expression data were hierarchically clustered and examined using Cluster and Treeview [68]. We also performed SOM (self-organizing map) clustering of the experimental samples (by array) for both the genotype and social form comparisons (data not shown). For the genotype comparisons, the samples clustered into two distinct groups according to genotype (BB and Bb) and no additional group was uncovered. Similarly, for the social form comparison, all the monogyne samples clustered together while the polygyne samples separated into two groups, those with high and those with low levels of viral gene expression. Because this analysis did not reveal any striking new patterns, the results are not presented in detail.

Annotation of Differentially Regulated Genes

Because previous annotations of the genes represented on the fire ant microarray [30] may be outdated, we performed new similarity searches against the non-redundant protein sequence database using the BLASTX algorithm [69,70]. All comparisons were performed on the Blast Network Service provided by the Swiss Institute for Bioinformatics (release July 17, 2007). The default settings were used with an E-value threshold of 1e-5, except where otherwise indicated. The accession number of the best match for each gene is reported in Tables 1 and 2, except when it was an Apis mellifera gene derived from the Ensembl automatic annotation. In this case, we chose the next best hit, because little is known about gene function in A. mellifera, and the genome of this species has been removed from the current Ensembl releases. All BLASTX matches with \( E \leq 1 \) (but limited to the top 20) are listed in Tables S3 and S4. Each fire ant gene was also manually assigned to a descriptive category (Tables 1 and 2 and Text S1).

The category putatively encapsulates the general function of each gene and is subjectively derived from examining the SwissProt or Genbank database entries of the five best hits (all \( E < 1e^{-5} \)), with an emphasis on Gene Ontology, Interpro, and PANTHER annotations. To determine which categories were overrepresented in each set of differentially expressed genes, we used a one-tailed hypergeometric test implemented in R [71,72].

Gene expression data meet Minimum Information About a Microarray Experiment (MIAME) standards and have been deposited at Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo/) with accession number GSE11694.

Supporting Information

Figure S1 Results of Mann-Whitney statistical tests for gene expression differences. Found at: doi:10.1371/journal.pgen.1000127.s001 (0.03 MB PDF)
Table S1 | Comparison of qRT-PCR and microarray expression ratios. Found at: doi:10.1371/journal.pgen.1000127.s002 (0.02 MB PDF)

Table S2 | List of gene categories significantly overrepresented among differentially expressed genes in the genotype and social form comparisons. Found at: doi:10.1371/journal.pgen.1000127.s003 (0.02 MB PDF)

Table S3 | BLASTX matches for genes differentially expressed between polygyne Gp-9 BB and Gp-9 Bb workers. Found at: doi:10.1371/journal.pgen.1000127.s004 (0.10 MB XLS)

Table S4 | BLASTX matches for genes differentially expressed between monogyne and polygyne Gp-9 BB workers. Found at: doi:10.1371/journal.pgen.1000127.s005 (0.18 MB XLS)

Table S5 | Genes, primer sequences, and primer concentrations used for qRT-PCR verification of microarray expression data. Found at: doi:10.1371/journal.pgen.1000127.s006 (0.02 MB PDF)

References

1. Hansen TF (2006) The evolution of genetic architecture. Annu Rev Ecol Evol Syst 37: 123–157.
2. Greenbaum RJ (2004) E pluribus unum, ex uno plura: quantitative and single-gene perspectives on the study of behavior. Annu Rev Neurosci 27: 79–105.
3. Orr HA (2008) The genetic theory of adaptation: a brief history. Nat Rev Genet 9: 119–127.
4. Orr HA, Coyne JA (1992) The genetics of adaptation - a reassessment. Am Nat 140: 725–742.
5. Sokolowski MB (2001) Drosophila genetics meets behaviour. Nat Rev Genet 2: 879–890.
6. Wolf JB, Brodie ED, Cheverud JM, Moore AJ, Wade MJ (1998) Evolutionary consequences of indirect genetic effects. Trends Ecol Evol 13: 64–69.
7. Wolf JB, Brodie ED, Moore AJ (1999) Interacting phenotypes and the evolutionary process. I. Selection resulting from social interactions. Am Nat 153: 254–266.
8. Robinson GE, Grozinger CM, Whitfield CW (2003) Sociogenomics: social life in molecular terms. Nat Rev Genet 6: 257–270.
9. Linksvayer TA, Wade MJ (2005) The evolutionary origin and elaboration of sociality in the aculeate hymenoptera: Maternal effects, social effects, and heterochrony. Q Rev Biol 80: 317–336.
10. Linksvayer TA (2006) Direct, maternal, and sibling genetic effects on individual and colony traits in an ant. Evolution 60: 2532–2561.
11. Nederlof AM, Michiel RE (2006) The evolutionary origin of an altruistic gene. Mol Biol Evol 23: 1480–1494.
12. Moore AJ, Brodie ED, Wolf JB (1997) Interacting phenotypes and the evolutionary process. I. Direct and indirect genetic effects of social interactions. Evolution 51: 1352–1362.
13. Gardner A, West SA, Barton NH (2007) The relation between multilocus population genetics and social evolution theory. Am Nat 169: 207–226.
14. Krieger MJ, Ross KG (2002) Identification of a major gene regulating complex social behavior. Science 295: 328–332.
15. Krieger MJ, Ross KG (1999) Major gene effects on phenotype and fitness: the relative roles of Pgn-3 and Gp-9 in introduced populations of the fire ant Solenopsis invicta. J Evol Biol 12: 672–680.
16. Krieger MJ, Ross KG (1995) Phenotypic Plasticity and Cultural Transmission of Alternative Social Organizations in the Fire Ant Solenopsis invicta. Behav Ecol Sociobiol 33: 121–129.
17. Krieger MJ, Ross KG (1995) Gene by Environment Interaction - Effects of a Single-Gene and Social-Environment on Reproductive Phenotypes of Fire Ant Queen Genotypes. Functional Ecology 9: 667–676.
18. DeHeer CJ, Goodman MAD, Ross KG (1999) Queen dispersal strategies in the multiple-queen form of the fire ant Solenopsis invicta. Am Nat 153: 660–675.
19. Keller L, Ross KG (1999) Selfish genes: a green beard in the red fire ant. Nature 394: 573–575.
20. Keller L, Ross KG (1995) Genetic control of social organization in an ant. Proc Natl Acad Sci USA 95: 14122–14127.
21. Gotzek D, Ross KG (2007) Genetic regulation of colony social organization in fire ants: an integrative overview. Q Rev Biol 82: 201–226.
22. Ross KG, Keller L (2002) Experimental conversion of colony social organization by manipulation of worker genotype composition in fire ants (Solenopsis invicta). Behav Ecol Sociobiol 51: 287–295.
23. Hallar BL, Krieger MJ, Ross KG (2007) Potential cause of lethality of an allele implicated in social evolution in fire ants. Genetica 131: 69–79.
24. Tschanke WR (2006) The fire ants. Cambridge, MA: Harvard Univ. Press.
25. Ross KG, Keller L (1995) Ecology and Evolution of Social-organization - Insights from fire ants and other highly eusocial insects. Annu Rev Ecol Syst 26: 631–656.
26. Shoemaker DD, Ross KG (1996) Effects of social organization on gene flow in the fire ant Solenopsis invicta. Nature 383: 613–616.
27. Ross KG, Shoemaker DD (1993) An unusual pattern of gene flow between the 2 social forms of the fire ant Solenopsis invicta. Evolution 47: 1595–1605.
28. Goodman MAD, Ross KG, Atsumu RA (2000) A formal assessment of gene flow and selection in the fire ant Solenopsis invicta. Evolution 54: 606–616.
29. Wang J, Jemielity S, Wu P, Wurm Y, Graff J, et al. (2007) An annotated cDNA library and microarray for large-scale gene-expression studies in the ant Solenopsis invicta. Genome Biol 8: R9.
30. McGraw LA, Gibson G, Clark AG, Wolfner MF (2004) Genes Regulated by Mating, Sperm, or Seminal Fluids in Mated Female Drosophila melanogaster. Curr Biol 14: 1509–1514.
31. Harbison ST, Chiang S, Kamdar KP, Mackay TFC (2003) Quantitative genomics of starvation stress responses in Drosophila. Genome Biol 6.
32. Morozova TV, Anholt RRH, Mackay TFC (2005) Quantitative trait locus mapping of adult alcohol intake in Drosophila melanogaster. Genes Brain Behav 4: 655–665.
33. Mascher MC (2001) Levels of selection and the evolution of social organization [PhD dissertation]: University of Georgia. vi, 104 leaves.
34. Rice WR (1987) Genetic hitchhiking and the evolution of reduced genetic activity of the V sex chromosome. Genetics 116: 161–167.
35. Backstrom D (2006) A dynamic view of sex chromosome evolution. Curr Opin Genet Dev 16: 578–585.
36. Charnov EL, Charlesworth B, Marais G (2005) Steps in the evolution of heteromorphic sex chromosomes. Heredity 95: 118–129.
37. Ross KG (1997) Multilocus evolution in fire ants: effects of selection, gene flow and recombination. Genetics 145: 961–974.
38. Krieger MJ, Ross KG (2005) Molecular evolutionary analyses of the odorant-binding protein gene Gp-9 in fire ants and other Solenopsis species. Mol Biol Evol 22: 2090–2103.
39. Mace RM, Porter SD (1996) Comparison of polygyny and monogamy in the fire ant Solenopsis invicta (Hymenoptera: Formicidae) populations densities. Ann Entomol Soc Am 89: 535–543.
40. Oh DH, Valles SM, Pereira RM (2004) Prevalence of Thelohania solanacearum (Microsporidia : Thelohaniidae) infection in monogyne and polygyne red imported fire ants (Hymenoptera : Formicidae). Environ Entomol 33: 340–345.
41. Valles SM, Strong CA, Hashimoto Y (2007) A new positive-strand RNA virus with unique genome characteristics from the red imported fire ant, Solenopsis invicta. Virology 365: 457–463.
42. Raj PA, Durmus AO (2002) Current status of defensive and their role in innate and adaptive immunity. FEBS Microbiol Lett 206: 9–18.
43. Goodman MAD, Sankovich KA, Kovacs JL (2007) Genetic and morphological variation over space and time in the invasive fire ant Solenopsis invicta. Biol Invasions 9: 571–584.
44. Goodman MAD, Mack PD, Pearse DE, Ross KG (1999) Effects of a single gene on worker and male body mass in the fire ant Solenopsis invicta (Hymenoptera : Formicidae). Annu Entomol Soc Am 92: 563–570.
48. Edwards AC, Rollmann SM, Morgan TJ, Mackay TF (2006) Quantitative genomics of aggressive behavior in Drosophila melanogaster. PLoS Genet 2: e154.

49. Roberge C, Einaum S, Guerlely H, Bernatchez L (2006) Rapid parallel evolutionary changes of gene transcription profiles in farmed Atlantic salmon. Molluscan Genet 15: 9–20.

50. Jarvis ED, Scharff C, Grossman MR, Ramos JA, Notebohm F (1998) For whom the bird sings: Context-dependent gene expression. Neuron 21: 775–788.

51. Whitfield CW, Cziko AM, Robinson GE (2003) Gene expression profiles in the brain predict behavior in individual honey bees. Science 302: 296–299.

52. Burmester SS, Jarvis ED, Fernald RD (2005) Rapid behavioral and genomic responses to social opportunity. PLoS Biol 3: 1996–2004.

53. Cummings ME, Larkins-Ford J, Reilly CR, Ramsey M, et al. (2008) Sexual and social stimuli elicit rapid and contrasting genomic responses. Proc Biol Sci 275: 393–402.

54. Linksvayer TA (2007) Ant Species Differences Determined by Epistasis between Brood and Worker Genomes. PLoS ONE 2: e694.

55. Rauter CM, Moore AJ (2002) Evolutionary importance of parental care performance, food resources, and direct and indirect genetic effects in a burying beetle. Proc Natl Acad Sci USA 99: 6829–6832.

56. Rauer CM, Moore AJ (2002) Evolutionary importance of parental care performance, food resources, and direct and indirect genetic effects in a burying beetle. J Evol Biol 15: 407–417.

57. Andam GV, Cosudes A, Fournik MK, Page RE (2006) Complex social behaviour derived from maternal reproductive traits. Nature 439: 76–78.

58. Grozinger CM, Fan Y, Hoover SF, Winston ML (2007) Genome-wide analysis reveals differences in brain gene expression patterns associated with caste and reproductive status in honey bees (Apis mellifera). Mol Ecol 16: 4837–4848.

59. Toth AL, Valera K, Newman TC, Miguez FE, Hutchison SK, et al. (2007) Wasp gene expression supports an evolutionary link between maternal behavior and eusociality. Science 318: 441–444.

60. Dawkins R (1999) The extended phenotype: the long reach of the gene. New York: Oxford University Press.

61. Roberge C, Einaum S, Guerlely H, Bernatchez L (1997) Survey for Pathogens of Fire Ants, Solenopsis-Spp (Hymenoptera-Formicidae) in Southeastern United-States. Fla Entomol 60: 275–279.

62. Shoemaker DI, DeHeer CJ, Krieger MB, Ross KG (2006) Population genetics of the invasive fire ant Solenopsis invicta (Hymenoptera: Formicidae) in the United States. Ann Entomol Soc Am 99: 1213–1233.

63. Jouvenaz DP, Allen GE, Banks WA, Wojcik DP (1977) Survey for Pathogens of Fire Ants, Solenopsis-Spp (Hymenoptera-Formicidae) in Southeastern United-States. Fla Entomol 60: 275–279.

64. Randolph JB, Waageber AS (1997) Stability, specificity and fluorescence brightness of multiply-labeled fluorescent DNA probes. Nucleic Acids Res 25: 2923–2929.

65. Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, et al. (2004) Bioconductor: open software development for computational biology and bioinformatics. Genome Biol 5: R80.

66. Benjamini Y, Hochberg Y (1995) Controlling the False Discovery Rate: a Practical and Powerful Approach to Multiple Testing. J R Stat Soc Ser B 57: 289–300.

67. Etien MB, Spellman PT, Brown PO, Botstein D (1998) Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci U S A 95: 14863–14868.

68. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215: 403–410.

69. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, et al. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25: 3389–3402.

70. Ihaka R, Gentleman R (1996) R: A language for data analysis and graphics. J Comp Graph Stat 5: 299–314.

71. Ihaka R, Personnaz L, Tating I, Potier MC (2007) Enrichment or depletion of a GO category within a class of genes: which test? Bioinformatics 23: 401–407.