MOLECULAR ECOLOGY

Molecular Ecology (2017) **26**, 1108–1117 doi: 10.1111/mec.13899

Convergent evolution of social hybridogenesis in *Messor* harvester ants

JONATHAN ROMIGUIER, AXEL FOURNIER, SZE HUEI YEK and LAURENT KELLER

Department of Ecology and Evolution, University of Lausanne, 1015 Lausanne, Switzerland

Abstract

Sexual reproduction generally requires no more than two partners. Here, we show convergent evolution of social hybridogenesis, a reproductive system requiring three reproductive partners in harvester ants. In this unorthodox reproductive system, two distinct genetic lineages live in sympatry and queens have to mate with males of their own lineage to produce queens along with males of the alternative lineage to produce workers. Using a large transcriptomic data set of nine species, we show that social hybridogenesis evolved at least three times independently in the genus *Messor*. Moreover, a study of 13 populations of *Messor barbarus* revealed that this mode of reproduction is fixed in the whole range of this ecologically dominant species. Finally, we show that workers can produce males carrying genes of the two genetic lineages, raising the possibility of rare gene flow between lineages contributing to the long-term maintenance of pairs of interdependent lineages. These results emphasize the evolutionary importance of social hybridogenesis, a major transition possibly linked to the peculiar ecology of harvester ants.

Keywords: behavior/social evolution, bioinformatics/phylogenomics, evolution of sex, genomics/proteomics

Received 27 July 2016; revision received 26 September 2016; accepted 12 October 2016

Introduction

Several organisms have evolved unusual modes of reproduction, comprising both components of asexual and sexual reproduction. In some species of social insects, by example, queens are produced by parthenogenesis while workers are produced by normal sexual reproduction (Pearcy *et al.* 2004, 2011; Fournier *et al.* 2005, 2016; Kobayashi *et al.* 2008; Matsuura *et al.* 2009). Using alternative modes of reproduction for the queen and worker castes, queens can increase the transmission rate of their genes to their reproductive female offspring while maintaining genetic diversity and social cohesion in the worker force (Smith *et al.* 2008). The ant *Cataglyphis hispanica* has evolved a similar mode of reproduction whereby parthenogenesis also produces new queens but workers are the product of hybrid matings between two distinct genetic lineages that coexist within a single population (Leniaud *et al.* 2012). This reproductive system relying on two dependent lineages to produce hybrid workers has been called social hybridogenesis and was first described in *Pogonomyrmex* harvester ants (Helms Cahan *et al.* 2002; Julian *et al.* 2002; Volny & Gordon 2002; Helms Cahan & Keller 2003). Contrary to the *Cataglyphis* case, social hybridogenesis in *Pogonomyrmex* does not involve parthenogenesis and the ensuing short-term advantage in terms of higher transfer of genes to the reproductive females. In some populations of *Pogonomyrmex*, two distinct genetic lineages coexist and queens mate multiply with males of their own and the alternate lineage; offspring produced from same-lineage matings always develop into queens, whereas interlineage hybrids develop into workers (Fig. 1A).

A remarkable feature of the *Pogonomyrmex* social hybridogenesis system is that the production of new fertile colonies requires the involvement of at least three mating partners – one female and two males, one from each lineage. Queens that only mate with males of different lineage fail to produce new queens, while queens that only mate with males of their own lineage fail to produce workers and the colony is therefore not viable. Understanding the origin and

Correspondence: Jonathan Romiguier and Laurent Keller, Fax: +41(0)21 692 41 73; Emails: jonathan.romiguier@gmail.com; laurent.keller@gmail.com

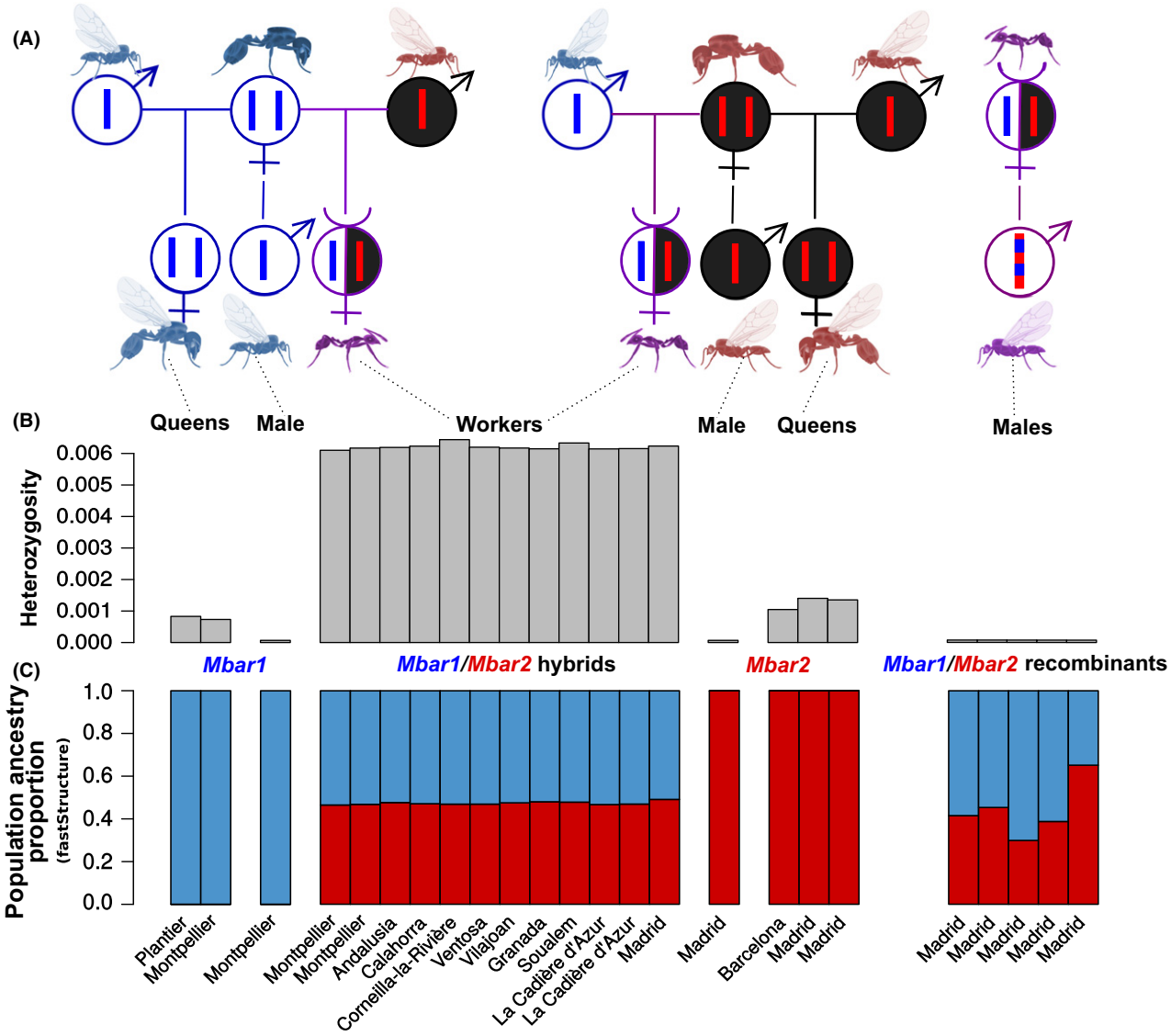

Fig. 1 Social hybridogenesis in *Messor barbarus*. (A) Social hybridogenesis reproductive system where two genetic lineages (blue and red) coexist within each population. Queens mate with males from both lineages. Pure lineage matings produce new queens while interlineage matings produce workers. (B) Workers of *M. barbarus* have higher genomewide heterozygosity levels (based on 90 951 SNPs) than queens (all pairwise comparisons significant, $P < 0.0001$, Wilcoxon tests). (C) Population ancestry proportions estimated by *fastStructure* in 24 individuals. Each bar corresponds to individuals from (A) and (B). Bars are named according to the population of origin of the individuals. Queens are pure *Mbar1* (blue) or *Mbar2* (red) lineages; workers are *Mbar1*/*Mbar2* hybrids. Males laid by queens are pure lineages, and males laid by workers are *Mbar1*/*Mbar2* recombinants.

maintenance of this unique system requiring three different mating partners has attracted considerable interest (Anderson *et al.* 2006, 2008; Linksvayer *et al.* 2006; Schwander *et al.* 2006, 2007a, b; Volny *et al.* 2006; Yamauchi & Yamamura 2006; Curry *et al.* 2009; Sirviö *et al.* 2010; Suni & Gordon 2010; Gordon *et al.* 2013; Mott *et al.* 2015). However, no new cases have been found since its discovery (Julian *et al.* 2002). So far, this reproductive system is restricted to populations thought to result from ancestral hybridization between *P. rugosus* and *P. barbatus*, two species with a standard mode of reproduction (Schwander *et al.* 2007a). It is unclear whether such systems are oddities resulting from rare historical events or whether they can be favoured under certain ecological conditions. Another important issue inherent to social hybridogenesis is the long-term maintenance of interdependent lineages (Yamauchi & Yamamura 2006). As interlineage matings give rise only to workers, interlineage gene flow is suppressed, possibly leading to genetic incompatibilities between the increasingly divergent lineages (Anderson *et al.* 2008).

In this study, we show that the evolution of social hybridogenesis is not as rare as initially thought. Our genetic analyses reveal that three species of the genus *Messor* have such a mode of reproduction. Moreover, we show that this mode of reproduction evolved at least four times independently in harvester ants, suggesting that the peculiar ecology of these seed-eating ants might facilitate such a transition. Finally, we demonstrate that workers can reproduce by parthenogenesis and produce males with a chimeric composition of the two genetic lineages, paving the way to long-term maintenance of social hybridogenesis.

Methods

Sampling for transcriptome sequencing

To uncover the reproductive system of *Messor* species, we sampled 52 individuals across 29 populations from nine species (*M. barbarus*, *M. capitatus*, *M. structor*, *M. cf decipiens*, *M. aciculatus*, *M. cf hellenius*, *M. minor hesperius*, *M. wasmanni* and *M. ebeninus*) for RNA sequencing (see Table S1, Supporting information for details). To determine whether queens are produced parthenogenetically or sexually in *M. barbarus*, we sampled the mother queen, one of her sons and two daughters (a new queen and a worker) from a monogyne colony raised in laboratory conditions since 15 years (colony noted *Madrid1* in Table S1, Supporting information). To check whether the genomes of the dependent lineages (hereafter named *Mbar1* and *Mbar2*) can recombine through worker parthenogenesis, we sequenced five males laid by orphan workers from the same *M. barbarus* laboratory colony. For all species except *M. barbarus*, we sampled at least one mated queen after a mating flight and waited for the emergence of the first workers in the laboratory to sequence a mother queen and one of her daughter worker.

Transcriptome sequencing and assembly

The whole body of each individual was flash-frozen before RNA extraction. Total RNA was extracted using specific protocols for ants (Gayral *et al.* 2011). Complementary libraries were prepared using an Illumina TrueSeq preparation kit. We sequenced these libraries on a HiSeq 2000 (Illumina) to produce 100-base pairs (bp) paired-end reads. We used *Trimmomatic* to remove adapters and reads with length less than 60 bp and average quality less than 30 (Bolger *et al.* 2014). *De novo* transcriptome assemblies of each species were performed using a combination of *Abyss* and *Cap3* to treat reads of all individuals of a given species, following previously described methods (Romiguier *et al.* 2014a).

Computing population genetic statistics

We followed the pipeline of Romiguier *et al.* (2014a). Illumina reads of each individual were mapped to the *de novo* transcriptome assembly of its corresponding species using the BWA program (Li & Durbin 2010), and we discarded contigs with a per-individual average coverage below ×2.5. We used the TRINITY package to predict open reading frames (ORFs). We discarded contigs carrying ORFs less than 200 bp. In contigs with ORFs longer than 200 bp, we deleted 5′ and 3′ flanking noncoding sequences, thus producing predicted coding sequences that are hereafter referred to as loci. We used the READ2SNP program and followed previously used methods to call genotypes (Romiguier *et al.* 2014a). To infer the reproductive system of *Messor* species, we calculated *Fis* and heterozygosity values for each locus and each species/individual using the tool DNDSPINPIS (http://kimura.univ-montp2.fr/PopPhyl/index.php?section=tools). In case of social hybridogenesis, we expect an excess of heterozygosity (high *Fis*) in workers which are hybrids between lineages compared to queens which are of pure lineage. *Fis* calculation was corrected for small sample size using the Weir–Cockerham correction (Weir & Clark Cockerham 1984). Confidence intervals were obtained by bootstrapping loci. We compared the average locus heterozygosity and the average locus *Fis* among individuals using R and nonparametric Wilcoxon tests. Finally, we conducted population structure analyses using *fastStructure* (Raj *et al.* 2014) with default parameters and a number of populations of $K = 2$.

Sampling and DNA sequencing

To quantify the dependent lineage distribution (hereafter called *Mbar1* and *Mbar2*) across colonies of two populations (Montpellier and Lunel, France), we sampled six colonies in each population during a mating flight in September 2013. We performed Sanger sequencing on a nuclear gene (*hsc70-4*) identified from transcriptomic data as carrying six fixed mutations between the *Mbar1* and *Mbar2* lineages. We sequenced the *hsc70*-4 nuclear gene in 56 males, 37 alate queens and 52 workers along with the *cox1* mitochondrial gene in 27 males and 21 alate queens. We extracted DNA using the 'BS96 DNA Tissue' protocol of the *BioSprint 96* extractor robot. We purified PCR products with the 'Wizzard Genomic DNA Purification' Promega kit.

Phylogenetic analyses

We used *OrthoMCL* (Li 2003) to retrieve 2070 1-1 ortholog genes among the nine *de novo* assemblies of *Messor*

species + the official gene set (version 1.2) of *Pogonomyrmex barbatus*. To build a *Messor* phylogeny, we divided species featuring social hybridogenesis (*M. barbarus*, *M. structor* and *M. ebeninus*) in two dependent lineages defined from the *fastStructure* analyses (hereafter called *Mbar1/Mbar2*, *Mstr1/Mstr2* and *Mebe1/Mebe2*). For this, we built consensus sequences of queens grouped according to their respective lineages (2 *Mbar1*, 3 *Mbar2*, 1 *Mstr1*, 3 *Mstr2* and 1 *Mebe1*) or species (1 *M. cf decipiens*, 2 *M. capitatus*, 1 *M. wasmanni*, 1 *M. minor*, 2 *M. cf hellenius* and 1 *M. aciculatus*). We used the paternal alleles of the *M. ebeninus* hybrid worker (daughter of the *Mebe1* queen) as *Mebe2* sequences. The resulting 2070 1-1 ortholog genes were aligned using MACSE (Ranwez *et al.* 2011) and then concatenated in a supermatrix that was cleaned using the automated method included in *trimal* (Capella-Gutierrez *et al.* 2009). All phylogenetic reconstructions were performed using RAXML (GTR + GAMMA model, 500 bootstrap replicates, *P. barbatus* used as the outgroup; Ranwez *et al.* 2011; Stamatakis 2014).

Results

Messor barbarus workers are hybrids between two distinct genetic lineages

The analysis of 19 *M. barbarus* transcriptomes (two males, five queens and 12 workers) from 12 different populations across the Mediterranean Basin (Table S1, Supporting information) revealed that workers are produced by social hybridogenesis. The workers exhibited extremely high levels of heterozygosity (*Fis* = −0.666 [−0.671, −0.662], 12 532 genes) in stark contrast to queens which exhibited a strong excess of homozygosity (*Fis* = 0.746 [0.738, 0.753], 5597 genes; Fig. 1B). An analysis of 126 967 SNPs with *fastStructure* (Raj *et al.* 2014) revealed that two queens and one male belonged to the same lineage (hereafter called *Mbar1*, synonymous nucleotide diversity of 0.0023), three queens and one male to another lineage (hereafter called *Mbar2*, synonymous nucleotide diversity of 0.0039) while all workers were inferred to be a 50/50 mix of these two genetic groups (Fig. 1C). This unusual pattern suggests social hybridogenesis, where queens belong to two distinct genetic lineages and workers are hybrids of both lineages.

To quantify the *Mbar1/Mbar2* lineage distribution across colonies of two populations (Montpellier and Lunel, France), we sampled males (n = 56), queens (n = 37) and workers (n = 52) from 12 colonies (six from each population) and sequenced a nuclear gene (*hsc70-4*) identified from transcriptomic data as carrying six fixed mutations between the *Mbar1* and *Mbar2* lineages. All reproductive individuals (23 males and 18 queens) from six colonies (three in each population) carried the six fixed mutations corresponding to the *Mbar1* lineage, while all reproductive individuals (33 males and 19 queens) from the six remaining colonies (three in each population) carried the six fixed mutations corresponding to the *Mbar2* lineage. Remarkably, all workers (n = 52) were *Mbar1/Mbar2* hybrids, with the six *hsc70-4* positions that carry *Mbar1/Mbar2* substitutions being heterozygous for every worker (Table S2, Supporting information). We also sequenced a mitochondrial gene (*cox1*) in a subset of reproductive individuals of the 12 colonies (27 of the 57 males and 21 of the 37 queens) to build a phylogenetic tree. This analysis confirmed the presence of two clear monophyletic clades (bootstrap of 100), grouping the individuals according to their *Mbar1/Mbar2* lineage assignation (Fig. 2).

Sexual origin of Messor barbarus queens and worker reproduction

To determine whether queens are parthenogenetically or sexually produced, we sequenced the whole transcriptome of a mother queen and a daughter queen. While the mother and daughter queens had similar heterozygosity levels (0.00140 and 0.00135, respectively), 50.1% of the heterozygous positions of the daughter (368/734 positions) were homozygous in the mother, indicating that the daughter was sexually produced. Sexual production of queens indicates that social hybridogenesis uncovered in *M. barbarus* is similar to the system of social hybridogenesis described in *Pogonomyrmex* (Fig. 1A; Volny & Gordon 2002).

To investigate whether orphan workers can lay eggs, we isolated in the laboratory 100 *M. barbarus* workers from a *Mbar2* colony after removing all the brood from the nest. We observed 16 eggs that gave rise to 16 males, hatching 25 days after the isolation. To check whether the genomes of dependent lineages *Mbar1* and *Mbar2* can recombine through worker parthenogenesis, we sequenced the transcriptome of five of the 16 males. These five males were all haploid and had a mix of alleles coming from the *Mbar1* and *Mbar2* lineages (*fastStructure* revealed that these males had 42%, 45%, 30%, 39% and 65% of their alleles corresponding to the *Mbar2* lineage, that is the lineage of the queen in their colony, Fig. 1C).

At least three independent origins of social hybridogenesis in Messor

To investigate whether other cases of social hybridogenesis occur in the genus *Messor*, we sequenced the transcriptome of a mother queen and a daughter worker in

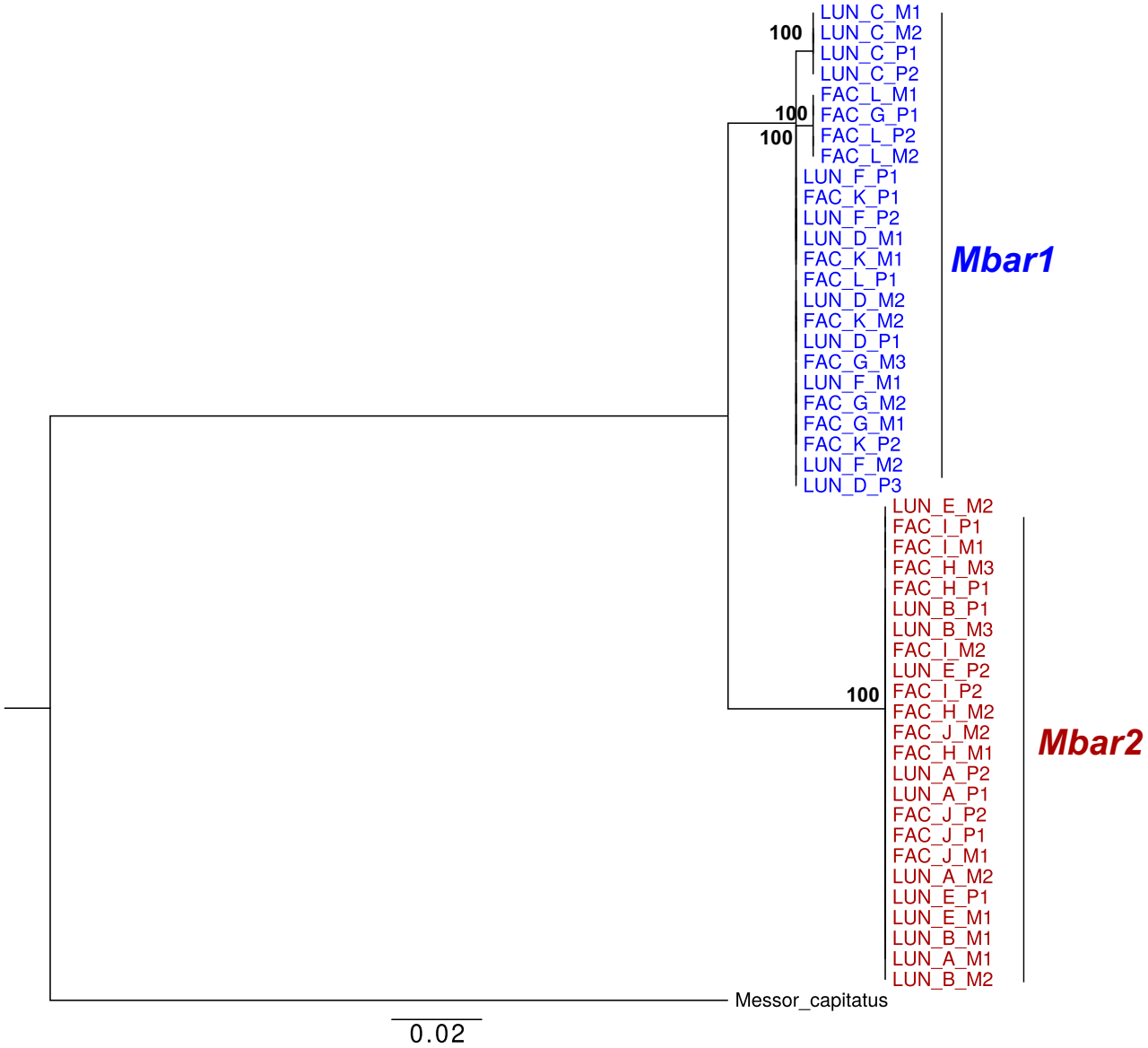

Fig. 2 Phylogenetic tree of 48 *M. barbarus* males and queens based on the *cox1* mitochondrial gene. Tip names correspond to the ID field in Table S2 (Supporting information). First three letters correspond to the population of origin (either FAC or LUN), 4th letter corresponds to the ID of one of the 12 colonies (from A to L) and the 5th letter corresponds to the caste (P for virgin queen and M for male). Individuals assigned to the *Mbar1* lineage are in blue, and individuals assigned to the *Mbar2* lineage are in red (assignment based on *hsc70-4* mutations, see Table S2, Supporting information). [Colour figure can be viewed at wileyonlinelibrary.com]

nine *Messor* species (including *M. barbarus*). Of the nine species, three (*M. barbarus*, *M. structor* and *M. ebeninus*) exhibited significant differences in the heterozygosity levels of the queen and the worker (Wilcoxon test, $P < 0.0001$ in the three cases), suggesting hybrid workers and pure lineage queens as typically seen in social hybridogenesis. The heterozygosity level of the queen and the worker was similar in the six remaining species (Wilcoxon test, NS in the six species, Fig. 3). To further explore the case of *M. structor*, we sequenced four additional queens and four additional workers from six different populations. Similar to *M. barbarus*, *fastStructure* revealed that queens belonged to two distinct genetic lineages (*Mstr1* with one queen, *Mstr2* with three queens) while workers were all hybrids between these two lineages (Fig. S1, Supporting information).

To determine the number of origins of social hybridogenesis, we built a molecular phylogeny of the *Messor* genus separating each hybridogenetic species (*M. barbarus*, *M. structor* and *M. ebeninus*) in two dependent lineages (respectively, *Mbar1*/*Mbar2*, *Mstr1*/*Mstr2* and *Mebe1*/*Mebe2*). Each pair of dependent

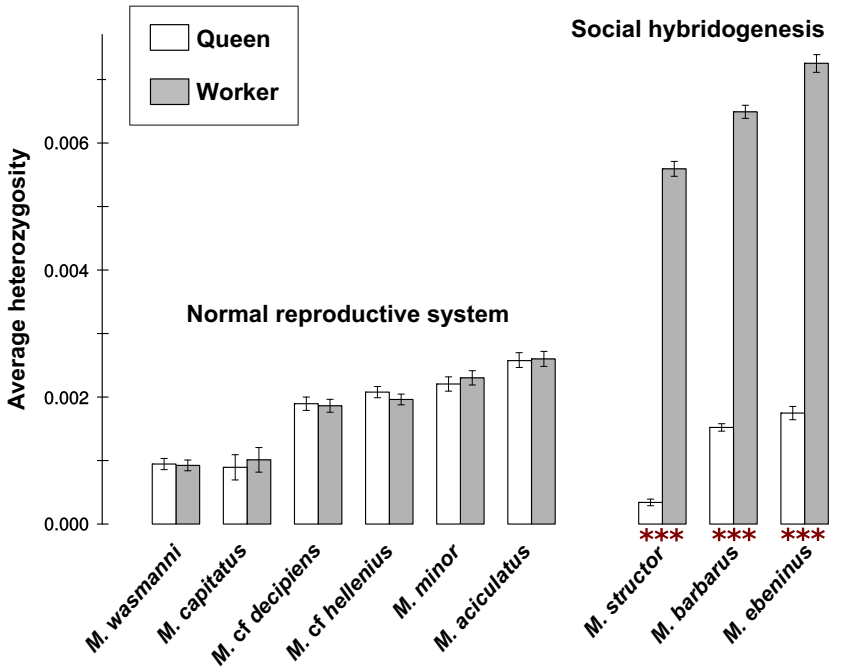

Fig. 3 Average locus heterozygosity comparisons between a mother queen and her daughter worker in nine *Messor* species. Three species have a significant heterozygosity difference characteristic of social hybridogenesis. [Colour figure can be viewed at wileyonlinelibrary.com]

lineages belonged to three distinct clades (Fig. 4), which suggests that social hybridogenesis (i.e. the occurrence of dependent genetic lineages) evolved three times independently. The two lineages of *M. barbarus* (*Mbar1* and *Mbar2*) clustered together (bootstrap support of 100). In contrast, one of the *M. structor* lineages (*Mstr2)* clustered with *M. cf hellenius* (bootstrap support of 100) rather than the other *M. structor* lineage (*Mstr1*). Similarly, the *M. ebeninus* lineage *Mebe2* clustered with *M. wasmanni* (bootstrap support of 100) instead of *Mebe1*.

Because social hybridogenesis results in a strict system of genetic caste determination (Schwander *et al.* 2010), we searched for genes where one given nucleotide was heterozygous in all hybrid workers (n = 17) of the three species with social hybridogenesis (*M. barbarus*, *M. structor* and *M. ebeninus*) and homozygous in queens of each of the six lineages (n = 11). There were only two such genomic regions, one for which the substitution induced a protein change. This was in *catalase* (position 1021 in the alignment available as Supplementary Material), a gene known to affect honeybee caste differentiation during early larval development (Cameron *et al.* 2013). The codon position varies between TTC and CTC across the phylogeny of *Messor*, which code, respectively, for the Phenylalanine and Leucine amino acids (Fig. 5). Interestingly, the *Messor* species with normal reproductive system (*M. cf decipiens*, *M. capitatus*, *M. wasmanni*, *M. minor*, *M. cf hellenius* and *M. aciculatus*) were always fixed for one of the two variants with all queens (n = 8) and workers (n = 10) being homozygous (Fig. 5).

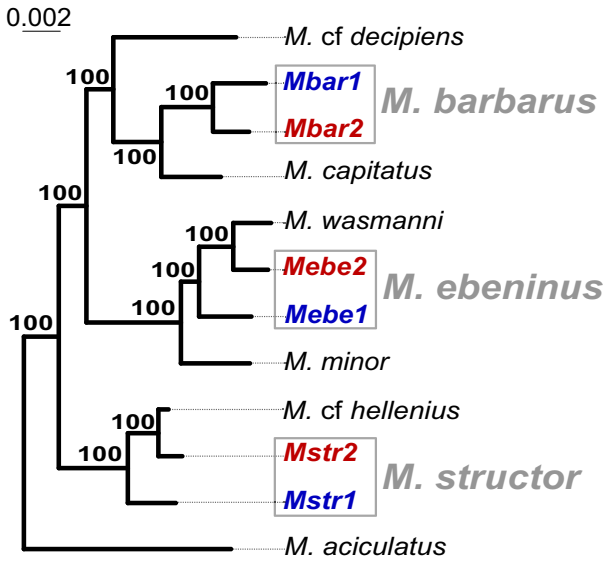

Fig. 4 Phylogenetic relationships among six *Messor* species and three pairs of dependent lineages (*Mbar1*/*Mbar2*, *Mstr1*/*Mstr2* and *Mebe1*/*Mebe2*) coexisting in species featuring social hybridogenesis (respectively, *M. barbarus*, *M. structor* and *M. ebeninus*). The phylogenetic tree is built from a supermatrix of 2070 orthologs genes, with sequences obtained from a consensus of all queens of a given lineage (for *M. barbarus*, *M. ebeninus* and *M. structor*) or all queens of a given species for *M. cf decipiens*, *M. capitatus*, *M. wasmanni*, *M. minor*, *M. cf hellenius* and *M. aciculatus*. [Colour figure can be viewed at wileyonlinelibrary.com]

Discussion

This study demonstrates that the reproductive system uncovered in *Pogonomyrmex* is not merely an

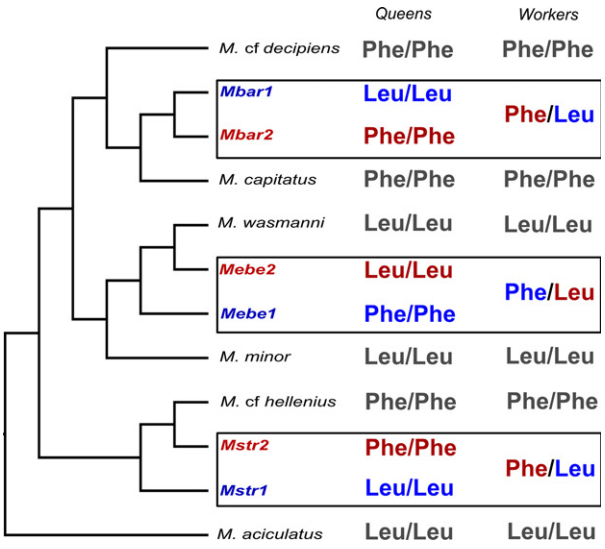

Fig. 5 Amino acid identity at position 1021 of the *catalase* gene across the *Messor* phylogeny. [Colour figure can be viewed at wileyonlinelibrary.com]

exceptional outcome of the complex and unusual hybridization history between *P. rugosus* and *P. barbatus*. A similar reproductive system has evolved independently in at least three species of the genus *Messor*, even if it is still unclear whether social hybridogenesis in two of these species (*M. ebeninus* and *M. structor*) is associated to parthenogenesis (as in *Cataglyphis*) or multiple matings with males from distinct genetic lineages (as in *Pogonomyrmex* and *M. barbarus*). Our data on *M. barbarus* reveal that this mode of reproduction can be the rule in a well-characterized morphospecies. The European harvester ant *M. barbarus* is one of the most dominant species of the western Mediterranean Basin, particularly in the Iberian Peninsula and Mediterranean France. The 13 populations that we investigated cover the whole distribution range of the species (Southern France, Iberian Peninsula and North Africa; Lebas *et al.* 2016) and they were all characterized by the presence of workers being hybrids between the two same distinct genetic lineages. This suggests the long-term persistence of this reproductive mode, possibly due to evolutionary advantages such as hybrid vigour of workers (Houle 1989; Cahan *et al.* 2010).

A surprising finding of this study is that three of the nine *Messor* species investigated features social hybridogenesis. Moreover, our analyses revealed that this mode of reproduction evolved independently in the three species, suggesting that the *Messor* genus is prone for such transitions. Interestingly, *Messor* and *Pogonomyrmex* have important convergences in their ecology. Species of these two genera live in dry climates and are obligate granivores (i.e. seed predators), which is unusual in ants. We suggest two explanations for the multiple origins of social hybridogenesis in ants with such an ecology.

The first is associated with the fact that species of these two genera live in environments favouring large synchronous mating flights. To successfully initiate a colony and produce reproductive females, queens have to mate with males of the two distinct genetic lineages. The likelihood of encountering males coming from colonies headed by queens of both lineages increases when many colonies participate in massive and synchronous mating flights instead of small mating flights spread over a long period of time. Such massive and synchronous mating flights are common after summer rainfalls (when the soil is moistened and new nests easy to excavate), which is the most typical environmental cue triggering mating flights in ants (Hölldobler & Wilson 1990). Both *Pogonomyrmex* and *Messor* ants occur mainly in Mediterranean climates where summer rainfalls are concentrated over a few days. Consequently, *Pogonomyrmex* species and *M. barbarus* are known to have highly synchronous mating flights (Hölldobler 1976; Markl *et al.* 1977; Davidson 1982; Gómez & Abril 2012), increasing the opportunity for reproductive females of these species to encounter males of both lineages during mating flights.

An alternative explanation considers the granivorous diet of *Messor* and *Pogonomyrmex*. In many ants, caste determination is influenced both by genetic and environmental factors (Schwander *et al.* 2010). Diet and more particularly larval nutrition seem to be key environmental factors for caste determination (Haydak 1943; Kapheim *et al.* 2011). In an ant species with a mixed seed/insect diet, larvae consuming higher proportions of insect/proteins compared to seeds tend to develop into queens (Smith & Suarez 2010). Strict granivory, as is typical for most harvester ants, might decrease the ability of workers to efficiently control the queen/worker ratio through larval feeding, thus paving the way for the evolution of genetic factors for caste determination and social hybridogenesis. Interestingly, another type of social hybridogenesis was found in hybrid zones between two fire ants, whereby *Solenopsis xyloni*/*Solenopsis geminata* hybrids develop into workers while pure *S. xyloni* offspring develop into queens (Cahan & Bradleigh Vinson 2003). While these fire ants do not consume seeds exclusively, both are classified as granivorous and seasonally store large quantities of seeds (Valone & Michael 2005; Tschinkel 2006), further supporting the view that granivory might be conducive of social hybridogenesis.

The exact evolutionary history of *Messor*-dependent lineages is difficult to assess. The two lineages of a pair may have split from a single ancestral gene pool or resulted from hybridization between two differentiated

species, as discussed in several studies on the *Pogonomyrmex* case (Anderson *et al.* 2006; Schwander *et al.* 2007a). In *M. barbarus*, the two dependent lineages are sister taxa (Fig. 4), which suggests that they derived from a single ancestral gene pool or from a closely related species that is extinct or not included in our analyses. In both *M. ebeninus* and *M. structor*, one of the dependent lineages is the closest relative of a species with a normal reproductive system. This could be explained by two hypothesis: (i) the closely related species with a normal reproductive system (*M. wasmanni* and *M. cf hellenius*) may have been derived from a dependent lineage (respectively, *Mebe2* and *Mstr2*) and reverted back to a classical reproductive system. Alternatively, (ii) the dependent lineage *Mebe2* may have arisen from ancestral hybridization between *M. ebeninus* and *M. minor* while *Mstr2* might stem from ancestral hybridization between *M. structor* and *M. cf hellenius*.

The analysis of males produced by workers also has important implications for our understanding of the maintenance of social hybridogenesis. A major cost associated with such a reproductive system is that dependent genetic lineages evolve independently, which could lead to an inability of colonies to produce workers in case of excessive divergence (Volny & Gordon 2002; Anderson *et al.* 2008). While inability to produce workers should theoretically be prevented by purifying selection, purifying selection is known to be particularly inefficient in eusocial insects due to their small effective population sizes (Romiguier *et al.* 2014b). Our finding that hybrid workers can produce viable males and that these males are the product of recombination between the hybrid genomes of workers raises the possibility of gene flow between the genetic lineages. Rare gene flow events between dependent lineages may prevent lineages from becoming too divergent, which could greatly contribute to the long-term persistence of these unusual reproductive systems, as it has been suggested in an analogous case for *Cataglyphis hispanica* (Darras & Aron 2015). On the other hand, introgression between lineages may also break the genetic caste determination system inherent to such reproductive systems, suggesting that gene flow must be rare and limited to genomic regions not involved in caste determination. Such rare gene flow limited to some parts of the genome has been recently reported between two highly divergent cryptic species (Roux *et al.* 2013).

Interestingly, the only heterozygous protein site common to all workers produced by social hybridogenesis (in *M. barbarus*, *M. structor* and *M. ebeninus*) belongs to the gene coding for *catalase*, an important enzyme in protecting the cell from oxidative damage by reactive oxygen species (ROS). This gene has higher expression in early larval development of honeybee queens compared to workers (Cameron *et al.* 2013), possibly to sustain the increased respiration rate of queen larvae (Melampy & Willis 1939). The same *catalase* position is always homozygous in *Messor* species with a normal reproductive system, both for workers and queens (Fig. 5), suggesting that homozygous females can develop into either workers or queens. *Catalase* is a tetramer of four polypeptide chains, raising the possibility that heterodimeric *catalase* proteins may compromise queen development in *Messor*, either because the enzyme is less efficient or has tightly co-evolved with another gene. Such an interaction between two loci has been proposed as the simplest possible model to explain genetic caste determination in *Pogonomyrmex* ants (Helms Cahan & Keller 2003). However, more work is required to determine the number and identity of the genes implicated in the strict mechanism of caste determination in social hybridogenesis in *Messor*.

In conclusion, this study reveals that social hybridogenesis is not as rare as previously thought. On the basis of our survey, as many as one-third of the species of the genus *Messor* (a widespread genus comprising 160 species) may be characterized by such a reproductive system. Our phylogenetic analyses revealed that social hybridogenesis can readily evolve several times from normal reproductive systems. Our study also suggests that there might be a link between unusual modes of reproduction and ecology, the currently known social hybridogenesis cases requiring three reproductive partners being restricted to two granivorous genera living in a dry climate. Finally, the analyses of males produced by hybrid workers demonstrate that recombination between lineages is possible, raising the possibility of rare but important gene flow between genetic lineages. This may contribute to long-term persistence of social hybridogenesis, as seems to be the case in *M. barbarus*.

Acknowledgements

We thank Nicolas Galtier and Camille Roux for useful discussions and P. Gayral for help during the sampling. We thank Tanja Schwander and Sean McGregor for comments on this manuscript. Some computations were performed at the Vital-IT (http://www.vital-it.ch) Center for high-performance computing (HPC) of the SIB Swiss Institute of Bioinformatics and the EPSRC-funded MidPlus HPC center. This work was supported by a Federation of European Biochemical Societies' (FEBS) long-term fellowship to J.R., the Swiss NSF, and an ERC Advanced Grant. SHY was supported by Swiss NSF to Prof. Michel Chapuisat (grant number: 31003A-146641).

References

Anderson KE, Gadau J, Mott BM *et al.* (2006) Distribution and evolution of genetic caste determination in *Pogonomyrmex* seed-harvester ants. *Ecology*, **87**, 2171–2184.

Anderson KE, Linksvayer TA, Smith CR (2008) The causes and consequences of genetic caste determination in ants (Hymenoptera: Formicidae). *Myrmecological News/Osterreichische Gesellschaft fur Entomofaunistik*, **11**, 119–132.

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for illumina sequence data. *Bioinformatics*, **30**, 2114–2120.

Cahan SH, Bradleigh Vinson S (2003) Reproductive division of labor between hybrid and nonhybrid offspring in a fire ant hybrid zone. *Evolution; International Journal of Organic Evolution*, **57**, 1562–1570.

Cahan SH, Daly AB, Tanja S, Arthur Woods H (2010) Genetic caste determination does not impose growth rate costs in *Pogonomyrmex* harvester ants. *Functional Ecology*, **24**, 301–309.

Cameron RC, Duncan EJ, Dearden PK (2013) Biased gene expression in early honeybee larval development. *BMC Genomics*, **14**, 903.

Capella-Gutierrez S, Silla-Martinez JM, Gabaldon T (2009) trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics*, **25**, 1972–1973.

Curry MM, Wheeler DE, Yang K, Anderson KE (2009) The potential for gene Ffow in a dependent lineage system of a harvester ant: fair meiosis in the F1 generation. *The Journal of Heredity*, **101**, 378–384.

Darras H, Aron S (2015) Introgression of mitochondrial DNA among lineages in a hybridogenetic ant. *Biology Letters*, **11**, 20140971.

Davidson DW (1982) Sexual selection in harvester ants (Hymenoptera: Formicidae: *Pogonomyrmex*). *Behavioral Ecology and Sociobiology*, **10**, 245–250.

Fournier D, Estoup A, Orivel J *et al.* (2005) Clonal reproduction by males and females in the little fire ant. *Nature*, **435**, 1230–1234.

Fournier D, Hellemans S, Hanus R, Roisin Y (2016) Facultative asexual reproduction and genetic diversity of populations in the humivorous termite *Cavitermes tuberosus*. *Proceedings of the Royal Society of London, Series B, Biological Sciences*, **283**, pii: 20160196. doi: 10.1098/rspb.2016.0196.

Gayral P, Philippe G, Lucy W *et al.* (2011) Next-generation sequencing of transcriptomes: a guide to RNA isolation in nonmodel animals. *Molecular Ecology Resources*, **11**, 650–661.

Gómez C, Abril S (2012) Nuptial flights of the seed-harvester ant *Messor barbarus* (LINNAEUS, 1767) (Hymenoptera: Formicidae) in the Iberian Peninsula: synchrony, spatial scale and weather conditions. *Myrmecological News/Osterreichische Gesellschaft fur Entomofaunistik*, **16**, 25–29.

Gordon DM, Anna P, De Bortoli N, Ingram KK (2013) Does an ecological advantage produce the asymmetric lineage ratio in a harvester ant population? *Oecologia*, **173**, 849–857.

Haydak MH (1943) Larval food and development of castes in the honeybee. *Journal of Economic Entomology*, **36**, 778–792.

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

Helms Cahan S, Parker JD, Rissing SW *et al.* (2002) Extreme genetic differences between queens and workers in hybridizing *Pogonomyrmex* harvester ants. *Proceedings of the Royal Society of London, Series B, Biological Sciences*, **269**, 1871–1877.

Hölldobler B (1976) The behavioral ecology of mating in harvester ants (Hymenoptera: Formicidae: *Pogonomyrmex*). *Behavioral Ecology and Sociobiology*, **1**, 405–423.

Hölldobler B, Wilson EO (1990) *The Ants*. Springer, Berlin.

Houle D (1989) Allozyme-associated heterosis in *Drosophila melanogaster*. *Genetics*, **123**, 789–801.

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

Kapheim KM, Bernal SP, Smith AR, Peter N, Wcislo WT (2011) Support for maternal manipulation of developmental nutrition in a facultatively eusocial bee, *Megalopta genalis* (Halictidae). *Behavioral Ecology and Sociobiology*, **65**, 1179–1190.

Kobayashi K, Kazuya K, Eisuke H, Kyohsuke O (2008) Clonal reproduction by males of the ant *Vollenhovia emeryi* (Wheeler). *Entomological Science*, **11**, 167–172.

Lebas C, Galkowski C, Blatrix R, Patrice W (2016) *Fourmis d'Europe Occidentale*. Delachaux et Niestlé, Paris.

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

Li L (2003) OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Research*, **13**, 2178–2189.

Li H, Durbin R (2010) Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics*, **26**, 589–595.

Linksvayer TA, Wade MJ, Gordon DM (2006) Genetic caste determination in harvester ants: possible origin and maintenance by cyto-nuclear epistasis. *Ecology*, **87**, 2185–2193.

Markl H, Hölldobler B, Hölldobler T (1977) Mating behavior and sound production in harvester ants (*Pogonomyrmex*, Formicidae). *Insectes Sociaux*, **24**, 191–212.

Matsuura K, Vargo EL, Kawatsu K *et al.* (2009) Queen succession through asexual reproduction in termites. *Science*, **323**, 1687.

Melampy RM, Willis ER (1939) Respiratory metabolism during larval and pupal development of the female honeybee (*Apis mellifica* L.). *Physiological Zoology*, **12**, 302–311.

Mott BM, Gadau J, Anderson KE (2015) Phylogeography of *Pogonomyrmex barbatus* and *P. rugosus* harvester ants with genetic and environmental caste determination. *Ecology and Evolution*, **5**, 2798–2826.

Pearcy M, Aron S, Doums C, Keller L (2004) Conditional use of sex and parthenogenesis for worker and queen production in ants. *Science*, **306**, 1780–1783.

Pearcy M, Goodisman MAD, Keller L (2011) Sib mating without inbreeding in the longhorn crazy ant. *Proceedings of the Royal Society of London, Series B, Biological Sciences*, **278**, 2677–2681.

Raj A, Stephens M, Pritchard JK (2014) fastSTRUCTURE: variational inference of population structure in large SNP data sets. *Genetics*, **197**, 573–589.

Ranwez V, Vincent R, Sébastien H, Frédéric D, Douzery EJP (2011) MACSE: Multiple Alignment of Coding SEquences accounting for frameshifts and stop codons. *PLoS ONE*, **6**, e22594.

Romiguier J, Gayral P, Ballenghien M *et al.* (2014a) Comparative population genomics in animals uncovers the determinants of genetic diversity. *Nature*, **515**, 261–263.

Romiguier J, Lourenco J, Gayral P *et al.* (2014b) Population genomics of eusocial insects: the costs of a vertebrate-like effective population size. *Journal of Evolutionary Biology*, **27**, 593–603.

Roux C, Tsagkogeorga G, Bierne N, Galtier N (2013) Crossing the species barrier: genomic hotspots of introgression

between two highly divergent *Ciona intestinalis* species. *Molecular Biology and Evolution*, **30**, 1574–1587.
Schwander T, Cahan SH, Keller L (2006) Genetic caste determination in *Pogonomyrmex* harvester ants imposes costs during colony founding. *Journal of Evolutionary Biology*, **19**, 402–409.
Schwander T, Cahan SH, Keller L (2007a) Characterization and distribution of *Pogonomyrmex* harvester ant lineages with genetic caste determination. *Molecular Ecology*, **16**, 367–387.
Schwander T, Keller L, Cahan SH (2007b) Two alternate mechanisms contribute to the persistence of interdependent lineages in *Pogonomyrmex* harvester ants. *Molecular Ecology*, **16**, 3533–3543.
Schwander T, Lo N, Beekman M, Oldroyd BP, Keller L (2010) Nature versus nurture in social insect caste differentiation. *Trends in Ecology & Evolution*, **25**, 275–282.
Sirviö A, Anu S, Pekka P *et al.* (2010) Origin and evolution of the dependent lineages in the genetic caste determination system of *Pogonomyrmex* ants. *Evolution; international Journal of Organic Evolution*, **65**, 869–884.
Smith CR, Suarez AV (2010) The trophic ecology of castes in harvester ant colonies. *Functional Ecology*, **24**, 122–130.
Smith CR, Toth AL, Suarez AV, Robinson GE (2008) Genetic and genomic analyses of the division of labour in insect societies. *Nature Reviews. Genetics*, **9**, 735–748.
Stamatakis A (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, **30**, 1312–1313.
Suni SS, Gordon DM (2010) Fine-scale genetic structure and dispersal distance in the harvester ant *Pogonomyrmex barbatus*. *Heredity*, **104**, 168–173.
Tschinkel WR (2006) *The Fire Ants*. Harvard University Press, Cambridge, Massachusetts.
Valone TJ, Michael K (2005) Interactions between granivorous and omnivorous ants in a desert grassland: results from a long-term experiment. *Ecological Entomology*, **30**, 116–121.
Volny VP, Gordon DM (2002) Genetic basis for queen-worker dimorphism in a social insect. *Proceedings of the National Academy of Sciences*, **99**, 6108–6111.
Volny VP, Greene MJ, Gordon DM (2006) Brood production and lineage discrimination in the red harvester ant (*Pogonomyrmex barbatus*). *Ecology*, **87**, 2194–2200.
Weir BS, Clark Cockerham C (1984) Estimating F-Statistics for the analysis of population structure. *Evolution; International Journal of Organic Evolution*, **38**, 1358–1370.
Yamauchi A, Yamamura N (2006) Persistence conditions of symmetric social hybridogenesis in haplo-diploid hymenoptera. *Journal of Theoretical Biology*, **240**, 24–31.

J.R. and L.K. conceived the project. J.R., A.F. and S.H.Y. performed sampling and laboratory work. J.R. developed the data analysis pipeline. J.R. and A.F. analysed the data. J.R. and L.K. wrote the paper.

Data accessibility

Raw ILLUMINA reads are freely available from the SEQUENCE READ ARCHIVE (SRA) database (http://www.ncbi.nlm.nih.gov/sra). See Table S1 (Supporting information) for each accession ID. Sequence alignments (*hsc70-4*, *cox1* and *catalase*) are available from the Dryad Digital Repository: http://dx.doi.org/10.5061/dryad.6gr08.

Supporting information

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

Table S1 List of *Messor* individuals with transcriptomic data sequenced. Species, caste, queen status (when relevant), colony/population/country of origin and ncbi SRA accession are provided.

Table S2 List of *M. barbarus* with the *hsc70-4* gene sequenced. Caste, population/colony of origin, sampling date, lineage, genotype at each diagnostic position of *hsc70-4* and presence/absence of the sequencing of the mitochondrial gene *cox1* are provided.

Table S3 Population genetic statistics from transcriptomic data. Each sheet corresponds to one different species. For each ORF (noted contigs), the number of bi-allelic/tri-allelic/quadri-allelic SNPs, the number of sequences (two per individuals), the number of complete sites, the GC3%, the non-synonymous nucleotidic diversity (piN), the synonymous nucleotidic diversity (piS), Fis values (with and without Weir-Cockerman correction) and individual heterozygosity are provided. Heterozygosity of each individuals is noted as H (*individual-id*), with *individual-id* corresponding to the ID field of Table S1.

Fig. S1 (A) Workers of *M. structor* have higher genome-wide heterozygosity levels (based on 90951 SNPs) than queens (all pairwise comparisons significant, $P < 0.0001$, Wilcoxon tests). (B) Population ancestry proportions estimated by *fastStructure* in 8 individuals. Each bar corresponds to individuals from (A). Bars are named according to the population of origin the individual. Queens are pure *Mstr1* (blue) or *Mstr2* (red) lineages, workers are *Mstr1*/*Mstr2* hybrids.