MOLECULAR ECOLOGY

Molecular Ecology (2013) **22**, 3797–3813 doi: 10.1111/mec.12346

A simple genetic basis for complex social behaviour mediates widespread gene expression differences

MINGKWAN NIPITWATTANAPHON,*† JOHN WANG,*‡ MICHIEL B. DIJKSTRA* and LAURENT KELLER*

*Department of Ecology and Evolution, University of Lausanne, Sorge, le Biophore, CH-1015 Lausanne, Switzerland, †Department of Genetics, Faculty of Science, Kasetsart University, Bangkok, Thailand, ‡Academia Sinica, Biodiversity Research Center, Nangang Taipei 115, Taiwan

Abstract

A remarkable social polymorphism is controlled by a single Mendelian factor in the fire ant *Solenopsis invicta*. A genomic element marked by the gene *Gp-9* determines whether workers tolerate one or many fertile queens in their colony. *Gp-9* was recently shown to be part of a supergene with two nonrecombining variants, SB and Sb. SB/SB and SB/Sb queens differ in how they initiate new colonies, and in many physiological traits, for example odour and maturation rate. To understand how a single genetic element can affect all these traits, we used a microarray to compare gene expression patterns between SB/SB and SB/Sb queens of three different age classes: 1-day-old unmated queens, 11-day-old unmated queens and mated, fully reproductive queens collected from mature field colonies. The number of genes that were differentially expressed between SB/SB and SB/Sb queens of the same age class was smallest in 1-day-old queens, maximal in 11-day-old queens and intermediate in reproductive queens. Gene ontology analysis showed that SB/SB queens upregulate reproductive genes faster than SB/Sb queens. For all age classes, genes inside the supergene were overrepresented among the differentially expressed genes. Consistent with the hypothesized greater number of transposons in the Sb supergene, 13 transposon genes were upregulated in SB/Sb queens. Viral genes were also upregulated in SB/Sb mature queens, consistent with the known greater parasite load in colonies headed by SB/Sb queens compared with colonies headed by SB/SB queens. Eighteen differentially expressed genes between reproductive queens were involved in chemical signalling. Our results suggest that many genes in the supergene are involved in regulating social organization and queen phenotypes in fire ants.

Keywords: fire ants, *Gp-9*, maturation, monogyne, polygyne, queen, queen odour, social form, *Solenopsis invicta*, supergene

Received 8 January 2013; revision received 29 March 2013; accepted 8 April 2013

Introduction

Differences in social behaviour can be caused by sequence variation or differential expression of the underlying genes (Robinson & Ben-Shahar 2002; Vosshall 2007). When differences in social behaviour have a genetic basis, they typically depend on variation in many genes, as for example the mating song and mating behaviour in fruit flies and crickets (Gleason 2005; Ruedi & Hughes 2008; Ellison *et al.* 2011). However, in some cases, variation in a single gene can have a pronounced effect on social behaviour. For example, males carrying certain *fruitless* mutations exhibit homosexual courtship in *Drosophila* (Goodwin *et al.* 2000), and an amino acid substitution in the *npr-1* gene coding for a neuropeptide Y receptor changes foraging behaviour in *Caenorhabditis elegans* (De Bono & Bargmann 1998).

The fire ant *Solenopsis invicta* provides one of the best-studied examples of a social polymorphism that is

Correspondence: Mingkwan Nipitwattanaphon, Fax: +41 21 692 4165; E-mail: mingkwan.nipitwattanaphon@gmail.com

perfectly explained by a single Mendelian element (Keller & Ross 1998; Ross & Keller 1998). In *S. invicta*, a single genomic element containing the odorant-binding protein *Gp-9* determines whether workers tolerate only one fertile queen (in the 'monogyne' social form) or many fertile queens ('polygyne' social form) in their colony (Ross & Keller 1998; Krieger & Ross 2002). Colonies that contain only $Gp\text{-}9^{BB}$ workers accept only one fertile queen, while colonies in which >15% of workers are $Gp\text{-}9^{Bb}$ will invariably accept multiple fertile queens, but only if these queens are $Gp\text{-}9^{Bb}$ (Ross & Keller 1998, 2002). Monogyne and polygyne colonies differ in many other aspects, such as the mode of colony founding (respectively, by a single queen without workers and by colony budding) and the level of between-colony aggression (greater in monogynes) (Ross & Keller 1995; DeHeer *et al.* 1999; Goodisman *et al.* 2000; DeHeer 2002). All differences between monogyne and polygyne colonies are perfectly associated with a suite of morphological, physiological and life history differences between individuals with different *Gp-9* genotypes, including the amount of fat accumulated by maturing queens, queen fecundity (Keller & Ross 1999), the proportion of saturated hydrocarbons on the queen's cuticle (Eliyahu *et al.* 2011), worker size (all greatest in $Gp\text{-}9^{BB}$ females) (Goodisman *et al.* 1999) and sperm count (greater in $Gp\text{-}9^{B}$ haploid males) (Lawson *et al.* 2012).

Recently, *Gp-9* was found to be one of approximately 616 genes in a supergene (Wang *et al.* 2013), which explains how this single odorant-binding protein can be associated with so many and so disparate traits, many of which are not implicated in odour. This supergene is a nonrecombining region spanning *c.* 13 Mb (55%) of one of the 16 chromosomes of *S. invicta*. This so-called social chromosome has two variants, the 'Social B' (SB) and 'Social b' (Sb) chromosomes named after the presence of a *B* or *b* allele at the *Gp-9* locus in its nonrecombining region. Previous work showed that females with two copies of the Sb chromosome (Sb/Sb) die soon after they eclose from the pupa (Ross 1997; Keller & Ross 1998; Ross & Keller 1998; Hallar *et al.* 2007). There are thus two types of reproductive queens: SB/SB queens individually heading monogyne colonies and SB/Sb queens jointly heading polygyne colonies.

To understand how the two variants of the social chromosome affect the phenotype and behaviour of queens, we used cDNA microarrays comprising 5956 genes (Wang *et al.* 2007) to compare gene expression of SB/SB and SB/Sb queens of three age classes: immature virgin queens collected 1 day after eclosion (hereafter 1-day-old queens), virgin queens collected 11 days after eclosion (11-day-old queens) and mated queens heading fully grown colonies collected from the field and raised about 8 months under controlled laboratory conditions (reproductive queens; these queens were mated and at least 17 months old, see Materials and methods). Our first aim was to determine whether there was an overrepresentation in the supergene of genes differently expressed between SB/SB and SB/Sb queens, because the supergene variants are thought to be the primary signal that set in motion the diverging gene cascades producing the differential phenotypes. Previous work reported such a pattern in 1-day-old queens (Wang *et al.* 2013). We therefore conducted a similar comparison for 11-day-old queens and reproductive queens. Additionally, we wanted to determine the similarity in the set of genes differentially expressed between the SB/SB and SB/Sb genotypes for queens and workers. Our second aim was to find genes that are associated with differences in maturation between SB/SB and SB/Sb queens, because previous work showed that SB/SB queens mature faster, initiate reproduction earlier and accumulate more fat than SB/Sb queens, enabling SB/SB queens to found colonies independently (Keller & Ross 1993, 1999; Ross & Keller 1998). A third aim was to test whether we would find a signature in terms of gene expression of the accumulation of transposable elements on the Sb chromosome, because Muller's ratchet and a reduced effective population size promotes the accumulation of transposons and other mildly deleterious mutations in nonrecombining genetic elements (e.g. Y chromosomes) (cf. Wang *et al.* 2013). Our fourth aim was to identify candidate genes that may account for odour differences between SB/SB and SB/Sb queens. Previous work showed that workers in polygyne colonies selectively eliminate SB/SB queens when they initiate reproduction, and that the discrimination is done on the basis of odours on the surface of the cuticle (Keller & Ross 1998). A study of cuticular hydrocarbon profiles also revealed that SB/SB reproductive queens have a higher proportion of saturated hydrocarbons than SB/Sb reproductive queens (Eliyahu *et al.* 2011). We thus predicted that genes implicated in chemical signalling would be differentially expressed between SB/SB and SB/Sb queens.

Our study is an important first step towards characterizing the pathways in queen physiology that regulate social form, as social form is an extended phenotype of the founding queen's supergene genotype (Keller & Ross 1998, 1999; Ross & Keller 1998; Ross *et al.* 1999; Goodisman *et al.* 2000; DeHeer 2002).

Materials and methods

Colony collection and maintenance

Solenopsis invicta colonies were collected in Athens, Georgia, USA, in March 2008 and 2009. We deduced

each colony's social form from the shape of its mound in the field (more pointed in monogynes), the size of the largest workers (larger in monogynes) and queen number; social form was always later confirmed by determining the *Gp-9* genotype of >50 alate queens per colony by PCR-RFLP (Krieger & Ross 2002) (polygyne colonies always included some SB/Sb individuals, while monogyne colonies only comprised SB/SB individuals). After transport to Lausanne, Switzerland, we reared them under standard conditions and diet (Jouvenaz *et al.* 1977) for 2–3 months before dividing them (see below).

Microarray samples

For our microarray experiments (below), we used virgin queens of both SB/SB and SB/Sb genotypes from polygyne colonies to control for possible genetic and environmental difference, for example, due to social form or different colonies. This was not possible for reproductive queens because SB/SB reproductive queens are never found in polygyne colonies; thus, we used SB/SB and SB/Sb reproductive queens from monogyne and polygyne colonies, respectively (more below). We needed to screen many queens to obtain sufficient queens of each genotype and age class. Thus, to maximize the yield in queens, we used the polygyne colonies to make queenless subcolonies consisting of approximately 10 000 haphazardly chosen workers with young brood, because reproductive queens partly inhibit the production of sexuals (Vargo & Fletcher 1986; Tschinkel 2006). We checked these subcolonies daily and transferred newly eclosed queens (0–24 h) to colony fragments with 300–500 workers but no brood. Brood was not included because the development of young queens is influenced by the presence of queen pheromone but not brood presence (Fletcher & Blum 1981, 1983). These virgin queens were allowed to mature for either 1 or 11 days and then snap-frozen in liquid nitrogen and transferred to −80 °C until DNA and RNA extraction. To avoid pseudoreplication and to minimize day effects, we used one SB/SB and one SB/Sb queen (always eclosed on the same day in the same subcolony) per colony of origin. Because polygyne SB/SB queens are eliminated by workers when they reach sexual maturity, it was laborious to obtain sufficient matching pairs of 11-day-old SB/SB and SB/Sb queens. We determined the *Gp-9* genotype with PCR-RFLP (Krieger & Ross 2002) of 1009 individuals to obtain sufficient numbers of SB/SB individuals (the following were genotypic frequencies: SB/Sb = 85.5%; SB/SB = 14.0%; Sb/Sb = 0.5%). Sb/Sb queens were rare because they typically die early in development (Keller & Ross 1998; Ross & Keller 1998; Hallar *et al.* 2007). Because we did not obtain enough Sb/Sb individuals to provide sufficient statistical power, they were not included in the study. Overall, we extracted RNA from one SB/SB and one SB/Sb queens for each of two age classes (1 day old and 11 days old) from each of 16 colonies (eight in 2008 and eight in 2009). There were thus 16 samples for a given genotype and age class.

To identify genes that are indicative of high fertility and high fecundity of queens, and to separate these genes from genes that regulate odour production, we also analysed highly fertile queens because SB/SB and SB/Sb queens always differ in odour, but SB/Sb are normally less fecund. Because it is impossible to mate queens in the laboratory, we collected reproductive queens from mature field colonies. Monogyne and polygyne colonies were kept for 3 months in the laboratory after the field collection. We then selected eight monogyne colonies with similar amounts of brood and workers. We also established 40 colonies with one polygyne queen, brood and haphazardly chosen workers from their native polygyne colony. This was done because monogyne colonies normally have a higher ratio of workers per queen, hence leading to a higher fecundity of queens in monogyne colonies. After about 5 months, we collected the eight SB/SB queens originating from monogyne colonies and eight SB/Sb queens from the polygyne colonies (we selected healthy colonies with comparable fecundity to the monogyne queens and containing all types of brood, hence avoiding unmated queens and diploid male-producing queens that are common in polygyne colonies; Ross & Fletcher 1986). The age of these queens was unknown. However, as mating flight occurs in May-June in Athens, Georgia, and because they were collected in March, they were at least 9 months old when collected. Given that colonies were kept 8 months in the laboratory before collecting queens, they were thus at least 17 months old (the maximum queen lifespan is about 6 years in *S. invicta*; Tschinkel 2006). Finally, while we cannot rule out that there is also an effect of social form (monogyne, polygyne) on gene expression in reproductive queens, at least social form alone did not affect gene expression in 1-day-old and 11-day-old queens (data not shown).

Overall, the number of individuals used for the microarray analyses was 16 SB/SB 1-day-old queens, 16 SB/Sb 1-day-old queens, 16 SB/SB 11-day-old queens, 16 SB/Sb 11-day-old queens, eight SB/SB reproductive queens and eight SB/Sb reproductive queens (see below for two samples that failed because of hybridization problems).

Molecular methods

We extracted nucleic acid from individual virgin queens by homogenizing their whole body in 2-mL tubes in

600 μL RLT buffer from the RNeasy kit (QIAGEN) with approximately eight ceramic beads for 1 min at maximum speed in a tissue homogenizer. To limit RNA extraction only to the 64 relevant samples (see above), we first genotyped all 1009 virgin queens as follows. We purified DNA by directly mixing a 6-μL aliquot of the RLT-buffered lysate with 200 μL of Lysis Master Mix from the Agencourt DNAdvance kit (Beckman Coulter) and then followed the remainder of the DNAdvance extraction protocol. Subsequently, we performed PCR-RFLP on this DNA to determine the genotype at *Gp-9* (Krieger & Ross 2002). For the 64 selected queens, we precipitated the remainder of the RLT-buffered lysate in 50% ethanol and extracted total RNA with the RNeasy kit (QIAGEN) following the manufacturer's protocol. We did not use the RNeasy kit for RNA extraction from the 16 reproductive queens, as their high protein and fat content inhibited the binding of RNA to the RNeasy column; instead, we followed a standard TRIzol RNA extraction protocol (Invitrogen), first homogenizing them in 1 mL TRIzol and then removing unwanted fractions with 160 μL chloroform. We assumed that this difference in extraction method between virgin queens and reproductive queens did not have pronounced effects on patterns of relative (as opposed to absolute) gene expression. For all samples, we digested DNA with the Turbo DNA-free kit (Life Technologies) and amplified mRNA 10 × to 100 × with the MessageAmp II kit (Life Technologies), following the manufacturers' protocol. We reverse-transcribed the amplified mRNA with random primers, labelled the resulting cDNA with a green (Cy3) dye and hybridized the cDNA against the microarray for 18–20 h at 64 °C before washing and scanning the slides, as described in more detail in the study by Ometto *et al.* (2011). Each sample was hybridized against a red-labelled (Cy5) standard reference that was reverse-transcribed from amplified mRNA from pooled total RNA from all castes and developmental stages in both social forms. The same total RNA pool was used to make the reference used in the studies by Wang *et al.* (2008); Wurm *et al.* (2010) and Ometto *et al.* 2011.

Microarray analysis

Our *S. invicta* microarray contains 11 024 cDNAs with a single PCR product representing an estimated 5956 unique genes (Wang *et al.* 2007). These cDNAs do not include 4108 cDNAs that were excluded from all analyses because their sequence did not match any transcripts from RNAseq on *S. invicta* adults of all castes and also did not match any sequence in the SwissProt and TrEMBL databases. These discarded cDNAs were probably not from *S. invicta* but from unknown pathogens and symbionts (see next section for details). We used two series of microarray slides that were printed in different years but that did not differ in layout or cDNA sequences: slides from 2007 and 2010 were used exclusively with samples from 2008 and 2009, respectively. All samples were hybridized against the standard reference (see above). Microarray slides were scanned using an Agilent Microarray Scanner, and the images were later transformed to the $\log_2$ of the ratio of the red to green signals with the GenePix Pro software (Applied Biosystems). The expression data for 1-day-old queens are the same as presented in the study by Wang *et al.* (2013). Expression data from a pair of 11-day-old SB/SB and SB/Sb queens from 2008 were excluded from all analyses because hybridization was poor, leaving 16 SB/SB 1-day-old queens, 16 SB/Sb 1-day-old queens, 15 SB/SB 11-day-old queens, 15 SB/Sb 11-day-old queens, eight SB/SB reproductive queens and eight SB/Sb mature queens. Background correction (with parameter 'normexp+offset = 50') and normalization ('NormalizedWithinArrays' with parameter 'Print-Tip Loess' followed by 'NormalizedBetweenArrays') were performed in the 'limma' package (Smyth 2004) for R 2.11.1 (R Development Core Team 2010). We fitted a general linear model to the spot-specific $\log_2$-transformed ratios of the normalized red to green signal, with the fixed factors 'Genotype' (SB/SB vs. SB/Sb), 'Age' (1-day-old queens, 11-day-old queens, reproductive queens), 'Slide series and year of collection' (see above) and the interaction 'Age × Genotype'. For comparisons between queens of the same age class, we only included the fixed factors 'Genotype' and 'Slide series and year of collection' without interactions. We used the limma function 'eBayes' to estimate a 'Bayesian moderated *t* statistic' for post hoc pairwise comparisons (Smyth 2004), setting the false discovery rate (FDR; Benjamini & Hochberg 1995) at 0.01 for the single *F* test per cDNA across all terms in the full model (Smyth 2004), and likewise at 0.01 for each post hoc pairwise moderated *t* test when testing for a single term. We only considered the *P* values from post hoc pairwise moderated *t* tests if the single limma *F* test per cDNA across all terms in the full model was significant. All *P* values are two-tailed unless otherwise specified.

Gene annotation

We used BLASTN to compare the cDNA sequences on our microarray to gene models derived from RNAseq on *S. invicta* queens, workers and males. We only considered alignments that fulfilled all of the following criteria: (i) cDNA and RNA sequences aligned over at least 100 successive base pairs; (ii) the alignment covered at least 40% of the cDNA sequence; and

(iii) sequence identity was at least 95%. For each RNA sequence that matched a cDNA according to these criteria, we obtained an annotation and gene ontology (GO) terms with Blast2GO (Conesa *et al.* 2005; Götz *et al.* 2008, 2011). Microarray cDNAs that did not match any RNAseq transcripts were compared against the *S. invicta* proteome using BLASTX (Wurm *et al.* 2011), and we retained the annotation and GO terms of the single best hit for each of these cDNAs. For the remaining cDNAs that neither matched RNAseq transcripts nor *S. invicta* proteome sequences (e.g. viral genes), we used the annotation and GO terms of the single best hit from a direct BLASTX comparison against the SwissProt (reviewed) and TrEMBL (unreviewed) databases. Thus, we assigned each cDNA to the single most likely corresponding gene, with an average of 1.9 (range 1–40) cDNAs per gene. We considered a gene to be differentially expressed if at least 1 cDNA corresponding to this gene was significantly differentially expressed. Different cDNAs from the same locus almost always showed qualitatively the same pattern of relative expression; rare exceptions to this general result might be due to alternative splicing as we found by RNA-seq analyses of the gene *doublesex* (M. Nipitwattanaphon, unpublished). We did an enrichment analysis on the GO terms of such differentially expressed genes with the R package 'TopGO' (Gentleman *et al.* 2004; Alexa & Rahnenführer 2010).

Real-time quantitative reverse transcription PCR (qRT-PCR)

We verified the expression differences for six desaturase and three elongase genes in 16 reproductive queens with qRT-PCR (eight SB/SB and eight SB/Sb reproductive queens). Half of these samples (four SB/SB and four SB/Sb reproductive queens) were samples used in our microarray experiment, but the other half were independent samples not used in the microarray experiment.

We used TaqMan Reverse Transcription Reagents (Life Technologies) to synthesize cDNA from 1 μg of the total RNA of individual queens with the following reagents: 5 μL of 10× RT buffer, 2.5 μL of 50 mM random hexamers, 1 μL of 25 mM dNTPs, 1 μL of 25 nM RNase inhibitor, 1 μL of reverse transcriptase (200 U) and water to a final volume of 50 μL. This mixture was incubated at 48 °C for 30 min and then heat-inactivated at 95 °C for 5 min. After the reverse transcription reaction, we added 30 μL of 10 mM Tris to adjust the volume to 80 μL. A fourfold dilution of these cDNA samples was used for the quantitative PCR.

The qRT-PCR amplification mixtures were made by combining 2 μL of diluted cDNA with 5 μL of Power SYBR Green PCR Master Mix (Applied Biosystems), 300 nM of each gene-specific forward and reverse primer (Table S5) and water to a final volume of 10 μL. We performed MAFFT multiple sequence alignment of 11 desaturase and 14 elongase candidate genes before selecting six desaturase and three elongase genes for designing gene-specific primers and qRT-PCR validation. Reactions were run in triplicate on an ABI PRISM 7900HT Sequence Detector (Applied Biosystems) using the default parameters defined by the manufacturer.

Results from Sequence Detection Systems software (Applied Biosystems; fluorescence intensity default threshold set at 0.2) were exported as tab-delimited files into qBasePLUS (Hellemans *et al.* 2007) for subsequent analysis. Sixteen amplified RNA samples (eight each for SB/SB and SB/Sb reproductive queen samples) were analysed for each gene, including the two control genes. We selected *GAPDH* and *RpS9* as controls because they exhibited no differential expression in this and other microarray and qRT-PCR studies (Wang *et al.* 2008). Outliers within each set of reaction triplicates were filtered (max $dC_t = 0.5$), and relative gene expression levels were calculated with PCR efficiency correction and reference gene normalization to the control genes. Statistical analysis (Wilcoxon test) of relative gene expression levels was performed in R.

Results

SB/SB queens mature faster than SB/Sb queens

Age class affected the expression of a greater number (4386) of genes than genotype (1028). Among these 4386 age class-dependent genes, 2033 were differentially expressed between 1-day-old and 11-day-old queens, 3744 between 11-day-old queens and reproductive queens, and 4020 between 1-day-old queens and reproductive queens. Only 1028 unique genes were significant for genotype, that is, differentially expressed between SB/SB queens compared with SB/Sb queens, irrespective of age class (Fig. 1a). The greater effect on gene expression of age class compared with genotype was confirmed by a principal component analysis (without scaling to variance), where age class was strongly correlated with PC1 and PC2, while genotype was only weakly correlated with PC1 and PC2. PC1 and PC2 jointly explained 90% of the variance in gene expression across all samples (Fig. 1b). The interaction Genotype × Age was significant for 550 genes, meaning that these genes were upregulated or downregulated faster or slower as queens became older for SB/SB queens than for SB/Sb queens, or that they were upregulated during development for queens of one genotype but downregulated for queens of the other genotype. Because our statistical analyses never gave qualitatively

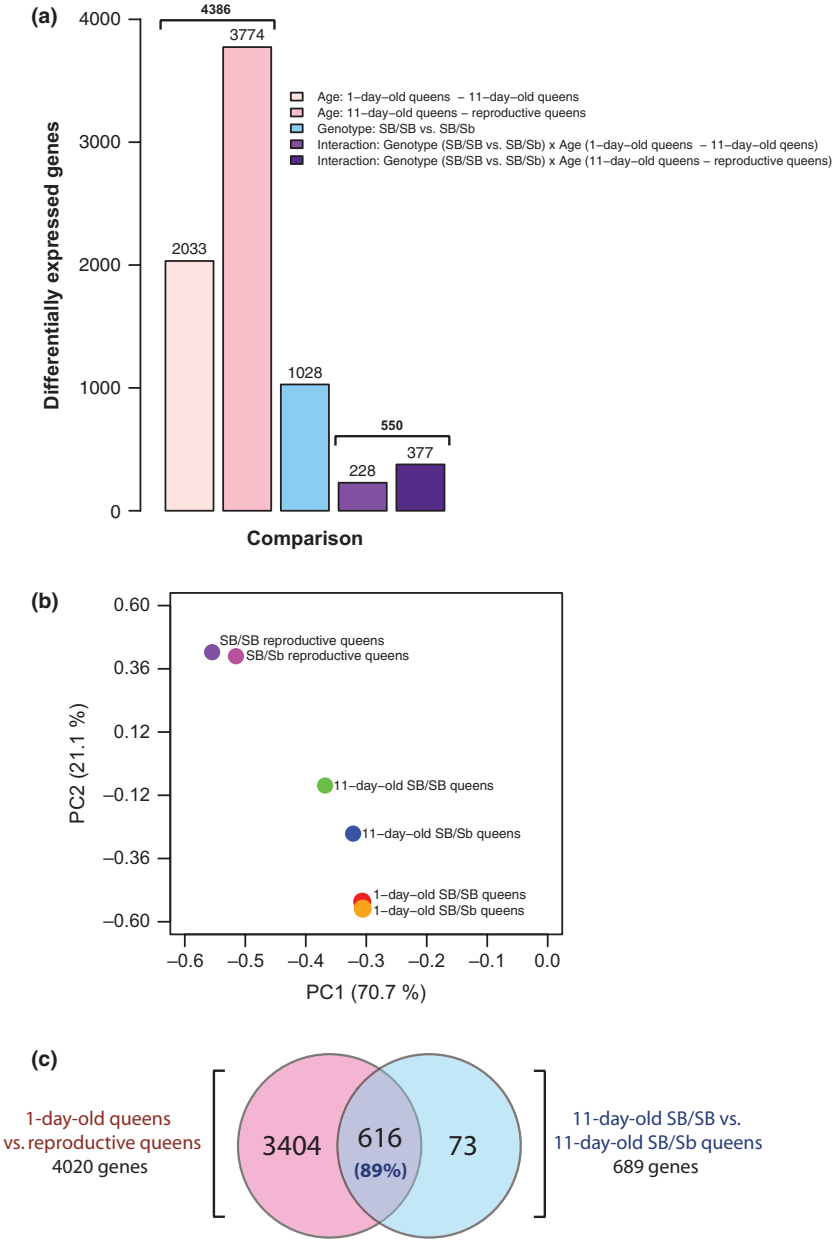

Fig. 1 Genes whose expression was different between Genotype and Age are highly associated. (a) Age class affected the expression of more genes than Genotype did. Bars represent the number of differentially expressed genes in post hoc moderated *t* tests (FDR = 0.01). (b) Age explains more between-queen variation in gene expression than Genotype does. Principal components 1 and 2 jointly explained 91.8% of the variance. (c) Most (89%) genes differentially expressed between genotypes were also differentially expressed between age classes. Genes differentially expressed between 1-day-old queens and reproductive queens were strongly and significantly overrepresented among genes differentially expressed between 11-day-old SB/SB and SB/Sb queens. The total number of genes on the microarray was 5956. Binomial test with H_0: $P_{success} = 4020/5956$; $n_{success} = 616$, $n_{trials} = 689$, $P < 10^{-5}$.

different results when the interaction was omitted (data not shown), we henceforth only give results on genes that were significant for age class and/or genotype when the GLM did not include the interaction term.

The number of differentially expressed genes between SB/SB and SB/Sb queens of the same age class was smallest in 1-day-old queens (38 genes), reached a maximum (689) in 11-day-old queens and decreased again (to 295 genes) for reproductive queens. Of the 689 genes that were differentially expressed between 11-day-old SB/SB and SB/Sb queens, 616 genes (89%) were also differentially expressed between 1-day-old queens and reproductive queens (Fig. 1c). This percentage overlap was significantly greater than expected by chance (binomial test with $P_{success} = 4020/5956$, $P < 10^{-5}$; see Fig. 1c). These results are consistent with SB/SB queens maturing faster and accumulating more fat than SB/Sb queens (Keller & Ross 1999). As expected, genes that were upregulated in all reproductive queens compared with all 1-day-old queens, irrespective of genotypes, were typically (404 of 429 genes) also upregulated in 11-day-old SB/SB queens compared with 11-day-old SB/Sb queens (Fig. S1, Table S1). Analogously, genes that were downregulated in reproductive queens compared with 1-day-old queens were typically (181 of 187 genes) also downregulated in 11-day-old SB/SB queens

compared with 11-day-old SB/Sb queens (Fig. S1, Table S2). In other words, 11-day-old SB/SB queens tended to upregulate the same genes as both SB/SB and SB/Sb reproductive queens (and vice versa for downregulated genes), while 11-day-old SB/Sb queens tended to upregulate the same genes as 1-day-old SB/SB and SB/Sb queens (and vice versa for downregulated genes).

There was considerable overlap in significant GO terms when separate enrichment analyses were performed on the set of genes that were upregulated in reproductive queens compared with 11-day-old queens and the set of genes that were upregulated in SB/SB queens compared with SB/Sb queens. In both enrichment analyses, GO terms implicated in reproduction in model organisms were significant (e.g. mitosis, genes localized in the nucleus, ubiquitin-dependent protein catabolic processes, nuclear mRNA splicing and proteosome complex, Table 1). Analogously, there was considerable overlap in significant GO terms when separate enrichment analyses were performed on the set of genes that belonged to the following categories: (1) downregulated in reproductive queens compared with 11-day-old queens; (2) downregulated in 11-day-old queens compared with 1-day-old queens; or (3) downregulated in SB/SB queens compared with SB/Sb queens. In each of analyses (1)–(3), at least some GO terms associated with oxidative metabolism were significant (e.g. mitochondrial genes, proton or electron transport, ubiquinone biosynthetic process and oxidation-reduction processes, Table 1). These results indicate that in terms of gene expression, SB/SB 11-day-old queens resemble reproductive queens more than SB/Sb individuals do, but that both SB/SB and SB/Sb queens eventually upregulate reproductive genes and downregulate oxidative metabolism genes as they become fully mature. Notably, the level of expression of the queen-specific paralog *vitellogenin-3*, which has peak expression in mature/reproductive queens, was nine times higher in 11-day-old SB/SB than in SB/Sb queens.

Despite the fact that SB/SB queens accumulate fat faster than SB/Sb queens, no GO terms implicated in fat storage (e.g. lipid storage, lipid particle, fat cell proliferation, fat body development) were significant in our enrichment analyses. Moreover, the only individual genes that were significant and obvious candidates for regulating fat storage were two homologs of *lipid storage droplets' surface-binding protein 2* (upregulated 2–4 times in SB/SB queens).

Genes inside the supergene are overrepresented among differentially expressed genes

For all age classes, genes inside the supergene were significantly overrepresented among genes that were differentially expressed between SB/SB and SB/Sb queens. Four of 38 genes in 1-day-old queens (hypergeometric test, $P = 0.005$), 42 of 689 genes in 11-day-old queens ($P = 0.006$) and 25 of 295 genes in reproductive queens ($P < 0.001$) lie inside the supergene. Genes in the supergene were overrepresented among genes that were upregulated in SB/Sb queens ($P = 0.003$ for 1-day-old queens; $P = 0.002$ for 11-day-old queens; $P = 0.0004$ for reproductive queens), but not among genes downregulated in SB/Sb queens ($P > 0.05$ for all three age classes) (Table 2; Fig. S2). However, there was no consistent trend towards either mostly upregulation or mostly downregulation of genes (irrespective of whether they lie inside or outside the supergene) in SB/Sb queens compared with SB/SB queens: there were significantly more upregulated genes in 1-day-old SB/Sb than in SB/SB queens (binomial test, $P < 0.001$), but significantly more downregulated genes in 11-day-old SB/Sb than in SB/SB queens ($P < 0.001$). The numbers of up- and downregulated genes did not differ significantly between SB/Sb and SB/SB reproductive queens ($P = 0.130$) (Table 2).

Genes that are consistently differentially expressed between queens and workers with SB/SB and SB/Sb genotypes

Figure 2 shows the correspondence between sets of genes differentially expressed between (1) 1-day-old SB/SB and SB/Sb queens, (2) 11-day-old SB/SB and SB/Sb queens, (3) SB/SB and SB/Sb reproductive queens and (4) adult SB/SB and SB/Sb workers whose age was not precisely known. Comparisons (1)–(3) are from the present study, while (4) is from the study by Wang *et al.* (2008), who used identical methods to measure gene expression in 20 replicates of 7–10 pooled adult polygyne workers per genotype, originally from the same population as in this study. Sixteen genes [1.7% of 934 genes differentially expressed in (1)–(4)] were differentially expressed between SB/SB and SB/Sb queens for each of the three age classes in the present study, and of these nine genes (1%) were always upregulated in SB/Sb females compared with SB/SB females in all four comparisons (centre of the Venn diagram, Fig. 2). Upregulation of these nine genes (*growth-arrest-specific protein 8*, *piggyBac transposon*, *nuclease harbi-like lc2_043034*, *BEL12-AG transposon*, *cytoplasmic tRNA2-thiolation protein 1-like*, *deoxyribonuclease tatD* and three uncharacterized genes) is therefore characteristic of SB/Sb females (Table S3). Thus, there was considerable agreement between results from workers and queens, which is also evident from the fact that among 38 genes differentially expressed between SB/SB and SB/Sb workers, 23 were also differentially

Table 1 Results of gene ontology enrichment analyses on age, genotype and genotype × age

Explanatory term	Direction	GO type*	GO identifier	GO term description	Annotated	Significant	Expected	*P*-value
Age (1-day-old – 11-day-old queens)	Up in 1-day-old queens of either genotype	BP	**GO:0015992**	**Proton transport**	**71**	**43**	**14.21**	**<0.001**
			GO:0006744	**Ubiquinone biosynthetic process**	**24**	**18**	**4.8**	**<0.001**
			GO:0006120	**Mitochondrial electron transport, NADH to ubiquinone**	**22**	**17**	**4.4**	**<0.001**
			GO:0055114	Oxidation-reduction process	335	119	67.05	**0.028**
			GO:0006814	**Sodium ion transport**	**34**	**17**	**6.81**	0.069
		CC	**GO:0005743**	**Mitochondrial inner membrane**	**71**	**44**	**14.51**	**<0.001**
			GO:0005739	Mitochondrion	216	103	44.13	**<0.001**
			GO:0005759	Mitochondrial matrix	46	22	9.4	**0.006**
			GO:0005811	Lipid particle	70	29	14.3	**0.007**
			GO:0070469	**Respiratory chain**	**36**	**26**	**7.35**	**0.008**
			GO:0005747	**Mitochondrial respiratory chain complex I**	**12**	**8**	**2.45**	0.078
		MF	**GO:0008137**	**NADH dehydrogenase (ubiquinone) activity**	**22**	**18**	**4.22**	**0.000**
	Up in 11-day-old queens of either genotype	BP	—	—	—	—	—	None significant
		CC	—	—	—	—	—	None significant
		MF	—	—	—	—	—	None significant
Age (11-day-old queens—reproductive queens)	Up in 11-day-old queens of either genotype	BP	GO:0055114	Oxidation-reduction process	335	142	83.1	**<0.001**
			GO:0006119	**Oxidative phosphorylation**	**70**	**37**	**17.36**	0.066
		CC	—	—	—	—	—	None significant
		MF	GO:0009055	Electron carrier activity	70	40	17.35	**<0.001**
			GO:0020037	Heme binding	58	35	14.38	**<0.001**
			GO:0004497	Monooxygenase activity	50	28	12.39	**0.001**
			GO:0016491	Oxidoreductase activity	340	139	84.29	**0.010**
			GO:0051539	Four iron, four sulfur cluster binding	11	9	2.73	**0.035**
	Up in reproductive queens of either genotype	BP	GO:0006511	Ubiquitin-dependent protein catabolic process	55	40	24.29	**0.043**
			GO:0000398	Nuclear mRNA splicing, via spliceosome	66	46	29.15	**0.043**
			GO:0007067	**Mitosis**	**52**	**37**	**22.97**	0.093
		CC	**GO:0005730**	**Nucleolus**	**76**	**58**	**36.21**	**0.000**
			GO:0005634	**Nucleus**	**609**	**368**	**290.12**	**0.008**
			GO:0071013	Catalytic step 2 spliceosome	38	30	18.1	**0.016**
			GO:0000502	Proteasome complex	34	27	16.2	**0.023**
		MF	**GO:0005515**	**Protein binding**	**938**	**497**	**404.73**	**<0.001**
			GO:0051082	Unfolded protein binding	39	29	16.83	**0.043**
			GO:0003723	RNA binding	240	132	103.56	**0.043**
			GO:0008026	ATP-dependent helicase activity	45	32	19.42	**0.047**

Table 1 *Continued*

Explanatory term	Direction	GO type*	GO identifier	GO term description	Annotated	Significant	Expected	*P*-value
Genotype	Up in SB/SB	BP	**GO:0015992**	**Proton transport**	**71**	**23**	**4.25**	**<0.001**
			GO:0006120	**Mitochondrial electron transport, NADH to ubiquinone**	**22**	**10**	**1.32**	**0.000**
			GO:0006814	**Sodium ion transport**	**34**	**11**	**2.04**	**0.003**
			GO:0006744	**Ubiquinone biosynthetic process**	**24**	**9**	**1.44**	**0.005**
			GO:0006119	**Oxidative phosphorylation**	**70**	**25**	**4.19**	**0.048**
		CC	**GO:0005747**	**Mitochondrial respiratory chain complex I**	**12**	**7**	**0.63**	**0.000**
			GO:0005743	**Mitochondrial inner membrane**	**71**	**23**	**3.75**	**0.028**
			GO:0070469	**Respiratory chain**	**36**	**17**	**1.9**	0.094
		MF	**GO:0008137**	**NADH dehydrogenase (ubiquinone) activity**	**22**	**9**	**1.29**	**0.003**
	Up in SB/Sb	BP	**GO:0007067**	**Mitosis**	**52**	**22**	**6.77**	**0.011**
		CC	**GO:0005730**	**Nucleolus**	**76**	**25**	**10.48**	**0.009**
			GO:0005634	**Nucleus**	**609**	**127**	**83.99**	**0.040**
		MF	**GO:0005515**	**Protein binding**	**938**	**163**	**120.36**	**0.024**
			GO:0003676	Nucleic acid binding	743	134	95.34	**0.036**

The full model was 'Relative gene expression' ~ 'Genotype' + 'Age' + 'Genotype × Age'. We did six separate GO enrichment analyses: one for genes that were significantly upregulated in SB/Sb queens and one for genes that were significantly upregulated in SB/SB queens for each of the terms Age, Genotype or Genotype × Age. 'Annotated' means the number of genes with this GO term present on the microarray; 'Expected' means the number of genes with this GO term on the microarray that would be significant if there was no association between GO terms and differences in expression; 'Significant' means the number of genes with this GO term on the microarray that were significantly differentially expressed. Significant GOs are shown in bold (FDR < 5%).
*indicates $P < 0.05$.

Table 2 Genes within the supergene were overrepresented among upregulated genes in SB/Sb queens of any age class

Age	Direction	No. of significant genes in				
		Inside supergene	Outside supergene	Unknown location	Two-tailed P*	Two-tailed $P^{\dagger}$
1-day-old queens	Down or upregulated in SB/Sb	4	11	23	**$P < 0.001$**	**$P = 0.005$**
	Downregulated in SB/Sb	0	2	1	More upregulation	—
	Upregulated in SB/Sb	4	9	22	associated with Sb	**$P = 0.003$**
11-day-old queens	Down or upregulated in SB/Sb	42	539	108	**$P < 0.001$**	**$P = 0.006$**
	Downregulated in SB/Sb	23	370	45	More	$P = 0.223$
	Upregulated in SB/Sb	19	169	63	downregulation associated with Sb	**$P = 0.002$**
Reproductive queens	Down or upregulated in SB/Sb	25	202	68	$P = 0.130$	**$P < 0.001$**
	Downregulated in SB/Sb	10	104	20	Equal up and	$P = 0.054$
	Upregulated in SB/Sb	15	98	48	downregulation associated with Sb	**$P < 0.001$**

There is no consistent pattern of more upregulation or more downregulation in SB/Sb queens.
*Binomial test with H_0: Differentially expressed genes are equally likely to be upregulated as downregulated in SB/Sb queens, with $P_{success} = 0.5$.
†Hypergeometric test with H_0: Differentially expressed genes are equally likely to lie inside as outside the supergene, with $P_{success} = 237/4797$.
Bold value indicates $P < 0.05$.

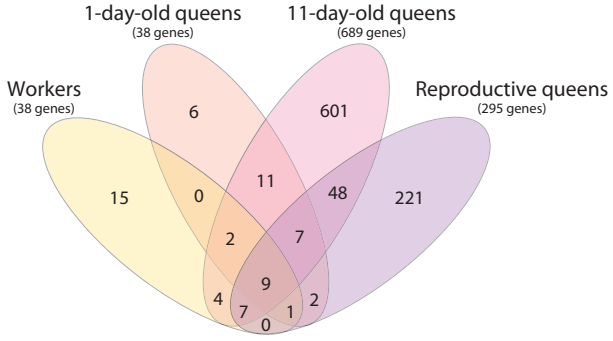

Fig. 2 Genes consistently differentially expressed between genotypes across queens and workers. Venn diagram showing the correspondence between four sets of differentially expressed genes. '1-day-old queens' refers to genes significantly differentially expressed between 1-day-old SB/SB and 1-day-old SB/Sb queens; and analogously for 11-day-old queens, reproductive queens and adult workers. Worker data are from a previous study (Wang *et al*. 2008).

expressed between SB/SB and SB/Sb queens for at least one age class (hypergeometric test, $P < 10^{-10}$), while 19 were also differentially expressed between SB/SB and SB/Sb queens for 2 age classes (hypergeometric test, $P < 10^{-10}$) (Table S3). In contrast to the pronounced agreement between results from workers and queens, there was little agreement between results from between-queen comparisons (1)–(3): among the 919 genes that were differentially expressed between SB/SB and SB/Sb females, 13 (1%) differed between 1-day-old queens and between 11-day-old queens but not between reproductive queens; 55 (6%) differed between 11-day-old queens and between reproductive queens but not between 1-day-old queens; and 3 (0.3%) differed between 1-day-old queens and between reproductive queens but not between 11-day-old queens (Fig. 2). These results indicate that differences in gene expression between SB/SB and SB/Sb queens for one age class generally poorly predict differences in gene expression between SB/SB and SB/Sb queens for another age class, except that genes inside the supergene are always overrepresented among differentially expressed genes for any age class. (Table 2, Fig. S2).

Viral genes and transposons are more highly expressed in SB/Sb queens

Seven viral genes (of estimated 10–30 unique viral genes on our microarray) were differentially expressed between SB/SB and SB/Sb reproductive queens. All seven genes were upregulated (2× to 74×) in SB/Sb reproductive queens (Table 3), consistent with the greater parasite load (including viruses) of polygyne colonies.

Consistent with the hypothesis of transposon accumulation in the Sb supergene, most (13 of 16) of the transposons and transposases that were differentially expressed between SB/SB and SB/Sb queens were upregulated in SB/Sb queens for at least one age class (Table 4); among these 13 genes, a *BEL-12 AG*

Table 3 Seven viral genes were upregulated in polygyne SB/Sb compared with monogyne SB/SB reproductive queens

Description	No. of genes	Fold difference
Integrase, catalytic core	1	3.6
Viral polyprotein	2	1.6
Viral polyprotein	2	2.9
Putative structural protein [SINV2]	1	2.8
Non-capsid protein	1	10.2
Non-structural protein [SINV2]	1	73.9
Viral a-type inclusion protein	1	2.4

Values denote the average of the simple ratios of expression in SB/Sb vs. in SB/SB across all replicate microarray spots per gene.

transposon, *nuclease harbi1-like lc2_043034*, a *piggyBac* transposon and a *ty1-copia* retrotransposon were upregulated in SB/Sb queens for all three age classes.

Candidate genes for chemical signalling by reproductive queens

Given that SB/SB and SB/Sb reproductive queens have different odours, it is noteworthy that at least 18 among 295 differentially expressed genes between SB/SB and SB/Sb reproductive queens are involved in chemical signalling (Table 5). Fourteen of these are involved in

Table 5 Eighteen candidate chemical signalling genes that may cause the social effect of Sb

Description	No. of genes	Fold difference*
Fatty acid/hydrocarbon biosythesis		
Acyl-delta desaturase	2	2.0–4.1
Elongation of very long chain fatty acids protein 1	1	1.8
Acyl-CoA dehydrogenase family	1	1.6
Cytochrome p450	7	1.6–2.2
Fatty acid synthase	1	2.0
Fatty acyl-reductase 1	2	2.2–2.5
OBP/CSP		
Odorant binding protein (OBP2, OBP12, OBP13)	3	2.4–4.2
Chemosensory protein CSP2	1	2.0

Single values denote the average of the simple ratios of expression in SB/Sb vs. in SB/SB, across all replicate microarray spots per gene. Ranges are given only when there is more than one differentially expressed gene with the same function.
*In case of multiple spots, we used the average. For multiple gene, we presented the range.

Table 4 Differentially expressed transposons between SB/SB and SB/Sb queens for at least 1 age class

Gene ID	Description	Linkage group	Upregulated in	Expression ratio		
				1-day-old queens	11-day-old queens	Reproductive queens
SI.MKN.04267	**BEL12_ag transposon polyprotein**	**?**	**SB/Sb**	**2.8**	**2.9**	**3.2**
SI.MKN.00039	Nuclease harbi1-like	?	SB/Sb	1.5	=	1.6
SI.MKN.00133	Nuclease harbi1-like	?	SB/Sb	=	1.6	=
SI.MKN.04299	Nuclease harbi1-like	?	SB/Sb	=	1.4	1.8
SI.MKN.04328	**Nuclease harbi1-like**	**?**	**SB/Sb**	**3.3**	**10.3**	**8.5**
SI.MKN.04583	**PiggyBac transposon**	**S**	**SB/Sb**	**9.3**	**28.9**	**24.1**
SI.MKN.04107	TE reverse transcriptase	?	SB/Sb	=	1.7	=
SI.MKN.02387	**Retrotransposon ty1-copia subclass**	**3**	**SB/Sb**	**1.5**	**2.8**	**5.3**
SI.MKN.04166	kda protein in nof-fb transposable element	?	SB/Sb	=	=	1.4
SI.MKN.05057	Transposon, probable	?	SB/Sb	=	1.4	=
SI.MKN.03796	Retrovirus-related pol polyprotein from transposon 412	?	SB/Sb	1.6	1.6	=
SI.MKN.00104	Transposable element tc3 transposase	?	SB/Sb	=	1.3	=
SI.MKN.05121	Transposase (Fragment)	?	SB/Sb	=	1.6	=
SI.MKN.90403	Mariner transposase	?	SB/SB	=	=	0.6
SI.MKN.00200	Mariner transposase	?	SB/SB	=	=	0.6
SI.MKN.04695	Mariner Mos1 transposase	?	SB/SB	=	0.6	0.5

Values denote the average of the simple ratios of expression in SB/Sb vs. in SB/SB across all microarray spots per gene. '=' means not differentially expressed for a given age class, while text in bold denotes transposons/transposases that were differentially expressed for all three age classes. 'S' denotes the 'social chromosome' that contains the supergene. Note that most transposons and transposases were upregulated in SB/Sb queens.
?indicates $P < 0.01$.

fatty acid or hydrocarbon biosynthesis (e.g. *fatty acid synthase, fatty-acyl reductase, fatty-acyl elongase, acyl-delta desaturase, acyl-CoA dehydrogenase and cytochrome P450*), while the others are the odorant-binding proteins *OBP2, OBP12, OBP13*, and the chemosensory protein *CSP2*. Notably, all of these 18 genes were more highly expressed in SB/Sb than in SB/SB reproductive queens.

We verified the expression of one desaturase gene and one elongase gene on microarrays and five additional desaturase and two elongase genes that are not on microarrays and found that five of six desaturase and two of three elongase genes including the one desaturase (Fig. 3e) and one elongase (Fig. 3h) genes also present on our microarrays showed significant higher expression in SB/Sb compared with SB/SB reproductive queens (Wilcoxon test, $P < 0.005$; Fig. 3).

Discussion

Our study shows that the differences in physiology and behaviour of queens that ultimately cause the social polymorphism (monogyny vs. polygyny) of fire ants are themselves due to widespread differences in gene expression. In particular, the higher rate and greater extent of fat accumulation by SB/SB queens enables them to found colonies independently and to lay unfertilized eggs sooner after losing their queen than SB/Sb queens (Keller & Ross 1993, 1999; K G Ross & Keller 1998). The trade-off is that SB/SB queens, but not SB/Sb queens, that reproduce in the queen's presence get executed by workers; the workers use odour as a cue to discriminate between queens in this way (Eliyahu *et al.* 2011; Keller & Ross 1998). Our study also shows that genes within the

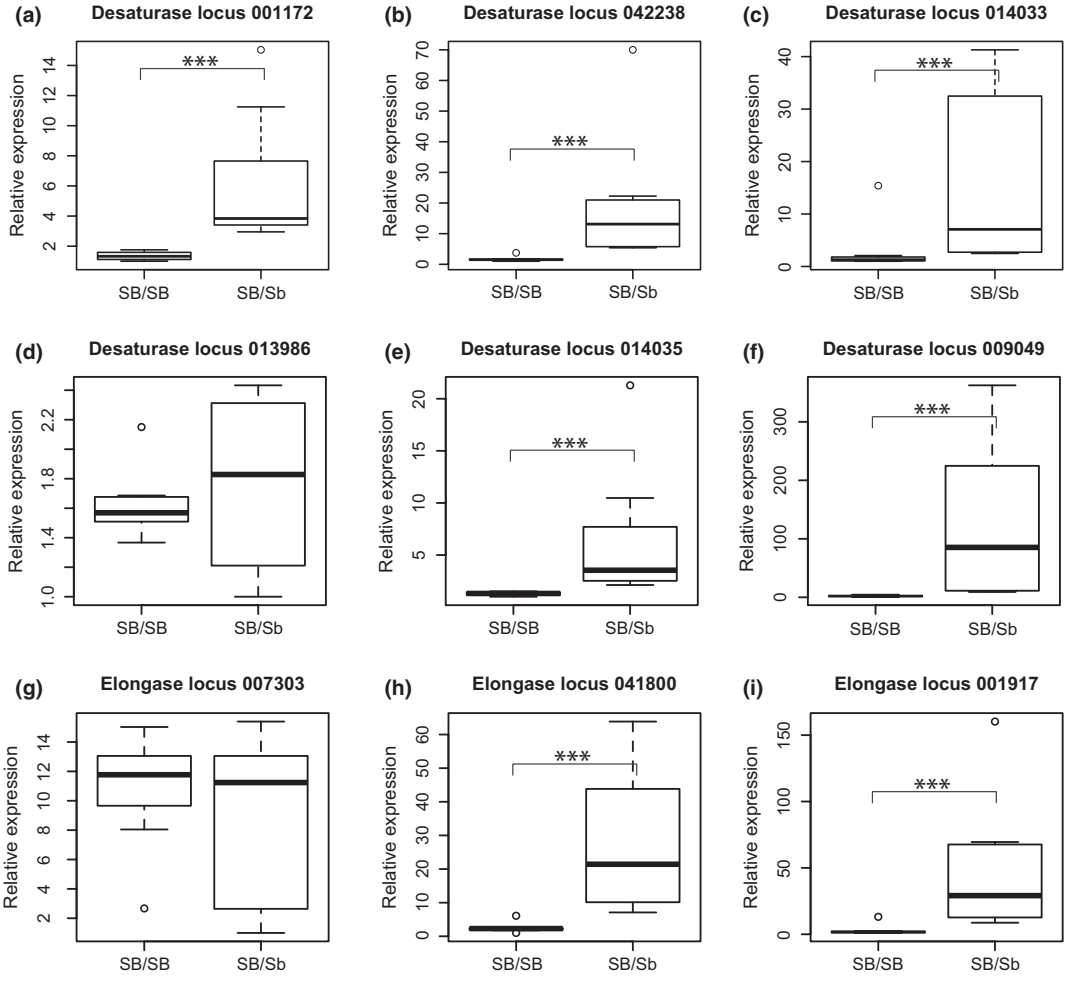

Fig. 3 Many desaturase and elongase genes are upregulated in SB/Sb compared with SB/SB reproductive queens. qRT-PCR analysis of gene expression levels in SB/SB ($n = 8$) and SB/Sb ($n = 8$) reproductive queens. (a–f) different desaturase genes; (g-i) different elongase genes; (e, h) also on the microarray. Box-and-whisker plots of relative gene expression levels: the top and bottom of the box are the first and third quartiles, respectively; horizontal bar within the box is median, whiskers extend from the box to the most extreme expression value length within 1.5× of the interquartile range of the box; data beyond the whiskers are outliers and plotted as points. ***$P < 0.005$, Wilcoxon test.

supergene are overrepresented among genes whose expression differs between SB/SB and SB/Sb queens.

Several lines of evidence are consistent with the known faster maturation of SB/SB queens (Keller & Ross 1993, 1999). First, principle component analysis showed that the differences between the two genotypes are most pronounced in 11-day-old queens and that 11-day SB/SB queens resemble reproductive queens more than 11-day SB/Sb queens do (in PC1 and PC2: Fig. 1b). Second, 90% of genes differentially expressed between 11-day-old SB/SB and SB/Sb queens are involved in reproduction (Fig. 1c, Table S4), including the queen-specific *vitellogenin-3*. Third, for all 3 age classes, there was an overrepresentation of genes that are likely involved in reproduction [e.g. mitosis; protein binding; localization in nucleolus or nucleus (Kocher *et al.* 2008; Baker & Russell 2009)] among genes upregulated in SB/SB compared with SB/Sb queens (Table 1). Three of five reproduction-associated GO terms that were significant for genes upregulated in 11-day-old SB/SB queens compared with 11-day-old SB/Sb queens were likewise significant for genes upregulated in reproductive queens compared with 11-day-old queens. Fourth, there was an overrepresentation of oxidative metabolism genes (mostly mitochondrial) among genes downregulated in SB/SB queens (Table 1). All eight oxidative-metabolism-associated GO terms that were significant for genes downregulated in 11-day-old SB/SB queens compared with 11-day-old SB/Sb queens were also significant for genes downregulated in 11-day-old queens compared with 1-day-old queens. A reduced oxidative metabolism in SB/SB queens is consistent with the earlier onset of fat accumulation in SB/SB queens than in SB/Sb queens (Keller & Ross 1993, 1999). However, no GO terms associated with fat storage were significant for genes differentially expressed between SB/SB and SB/Sb queens for any age class, while two homologs of *lipid storage droplets surface-binding protein 2* were the only obvious candidate genes for fat storage that were upregulated in 11-day-old SB/SB queens.

Genes inside the supergene were significantly overrepresented among genes differentially expressed between individuals of all castes [queens (this study and Wang *et al.* 2013), workers (Wang *et al.* 2008), males (Wang *et al.* 2013)] carrying the Sb supergene compared with individuals without the Sb supergene. A difference is that for both workers and males, the majority of differentially expressed genes lie inside the supergene (Wang *et al.* 2008, 2013), while for queens, genes inside the supergene are a minority (1–27%) among all differentially expressed genes, even if this percentage is greater than expected by chance for all age classes (all $P < 0.01$; Table 2). This implies that some of the translation products from the supergene transregulate genes outside the supergene, especially in queens. In animals with XY sex determination, where the Y chromosome is always evolutionarily derived, the male phenotype generally involves more downregulation than upregulation (relative to the corresponding X-linked genes) of those Y-linked genes that are still functional (Charlesworth & Charlesworth 2000; Charlesworth *et al.* 2005). In fire ants, the Sb supergene can be assumed to be evolutionary derived. However, for all age classes, differentially expressed genes lying inside the supergene were equally often upregulated as downregulated in SB/Sb compared with SB/SB queens. Likewise, genes within the supergene that were differentially expressed between SB/SB and SB/Sb workers were equally often upregulated as downregulated in SB/SB workers, and analogously for genes within the supergene that were differentially expressed between SB and Sb haploid males (Wang *et al.* 2013). It is important to note that the gene *Gp-9*, which is a perfect marker of the supergene, was not differentially expressed between SB/SB and SB/Sb queens of any age class. This suggests that the functional differences between $Gp\text{-}9^B$ and $Gp\text{-}9^b$ are likely mainly due to amino acid substitutions rather than to differences in levels of allele-specific gene expression.

Only 16 genes were differentially expressed between SB/SB and SB/Sb queens for all three age classes, suggesting that the supergene affects different genes across development. Given the known faster maturation of SB/SB queens, some of the 38 genes differentially expressed between 1-day-old queens might be regulatory factors, inducing transcriptional differences in many downstream genes necessary for reproduction. Many of these 38 genes are indeed involved in transcription regulation (e.g. *pre-mRNA branch site p14*, *sequestosome*, *oskar*, *enhancer of mRNA-decapping protein*, *deoxyribonuclease TATDN1*, *growth-arrest-specific protein 8*, *ubiquitin-like domain-containing CTD phosphatase 1*, *ectopic-P granules protein 5 homolog*). On the other hand, that only 71 of the 689 genes differentially expressed between 11-day-old queens were also differentially expressed between reproductive queens is not surprising given that the presence or absence of the Sb chromosome strongly affects fecundity (and thus the expression of reproductive genes) of queens when they are approximately 11 days old but not when they have become reproductive queens. Consistent with this, five reproduction-associated GO terms were significant for genes upregulated in 11-day-old SB/SB compared with 11-day-old SB/Sb queens (Table S4), but no reproduction-associated GO terms were significant for genes differentially expressed between SB/SB and SB/Sb reproductive queens.

More than half (23/38 genes) of the genes differentially expressed between SB/SB and SB/Sb workers

were also differentially expressed between SB/SB and SB/Sb queens for at least one age class (Fig. 2), and 19 of these 38 genes were also differentially expressed between queens for at least two age classes (Table S3). This agreement indicates that the supergene affects expression in a similar manner in queens and workers, for example SB/SB workers and queens are generally bigger/heavier compared with SB/Sb and/or the ability of SB/Sb individuals to recognize individuals with the Sb supergene.

The upregulation of seven viral genes in SB/Sb (polygyne) reproductive queens compared with SB/SB (monogyne) reproductive queens, which has likewise been reported for polygyne compared with monogyne workers (Wang *et al.* 2008), is consistent with the known greater parasite load (viruses, microsporidia) of polygyne colonies (Fuxa *et al.* 2005; Preston *et al.* 2007; Hashimoto & Valles 2007, 2008a,b; Valles *et al.* 2010; Allen *et al.* 2011). The lower parasite load of monogyne colonies is thought to result from intense, recurrent selection against infected queens during the independent founding stage, which prevents infected queens from successfully establishing colonies, while infected queens in polygyne bud nests may survive because the cost of their infection gets partitioned among infected and noninfected queens.

Most of the transposons (13 of 16) differentially expressed between queens for at least one age class were more highly expressed in SB/Sb queens (Table 4). Assuming that the higher expression of transposons in SB/Sb queens results from them carrying more transposon copies (cf. Pasyukova *et al.* 1998), this result is again consistent with the hypothesis that the Sb supergene evolved in a similar fashion as Y chromosomes, which are known to accumulate transposons (Charlesworth 1991). Despite large differences between queens and workers in morphology, physiology and overall gene expression, three of nine transposons that were more expressed in SB/Sb workers compared with SB/SB workers were also more expressed in SB/Sb queens compared with SB/SB queens (Table S3).

The only exception to the pattern of greater expression of transposons in SB/Sb females, namely the upregulation of *mariner transposase* in SB/SB queens (Table 4), does not necessarily contradict the hypothesis that the Sb supergene has accumulated transposons, provided that a significant proportion of mariner copies lies in the Sb supergene and that some of the Sb-linked copies are defective: this is because transcription of *mariner* transposons is known to decrease due to autoinhibition, through overproduction inhibition when there are many copies, or through dominant negative complementation when some copies contain hypomorphic missense mutations (Lohe & Hart 1996).

The higher expression due to a higher copy number of transposons in individuals with the Sb supergene would have automatically generated more genetic variation that could be screened by selection to adaptively upregulate or downregulate genes in the Sb supergene. This is because the insertion of transposons can modify the expression of a gene by disrupting the *cis*-regulation of its native promoter, by donating a new *cis*-regulatory sequence or by altering the *trans*-regulation by an upstream gene (Lerat & Sémon 2007; Huang *et al.* 2008; Deloger *et al.* 2009; Herpin *et al.* 2010; Magwire *et al.* 2011). Indeed, half of the transposons in the *Drosophila* genome lie inside genes, especially within introns and 5′ UTRs (Deloger *et al.* 2009), presumably because selection has favoured transposons that caused adaptive changes in gene regulation.

Previous studies have shown that workers use differences in queen odour to discriminate between SB/SB and SB/Sb queens (Keller & Ross 1998; Ross & Keller 1998), so that chemical signalling genes that differ in expression between reproductive queens are obvious candidates for genes that elicit discrimination by workers. Of the 18 chemical signalling genes that were differentially expressed between reproductive queens, 14 were involved in hydrocarbon synthesis. Among these 14 genes, two desaturases (cDNAs SiJWA08ACM; and SiJWE01ADR, SiJWF06BCP, SiJWF06BCP.scf, SiJWG02-CAJ, SiJWG03BAW) and one elongase (SiJWA02ABI) are our best candidates for the genes that produce the cue that elicits worker discrimination, because differential expression of desaturases and elongases is known to yield different pheromones or hydrocarbon profiles in other insects (see Wicker-Thomas & Chertemps 2010 for review). Differential expression of these candidate genes was also supported by the qPCR results (Fig. 3) where most of these genes (7/9), including the genes on our microarrays, were highly expressed in SB/Sb reproductive queens, suggesting that many of these genes that are not present on our microarrays are also involved in queen's odour.

The upregulation of desaturases in SB/Sb reproductive queens is consistent with them having a higher proportion of unsaturated hydrocarbons on their cuticle than SB/SB reproductive queens (Eliyahu *et al.* 2011). As the precursors of hydrocarbons are fatty acids, while desaturases transform saturated fatty acids into unsaturated fatty acids (see Blomquist & Bagnères 2010 for review), the upregulation of two desaturases in SB/Sb reproductive queens may directly yield a greater proportion of unsaturated cuticular hydrocarbons. On the other hand, the upregulation of elongase SiJWA02ABI might be necessary for the elongation of unsaturated hydrocarbons in SB/Sb reproductive queens, because elongases can be coregulated with desaturases to

elongate certain types of hydrocarbon chains (Jakobsson *et al.* 2006; Green *et al.* 2010). The equal expression of desaturases in 11-day-old SB/SB and SB/Sb queens is likewise consistent with the equal proportion of unsaturated compared with saturated hydrocarbons in SB/SB and SB/Sb 14-day-old queens (Eliyahu *et al.* 2011). The only known difference in cuticular hydrocarbons between 14-day-old SB/SB and SB/Sb queens is the higher proportion of *cis*-alkaloids (possibly a cue used by workers to recognize reproductive queens) in SB/SB queens (Eliyahu *et al.* 2011). This would imply that, even 14 days after eclosion, SB/Sb queens are still not fully mature and therefore have less *cis*-alkaloids and unsaturated hydrocarbons.

The remaining four (of 18) chemical signalling genes that are good candidates for determining social form are three odorant-binding proteins and one chemical signalling protein, all of which were more highly expressed in SB/Sb than in SB/SB reproductive queens. Odorant-binding proteins and chemical signalling proteins are prime candidates for genes that are important in queen-to-queen, queen-to-worker and worker-to-queen signalling because they are important for odour and pheromone detection (Steinbrecht 1998; Pelosi *et al.* 2005). Their upregulation in SB/Sb queens might be an adaptive response induced by a more complex mix of odours in polygyne colonies. For example, a polygyne queen may need to sense queen number to manipulate workers into raising her brood at the expense of the brood of nest-mate queens (cf. Keller & Nonacs 1993; Hannonen *et al.* 2002). In particular, the upregulation of the odorant-binding protein *OBP2* seems to be universal in females (workers and queens) with the Sb supergene. This suggests that, as previously proposed (Wang *et al.* 2008), OBP2 might regulate social form by enabling SB/Sb workers to discriminate in favour of SB/Sb queens. Furthermore, the odorant-binding proteins *OBP2, OBP12, OBP13*, and the chemical signalling protein *CSP2* lie inside the supergene, which supports the hypothesis that the Sb supergene has expanded through evolution to include genes that benefit polygyny.

In conclusion, this study highlights the unique importance of the supergene for social organization. A queen's supergene genotype directly and indirectly affects downstream genes that regulate her oxidative metabolism, sexual maturation and odour. These effects of the supergene on the founding queen's physiology then get amplified, because the supergene inside her worker daughters determines which queens they will tolerate, using the supergene-controlled odour of queens as a cue. Our findings thus illustrate how simple genetic differences in individuals can translate into complex phenotypic differences at higher levels of organization.

Acknowledgements

We thank Frédéric Schütz, Darlene Goldstein, and Colby Tanner for statistical advice; Oksana Grognuz-Riba for sharing RNAseq data; Kenneth G. Ross for help during ant collection and Kenneth G. Ross as well as three reviewers for comments on the manuscript; Laelia Maumary and Jezaelle Rufener for ant care; Christine LaMendola for laboratory assistance; Céline Stoffel, Hannes Richter, and Eyal Privman for qPCR help; and the staff of the DNA Array Facility of the University of Lausanne for microarray printing, advice, and access to software.

References

Alexa A, Rahnenführer J (2010) topGO: Enrichment analysis for Gene Ontology. 2010. Bioconductor package version 2.6.0. http://www.bioconductor.org/packages/release/bioc/html/topGO.html.

Allen C, Valles SM, Strong CA (2011) Multiple virus infections occur in individual polygyne and monogyne *Solenopsis invicta* ants. *Journal of invertebrate pathology*, **107**, 107–111.

Baker DA, Russell S (2009) Gene expression during *Drosophila melanogaster* egg development before and after reproductive diapause. *BMC Genomics*, **10**, 242.

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B (Methodological)*, **57**, 289–300.

Blomquist GJ, Bagnères A-G (2010) *Insect Hydrocarbons: Biology, Biochemistry, and Chemical Ecology*. Cambridge University Press, Cambridge.

Charlesworth B (1991) The evolution of sex chromosomes. *Science*, **251**, 1030–1033.

Charlesworth B, Charlesworth D (2000) The degeneration of Y chromosomes. *Philosophical transactions of the Royal Society of London. Series B, Biological Sciences*, **355**, 1563–1572.

Charlesworth D, Charlesworth B, Marais G (2005) Steps in the evolution of heteromorphic sex chromosomes. *Heredity*, **95**, 118–128.

Conesa A, Götz S, García-Gómez JM, *et al.* (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics*, **21**, 3674–3676.

De Bono M, Bargmann CI (1998) Natural variation in a neuropeptide Y receptor homolog modifies social behavior and food response in *C. elegans*. *Cell*, **94**, 679–689.

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

DeHeer CJ, Goodisman MAD, Ross KG (1999) Queen dispersal strategies in the multiple-queen form of the fire ant *Solenopsis invicta*. *The American naturalist*, **153**, 660–675.

Deloger M, Cavalli FMG, Lerat E, *et al.* (2009) Identification of expressed transposable element insertions in the sequenced genome of *Drosophila melanogaster*. *Gene*, **439**, 55–62.

Eliyahu D, Ross KG, Haight KL, 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**, 1242–1254.

Ellison CK, Wiley C, Shaw KL (2011) The genetics of speciation: genes of small effect underlie sexual isolation in the

Hawaiian cricket *Laupala*. *Journal of evolutionary biology*, **24**, 1110–1119.

Fletcher DJ, Blum MS (1981) Pheromonal control of dealation and oogenesis in virgin queen fire ants. *Science*, **212**, 73–75.

Fletcher DJC, Blum MS (1983) The inhibitory pheromone of queen fire ants: effects of disinhibition on dealation and oviposition by virgin queens *Journal of Comparative Physiology A*, **153**, 467–475.

Fuxa JR, Milks ML, Sokolova YY, Richter AR (2005) Interaction of an entomopathogen with an insect social form: an epizootic of Thelohania solenopsae (Microsporidia) in a population of the red imported fire ant, *Solenopsis invicta*. *Journal of invertebrate pathology*, **88**, 79–82.

Gentleman RC, Carey VJ, Bates DM, *et al.* (2004) Bioconductor: open software development for computational biology and bioinformatics. *Genome biology*, **5**, R80.

Gleason JM (2005) Mutations and natural genetic variation in the courtship song of *Drosophila*. *Behavior Genetics*, **35**, 265–277.

Goodisman 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). *Annals of the Entomological Society of America*, **92**, 563–570.

Goodisman MAD, Deheer CJ, Ross KG (2000) Unusual behavior of polygyne fire ant queens on nuptial flights. *Journal of Insect Behavior*, **13**, 455–468.

Goodwin SF, Taylor BJ, Villella A, *et al.* (2000) Aberrant splicing and altered spatial expression patterns in fruitless mutants of *Drosophila melanogaster*. *Genetics*, **154**, 725–745.

Götz S, García-Gómez JM, Terol J, *et al.* (2008) High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic acids research*, **36**, 3420–3435.

Götz S, Arnold R, Sebastián-León P, *et al.* (2011) B2G-FAR, a species-centered GO annotation repository. *Bioinformatics*, **27**, 919–924.

Green CD, Ozguden-Akkoc CG, Wang Y, Jump DB, Olson LK (2010) Role of fatty acid elongases in determination of de novo synthesized monounsaturated fatty acid species. *Journal of lipid research*, **51**, 1871–1877.

Hallar BL, Krieger MJB, Ross KG (2007) Potential cause of lethality of an allele implicated in social evolution in fire ants. *Genetica*, **131**, 69–79.

Hannonen M, Sledge MF, Turillazzi S, Sundström L (2002) Queen reproduction, chemical signalling and worker behaviour in polygyne colonies of the ant Formica fusca. *Animal Behaviour*, **64**, 477–485.

Hashimoto Y, Valles SM (2007) *Solenopsis invicta* virus-1 tissue tropism and intra-colony infection rate in the red imported fire ant: a quantitative PCR-based study. *Journal of invertebrate pathology*, **96**, 156–161.

Hashimoto Y, Valles SM (2008a) Detection and quantitation of *Solenopsis invicta* virus-2 genomic and intermediary replicating viral RNA in fire ant workers and larvae. *Journal of invertebrate pathology*, **98**, 243–245.

Hashimoto Y, Valles SM (2008b) Infection characteristics of *Solenopsis invicta* virus 2 in the red imported fire ant, *Solenopsis invicta*. *Journal of invertebrate pathology*, **99**, 136–140.

Hellemans J, Mortier G, De Paepe A, Speleman F, Vandesompele J (2007) qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. *Genome biology*, **8**, R19.

Herpin A, Braasch I, Kraeussling M, *et al.* (2010) Transcriptional rewiring of the sex determining dmrt1 gene duplicate by transposable elements. *PLoS genetics*, **6**, e1000844.

Huang X, Lu G, Zhao Q, Liu X, Han B (2008) Genome-wide analysis of transposon insertion polymorphisms reveals intraspecific variation in cultivated rice. *Plant physiology*, **148**, 25–40.

Jakobsson A, Westerberg R, Jacobsson A (2006) Fatty acid elongases in mammals: their regulation and roles in metabolism. *Progress in lipid research*, **45**, 237–249.

Jouvenaz DP, Allen GE, Banks WA, *et al.* (1977) A survey for pathogens of fire ants, *Solenopsis* spp., in the Southeastern United States. *Florida Entomologist*, **60**, 275–279.

Keller L, Nonacs P (1993) The role of queen pheromones in social insects: queen control or queen signal? *Animal Behaviour*, **45**, 787–794.

Keller L, Ross KG (1993) Phenotypic plasticity and "cultural transmission" of alternative social organizations in the fire ant *Solenopsis invicta*. *Behavioral Ecology and Sociobiology*, **33**, 121–129.

Keller L, Ross KG (1998) Selfish genes: a green beard in the red fire ant. *Nature*, **251**, 573–575.

Keller L, Ross KG (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**, 672–680.

Kocher SD, Richard F-J, Tarpy DR, Grozinger CM (2008) Genomic analysis of post-mating changes in the honey bee queen (*Apis mellifera*). *BMC Genomics*, **9**, 232.

Krieger MJB, Ross KG (2002) Identification of a major gene regulating complex social behavior. *Science*, **295**, 328–332.

Lawson LP, Vander Meer RK, 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 Biological Sciences/The Royal Society*, **279**, 3217–3222.

Lerat E, Sémon M (2007) Influence of the transposable element neighborhood on human gene expression in normal and tumor tissues. *Gene*, **396**, 303–311.

Lohe AR, Hart DL (1996) Autoregulation of mariner transposase complementation activity by overproduction. *Molecular biology and evolution*, **13**, 549–555.

Magwire MM, Bayer F, Webster CL, Cao C, Jiggins FM (2011) Successive increases in the resistance of *Drosophila* to viral infection through a transposon insertion followed by a Duplication. *PLoS genetics*, **7**, e1002337.

Ometto L, Shoemaker D, Ross KG, Keller L (2011) Evolution of gene expression in fire ants: the effects of developmental stage, caste, and species. *Molecular biology and evolution*, **28**, 1381–1392.

Pasyukova EG, Nuzhdin SV, Filatov DA (1998) The relationship between the rate of transposition and transposable element copy number for copia and Doc retrotransposons of *Drosophila melanogaster*. *Genetical research*, **72**, 1–11.

Pelosi P, Calvello M, Ban L (2005) Diversity of odorant-binding proteins and chemosensory proteins in insects. *Chemical Senses*, **30**(Suppl 1), i291–i292.

Preston CA, Fritz GN, Vander Meer RK (2007) Prevalence of Thelohania solenopsae infected *Solenopsis invicta* newly mated queens within areas of differing social form distributions. *Journal of invertebrate pathology*, **94**, 119–124.

R Development Core Team (2010) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Robinson GE, Ben-Shahar Y (2002) Social behavior and comparative genomics: new genes or new gene regulation? *Genes, Brain, and Behavior*, **1**, 197–203.

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

Ross KG, Fletcher DJC (1986) Diploid male production: a significant colony mortality factor in the fire ant *Solenopsis invicta* (Hymenoptera:Formicidae). *Behavioral Ecology and Sociobiology*, **19**, 283–291.

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

Ross KG, Keller L (1998) Genetic control of social organization in an ant. *Proceedings of the National Academy of Sciences of the United States of America*, **95**, 14232–14237.

Ross K, Keller L (2002) Experimental conversion of colony social organization by manipulation of worker genotype composition in fire ants (*Solenopsis invicta*). *Behavioral Ecology and Sociobiology*, **51**, 287–295.

Ross KG, Shoemaker DD, Krieger MJ, DeHeer CJ, 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**, 525–543.

Ruedi EA, Hughes KA (2008) Natural genetic variation in complex mating behaviors of male *Drosophila melanogaster*. *Behavior Genetics*, **38**, 424–436.

Smyth GK (2004) Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Statistical Applications in Genetics and Molecular Biology*, **3**, Article3.

Steinbrecht RA (1998) Odorant-binding proteins: expression and function. *Annals of the New York Academy of Sciences*, **855**, 323–332.

Tschinkel WR (2006) *The fire ants*. Belknap Pr, Cambridge, MA, and London, UK.

Valles SM, Oi DH, Porter SD (2010) Seasonal variation and the co-occurrence of four pathogens and a group of parasites among monogyne and polygyne fire ant colonies. *Biological Control*, **54**, 342–348.

Vargo EL, Fletcher DJC (1986) Evidence of pheromonal queen control over the production of male and female sexuals in the fire ant, *Solenopsis invicta*. *Journal of Comparative Physiology A*, **159**, 741–749.

Vosshall LB (2007) Into the mind of a fly. *Nature*, **450**, 193–197.

Wang J, Jemielity S, Uva P, *et al*. (2007) An annotated cDNA library and microarray for large-scale gene-expression studies in the ant *Solenopsis invicta*. *Genome biology*, **8**, R9.

Wang J, Ross KG, Keller L (2008) Genome-wide expression patterns and the genetic architecture of a fundamental social trait. *PLoS genetics*, **4**, e1000127.

Wang J, Wurm Y, Nipitwattanaphon M, *et al*. (2013) A Y-like social chromosome causes alternative colony organization in fire ants. *Nature*, **493**, 664–668.

Wicker-Thomas C, Chertemps T (2010) Molecular biology and genetics of hydrocarbon production. In: *Insect Hydrocarbons: Biology, Biochemistry, and Chemical Ecology* (eds Blomquist GJ, Bagnéres A-G), pp. 53–74. Cambridge University Press, Cambridge.

Wurm Y, Wang J, Keller L (2010) Changes in reproductive roles are associated with changes in gene expression in fire ant queens. *Molecular ecology*, **19**, 1200–1211.

Wurm Y, Wang J, Riba-Grognuz O, *et al*. (2011) The genome of the fire ant *Solenopsis invicta*. *Proceedings of the National Academy of Sciences of the United States of America*, **108**, 5679–5684.

M.N., J.W. and L.K. planned the experiments, M.N. collected and analyzed the data, and M.B.D., M.N. and L.K. wrote the paper.

Data accessibility

Gene expression data meet MIAME standards and have been deposited at Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo/) under accession number GSE42062.

Supporting information

Additional supporting information may be found in the online version of this article.

Fig. S1 Most (95%) genes upregulated or downregulated in 11-day-old SB/SB queens are involved in reproductive maturation.

Fig. S2 Overrepresentation of genes inside the supergene among genes differentially expressed between genotypes.

Table S1 Upregulated genes in more mature queens.

Table S2 Upregulated genes in immature queens.

Table S3 Genes consistently differentially expressed in all between-genotype comparisons for queens and workers.

Table S4 Gene Ontology categories overrepresented among differentially expressed genes between 11-day-old SB/SB and SB/Sb queens.

Table S5 Genes and primer sequences used for qRT-PCR verification of microarray expression data.