Ants exhibit asymmetric hybridization in a mosaic hybrid zone

JESSICA PURCELL,*† SACHA ZAHND,* ANOUK ATHANASIADES,* REBECCA TÜRLER,* MICHEL CHAPUISAT*[1] and ALAN BRELSFORD*‡[1]
*Department of Ecology and Evolution, University of Lausanne, 1015 Lausanne, Switzerland, †Department of Entomology, University of California, Riverside, Riverside, CA 92521, USA, ‡Department of Biology, University of California, Riverside, Riverside, CA 92521, USA

Abstract

Research on hybridization between species provides unparalleled insights into the pre- and postzygotic isolating mechanisms that drive speciation. In social organisms, colony-level incompatibilities may provide additional reproductive barriers not present in solitary species, and hybrid zones offer an opportunity to identify these barriers. Here, we use genotyping-by-sequencing to sequence hundreds of markers in a hybrid zone between two socially polymorphic ant species, *Formica selysi* and *Formica cinerea*. We characterize the zone, determine the frequency of hybrid workers, infer whether hybrid queens or males are produced and investigate whether hybridization is influenced by colony social organization. We also compare cuticular hydrocarbon profiles and aggression levels between the two species. The hybrid zone exhibits a mosaic structure. The asymmetric distribution of hybrids skewed towards *F. cinerea* suggests a pattern of unidirectional nuclear gene flow from *F. selysi* into *F. cinerea*. The occurrence of backcrossed individuals indicates that hybrid queens and/or males are fertile, and the presence of the *F. cinerea* mitochondrial haplotype in 97% of hybrids shows that successful F1 hybrids will generally have *F. cinerea* mothers and *F. selysi* fathers. We found no evidence that social organization contributes to speciation, because hybrids occur in both single-queen and multiple-queen colonies. Strongly differentiated cuticular hydrocarbon profiles and heightened interspecific aggression further reveal that species recognition cues are both present and perceived. The discovery of fertile hybrids and asymmetrical gene flow is unusual in ants, and this hybrid zone will therefore provide an ideal system with which to investigate speciation in social insects.

Keywords: bimodal hybrid zone, formicinae, genomic cline, mosaic hybrid zone, reproductive isolation, speciation, supergene

Received 13 March 2016; revision received 13 July 2016; accepted 14 July 2016

Introduction

Eusocial insects form several speciose clades, yet researchers still know little about the reproductive barriers leading to speciation in these lineages. Due to their social group structure, speciation and reinforcement processes in eusocial organisms have the potential to differ from those of nonsocial organisms (e.g. Kulmuni *et al.* 2010; Kulmuni & Pamilo 2014). Individual-level incompatibilities, including pre- and postzygotic isolating mechanisms, should contribute to reproductive barriers in both social and solitary organisms. Social species could also have novel barriers due to colony-level incompatibilities. These may include diverging social strategies, such as transitions towards increasing queen numbers, which are often accompanied by a shift from nuptial flight mating to intranidal mating (e.g.

Correspondence: Jessica Purcell, Fax: +1 951 827 3086; E-mail: Jessica.Purcell@unil.ch
[1]Co-last authors

Seifert 2010). Colony-level incompatibilities may also involve intracolony conflicts between nestmate queens mated with conspecific *versus* heterospecific males (Helms Cahan & Vinson 2003). More research is needed to determine the prevalence of colony-level mechanisms in speciation and reproductive isolation (Steiner *et al.* 2011).

Hybrid zones can provide insights into reproductive barriers that develop between diverging species (reviewed by Abbott *et al.* 2013). In eusocial insects, hybridization is relatively common and widespread (reviewed by Feldhaar *et al.* 2008), with well-documented examples between subspecies of honeybees (Rinderer *et al.* 1991; Chávez-Galarza *et al.* 2015), in termites (Lefebvre *et al.* 2008) and in many ant genera (e.g. Ross *et al.* 1987; Shoemaker *et al.* 1994; Julian *et al.* 2002; Helms Cahan & Keller 2003; Kulmuni *et al.* 2010; Steiner *et al.* 2011; Kulmuni & Pamilo 2014). Because of the haplodiploid genetic system, hybridization can lead to unusual genetic patterns in Hymenoptera. As haploids, males are particularly likely to show hybrid incompatibilities (Schilthuizen *et al.* 2011), and some species have evolved elaborate mechanisms to avoid producing hybrid males (e.g. Kulmuni *et al.* 2010; Kulmuni & Pamilo 2014). In other social Hymenoptera, hybridization has been co-opted to generate a genetic caste determination mechanism, with hybrid offspring becoming sterile workers and intralineage matings producing sexual females (Helms Cahan & Keller 2003; Helms Cahan & Vinson 2003; Schwander *et al.* 2007; Leniaud *et al.* 2012).

When fertile hybrid sexuals are produced, what factors preserve species differences? Some hybrid zones in *Solenopsis* ants exhibit ongoing gene flow between species (e.g. Shoemaker *et al.* 1996). In these systems, hybrids are relatively rare and F1 hybrids are more common than backcrosses, suggesting that strong selection against hybrid queens and males helps to maintain the hybrid zones. Systems like these, wherein some gene flow is maintained between divergent lineages, provide an ideal opportunity to study the process of speciation in eusocial organisms and to examine whether it differs from nonsocial organisms.

Here, we investigate a hybrid zone between two species of ants, *Formica selysi* Bondroit and *Formica cinerea* Mayr. *Formica cinerea* is broadly distributed across Europe, extending into western Siberia. *Formica selysi* is restricted to the Alps and Pyrenees, where it generally lives near streams and rivers (Seifert 2002). Variation in social organization has been extensively studied in *F. selysi*. Colonies headed by a single queen (=monogynous) and colonies containing multiple queens (=polygynous) are found in the same populations (Chapuisat *et al.* 2004; Purcell & Chapuisat 2013). Moreover, colony

organization in *F. selysi* is associated with a supergene, that is a large group of linked genes showing two well-differentiated haplotypes, Sm and Sp (Purcell *et al.* 2014). Individuals from monogynous colonies have a Sm/Sm or Sm genotype (females and males, respectively) at this supergene, whereas females from polygynous colonies have a Sp/Sp or Sp/Sm genotype and males have a Sp genotype (Purcell *et al.* 2014, 2015). Social organization is also polymorphic in *F. cinerea*, with some populations in northern Europe containing only monogynous and weakly polygynous colonies, and others containing highly polygynous colonies (Goropashnaya *et al.* 2001).

To characterize the genetic structure of the contact zone between these two species along the Rhône River in Vaud and Valais, Switzerland, we surveyed hundreds of nuclear markers throughout the genome (obtained through genotyping-by-sequencing, GBS; Parchman *et al.* 2012; Purcell *et al.* 2014; Brelsford *et al.* 2016). In addition, we genotyped a diagnostic single nucleotide polymorphism (SNP) in the mitochondrial COI gene. We had four primary goals: (i) investigate the frequency of hybrids between *F. selysi* and *F. cinerea*; (ii) determine whether hybrid queens or males ever reproduce in this system; (iii) examine whether any portion of the genome introgresses across this zone; and (iv) ask whether gene flow is limited to one social form (i.e. monogynous or polygynous colonies). Secondarily, we investigated the strength of species discrimination in the contact and in allopatric zones, by comparing cuticular hydrocarbon profiles and measuring intra- and interspecific worker aggression. Taken together, these analyses provide insights into the structure of this hybrid zone and identify some factors that may prevent these two closely related species from collapsing into a single species.

Methods

A recent population genetic analysis provided preliminary evidence of the presence of a contact zone between two divergent genetic lineages in the Rhône Valley (Purcell *et al.* 2015). An independent morphology assessment carried out by Dr. Bernhard Seifert (personal communication) revealed that these genetic lineages are in fact two closely related species, *Formica selysi* and *Formica cinerea*. Within the Alps, our surveys revealed that these species differ in their distribution: *F. selysi* is found predominantly in north and west of the main range of the Alps (the two mitochondrial lineages from the Mediterranean and North Sea drainage basins, as reported by Purcell *et al.* (2015), were confirmed to be *F. selysi*), while *F. cinerea* is found primarily south and east of the central Alps (in the Adriatic Sea
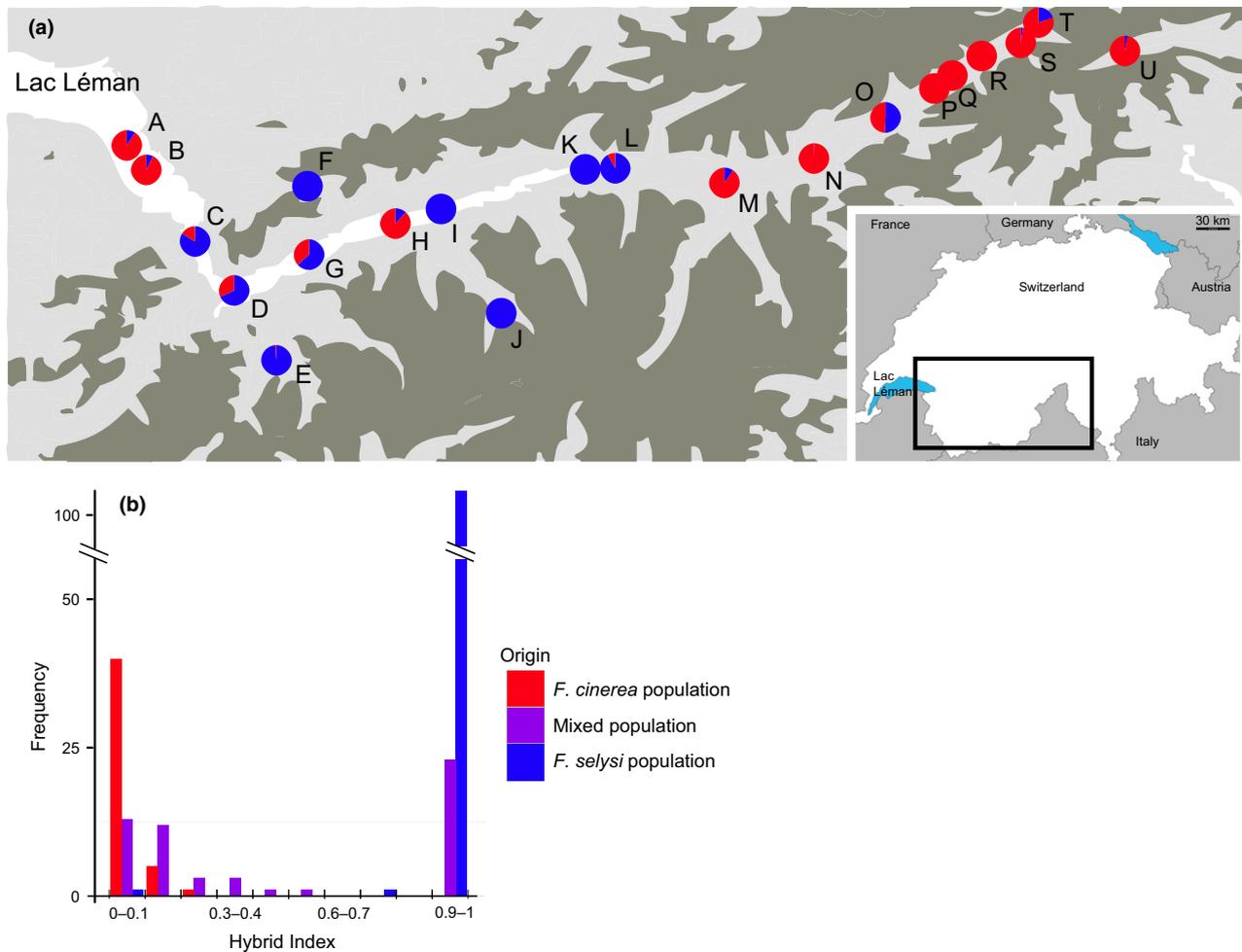
Fig. 1 (a) Map of the genetic background of each population. *Formica cinerea, Formica selysi* and their hybrids display a mosaic distribution across the Swiss Rhône Valley. The pie charts represent the proportion of *F. selysi* (blue) and *F. cinerea* (red) genetic background, estimated from the GBS-based STRUCTURE analysis of 493 SNP markers. Locality names are as follows: A: Noville, B: Aigle, C: St. Maurice, D: Martigny, E: Orsières, F: Derborence, G: Riddes, H: Conthey, I: Sion, J: Les Haudères, K: Finges, L: Leuk, M: Visp, N: Naters, O: Fiesch, P: Blitzingen, Q: Biel, R: Goms, S: Ulrichen, T: Obergestein and U: Nufenen. (b) Distribution of hybrid indexes. The distribution is strongly bimodal, with the majority of individuals sampled being predominantly *F. selysi* or *F. cinerea*. These individuals were sampled from populations containing a mostly *F. selysi* genetic background (blue bars, less than 10% admixture with *F. cinerea* at the population level), a mostly *F. cinerea* genetic background (red bars, less than 10% admixture with *F. selysi* at the population level), or a mix of the two backgrounds (purple bars). Both intermediate hybrids and backcrosses are observed in this sample; the distribution of backcrosses is asymmetric and biased towards the *F. cinerea* genetic background.

drainage basin, Purcell *et al.* 2015; this study). *Formica selysi* and *F. cinerea* co-occur in several populations along the Rhône River in Valais and Vaud, Switzerland, which prompted us to investigate whether they hybridize in this area.

Genetic characterization of the hybrid zone

We used a GBS approach to sequence loci distributed across the genome, following the protocol from Brelsford *et al.* (2016). The combination of enzymes selected (SbfI and MseI) produces a lower marker density, but

allows more individuals to be multiplexed, than our previous work in this system (Purcell *et al.* 2014). We genotyped 208 individual workers, each from a different colony, from 21 populations in the putative hybrid zone (Fig. 1a). We processed the GBS data using Bowtie 2.2.0 (Langmead & Salzberg 2012) to align the sequences to the reference genome of *F. selysi* (Purcell *et al.* 2014) and SAMtools/BCFtools (Li *et al.* 2009) to identify SNPs and call genotypes. We then used VCFtools 0.1.11 (Danecek *et al.* 2011) to filter the resulting genotypes based on quality score (minimum 20), missing data (maximum 0.20 per locus), minor allele

frequency (minimum 0.05) and to exclude indels. This resulted in 493 polymorphic loci used in subsequent analyses. 21 of these loci were located within the supergene associated with social organization (Purcell *et al.* 2014).

To assess the admixture of individuals and genetic structure of populations, we applied a Bayesian clustering approach to the full GBS data (493 loci), implemented in STRUCTURE 2.3.2 (Pritchard *et al.* 2000). The model was replicated 10 times for each K-value from 1 to 10, with a burn-in period of 50 000 iterations, followed by 100 000 iterations per test. We used the admixture model setting. We inferred the number of genetic clusters in our data following Evanno *et al.* (2005). STRUCTURE admixture coefficients (K = 2) were used to infer ancestry for each individual (hereafter 'hybrid index' or 'HI'). To assess the haplotypes at the supergene for social organization, we reran the STRUCTURE analyses using the 21 SNPs located within this supergene.

We then identified species-diagnostic nuclear markers (n = 33, on 30 unique scaffolds; defined as frequency differences >0.95 between allopatric *F. selysi* and *F. cinerea*). We carried out a genomic cline analysis using one diagnostic marker per scaffold, as implemented in the HIest package in R 3.1.1 (Fitzpatrick 2013; R Core Team 2015). Genomic cline analysis examines the relationship between locus-specific and genome-wide ancestry, seeking to identify genetic markers that are more likely or less likely than the genome-wide average to introgress between hybridizing taxa (Szymura & Barton 1986; Gompert & Buerkle 2009). HIest fits genotypic data to genome-wide hybrid index with a logit-logistic model, using maximum likelihood. The hybrid index values used in this analysis were based on STRUCTURE ancestry estimates from the full data set of 493 loci.

To further characterize the hybrid zone and increase our sample size for cuticular hydrocarbon and behavioural analyses, we genotyped some of the individuals from the GBS sample, as well as individuals from additional colonies, at a mitochondrial SNP. This SNP, located in the COI gene, was species diagnostic for *F. selysi* and *F. cinerea* and is informative for the maternal ancestry of hybrids (Purcell *et al.* 2015). We first genotyped 70 individuals belonging to the GBS sample and coming from thirteen populations located within the contact zone (Table S1, Supporting information). We then sampled workers from 82 additional colonies in two mixed populations. These populations are Branson (46°7′27.16″N, 7°5′8.38″E) and Riddes (46°10′42.89″N, 7°13′7.90″E). We genotyped four workers per colony at the mitochondrial SNP, which allowed us to identify a sample of colonies for subsequent assessment of cuticular hydrocarbon differentiation and aggression. We also assessed the social organization of these colonies by genotyping two workers per colony at two distinct genetic markers, each diagnostic for the supergene haplotype in both species (Sm vs. Sp; Purcell *et al.* 2014; diagnostic in *F. cinerea*, Box S1).

Species discrimination: cuticular hydrocarbons and aggression

To evaluate the strength of species discrimination in sympatry and allopatry, we collected workers from two populations where both species occur (Branson and Riddes) and two populations away from the contact zone: Aigle for *F. cinerea* (46°19′8.72″N, 6°55′59.59″E, but note that some individuals in this population were admixed with up to 15% F. selysi ancestry) and Finges for *F. selysi* (46°18′48.10″N, 7°37′11.35″E). We selected only polygynous colonies, in order to exclude any confounding effect due to phenotypic differences between the monogynous and polygynous social forms (Schwander *et al.* 2005; Rosset & Chapuisat 2007; Rosset *et al.* 2007). We measured two phenotypic traits linked to species discrimination: cuticular hydrocarbon profile and aggression between non-nestmate workers.

Cuticular hydrocarbons

We assessed the cuticular hydrocarbon profile of groups of 10 workers per colony from four *F. cinerea* colonies (and/or *F. cinerea*-like hybrids) and six *F. selysi* colonies. The worker groups were frozen at −20 °C the day before the GC-MS analysis. We extracted cuticular components by immersing the 10 workers in 150 µL of pentane solution (93.75 parts pentane, 6.25 parts internal standard) for 15 min and transferring 100 µL of this pentane solution to a new vial. We injected 4 µL of the solution into a Thermo Scientific™ gas chromatograph with an internal standard consisting of alkane standard solution C8–C20 (40 mg/L) and *n*-Pentane (Sigma). We operated the gas chromatograph in splitless injection mode and used helium as the carrier gas. We set the oven temperature to 70 °C and ramped the temperature at 10 °C/min to a final temperature of 320 °C, which was held for 15 min. The position and corrected height of each CHC peak was assessed in Xcalibur 1.4 SR1™ (Thermo Scientific).

Aggression between workers

Ant aggressiveness was measured by monitoring 5-min pairwise interactions between non-nestmate workers. We tested the following combinations of workers: same species, same population; same species, different

populations; different species, same population; different species, different populations. At least 24 h before the test, workers were marked with a dot of paint on the thorax, with the colour randomized with respect to species and population. Pairs of ants were transferred to a Petri dish side-lined with Fluon. We filmed the interactions for five minutes using a webcam (Logitech C910). For each interaction, aggressiveness was defined on the following scale: 0 = no contact, 1 = antennation, 2 = mandible open when workers are adjacent, 3 = biting, 4 = biting plus acid ejection posture. For each aggression test ($n = 94$), we retained the highest aggression score (Rosset *et al.* 2007). In interspecific tests, we attempted to identify the individual initiating aggressive interactions of score 3 or 4 ($n = 29$ interactions). Scoring was carried out by two observers who were blind with respect to the origins of each combination, with a subset of cases scored by both to ensure consistency.

Statistical analysis

We compared the cuticular hydrocarbon profiles of *F. selysi* and *F. cinerea* colonies using a principal component analysis (PCA), implemented in R 3.1.1. To assess the statistical significance of differences in aggression within and between species, and within and between populations, we performed permutation tests using the *boot* and *coin* packages in R. We assessed whether one species was more likely to initiate aggression than the other using a binomial test.

Results

Genetic characterization of the hybrid zone

Our genomic analyses confirmed the presence of a hybrid zone between *Formica selysi* and *Formica cinerea* in the Swiss Rhône Valley and revealed that the zone has a mosaic structure. Ten sites contained predominantly *F. cinerea* colonies, six sites contained *F. selysi* colonies, and five sites contained colonies of each species as well as hybrids (Fig. 1a). STRUCTURE admixture coefficients of workers in mixed populations, based on 493 polymorphic nuclear loci in one worker per colony, were bimodally distributed, indicating that this sample was composed of mainly pure *F. cinerea* ($n = 43$) and *F. selysi* ($n = 124$), with 20% of individuals ($n = 41$) having a hybrid index (HI) between 0.05 and 0.95 (Fig. 1b). Few of these hybrids had HI values close to 0.5, as would be expected for first-generation hybrids. The presence of backcrosses in this distribution clearly demonstrates that hybrid queens and/or males can produce viable offspring in this system.

Overall, we identified a signature of asymmetrical hybridization, with most sampled hybrids exhibiting a HI between 0.05 and 0.5 (*F. cinerea*-like hybrids; Fig. 1b). The paucity of individuals with HI between 0.5 and 0.95 (*F. selysi*-like hybrids) suggests a pattern of genome-wide introgression from *F. selysi* into *F. cinerea*. The genomic cline analysis of 30 nuclear loci identified two markers as exceptions to this genome-wide pattern (Fig. 2). In these two instances, *F. selysi* alleles were underrepresented in *F. cinerea*-like hybrids relative to other loci. Because we only found three individuals with HI between 0.5 and 0.95, we cannot draw conclusions about the behaviour of any loci in *F. selysi*-like hybrids. To further characterize this pattern of asymmetric hybridization, we genotyped mtDNA in 40 individuals with HI between 0.05 and 0.95, of which 39 (97%) carried the *F. cinerea* haplotype and one (3%) carried the *F. selysi* haplotype. One hybrid individual failed to amplify. An additional 30 individuals had HI less than 0.05 (=*F. cinerea*) or greater than 0.95 (=*F. selysi*), and the mtDNA haplotype was perfectly associated with species in these samples (Table S1, Supporting information). When added to a genomic cline analysis, mtDNA was a highly significant outlier, with the *F. selysi* haplotype being underrepresented in hybrids (Fig. S1, Supporting information).
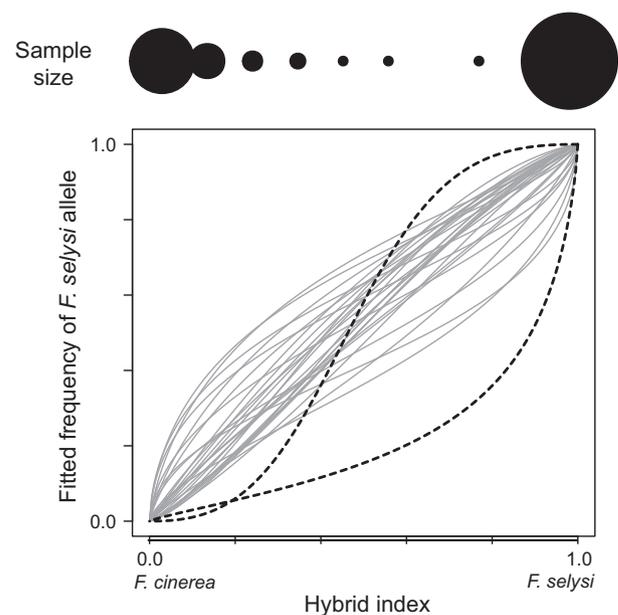

Fig. 2 A genomic cline analysis of 30 species-diagnostic SNPs reveals two significant outliers (black dotted lines). Both outlier clines indicate that *F. cinerea*-like hybrids are more likely to harbour *F. cinerea* alleles at these loci. Cline fits may be unreliable for hybrid index values between 0.5 and 0.9, due to low representation of individuals within this HI range (shown in the sample size chart above the panel).

The social organization of colonies was inferred from the haplotypes at the social supergene, as determined by a STRUCTURE analysis using the 21 loci located in the supergene (Table S2, Supporting information). This analysis revealed that colonies containing hybrid individuals (HI of 0.05–0.95) can be either monogynous or polygynous. Indeed, among 40 hybrids genotyped, 15 individuals had the monogyne-associated Sm/Sm genotype, while 17 and 8 individuals had polygyne-associated genotypes, Sm/Sp and Sp/Sp, respectively.

Species discrimination: cuticular hydrocarbons and aggression

We found significant differences between the two species in cuticular hydrocarbon profile, as well as an increased level of aggression in interactions between individuals from different species (Fig. 3). *F. cinerea* and *F. selysi* had very different hydrocarbon profiles. The first principal component axis, explaining 73% of the variation, showed no overlap between the two species. There was some variation among colonies within species in the second principal component, which explained an additional 16% of the variance. The PCA revealed no clear differences between workers from different populations within species (not shown). Workers showed strong species discrimination. Biting and acid projection postures occurred frequently between species, but were almost absent within species, and overall there was much more aggression in interspecies interactions than in intraspecies interactions (Fig. 3c vs. b and d; permutation test, $P < 0.0001$). We found no difference in the proportion of aggressive interactions initiated by *F. selysi* and *F. cinerea* (binomial test, $P = 0.64$). Moreover, we found no significant difference in aggression between workers from sympatric or allopatric zones, and therefore, these were pooled for this analysis (Table S3, Supporting information).

Discussion

We have identified a geographically broad mosaic hybrid zone between *Formica selysi* and *Formica cinerea* along the Rhône River upstream of Lac Léman in Switzerland. Early-generation hybrids (hybrid index between 0.2 and 0.8) are infrequent compared to parental species, and nearly all hybrids are genetically closer to *F. cinerea* than to *F. selysi*. This suggests that there are mechanisms reducing interspecific mating or the success of colonies containing hybrid individuals and that these reproductive barriers are asymmetric. The presence of different cuticular hydrocarbon signatures in the two species and their increased interspecific aggression tendencies may be factors maintaining distinct species within this contact zone.
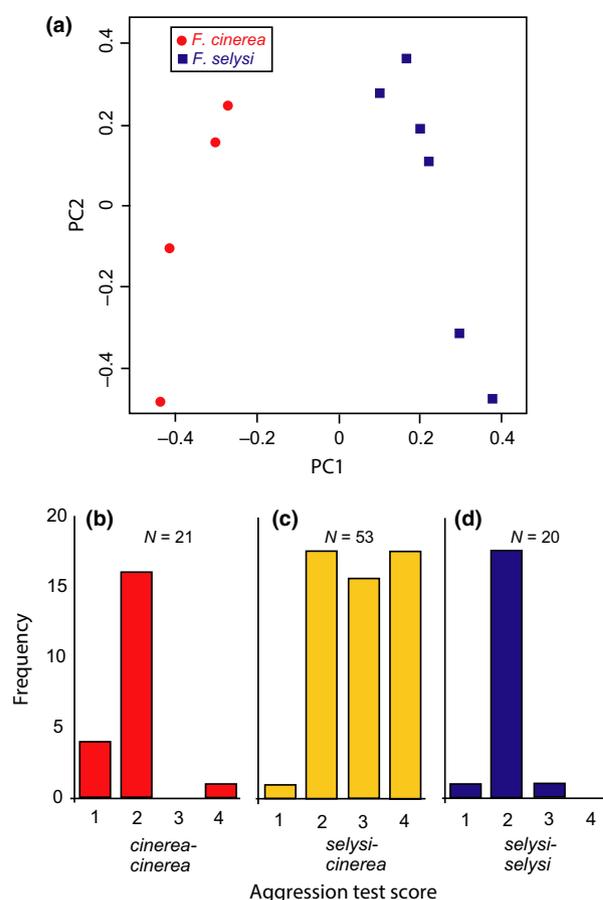

Figure 3 (a) PCA of the cuticular hydrocarbon profiles. Each point represents a colony, while symbols represent different species. The profiles of the two species differ significantly along the first principal component axis, which explains 73% of the total variance. (b, c and d) Aggression scores in pairwise interactions between non-nestmates. The level of aggression was much higher in interspecific trials (c) than in intraspecific trials (b and d).

The asymmetry in the distribution of hybrid index (HI), which is biased towards *F. cinerea*, and the fact that 97% of hybrids have the *F. cinerea* mitochondrial haplotype reveal an intriguing pattern of asymmetrical gene flow with a genome-wide nuclear introgression from *F. selysi* into *F. cinerea* (e.g. White *et al.* 2015). Taken together, the nuclear and mitochondrial data suggest that *F. selysi* queens mated with *F. cinerea* males rarely produce fertile offspring. The majority of hybrid individuals with HI between 0.1 and 0.4 would then emerge through hybrid queens (from a *F. cinerea* queen mated with a *F. selysi* male) mating with *F. cinerea* males or through *F. cinerea* queens mating with males of hybrid origin. Either way, hybrid sexuals must be produced and (at least occasionally) have some reproductive success to explain the distribution of HIs. Overall, this system should provide an ideal opportunity to

investigate the relative effects of pre- and post-zygotic isolating mechanisms in two ant species that differ in the direction and degree of their mutual reproductive barrier.

Two nuclear markers do not follow the general pattern of unidirectional gene flow from *F. selysi* to *F. cinerea*. These outliers in the genomic cline analysis show an underrepresentation of *F. selysi* alleles in hybrids, indicating that these loci might be associated with reproductive barriers. We searched the scaffolds containing the two outliers against the *Camponotus floridanus* and *Apis mellifera* genomes. The first outlier is in a region present in multiple copies in *C. floridanus*, with no known genes. The second outlier is located within the gene TRPA5, which belongs to a gene family involved in sensory integration (Matsuura *et al.* 2009). Further investigation is needed to determine why the *F. selysi* alleles at these markers are underrepresented in hybrids compared to the rest of the genome, and whether these loci contribute to species boundaries in this system.

Interestingly, we observe hybrids in both monogynous and polygynous colonies. This rules out a complete barrier to hybridization in either social form. Our genomic analysis reveals that both species share the supergene haplotypes that are associated with social organization in *F. selysi* (Purcell *et al.* 2014). In our surveys of *F. selysi*, polygynous colonies never produce Sm males or Sm/Sm queens and workers, and monogynous colonies never contain Sp genotypes. Thus, if queens of monogynous origin did not hybridize, we would not encounter hybrid Sm/Sm workers. Similarly, if sexuals of polygynous origin did not hybridize, we would never encounter hybrid Sm/Sp or Sp/Sp workers. We do not yet know, however, whether haplotypes of the social supergene are shared through introgression within the hybrid zone, or whether the divergence of the Sm and Sp haplotypes of the supergene predates the divergence between the two species.

We identified significant differences in the cuticular hydrocarbon profiles of the two species, and higher levels of aggression between workers in interspecific interactions, compared to intraspecific interactions. The increased aggression towards heterospecifics may play a role in speciation, as high levels of aggression between species may result in interspecific exclusion, biased mating and possibly reduced fitness of hybrid colonies.

How does this system fit into the broader picture of ant hybrid zones? In several well-studied species, hybridization has been co-opted to generate genetic caste determination, whereby hybrid offspring are workers, while nonhybrid diploid offspring become new queens (Helms Cahan & Keller 2003; Schwander

et al. 2007; Leniaud *et al.* 2012). This system results in a different pattern from the one that we observed—hybrid workers would have HI of 0.5, while queens and males would not be hybrids. Pearson (1983) observed an asymmetrical hybrid zone between *Lasius alienus* and *L. niger* in the UK, wherein *L. alienus* queens mated with *L. niger* males produced hybrid workers and males. Thus, only F1 hybrids (workers) were detected, because haploid males were produced parthenogenetically. In the hybrid zone that we investigate here, the signature of backcrossing in the distribution of hybrid indexes, with the majority of hybrids having HIs between 0.1 and 0.4, shows that hybrid sexuals are produced and can themselves produce viable offspring in the *Formica selysi/cinerea* hybrid zone. The hybrid zone described in this study bears a stronger resemblance to the mosaic hybrid zone between *Solenopsis invicta* and *S. richteri* in North America (Shoemaker *et al.* 1996). Shoemaker *et al.* (1996) suggested that the hybrid zone in *Solenopsis* was likely to be maintained by reduced fitness of hybrid sexuals and/or reduced competitive ability of hybrid colonies. Similar mechanisms may be in place in the *F. selysi/cinerea* hybrid zone if, for example, fertility is reduced or nestmate recognition is disrupted in hybrids.

Beyond ants, mosaic hybrid zones and asymmetric hybridization have been identified in many other systems. Mosaic hybrid zones can arise from species-specific adaptation to patchily distributed habitat types (Harrison 1986; Rand & Harrison 1989), or from stochastic patterns of colonization (Harrison 1986), especially in combination with assortative mating (M'Gonigle & FitzJohn 2010). Asymmetric introgression in hybridizing species is common (e.g. Johnson *et al.* 2015; Kenney & Sweigart 2016; Sardell & Uy 2016) and may result from many causes including hybrid zone movement (Barton & Hewitt 1985; Excoffier *et al.* 2009), patterns of mate choice (Svensson *et al.* 2007), species differences in dispersal (Currat *et al.* 2008) or the genetic architecture of reproductive incompatibilities (Tiffin *et al.* 2001). Future work on patterns of mate choice, ecological differences and genetic differentiation between *F. selysi* and *F. cinerea* will help to elucidate the causes of the mosaic structure and asymmetric introgression that we have identified.

The hybrid zone between *F. selysi* and *F. cinerea* that we describe here provides a promising opportunity to investigate how speciation functions in eusocial insects. Because hybrids are present, we should be able to tease apart the relative roles of mate choice, reduced fitness or viability in hybrid queens or males, and colony-level incompatibilities. Overall, investigating this and similar systems will provide valuable insights into how speciation works in social organisms.

Acknowledgements

References

Abbott R, Albach D, Ansell S *et al.* (2013) Hybridization and speciation. *Journal of Evolutionary Biology*, **26**, 229–246.

Barton NH, Hewitt G (1985) Analysis of hybrid zones. *Annual Review of Ecology and Systematics*, **16**, 113–148.

Brelsford A, Dufresnes C, Perrin N (2016) High-density sex-specific linkage maps of a European tree frog (*Hyla arborea*) identify the sex chromosome without information on off-spring sex. *Heredity*, **116**, 177–181.

Chapuisat M, Bocherens S, Rosset H (2004) Variable queen number in ant colonies: no impact on queen turnover, inbreeding, and population genetic differentiation in the ant *Formica selysi*. *Evolution*, **58**, 1064–1072.

Chávez-Galarza J, Henriques D, Johnston JS *et al.* (2015) Revisiting the Iberian honey bee (*Apis mellifera iberiensis*) contact zone: maternal and genome-wide nuclear variations provide support for secondary contact from historical refugia. *Molecular Ecology*, **24**, 2973–2992.

Currat M, Ruedi M, Petit RJ, Excoffier L (2008) The hidden side of invasions: massive introgression by local genes. *Evolution*, **62**, 1908–1920.

Danecek P, Auton A, Abecasis G *et al.* (2011) The variant call format and VCFtools. *Bioinformatics*, **27**, 2156–2158.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, **14**, 2611–2620.

Excoffier L, Foll M, Petit RJ (2009) Genetic consequences of range expansions. *Annual Review of Ecology, Evolution, and Systematics*, **40**, 481–501.

Feldhaar H, Foitzik S, Heinze J (2008) Lifelong commitment to the wrong partner: hybridization in ants. *Philosophical Transactions of the Royal Society B*, **363**, 2891–2899.

Fitzpatrick BM (2013) Alternative forms for genomic clines. *Ecology and Evolution*, **3**, 1951–1966.

Gompert Z, Buerkle CA (2009) A powerful regression-based method for admixture mapping of isolation across the genome of hybrids. *Molecular Ecology*, **18**, 1207–1224.

Goropashnaya AV, Seppä P, Pamilo P (2001) Social and genetic characteristics of geographically isolated populations in the ant *Formica cinerea*. *Molecular Ecology*, **10**, 2807–2818.

Harrison RG (1986) Pattern and process in a narrow hybrid zone. *Heredity*, **56**, 337–349.

Helms Cahan S, Keller L (2003) Complex hybrid origin of genetic caste determination in harvester ants. *Nature*, **424**, 306–309.

Helms Cahan S, Vinson SB (2003) Reproductive division of labor between hybrid and non-hybrid offspring in a fire ant hybrid zone. *Evolution*, **57**, 1562–1570.

Johnson BB, White TA, Phillips CA, Zamudio KR (2015) Asymmetric introgression in a spotted salamander hybrid zone. *Journal of Heredity*, **5**, 608–617.

Julian GE, Fewell JH, Gadau J, Johnson RA, Larrabee D (2002) Genetic determination of the queen caste in an ant hybrid zone. *Proceedings of the National Academy of Sciences of the USA*, **99**, 8157–8160.

Kenney AM, Sweigart AL (2016) Reproductive isolation and introgression between sympatric *Mimulus* species. *Molecular Ecology*, **11**, 2499–2517.

Kulmuni J, Pamilo P (2014) Introgression in hybrid ants is favored in females but selected against in males. *Proceedings of the National Academy of Sciences of the USA*, **111**, 12805–12810.

Kulmuni J, Seifert B, Pamilo P (2010) Segregation distortion causes large-scale differences between male and female genomes in hybrid ants. *Proceedings of the National Academy of Sciences of the USA*, **107**, 7371–7376.

Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. *Nature Methods*, **9**, 357–359.

Lefebvre T, Châline N, Limousin D, Dupont S, Bagnères A-G (2008) From speciation to introgressive hybridization: the phylogeographic structure of an island subspecies of termite, *Reticulitermes lucifugus corsicus*. *BMC Evolutionary Biology*, **8**, 38.

Leniaud L, Darras H, Boulay R, Aron S (2012) Social hybridogenesis in the clonal ant *Cataglyphis hispanica*. *Current Biology*, **22**, 1188–1193.

Li H, Handsaker B, Wysoker A *et al.* 1000 Genome Project Data Processing Subgroup (2009) The sequence alignment/map format and SAMtools. *Bioinformatics*, **25**, 2078–2079.

Matsuura H, Sokabe T, Kohno K, Tominaga M, Kadowaki T (2009) Evolutionary conservation and changes in insect TRP channels. *BMC Evolutionary Biology*, **9**, 228.

M'Gonigle LK, FitzJohn RG (2010) Assortative mating and spatial structure in hybrid zones. *Evolution*, **64**, 444–455.

Parchman TL, Gompert Z, Mudge J, Schilkey FD, Benkman CW, Buerkle A (2012) Genome-wide association genetics of an adaptive trait in lodgepole pine. *Molecular Ecology*, **21**, 2991–3005.

Pearson B (1983) Hybridisation between the ant species *Lasius niger* and *Lasius alienus*: the genetic evidence. *Insectes sociaux*, **30**, 402–411.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.

Purcell J, Chapuisat M (2013) Bidirectional shifts in colony queen number in a socially polymorphic ant population. *Evolution*, **67**, 1169–1180.

Purcell J, Brelsford A, Wurm Y, Perrin N, Chapuisat M (2014) Convergent genetic architecture underlies social organization in ants. *Current Biology*, **24**, 2728–2732.

Purcell J, Pellissier L, Chapuisat M (2015) Social structure varies with elevation in an alpine ant. *Molecular Ecology*, **24**, 498–507.

R Core Team (2015) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. Available at: http://www.R-project.org/.

Rand DM, Harrison RG (1989) Ecological genetics of a mosaic hybrid zone: mitochondrial, nuclear, and reproductive differentiation of crickets by soil type. *Evolution*, **43**, 432–449.

Rinderer TE, Stelzer JA, Oldroyd BP, Buco SM, Rubink WL (1991) Hybridization between European and Africanized honey bees in the neotropical Yucatan Peninsula. *Science*, **253**, 309–311.

Ross KG, Vander Meer RK, Fletcher DJC, Vargo EL (1987) Biochemical phenotypic and genetic studies of two introduced fire ants and their hybrid (Hymenoptera: Formicidae). *Evolution*, **41**, 280–293.

Rosset H, Chapuisat M (2007) Alternative life-histories in a socially polymorphic ant. *Evolutionary Ecology*, **21**, 577–588.

Rosset H, Schwander T, Chapuisat M (2007) Nestmate recognition and levels of aggression are not altered by changes in genetic diversity in a socially polymorphic ant. *Animal Behaviour*, **74**, 951–956.

Sardell JM, Uy JAC (2016) Hybridization following recent secondary contact results in asymmetric genotypic and phenotypic introgression between island species of Myzomela honeyeaters. *Evolution*, **70**, 257–269.

Schilthuizen M, Giesbers MCWG, Beukeboom LW (2011) Haldane's rule in the 21st century. *Heredity*, **107**, 95–102.

Schwander T, Rosset H, Chapuisat M (2005) Division of labour and worker size polymorphism in ant colonies: the impact of social and genetic factors. *Behavioral Ecology and Sociobiology*, **59**, 215–221.

Schwander T, Keller L, Helms Cahan S (2007) Two alternate mechanisms contribute to the persistence of interdependent lineages in *Pogonomyrmex* harvester ants. *Molecular Ecology*, **16**, 3533–3543.

Seifert B (2002) A taxonomic revision of the *Formica cinerea* group (Hymenoptera: Formicidae). *Abhandlungen und Berichte des Naturkundemuseseums Görlitz*, **74**, 245–272.

Seifert B (2010) Intranidal mating, gyne polymorphism, polygyny, and supercoloniality as factors for sympatric and parapatric speciation in ants. *Ecological Entomology*, **35**, 33–40.

Shoemaker DD, Ross KG, Arnold ML (1994) Development of RAPM markers in two introduced fire ants, *Solenopsis invicta* and *S. richteri*, and their application to the study of a hybrid zone. *Molecular Ecology*, **3**, 531–539.

Shoemaker DD, Ross KG, Arnold ML (1996) Genetic structure and evolution of a fire ant hybrid zone. *Evolution*, **50**, 1958–1976.

Steiner FM, Seifert B, Grasso DA *et al.* (2011) Mixed colonies and hybridization of *Messor* harvester ant species (Hymenoptera: Formicidae) organisms. *Diversity and Evolution*, **11**, 107–134.

Svensson EI, Karlsson K, Friberg M, Eroukhmanoff F (2007) Gender differences in species recognition and the evolution of asymmetric sexual isolation. *Current Biology*, **17**, 1943–1947.

Szymura JM, Barton NH (1986) Genetic analysis of a hybrid zone between the fire-bellied toads, *Bombina bombina* and *B. variegate*, near Cracow in southern Poland. *Evolution*, **40**, 1141–1159.

Tiffin P, Olson S, Moyle LC (2001) Asymmetrical crossing barriers in angiosperms. *Proceedings of the Royal Society of London B: Biological Sciences*, **268**, 861–867.

White GM, Michaelides S, Heathcote RJP *et al.* (2015) Sexual selection drives asymmetric introgression in wall lizards. *Ecology Letters*, **18**, 1366–1375.

Data accessibility

GBS data are archived on the NCBI Sequence Read Archive (SRA) in BioProject PRJNA336636.

This BioProject contains Illumina reads for each individual.

-Genotypes and aggression data are presented in the Supplementary Materials file.
-CHC and additional aggression data are archived in DOI: 10.5061/dryad.sj169

The CHC data include retention time plots and raw data from each GCMS run. The aggression data include a spreadsheet listing the aggression responses in each of our pairwise aggression tests.

Supporting information

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

Table S1. Distribution of COI genotypes by hybrid index (Fc for *F. cinerea* and Fs for *F. selysi*).

Table S2. Results of STRUCTURE analysis based on 21 SNPs from the social chromosome for each individual in each locality.

Table S3. Aggression trial scores for intra- and interspecific pairwise aggression trials with workers from the same population (=sympatric) or different populations (=allopatric).

Fig. S1. Genomic cline analysis of 30 species diagnostic nuclear SNPs and one mtDNA SNP.

Box. S1. PCR-RFLP Protocol for genotyping markers on the social chromosome.