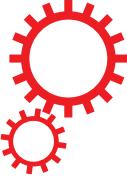


# SCIENTIFIC REPORTS



Corrected: Publisher Correction

OPEN

## The Interplay between Incipient Species and Social Polymorphism in the Desert Ant *Cataglyphis*

Received: 22 January 2019

Accepted: 19 June 2019

Published online: 01 July 2019

Tali Reiner Brodetski<sup>1</sup>, Shani Inbar<sup>2</sup>, Pnina Cohen<sup>2</sup>, Serge Aron<sup>3</sup>, Eyal Privman<sup>2</sup> & Abraham Hefetz<sup>1</sup>

In social insects, due to considerable polyphenism as well as high level of hybridization, the delimitation of species can be challenging. The genus *Cataglyphis* presents a high level of diversification, making it an excellent model with which to study evolutionary paths. Israel appears to be a "hot spot" for recent speciation in this genus. Although previous studies have described multiple species of *Cataglyphis* in Israel, a recent genetic study has questioned the existence of some of these historically described species. The present study focuses on an apparent species complex, the *C. niger* species complex which includes *C. niger*, *C. savignyi*, and *C. drusus* that are distinguishable by their mitochondrial DNA (and therefore named mitotypes) but not by their nuclear DNA. Using a multi-method approach (genetics, chemistry and behavior), we show that these mitotypes also differ in their social structures and are readily distinguishable by their cuticular hydrocarbons profiles. While most populations of the different mitotypes are allopatric, at our study site they are sympatric, but nonetheless maintain the observed differences between them. This raises the evolutionary question: Are these incipient species that have diverged with gene flow, or is this a case of social and chemical polymorphism that is maintained within a single species? Unveiling the interplay between social polyphenism and species segregation is at the core of evolutionary biology.

Although the biological species concept requires a reproductive barrier between species<sup>1</sup>, many alternative concepts (ecological, evolutionary, cohesion, phylogenetic, pragmatic and others) that are based on different common characteristics within species, have been proposed<sup>2,3</sup>. Much has also been discussed about whether sympatric speciation can occur and what mechanisms might facilitate it<sup>4</sup>. Sympatric speciation assumes adaptive divergence of multiple traits in order to generate strong ecologically-based reproductive isolation. However, the possibility of sympatric speciation without some gene flow does seem unlikely and, indeed, recent studies have reported that speciation with gene flow is more common than previously thought<sup>5,6</sup>. In insects, incomplete divergence in mating biology can lead to relatively high levels of hybridization between species, challenging the biological species concept and facilitating the possibility of cryptic speciation and speciation with gene flow<sup>7–9</sup>. Speciation can be considered as a continuous spectrum of various transitions or diversifications in morphological, genetic, and behavioral traits that at different time points might not yet show full evolutionary independence (i.e. species without gene flow between them). This phenomenon is known as incomplete lineage sorting<sup>10,11</sup>. In order to accurately describe diversifications between closely-related species that might be on the species continuum, classical morphology-based taxonomy is not always sufficient, but requires additional methods of characterization such as chemical, behavioral, ecological, and molecular phenotyping<sup>11–15</sup>. Differences between species, unlike those between populations within a species, should encompass vast changes that cannot be explained as simple polymorphism<sup>16</sup>. For example, cuticular hydrocarbon profiles, which are the presumed nestmate recognition cues in social insects, differ slightly between nests (setting them as colony insularity markers and exemplifying normal polymorphism) but greatly differ between species (exemplifying major changes). Likewise, behavioral changes that affect either mating preference or mating systems, foraging, or habitat selection can all lead to niche displacement and a drive for sympatric speciation<sup>17</sup>. In ants several factors have been suggested to lead to sympatric speciation. Examples include cases of intranidal mating, gyne (unmated ant queens) morphological polymorphism,

<sup>1</sup>School of Zoology, George S. Wise Faculty of Life Sciences, Tel Aviv University, Tel Aviv, 6997801, Israel. <sup>2</sup>Institute of Evolution, Department of Evolutionary and Environmental Biology, University of Haifa, Haifa, Israel. <sup>3</sup>Evolutionary Biology and Ecology, Université Libre de Bruxelles, Université D'Europe, Brussels, Belgium. Correspondence and requests for materials should be addressed to T.R.B. (email: [talireiner@gmail.com](mailto:talireiner@gmail.com))

a shift from monogyne (single queen) to polygyne (multiple queen) colonies, and host-parasite interactions<sup>15,18</sup>. These scenarios leads to mating place separation and facilitates further genotype divergence, eventually leading to novel genotypes and new species and describing a possible evolutionary path which new species partake in<sup>19</sup>. Social polymorphism (both monogyne and polygyne colonies in the same species and the same population) has been described in several ant species, including *Pheidole*, *Formica*, *Leptothorax*, etc.<sup>18</sup>. Finally, recent studies have shown that there are few cases in which social polymorphism and other traits do not lead to speciation but are acquired by means of a 'super-gene', also named 'social chromosome' distributed in the population<sup>20–22</sup>. However, this 'social chromosome' mechanism is rather rare and has been described only twice, while most incidents of social polymorphism are mainly believed to depend on ecological constraints.

The ant genus *Cataglyphis* comprises over a hundred species<sup>23</sup> displaying various social and population structures, ranging from the basic monogynous multicolonial population to polygynous supercolonies, making it an excellent model to study evolutionary paths<sup>23–27</sup>. The genus also expresses various reproductive strategies, from sexual to the asexual (parthenogenetic) reproduction of new reproductive queens<sup>24–28</sup>. Israel appears to be a "hot spot" for recent speciation in this genus: in the *bicolor* group, we identified eight mitochondrial types<sup>29</sup>, and in the *albicans* group there are nine identified mitochondrial types<sup>30</sup>. These mitochondrial types do not always concur with nuclear delimitation methods, and it is not clear whether mitochondrial types that have been previously used to delimit species can also serve as identifiers in this case of the *bicolor* group.

Here, we examined an apparent species complex in the *Cataglyphis bicolor* group of the previously described putative species *C. niger* and *C. savignyi*<sup>23,31</sup>, and a population tentatively identified as *C. drusus* (henceforth termed the "niger complex"). In a previous study, these putative species were found to exhibit different mitochondrial haplotypes (hereafter called mitotypes), but could not be differentiated on the basis of the nuclear genes used<sup>29</sup>. One explanation for this discrepancy is that the genetic markers employed in Eyer *et al.* 2017<sup>29</sup> were not sensitive enough to detect differences between the mitotypes. We therefore decided to clarify this point by using Double Digest Restriction Associated DNA (**ddRAD**) Sequencing, which provides many more genomic markers, enabling better accuracy for delimiting species. Furthermore, we wanted to determine the social and population structures of the mitotype's populations (both in allopatric and sympatric populations) using a multi-method approach that included: microsatellite markers, chemistry of cuticular hydrocarbons, and behavioral assays that were designed to examine colony insularity (indicating the colony structure from a colony composed of a single nest to a colony composed of multiple nests). Finally, we set out to infer if assortative mating occurs between the different mitotypes.

## Methods

**Sample collection.** Sample collection was divided according to the different analysis's methods, as different analyses were carried out at different time points. Disentangling the species complex as well as the deciphering the social and population structures were gradual processes as follows: 1. Samples collected from different geographic areas across Israel for the RAD sequencing analysis; 2. Samples collected from the Betzet and the Tel-Baruch populations for the microsatellite and the CHC analyses; 3. Samples collected from the Betzet and the Tel-Baruch populations for behavioral experiments; and 4. Samples of queens, newly mated gynes, and their spermatheca contents for mitotype analysis from the Tel-Baruch population.

1. In order to clarify further the taxonomic status and species delimitation of the *bicolor* group in Israel<sup>29</sup> we performed RAD sequencing on samples from different localities, which were further divided into groups based on their mitotype *C. drusus* ( $n=9$ ), *C. niger* ( $n=14$ ), *C. savignyi* ( $n=47$ ), *C. israelensis* ( $n=35$ ), *C. isis* ( $n=4$ ), *C. holgerseni* ( $n=3$ ), and samples from Turkey ( $n=6$ ) identified as *C. nodus*. Individual ants were collected from 55 locations across Israel (sampling map in Supplementary Fig S1), and from Turkey (kindly collected by Dr. Kadri Kiran, Trakya University, Turkey), 76 samples were originally collected along with samples for Eyer *et al.*<sup>29</sup> and the rest 42 samples were specifically collected for this project (total of 118 samples collections are detailed in Table S1). All samples for genetic analysis were kept in absolute ethanol until extraction.
2. To determine social and population structure we used both microsatellite markers and CHC on samples collected from all visible nests at two plots as follows: Nests were collected from two coastlines locations: Betzet ( $33^{\circ}4'36.09''N$ ,  $35^{\circ}6'32.05''E$ ) and Tel Baruch ( $32^{\circ}7'33.76''N$ ,  $34^{\circ}47'9.04''E$ ). These locations present similar ecological niches that are characterized by semi-stabilized sand dunes. We sampled all visible nests in a plot of  $200 \times 200$  m, as well as a transect of 1 km (at 100 m intervals). From the Betzet and Tel Baruch sites respectively, we collected partial nests-queenless nests (QL;  $n=57/39$ ), and whole nests/queenright nests (QR;  $n=2/11$ ; at). All the nests in the experimental plot were excavated over four consecutive days to avoid pseudo-replication due to nest migration. Ants were dissected as follows: the head, for CHC content, was kept in hexane; the thorax and legs, for DNA extraction, were kept in 100% ethanol; and the abdomen was frozen at  $-20^{\circ}C$  and kept for future analysis. Table S2 details the number of nests used in each analysis and collection site.
3. Because some of the results from the microsatellite and the CHC data were undecisive, we decided to take a multi-method approach<sup>32</sup> and use behavioral assays. Nests at the above-mentioned locations were thoroughly excavated to include the queen, all workers, and brood. To ensure that the entire nest was collected, excavation was further extended by an extra 0.5–1 m in all directions after having observed the last gallery. In the Betzet location, despite our thorough excavations the queens were difficult to locate, so we used both QR ( $n=3$ ) and QL nests ( $n=15$ ) for the experiments. At Tel Baruch we successfully excavated and used 10 QR nests.
4. In order to determine if mating occurs between the different mitotypes we collected newly-mated queens and analyzed their spermatheca content. Established queens were collected from the above excavated

nests. Newly mated gynes were collected from the Tel Baruch population during the 21 consecutive days of nuptial flight. Some of the gynes had shed their wings and some had also begun to dig a new nest. All were dissected and their spermatheca content was extracted. Sperm was kept in Ringer's buffer and frozen at  $-80^{\circ}\text{C}$  until DNA extraction.

**Genetic analysis.** *Mitochondrial haplotypes.* were determined by genotyping Cytochrome B using one worker per nest<sup>29</sup>. CytB primers used: CB1 (Forward) TATGTAACCATGAGGACAAATTC and CB2 (Reverse) ATTACACCTCCTAATTATTAGGAAT, annealing temperature was  $48^{\circ}\text{C}$ <sup>33</sup>. PCR products were purified with ExoSAP-IT™ PCR purification kit (Applied Biosystems). Sequencing was performed on an ABI 3730 Genetic Analyzer (Applied Biosystems). Base calling and sequence reconciliation were inspected by eye and performed using Unipro UGENE v 1.3. Sequences were aligned using MUSCLE algorithm<sup>34</sup> implemented in UGENE Aligner, together with sequences previously used in Eyer *et al.*<sup>29</sup>.

*RAD sequencing.* All the DNA for this analysis was extracted using Qiagen blood and tissue DNA extraction kit following the standard protocol. We used a modified protocol of<sup>35,36</sup>, to construct double-digest RAD libraries, modifications added by A. Brelsford, A. Mastretta-Yanes, J. Leuenberger, R. Sermier. Changes were as follows: Added dual-index barcoding to allow multiplexing >96 samples per library. The Y-adapter for MseI from Peterson *et al.* (ddRAD) was used to prevent amplification of MseI-MseI fragments. Simplified restriction and ligation mixes using CutSmart buffer. Purification post-ligation was done using Agencourt AMPure beads. PCR using Q5 Hot Start Polymerase (New England BioLabs). PCR modified to decrease number of cycles by increasing number of replicates and starting DNA volume. Addition of primers and dNTPs for a final thermal cycle, in order to reduce production of single-stranded or heteroduplex PCR products. Final cleanup and size selection were carried out with AMPure beads. Three samples were omitted from the final results due to poor sequencing yield. All samples were multiplexed and sequenced by single-end 100 bp reads on 2 lanes of an Illumina HiSeq4000 sequencer.

*Processing of RAD sequence data.* A total of 648,303,541 reads were sequenced, with an average of 5,026,520 reads per sample (275,586 reads for the lowest coverage sample, pol633\_u). Dividing the number of reads by the 67,687 EcoRI restriction sites in the reference genome, it is an average coverage of 37X. The raw reads were initially processed using the Stacks pipeline<sup>37,38</sup> and low quality reads were discarded. These were defined as reads in which the score drops below an averaged phred score of 10 in a sliding window of 15% of the reads' length, that is 14 bp for a read length of 93 bp. After this initial processing, 583,275,943 reads that were assigned to a sample (using the specific barcode sequence that was attached to each sample) remain (92%). A total of 459,081,529 reads were mapped to the reference genome of *C. niger* (version Cnig\_gn1) using Bowtie2<sup>39</sup>. We filtered the alignments allowing for a maximum of four mismatches. To filter non-unique mappings, we removed alignments where the second-best hit had less than twice the number of mismatches found in the best hit. After this filtering, 369,719,665 alignments remained. At this point, the actual coverage of mapped loci was calculated, with an average of  $22\times$ . We continued to analyze the mapped sequences to identify single nucleotide polymorphic sites (SNPs) using the Stacks pipeline and created a catalog containing 300,509 SNPs for the 126 samples.

*STRUCTURE analysis for RAD samples.* Individual samples were divided to seven populations according to their mitotype: *israelensis*, *drusus*, *niger*, *savignyi*, *holgerseni*, *isis* from Israel, and from Turkey- *noda*. The data in the catalogue was filtered as follows: 1) loci were removed if they had more than 7% missing data (samples where a genotype could not be called at that locus). 2) Minimal minor allele frequency (MAF) of 1% (a minimum of 3 allele copies in the 230 chromosomes of 115 diploid individuals). 3) A minimum of five reads per individual for calling a genotype. Each sample was required to have genotype calls in at least 80% of the loci. This resulted in a catalogue of 19,995 high confidence SNPs that were used in the population structure analysis (Supplementary Fig. S2).

*STRUCTURE*<sup>40,41</sup> was run over  $K$  (number of expected clusters) values of 3–10; for each  $K$  value, 100 iterations of *STRUCTURE* were run, each over a different set of 1,000 SNPs randomly chosen out of the high confidence SNPs. In each set, a minimal genomic distance of 2,000 nucleotides between sites was maintained to comply with *STRUCTURE*'s assumption of no linkage disequilibrium. The program was run with 1,100,000 MCMC iterations, and the first 100,000 were discarded (burn-in period). All the other parameters were kept at default. The results were analyzed using *CLUMPAK*<sup>42</sup> and combined for each  $K$  value.

To test whether the three species in the complex are genetically separated or not we used full maximum likelihood calculations to compare between two multi-species coalescent models as implemented by 3s program<sup>43–45</sup>. The null model which does not allow gene flow between two closely related species and an alternative model that does. Both models include a third species that is used as an out-group. This species is assumed not to have genetic flow with the species tested. We preformed multiple such likelihood ratio tests (LRT) between different combinations of individuals of the populations of *C. savignyi*, *C. drusus*, *C. niger*, *C. israelensi* and *C. isis*, first to check for gene flow between *C. israelensis* and any one of the species in the complex with the out-group of *C. isis* and then between the species of the complex with the out-group of *C. israelensis*.

The reference genome of *C. drusus* was aligned to the 93 base pairs long sequences of individual ants, the RAD sequencing product. Allowing a maximum of 4 mismatches for the best hit and a minimum of twice as many mismatches as the best hit for the second-best hit, only this time no gaps were allowed for the alignment. Based on the position of the sequences in relation to the reference genome, three sequences from three different individuals were aligned together in one of the following configurations: 123 – a sequence from individual of species 1, a sequence from an individual of species 2 and a sequence from individual of species 3; 223 – two sequences

from two different individuals of species 2 and a sequence from individual of species 3; and similarly, 113. This is assuming species 1 and 2 are tested for gene flow while species 3 is used as an out-group and is assumed to have no gene flow. At least 2000 bp gap was maintained between the triplets of aligned sequences (as inferred from the reference genome). For each locus one of the three possible alignment configurations was chosen at random.

For each combination of species, 3 s was run three times over the corresponding aligned sequences, using the Gaussian quadrature number of points = 16 and different seed values. Then LRT was performed between the highest scoring null and alternative models. The null model of no genetic flow was rejected if  $2D\ln L > 5.99$  (corresponding to  $\chi^2$  distribution with two degrees of freedom and significance level of 0.05 for type I mistake).

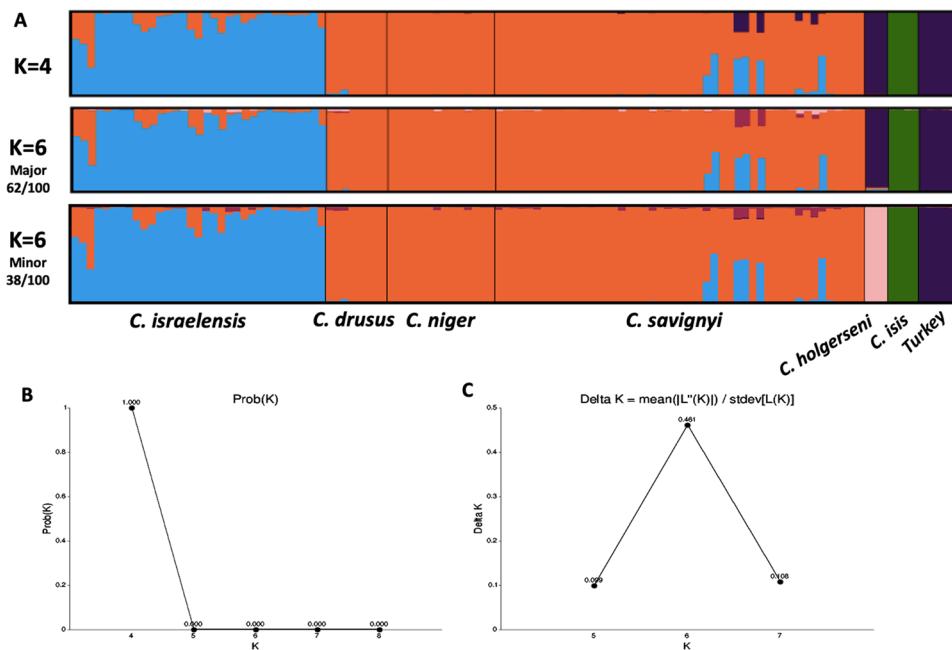
**Microsatellites.** To understand the genetic population structure as well as the social structure of the Betzet and Tel Baruch populations we used a microsatellite markers analysis. The analysis was performed on 648 workers of the Betzet population (*C. drusus* only;  $X + SE = 11.42 \pm 0.1$ ,  $N = 57$  nests). DNA was extracted with 5% CHELEX (BIO-RAD) and then amplified with 11 microsatellite markers (previously designed for: *C. hispanica* - Ch01, Ch08, Ch10, Ch11, Ch12; Darras, Kuhn and Aron, 2014, *C. cursor* - Cc54, Cc65, Cc96 and Cc11; Pearcy *et al.*, 2004, and *C. niger* - Cn02, Cn04; Saar *et al.*, 2014). Similarly, genetic analysis was performed on 548 workers of the Tel Baruch population ( $X + SE = 21.08 \pm 0.45$ ,  $N = 26$  nests). DNA was amplified with 7 microsatellite markers (previously designed for *C. hispanica* - Ch23; Darras *et al.* 2014, *C. cursor* - Cc99, Cc54, Cc51; Pearcy *et al.* 2004, and *C. niger* - Cn02, Cn04, Cn08; Saar *et al.* 2014). Different markers were used for the different populations because some of the makers did not work well in some of the populations (null alleles, linkage disequilibrium, and low amplification yield). Amplification was done using Type-it PCR mix (QIAGEN). PCR products were sequenced with ABI3500, genotypes were then analyzed with GeneMarker. Microsatellite data were checked for linkage disequilibrium, fixation index, and G-test using GENEPOLP<sup>46</sup>, heterozygosity was checked using both GENEPOLP and FSTAT<sup>47</sup>, and  $F_{st}$  was checked using FSTAT. To assess the number of queens, queen genotypes, and the number of matings we used the program COLONY<sup>48</sup>. The program STRUCTURE<sup>40</sup> was used to evaluate the number of colonies in the population. The number of K (number of possible colonies, from 1 to n size) was evaluated by the Evanno method using STRUCTURE Harvester<sup>49</sup>. PCA of microsatellite data was performed as described by Ryan *et al.*, 2017<sup>50</sup> (using R version 3.3.3; library adegenet and ggplot2 for visualization), and relatedness was estimated using Coancestry<sup>51</sup> (Queller GT method).

**Behavioral assays.** In order to examine nest insularity and colony composition at the population level we used nest-merging experiments<sup>32</sup>. We composed queenright (QR) colonies ( $n = 3$  and  $n = 10$  for Betzet and Tel Baruch, respectively) comprising approximately 200 workers maintained in nesting boxes prior to the experiment, as well as freshly excavated queenless (QL) nests ( $n = 15$ ). Workers and queens of each colony were marked on the thorax with different colors of nail polish, and then left for a day to get acclimated. At the onset of the merging experiment the two nests were connected to a common foraging arena (size  $20 \times 20$  cm). Sand and food were placed in the foraging arena and water and cotton balls in the enclosed nests. Scan observations were performed for up to three days (merging was stopped on two occasions in order to save the queens), while recording interactions between non-nestmates, including aggression, trophallaxis, adult carrying, and antennation. For the Betzet population, merging experiments were carried out between 18 nests, three of which were QR and 15 were QL. The merging tests between the QL nests were performed up to a week after excavation, in order to avoid a possible 'QL nest behavior', in which nests that are rendered queenless tend to lose their specific nest 'signature' over time. All nests were tested against all nests. If nests merged together, they were left as one colony, and at times were used again as a single colony for further analyses. If we observed over ten aggressive interactions within ten minutes, we stopped the test and separated the two nests. In Tel Baruch merging was done using 10 QR nests of the different mitotypes, as detailed in the results. If nests merged, they were left together as one colony with multiple queens.

**Cuticular hydrocarbon analysis.** In order to distinguish colony boundaries according to their cuticular hydrocarbon profiles, we immersed individual heads in hexane (5–10 ants from each nest), containing 100 ng/ $\mu$ l of eicosane ( $C_{20}$ ) as internal standard. Initial analysis was conducted by gas chromatography/mass spectrometry (GC/MS), using a VF-5ms capillary column that was temperature-programmed from 60 °C to 300 °C (with 1 min initial hold) at a rate of 10 °C per min, with a final hold of 15 min. Compound identification was achieved according to their fragmentation pattern and respective retention indices compared to authentic compounds. We identified 74 and 60 compounds for the Tel Baruch and Betzet population samples respectively (Supplementary Fig. S3A,B). Quantitative analyses were performed by flame ionization gas chromatography (GC/FID), using the above running conditions. Peak integration was performed using the program Galaxie Varian 1.9. A total of 37/35 hydrocarbons where used in the Tel Baruch and Betzet analysis respectively (highlighted in bold in Supplementary Fig. S3A,B). Hydrocarbon profile specificity was determined using multivariate statistics, i.e., discriminant analysis using the stepwise forward mode. This module of the program works initially as a principle component analysis for reducing the number of variables to fit the number of cases, and then performs the discriminant function. We then used k-means clustering in order to determine the structuring of the population, and assessed the most probable K using kmeans function in stats, as well as mclust in mclust and nbclust in NbClust libraries in R. Plot was done by clusplot function in cluster library.

## Results

**RAD sequencing.** STRUCTURE analysis for the possible K clusters was examined between 3 to 10 numbers of K's, resulting in the most probable Ks of 4 and 6 (Fig. 1). There was a clear separation between *C. isis*, *C. holgerseni* and *C. israelensis* at all possible Ks, and in the minor runs of K = 6 also *C. nodus* was clearly distinct. However, at K = 4 or major runs of K = 6 distinction between the *C. holgerseni* and *C. nodus* samples was not clear



**Figure 1.** (A) STRUCTURE results of the ddRADseq, possible K outcomes 4 and 6, major and minor runs showed differences in the *C. holgerseni* cluster. *C. israelensis* in light blue, the *C. niger* complex: *C. drusus*, *C. niger*, and *C. savignyi* in orange, *C. holgerseni* in dark blue or pink, *C. isis* in green, and *C. noda*- from Turkey in dark blue. (B) Analysis of most probable K by the Pritchard method. (C) Analysis of the most probable K using the Evanno method.

Species combination	loci Nr.	2DlnL	Pvalue
<i>C. israelensis</i> : <i>C. savignyi</i> , <i>C. isis</i>	23,031	0	>0.99
<i>C. israelensis</i> : <i>C. niger</i> , <i>C. isis</i>	22,770	4.933	0.1
<i>C. israelensis</i> : <i>C. drusus</i> , <i>C. isis</i>	23,662	0.55	0.975
<i>C. savignyi</i> : <i>C. niger</i> , <i>C. israelensis</i>	27,769	103.99	<0.001
<i>C. savignyi</i> : <i>C. drusus</i> , <i>C. israelensis</i>	28,830	384.18	<0.001
<i>C. niger</i> : <i>C. drusus</i> , <i>C. israelensis</i>	28,626	57.399	<0.001
<i>C. Holgerseni</i> : <i>C. savignyi</i> , <i>C. isis</i>	22,771	21.17	<0.001
<i>C. savignyi</i> : <i>C. isis</i> , <i>C. holgerseni</i>	23,561	0.008	>0.99

**Table 1.** Gene flow 3S test. 3S run results comparing two coalescent models for different triplets of species. The difference shown is between the highest scoring alternative model and the highest scoring null model between 3 runs for each of the species combination. The bold lines are for species triplets, for which the null model was rejected, suggesting gene flow between the first two species.

cut (blue or pink and blue respectively). In the *C. israelensis* population, some individuals showed genetic polymorphism that suggests hybridization between *C. israelensis* and the *C. niger* complex, mainly within the *savignyi* mitotype. STRUCTURE analysis for any attempted K (between 3 to 10) could not separate the three mitotypes populations of *C. drusus*, *C. niger* and *C. savignyi* and they are all clustered together (Fig. 1; colored in orange).

Species delimitations within the *niger* complex were put to the test using the maximum likelihood approach as implemented by 3 s program, comparing a coalescent null model that does not allow for gene flow between tested species and an alternative model that does. At first *C. israelensis* was tested against each of the species in the complex with the out-group of *C. isis*. As the null model could not be rejected for any of the tests, we used *C. israelensis* as the out-group in the following tests to check for gene flow within the *niger* complex. In all three tests the null model was rejected and the *savignyi*, *niger* and *drusus* mitotypes were found to have gene flow between them. For the combination of *C. Holgerseni*, *C. savignyi* and *C. isis*, *C. Holgerseni* was found to have gene flow with the *savignyi* mitotype (Table 1).

**The Betzet population.** *Genetic analysis.* We examined a subset of 25 nests from the Betzet population for their mitotypes, and found that all of them presented the same mitotype, similar to the *drusus* group described in Eyer *et al.* (2017).

Microsatellite DNA analyses revealed that out of the 11 markers amplified and sequenced, three (Cc65, Cc96 and Cc11) were at linkage disequilibrium and therefore omitted from the analysis. None of the markers used

in this study showed indications of null alleles. The number of alleles in the remaining eight microsatellites loci ranged from 4 to 16. Mean heterozygosity was  $H_o = 0.65$  (range 0.31–0.84). The fixation index  $F_{IT}$  was slightly higher and significantly different from zero (mean  $\pm$  SE jackknife = 0.09 + 0.06; Z test = 11 p-value < 0.0001), indicating deviation from Hardy-Weinberg equilibrium and suggesting that mating was not completely random in the population studied.  $F_{ST}$  across the 59 nests sampled at Betzet was 0.19, while within nest relatedness was 0.35 (SE jackknife: 0.02). The G-test showed a result significantly different from zero between some of the nests, indicating that they were genetically differentiated and belonged to distinct colonies (21 nests), while for other nests the G test was not significantly different from zero and they are considered to belong to the same colonies (36 nests divided into 13 colonies). STRUCTURE analysis indicated that the population consists in 2, 3, or 13 genetic colonies (most probable K numbers; Fig. 2A,B) out of 57 nests analyzed. However, this result of 13 colonies did not seem very probable due to the dispersed location within the plot of the nests assigned to a particular genetic colony, and the fact that mixed genotypes were present within a single nest. K = 2 did not seem very probable either, as it suggests a uniclonal population (we did not see any evidence for such a population structure when observing the population in the field). We therefore conducted a Principle Component Analysis using the genetic colonies assigned by STRUCTURE division of K = 3, which revealed only a very small differentiation in genotypes, which were mainly very similar (PC1 = 3.4%, PC2 = 3.1%; Fig. 2C).

Genotyping of workers from 42 nests showed that within each nest all of the workers' genotypes matched to a single queen; this conclusion was also supported by COLONY, unequivocally confirming monogyny. In the remaining 15 nests some of the worker genotypes (average of 1.7 workers per nest; SE = 0.22) did not perfectly match the queen's assigned genotype, suggesting either multiple queens (which is possible but less likely in light of the above data indicating monogyny), queen turn-over with workers that had remained in the nest from previous residing queens or the presence of drifting workers (workers who mistakenly enter and remain in another colony; a phenomenon documented in other *Cataglyphis* species; M. Knaden, personal communications). Comparing the reconstructed matrilines of different nests based on the COLONY results, we assumed that nests with similar matrilines belong to the same colony, resulting in several polydomous colonies found in the population (Table S3 and in Fig. 2D). Considering monogyny, the COLONY analysis suggested polyandry with an average of 5.3 (SE = 0.22) males contributing to worker production (range: 2–9 males).

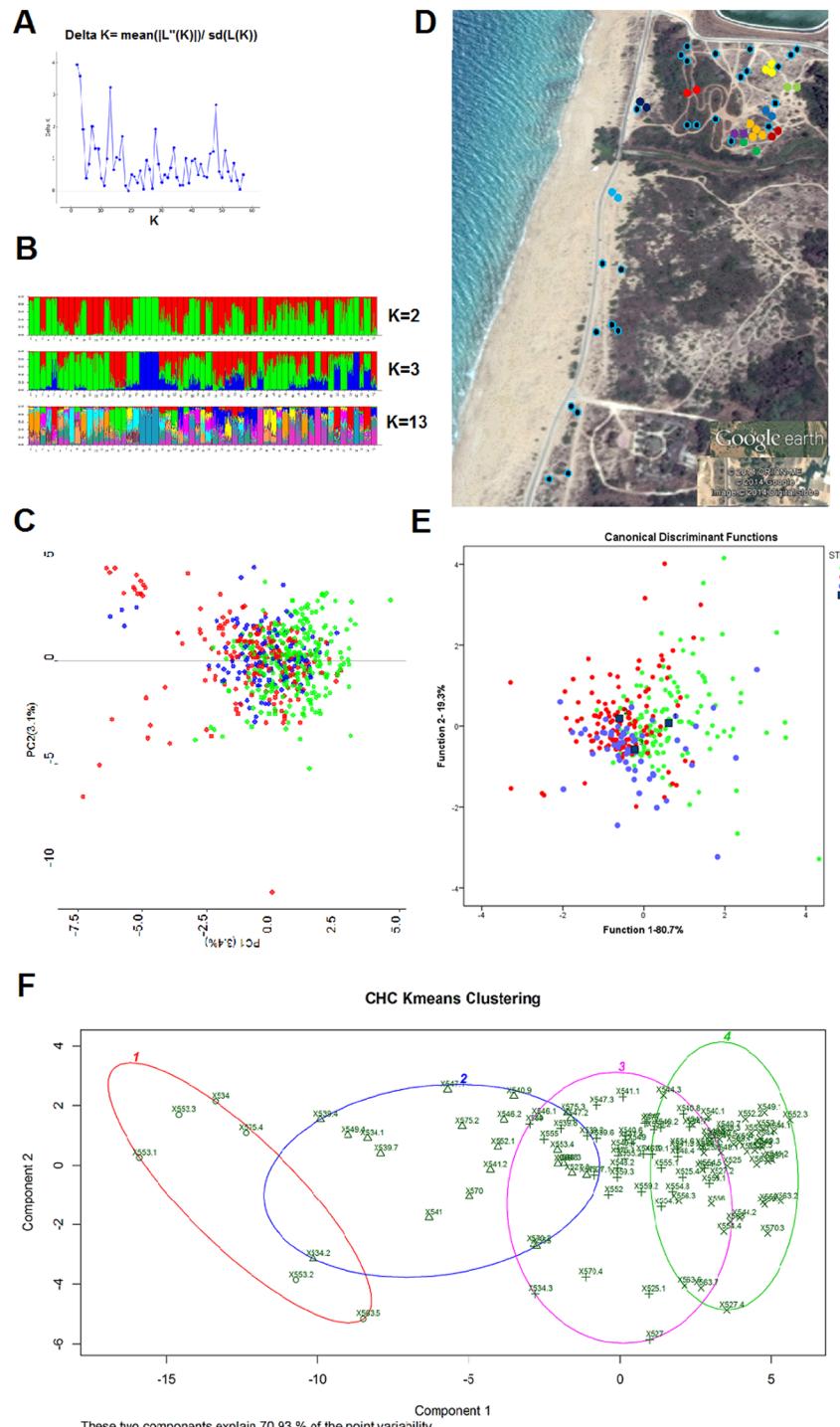
**Cuticular hydrocarbons.** Cuticular washes contained 60 identified hydrocarbons, of which 35 were used in the following analyses (Supplementary Fig. S3A). CHC profiles were analyzed by calculating the relative amounts of each compound (peaks comprising inseparable compounds were considered as a single compound), for 49 nests (275 individuals). We then used the k-means approach to determine the level of structuring in the samples. The most probable number of K's was 4 according to all three analyses methods used (k means, mclust, and nbclust), with the two components explaining 70.93% of the variability (Fig. 2F). We further performed a discriminant analysis based on the CHC profiles of the genetic colonies assigned by 'STRUCTURE' at K = 3. Colony assignment of three explains a high percentage of the variance (function 1–80.7% and function 2–19.3%), although the profiles are very similar and overlap (Fig. 2E).

**Behavior.** A possible resolution of the presumed discrepancy between CHC and genetic data lies in performing behavioral assays, that is, in examining whether two distinct nests will merge to a single colony when given the opportunity. Out of the various 27 trials performed, 14 nest-pairs successfully merged, culminating in seven colonies, each encompassing one to five nests (Table 2 and Fig. 3). During the successful merges there was no aggression at all, either between workers or between workers and queens. The remaining 13 trials resulted in mutual aggression between workers that belonged to different nests, as well as aggression towards the queens (in QR nests). These trials were stopped after ten minutes to avoid further injuries to workers and queens. When merging occurred, it was between either a QR and a QL nest or two QL nests, and generally when nests were located in close proximity (except for one case), suggesting that colonies with the *C. drusus* mitotype in the Betzet population are generally polydomous, having one central QR nest and several satellite QL nests.

**Tel Baruch Population.** *Genetic analysis.* Genetic analysis ( $N = 26$  nests) using CytB identified representatives of all three mitotypes belonging to the 'niger complex' in the sampled plot (*niger*, *savignyi* and *drusus*;  $n = 6$ , 17 and 3 nests, respectively).

Microsatellite DNA analysis was performed on 548 workers of the Tel Baruch population (mean number of workers typed per nest  $\pm$  SE =  $21.08 \pm 0.45$ ,  $N = 26$  nests) at 7 polymorphic loci. None of the seven microsatellite markers showed indications of null alleles or linkage disequilibrium. First, we analyzed all the nests in the population, irrespective of their mitotypes, assuming that they comprise a single population. The number of alleles ranged from five to 25. Mean heterozygosity was  $H_o = 0.69$  (range: 0.4–0.87). The fixation index  $F_{IT}$  was low, even though slightly higher than zero (mean  $\pm$  SE jackknife = 0.02 + 0.01) indicating that random mating predominates in the tested population.  $F_{ST}$  was low at 0.12 while the within-nest relatedness was 0.23 (SE jackknife: 0.01). STRUCTURE indicated that the 26 nests analyzed consisted in 2, 3, or 6 colonies (Fig. 4A,B). Comparing colony distribution according to the above results (Fig. 4C) with that of the mitotype distribution (Fig. 4D) showed incongruence. Furthermore, PCA distribution of the genotypes showed that, overall, they were very similar throughout the population (Fig. 4E).

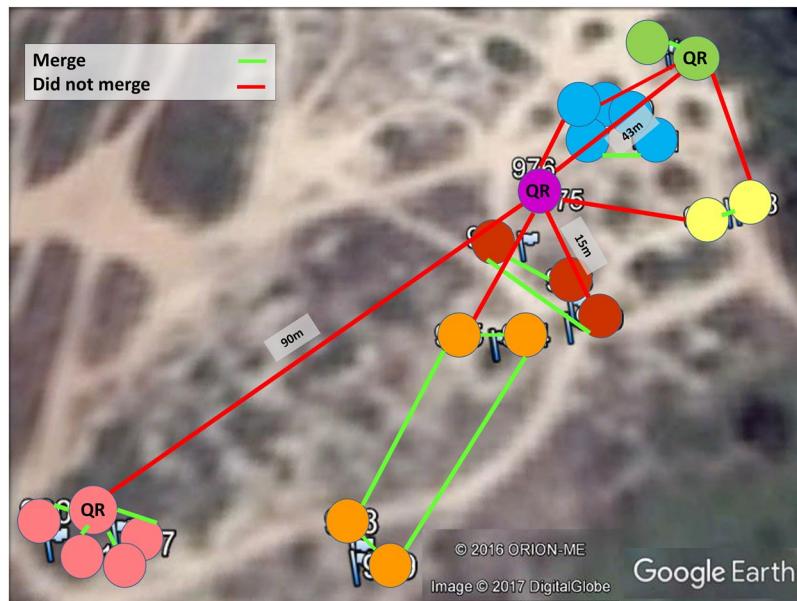
Genetic analysis of the population structure considering the mitotypes separately gave a slightly different picture. Each mitotype had its distinct characteristics (Table 3); the *niger* mitotype had a very low  $F_{ST}$  and very low within-nest relatedness, as predicted for polygynous colonies. The *drusus* and *savignyi* mitotypes, on the other hand, exhibited higher  $F_{ST}$  and relatedness values, as predicted for monogynous colonies. However, perhaps due to the rather small sample size or the unique situation of this sympatric population, we obtained slightly different

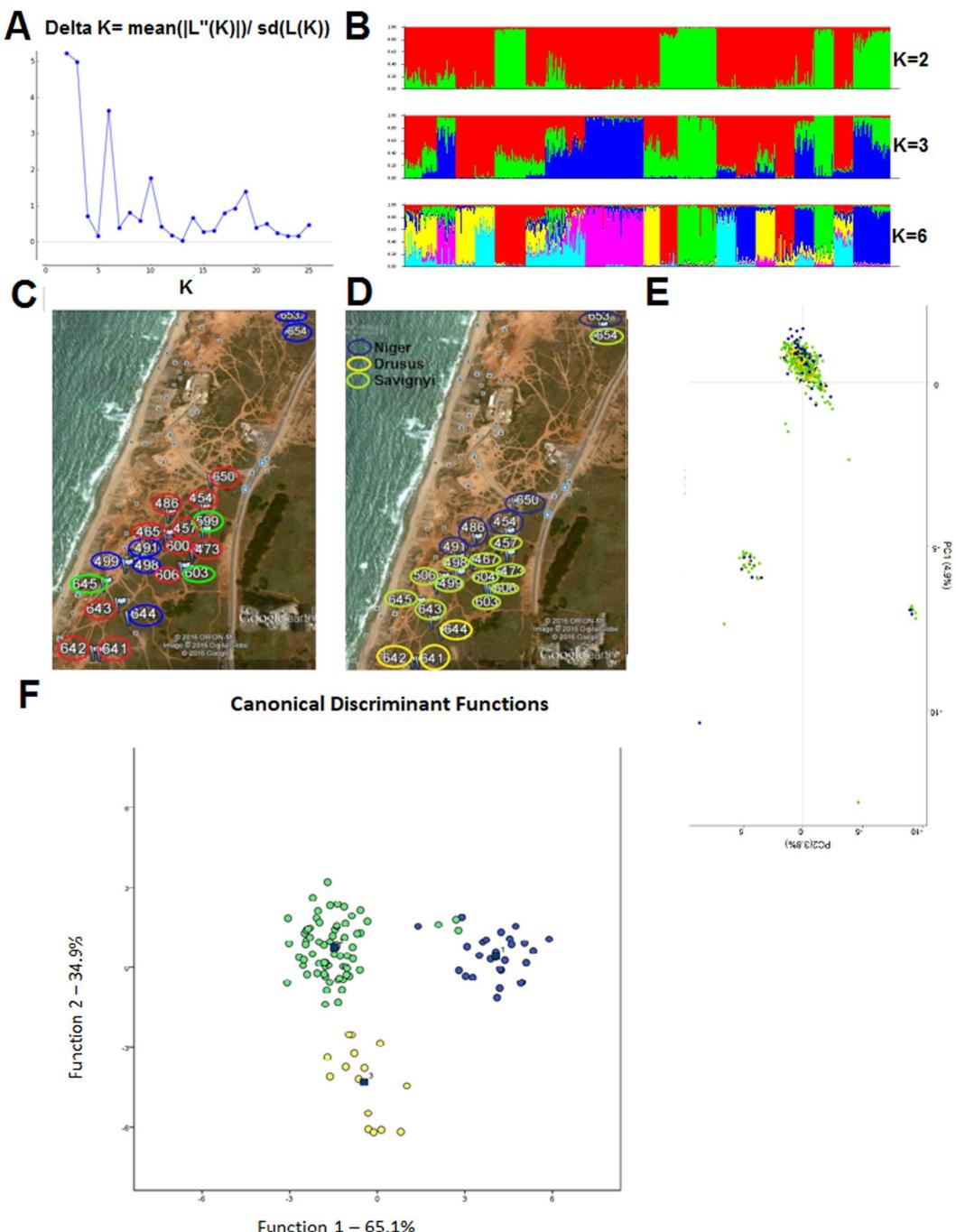


**Figure 2.** (A) STRUCTURE harvester results for the Betzet population indicating the best probable  $K$  is  $K = 2$ , then  $K = 3$  and  $K = 13$ . (B) The population STRUCTURE divisions according to the most probable  $K$  possibilities. (C) Principle Components Analysis of genotypes colored according to STRUCTURE result  $k = 3$  (D) Locations of estimated colonies based on similar assumed matrilines according to COLONY; same color indicates the same matriline, light blue circle with black edge indicates a colony with no other similar matriline. (E) Discriminant Analysis of CHC groups are set according to  $k = 3$ . (F) K- means of CHC profiles  $K = 4$ , individuals are marked with nest number (first 3 digits) and individual ID. Circle indicates  $k-1$ , triangle  $k-2$ , cross  $k-3$ , and X  $k-4$ . The two components explain 70.93% of the variation.

results to those for the other studied populations containing a single mitotype (data from: the Betzet population- only *drusus* mitotype in this paper, the Nitzanim population- only *niger* mitotype from Leniaud *et al.*, 2011, and the Rishon and Arad populations- only *savignyi* mitotype from<sup>52,53</sup>; Table 3).

Nest A	Nest B	Merging	Between-colony trophallaxis	Between-colony aggression	Maximal distance between colonies (meters)
1	9	+	+	-	1.5
2	3	+	+	-	2
4	5	+	+	-	1.5
14	15	+	+	+	2
10	11	+	+	+	2.3
16	17	+	+	+	10.4
6	4+5	+	+	+	4.5
8	13	-	-	+	14
12	13	-	-	+	20
8	12	+	+	-	18
2+3	4+5+6	-	-	+	34
1+9	4+5+6	+	+	-	8.5
1+9	2+3	-	-	+	10
16+17	14+15	-	-	+	37
10+11	2+3	-	-	+	36
16+17	18	+	+	-	9.4
14+15	12+8	-	-	+	15–50
1+9+4+5+6	10+11	-	-	+	20
1+9+4+5+6	12+8	-	-	+	37
1+9+4+5+6	7	+	+	-	22
10+11	12+8	+	+	-	55
14+15	13	+	+	-	42
14+15	10+11	-	-	+	75–112
10+11	16+17+18	+	+	-	140
2+3	7	-	-	+	34
7 (QR)	10+11 (QR)	-	-	+	43
7 (QR)	16+17+18 (QR)				90

**Table 2.** Merging assays of the Betzet population.**Figure 3.** 22 nests that were collected for the behavioral assays. Each circle indicates a physical nest, while each color indicates a different colony obtained by the merging assays. Queen-right nests are indicated by QR. Assays were performed between all colonies as indicated in Table 2. Nests that merged are indicated by a green line; nests that did not merge are indicated by a red line.



**Figure 4.** (A) STRUCTURE harvester results for the Tel Baruch population indicating the best probable  $K$  is  $K = 2$ , then  $K = 3$  and  $K = 6$ . (B) The population STRUCTURE divisions according to the most probable  $K$  possibilities. (C) STRUCTURE results for the Tel Baruch population of  $K = 3$  plotted on the map, the compression with the mitotypes plotted in (D) exemplifies that the STRUCTURE division is not related to the mitochondrial DNA. (D) Plotted on the map of the Tel Baruch sampling site Cytochrome B phylogenetic clades. (E) Principle Components Analysis of genotypes for Tel Baruch population colored according to the different mitotypes. (F) Discriminant analysis results of CHC's composition for the Tel Baruch population according to the three mitotypes.

**Cuticular hydrocarbon analyses.** For chemical analyses, we used head extracts of 76 ants belonging to 28 nests. We identified 74 compounds in the cuticular washes, of which 37 were used in the following analysis (Supplementary Fig. S3B). Profile determination was constructed by calculating the relative amounts of each of the hydrocarbon components, and profile comparison was done by discriminant analysis, while grouping the ants according to the CytB mitotype. As can be seen in Fig. 4F, there was a clear discrimination of CHC profiles among the three mitotypes.

Haplotype	Fit	Fit in a single mitotype population	Fst	Fst in a single mitotype population	Fis	Within nest Relatedness	Within nest Relatedness in a single mitotype population
All nests/ haplotypes n = 26	0.022 + 0.013		0.117 + 0.008		0.108 + 0.008	0.229 + 0.013	
<i>niger</i> n = 6	0.051 + 0.02	0.023 + 0.008 <sup>a</sup>	0.059 + 0.005	0.005 + 0.003 <sup>a</sup>	0.008 + 0.019	0.112 + 0.009	0.02 + 0.007 <sup>a</sup>
<i>drusus</i> n = 3	0.027 + 0.033	0.09 + 0.06 <sup>b</sup>	0.092 + 0.008	0.19 + 0.015 <sup>b</sup>	0.131 + 0.029	0.19 + 0.013	0.35 + 0.02 <sup>b</sup>
<i>savignyi</i> n = 17	0.03 + 0.011	0.031 <sup>c</sup>	0.143 + 0.013	0.151 <sup>c</sup>	0.132 + 0.015	0.278 + 0.023	0.247 <sup>c</sup>

**Table 3.** Microsatellites analysis of the Tel Baruch population when divided according to the different mitotypes. <sup>a</sup>Population is located in Nitzanim<sup>53</sup>. <sup>b</sup>Population is located in Betzet (This paper). <sup>c</sup>Population is located in Arad, Rishon and Palmachim<sup>52,53</sup>.

**Behavioral assay.** Two alternative hypotheses emerged from the genetics data regarding the social structure of the Tel Baruch population: a single, highly polymorphic species, as suggested by the nuclear gene analyses; or multiple species, as suggested by the mtDNA and CHC analyses. We therefore attempted to differentiate between the two hypotheses using behavioral assays comprising the possible merging of two distinct nests. Seven merging tests were performed using various pairs of nests (Fig. 5 and Table 4). First, we used nests of the *niger* mitotype located at distances between 100 and 500 meters from one another. In all cases the nests merged after a few hours, during which we observed both trophallaxis and mutual carrying between workers from the different nests. After 24 hours, the nests had fully merged, with the queens and brood having been moved into one of the nest boxes. All other pairing combinations, i.e., *drusus-niger*, *drusus-savignyi*, *savignyi-savignyi*, and *savignyi-niger*, resulted in aggressive behavior between the workers and none of these nests merged (Fig. 5).

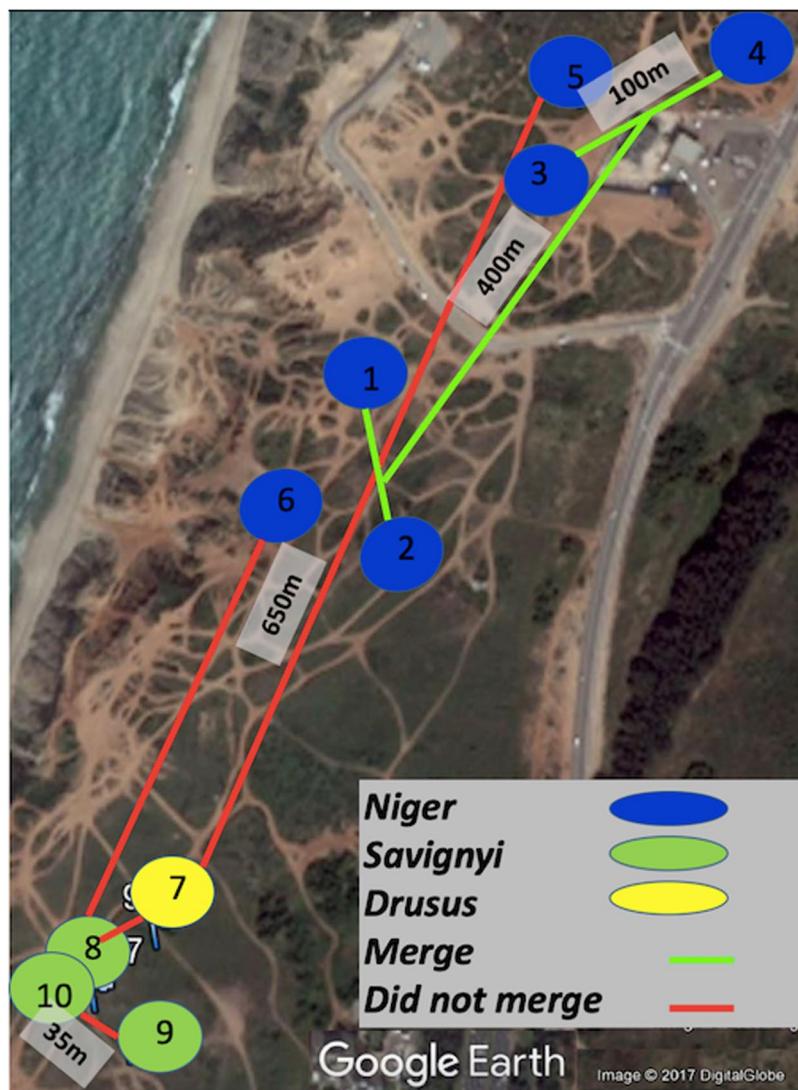
**Sperm analysis.** To examine whether assortative mating between the different mitotypes occurs, we genotyped, using CytB, the sperm extracted from the spermatheca of ten queens of the *niger* mitotype. Nine out of the ten sperm samples matched the *savignyi* mitotype and one matched the *niger* mitotype (Fig. S4). In some of the sequences there appeared to be more than one possibility of a nucleotide in certain positions, possibly indicating multiple sequences/mitotypes. This suggests that multiple males of multiple mitotypes inseminated a specific queen.

## Discussion

A recent phylogeographic study of species delimitation in the *bicolor* group of the genus *Cataglyphis* in Israel<sup>29</sup> pointed out discrepancies between mitochondrial DNA and nuclear DNA sequences within the putative species *C. niger*, *C. savignyi*, and *C. drusus*, i.e., the *C. niger* complex. The three mitotypes also differ in their social structure (Table 5). The *savignyi* mitotype is monogyne demonstrating either monodomous<sup>53</sup> or polydomous (<sup>47</sup> mistakenly identified as *niger* when published) nesting behavior; whereas the *niger* mitotype is polygyne and is organized as a supercolony<sup>53</sup>. The social and population structure of the *drusus* mitotype has not previously been investigated. Furthermore, this study revealed that all three mitotypes occur sympatrically in the Tel Baruch site. Therefore, disentangling the species identity of this complex is expected to provide insight not only into the process of sympatric speciation, but also into the evolution of advanced social structures in ants.

First, we sought to rule out the possibility that the amplicons of the nuclear genes employed in Eyer *et al.*<sup>29</sup> were not sensitive enough to detect differences between the three mitotype. We therefore conducted RAD sequencing on representatives of the above complex, along with samples from most of the locations described in Eyer *et al.*<sup>29</sup>, comprising other species, namely *C. holgerseni*, *C. isis*, *C. israelensis* and *C. nodus* (collected in Turkey). The RAD results were compatible with the nDNA data of the Eyer *et al.*<sup>29</sup> study, and ruling out the possibility of errors due to the limited sequence information for the study. *STRUCTURE* results confirmed the existence of five bicolor species in Israel, and most importantly confirmed the existence of gene flow between the three mitotypes of the *C. niger* complex (i.e. *niger*, *savignyi*, and *drusus*), suggesting that they constitute a single species. The results also indicate hybrid zones between *C. israelensis* and the *savignyi* mitotype of the *C. niger* complex.

The case of the *C. niger* complex is interesting and somewhat enigmatic. The three mitotypes have a peculiar distribution along the coastal plains of Israel<sup>32</sup>. The northern part, from the Lebanese border to Mount Carmel, is inhabited exclusively by the *drusus* mitotype. To the south of Mount Carmel, the *savignyi* mitotype predominates, except for a patchy distribution of another species, *C. israelensis*<sup>29,54</sup>. Curiously, at the border of their distribution, we found hybrids between *C. israelensis* and the *savignyi* mitotype of the *C. niger* complex. The *savignyi* mitotype has the largest distribution. It occupies a large part of the sandy areas along the coast and it dominates in areas of heavier soil, extending to the more desert areas of Israel. The *niger* mitotype, in contrast, is present only in the central coastal plain, where it has a patchy occurrence. Thus, at least in a large proportion of their distribution, although the *drusus* and *savignyi* mitotypes are allopatric, the nuclear sequencing data indicates free gene flow. This raises the question of whether we are encountering three incipient species that demonstrates rapid mitochondrial evolution but slower nuclear evolution; or a single species with long-range dispersal of sexuals (mostly through males because *Cataglyphis* females are generally known to have lower flying capability<sup>55</sup>), to account for the apparent gene flow among the highly allopatric populations.



**Figure 5.** Location of nests used for the merging experiment. Nests are colored according to the mitotype: blue- *niger*, green- *savignyi*, and yellow- *drusus*. The lines drawn between nests indicate if the nests merged (green) or not (red).

Nest A (mitochondrial haplotype)	Nest B (mitochondrial haplotype)	Merging	Between- colony trophallaxis	Between- colony aggression	Maximal distance between colonies (meters)
1 ( <i>Niger</i> )	2 ( <i>Niger</i> )	+	+	-	100
3 ( <i>Niger</i> )	4 ( <i>Niger</i> )	+	+	-	100
1+2 ( <i>Niger</i> )	3+4 ( <i>Niger</i> )	+	+	-	400
5 ( <i>Niger</i> )	7 ( <i>Drusus</i> )	-	-	+	650
6 ( <i>Niger</i> )	8 ( <i>Savignyi</i> )	-	-	+	450
7 ( <i>Drusus</i> )	8 ( <i>Savignyi</i> )	-	-	+	35
9 ( <i>Savignyi</i> )	10 ( <i>Savignyi</i> )	-	-	+	35

**Table 4.** Merging assays of the Tel Baruch population.

The sympatric cooccurrence of all three mitotypes in the Tel Baruch population enables us to investigate the differences among the complex mitotypes within a natural setting without a confounding ecological variation. The microsatellite analysis reveals an *Fst* value of 0.12 across all the nests of the Tel-Baruch population, suggesting some level of diversification in the highly polymorphic markers<sup>55,56</sup>. However, this estimate

Species		Mating system <sup>a</sup>	Queen number <sup>b</sup>	Thelytoky <sup>c</sup>		Isolation by Distance (based on nDNA)	Reference(s)
				Workers	Queens		
<i>Niger complex</i>	<i>C. niger</i>	m	P	No	No	Yes (nDNA)	<a href="#">53,63</a> , This study
	<i>C. savignyi</i>	m	M	Yes	No	—	<a href="#">52,53</a> , This study
	<i>C. drusus</i>	m	M	No	No	No (nDNA)	This study
<i>C. israelensis</i>		m	M	No	No	Yes (nDNA)	<a href="#">32</a>
<i>C. viatica</i>		m	M	Yes	No	—	<a href="#">69,70</a>

**Table 5.** Social traits in *Cataglyphis bicolor* species group. Abbreviations: m, obligatory multiple mated (paternity frequency  $\geq 2$ , and often  $\geq 4$ ); M, monogyny (one queen); (—) no data available; nDNA, nuclear microsatellite DNA; P, polygyny (several queens). <sup>a</sup>Mating frequency is determined from the arithmetic mean number of patrilines found from mother-offspring genetic combinations and/or from direct genotyping of sperm stored in the spermatheca of the queens. <sup>b</sup>Queen number is determined from field observations and/or from workers genotypes. <sup>c</sup> Thelytoky: occurrence of (yes) or failure to detect (no) thelytokous parthenogenesis by workers and queens (i.e., asexual production of new daughter queens). Thelytokous parthenogenesis is determined from laboratory observations (i.e., worker-produced females) and from queen-daughter genetic combinations. Worker reproduction usually occurs in queenless colonies.

might not be accurate because some nests belong to the same colony. As reported previously, the *niger* mitotype is polygynous and supercolonial<sup>53</sup>. Here, we further established that this is an exclusively *niger* mitotype characteristic. Both *savignyi* and *drusus* at the study sites were monogynous and facultatively polydomous. The three mitotypes inhabiting the same plot were also distinctively different in their CHC composition. CHC's are the main mechanism through which ants maintain nest insularity<sup>57–59</sup>, thus making it highly important in such a diverse population. Within each of the mitotypes, differences in CHC profiles were rather small, consistent with populations exhibiting limited dispersal. Thus, the question of whether we have 3 incipient species (as indicated by morphology<sup>33</sup>, mitochondrial DNA, chemistry, and behavioral data), or a single species (as indicated by the nDNA genotyping) that is highly polymorphic (morphologically, chemically and to a certain degree genetically), remains open. While this does not pose a problem regarding mitotype stability at the site (being maternally inherited), how the social structure (i.e., polygyny and supercoloniality) is maintained remains elusive.

According to the incipient species hypothesis, we expected assortative mating according to the different mitotypes. However, sperm analysis of mated gynes of the *niger* mitotype revealed a high presence of the *savignyi* mitotype (matching its abundance at the site), indicating the absence of mitotype-associated assortative mating. Such non-assortative mating can be the result of either recent admixture of the 3 species due to anthropogenic-associated geographic barrier destruction, leading the 3 species to merge into a single species, or supporting the single species hypothesis having 3 mitochondrial lineages, each of which associated with different social structure. Cross-mating between lineages within a species has been described in *Pogonomyrmex*, resulting in caste differentiation: eggs fertilized by males of the queen's lineage develop into new queens, while eggs fertilized by males of the opposite lineage develop into sterile workers<sup>60</sup>. Another report of cross-mating between two social forms is that for *Solenopsis invicta*<sup>61,62</sup>, in which polygynous queens mate with males from monogynous colonies. In *Cataglyphis*, several species of the closely-related *C. altisquamis* group exhibit cross-mating between two hybridogenetic lineages, which facilitates worker production and genetic diversity (*C. velox*, *C. altisquamis*, *C. hispanica* and *C. mauritanica*)<sup>24,27,28</sup>. Here we showed that the polygynous *niger* queens partake in nuptial flights, which enables mating with males of the *savignyi* mitotype. However, since within-colony genetic diversity is satisfied in the case of non-related co-inhabiting queens<sup>63</sup>, it is not clear whether there is a genetic necessity for cross-mitotype mating, paving the way for the evolution of assortative mating among mitotypes. In other *Cataglyphis* polygynous species, gyne morphology reveals the loss of wing muscles<sup>64</sup>, confining the queens to mating near or within the nest. Such a change in morphology would lead to mating place separation and sympatric speciation.

Another possible mechanism that could explain the social polymorphism exhibited by the *niger* complex, if considering the complex as a single species, is the occurrence of a non-combining 'social chromosome', as described in two other ant species, *Solenopsis invicta* and *Formica selysi*<sup>20,21</sup>. The independent evolution of a 'social-chromosome' mechanism may suggest that such a mechanism may also have evolved in *Cataglyphis*. However, the apparent differences in the mitochondrial DNA may indicate a maternal inheritance of these social traits as well, which raises the possibility of alternative explanations, not previously described in ants, such as a maternal genetic imprinting on social traits.

When considering possible mechanisms and road maps for sympatric speciation, we assume traits evolving in a single population in different directions. These traits should have major life-history effects that would drive for assortative mating, stopping gene flow and promoting speciation. Social behavior and recognition cues are such traits. CHC diversification has been shown to result in speciation through pre-mating isolation across phyla<sup>65</sup>. Likewise, differences in social forms can cause isolation through both pre-mating and post mating processes<sup>66,67</sup>. The clear separation in mDNA shows the lack of female gene flow between social forms suggesting the initiation of species diversification<sup>68</sup>. The *C. niger complex*, with its differences in CHC composition, social structures and mDNA between the three incipient species, exemplifies a population in which some major traits have diverged while others have not, raising the possibility of a sympatric speciation continuum based on a differentiation in sociality.

## References

- Mayr, E. The biological species concept. Species concepts phylogenetic theory a debate. *Columbia Univ.* (pp. 17–29. Press, New York, 2000).
- Seifert, B. A pragmatic species concept applicable to all eukaryotic organisms independent from their mode of reproduction or evolutionary history. *Soil Org.* **86**, 85–93 (2014).
- De Queiroz, K. Species concepts and species delimitation. *Syst. Biol.* **56**, 879–886 (2007).
- Bolnick, D. I. & Fitzpatrick, B. M. Sympatric Speciation: Models and Empirical Evidence. *Annu. Rev. Ecol. Evol. Syst.* **38**, 459–487 (2007).
- Feder, J. L., Egan, S. P. & Nosil, P. The genomics of speciation-with-gene-flow. *Trends Genet.* **28**, 342–350 (2012).
- Nosil, P. Speciation with gene flow could be common. *Mol. Ecol.* **17**, 2103–2106 (2008).
- Matsubayashi, K. W., Ohshima, I. & Nosil, P. Ecological speciation in phytophagous insects. *Entomol. Exp. Appl.* **134**, 1–27 (2010).
- Nosil, P., Harmon, L. J. & Seehausen, O. Ecological explanations for (incomplete) speciation. *Trends Ecol. Evol.* **24**, 145–156 (2009).
- Cognate, A. I., Seybold, S. J. & Sperling, F. A. H. Incomplete barriers to mitochondrial gene flow between pheromone races of the North American pine engraver, *Ips pini* (Say) (Coleoptera, Scolytidae). *Proc. R. Soc. London B Biol. Sci.* **266**, 1843–1850 (1999).
- Hey, J. & Pinho, C. Population genetics and objectivity in species diagnosis. *Evolution (N. Y.)* **66**, 1413–1429 (2012).
- Fisher, C., Robinson, C. T. & Malard, F. Cryptic species as a window into the paradigm shift of the species concept. *Mol. Ecol.* **27**, 613–635 (2018).
- Ronque, M. U. V., Azevedo-Silva, M., Mori, G. M., Souza, A. P. & Oliveira, P. S. Three ways to distinguish species: Using behavioural, ecological, and molecular data to tell apart two closely related ants, *Camponotus rengersi* and *Camponotus rufipes* (Hymenoptera: Formicidae). *Zool. J. Linn. Soc.* **176**, 170–181 (2016).
- Ross, K. G., Gotzek, D., Ascunce, M. S. & Shoemaker, D. D. Species delimitation: A case study in a problematic ant taxon. *Syst. Biol.* **59**, 162–184 (2010).
- Kather, R. & Martin, S. J. Cuticular hydrocarbon profiles as a taxonomic tool: Advantages, limitations and technical aspects. *Physiol. Entomol.* **37**, 25–32 (2012).
- Seifert, B. Cryptic species in ants (Hymenoptera: Formicidae) revisited: we need a change in the alpha taxonomic approach. *Myrmecol. Nachr.* **12**, 149–166 (2009).
- Seifert, B. Analyzing large-scale and intranidal phenotype distributions in eusocial Hymenoptera - A taxonomic tool to distinguish intraspecific dimorphism from heterospecificity. *Myrmecological News* **23**, 41–59 (2016).
- Arnqvist, G., Edvardsson, M., Friberg, U. & Nilsson, T. Sexual conflict promotes speciation in insects. *Proc. Natl. Acad. Sci.* **97**, 10460–10464 (2000).
- Torres, C. W., Tonione, M. A., Ramírez, S. R., Sapp, J. R. & Tsutsui, N. D. Genetic and chemical divergence among host races of a socially parasitic ant. *Ecol. Evol.* **0** (2018).
- Seifert, B. Intranidal mating, gyne polymorphism, polygyny, and supercoloniality as factors for sympatric and parapatric speciation in ants. *Ecol. Entomol.* **35**, 33–40 (2010).
- Wang, J. *et al.* A Y-like social chromosome causes alternative colony organization in fire ants. *Nature* **493**, 664–668 (2013).
- Purcell, J., Brelsford, A., Wurm, Y., Perrin, N. & Chapuisat, M. Convergent Genetic Architecture Underlies Social Organization in Ants. *Curr. Biol.* **24**, 2728–2732 (2014).
- Taylor, S. & Campagna, L. Avian supergenes. *Science (80-)* **351**, 446 LP–447 (2016).
- Agosti, D. Review and reclassification of Cataglyphis (Hymenoptera, Formicidae). *J. Nat. Hist.* **24**, 1457–1505 (1990).
- Aron, S., Mardulyn, P. & Leniaud, L. Evolution of reproductive traits in Cataglyphis desert ants: mating frequency, queen number, and thelytoky. *Behav. Ecol. Sociobiol.* **70**, 1367–1379 (2016).
- Eyer, P.-A., Leniaud, L., Darras, H. & Aron, S. Hybridogenesis through thelytokous parthenogenesis in two Cataglyphis desert ants. *Mol. Ecol.* **22**, 947–955 (2013).
- Darras, H., Kuhn, A. & Aron, S. Genetic determination of female castes in a hybridogenetic desert ant. *J. Evol. Biol.* **27**, 2265–2271 (2014).
- Leniaud, L., Darras, H., Boulay, R. & Aron, S. Social hybridogenesis in the clonal ant Cataglyphis hispanica. *Curr. Biol.* **22**, 1188–1193 (2012).
- Boulay, R. *et al.* Social Life in Arid Environments: The Case Study of Cataglyphis Ants. *Annu. Rev. Entomol.* **62**, 305–321 (2017).
- Eyer, P. A., Seltzer, R., Reiner-Brodetzki, T. & Hefetz, A. An integrative approach to untangling species delimitation in the Cataglyphis bicolor desert ant complex in Israel. *Mol. Phylogenet. Evol.* **115**, 128–139 (2017).
- Eyer, P. A. & Hefetz, A. Cytonuclear incongruences hamper species delimitation in the socially polymorphic desert ants of the Cataglyphis albicans group in Israel. *Journal of Evolutionary Biology*, <https://doi.org/10.1111/jeb.13378> (2018).
- Radchenko, A. G. The phylogeny and faunogenesis of the genus Cataglyphis Foerster (Hymenoptera, Formicidae). *Entomol. Rev.* **81**, 951–958 (2001).
- Reiner Brodetzki, T. & Hefetz, A. Determining social and population structures requires multiple approaches: A case study of the desert ant Cataglyphis israelensis. *Ecol. Evol.* **0**, ece3.4535 (2018).
- Simon, C. *et al.* Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. *Ann. Entomol. Soc. Am.* **87**, 651–701 (1994).
- Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797 (2004).
- Parchman, T. L. *et al.* Genome-wide association genetics of an adaptive trait in lodgepole pine. *Mol. Ecol.* **21**, 2991–3005 (2012).
- Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S. & Hoekstra, H. E. Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS One* **7**, e37135 (2012).
- Catchen, J. M., Amores, A., Hohenlohe, P., Cresko, W. & Postlethwait, J. H. Stacks: Building and Genotyping Loci De Novo From Short-Read Sequences. *G3&#39;Genes|Genomes|Genetics* **1**, 171–182 (2011).
- Catchen, J. M. Stacks: an analysis tool set for population genomics. *Mol. Ecol.* **22**, 3124–3140 (2013).
- Alexander, D. H., Novembre, J. & Lange, K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* **19**, 1655–1664 (2009).
- Pritchard, J. K., Stephens, M. & Donnelly, P. Inference of population structure using multilocus genotype data. *Genetics* **155**, 945–959 (2000).
- Falush, D., Stephens, M. & Pritchard, J. K. Inference of population structure using multilocus genotype data: Dominant markers and null alleles. *Mol. Ecol. Notes* **7**, 574–578 (2007).
- Kopelman, N. M., Mayzel, J., Jakobsson, M., Rosenberg, N. A. & Mayrose, I. Clumpak: A program for identifying clustering modes and packaging population structure inferences across K. *Mol. Ecol. Resour.* **15**, 1179–1191 (2015).
- Yang, Z. Likelihood and Bayes estimation of ancestral population sizes in hominoids using data from multiple loci. *Genetics* **162**, 1811–1823 (2002).
- Zhu, T. & Yang, Z. Maximum likelihood implementation of an isolation-with-migration model with three species for testing speciation with gene flow. *Mol. Biol. Evol.* **29**, 3131–3142 (2012).
- Dalquen, D. A., Zhu, T. & Yang, Z. Maximum likelihood implementation of an isolation-with-migration model for three species. *Syst. Biol.* **66**, 379–398 (2017).
- Raymond, M. & Rousset, F. GENEPOLP on the Web (Version 3.4). URL, <http://wbiomed.curtin.edu.au/genepop/> Updated from Raymond Rousset (1995).
- Goudet, J. FSTAT (version 1.2): a computer program to calculate F-statistics. *J. Hered.* **86**, 485–486 (1995).
- Jones, O. R. & Wang, J. COLONY: a program for parentage and sibship inference from multilocus genotype data. *Mol. Ecol. Resour.* **10**, (551–555 (2010).

49. Earl, D. A. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv. Genet. Resour.* **4**, 359–361 (2012).
50. Ryan, S. F. *et al.* Patterns of divergence across the geographic and genomic landscape of a butterfly hybrid zone associated with a climatic gradient. *Mol. Ecol.* (2017).
51. Wang, J. COANCESTRY: a program for simulating, estimating and analysing relatedness and inbreeding coefficients. *Mol. Ecol. Resour.* **11**, 141–145 (2011).
52. Saar, M., Leniaud, L., Aron, S. & Hefetz, A. At the brink of supercoloniality: genetic, behavioral, and chemical assessments of population structure of the desert ant *Cataglyphis niger*. *Front. Ecol. Evol.* **2**, 1–10 (2014).
53. Leniaud, L., Hefetz, A., Grumiau, L. & Aron, S. Multiple mating and supercoloniality in *Cataglyphis* desert ants. *Biol. J. Linn. Soc.* **104**, 866–876 (2011).
54. Ionesco, A. & Eyer, P. A. Notes on *Cataglyphis* Foerster, 1850 of the bicolor species-group in Israel, with description of a new species (Hymenoptera: Formicidae). *Isr. J. Entomol.* **46**, 109–131 (2016).
55. Leppänen, J., Seppä, P., Vepsäläinen, K. & Savolainen, R. Genetic divergence between the sympatric queen morphs of the ant *Myrmica rubra*. *Mol. Ecol.* **24** (2015).
56. Wondji, C., Simard, F. & Fontenille, D. Evidence for genetic differentiation between the molecular forms M and S within the Forest chromosomal form of *Anopheles gambiae* in an area of sympatry. *Insect Mol. Biol.* **11**, 11–19 (2002).
57. Soroker, V., Vienne, C., Hefetz, A. & Nowbahari, E. The postpharyngeal gland as a “Gestalt” organ for nestmate recognition in the ant *Cataglyphis niger*. *Naturwissenschaften* **81**, 510–513 (1994).
58. Hefetz, A. The evolution of hydrocarbon pheromone parsimony in ants (Hymenoptera: Formicidae) – interplay of colony odor uniformity and odor idiosyncrasy. A review. *Regulation* **10**, 59–68 (2007).
59. Kidokoro-Kobayashi, M. *et al.* Chemical Discrimination and Aggressiveness via Cuticular Hydrocarbons in a Supercolony-Forming Ant, *Formica yessensis*. *PLOS One* **7** (2012).
60. Schwander, T., Keller, L. & Cahan, S. H. Two alternate mechanisms contribute to the persistence of interdependent lineages in *Pogonomyrmex* harvester ants. *Mol. Ecol.* **16**, 3533–3543 (2007).
61. Ross, K. G. & Shoemaker, D. D. Nuclear and mitochondrial genetic structure in two social forms of the fire ant *Solenopsis invicta*: Insights into transitions to an alternate social organization. *Heredity (Edinb.)* **78**, 590–602 (1997).
62. Shoemaker, D. D. & Ross, K. G. Effects of social organization on gene flow in the fire ant *Solenopsis invicta*. *Nature* **383**, 613–616 (1996).
63. Reiner Brodetski, T., Brodetski, G., Feinerman, O. & Hefetz, A. Inside the super-organism; Behavior and demography of the supercolonial ant *Cataglyphis niger*. *bioRxiv* 392522 (2018).
64. Peeters, C. & Aron, S. Evolutionary reduction of female dispersal in *Cataglyphis* desert ants. *Biol. J. Linn. Soc.* **122**, 58–70 (2017).
65. Smadja, C. & Butlin, R. K. On the scent of speciation: The chemosensory system and its role in premating isolation. *Heredity (Edinb.)* **102**, 77–97 (2009).
66. Goropashnaya, A. V., Fedorov, V. B. & Pamilo, P. Recent speciation in the *Formica rufa* group ants (Hymenoptera, Formicidae): Inference from mitochondrial DNA phylogeny. *Mol. Phylogenet. Evol.* **32**, 198–206 (2004).
67. Avril, A., Purcell, J., Brellsford, A. & Chapuisat, M. Asymmetric assortative mating and queen polyandry are linked to a supergene controlling ant social organization. *Mol. Ecol.* **0** (2018).
68. Gyllenstrand, N., Seppä, P. & Pamilo, P. Restricted gene flow between two social forms in the ant *Formica truncorum*. *J. Evol. Biol.* **18**, 978–984 (2005).
69. Aron, S., Darras, H., Eyer, P.-A., Leniaud, L. & Pearcy, M. Structure génétique des sociétés et systèmes d'accouplement chez la fourmi *Cataglyphis viatica* (Fabricius 1787). *Bull. l'Institut Sci. Rabat* **35**, 103–109 (2013).
70. Pearcy, M., Clémencet, J., Chameron, S., Aron, S. & Doums, C. Characterization of nuclear DNA microsatellite markers in the ant *Cataglyphis cursor*. *Mol. Ecol. Resour.* **4**, 642–644 (2004).

## Acknowledgements

This research was funded by the Israel Science foundation (ISF grant # 844/13 to AH). The research was also supported by the Naomi Prawer Kadar Foundation through the Tel Aviv University GRTF Program (travel grant to TRB). We would like to thank Raya Zeltzer, Maayan Drin, and PA Eyer for their help with sample collection; Guy Sobol, Yafit Brenner, and Tovit Simon, for technical assistance and data process; Laurent Grumiau and Hugo Darras for help with genetic data; and Naomi Paz for editing.

## Author Contributions

T.R.B., S.A. and A.H. designed the project. T.R.B., P.C., S.I. and E.P. conducted the ddRAD analysis. T.R.B. and A.H. conducted the C.H.C. analysis. T.R.B. and S.A. conducted the microsatellite analysis. T.R.B. conducted the behavioral and mtDNA analysis. T.R.B. and A.H. lead the writing of the manuscript.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-019-45950-1>.

**Competing Interests:** The authors declare no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.