Recurring adaptive introgression of a supergene that determines social organization

Eckart Stolle[1,2§*], Rodrigo Pracana[1§*], Marian K. Priebe[1], Gabriel Luis Hernández[1], Claudia Castillo-Carrillo[1], Maria Cristina Arias[3], Carolina Ivon Paris[4], Martin Bollazzi[5], Anurag Priyam[1], Federico López-Osorio[1], and Yannick Wurm[1,6*]

[1] School of Biological and Chemical Sciences, Queen Mary University of London, London, United Kingdom

[2] Leibniz Institute of Animal Biodiversity, Zoological Research Museum Alexander Koenig, Bonn, Germany

[3] Departamento de Genética e Biologia Evolutiva, Instituto de Biociências, Universidade de São Paulo, São Paulo, Brazil

[4] Departamento Ecología, Genética y Evolución, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Buenos Aires, Argentina

[5] Unidad de Entomología, Departamento de Protección Vegetal, Facultad de Agronomía, Universidad de la República, Montevideo, Uruguay

[6] Alan Turing Institute, London, United Kingdom

§ These authors contributed equally to this work

* Correspondence to e.stolle@leibniz-zfmk.de, rodrigopracana@gmail.com, y.wurm@qmul.ac.uk

Abstract

Introgression has been proposed as an essential source of trait adaptation. However, a key barrier to adaptive introgression is that recombination can break down combinations of alleles that underpin many traits. Supergene regions, where suppressed recombination leads to joint inheritance of alleles at many loci, could overcome this challenge. Here, we study the evolution of a large supergene region that determines a major social and ecological trait in *Solenopsis* fire ants: whether colonies have one queen or multiple queens. Using coalescent-based phylogenies built from the genomes of 365 haploid fire ant males, we show that the supergene variant responsible for multiple-queen colonies evolved in one species and repeatedly spread to other species through introgressive hybridization. This finding highlights how supergene architecture can enable a complex adaptive phenotype to recurrently permeate species boundaries.

Introduction

Alleles can be transferred between species by introgressive hybridization (*1*, *2*). Is it possible for a complex trait involving alleles at many loci to evolve in one species and then introgress into another? One of the main barriers to this type of introgression is recombination, which breaks up stretches of DNA, producing combinations of alleles from both the original and recipient species (*3*). Genomic inversions can impede recombination to create supergene variants that are inherited as units (*4*, *5*). These supergenes can spread among populations unhindered by recombination, but the extent to which they introgress across species boundaries remains unresolved.

In this study, we test competing hypotheses regarding the origin and evolutionary history of a supergene that determines whether colonies of the fire ant *Solenopsis invicta* have one queen or multiple queens. Several behavioral, morphological and physiological traits co-vary with queen number, with the single-queen social form thought to be better at dispersing, and the multiple-queen social form thought to be more competitive when population densities are high (*6*). The social dimorphism is controlled by the two variants (SB and Sb) of a 27.5 Mb supergene that includes more than 470 protein coding genes on “social chromosome” 16 (*7*–*9*). Single-queen colonies contain only the SB variant of the supergene, while colonies that include the Sb variant accept multiple queens. Interestingly, close relatives of *S. invicta* are also socially dimorphic and carry the same social supergene system (*8*, *10*, *11*). Although this supergene system could be a trans-species polymorphism that evolved in the common ancestor of these species (*8*, *10*, *11*), or even a set of polymorphisms that originated independently in each species, it is possible that its emergence (*12*) or spread (*8*, *11*, *13*) were instead mediated by introgressive hybridization.

Results

To understand the evolutionary history of the social supergene, we compare a phylogenetic tree representing the relationship among *Solenopsis* species with a phylogenetic tree representing the history of the supergene region (Fig. 1a to 1d illustrate potential scenarios). For this, we first

obtained whole-genome sequences of 368 samples of *Solenopsis*. These include 261 samples published by ourselves and others (*7*, *8*, *11*, *14–17*) and 107 that we additionally collected to expand taxonomic and geographical coverage (Table S1; Fig. S1). Most of the samples are haploid males (365 out of 368). Using these genomes, we constructed a phylogenetic tree from each of 1728 informative single-copy genes (*18*) (Table S2). We then applied a coalescent-based method (*19*) to reconstruct a species tree from 1631 single-copy genes mapping to chromosomes 1 to 15 and a supergene tree from 97 single-copy genes mapping to the supergene region of the social chromosome (Fig. 1e). In both trees, nodes critical for interpreting the histories of the species and the supergene are highly supported (*19*) (bootstrap support 98-100%, local posterior probabilities 0.89-1.00).

The species tree (left side of Fig. 1e) shows that males from the socially dimorphic species *S. invicta*, *S. macdonaghi* and *S. richteri* are organized into two sister clades. One clade includes all *S. richteri* males. The second clade includes *S. invicta* and *S. macdonaghi* males in a paraphyletic relationship, suggesting that this clade may represent a single species. We hereafter refer to it as *S. invicta/macdonaghi*. All "outgroup" samples, representing *S. geminata, S. pusillignis*, *S. saevissima*, *S. interrupta*, *S. megergates* and undescribed species *S. AdRX (11, 20)*, are organized into five clades branching off the base of the species tree.

Comparing the species tree with the supergene tree (right side of Fig. 1e) allows us to infer the evolutionary history of the supergene. For males carrying the SB variant of the supergene and all but five outgroup males (see below), the general topologies of the supergene tree and the species tree are congruent. However, all Sb males are grouped into a clade sister to the SB *S. invicta/macdonaghi* clade. We can thus reject the hypothesis that the supergene is a trans-species polymorphism that originated before the speciation of *S. richteri* and *S. invicta/macdonaghi*. Indeed, if that were the case, each species would form one monophyletic group within the SB clade and another within the Sb clade (scenario in Fig. 1b). We can also reject the hypothesis that there were multiple origins of the supergene, whereby the *S. richteri* Sb samples would form a distinct sister group to the *S. richteri* SB samples (scenario in Fig. 1c). Instead, the grouping of Sb males as sister to the *S. invicta/macdonaghi* SB clade shows that the Sb supergene variant originated once in the common ancestor of this clade, and then introgressed into other species (scenario in Fig. 1d). Comparing tree topologies also enables us to test and refute the hypothesis that the supergene arose through hybridization of two species with opposing orientations (*12*) of the supergene-equivalent region on chromosome 16. Instead, the differentiation between Sb and SB suggests that the inversions that suppress recombination between Sb and SB (*7*, *8*, *11*) are conserved across species.

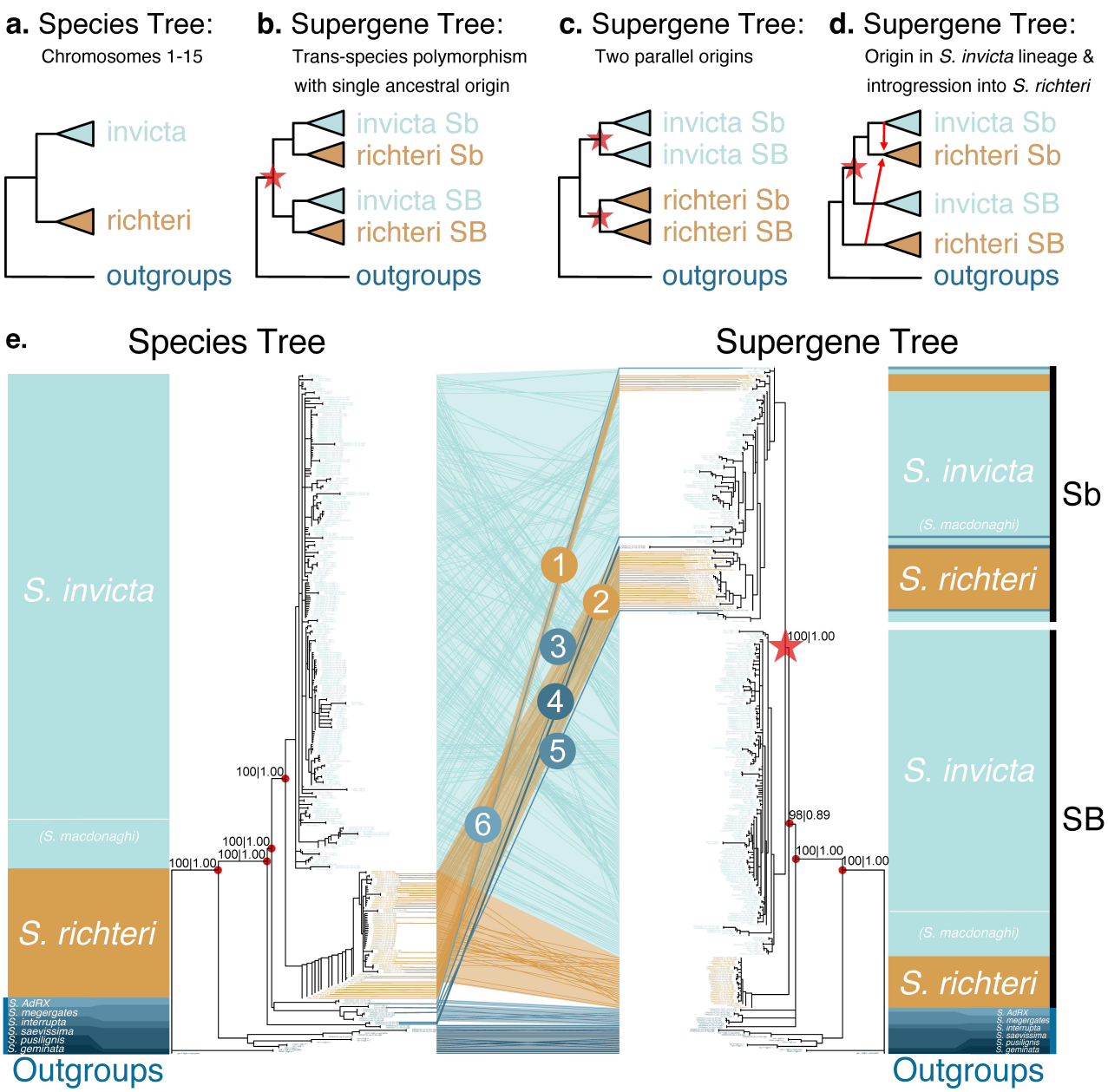

Fig. 1. Hypothetical and empirical species and supergene phylogenetic trees. **a-d,** A simplified species tree (**a**), and hypothetical scenarios for the evolutionary history of the supergene (**b-d**). **b**, SB and Sb supergene variants diverged (star) in the common ancestor of *S. invicta* and *S. richteri;* the supergene is thus a trans-species polymorphism. **c**, Sb evolved twice from SB, representing independent origins (stars) after the separation of the two species. **d**, Sb diverged in *S. invicta* and spread to *S. richteri* through introgression (arrows). **e**, Empirical coalescent-based trees of 368 *Solenopsis* samples based on 1631 single-gene trees from chromosomes 1-15 (left; species tree) and 97 single-gene trees from the supergene region of chromosome 16 (right; supergene tree). Branches shorter than 0.05 were collapsed into polytomies. A tanglegram (middle) indicates the relative positions of each sample in both trees. Circled numbers highlight patterns consistent with introgression of Sb from *S. invicta/macdonaghi* into other species. Support values (ASTRAL bootstrap support | local posterior probability) are provided for key nodes of speciation and supergene differentiation.

The Sb males from species other than *S. invicta/macdonaghi* are positioned into six Sb subclades within the supergene tree, suggesting that six independent introgression events occurred. Introgression of Sb from *S. invicta/macdonaghi* into *S. richteri* appears to have occurred twice given the presence of two groups of *S. richteri* Sb samples within the Sb subtree (numbers 1 and 2 in Fig. 1e). Strikingly, the other four cases of Sb introgression (numbered 3 to 6 in Fig. 1e) involve five Sb males positioned in early diverging outgroup clades of the species tree: one case involving *S. interrupta* (numbered 4; with two males), two cases involving *S. megergates* (numbered 3 and 5), and one case involving *S. AdRX* (numbered 6). All Sb haplotypes ultimately originate from the *S. invicta/macdonaghi* clade. However, in the supergene tree, one *S. megergates* Sb male (numbered 5) is adjacent to one of the *S. richteri* Sb clades (numbered 2), raising the possibility that Sb introgressed from one of these two species to the other secondarily. Also of note is that one of the introgressions into *S. megergates* (Fig. 1e, numbered 3) likely occurred very recently because its Sb haplotype is highly similar to those found in three *S. invicta*/*macdonaghi* Sb males we collected within 5.2 km (Supplementary Text S1; Fig. S2; Table S3). In contrast, both *S. richteri* Sb clades (clades 1 and 2 of Fig. 1e) include high nucleotide diversities and long branch lengths, suggesting that the introgressions into *S. richteri* are much more ancient (Fig. S3, Supplementary Text S2). Altogether our results suggest that introgression of the Sb supergene variant is a common mechanism for rapid adaptive evolution among fire ants and that this mechanism is not restricted to a particular Sb source haplotype or recipient species.

Discussion

Our results contradict previous suggestions that the Sb supergene variant evolved in a common ancestor of *S. invicta* and *S. richteri (11)* or within *S. richteri (13).* By incorporating twice as many samples as the most recent large-scale study (*11*), our work adds clades that had not yet been sampled. Furthermore, we focused on single-copy genes, which allowed us to avoid complex genealogies and potential genotyping artifacts caused by read mis-mapping of copy-number variants of genes and repetitive elements; such variation is particularly common in the degenerating Sb variant of the social chromosome (*8*, *15*, *17*, *21*). Finally, rather than implying that all sequences share one evolutionary history, our coalescent-based tree reconstruction approach should be robust to gene tree discordance due to processes including tree reconstruction errors, incomplete lineage sorting and gene conversion (*19*, *22*). Nevertheless, we performed a control analysis following the concatenated alignment similar approach to the first study (*11*), but after having excluded ambiguous regions of the genome (Supplementary Text S3; Supplementary Fig. S4). The resulting phylogenies are consistent with those from the coalescent-based tree reconstructions (Fig. 1e) and not the previous studies, supporting the introgression of Sb from *S. invicta*/*macdonaghi* into each of the four *Solenopsis* species mentioned above. The only notable difference was that the *S. richteri* Sb haplotypes are placed into one monophyletic group in the concatenated alignment tree and two groups in the coalescent-based tree (discussed in Supplementary Text S3).

Particular aspects of fire ant biology may have facilitated the spread of Sb through introgressive hybridization. First, flooding can transport entire floating fire ant colonies along the major South American river systems (*23*, *24*). Furthermore, mating flights can reach altitudes of 300 meters, with winds carrying young queens over great distances (*24*). These processes are likely to have created opportunities for admixture between lineages that were geographically distant. Once hybridization occurred, we hypothesize that the type of “green-beard” phenotype encoded by Sb facilitated the acceptance of hybrid Sb-carrying queens into established multiple-queen colonies of *S. invicta/macdonaghi*. This would have provided a crucial social buffer to any allelic incompatibilities of initial hybrids (see Supplementary Text S4). Over longer timescales, the ecological advantages of the multiple-queen social form likely allowed Sb to persist.

The strong support for the species tree for chromosomes 1 to 15 is in line with the absence of detectable contemporary gene flow between sympatric populations of *S. richteri* and *S. invicta/macdonaghi* in their native range (*25*). In contrast, the long-term retention of introgressed Sb supergene variants suggests that the adaptive benefits of Sb outweigh the costs of hybridization. Future studies of the evolution and function of the social chromosome supergene system will have to take the introgression of Sb into consideration. An important line of research will be to understand the dynamics of selection on *Sb/Sb* queens. Such queens are smaller and less fertile than *SB/Sb* or *SB/SB* queens (*26*, *27*) and are virtually absent in *S. invicta*. However, *Sb/Sb* queens are present in at least some *S. richteri* populations (*28*), thus enabling recombination between Sb haplotypes. Because of the differences in recombination regime and the bottlenecks associated with introgression, the Sb variants carried by each species may have different mutational loads or have different epistatic interactions with the rest of the genome.

Overall, our results clearly show that Sb evolved in an ancestor of *S. invicta/macdonaghi* that existed after divergence from the lineage leading to *S. richteri.* Through adaptive introgression, the Sb supergene variant spread to *S. richteri*, *S. megergates*, *S. interrupta* and *S. AdRX*. Supergene architecture can thus enable a complex social phenotype encoded by alleles at multiple genes to permeate species boundaries repeatedly. Our finding complements recently documented introgressions of smaller supergenes encoding simpler phenotypes (*12*, *29–31*) in other systems. The introgression of the Sb supergene variant shows how a species can rapidly gain the genetic basis for a complex phenotype, and how selection for changes in social organization can overcome the constraints of genomic architecture.

Materials and Methods

Collection and sequencing of fire ants from across their native range

We collected and flash-froze *Solenopsis* fire ants from 88 South American colonies (permit numbers: Uruguay Nº 001/2015, Argentina 007/15, 282/2016, 20.911.387, 025052053-115, 433/02101-0014449-4 and 25253/16, Brazil 14BR015531/DF). After phenol-chloroform DNA extraction on a pool of ten workers from each colony, we performed a Restriction Fragment Length Polymorphism (RFLP) assay (*10*) to determine whether both social chromosome variants are present (Table S1). From each of the 41 colonies that included only the SB variant, we retained one male. From 47 colonies that included both supergene variants, we either retained one SB and one Sb male (for 14 colonies), one Sb male (for 27 colonies), one SB male (for 2 colonies), two Sb males (for 3 colonies) or three Sb males (from 1 colony). We tentatively determined species identity through partial sequencing of the mitochondrial cytochrome oxidase I gene. Two colonies from outgroup species *S. pusillignis* and *S. saevissima* lacked males; thus, we instead retained a pool of 10 workers. From each of the resulting 107 samples, we prepared an individually barcoded Illumina TruSeq PCR-Free library with 350 bp insert size, which we sequenced on Illumina HiSeq, providing 10.3-fold average genome coverage per sample. We additionally obtained sequence reads from 260 *Solenopsis* males and 1 pool of workers (*S. geminata*) from seven other studies: Bioprojects PRJNA396161 (*8*, *17*), PRJNA542606 (*14*), PRJNA421367 (*11*), PRJNA450756 (*16*), PRJNA182127 (*17*), SRR621118 (*7*), SRX021921 (*15*). We performed optical duplicate filtering (minimum optical distance 2500), quality filtering and adapter trimming using clumpify (BBmap, v37.50) and skewer (v0.2.2, --mean-quality 20, --end-quality 15, minimum length of 80 for 100 bp reads and of 100 for 150 bp reads; removing degenerate reads). Table S1 provides further details on sample collection and sequencing.

Genotyping through mapping reads to the reference genome

We aligned the filtered reads for each sample to the *S. invicta* reference genome assembly Si_gnGA (*8*) and bait sequences including the *S. invicta* mitochondrion, *phi*X phage and *Wolbachia* sequences (NC_014672.1, NC_001422.1; AF243435.1, AF243436.1, CP001391.1, AM999887.1, GCA_003704235.1) using bwa-mem2 (v2.0pre2 with parameters -B 6 -E 2 -L25,25 -U 50 -T 50 -h 4,200 -a -V -Y -M). Sequence divergence within our dataset is sufficiently low for this approach as, for example, 98.95% of individual reads and 97.16% of pairs of reads from the most distant outgroup species used (*S. geminata*; sequence divergence from *S. invicta* ~3 million years ago (*8*)), map directly to the *S. invicta* reference genome (Table S1). For samples from PCR-based library preparation protocols, we marked duplicate reads with sambamba v0.7.1. We identified and subsequently excluded 23.05 Mb of regions which had coverage higher than median coverage plus 3 standard deviations of single-copy gene coverage across all samples because such high-coverage regions likely represent collapsed repeats (mosdepth v0.2.9, bedtools v2.27.1, bedops v2.4.30). We jointly called variants in all individuals using freebayes v1.2.0 (*32*), with parameters --

min-alternate-count 4 --min-alternate-fraction 0.4 --min-coverage 4 --ploidy 2 --use-best-n-alleles 8 --min-mapping-quality 40 --min-base-quality 30 --use-reference-allele. Although male ants are normally haploid, the genotyping process considered that samples are diploid so that we could identify and remove diploid males and apparently heterozygous sites that could be due to copy number variation or sequencing artefacts. After decomposing multi-nucleotide variants, we retained only Single Nucleotide Polymorphisms (SNPs) and excluded low-quality variants (Q<30), low coverage variants represented by either only forward or only reverse reads, and variants overlapping high-coverage genomic regions or simple sequence repeats (trf v4.09, bcftools v1.10.2). After this first round of genotyping to identify high confidence variants, we performed relaxed targeted genotyping using freebayes --haplotype-basis-alleles of each sample with the following parameters: --ploidy 2 --haplotype-length -1 --use-best-n-alleles 4 --min-mapping-quality 30 --min-base-quality 28 --min-alternate-fraction 0.35 --min-alternate-total 2 --min-coverage 2 --use-reference-allele. After decomposition, individual genotypes where coverage was greater than median plus 3 standard deviations of coverage on single-copy genes or with an allelic balance between 0.25 and 0.75 were set to missing because such sites are not expected in haploids and thus are untrustworthy. We created a haploid consensus from each of the three pools of workers by randomly selecting one allele at each heterozygous site. We had initially collected and sequenced 18 additional males (Table S1, not included in the sample numbers used above), but for nine of these, more than 25% of variant sites could not be genotyped, and the nine others had particularly high numbers of heterozygous sites indicating that they are likely diploid. All 18 males were removed from further analysis. We retained variant sites for which more than 75% of the remaining 368 samples were genotyped. We analyzed sequences of the Gp-9 marker (Table S1) and performed Principal Component Analysis (Fig. S5) to confirm supergene variants.

Coalescent-based phylogenetic inference

The *S. invicta* genome assembly (Annotation Release 100) contains exactly one copy of 5851 of the 5991 genes expected to be in a single copy across Hymenoptera (BUSCO v4.0.5 (*18*)). We made an initial consensus tree (not shown) based on genes mapped to chromosomes 1 to 15, rooted to *S. geminata*, which we used to label species of each sample according to Yan et al (*11*). Taxonomic assignment based on a mitochondrial phylogenetic tree was unreliable (Supplementary Text S5, Fig. S6). We retained the most informative 1881 out of 5851 genes for which the 342 males identified as *S. invicta, S. macdonaghi* and *S. richteri* (the expected ingroup species in the phylogenetic tree of *Solenopsis*) included an average of at least 10 single-nucleotide differences relative to the reference genome. We obtained a consensus sequence for each of these genes, including coding sequence and introns for each sample using bcftools consensus v1.10.2 from the VCF file of all genotypes. Modeltest-ng v0.1.6 (*33*) identified the best substitution models for intronic sequences and for coding sequences (Table S2). For each of the 1881 single-copy genes, we then used RAxML-NG v0.9.0 (*34*) with 50 random and 50 parsimony starting trees to build maximum-likelihood trees with 100 bootstraps, using separate partitions for introns and exons.

After removing the trees for the 153 genes that did not converge, we used ASTRAL-III v5.14.3 (*19*) to create two coalescent-based trees supported 100 bootstraps: one tree based on 97 single-copy genes in the supergene region of social chromosome 16, and one tree based on 1631 single-copy genes from chromosome 1 to 15. Trees were then rooted to the outgroup *S. geminata* and extremely short branches (branch length <0.05) were collapsed into polytomies. The topologies of these two trees were largely congruent with trees constructed from coding sequence alone, or constructed from the full set of 5851 genes with either a coalescent-based tree inference from gene trees (ASTRAL-III) or from concatenation (RAxML-NG, Fig. S4, Supplementary Text S2). ABBA-BABBA type tests to distinguish introgression from incomplete lineage sorting were unnecessary because the evolutionary scenario in which the supergene evolved before the speciation events leading to *S. richteri* and *S. invicta* (Fig. 1b) was rejected (Fig. 1e). To generate Fig. 1e, we used the R package dendextend and subsequently adjusted tree layouts, labels and background colors using Affinity Designer.

Acknowledgements

We thank S. Coelho, D. Pereira Nogueira Da Silva, N. Curi, N. Souza Araujo (Universidade de São Paulo), R. Jaffé (Instituto Tecnológico Vale), Pablo Álvarez Osorio (Buenos Aires), E. Boné (Universidad de Buenos Aires), Y. Guillij (Dir. de Usos Sust. de los Recursos Naturales, Entre Ríos), Dir. de Conservación de la Biodiversidad Argentina, C. Durrant, M. dos Reis, L. Rodrigues Santiago, C. Martinez-Ruiz, R. Nichols, and C. Eizaguirre (Queen Mary University of London), for their help with organisation, sampling, permits, preparation, sequencing or analysis, useful discussions, and comments on the manuscript.

Funding

European Union (FP7) Marie-Curie-Fellowship PIEF-GA-2013-623713 (ES, YW)

German Academic Exchange (DAAD) short term postdoctoral fellowship 570704 83 (ES)

Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) BEPE grant 2014/04943-0 (MCA)

Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) Ciência sem Fronteiras grant 248391/2013-5 (YW)

Biotechnology and Biological Sciences Research Council grant BB/K004204/1 (YW)

Biotechnology and Biological Sciences Research Council grant BB/T015683/1 (YW)

Biotechnology and Biological Sciences Research Council grant BB/S507556/1 (YW)

Natural Environment Research Council grant NE/L00626X/1 (YW)

Author contributions

Conceptualization: ES, YW

Sample collection, identification and processing: ES, CCC, MCA, CIP, MB, YW

DNA extraction and sequence library construction: ES, CCC

Initial data analysis: ES, RP, AP

Main data analyses: ES, RP, YW

Additional data analyses and visualization: ES, RP, MKP, GLH, AP, FLO

Supervision: YW

Writing — original draft: ES, RP, YW

Writing — review & editing: ES, RP, MKP, GLH, AP, FLO, YW

Competing interests

The authors declare that they have no competing interests.

Data and materials availability

All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Additional data and code to obtain the data are available as follows:

Code, analysis scripts and Newick tree files are available on at:

https://www.dropbox.com/s/73ps5h1v221olrr/GitHub_repository.zip?dl=0 (during review);

https://github.com/wurmlab/2020-06-introgression (long term).

DNA sequences are at SRA PRJNA685290 (Table S1 provides accession numbers):

https://dataview.ncbi.nlm.nih.gov/object/PRJNA685290?reviewer=pnqvkmbf7lljhvk0744oa67l8i

The VCF genotype matrix we generated is available at:

https://www.dropbox.com/s/xsfsp934hjg0j0w/gt.vcf.gz?dl=0

Supplementary Materials

Supplementary text, supplementary figures and the legends for the supplementary tables are given in an additional PDF. Supplementary tables were supplied as additional files.

References

1. E. Anderson, *Introgressive hybridization* (J. Wiley, New York, 1949).
2. R. G. Harrison, E. L. Larson, Hybridization, introgression, and the nature of species boundaries. *J. Hered.* **105 Suppl 1**, 795–809 (2014).
3. S. Yeaman, M. C. Whitlock, The genetic architecture of adaptation under migration-selection balance. *Evolution*. **65**, 1897–1911 (2011).
4. M. J. Thompson, C. D. Jiggins, Supergenes and their role in evolution. *Heredity*. **113**, 1–8 (2014).
5. M. Kirkpatrick, How and why chromosome inversions evolve. *PLoS Biol.* **8**, e1000501 (2010).
6. K. G. Ross, L. Keller, Ecology and evolution of social organisation: Insights from fire ants and other highly eusocial insects. *Annu. Rev. Ecol. Syst.* **26**, 631–656 (1995).

7. J. Wang, Y. Wurm, M. Nipitwattanaphon, O. Riba-Grognuz, Y.-C. Huang, D. Shoemaker, L. Keller, A Y-like social chromosome causes alternative colony organization in fire ants. *Nature*. **493**, 664–668 (2013).
8. E. Stolle, R. Pracana, P. Howard, C. I. Paris, S. J. Brown, C. Castillo-Carrillo, S. J. Rossiter, Y. Wurm, Degenerative expansion of a young supergene. *Mol. Biol. Evol*. **36**, 553–561 (2019).
9. L. Keller, K. G. Ross, Selfish genes: a green beard in the red fire ant. *Nature*. **394**, 573–575 (1998).
10. M. J. B. Krieger, K. G. Ross, Identification of a major gene regulating complex social behavior. *Science*. **295**, 328–332 (2002).
11. Z. Yan, S. H. Martin, D. Gotzek, S. V. Arsenault, P. Duchen, Q. Helleu, O. Riba-Grognuz, B. G. Hunt, N. Salamin, D. Shoemaker, K. G. Ross, L. Keller, Evolution of a supergene that regulates a trans-species social polymorphism. *Nat. Ecol. Evol.* **4**, 240–249 (2020).
12. P. Jay, A. Whibley, L. Frézal, M. Á. Rodríguez de Cara, R. W. Nowell, J. Mallet, K. K. Dasmahapatra, M. Joron, Supergene evolution triggered by the introgression of a chromosomal inversion. *Curr. Biol.* **28**, 1839–1845.e3 (2018).
13. P. Cohen, E. Privman, The social supergene dates back to the speciation time of two *Solenopsis* fire ant species. *Sci. Rep.* **10**, 11538 (2020).
14. C. Martinez-Ruiz, R. Pracana, E. Stolle, C. I. Paris, R. A. Nichols, Y. Wurm, Genomic architecture and evolutionary antagonism drive allelic expression bias in the social supergene of red fire ants. *eLife*. **9**, e55862 (2020).
15. Y. Wurm, J. Wang, O. Riba-Grognuz, M. Corona, S. Nygaard, B. G. Hunt, K. K. Ingram, L. Falquet, M. Nipitwattanaphon, D. Gotzek, M. B. Dijkstra, J. Oettler, F. Comtesse, C.-J. Shih, W.-J. Wu, C.-C. Yang, J. Thomas, E. Beaudoing, S. Pradervand, V. Flegel, E. D. Cook, R. Fabbretti, H. Stockinger, L. Long, W. G. Farmerie, J. Oakey, J. J. Boomsma, P. Pamilo, S. V. Yi, J. Heinze, M. A. D. Goodisman, L. Farinelli, K. Harshman, N. Hulo, L. Cerutti, I. Xenarios, D. Shoemaker, L. Keller, The genome of the fire ant *Solenopsis invicta*. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 5679–5684 (2011).
16. E. Privman, P. Cohen, A. B. Cohanim, O. Riba-Grognuz, D. Shoemaker, L. Keller, Positive selection on sociobiological traits in invasive fire ants. *Mol. Ecol.* **27**, 3116–3130 (2018).
17. R. Pracana, I. Levantis, C. Martínez-Ruiz, E. Stolle, A. Priyam, Y. Wurm, Fire ant social chromosomes: Differences in number, sequence and expression of odorant binding proteins. *Evol. Lett.* **1**, 199–210 (2017).
18. M. Seppey, M. Manni, E. M. Zdobnov, BUSCO: Assessing genome assembly and annotation completeness. *Methods Mol. Biol.* **1962**, 227–245 (2019).
19. C. Zhang, M. Rabiee, E. Sayyari, S. Mirarab, ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinform.* **19**, 153 (2018).
20. D. D. Shoemaker, M. E. Ahrens, K. G. Ross, Molecular phylogeny of fire ants of the *Solenopsis saevissima* species-group based on mtDNA sequences. *Mol. Phylogenet. Evol.* **38**, 200–215 (2006).

21. S. Fontana, N.-C. Chang, T. Chang, C.-C. Lee, V.-D. Dang, J. Wang, The fire ant social supergene is characterized by extensive gene and transposable element copy number variation. *Mol. Ecol.* **29**, 105–120 (2020).
22. L. S. Kubatko, J. H. Degnan, Inconsistency of phylogenetic estimates from concatenated data under coalescence. *Syst. Biol.* **56**, 17–24 (2007).
23. B. J. Adams, L. M. Hooper-Bùi, R. M. Strecker, D. M. O'Brien, Raft formation by the red imported fire ant, *Solenopsis invicta*. *J. Insect Sci.* **11**, 171 (2011).
24. W. R. Tschinkel, *The Fire Ants* (Harvard Univ. Press, 2006).
25. P. Cohen, E. Privman, Speciation and hybridization in invasive fire ants. *BMC Evol. Biol.* **19**, 111 (2019).
26. C. J. DeHeer, M. A. D. Goodisman, K. G. Ross, Queen dispersal strategies in the multiple-queen form of the fire ant *Solenopsis invicta*. *Am. Nat.* **153**, 660–675 (1999).
27. C. J. DeHeer, A comparison of the colony-founding potential of queens from single- and multiple-queen colonies of the fire ant *Solenopsis invicta*. *Anim. Behav.* **64**, 655–661 (2002).
28. B. L. Hallar, M. J. B. Krieger, K. G. Ross, Potential cause of lethality of an allele implicated in social evolution in fire ants. *Genetica*. **131**, 69–79 (2007).
29. N. B. Edelman, P. B. Frandsen, M. Miyagi, B. Clavijo, J. Davey, R. B. Dikow, G. García-Accinelli, S. M. Van Belleghem, N. Patterson, D. E. Neafsey, R. Challis, S. Kumar, G. R. P. Moreira, C. Salazar, M. Chouteau, B. A. Counterman, R. Papa, M. Blaxter, R. D. Reed, K. K. Dasmahapatra, M. Kronforst, M. Joron, C. D. Jiggins, W. O. McMillan, F. Di Palma, A. J. Blumberg, J. Wakeley, D. Jaffe, J. Mallet, Genomic architecture and introgression shape a butterfly radiation. *Science*. **366**, 594–599 (2019).
30. L. C. Norris, B. J. Main, Y. Lee, T. C. Collier, A. Fofana, A. J. Cornel, G. C. Lanzaro, Adaptive introgression in an African malaria mosquito coincident with the increased usage of insecticide-treated bed nets. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 815–820 (2015).
31. C. Brion, C. Caradec, D. Pflieger, A. Friedrich, J. Schacherer, Pervasive phenotypic impact of a large nonrecombining introgressed region in yeast. *Mol. Biol. Evol.* **37**, 2520–2530 (2020).
32. E. Garrison, G. Marth, Haplotype-based variant detection from short-read sequencing. http://arxiv.org/abs/1207.3907 [q-bio.GN] (2012).
33. D. Darriba, D. Posada, A. M. Kozlov, A. Stamatakis, B. Morel, T. Flouri, ModelTest-NG: a new and scalable tool for the selection of DNA and protein evolutionary models. *Mol. Biol. Evol.* **37**, 291–294 (2020).
34. A. M. Kozlov, D. Darriba, T. Flouri, B. Morel, A. Stamatakis, RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics*. **35**, 4453–4455 (2019).
35. J. Oksanen, F. G. Blanchet, M. Friendly, R. Kindt, P. Legendre, D. McGlinn, P. R. Minchin, R. B. O'Hara, G. L. Simpson, P. Solymos, M. H. H. Stevens, E. Szoecs, H. Wagner, *vegan: Community Ecology Package* (2020; https://CRAN.R-project.org/package=vegan).
36. B. Pfeifer, U. Wittelsbürger, S. E. Ramos-Onsins, M. J. Lercher, PopGenome: an efficient Swiss army knife for population genomic analyses in R. *Mol. Biol. Evol.* **31**, 1929–1936 (2014).

37. A. R. Quinlan, I. M. Hall, BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*. **26**, 841–842 (2010).
38. E. M. Ortiz, vcf2phylip v2.0: convert a VCF matrix into several matrix formats for phylogenetic analysis (2019), doi:10.5281/zenodo.2540861.
39. J. W. Brown, J. F. Walker, S. A. Smith, Phyx: phylogenetic tools for unix. *Bioinformatics*. **33**, 1886–1888 (2017).
40. S. Lutteropp, A. M. Kozlov, A. Stamatakis, A fast and memory-efficient implementation of the transfer bootstrap. *Bioinformatics*. **36**, 2280–2281 (2020).
41. H. Li, A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics*. **27**, 2987–2993 (2011).
42. T. Nakamura, K. D. Yamada, K. Tomii, K. Katoh, Parallelization of MAFFT for large-scale multiple sequence alignments. *Bioinformatics*. **34**, 2490–2492 (2018).
43. D. Gotzek, J. Clarke, D. Shoemaker, Mitochondrial genome evolution in fire ants (Hymenoptera: Formicidae). *BMC Evol. Biol.* **10**, 300 (2010).