Disruption of gene expression in hybrids of the fire ants *Solenopsis invicta* and *Solenopsis richteri*

LINO OMETTO,*§ KENNETH G. ROSS,† DEWAYNE SHOEMAKER‡ and LAURENT KELLER*

*Department of Ecology and Evolution, University of Lausanne, Biophore, 1015 Lausanne, Switzerland, †Department of Entomology, University of Georgia, Athens, GA 30602-2603, USA, ‡Center for Medical, Agricultural, and Veterinary Entomology, United States Department of Agriculture – Agricultural Research Service, Gainesville, FL 32608, USA

**Abstract**

Transcriptome analysis is a powerful tool for unveiling the distribution and magnitude of genetic incompatibilities between hybridizing taxa. The nature of such incompatibilities is closely associated with the evolutionary histories of the parental species and may differ across tissues and between the sexes. In eusocial insects, the presence of castes that experience divergent selection regimes may result in additional distinct patterns of caste-specific hybrid incompatibilities. We analysed levels of expression of >14 000 genes in two life stages of each caste in the fire ants *Solenopsis invicta* and *Solenopsis richteri* and in their hybrids. We found strong contributions of both developmental stage and caste to gene expression patterns. In contrast, variability in expression was only weakly associated with taxonomic identity, with hybrid scores falling between those of the two parental species. Hybrid incompatibilities were surprisingly modest, with only 32 genes being mis-expressed, indicating low levels of disruption in gene regulation in hybrids; males and workers each mis-expressed at least seven times as many genes as queens. Interestingly, homologues of many of the mis-expressed genes have been implicated in behavioural variation in *Drosophila melanogaster*. General expression profiles of hybrids consistently were more similar to those of *S. richteri* than *S. invicta*, presumably because *S. richteri* trans-regulatory elements tend to be dominant and/or because there is an overall bias in the genetic composition of the hybrids towards *S. richteri*. Altogether, our results suggest that selection acting on each caste may contribute differently to interspecific divergence and speciation in this group of ants.

**Keywords**: ants, caste, gene expression, hybrids, *Solenopsis invicta*, *Solenopsis richteri*

Received 20 November 2011; revision received 27 January 2012; accepted 8 February 2012

**Introduction**

The process of speciation in eukaryotes typically involves divergence of two lineages from an ancestral lineage, resulting in their genomes becoming sufficiently incompatible that hybridization produces inviable or relatively less fit individuals. Such divergence can involve variable numbers of loci and may stem from neutral stochastic pressures or from extrinsic or intrinsic selective forces (Coyne & Orr 2004). For instance, when species-specific loci have evolved to confer adaptation to divergent niches, hybrids may suffer fitness losses from having an intermediate phenotype poorly suited to either of the parental habitats (Schluter 2001; Taylor *et al.* 2012). Similarly, reduced hybrid fitness may derive from deleterious epistatic interactions between divergent parental alleles (i.e. Dobzhansky–Muller incompatibilities, Dobzhansky 1937; Muller 1942). These latter interactions typically involve co-adapted gene complexes and networks, implying that incompatibilities can involve both structural genes and regulatory regions. Consequently, accumulation of species-specific genetic variation can result in divergent regulatory networks that are disrupted in introgressed genomes, causing profound alterations in gene expres-
sion profiles compared to the parental species (e.g. Ranz et al. 2004; Landry et al. 2007; Ortiz-Barrientos et al. 2007; Renaut & Bernatchez 2011).

The genomic distribution and magnitude of such hybrid incompatibilities are influenced by the degree of genetic divergence at particular genomic elements, the effect sizes of diverging elements, dominance relationships among the parental alleles and patterns of epistasis (e.g. Orr 1993a; Turelli & Orr 1995). Thus, hybrid incompatibilities are tied to the evolutionary histories of the interacting genes in the parental species. Because the selective history of a gene is linked to its expression pattern and, ultimately, its phenotypic effect, the degree of incompatibility also may differ among environments, across tissues, and/or between the sexes. An additional layer of phenotypic differentiation often occurs in social hymenopteran insects as a result of the caste system, whereby females conditionally develop into fertile queens that are morphologically, physiologically and behaviourally distinct from sterile workers.

In highly eusocial Hymenoptera such as most ants and the honey bee, the striking behavioural and morphological differences between queens, workers and males are associated with important differences in historical and current selection regimes acting on each caste (Hunt et al. 2010, 2011; Ometto et al. 2011). Not surprisingly, previous studies have revealed that differences between social insect castes are associated with striking differences in gene expression during immature development (Evans & Wheeler 2000; Scharf et al. 2003; Pereboom et al. 2005; Hoffman & Goodisman 2007; Weil et al. 2009) as well as in the adult stage (Judice et al. 2006; Sumner et al. 2006; Gräff et al. 2007; Sen Sarma et al. 2007; Hunt & Goodisman 2010; Colgan et al. 2011; Ometto et al. 2011). Thus, interspecific hybridization may result in distinct caste-specific and sex-specific patterns and levels of hybrid incompatibilities in these insects. Accordingly, comparison of the nature of incompatibilities between the two female castes is expected to provide unique information on the relative importance of sexual (queen) versus somatic (worker) divergence in driving hybrid incompatibilities and reproductive isolation. Notably, incompatibilities arising solely in workers have only an indirect effect on hybrid fitness, because members of this caste generally do not reproduce, yet they profoundly affect colony survival and productivity.

Hybridization in the wild has been well documented in several species of ants (e.g. Pearson 1983; Douwe & Stille 1991; Shoemaker et al. 1996; Seifert 1999; Helms Cahan et al. 2002; Feldhaar et al. 2003; Helms Cahan & Vinson 2003; Umphrey 2006; Kulmuni et al. 2010; reviewed in Feldhaar et al. 2008). However, excluding unusual reproductive systems such as those in which historical hybridization has led to the stable coexistence of distinct interbreeding lineages (Helms Cahan & Keller 2003; Helms Cahan et al. 2004; Schwander et al. 2006), very few studies have investigated any of the phenotypic or fitness consequences of hybridization in ants (Jessen & Klinkicht 1990; Ross & Robertson 1990; Pusch et al. 2006). Thus, the genetic basis of hybrid incompatibilities manifested in each caste, and the contributions of these incompatibilities to reproductive isolation between hybridizing species, remains largely unknown for ants as well as other social insects.

A valuable study system for investigating patterns of hybrid incompatibility in social Hymenoptera is represented by the closely related fire ant species Solenopsis invicta and Solenopsis richteri. These two species are reproductively isolated in their native South American ranges (Ross & Shoemaker 2005) but exhibit extensive introgression and evidently produce relatively fit hybrids in their invasive North American ranges (Ross & Robertson 1990; Shoemaker et al. 1996). The absence of premating reproductive isolation between the USA populations may stem in part from the fact that they originated from allopatric South American populations that lack the behavioural isolating mechanisms produced by reinforcement in areas of sympathy in their native ranges. Alternatively, specific ecological or environmental features constituting barriers to hybridization in the native ranges may be lacking in the introduced range, thus compromising reproductive isolation caused by extrinsic factors. Finally, the genetic bottlenecks that each species experienced upon introduction to the USA may have compromised genetically based mate recognition systems that enforce reproductive isolation in the native ranges. Regardless of the nature of the premating barriers that have been compromised, the fact that post-mating barriers between the species are incompletely developed provides the opportunity for in-depth study of the genomics of hybrid incompatibility in a wild setting.

Previous genetic, biochemical and morphological studies of the large zone of hybridization between S. invicta and S. richteri in the USA revealed that it features a gradual changeover in allele frequencies and character sizes from those typical of S. invicta in the south to those typical of S. richteri in the north (Ross et al. 1987; Ross & Robertson 1990; Shoemaker et al. 1996). Highly admixed hybrid genotypes predominate in the centre of the zone, consistent with the occurrence of hybridization over many generations. Moreover, genotypes in these highly admixed populations typically occur at frequencies expected under Hardy–Weinberg equilibrium, and overall significant pairwise linkage disequilibria are effectively absent. Although large-scale clinal changes in genetic and phenotypic characters are the norm
along study transects spanning the zone, pockets of pure parental forms and hybrids of varying genetic constitution are interspersed throughout it; this mosaic small-scale distribution may reflect chance historical colonization of suitable habitat patches or preferential nesting success of hybrids of varying genomic composition among ecologically distinct microhabitats (Shoemaker et al. 1996; the parental species occupy different habitats in their native ranges – Trager 1991). The ubiquity of advanced-generation hybrid genotypes in the USA hybrid zone, combined with the general lack of single-locus and linkage disequilibria, suggests not only that F1 hybrids are viable and fertile, but also that any general breakdown in hybrid fitness may be modest.

The possibility that hybrid incompatibility may cause disrupted development in hybrid S. invicta/S. richteri was earlier studied by comparing the developmental stability of several queen morphological characters between hybrids and the parental species (Ross & Robertson 1990). Hybrid ants displayed significantly reduced developmental stability, but the effect was weak. Thus, genomic incompatibility between the parental species may create only modest viability selection against hybrid genotypes that is not sufficient to preclude extensive hybridization. Moreover, the finding that some diagnostic genetic markers introgress more freely across the hybrid zone than species-typical morphological traits (Shoemaker et al. 1996) suggests that genes underlying the latter recombine less readily into a heterospecific genomic background, consistent with mild intrinsic fitness penalties to at least some recombinant genotypes.

A previous study of genome-wide expression in S. invicta and S. richteri revealed low divergence in the expression patterns of the two species (Ometto et al. 2011). However, this differentiation varies greatly between castes and developmental stages, with the largest differences in gene expression occurring between adult workers of the two species. The aim of this study was to determine levels of gene expression in hybrids to learn whether the various castes, sexes and life stages also differ in the degree of compatibility between the S. invicta and S. richteri regulatory machineries. In particular, we were interested in investigating whether differences in life history between castes result in greater numbers of genes being mis-expressed in workers than in males and queens. Because fire ants workers do not reproduce, and thus have only an indirect fitness component, gene expression may be under more relaxed selection in this caste than in queens and males (Links-vayer & Wade 2009; Ometto et al. 2011). If we assume that regulation of gene expression has a genetic basis and is under stronger purifying selection in reproductively competent individuals, genes preferentially expressed in workers should be regulated by elements exhibiting a greater between-species divergence than elements imposing similar gene expression levels across castes. Moreover, there should also be relatively greater interspecific divergence in the patterns of selection acting on workers; this is because adult members of this caste most directly experience the distinct environments characterizing the different ranges occupied by the two species (namely, tropical and subtropical habitats for S. invicta and somewhat more temperate habitats for S. richteri – Trager 1991), whereas queens and males are almost exclusively confined to the buffered, within-nest environment. As a result, we predict that there should be more genes mis-expressed in workers than in queens and males. Similarly, pupae are entirely confined within the nest, where they are cared for by the workers, and therefore, they presumably also experience more similar environments across the two species than do adults, leading to the prediction that pupae should have fewer genes mis-expressed than adults.

Materials and methods

We analysed gene expression levels of workers, queens and males in the pupal and adult stages of S. invicta X S. richteri hybrids as well as the two parental species. As typically is the case in ants, it is impossible to conduct controlled matings with fire ants under laboratory conditions. Therefore, colonies of hybrids were collected directly from the field in Grenada and Alcorn Counties, Mississippi, USA (a central location in the hybrid zone) in May of 2005 and 2007. Upon return to the laboratory, hybrid as well as pure-species colonies were maintained in a single rearing room under identical standard conditions, including the same artificial nests, diets, and temperature and humidity regimes (Jouvenaz et al. 1977).

Taxonomic status of every study colony was confirmed by genotyping 8–10 individuals from each at five informative allozyme loci that feature strong or fixed allelic differences between the species (see Shoemaker et al. (1996) for details of the allozyme methods).

We conducted analyses of gene expression patterns in hybrids as previously described for pure S. invicta and S. richteri (Ometto et al. 2011). Briefly, we randomly sampled, and pooled, 1–6 nestmate individuals (mean = 4.6, median = 5.0) for each replicate of the eighteen specific categories, or nodes, of interest (two developmental stages X 3 castes X 3 taxa (two species + one hybrid); Table S1, Supporting information). Specimens of both developmental stages were standardized by age (pupae – first appearance of pink eyes; adults – <12 h posteclosion), and specimens of the worker caste were further standardized by including only majors so as to minimize possible allometric
differences associated with intracaste polyphenism, which could affect gene expression levels. The cDNA derived from each experimental sample was hybridized against a common reference RNA on our custom-made spotted cDNA microarray after being randomly labelled with either Cy3 or Cy5 (Wang et al. 2007; Wurm et al. 2009). Because we used each colony only once for a given caste and developmental stage, each sample was equivalent to an independent biological replicate. Microarray intensity analyses were performed using the software package limma in R (R Development Core Team 2009). The power of the between-array normalization was enhanced by adding data from arrays hybridized with cDNA originating from larvae, which underwent the same experimental protocols as the samples used in the present study and in Ometto et al. (2011). Raw data for all 140 hybridizations (i.e. pure species and hybrid data for workers, queens and males at the larval, pupal and adult stages) are available in the Gene Expression Omnibus database under accession number GSE35217.

The Bayesian approach implemented in the program Bayesian Analysis of Gene Expression Levels (BAGEL; Meiklejohn & Townsend 2006) was used to estimate the relative expression level of each clone for each node (stage/caste/taxon) and the significance of the differential expression of such clones among nodes (data deposited in the Dryad repository: doi:10.5061/dryad. m0r5qv24). Significance was estimated after correcting for multiple testing using the false discovery rate approach as described in Ometto et al. (2011). If at least one of the clones belonging to the same contig was differentially expressed, the contig was considered differentially expressed as well. For consistency with our previous study (Ometto et al. 2011), we report only results based on contigs, and we use contig synonymously with gene throughout the remainder of this text. Following Ometto et al. (2011), a given gene was categorized as caste-biased in its expression when significantly over- or under-expressed in one caste compared to both of the other castes for a given developmental stage.

Some regulatory mechanisms/factors that have diverged between S. invicta and S. richteri may not necessarily translate into different gene expression levels or different phenotypes between the two species (True & Haag 2001). However, such divergence could be apparent in hybrids that contain a mix of different and possibly incompatible regulatory components (Riddle & Birchler 2003), resulting in consistently over- or under-expressed transcripts relative to the pure parental species (e.g. Michalak & Noor 2003; Ranz et al. 2004; Haerty & Singh 2006). We therefore quantified the degree of hybrid incompatibilities in gene regulation by comparing the level of expression of each gene between hybrids and each of the two parental species. Genes were defined as mis-expressed when either significantly over- or under-expressed in hybrids compared to both parental species. These comparisons were performed separately for each developmental stage and caste.

The relative expression of each contig (measured as the average of the values calculated by BAGEL for the associated clones) was used to estimate the relative importance of developmental stage, caste and taxonomic identity on the observed gene expression levels across all genes. First, we estimated pairwise gene expression distances between each of the eighteen nodes and constructed the corresponding distance tree as described in Ometto et al. (2011). In a second approach, we performed principal components analyses on the scaled data matrix using the singular value matrix decomposition approach of the prcomp package implemented in R (R Development Core Team 2009).

The normalized gene expression data were used to estimate the variability in gene expression, $V_X$, within S. invicta, S. richteri, and hybrids, and to evaluate $H_p$, the degree of similarity in gene expression between hybrids and the two parental species. Specifically, $V_X$ was estimated as the ratio between the mean and the standard deviation of the log$_2$ ratios of the normalized dye intensities calculated across all samples for the species, caste and developmental stage of interest (see Ometto et al. (2011) for details of the methods). The similarity index $H_p$ was estimated for each clone and hybrid sample as, $\frac{(\bar{r}_{i} - \bar{r}_{h})}{\max_{i} |\bar{r}_{i}| - \min_{i} |\bar{r}_{i}|}$, where $h$ is the normalized gene expression for the sample of interest, $i$ and $r$ are the across-samples mean expression values for S. invicta and S. richteri, respectively.

We used gene ontology (GO) terms to test whether genes differentially expressed between hybrids and the parental species were over-represented (enriched) for some categories. Because no GO terms are available for S. invicta, we used the Blast2GO web tool (Conesa et al. 2005) to functionally annotate the S. invicta clone sequences available from Fourmidable, http://fourmidable.unil.ch (Wang et al. 2007; Wurm et al. 2009). Searching within the National Center for Biotechnology Information (NCBI) database, the tool could assign a putative orthologue to 10 527 clones, 6197 of which had associated GO terms. In a second approach, mapping information (Hunt et al. 2011) allowed us to assign 4767 of the 22 856 clones spotted onto the microarray to 2066 S. invicta genes (official gene set version 2.2.0, available from Fourmidable, http://fourmidable.unil.ch; Wurm et al. 2009, 2011). We then searched for putative orthologues in the well-annotated Drosophila melanogaster genome (official gene set release 5.9, available at http://flybase.org – Tweedie et al. 2009).
Orthology was determined by BLASTp searches (Camacho et al. 2009), and custom Perl scripts were used to identify a total of 1403 reciprocal best hits, which were retained as pairs of putative orthologues. The annotations of the D. melanogaster orthologues were subsequently used to search for potential enrichment by GO terms using the DAVID bioinformatics tool (Huang et al. 2009). Functional prediction based on orthology may be compromised when conducted with distantly related species (e.g. fruit flies and fire ants) at the single gene level (e.g. Nehrt et al. 2011), but this potential problem should only minimally affect global conclusions.

Results and Discussion

Principal components analysis revealed a strong contribution of developmental stage to gene expression, with the first component, on which this variable loads heavily, explaining 42% of the total gene expression variability (Fig. S1, Supporting information). The second, third and fourth components, associated with caste and sex, together accounted for 39% of the expression variability. The fifth component, explaining 5% of the gene expression variability, was not distinctly associated with any of the variables. Only the sixth component, which accounts for just 4.2% of the expression variability, was associated with taxonomic identity, with hybrid scores falling between the two parental species’ scores in both life stages of all castes (Fig. S1, Supporting information). The fact that taxonomic group explains so little of the variability in the expression data implies strong conservation in gene expression patterns between the two parental species. This conclusion also is supported by the close clustering of hybrids with the two parental species within each stage and caste in the gene expression-based distance tree (Fig. 1).

Gene expression differentiation between hybrids and the parental species

Explicit comparisons of gene expression levels between hybrids and the two parental species support the view that the conserved expression patterns translate into limited incompatibilities in gene regulation between S. invicta and S. richteri. In male pupae, there were only 147 genes (1.0% of the 14,467 genes analysed) that were differentially expressed between hybrids and either of the two parental species (Fig. 2). In female pupae, the numbers were similarly low, with 239 (1.7%) genes differentially expressed in hybrid queens and 319 (2.2%) in hybrid workers. In adults, there were slightly more genes differentially expressed between hybrids and the parental species, with 310 (2.1%) such genes in males, 247 (1.7%) in queens and 662 (4.6%) in workers (Fig. 2).

When considering all three castes in aggregate for the pupal stage, 4.0% of genes were differentially expressed between hybrids and S. invicta and 1.1% between hybrids and S. richteri. A similar comparison previously revealed that 5.8% of genes were differentially expressed between pupae of the two parental species (Ometto et al. 2011). In adults, the patterns were similar, with 5.4% of genes differentially expressed in the hybrid-invicta comparison, 3.5% in the hybrid-richteri comparison and 10.7% in the invicta-richteri comparison. Importantly, our finding of a lower percentage of genes being differentially expressed between hybrids and either parental species than between the two species is conservative, because the false discovery rate for the hybrid vs. parental species analyses is at least twice as high as the false discovery rate for the interspecific analyses for both life stages (data not shown; Ometto et al. 2011). This finding of much lower proportions of genes being differentially expressed between hybrids and either parental species than between the species (about 1/2) suggests relatively few incompatibilities and thus quite recent divergence between S. invicta and S. richteri. For comparison, approximately equal numbers of genes are differentially expressed between F1 Drosophila simulans X Drosophila sechellia hybrids and the parental species as between the two species (Haerty & Singh 2006; Moehring et al. 2007; Wurmsen et al. 2011), which evidently separated some 250 000 years ago (Mcdermott & Kliman 2004). Moreover, about twice as many genes are differentially expressed between F1 Drosophila melanogaster X D. simulans hybrids and the parental species as between the two species (Ranz et al. 2004), which are
estimated to have diverged about 2–3 million years ago (Hey & Kliman 1993). While direct extrapolation of times of divergence across these insect orders is not appropriate, the fire ant regulatory element divergence clearly is comparatively low. This is in line with other evidence that the clade to which our two study species belong (the Solenopsis saevissima species group) is a youthful group in an active phase of evolutionary radiation (Ross & Trager 1990; Shoemaker et al. 2005).

One factor complicating explanations of the relatively low divergence in gene expression levels between the fire ant hybrids and the two parental species is that the sampled hybrids were not F1 hybrids, but rather a mix of various advanced-generation hybrids and backcrosses. Thus, individuals (colonies) with high levels of hybrid incompatibilities may have been selectively purged in the wild and thus unavailable for sampling. While this effect may contribute to an underestimate of the extent of transcriptome divergence between S. invicta and S. richteri, such estimates are most relevant to the ecological context of natural hybridization. Clearly, the ecologically effective level of genetic incompatibilities between the two parental genomes is sufficiently low that (i) hybrids are widespread and abundant in the introduced (USA) range; (ii) introgression has proceeded through multiple generations to yield highly admixed recombinant genotypes that predominate in the hybrid zone centre; and (iii) developmental stability has been little disrupted in extant hybrid queens (Ross & Robertson 1990; Shoemaker et al. 1996).

**Mis-expressed genes in hybrids**

Consistent with the finding of low divergence in gene expression profiles between hybrids and the parental species, we also found low levels of disruption in gene regulation in hybrids. There were only 16 mis-expressed genes (i.e. genes significantly over- or under-expressed in hybrids compared to both parental species) in pupae and 18 in adults (Fig. 2), only one of which was mis-expressed in both stages. When analysed individually, however, the three castes differed greatly in the number of genes mis-expressed, with the male and worker castes each mis-expressing at least seven times more genes than the queen caste in both life stages (goodness-of-fit test comparing the observed data to the expectation of equal numbers of mis-expressed genes across castes; \( P < 0.05 \) for all such comparisons; \( P > 0.8 \) for the comparisons between males and workers for both stages). The haploid genomic state of males may contribute to the large number of mis-expressed genes in this caste, because haploidy should result in admixed males expressing all recessive incompatibilities after the F1 female generation (i.e. after a first round of meiotic recombination between the paternal and maternal homologues; Haldane 1922; Orr 1993b; Turelli & Orr 1995; Schilthuizen et al. 2011; Koevoets et al. 2012). On the other hand, the significant difference in the number of mis-expressed genes between the two female castes can only be explained by higher levels of regulatory incompatibilities in hybrid workers than in hybrid queens. This finding is in line with the prediction that, because workers are sterile and experience more divergent environments than queens and males, they should exhibit higher levels of regulatory incompatibilities (see Introduction). The observed lower number of differentially expressed (as opposed to mis-expressed) genes in hybrid pupae than in hybrid adults, particularly notable

---

**Fig. 2** Venn diagrams depicting numbers of genes significantly overexpressed (>) or under-expressed (<) in pupal and adult hybrids (hyb) relative to Solenopsis invicta (inv) and Solenopsis richteri (ric). Blue-shaded overlapping areas identify genes consistently over- or under-expressed in hybrids relative to both parental species (i.e. mis-expressed genes). Grey-shaded overlapping areas identify genes that are significantly overexpressed in hybrids compared to one parental species while significantly under-expressed compared to the other. Sizes of circles are proportional to the numbers of genes.
in males and workers (Fig. 2), agrees with the greater conservation of gene expression profiles observed in pupae of the pure species compared to adults (Ometto et al. 2011); these patterns are consistent with pupae experiencing a more buffered environment than adults as well as with a generally greater conservation of regulatory networks involved in early than in late development during cladogenesis (Artieri & Singh 2010; Domazet-Loso & Tautz 2010).

One potentially confounding factor in interpreting patterns of mis-expression based on whole-body mRNA extracts, such as used in our study, is the possible atrophy of various tissue/organ systems owing to hybrid breakdown, one result of which may be differences in mRNA abundance attributable solely to gross differences in tissue abundance between hybrids and pure-species individuals (e.g. Ranz et al. 2004; Blumenstiel & Hartl 2005). While we cannot entirely rule out a role for such effects, several lines of evidence suggest they may not be important. First, these effects are not expected to produce the commonly observed pattern of hybrid fire ant expression exceeding parental expression, barring the occurrence of hypertrophy rather than atrophy of hybrid tissues/organs. Second, fire ant workers effectively lack a reproductive system (Tschinkel 2006), so that the atrophy of reproductive tissues in hybrids commonly underlying artifactual expression differences between them and parentals (Hollocher et al. 2000; Ranz et al. 2004; Graveley et al. 2011) is not possible in this caste. Finally, our hybrid study samples did not generally exhibit greater variation in expression levels across replicate colonies than did the parental species, as would be predicted given that the hybrids represent various levels of backcrossing and assuming that this would yield corresponding variation in tissue dystrophy. Among the six developmental stage/caste combinations, the degree of gene expression variability, $V_X$, in hybrids significantly surpassed that in both parental species in only half of the cases (pupal workers and males and adult queens; Wilcoxon test, all $P < 0.001$), suggesting minimal general effects of any hybrid tissue atrophy/hypertrophy on expression differences.

Functional characterization of mis-expressed genes

Functional annotation of the 33 mis-expressed genes was conducted by searching for their putative orthologues and related GO terms in available annotated genomes. In a first approach, we probed the sequences of the clones associated with the mis-expressed genes against the National Center for Biotechnology Information database using the Blast2GO tool (Conesa et al. 2005). We could assign putative orthologues to 14 of the 33 mis-expressed genes, most of which matched genes from ants (Table 1) (Bonasio et al. 2010; Nygaard et al. 2011; Wurm et al. 2011). It was possible to associate GO terms with only five of these genes, and no significant over-representation of any functional class was apparent in these annotation data. The limited success of our functional annotation efforts may stem from the fact that clone sequences represent only partial genes, which reduces the efficiency of the orthologue searches.

In a second approach specifically aimed at overcoming this limitation and targeting well-annotated genomes, we were able to identify seven of the 33 mis-expressed genes by reference to the S. invicta genome, and five of these were inferred to have an identifiable D. melanogaster orthologue (Table 1). Two of the S. invicta genes matched the same D. melanogaster orthologue, suggesting a possible duplication event in an ancestral fire ant lineage and reducing the number of putative orthologue pairs to four. Two of these four genes are annotated in D. melanogaster as being involved in nitrogen compound biosynthetic processes, a significant over-representation of this gene class according to the gene ontology analysis of the S. invicta-D. melanogaster orthologue pairs (Fisher’s exact test; $P = 0.03$).

Interestingly, four of the genes for which we could infer putative functions have been implicated as being involved in behavioural variation in D. melanogaster (Table 1). For instance, gene CG3011 is related to the response to ethanol exposure (Kong et al. 2010), while gene $ade5$ is involved in between-male aggression (Edwards et al. 2009). The S. invicta orthologues of both genes were under-expressed in hybrids compared to both parental species in adult workers. The third gene codes for the cuticular protein 72Ec, while the fourth gene is a putative homologue of the odorant/hormone-binding gene CG14661 in D. melanogaster. Both of these genes show circadian oscillations in D. melanogaster (Ceriani et al. 2002; Wijnen et al. 2006), and in Solenopsis, they were overexpressed in hybrids compared to both parental species (Cpr72Ec in male pupae and CG14661 in worker pupae). Given the paramount importance of behavioural interactions in the social lives of ants, both between adults and between pupae and nursing adult workers, mis-expression of genes influencing the performance of, or responses to, behaviours can be expected to have profound fitness consequences in hybrid fire ants not necessarily evident at the physiological or morphological levels (cf. Ross & Robertson 1990).

Asymmetries in hybrid gene expression profiles

The expression profiles of hybrids in both developmental stages were significantly more similar to those of
<table>
<thead>
<tr>
<th>Name*</th>
<th>Mis-type†</th>
<th>Selenops invicta gene</th>
<th>Blast2GO‡</th>
<th>Gene ontology (GO) terms/function</th>
<th>Gene ontology (GO) terms/function</th>
</tr>
</thead>
<tbody>
<tr>
<td>SijWA08BAL</td>
<td>+aM</td>
<td>Hypothetical protein G5I_14350 (Acromyrmex echinatior)</td>
<td>Binding</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>SijWB02BCV</td>
<td>-aW</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>SijWB10AEA</td>
<td>-aW</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>SijWC03ACC</td>
<td>+aM</td>
<td>Alpha-tocopherol transfer (S. invicta)</td>
<td>Transporter activity</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>SijWC04BCM</td>
<td>+pW</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>SijWC04BDK</td>
<td>-pM</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>SijWC07ADV</td>
<td>-pM</td>
<td>SI2.2.0_06746</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>SijWC10CAM</td>
<td>+aM</td>
<td>SI2.2.0_05389</td>
<td>Hypothetical protein SINV_05389 (S. invicta)</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>SijWD04BCT</td>
<td>-pM</td>
<td>SI2.2.0_06746</td>
<td>Circadian clock-controlled protein (S. invicta)</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>SijWD10ACV</td>
<td>+aM</td>
<td>NA</td>
<td>Transmembrane protein 39A-A-like (Acromyrmex echinatior)</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>SijWD10BAK</td>
<td>+pW</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>SijWE06BAD</td>
<td>+aW</td>
<td>SI2.2.0_07244</td>
<td>Multifunctional protein ade2 (S. invicta)</td>
<td>Nitrogen compound biosynthetic process</td>
<td>Nitrogen compound biosynthetic process; involved in intermale aggression</td>
</tr>
<tr>
<td>SijWE10AAV</td>
<td>-aW</td>
<td>SI2.2.0_09632</td>
<td>Serine hydroxymethyltransferase (S. invicta)</td>
<td>L-serine metabolic process; methylation; nitrogen compound biosynthetic process</td>
<td>CG3011 Serine hydroxymethyltransferase; nitrogen compound biosynthetic process; related to the response to ethanol exposure</td>
</tr>
<tr>
<td>SijWE12BAI</td>
<td>+pM</td>
<td>NA</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>SijWE12BDW</td>
<td>+aW</td>
<td>NA</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>SijWF01ABR</td>
<td>-aQ</td>
<td>SI2.2.0_09632</td>
<td>Serine hydroxymethyltransferase (S. invicta)</td>
<td>L-serine metabolic process; methylation; nitrogen compound biosynthetic process</td>
<td>CG3011 Serine hydroxymethyltransferase; nitrogen compound biosynthetic process; related to the response to ethanol exposure</td>
</tr>
<tr>
<td>SijWF08BBH</td>
<td>-aM</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Name*</td>
<td>Mis-type†</td>
<td>Solenopsis invicta gene</td>
<td>Gene</td>
<td>Gene ontology (GO) terms/function</td>
<td>Blast2GO‡</td>
</tr>
<tr>
<td>-------------</td>
<td>-----------</td>
<td>-------------------------</td>
<td>---------------</td>
<td>-----------------------------------</td>
<td>-----------</td>
</tr>
<tr>
<td>SijWF09BDN</td>
<td>–aW</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>SijWG01BBA</td>
<td>–aW</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>SijWG04AAG</td>
<td>–aW</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>SijWG06CAG</td>
<td>pM</td>
<td>S12.2.0_01800</td>
<td>PR domain zinc finger protein 13 (S. invicta)</td>
<td>Zinc ion binding; nucleic acid binding</td>
<td>NA</td>
</tr>
<tr>
<td>SijWG09BCP</td>
<td>–pW</td>
<td>S12.2.0_13401</td>
<td>Cuticle protein 6 (S. invicta)</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>SijWH02BCR</td>
<td>–aM</td>
<td>NA</td>
<td>Hypothetical protein SINV_00721 (S. invicta)</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>SijWH03AAAT</td>
<td>–pQ</td>
<td>NA</td>
<td>Hypothetical protein SINV_16046 (S. invicta)</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>SijWH04AAK</td>
<td>–pW</td>
<td>NA</td>
<td>Hypothetical protein EAI_10442 (Harpegnathos saltator)</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>SijWH04ACW</td>
<td>–pW</td>
<td>S12.2.0_04262</td>
<td>Hypothetical protein EAI_10442 (Harpegnathos saltator)</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>SijWH05BCO</td>
<td>–p/aM</td>
<td>NA</td>
<td>Chymotrypsin-2 (S. invicta)</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>SijWH06CAH</td>
<td>–aM</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>SijWH07ACG</td>
<td>–pM</td>
<td>NA</td>
<td>WD-repeat-containing protein 16 (S. invicta)</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>SijWH09ABR</td>
<td>–pW</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>SijWH09ADK</td>
<td>–pW</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>

NA, not available.

*Name of the clone spotted onto the microarray (Wurm et al. 2009).
†Type of mis-expression pattern of the clone. The three-letter code indicates whether a clone was overexpressed (+) or under-expressed (−) in hybrids compared to both parental species in pupal (p) or adult (a) workers (W), queens (Q), or males (M).
‡Putative orthologues of the clones and associated GO terms.
§Putative orthologues of the *S. invicta* genes identified by a reciprocal best hit approach in *D. melanogaster* and associated GO terms.
S. richteri than S. invicta (goodness-of-fit test, both \( P < 10^{-13} \)). In pupae, the degree of hybrid expression bias towards S. richteri was strong in all three castes (Fig. 2). In queens of this stage, more than twice as many genes were differentially expressed between hybrids and S. invicta than between hybrids and S. richteri (127 vs. 28 for males and 276 vs. 56 for workers; both \( P < 10^{-11} \)), while in males and workers, there were four times more genes differentially expressed between hybrids and S. invicta than between hybrids and S. richteri (177 vs. 41; \( P < 10^{-15} \)) and adult males (261 vs. 617; \( P < 10^{-15} \)), the expression bias in hybrid adult workers was towards S. invicta, with more genes differentially expressed between hybrids and S. richteri than between hybrids and S. invicta (398 vs. 300; \( P < 10^{-15} \)). The distinctive gene expression profile of hybrid adult workers, which departed from the overall pattern of bias towards S. richteri, was further examined by two additional analyses.

In the first analysis, we compared the ratio of expression for hybrids and S. richteri to the ratio for hybrids and S. invicta (Fig. 3). The expression levels in hybrid pupae of each of the three castes were closer to those of S. richteri than to those of S. invicta (paired Wilcoxon test; \( P < 10^{-13} \) for all three castes). The same pattern was found for adult queens (\( P < 10^{-10} \)) but was reversed for adult males (\( P < 10^{-10} \)). No significant pattern emerged for adult workers (\( P = 0.383 \)).

In the second analysis, we examined the degree of conservation of caste-biased gene expression between hybrids and each parental species by comparing the numbers of genes over- and under-expressed in a particular caste in both hybrids and S. invicta to the numbers of genes over- and under-expressed in that same caste in both hybrids and S. richteri. In pupae, genes that were caste-biased in hybrids were significantly more likely to be caste-biased as well in S. richteri than in S. invicta (Fisher’s exact tests, for overexpressed genes \( P < 0.0001 \) for workers, queens and males; for under-expressed genes \( P < 0.05 \) for workers and queens and \( P > 0.17 \) for males; Figs S2a and S3a, Supporting information respectively). In adults, such differences, while apparent for each caste, were statistically significant only for overexpressed genes in males (\( P < 10^{-9} \); \( P > 0.183 \) for overexpressed genes in workers and queens and \( P < 0.05 \) for under-expressed genes in all three castes; Figs S2b and S3b, Supporting information). Overall, these results can be taken to indicate different interactions of the parental regulatory machineries both among castes and between developmental stages.

Two hypotheses may explain the finding that hybrids generally display gene expression patterns closer to S. richteri than to S. invicta. The first is directly linked to the relative contribution of S. richteri to the hybrid genome pool. The allozyme data revealed an overall bias in the genetic composition of the hybrid colonies towards S. richteri; assuming independence of the five loci analysed, the average hybrid index (Buerkle 2005) was 0.69 (range, 0.58–0.94 for the seven colonies), where a value of 0 corresponds to pure S. invicta and a value of 1 to pure S. richteri. Thus, it is probable that the hybrids analysed in this study had a larger fraction of their genomes originating from S. richteri than from S. invicta. In females, this would mean that an increased fraction of loci were homozygous for S. richteri cis or trans elements regulating the expression of associated genes compared to a population with a hybrid index of 0.5, and hence, the fraction of genes with expression patterns most similar to those of S. richteri was seen in hybrid pupae of each of the three castes, while in workers and males, there were four times more genes differentially expressed between hybrids and S. invicta than between hybrids and S. richteri (Fig. 2).

---

**Fig. 3** Ratios between the relative gene expression levels in hybrids and either Solenopsis invicta (inv) or Solenopsis richteri (ric). Values closer to one indicate greater similarity of gene expression levels between hybrids and a parental species. Error bars denote the standard error of the mean. Paired Wilcoxon test, **** \( p < 0.0001 \), NS = not significant.
S. richteri would exceed the fraction similar to S. invicta. To further test this hypothesis, we conducted correlation analyses between the colony-specific hybrid index and $H_M$, an index of similarity in expression between hybrids and each parental species. Correlations were not significant ($P > 0.13$ for pupae and adults in all three castes). This indicates either that the genomic composition has only a marginal effect on the gene expression patterns of hybrids, or perhaps more likely, that the five allozyme loci in aggregate do not provide sufficient information to accurately predict the genome pool composition of individual colonies (and thus cannot be used to reliably test whether the bias in gene expression patterns of hybrids towards S. richteri stems from a higher contribution of S. richteri to the hybrid genome pool).

The second hypothesis is that S. richteri trans-regulatory elements are on average dominant to those of S. invicta, thus imposing an expression profile closer to that of S. richteri even with relatively even mixes of the two genomes. In support of this, the gene expression-based distance tree revealed similar overall divergence in gene expression patterns of hybrids from those of each parental species (Fig. 1), in contrast to the greater similarity of hybrids to S. richteri than S. invicta at the relatively small fraction of genes with significantly divergent expression in hybrids. Asymmetry in gene expression differences towards one of the parental species has also been observed in hybrids between sympatric anadromous and resident populations of brook charr (Mavarez et al. 2009) and in hybrids between lake whitefish species pairs (Renaut et al. 2009), pointing to the common existence of complex interactions between parental species’ regulatory machineries when divergent genomes are admixed (Gibson & Weir 2005; Rockman & Kruglyak 2006).

Under both hypotheses, the asymmetry in expression profiles might be expected to hold across all three castes in both developmental stages, yet our results show that adult workers consistently differ in this regard from the other classes of individuals. This incongruity may stem from either of two nonmutually exclusive causes. First, adult worker gene expression may be under the control of regulatory elements that generally are more sensitive to genetic perturbations, such that expression levels in these individuals often differ from those of the parental species irrespective of the species origin of the regulatory elements. Alternatively, purifying selection may operate less efficiently in workers than in queens and males, resulting in delayed purging of incompatibilities in this caste as introgression proceeds (Linksvayer & Wade 2009). This explanation is appealing because workers are obligately sterile in these ants and so possess only indirect fitness components manifested by the effects of their activities on the survival and reproduction of related reproductive nestmates.

Conclusion

This study reveals evidence of surprisingly modest hybrid incompatibilities in gene expression between two invasive fire ant species, S. invicta and S. richteri. This low level of incompatibility probably accounts in part for the widespread and persistent hybridization between the two species in the USA, where premating barriers have been compromised following their introductions in the early part of the last century. This study further highlights significant heterogeneity in the degree of hybrid incompatibilities across life stages and castes, including elevated numbers of genes misexpressed in workers compared to males and queens and discordance among classes of individuals in the similarity of their gene expression profiles to those of each parental species. Altogether, these results suggest that selection pressures acting specifically on each caste may contribute differently to interspecific divergence and the processes of speciation in these ants as well as in other social insects.

Acknowledgements

We thank Christine La Mendola and the Lausanne Genomic Technologies Facility for technical support, the Keller laboratory for helpful discussions and two reviewers for helpful comments on the manuscript. This work was supported by grants from the Roche Research Foundation (to L.O.) and Swiss NSF (L.K.) as well as an ERC advanced grant (L.K.).

References


Linksvayer TA, Wade MJ (2009) Genes with social effects are expected to harbor more sequence variation within and between species. Evolution, 63, 1685–1696.


L. Ometto’s main research focus is the molecular basis of adaptation. This work is part of his post-doctoral project on the evolution of gene expression in fire ants. K. Ross’s research focuses on the evolutionary genetics of social Hymenoptera and other insects, with much of his work concerning native and introduced fire ant populations. The main research interest of D. Shoemaker is the population and evolutionary genetics of various insects, focusing on fire ants and Wolbachia-insect interactions. L. Keller works on various aspects of evolutionary ecology and social behavior in ants.

**Data accessibility**

Microarray data are deposited in NCBI’s Gene Expression Omnibus database under accession number GSE35217.

BAGEL output data is deposited in the Dryad repository: doi:10.5061/dryad.m0r5qv24.

**Supporting information**

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

**Fig. S1** Principal components analysis of gene expression levels. Three-letter labels identify taxonomic group (h = hybrids; i = *S. invicta*, r = *S. richteri*), caste (M = males; Q = queens; W = workers), and developmental stage (p = pupae; a = adults). Percentages of variability explained by each component are shown in parentheses.

**Fig. S2a** Venn diagrams depicting numbers of genes overexpressed in pupal males, workers, and queens of hybrids (hyb), *S. invicta* (inv), and *S. richteri* (ric). Overlapping areas indicate differentially expressed genes common to two or three of the taxonomic groups. Circle size is proportional to the number of genes.

**Fig. S2b** Venn diagrams depicting numbers of genes overexpressed in adult males, workers, and queens of hybrids (hyb), *S. invicta* (inv), and *S. richteri* (ric). See Fig. S2a (Supporting information) caption for additional explanation.

**Fig. S3a** Venn diagrams depicting numbers of genes underexpressed in pupal males, workers, and queens of hybrids (hyb), *S. invicta* (inv), and *S. richteri* (ric). See Fig. S2a (Supporting information) caption for additional explanation.

**Fig. S3b** Venn diagrams depicting numbers of genes underexpressed in adult males, workers, and queens of hybrids (hyb), *S. invicta* (inv), and *S. richteri* (ric). See Fig. S2a (Supporting information) caption for additional explanation.

**Table S1.** Complete list of the microarray hybridizations performed. Each hybridization contrasted the expression levels in the sample of interest with those in a common reference.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.