Social antagonism facilitates supergene expansion in ants

Giulia Scarparo[1,3]*, Marie Palanchon[2], Alan Brelsford[2], and Jessica Purcell[1]*

[1]Department of Entomology, University of California Riverside, 165 Entomology Bldg. Citrus Drive, Riverside, CA 92521, USA

[2]Department of Evolution, Ecology, and Organismal Biology, University of California Riverside, 2710 Life Science Bldg., Riverside, CA 92521 , USA

[3] Lead contact

*Correspondence to Giulia Scarparo and Jessica Purcell.

Emails: giulias@ucr.edu; jpurcell@ucr.edu

Keywords: sex chromosome evolution; queen size dimorphism; queen number; genetic architecture; social parasites**;** colony social form; sex ratio; social insects.

Summary

Antagonistic selection has long been considered a major driver of the formation and expansion of sex chromosomes. For example, sexually antagonistic variation on an autosome can select for suppressed recombination between that autosome and the sex chromosome, leading to a neo-sex chromosome. Autosomal supergenes, chromosomal regions containing tightly linked variants affecting the same complex trait, share similarities with sex chromosomes, raising the possibility that sex chromosome evolution models can explain the evolution of genome structure and recombination in other contexts. We tested this premise in a *Formica* ant species wherein we identified four supergene haplotypes on chromosome 3 underlying colony social organization and sex ratio. We discovered a novel rearranged supergene variant (9r) on chromosome 9 underlying queen miniaturization. The 9r is in strong linkage disequilibrium with one chromosome 3 haplotype (P_2) found in multi-queen (polygyne) colonies. We suggest that queen miniaturization is strongly disfavored in the single queen (monogyne) background, and thus socially antagonistic. As such, divergent selection experienced by ants living in alternative social 'environments' (monogyne and polygyne) may have contributed to the emergence of a genetic polymorphism on chromosome 9 and associated queen-size dimorphism. Consequently, an ancestral polygyne-associated haplotype may have expanded to include the polymorphism on chromosome 9, resulting in a larger region of suppressed recombination spanning two chromosomes. This process is analogous to the formation of neo-sex chromosomes and consistent with models of expanding regions of suppressed recombination. We propose that miniaturized queens, 16-20% smaller than queens without 9r, could be incipient intraspecific social parasites.

Introduction

When only certain combinations of alleles at different genes yield positive fitness outcomes, theory predicts the formation or expansion of regions of suppressed recombination to ensure that beneficial combinations are co-transmitted. This idea is well established in the sex chromosome evolution theory literature[1,2], where sexually antagonistic selection is an integral part of the 'canonical' sex chromosome evolution scenario[3]. Specifically, the model proposes that recombination is suppressed in incipient sex chromosomes when sex determining genes are linked with alleles that are advantageous in one sex, but detrimental in the other. Empirical evidence supporting sexually antagonistic selection can be found in neo-sex chromosomes, where sex chromosomes recently expanded to include a previously

autosomal region[4–6]. Extending the search for a role of antagonistic selection outside sex chromosomes will help to reveal the broader prevalence of this mechanism in other regions of suppressed recombination.

The importance of reduced or suppressed recombination has been widely recognized in contexts beyond sex chromosome evolution, such as in the emergence of local adaptation (e.g. Charlesworth and Charlesworth[7]; Kirkpatrick and Barton[8]; Yeaman[9]). Researchers propose that selection should favor suppressed recombination between combinations of alleles that work well together in specific environments (achieved through inversion[8] or other chromosomal rearrangement[9]). In parallel to the idea of sexually antagonistic selection, mismatched alleles would be subject to strong negative selection in either environment. Given the similarities between models of suppressed recombination around sex-determining loci and locally adapted loci, both bodies of research are relevant to understanding supergene evolution.

Some studies of autosomal supergenes have drawn inspiration from models of sex chromosome evolution and inversion formation (e.g. Branco et al.[10]; Branco et al.[11]; Brelsford et al.[12]; Duhamel et al.[13]). Phenotypic traits controlled by supergenes include alternative mating systems[11], migratory behavior[14], mimetic coloration[15], and social organization[16,17]. Here, we explore the possibility that supergenes form and expand through a process that parallels the canonical model of sex chromosome evolution. Working with an autosomal supergene that controls colony queen number in *Formica* ants[12,17–19], we propose that alternative social contexts shaped by a supergene could favor the expansion of regions of suppressed recombination to include "socially antagonistic" loci. We define socially antagonistic loci as alleles that have beneficial fitness outcomes in one social environment, but detrimental outcomes in the other social environment (see also Chapuisat[19]; Martinez-Ruiz et al.[20]). Because the *Formica* supergene is approximately 23 MY old, determining what genetic polymorphisms were present during supergene formation would be difficult in this system (see also Coughlan and Willis[21]). Instead, we focus on elaborations to the existing supergene system.

The *Formica* social supergene was initially described in the Alpine silver ant *Formica selysi*[17]. Alternative haplotypes of the supergene are associated with colony queen number, thus determining whether a colony is monogyne (with only one queen) or polygyne (with two

or more queens). In *Formica* ants (and other socially polymorphic ants), a suite of other traits is frequently associated with variation in colony queen number, including body size of queens and workers[22–24], colony size[23], dispersal probability[22,25], and investment in sexual offspring[26,27]. Genes underlying extreme versions of these traits could be candidates in the search for socially antagonistic loci. In *F. selysi*, monogyne colonies exclusively harbor individuals carrying the monogyne-associated haplotype, M, whereas polygyne colonies always contain individuals bearing at least one copy of the alternative polygyne-associated haplotype, P[17,28]. The P haplotype acts as a maternal-effect killer, causing the early death of any offspring of heterozygous mothers that do not bear the P haplotype[29]. Recently, Tafreshi et al.[30] proposed that this polymorphism is only stable in the presence of both assortative mating and large fitness differences between supergene genotypes, both of which have recent empirical support[28,31].

Brelsford et al.[12] showed that our focal species, *Formica cinerea,* has (at least) three supergene haplotypes. In the present study, we set out to characterize differences between the supergene haplotypes in *F. cinerea.* Since our limited preliminary evidence suggested that two of the three haplotypes are found in polygyne colonies, we hypothesized that the two P haplotypes would contain different gene sets and control distinct phenotypic traits. Certain phenotypes might be beneficial in the polygyne context, but detrimental in the monogyne context. If so, one or both of the P haplotypes could harbor socially antagonistic loci. We did not predict *a priori* the mechanism through which the gene sets would differ, although we were open to the possibility of an expansion analogous to neo-sex chromosome formation[4–6] or of translocation of genes into the supergene region[9,32]. To test these predictions and detect signatures of socially antagonistic selection, we collected workers, gynes, males, and queens from 172 *F. cinerea* colonies from northern Italy and characterized the genetic architecture and phenotype associated with alternative supergene haplotypes. We examined whether the two putative P haplotypes control alternative traits, in addition to colony queen number.

Results

Characterization of supergene haplotypes

As expected based on a preliminary assessment of genetic variation on the *F. cinerea* supergene[12], we detected more than two haplotypes. Overall, a principal component analysis (PCA) of chromosome 3 revealed that *F. cinerea* populations in northern Italy have

four supergene haplotypes. Within these plots, we identified ten clusters, which is consistent with the expected number of genotype combinations for a system with four distinct haplotypes (four homozygous genotypes and six heterozygous combinations). PC axes 1 and 2 separate all individuals into six clear clusters (Figure 1A) and PC axis 3 further reveals variation in the putative M haplotypes (Figure 1B). We examined F_{IS} in each of the clusters to determine whether individuals are homozygous or heterozygous for the supergene; heterozygous individuals exhibit negative F_{IS} across the supergene region (plotted as split circles in the figure), while homozygous individuals exhibit positive F_{IS} (full circles). Haploid males (half circles in the figure) cluster with homozygous females. We then examined the frequency of reference and alternative alleles relative to the *F. selysi* reference genome, which was constructed from a pool of M males[12], to make a preliminary determination about whether haplotypes were M-like or P-like. Based on this analysis, two of the four haplotypes are M (M_A and M_D) and two are P (P_1 and P_2). The M_D is mostly found in a heterozygous state, with the exception of one newly-mated queen (M_DM_D) and seven males (M_D) out of 239 total individuals carrying M_D.

To investigate genetic differences between alternative haplotypes, we looked at the F_{ST} between haplotype pairs, revealing that the M_D compared to the M_A spans only the first half of chromosome 3, from ~2.0-7.5 Mbp (Figure 1C), while the genetic differentiation between M haplotypes and P_1 spans the same supergene region discovered in *F. selysi*[17], from ~2.0-12.5 Mbp (Figures 1D-E). The F_{ST} plots show high differentiation that spans almost all of chromosomes 3 and 9 when comparing P_2 with the other haplotypes (Figures 1F–H).

Assessment of a newly discovered supergene region on chromosome 9

Given the second region of high differentiation between the P_2 haplotype and all other supergene haplotypes, we investigated variation on chromosome 9. Here, we detected two alternative haplotypes. The PCA (Figure 2A) displays three distinct clusters of individuals along PC 1. Individuals in the left and right clusters appear to be homozygous based on positive F_{IS} values, while individuals in the central cluster are heterozygous. We named the two alternative haplotypes as follows: "9a" referring to the ancestral chromosome structure, and "9r" referring to the rearranged chromosome structure relative to the *F. selysi* genome, as revealed by analysis of linkage disequilibrium (LD) within each homozygous genotype (Figures S1A-B). A comparison of previously published linkage maps from *F. selysi* and the distantly related *F. exsecta* confirms that the 9a chromosome structure and the lack of LD

between chromosomes 9 and 3 are ancestral[18]. This region of suppressed recombination on chromosome 9 spans from ~2.4- 9.4 Mbp and contains hundreds of genes.

The P_2 haplotype on chromosome 3 and the 9r on chromosome 9 are almost always transmitted together (χ^2 = 1273.8, df = 1, p-value < 0.0001). No P_2P_2 individual has been found to be 9a9a homozygous (Figure 2B). In contrast, individuals without the P_2 almost always bear only the 9a haplotype. However, we noticed some mismatches in this pattern showing an imperfect association between the P_2 and 9r (Figure 2B). The mismatches occur disproportionately in workers (19 out of 20 observations, $\chi^{2\,=}$ 14.3, df = 1, p-value < 0.0001).

Colony social form is associated with chromosome 3 haplotypes

In other *Formica* species, colony social form is controlled by the social supergene on chromosome 3[12]. To verify that queen number is associated with variation on chromosome 3 in *F. cinerea*, we assessed colony social form and supergene genotype distribution within colonies. Of the 120 analyzed colonies, half are monogyne (39 monogyne monandrous, 21 monogyne polyandrous), and half polygyne (Figure 3A). We found a significant association of the M haplotypes with the monogyne form, where 55 out of 60 colonies contain exclusively M_AM_A and/or M_DM_A individuals (Figures 3B-C; Figure S2A). Similarly, we observed a strong association between P haplotypes and the polygyne form, with 56 out of 60 polygyne colonies having members with at least one P haplotype (Figure 3D; Figure S2B). Despite several exceptions, the association of M haplotypes with the monogyne form and the P haplotypes with the polygyne form is still significant (z_{115}= 2.4, p-value < 0.05, GLMM).

The M_D influences colony sex ratio

Some species of social insect show a pattern of split sex ratio at the population level, in which some colonies specialize in the production of future queens and others in the production of males[33–36]. This also occurs in *F. cinerea*, especially in monogyne colonies (Figure S3). In contrast, polygyne colonies more often produce a mix of males and gynes or exclusively males (Figure S3). Our data show that the M_D haplotype is associated with the production of gynes (z_{32}= 2.3, p < 0.05, GLMM). The M_D haplotype is rarely present in males although we found some exceptions (7 M_D males).

The effect of the two supergenes on body size: chromosome 9 harbors a miniaturizing haplotype

Based on field observations that *F. cinerea* alates vary substantially in size, we measured the head width of gynes, queens and males. Our results revealed that alates with at least one copy of the P_2 haplotype have significantly smaller heads than alates without the P_2 (gynes and queens: all Tukey post-hoc comparisons $p< 0.0001$; males: all Tukey post-hoc comparisons $p< 0.0001$; linear mixed model) (Figures 4A, C). However, this size reduction is caused by the 9r haplotype on chromosome 9 rather than the P_2, as demonstrated by a genome-wide association study (GWAS) that identified numerous loci associated with alate size, all on chromosome 9 (Figure 4E). The presence of a single small gyne without the P_2 but with the 9r is consistent with this pattern (Figure 4A). Overall, gynes and queens with at least one 9r copy are significantly smaller than 9a9a gynes (F_2=232.2, $p< 0.0001$) (Figure 4B). On average, 9a9r gynes are 15.7% smaller than 9a9a gynes (t_{146} = 16.1, $p < 0.0001$, linear mixed model). This size reduction is 20.3% in 9r homozygous gynes (t_{256}= 17.3, $p< 0.0001$; linear mixed model). 9r homozygous gynes are 5.42% smaller than 9r heterozygous gynes (t_{270}= 4.4, $p< 0.0001$; linear mixed model). Males exhibit a similar pattern, although male miniaturization appears to be less drastic, with 9r males being 8.6% smaller than 9a males ($t_{89.9}$= 12.7, $p< 0.0001$; linear mixed model) (Figure 4D). We also observed a significant reduction in body size of 4.9% in 9a9a gynes with at least one P_1 haplotype compared to 9a9a gynes without a P haplotype (M_AM_A-M_AP_1, $t_{197.5}$=6.7, $p< 0.0001$; M_AM_D-M_AP_1, $t_{202.4}$=6.8; $p< 0.0001$; M_AM_A-P_1P_1,$t_{140.6}$=4.7, $p< 0.001$; M_AM_D-P_1P_1, $t_{144.2}$=5.4, $p< 0.0001$; linear mixed model; we did not observe a significant difference between M_AM_D and M_AM_A gynes and queens and M_DP_1 gynes and queens due to the small sample size of the latter [n=6], although they are smaller on average). Interestingly, P1 males are on average 2.6% bigger than M_A males ($t_{-119.8}$=-3.1, $p< 0.05$; linear mixed model).

Discussion

Most ant species harboring a social supergene have only two alternative haplotypes, one associated with monogyny and the other associated with polygyny[37]. Here we describe for the first time a species, *Formica cinerea*, that bears four supergene haplotypes on chromosome 3, all co-occurring in a single population. As found in several congeneric species so far[17,36,38,39], the social form in *F. cinerea* is genetically controlled. Two M haplotypes (M_A and M_D) are strongly associated with single-queen colonies, while two P haplotypes (P_1 and P_2) are almost exclusively present in multi-queen colonies. We

discovered a novel rearranged supergene variant (9r) on chromosome 9 underlying queen miniaturization, in strong LD with the P_2 polygyne-associated haplotype.

Socially antagonistic alleles and supergene expansion

Alternative social forms in ants generally conform to the "polygyny syndrome" in which gynes of polygyne colonies are about 10% smaller and have lower relative fat content than those produced by monogyne colonies[40]. In *F. cinerea* polygyne colonies, we observed two distinct gyne sizes: 9a9a gynes are relatively large (though still 5% smaller on average than monogyne-produced gynes); in contrast, gynes with a 9r haplotype are 16-20% smaller than 9a9a gynes (Figures 4A-B). This aligns with other cases of extreme queen-size dimorphism (microgynes and macrogynes)[41]. No 9r *F. cinerea* gynes or queens have been observed in monogyne colonies. Polygyny, therefore, appears to be a precondition for microgyny in this species.

We suggest that fitness epistasis initially emerged between an ancestral P haplotype and an incipient mutation on chromosome 9 that caused reduced body size in queens. In the process of establishing a new colony, macrogynes rely solely on their body reserves (wing muscles and fat bodies) to raise their first brood[42,43]. In order to be successful it is essential that they produce a worker caste in a short time, before depleting all their body reserves. The independent colony founding strategy is highly risky, and founding queens often suffer high mortality[49]. Microgynes lack large fat reserves necessary to establish new nests[40,41,44], and thus, would be severely disadvantaged in an independent founding monogyne context. Conversely, in the polygyne background, colony foundation risks are reduced because queens can join existing colonies. Based on studies of microgynes in other species[41,44], we hypothesize that *F. cinerea* microgynes are less costly to produce. Although they are expected to lay fewer eggs than macrogynes[41,44] their lower fecundity could be buffered by coexistence with other reproductive queens.

Genetic mismatches between body size and colony social form would have a high cost, leading to strong selection for LD between alleles on P2 and 9r (Figure S1C). This expanded region of suppressed recombination spanning two chromosomes would include socially antagonistic alleles, beneficial in polygyne colonies but detrimental in monogyne colonies.

We also observed miniaturization in males (Figures 4C-D), where 9r males are 8.6% smaller than 9a males. Miniaturization of males may be a byproduct of selection on queen body size, or small males may be favored in some contexts. For example, body size reduction in both

sexes may lead to assortative mating between morphotypes. Incomplete assortative mating with respect to social form has been documented in *F. selysi*, where 80% of queens of monogyne origin mated with males from monogyne colonies, while the remaining 20% mated with males from polygyne colonies[25,28]. Here we speculate that small 9r males from polygyne colonies are disadvantaged during mating flights, where they must compete with large 9a males (M_A and P_1). If true, they may adopt alternative strategies by mating close to nests, with small 9r gynes.

Ultimately, our results are consistent with predictions of the canonical model of sex chromosome evolution in which antagonistic selection leads to the expansion of regions of suppressed recombination between advantageous combinations of alleles. Two novel features are present in our system. First, the expansion of LD is occurring in an autosomal supergene instead of in a sex chromosome[1,2]. Second, the environment that shapes alternative traits is the social context determined by colony queen number as opposed to sex[1] or the extrinsic environment[9]. We note that we cannot rule out the possibility that the initial mutation leading to an association between chromosome 9 and the P_2 haplotype was selectively neutral, and that this association enabled the invasion of a queen miniaturizing mutation on chromosome 9[45].

Although they exhibit high levels of LD, P_2 and 9r are not perfectly correlated (Figure 2B). The occasional decoupling of alleles suggests that recombination occasionally happens between chromosomes 3 and 9, raising questions about how these two supergene regions are associated. Several alternative mechanisms could mediate the incomplete association between P_2 and 9r. We speculate that P_2 and 9r may be physically linked by the fusion of chromosomes 3 and 9 or, alternatively, that they are linked through a reciprocal translocation. Neo-sex chromosomes shaped by suppression of recombination between an autosome and an ancestral sex chromosome have been documented in a variety of eukaryotic organisms [e.g., reptiles[5]; fish[4,5]; birds[46]; insects[47]; and plants[6]]. These neo-sex chromosomes often arise from Robertsonian fusion of acrocentric chromosomes[48,49] or the reciprocal translocation of genetic material between non-homologous chromosomes[6,50]. We suggest that our system may be analogous to the formation of neo-sex chromosomes. Suppression of recombination between 9r and P_2 could initially be achieved through chromosomal fusion, reciprocal translocation, or very strong epistasis. Strong epistasis without physical linkage can occur if individuals with mismatched genotypes (e.g. P_2 without 9r or 9r without P_2) experience a high mortality rate during development. We observed 20

cases where the 9r and P_2 were not co-transmitted to offspring. These genotypic mismatches occurred in both directions (9r without P_2 and P_2 without 9r). If the strong association between the two is derived from physical changes in the chromosome structure, these exceptions could result from rare double-recombination events. Alternatively, if selection against mismatches is present, the exceptions reveal that such a system is not 100% lethal. In either scenario, the observation that recombinant individuals were significantly more likely to be workers suggests that decoupling P_2 and 9r could bias the development of immature offspring toward workers instead of gynes. Further research is needed to identify the mechanism that locked these two regions of suppressed recombination together.

Microgyny as an incipient form of intraspecific social parasitism?

Queen-size dimorphism associated with polygyny may lead to intraspecific parasitism, where the microgynes take advantage of the macrogynes by specializing in sexual offspring production[51]. Researchers are generally interested in the origins of social parasitism (e.g. Trible et al.[52]), and Linksvayer et al.[53] predicted that a supergene might underlie the transition from a free-living to a socially parasitic lifestyle in ants. In *Formica,* queen miniaturization was previously described only in species that parasitize other *Formica* species (*difficilis*, *dakotensis*, and *exsecta* clades)[54], although it has not been linked with colony social organization. Here we describe microgyny in a non-parasitic *Formica* species for the first time and speculate that 9r microgynes could be incipient intraspecific social parasites. The best-known case of intraspecific parasitism occurs in *Myrmica rubra: w*hen microgynes and macrogynes coexist in the same nest, microgynes produce very few worker offspring, focusing their reproductive effort mostly on sexuals[51,55]. During our field collections, we tried to minimize damage to nests, so we did not observe mature microgynes and macrogynes occurring together in the same nest. However, we found four colonies where all the workers were 9a9a homozygotes while alates were 9a9r and 9r9r microgynes. We also noticed that virgin microgynes and macrogynes were never produced by the same colony. Although preliminary, these findings could represent the first hint that *F. cinerea* microgynes are intraspecific social parasites.

Supergene variation associated with three complex traits

Our results reveal that four haplotypes on chromosome 3 detected in *F. cinerea* are associated with at least three complex traits: social structure, alate size, and sex ratio. We

show that microgyny is controlled by a newly discovered supergene region on chromosome 9. As already studied in other *Formica* species, we confirm that M haplotypes are associated with monogyne colonies, while P haplotypes are associated with polygyny. However, we found a few exceptions to this pattern: several apparently monogyne colonies include individuals with a P haplotype, and several apparently polygyne colonies lack P haplotypes (Figures S2A-B). We suggest that these exceptions could be an expression of the natural decline of the colony. Polygyne colonies may be functionally monogyne at the end of their lifespan if only one queen is effectively reproductive or has survived[56]. On the other hand, declining monogyne colonies may be more inclined to accept a new, most likely related, queen if their queen stops producing enough eggs[57].

A third complex phenotypic trait, colony sex ratio, is associated with the M_D haplotype, aligning with recent discoveries in *F. glacialis* and *F. podzolica*[36]. We find that *F. cinerea* monogyne colonies, regardless of the number of matings, specialize in the production of gynes or males. We show that split sex ratio is mediated by M_D and M_A haplotypes. Based on inferences from offspring genotypes, queens heterozygous for M_D tend to produce gynes, while queens homozygous for M_A tend to produce males. In contrast, polygyne colonies are mostly male-producing or produce both males and gynes. Structurally, the M_D haplotype in *F. cinerea* spans the first half of chromosome 3 as in *F. glacialis* and *F. podzolica*. In a further parallel, we mainly found the M_D haplotype in heterozygous females, and observed a very low frequency of M_D homozygotes and haploids. We do not yet have enough information to determine whether these M_D haplotypes share a common origin or originated independently.

Conclusions

A novel supergene variant (9r) on chromosome 9 underlying a 16-20% reduction of queen body size (microgyny) is highly associated with the polygyne-associated P_2 haplotype on chromosome 3. Microgynes are absent from *F. cinerea* monogyne colonies, consistent with previous hypotheses that polygyny is a precondition for microgyny[41,43]. Here we propose that socially antagonistic selection favored the suppression of recombination between a P haplotype and a miniaturizing allele on chromosome 9, consistent with the canonical model of sex chromosome evolution[2]. While models of sex chromosome evolution have been used as a source of inspiration for supergene research for more than a decade, many studies have applied these models to try to understand degeneration of a non-recombining supergene haplotype (e.g. Wang et al.[16]; Tuttle et al.[58]; Stolle et al.[59]). Here, we add

empirical support to the idea that such models can also provide a useful starting point for understanding the origin and expansion of autosomal supergenes (reviewed by Gutierrez-Valencia et al.[60]). In recent years, new models from both the sex chromosome and local adaptation research fields describe additional hypotheses for the emergence of regions of suppressed recombination[9,61,62]. The hypotheses developed in these models should also be tested in autosomal supergene systems, as this may lead to new breakthroughs in our understanding of the evolution of recombination.

Acknowledgments

This material is based upon work supported by the US National Science Foundation DEB Grant No. 1754834 to A.B. and J.P. and DEB Grant No. 1631776 to J.P. Computations were performed using the computer clusters and data storage resources of the UCR HPCC, which were funded by grants from NSF (MRI-2215705, MRI-1429826) and NIH (1S10OD016290-01A1). This publication includes data generated at the UC San Diego IGM Genomics Center utilizing an Illumina NovaSeq 6000 that was purchased with funding from a NIH SIG grant #S10 OD026929. We thank D. Zarate, G. Lagunas- Robles, Z. Alam, D. Pierce, N. Najar, and the FoG discussion group (Campbell, Samuk, and Ostevik labs) for providing helpful feedback on an earlier version of the manuscript and M. Molfini and B. Purcell for assistance in the field.

Author contribution

Conceptualization, G.S, A.B., and J.P.; Methodology, G.S, A.B., and J.P.; Formal Analysis, G.S.; Investigation, G.S., M.P., A.B., and J.P.; Resources, A.B. and J.P.; Writing – Original Draft, G.S. and J.P.; Writing – Review & Editing, G.S., M.P., A.B., and J.P.; Visualization, G.S. and J.P.; Supervision, A.B. and J.P.; Funding Acquisition, A.B. and J.P.

Declaration of interests

The authors declare no competing interests.

Figure legends

Figure 1. **Principal component analysis and genetic differentiation identify four supergene haplotypes, including one encompassing variation on both chromosome 3 and chromosome 9**.

Principal component axes 1 and 2 (A) distinguish six groups of individuals. The solid colored circles show homozygous individuals (based on positive F_{IS} values). Dual-colored circles show heterozygous individuals (based on negative F_{IS} values). Each half circle represents a haplotype on chromosome 3, and haploid males are represented by half circles. PC axis 3 reveals a fourth haplotype (M_D) that is distinct from M_A over a smaller region of chromosome 3 (B). Individuals with at least one copy of M_D exhibited relatively high PC3 values. The M_D compared to M_A spans only the first half of chromosome 3 (C). Elevated differentiation (FST) occurred between the M and P haplotypes across most of chromosome 3 when comparing haploid males (D, E). High differentiation was also evident on chromosome 9 (F–H), when comparing the P_2 haplotype to the other three haplotypes. PC1 explains 51% of the total variance, while PC2 and PC3 explain respectively 29% and 3.8%.

Figure 2. PCA of variants on chromosome 9 identified three clusters corresponding to three supergene genotypes.

The left cluster contains 9a9a individuals, while middle and right clusters show respectively 9a9r and 9r9r individuals (A). The colors of half circles in the PCA indicate chromosome 3 haplotypes to reveal mismatches between chromosomes 3 and 9. PC1 explains 75% of the total variance, and PC2 2.1%. Individuals with the 9r haplotype on chromosome 9 almost always have the P2 haplotype on chromosome 3, although we found some mismatches (13 out of 1151 9a9a individuals, all workers, harbor at least one copy of the P_2; 7 out of 130 9a9r individuals, one gyne and six workers, do not carry the P_2 haplotype). The chord diagram (B) shows associations between genotypes on chromosome 9 (left segments) and genotypes on chromosome 3 (right segments). Note that the ribbons connect chromosome 9 with chromosome 3 genotypes in the same individuals. Related to Figure S1.

Figure 3. Association of the haplotypes on chromosome 3 with colony social form.

Ridgeline plots of the distribution of pairwise relatedness among nestmates reveal variation in colony social structure. Of the 120 colonies analyzed, half were monogyne (39 monandrous monogynous, 21 polyandrous monogynous) and half were polygyne (A). In monogyne monandrous (B) and polyandrous (C) colonies, most colonies contained either exclusively M_AM_A workers or a mix of M_AM_A and M_AM_D workers. In contrast, most polygyne

colonies (D) contained individuals with at least one copy of one of the P haplotypes (P_1 and P_2). A few exceptions to this general pattern are observed in both monogyne and polygyne colonies (Related to Figure. S2). Vertical lines at 0.7 and 0.2 show approximately where we expect peaks of full- and half-siblings, respectively, considering the downward bias typical of relatedness estimates based on RADseq markers.

Figure 4. The 9r haplotype is strongly associated with reduced head width in gynes and males.

Individual gynes and males with the P_2 haplotype are significantly smaller than those without (A, C). Likewise, gynes and males with at least one copy of the 9r haplotype are significantly smaller than those without a 9r haplotype (B, D). GWAS analysis (E) confirms that a large region on chromosome 9 is most strongly associated with body size miniaturization in *F. cinerea*. The blue line shows the significance threshold adjusted for Bonferroni correction.

STAR Methods

LEAD CONTACT AND MATERIALS AVAILABILITY

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Jessica Purcell (jpurcell@ucr.edu).

Materials availability

This study did not generate new unique reagents. There are restrictions to the availability of tissue and DNA samples due to the lack of an external centralized repository for their distribution and our need to maintain the stock. We are glad to share oligonucleotides with reasonable compensation by requestor for processing and shipping.

Data and code availability

- Raw Illumina sequencing reads are available at the National Center for Biotechnology Information Short Reads Archive, BioProject PRJNA966702.
- Phenotypic data and supergene genotypes for each individual ant are available in Dryad, https://doi.org/doi:10.5061/dryad.02v6wwq8s.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Formica cinerea is a socially polymorphic species with a wide distribution across Europe[63]. This species nests preferentially along sand and gravel banks of rivers and open sand dunes. We collected *F. cinerea* workers and alates (gynes and males) from colonies in northern Italy (Aosta Valley and Piedmont) in June-July across several years, 2014, 2018-2021 (Table S1). Whenever possible, we sampled up to 10 gynes and males, and about 15 workers from each colony, and noted the observed sex-ratio. When multiple mature queens were found within colonies, we also sampled a subset of them. During 2019-2021, we collected newly mated wingless queens that were either looking for suitable locations to start new colonies or were under stones in self-dug chambers with no workers. We stored samples in 96-100% ethanol.

METHOD DETAILS

Library preparation

We extracted DNA from the head and thorax of workers, and only the head of gynes, queens and males. For the 2014 and 2018-2020 samples, we used the QIAGEN DNeasy Blood & Tissue Kit with modifications described in McGuire et al.[38]. Specifically,we manually ground the tissue with sterile pestles in a 1.7 mL tube while immersed in liquid nitrogen, and left the pulverized samples overnight in a solution of 180 µL of buffer ATL and 20 µL of proteinase K at 56°C. The day after we added 200 µL of buffer AL and 200 µL of 100% ethanol. We then transferred the supernatant into alternatively sourced spin columns (BPI-tech.com), added 70% ethanol for DNA wash, and eluted the DNA in 30 µL of buffer EB. We extracted individuals collected in 2021 using the QiaAmp 96 DNA QiaCube HT kit. We manually ground the ant tissues as described above, and, following the overnight digestion in 180 µL of buffer ATL and 20 µL of proteinase K, we transferred the supernatant to the QIAcube HT/QIAxtractor robot to complete the extraction. We followed the automatized protocol QiaAmp 96 DNA. We eluted the DNA in 100 µL of buffer EB.

We sequenced all samples using a double-digest restriction site-associated DNA sequencing (RADseq) approach (protocol from Brelsford et al.[64]). We digested 6 µL DNA per sample using restriction enzymes MseI and PstI and incubated the samples at 37°C for 3 hours on a thermal cycler with a heated lid. We then ligated a universal MseI adapter and uniquely barcoded PstI adapter to each sample. After an incubation of 3 hours at 16° C on a thermal cycler, we diluted the product adding 40 µL of water. We then removed small DNA

fragments using Serapure magnetic beads[65] or Omega magnetic beads (Omega Bio-tek, 2021) in a 0.8:1 ratio (beads: sample solution) and removed impurities with two consecutive 70% ethanol washes. We air-dried the magnetic beads for 10-15 minutes to remove all traces of ethanol. Finally, we resuspend the DNA adding 40 µL of water. We amplified each sample in four separate PCR reactions with indexed Illumina primers and then pooled the replicate PCR products for each sample for a final PCR cycle, with added primers and dNTP. We ran each PCR product on a 1.5% agarose gel for 20 minutes. Finally, we pooled the samples that were successfully amplified in a tube and did a final round of small fragment removal using the magnetic beads. We sequenced all libraries using 150 bp paired-end reads on Illumina Novaseq 6000 or HiSeq X sequencers. Sample sizes and sequencing details for each batch are provided (Table S2).

Bioinformatics

We used *Stacks* 2.60 to demultiplex our data with default parameters[66], *PEAR* v0.9.10[67] to merge paired-end reads and remove adaptor sequences, and *BWA-mem2*[68] to align reads to the *Formica selysi* genome[12]. We called SNPs using *BCFtools mpileup*[69] and filtered the genotypes for a minimum read depth of 7 (--minDP), a minor allele frequency of 5% (--maf) and excluded indels (--remove-indels) and sites with over 80% missing data (--max-missing) using *VCFtools* 0.1.16-18[70].

Excluding duplicated regions

Ant males are haploid, and this feature provides an opportunity to identify and omit duplicated genomic regions. Males were treated as diploid in our initial pipeline, and loci that appeared heterozygous in at least 5% of males were flagged for removal from the complete dataset, because these reflect variable sequences in duplicated regions instead of alternative alleles in a single region of the genome.

Mitigating the batch effect

In order to have an adequate sample size for all supergene genotypes in all castes (particularly gynes and males, which are sampled opportunistically), we added data incrementally across years. Differences in extraction protocols and variation among sequencing lanes caused a batch effect (Figure S4A). To mitigate this issue, we calculated the Weir and Cockerham's F_{ST} between batch pairs at each locus. We then removed all

SNPs showing F_{ST} values ≥ 0.3 in the comparison of at least one pair of batches (because the geographic scope of sampling was similar across years, we would not expect to find true changes in allele frequency of this magnitude) (Figure S4B). Our final dataset resulted in 15129 SNPs and 1415 individuals. Workers, gynes, males and mature queens were collected from 172 colonies, and 95 newly mated queens were collected as they sought a suitable place to start their colony.

Population structure

Formica cinerea samples were collected from 13 localities in northern Italy (Table S1), ranging from 1 km to 82 km apart. To assess the genetic structure of the sampled individuals, we randomly selected 1 worker per colony and filtered the genotypes for a minimum read depth of 7 (--minDP), a minor allele count of 2 (--mac) and excluded indels (--remove-indels) and sites with over 80% missing data (--max-missing) using VCFtools[70]. We removed all the loci suspected to be responsible for the batch effect (see 'Mitigating the batch effect' paragraph) and misaligned due to duplicated regions (see 'Excluding duplicated regions' paragraph). Finally, we excluded markers on chromosomes 3 and 9. This dataset resulted in 139 workers and 27398 SNPs.

We performed a PCA in PLINK v1.90b6.25[71] and plotted the first two principal components in R v3.4.0[72] using the function *ggplot*[73]. Using the same dataset, we ran ADMIXTURE v1.3.0[74] to infer genetic clusters in our dataset for K values from 1 to 13 and assessed the best K value using the cross-validation error. The PCA and ADMIXTURE result (K=1) suggested the absence of population structure and that the samples analyzed in this study belong to a panmictic population (Figures S4C-D).

Determination of colony social form

We used COANCESTRY 1.0.1.10[75] to determine pairwise relatedness using workers and gynes (using Wang[76] estimator) and infer colony social form. To ensure that these analyses were independent of our assessments of supergene variation, we created a dataset that excluded chromosome 3 and chromosome 9. To have a robust assignment, we kept only colonies with at least 5 diploid individuals and excluded haploid males. The final dataset resulted in 761 individuals from 120 colonies. A recent literature review and simulation study confirmed that relatedness estimates tend to be downward biased, yet more precise, in SNP-based datasets with hundreds or thousands of loci compared to microsatellite-based

datasets with fewer loci[77]. Given the known biases in datasets like ours, we called colonies with all pairwise relatedness estimates ≥ 0.6 as monogyne monandrous, colonies with bimodal distribution of pairwise relationships with at least 40% ≥ 0.6, but none <0.2 as monogyne polyandrous, and colonies with at least one pairwise relationship ≤ 0.1 as polygyne. We visualized the distribution of within-colony relatedness estimates with a ridgeline plot produced in R[72] using the function *ggplot* (package ggplot2[73]).

To investigate the association of the colony social organization with the supergene, we first performed a principal component analysis (PCA) for all individuals (workers, males, gynes and queens) using only the 1235 SNPs on chromosome 3, which contains the known *Formica* social supergene[12,18]. We then assigned the genotypes to each individual based on clusters in PCA and F_{IS} value (heterozygous individuals have negative F_{IS} values across the supergene, while homozygotes have positive values). To further investigate the genetic differentiation between each haplotype, we selected haploid males and calculated Weir and Cockerham's F_{ST} for all pairwise combinations of supergene haplotypes. The PCA was calculated in PLINK[71] with the --pca flag, while the F statistics were calculated in VCFtools[70], using the --het flag (F_{IS}) and the --weir-fst-pop flag. Finally, we examined haplotype distribution in monogyne and polygyne colonies.

From the F_{ST} plot we noticed a second supergene on chromosome 9 visible when comparing the P_2 haplotype (see Results section) on chromosome 3 with the other haplotypes. For chromosome 9, we performed a PCA and analysis of F_{IS} using only the 983 loci on that chromosome to assign genotypes to each individual. To identify which of the variants is rearranged relative to the *F. selysi* reference genome, we built two within-haplotype heatmaps of linkage disequilibrium using only homozygous individuals at each haplotype on chromosome 9. We also constructed a third heatmap of linkage disequilibrium between P_2 and 9r using only those individuals that were P_2P_2 on chromosome 3 and 9r9r on chromosome 9. For this analysis, we used the LDheatmap package[80] from R.

Morphometrics

To assess whether polygyne *Formica cinerea* alates (gynes, queens and males) exhibit the reduction in size typical of polygyny syndrome[40], we measured the maximum width across the eyes in 281 gynes and queens and 374 males using a Leica DMC2900 camera mounted on a Leica S8APO at 25× magnification. We used head width because it is known to have

a strong positive correlation with several body segment dimensions in *Formica* species[23,81], and thus serves as a good proxy for body size within caste.

Sex ratio

While inspecting *F. cinerea* colonies during sampling, we took note of whether they exhibited a strongly skewed sex ratio, i.e. whether the colony preferentially produced gynes or males, or both sexes. We attributed the sex ratio to colonies observed with at least seven alates. Gyne producing colonies had at least seven gynes and no more than two males, male producing colonies had at least seven males and no more than two gynes, and mixed colonies were intermediate between the two. In total, 23 *F. cinerea* colonies were male-producing, 13 gyne-producing, and 6 were mixed. For each of these colonies, we looked at the haplotype counts on chromosome 3.

QUANTIFICATION AND STATISTICAL ANALYSIS

Determination of colony social form

We tested the significance of the association between haplotypes on chromosome 3 and colony social form by fitting a generalized linear mixed model (GLMM) with binomial distribution[78], where monogyny is 0 and polygyny is 1. The "presence of P haplotypes" was defined as 0 if no individual in the colony carries a P haplotype, and 1 if at least one individual in the colony harbors a P haplotype (regardless if P_1 or P_2). The variable "presence of P haplotypes" was included as a fixed factor; year, and locality as random factors. Since not all colonies produced alates, we considered only workers. For this analysis, we used the glmer function in R (package lme4[79]). The analysis included 59 monogyne colonies and 60 polygyne colonies.

Testing the association between P_2 and 9r

We verified the association between P_2 and 9r by performing a chi-squared test (levels: P_2 present or absent, 9r present or absent). In total our dataset counted 1134 individuals (605 workers, 257 gynes/queens and 272 males) without P_2 and 9r; 254 individuals (88 workers, 54 gynes/queens and 112 males) with both P2 and 9r; 7 individuals (1 gyne and 6 workers) with 9r but not P_2; 13 individuals (all workers) with P_2 but not 9r. Since we found some mismatches in the co-transmission of P_2 and 9r, we checked whether these mismatches were mainly present in the workers rather than in the reproductive individuals (queens,

gynes, and males) by performing a second chi-squared test (levels: workers or alates, presence or absence of P_2-9r mismatches). Nineteen out of 712 workers and 1 out of 696 alates showed mismatches. Both chi-squared tests were performed in R[72].

Body size association with chromosome 9

To test whether gynes and queens (n= 281) with different supergene genotypes have significantly different sizes, we fit two independent linear mixed models for chromosome 3 and chromosome 9 using colony as a random effect and genotype as a fixed effect. We repeated the same analyses for males (n= 373). For these analyses, we used the R package lme4[79]. Pairwise p-values were obtained after performing Tukey post hoc tests using the *emmeans* function[82] in R.

To identify genomic regions associated with body size, we performed a Genome Wide Association Study (GWAS) using a univariate linear mixed model implemented in *Gemma* v0.94[83]. Males were excluded from this analysis. Since *Gemma* requires that no missing genotypes are present in the data, we imputed missing genotypes with *Beagle* v4.1[84] using the full dataset of SNPs that passed previously mentioned filters. *Gemma* uses a relatedness matrix generated from the sample genetic data to correct for non-independence of the samples due to population structure. We applied a Bonferroni correction to calculate the significance threshold.

Sex ratio

We tested the significance of the association between M_D and gyne production by fitting a generalized linear mixed model (GLMM) with binomial distribution[78], where male-producing colonies (n= 23) are 0 and gyne-producing colonies (n= 13) are 1. Mixed colonies were excluded from this analysis. We also transformed "presence of M_D haplotype" into a binomial variable, where colonies without M_D haplotype were coded as 0, and colonies with at least one individual with a M_D haplotype were coded as 1. The variable "presence of M_D haplotype" was included as a fixed factor; year, and locality as random factors. We considered only workers for this analysis. We used the R package lme4[79] and the glmer function.

References

1. Charlesworth, B., and Charlesworth, D. (1978). A model for the evolution of dioecy and gynodioecy. Am. Nat. *112*, 975–997.
2. Charlesworth, D., Charlesworth, B., and Marais, G. (2005). Steps in the evolution of heteromorphic sex chromosomes. Heredity *95*, 118–128 (2005).
3. Kratochvíl, L., Stöck, M., Rovatsos, M., Bullejos, M., Herpin, A., Jeffries, D. L., Peichel, C. L., Perrin, N., Valenzuela, N. and Pokorná, M. J. (2021). Expanding the classical paradigm: what we have learnt from vertebrates about sex chromosome evolution. Philos. Trans. R. Soc. Lond., B, Biol. Sci. *376*(1833), 20200097.
4. Kitano J., et al. (2009) A role for a neo-sex chromosome in stickleback speciation. Nature *461*(7267), 1079-1083. 10.1038/nature08441
5. Pennell, M. W., Kirkpatrick, M., Otto, S. P., Vamosi, J. C., Peichel, C. L., Valenzuela, N., and Kitano, J. (2015). Y fuse? Sex chromosome fusions in fishes and reptiles. PLoS genet. *11*(5), e1005237.
6. Rifkin, J. L., Beaudry, F. E., Humphries, Z., Choudhury, B. I., Barrett, S. C., and Wright, S. I. (2021). Widespread recombination suppression facilitates plant sex chromosome evolution. MBE *38*(3), 1018-1030.
7. Charlesworth, D., and Charlesworth, B. (1979). Selection on recombination in clines. Genetics *91*(3), 581-589. 10.1093/genetics/91.3.581
8. Kirkpatrick, M., and Barton, N. (2006). Chromosome inversions, local adaptation and speciation. Genetics *173*, 419–434. 10.1534/genetics.105.047985
9. Yeaman, S. (2013). Genomic rearrangements and the evolution of clusters of locally adaptive loci. PNAS *110*, E1743–E1751. 10.1073/pnas.1219381110
10. Branco, S., et al., (2017) Evolutionary strata on young mating–type chromosomes despite the lack of sexual antagonism. PNAS *114*, 7067–7072. 10.1073/pnas.1701658114
11. Branco, S., et al. (2018). Multiple convergent supergene evolution events in mating–type chromosomes. Nat. Commun. *9*, 1–13. 10.1038/s41467-018-04380-9
12. Brelsford, A., Purcell, J., Avril, A., Van, P.T., Zhang, J., Brütsch, T., Sundström, L., Helanterä, H. and Chapuisat, M. (2020). An ancient and eroded social supergene is widespread across *Formica* ants. Curr. Biol. *30*, 304–311. 10.1016/j.cub.2019.11.032
13. Duhamel, M., Carpentier, F., Begerow, D., Hood, M. E., Rodríguez de la Vega, R. C., and Giraud, T. (2022). Onset and stepwise extensions of recombination suppression are common in mating-type chromosomes of *Microbotryum* anther-smut fungi. J. Evol. Biol. *35*, 1619–1634. 10.1111/jeb.13991

14. Kess, T., et al. (2019). A migration–associated supergene reveals loss of biocomplexity in Atlantic cod. Sci. Adv. 5, eaav2461. 10.1126/sciadv.aav2461
15. Joron, M., et al. (2011). Chromosomal rearrangements maintain a polymorphic supergene controlling butterfly mimicry. Nature *477*, 203–206. 10.1038/nature10341
16. Wang, J., Wurm, Y., Nipitwattanaphon, M., Riba-Grognuz, O., Huang, Y. C., Shoemaker, D., & Keller, L. (2013). A Y–like social chromosome causes alternative colony organization in fire ants. Nature *493*, 664–668. 10.1038/nature11832
17. Purcell, J., Brelsford, A., Wurm, Y., Perrin, N., and Chapuisat, M. (2014). Convergent genetic architecture underlies social organization in ants. Curr. Biol. *24*, 2728–2732. 10.1016/j.cub.2014.09.071
18. Purcell, J., Lagunas-Robles, G., Rabeling, C., Borowiec, M. L., and Brelsford, A. (2021). The maintenance of polymorphism in an ancient social supergene. Mol. Ecol. *30*, 6246–6258. 10.1111/mec.16196
19. Chapuisat, M. (2023). Supergenes as drivers of ant evolution. Myrmecol. News *33*, 1–1
20. Martinez-Ruiz, C., Pracana, R., Stolle, E., Paris, C. I., Nichols, R. A., and Wurm, Y. (2020). Genomic architecture and evolutionary antagonism drive allelic expression bias in the social supergene of red fire ants. eLife 9, e64678. 10.7554/eLife.55862
21. Coughlan, J. M., and Willis, J. H. (2019). Dissecting the role of a large chromosomal inversion in life history divergence throughout the *Mimulus guttatus* species complex. Mol. Ecol. *28*(6), 1343-1357.
22. Sundström, L. (1995a). Dispersal polymorphism and physiological condition of males and females in the ant, *Formica truncorum*. Behav. Ecol. *6*, 132–139. 10.1093/beheco/6.2.132
23. Schwander, T., Rosset, H., and Chapuisat, M. (2005). Division of labour and worker size polymorphism in ant colonies: the impact of social and genetic factors. Behav. Ecol. Sociobiol. *59*, 215–221. 10.1007/s00265-005-0027-6
24. Rosset, H., and Chapuisat, M. (2007). Alternative life–histories in a socially polymorphic ant. Evol. Ecol. *21*, 577–588. 10.1007/s10682-006-9139-3
25. Fontcuberta, A., De Gasperin, O., Avril, A., Dind, S., and Chapuisat, M. (2021). Disentangling the mechanisms linking dispersal and sociality in supergene-mediated ant social forms. Proc. R. Soc. B *288(*1949), 20210118.
26. Sundström, L. (1995b). Sex allocation and colony maintenance in monogyne and polygyne colonies of *Formica truncorum* (Hymenoptera: Formicidae): the impact of kinship and mating structure. Am. Nat. *146*, 182–201.
27. Rosset, H., and Chapuisat, M. (2006). Sex allocation conflict in ants: when the queen rules. Curr. Biol. *16*, 328–331. 10.1016/j.cub.2005.12.036
28. Avril, A., Purcell, J., Brelsford, A., and Chapuisat, M. (2019). Asymmetric assortative mating and queen polyandry are linked to a supergene controlling ant social organization. Mol. Ecol. *28*, 1428–1438. 10.1111/mec.14793
29. Avril, A., Purcell, J., Béniguel, S., and Chapuisat, M. (2020). Maternal effect killing

by a supergene controlling ant social organization. PNAS *117*, 17130–17134.

30. Tafreshi, A. G., Otto, S. P., and Chapuisat, M. (2022). Unbalanced selection: the challenge of maintaining a social polymorphism when a supergene is selfish. Philos. Trans. R. Soc. B: Biol. Sci. *377*, 20210197. 10.1098/rstb.2021.0197
31. Blacher, P., De Gasperin, O., Grasso, G., Sarton-Lohéac, S., Allemann, R., and Chapuisat, M. (2023). Cryptic recessive lethality of a supergene controlling social organization in ants. Mol. Ecol. *32*, 1062–1072. 10.1111/mec.16821
32. Li, Q., et al. (2022). Local adaptation and the evolution of genome architecture in threespine stickleback. Genome Biol. Evol. *14*, evac075. 10.1093/gbe/evac075
33. Trivers, R. L., and Hare, H. (1976). Haplodiploidy and the Evolution of the Social Insect: The unusual traits of the social insects are uniquely explained by Hamilton's kinship theory. Science *191*(4224), 249-263.
34. Pamilo, P., and Rosengren, R. (1984). Evolution of nesting strategies of ants: genetic evidence from different population types of *Formica* ants. Biol. J. Linn. Soc. *21*(3), 331-348 (1984). 10.1111/j.1095-8312.1984.tb00370.x
35. Boomsma, J. J., and Grafen, A. (1990). Intraspecific variation in ant sex ratios and the Trivers-Hare hypothesis. Evol. *44*(4), 1026-1034. 10.1111/j.1558-5646.1990.tb03823.x
36. Lagunas-Robles, G., Purcell, J., and Brelsford, A. (2021). Linked supergenes underlie split sex ratio and social organization in an ant. PNAS *118*, e2101427118 (2021). 10.1073/pnas.2101427118
37. Kay, T., Helleu, Q., and Keller, L. (2022). Iterative evolution of supergene–based social polymorphism in ants. Philos. Trans. R. Soc. B: Biol. Sci. *377*, 20210196. 10.1098/rstb.2021.0196
38. McGuire, D., Sankovitz, M., and Purcell, J. (2022). Purcell, A novel distribution of supergene genotypes is present in the socially polymorphic ant *Formica neoclara*. BMC Ecol. Evol. *22*, 1–12. 10.1186/s12862-022-02001-0
39. Pierce, D., Sun, P., Purcell, J., and Brelsford, A. (2022). Brelsford, A socially polymorphic *Formica* ant species exhibits a novel distribution of social supergene genotypes. J. Evol. Biol. *35*(8), 1031-1044 (2022). 10.1111/jeb.14038
40. Keller, L. (1993). The assessment of reproductive success of queens in ants and other social insects. Oikos 177-180. 10.2307/3545107
41. Wolf, J. I., and Seppä, P. (2016). Queen size dimorphism in social insects. Insectes Soc*. 63*(1), 25-38. 10.1007/s00040-015-0445-z
42. Wheeler, D. E., and Buck, N. A. (1996). Depletion of reserves in ant queens during claustral colony founding. Insectes Soc. *43*, 297-302. 10.1007/BF01242930
43. Peeters, C., and Ito, F. (2001). Colony dispersal and the evolution of queen morphology in social Hymenoptera. Annu. Rev. Entomol. *46*(1), 601-630.
44. Lachaud, J. P., Cadena, A., Schatz, B., Pérez-Lachaud, G., and Ibarra-Núñez, G. (1999). Queen dimorphism and reproductive capacity in the ponerine ant, *Ectatomma ruidum* Roger. Oecologia *120*(4), 515-523. 10.1007/s004420050885

45. Ponnikas, S., Sigeman, H., Abbott, J. K., and Hansson, B. (2018). Why do sex chromosomes stop recombining?. Trends Genet. *34*(7), 492-503. 10.1016/j.tig.2018.04.001

46. Sigeman, H., et al. (2021). Avian neo-sex chromosomes reveal dynamics of recombination suppression and W degeneration. MBE *38*(12), 5275-5291.

47. Nguyen, P., Sýkorová, M., Šíchová, J., Kůta, V., Dalíková, M., Čapková Frydrychová, R., Neven L. G., Sahara, K., and Marec, F. (2013). Neo-sex chromosomes and adaptive potential in tortricid pests. PNAS *110*(17), 6931-6936. 10.1073/pnas.1220372110

48. Gruetzner, F., Ashley, T., Rowell, D. M., and Marshall Graves, J. A. (2006). How did the platypus get its sex chromosome chain? A comparison of meiotic multiples and sex chromosomes in plants and animals. Chromosoma *115*, 75-88.

49. Palacios-Gimenez, O. M., Marti, D. A., and Cabral-de-Mello, D. C. (2015). Neo-sex chromosomes of *Ronderosia bergi*: insight into the evolution of sex chromosomes in grasshoppers. Chromosoma *124*, 353-365.

50. Toups, M. A., Rodrigues, N., Perrin, N., & Kirkpatrick, M. (2019). A reciprocal translocation radically reshapes sex-linked inheritance in the common frog. Mol. ecol. *28*(8), 1877-1889.

51. Schär, S., and Nash, D. R. (2014). Evidence that microgynes of *Myrmica rubra* ants are social parasites that attack old host colonies. J. Evol. Biol. *27*(11), 2396-2407. 10.1111/jeb.12482

52. Trible, W., Chandra, V., Lacy, K.D., Limón, G., McKenzie, S.K., Olivos-Cisneros, L., Arsenault, S.V. and Kronauer, D.J. (2023). A caste differentiation mutant elucidates the evolution of socially parasitic ants. Curr. Biol. *33*(6), 1047-1058. 10.1016/j.cub.2023.01.067

53. Linksvayer, T.A., Busch, J.W. and Smith, C.R. (2013). Social supergenes of superorganisms: do supergenes play important roles in social evolution?. *BioEssays*, *35*(8), pp.683-689. doi.org/10.1002/bies.201300038

54. Borowiec, M. L., Cover, S. P., and Rabeling, C. (2021). The evolution of social parasitism in *Formica* ants revealed by a global phylogeny. PNAS *118*(38), e2026029118. 10.1073/pnas.2026029118

55. Leppänen, J., Seppä, P., Vepsäläinen, K., and Savolainen, R. (2015). Genetic divergence between the sympatric queen morphs of the ant *Myrmica rubra*. Mol. Ecol. *24*(10), 2463-2476 (2015). 10.1111/mec.13170

56. Purcell, J., and Chapuisat, M. (2013). Bidirectional shifts in colony queen number in a socially polymorphic ant population. Evol. *67*(4), 1169-1180

57. Al-Lawati, H., and Bienefeld, K. (2009). Maternal age effects on embryo mortality and juvenile development of offspring in the honey bee (Hymenoptera: Apidae). Ann. Entomol. Soc. Am. *102*, 881–888.

58. Tuttle, E.M., Bergland, A.O., Korody, M.L., Brewer, M.S., Newhouse, D.J., Minx, P., Stager, M., Betuel, A., Cheviron, Z.A., Warren, W.C. and Gonser, R.A., 2016.

Divergence and functional degradation of a sex chromosome-like supergene. Curr. Biol. *26*(3), 344-350. doi.org/10.1016/j.cub.2015.11.069

59. Stolle, E., Pracana, R., Howard, P., Paris, C.I., Brown, S.J., Castillo-Carrillo, C., Rossiter, S.J. and Wurm, Y. (2019). Degenerative expansion of a young supergene. MBE *36*(3), 553-561. doi.org/10.1093/molbev/msy236
60. Gutiérrez-Valencia, J., Hughes, P.W., Berdan, E.L. and Slotte, T. (2021). The genomic architecture and evolutionary fates of supergenes. *GBE 13*(5), p.evab057. doi.org/10.1093/gbe/evab057
61. Lenormand, T., and Roze, D. (2022). Y recombination arrest and degeneration in the absence of sexual dimorphism. Science *375*, 663–666. 10.1126/science.abj1813
62. Jay, P., Tezenas, E., Véber, A., and Giraud, T. (2022). Sheltering of deleterious mutations explains the stepwise extension of recombination suppression on sex chromosomes and other supergenes. PLoS Biol. *20*, e3001698. 10.1371/journal.pbio.3001698
63. Seifert, B. (2018). The Ants of Central and North Europe. lutra Verlags – und Vertriebsgesellschaft Tauer, Germany, 310–312 pp.
64. Brelsford, A., Rodrigues, N., and Perrin, N. (2016). High-density linkage maps fail to detect any genetic component to sex determination in a *Rana temporaria* family. J. Evol. Biol. *29*, 220–225. 10.1111/jeb.12747
65. Rohland, N., and Reich, D. (2012). Cost-effective, high-throughput DNA sequencing libraries for multiplexed target capture. Genome Res. *22*, 939–946. 10.1101/gr.128124.111
66. Catchen, J. M., Amores, A., Hohenlohe, P., Cresko, W., and Postlethwait, J. H. (2011).Stacks: building and genotyping loci de novo from short-read sequences. G3: Genes| genomes| genetics *1*, 171–182. 10.1534/g3.111.000240
67. Zhang, J., Kobert, K., Flouri, T., and Stamatakis, A. (2014). PEAR: a fast and accurate Illumina Paired-End reAd mergeR. Bioinformatics, 30, 614–620. 10.1093/bioinformatics/btt593
68. Vasimuddin, M., Misra, S., Li, H., and Aluru, S. (2019). “Efficient architecture-aware acceleration of bwa-mem for multicore systems” in 2019 IEEE International Parallel and Distributed Processing Symposium (IPDPS)., pp. 314–324. 10.1109/IPDPS.2019.00041
69. Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics *25*, 1754–1760. 10.1093/bioinformatics/btp324
70. Danecek, P., et al. (2011) The variant call format and VCFtools. Bioinformatics 27, 2156–2158. 10.1093/bioinformatics/btr330
71. Purcell, S., et al. (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. *81*(3), 559-575. 10.1086/519795
72. R Core Team, R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available at *http://www. R-*

project. org/.Deposited 2016

73. Wickham, H. (2009). Ggplot: using the grammar of graphics with R. New York, 1076.
74. Alexander, D. H., Novembre, J., and Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. Genome Res. *19*(9), 1655–1664
75. Wang, J. (2011). COANCESTRY: a program for simulating, estimating and analysing relatedness and inbreeding coefficients. Mol. Ecol. Resour. *11*, 141–145. 10.1111/j.1755-0998.2010.02885.x
76. Wang, J. (2002). An estimator for pairwise relatedness using molecular markers. Genetics *160*(3), 1203–1215. 10.1093/genetics/160.3.1203
77. Attard, C. R., Beheregaray, L. B., and Möller, L. M. (2018). Genotyping-by-sequencing for estimating relatedness in nonmodel organisms: Avoiding the trap of precise bias. Mol. Ecol. Resour. *18*(3), 381-390. 10.1111/1755-0998.12739
78. Bolker, B.M., Brooks, M.E., Clark, C.J., Geange, S.W., Poulsen, J.R., Stevens, M.H.H. and White, J.S.S. (2009). Generalized linear mixed models: a practical guide for ecology and evolution. TREE *24*(3), pp.127-135.
79. Bates, D., Mächler, M., Bolker, B., and Walker, S. (2014). Fitting linear mixed-effects models using lme4. Availabe at *https://arxiv.org/abs/1406.5823*. Deposited 23 Jun 2014
80. Shin, J.H., Blay, S., Lewin-Koh, N., McNeney, B., Yang, G., Reyers, M., Yan, Y. and Graham, J. (2016). Package 'LDheatmap'. R package.
81. Tawdros, S., West, M., and Purcell, J. (2020). Scaling relationships in *Formica* ants with continuous worker size variation. Insectes Soc. *67*(4), 463-472. 10.1007/s00040-020-00779-0
82. Lenth, R., Singmann, H., Love, J., Buerkner, P., and Herve, M. (2019). Package 'emmeans'.
83. Zhou, X., and Stephens, M. (2012). Genome-wide efficient mixed-model analysis for association studies. Nat. Genet. *44*(7), 821-824. 10.1038/ng.2310
84. Browning, B. L., and Browning, S. R. (2016). Genotype imputation with millions of reference samples*. Am. J. Hum. Genet. 98*(1), 116-126. 10.1016/j.ajhg.2015.11.020
85. Toews, D. P., Brelsford, A., Grossen, C., Milá, B., and Irwin, D. E. (2016). Genomic variation across the Yellow-rumped Warbler species complex. Auk *133*(4), 698-717.