MOLECULAR ECOLOGY

Molecular Ecology (2014) 23, 2529–2542 doi: 10.1111/mec.12749

Recent speciation and secondary contact in endemic ants

MICHAEL J. JOWERS,*†[1] FERNANDO AMOR,*[1] PATROCINIO ORTEGA,* ALAIN LENOIR,‡ RAPHAËL R. BOULAY,†‡ XIM CERDÁ,* and JUAN A. GALARZA,§

*Estación Biológica de Doñana (CSIC), Av. Américo Vespucio, 41092 Sevilla, Spain, †Departamento de Biología Animal, Facultad de Ciencias, Universidad de Granada, 18071 Granada, Spain, ‡IRBI, UMR CNRS 6035, Faculté des Sciences, Université François Rabelais, 37200 Tours, France, §Centre of Excellence in Biological Interactions, Department of Biological and Environmental Science, University of Jyväskylä, Survontie 9, 40500 Jyväskylä, Finland

Abstract

Gene flow is the main force opposing divergent selection, and its effects are greater in populations in close proximity. Thus, complete reproductive isolation between parapatric populations is not expected, particularly in the absence of ecological adaptation and sharp environmental differences. Here, we explore the biogeographical patterns of an endemic ant species, *Cataglyphis floricola*, for which two colour morphs (black and bicolour) coexist in parapatry throughout continuous sandy habitat in southern Spain. Discriminant analyses of six biometric measurements of male genitalia and 27 cuticular hydrocarbons reveal high differentiation between morphs. Furthermore, the low number of shared alleles derived from nuclear markers (microsatellites) between the morphs at their contact zone suggests the absence of recent gene flow. Mitochondrial DNA (COI) phylogenetic analysis and median-joining networks show that the black morph is basal to the bicolour morph, with unique haplotypes recovered for each morph. Mismatch distribution analysis and Bayesian skyline plots suggest that they are undergoing different demographic changes, with the bicolour and black morphs at demographic equilibrium and expansion, respectively. Thus, our results show complete reproductive isolation between the two colour morphs as evidenced from genetic, chemical and morphological data. We suggest that these divergence events could be explained by historical vicariance during the Pleistocene, in which reproductive traits experienced strong divergent selection between the morphs initiating or culminating speciation.

Keywords: behaviour/social evolution, insects, phylogeography, population genetics, speciation

Received 31 December 2012; revision received 27 March 2014; accepted 27 March 2014

Introduction

Defining species and understanding how they are formed is of paramount importance to biodiversity science and the field of evolutionary biology (Seifert 2009). After years of debate on the necessary criteria to define species, the unified species concept proposes to consider separately evolving lineages as different species (de Queiroz 2007). Thus, a species can be viewed as the largest taxon in which gene flow and allele expansion are possible, defining a unit of evolution (Wiley 1977). Yet, delimiting a species unit is generally based on phenetic and/or phylogenetic relationships. In this regard, geographical variation in phenotypes is often linked to speciation processes (Endler 1977). This is particularly true when colour polymorphism is observed across a species' distribution (Gray & McKinnon 2007). Thus, the study of colour polymorphisms can yield valuable information about possible speciation processes.

The formation of new species can be driven by both ecological and nonecological processes. Nonecological speciation derives from the action of natural selection on pleiotropic genes leading indirectly to reproductive isolation (Bradshaw & Schemske 2003), assortative mating and random genetic drift (Kimura & Ohta 1971; Rundle & Nosil 2005; Nosil *et al.* 2008). Ecological speci-

Correspondence: Juan A. Galarza, Fax: +358 (0) 14 617 239; E-mail: juan.galarza@jyu.fi

[1]These authors contributed equally to this work.

ation, on the other hand, is driven by divergent natural selection, promoted by environmental differences and local adaptation, resulting in reproductive isolation (Rundle & Nosil 2005). Gene flow is one of the main opposing forces of divergent selection, and its effects are expected to be greater between nearby populations rather than between distant ones. Therefore, divergent selection increases as a function of local adaptation to different environments and the geographical distance between them (Rice & Hostert 1993; Schluter 2009). Thus, complete reproductive isolation is not expected between neighbouring populations, particularly in the absence of obvious barriers to dispersal (Turelli *et al.* 2001).

Cryptic species seem to be particularly frequent in some groups of animals such as ants (Seifert 2009). Their identification using morphology-based approaches is limited by the great level of adaptive polymorphism within the same nest. For instance, individuals with different behavioural specializations within the same nest may have very different morphologies. Similarly, the range of colour variation is often considered to be too high between individuals, nests and populations to be reliably used for species diagnosis. Therefore, an integrated approach using data from multiple independent sources seems necessary to confidently assess possible speciation events in polymorphic species.

The ant *Cataglyphis floricola* Tinaut 1993 is endemic to the Doñana National Park (hereafter, DNP) and its surroundings (Southwestern Spain, Fig. 1). Two colour morphs, black and black/red bicolour (Fig. S1, Supporting information), coexist over what seems a homogeneous habitat in this restricted area. Black and red are frequent colours for ants, but their adaptive significance is not known for this species. Thus far, both morphs have been regarded as a single species (Tinaut 1993). However, the possibility that they might constitute two different species that have recently diverged cannot be ruled out. The present-day topography of the area presents no apparent barriers to gene flow, which could indicate an ongoing speciation event (i.e. in parapatry). However, geological data shows that the distribution area has suffered important geomorphological modifications since the Pleistocene, with significant alterations to the coastal area (Zazo *et al.* 2005). This could suggest a single recent speciation event in time (i.e. in allopatry) due to past geological processes absent today.

Here, we investigate these hypotheses through an integrative approach consisting of morphological, genetic and chemical data. We first determined the whole distribution area of both morphs through field surveys. We then assessed the phylogenetic relationship between populations of both morphs across their distribution range using mitochondrial DNA. Subsequently, we tested for reproductive isolation between populations located in the contact zone of both morphs by means of nuclear microsatellite data. Finally, we analysed differences in traits (male genitalia and cuticular hydrocarbons) of both morphs that have previously been considered important to separate cryptic ant species (Tinaut 1993; Dahbi *et al.* 1996). We discuss the importance of such an integrated approach, not only to understand the mechanisms of speciation, but also for

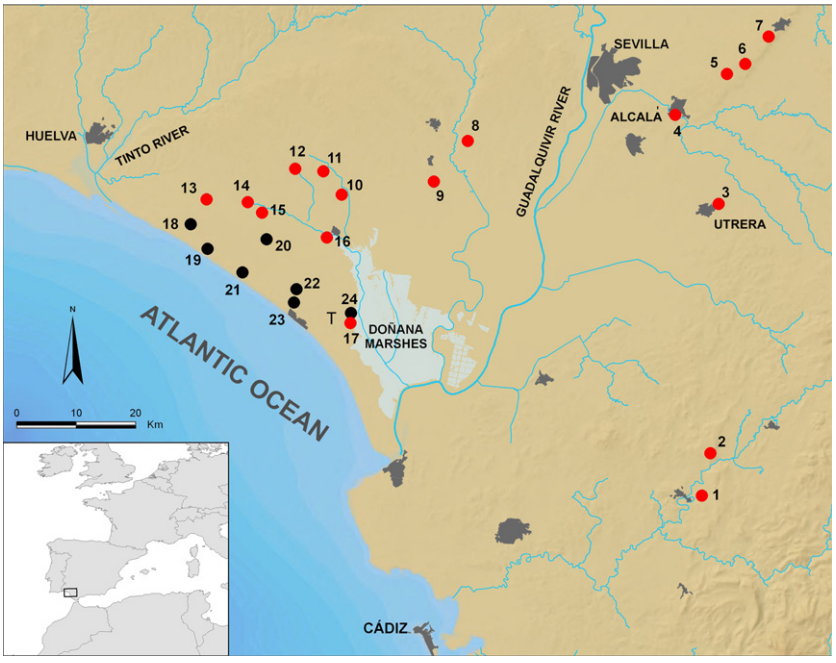

Fig. 1 Distribution of *Cataglyphis floricola*. Red and black points indicate bicolour and black morph nest presence, respectively. The numbering corresponds to the populations used in phylogeographical analysis. 'T' corresponds to the transect locality.

the conservation of endemic and cryptic species that occupy habitats threatened by fragmentation.

Materials and methods

Distribution survey

Cataglyphis floricola is a strictly monogynous species that reproduces by colony fission where the young queen is transported by workers to a new nest located at a short distance to found a new colony (Amor *et al.* 2011). Long-distance dispersal occurs in males, while females are philopatric.

To assess the distribution of both morphs, we surveyed 400 locations between 2006 and 2011 in the South of the Iberian Peninsula. All localities were distributed as regularly as possible in a triangle of approximately 210 km per side, which vertices were 37°54′N, 4°23′W; 37°05′N, 8°15′W; and 36°01′N, 5°36′W, with the Guadalquivir River in its centre (Figs 1 and S3, Supporting Information). Localities were visited during the hours of maximum activity of *C. floricola*. The presence of the species was checked by two persons searching for foragers in periods of approximately 10 min. All GPS locations were systematically recorded.

To identify possible interactions between the morphs, a transect (1056 m) was outlaid in 2008 at DNP where both morphs where both morphs were found living in parapatry. The transect was divided into seven approximately equidistant points (mean 176 m) (Fig. S2, Supporting information). The three eastern points (P1, P2 and P3) contained only nests of the bicolour morph, while the three western points (P5, P6 and P7) contained only the black morph. In the middle of the transect, P4 contained nests of both morphs. Between two and five nests were sampled from each point, with the exception of the central point where 10 nests were sampled, five from each colour morph. To avoid taking samples from nests belonging to a possible fission event (average distance between fission nests in *C. floricola*: 7.7 ± 0.9 m, Amor *et al.* 2011), the sampling distance between nests was up to 19.5 m. A total of 58 and 60 bicolour and black workers, respectively, were collected from 30 different nests (15 nests per morph).

Sample preparation for genetic analyses

DNA was extracted from the brain and surrounding musculature from workers of both morphs. This soft tissue proved more suitable for efficient PCR amplification than other body parts containing chitin, such as legs or thorax which inhibited polymerase chain reactions (PCR). DNA extraction followed the HotShot method (Truett *et al.* 2000) and was then stored at −20 °C.

PCRs were carried out in pairs or individually for the microsatellites. Each 20 μL PCR volume (for sequencing and microsatellites) contained approx. 50 ng DNA, 200 μM of each dNTP, 0.15 μM of each primer, 2 μL 10× buffer, 0.8 μL $Mgcl_2$ and 0.1 unit of taq polymerase (Qiagen). The thermal cycle profile was as follows: an initial denaturation step of 2 min at 94 °C; 35 cycles of denaturation at 30 s at 94 °C, annealing for 30 s at 52 °C and extension for 45 s at 72 °C; and a final extension for 5 min at 72 °C. Following the PCR, excess primers and dNTPs were removed using enzymatic reaction of *E. coli* exonuclease I, Antartic phosphatase and Antartic phosphatase buffer (all New England Biolabs). Sequencing was carried out in both strands using the BigDye® Terminator v1.1 cycle sequencing kit (Applied Biosystems) according to the manufacturer's instructions. Labelled fragments were resolved on an automated A3130*xl* genetic analyzer (Applied Biosystems). Mitochondrial DNA variation was assessed from all nests sampled from the transect and from throughout the entire distribution range (Fig. 1, S2, Supporting information). The primers LCO (L) 5′-GGTCAACAAATCATAAAGATATTGG-3′ and HCO (R) 5′-TAAACTTCAGGGTGACCAAAAAATCA-3 (Folmer *et al.* 1994) amplified a 950-base pair fraction of the cytochrome oxidase subunit I. When DNAs failed to amplify with such primer combinations, we used two other primer combinations resulting in two shorter overlapping amplification fragments; primers Cflor (L) 5′-TGCAGGAACAGGATGAACAA-3′ and Cflor (R) 5′-TGGCCCATCATAAAGATGAA-3′ amplified approximately a 500-base pair fraction, while the primers LCO (L) and HCO (R) amplified approximately 700 base pairs of the COI. PCR conditions were exactly as those described for the nuclear microsatellites. Templates were sequenced on both strands, and the complementary reads were used to resolve ambiguous base-calls in SEQUENCHER version 4.9 (Gene Codes). After removing PCR primers and incomplete terminal sequences, 884 base pairs were available for analyses. All nucleotide sequences could be aligned without gaps, when translated into amino acids using the invertebrate mitochondrial code, stop codons were absent. Sequences were aligned in SEAVIEW version 4.2.11 (Gouy *et al.* 2010) under CLUSTALW2 (Larkin *et al.* 2007) with default settings. Nucleotide differences and *p*-uncorrected distances (%) analyses were calculated using MEGA version 5 (Tamura *et al.* 2011). Six microsatellite markers developed for *C. cursor* (Ccur11, Ccur 26, Ccur 51, Ccur 61, Ccur 99 and Ccur 100 Pearcy *et al.* 2004) were used to study nuclear polymorphisms in *C. floricola* of both morphs. Control for genotyping errors due to null alleles and allele dropouts was performed with Micro-checker (Van Oosterhout *et al.* 2004). Linkage disequilibrium

analyses and descriptive statistics were run in GENEPOP on the Web (Raymond & Rousset 1995).

A total of 144 workers were genotyped throughout 2008, 79 of the bicolour morph and 65 of the black morph. From these, 118 were collected from 15 bicolour (n = 60) and black (n = 58) nests, respectively, along the seven sampling points of the transect. The remaining 26 (19 bicolour and seven black) were from 24 localities representing the distributional area of the species (Figs S2 and S3, Supporting information). For the population structure analyses (mitochondrial sequences), one worker was sequenced from a single nest at each locality, with the only exception of Utrera and Alcalá de Guadaría, where two nests were sampled from each locality and one worker was sequenced from each nest (n = 2). One worker per nest was also sequenced from each of the 30 nests (15 nests for each morph) from the seven sampled points at the transect area (Fig. S2, Supporting information). Thus, for the study regarding all populations, a total of 56 workers were sequenced (30 from the transect and 26 from 24 different localities, Figs 1 and 2).

Microsatellite analyses

Descriptive statistics, namely the number of alleles, allele frequencies, observed and expected heterozygosity and Wright's *F*-statistics were computed with FSTAT (Goudet 1995) and GENEPOP on the Web (Raymond & Rousset 1995) for all samples of each morph separately. Likewise, the allelic frequency differentiation (F_{ST}) between the morphs was computed using FSTAT (Goudet 1995). The Bayesian clustering software STRUCTURE 2.1 (Pritchard *et al.* 2000) was used to infer the number of populations (K) independent of spatial sampling. Analyses were performed using the admixture model with correlated allele frequencies in 20 independent runs from K = 1 to K = 20, with a burn-in of 100 000 iterations followed by another 1 000 000 iterations. Selection of K was determined using two methods which were run in HARVEST version 0.6.1 (Earl 2011): (i) by plotting the negative log-likelihoods [(ln $P(D)$)] vs. K; and (ii) using the ΔK method described in Evanno *et al.* (2005).

Mitochondrial analyses

We employed a Mantel test to assess isolation by distance by plotting ($F_{ST}/(1-F_{ST})$) coefficients between pairs of nests against the geographical distance between them (Slatkin 1993). We used one sequence per colony (as they are monogynous) to generate the genetic distance matrix (Jukes-Cantor) for the bicolour morph. This test was not possible for the black morph due to the low sequence variability and short geographical

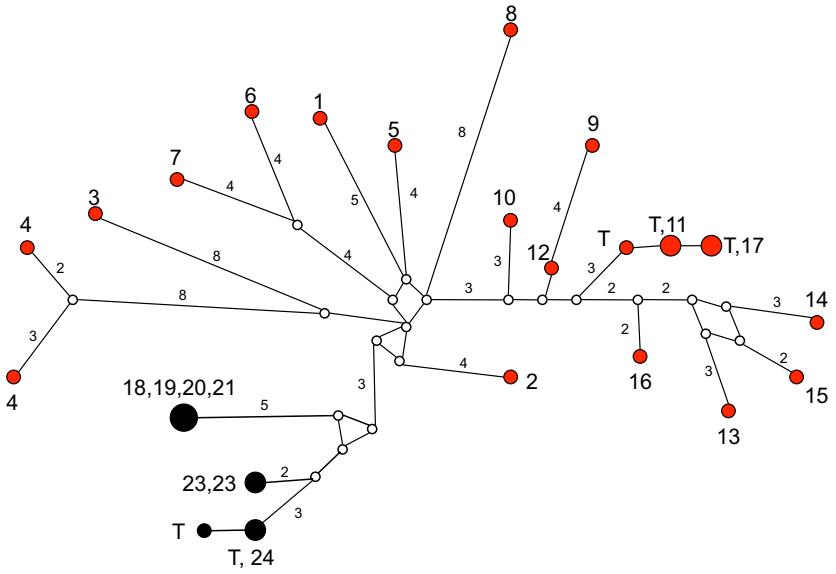

Fig. 2 Median-joining network of all *Cataglyphis floricola* haplotypes based on partial COI sequences (884 bp) showing the frequency of each haplotype. The size of the circles corresponds to the haplotype frequency (from a minimum frequency of 1, then 2 and maximum of 4). Numbers next to circles indicate the locality of the sample. Red and black circles indicate bicolour and black morph nests, respectively. Small white circles indicate an undetected intermediate haplotype state. Numbers next to lines indicate the number of mutations between haplotypes. Because of the much higher sampling effort conducted in the transect, only a maximum frequency of one was included for each of the three haplotypes recovered. Localities: Arcos de la Frontera (1–2), Utrera (3), Alcalá de Guadaíra (4), Carmona (5–7), Aznalcázar (8–9), Doñana National Park (10–24). 'T' corresponds to haplotypes recovered from the transect.

distance between colonies. The Spearman rank correlation coefficient (two-tailed) significance was evaluated through 10 000 permutations in ARLEQUIN 3.5 (Excoffier & Lischer 2010). Mitochondrial COI haplotype frequencies were estimated in DnaSP version 5.10 (Librado & Rozas 2009) using a median-joining network constructed with NETWORK 4.5 (Bandelt *et al.* 1999) setting weighs = 10 and epsilon = 0.

The most appropriate substitution model for the Bayesian inference (BI) analyses was determined by the Bayesian information criterion (BIC) in JMODELTEST version 0.1.1 (Posada 2008). The tree was constructed using the BI optimality criteria under the best fitting model (HKY+I). MrBayes (Ronquist & Huelsenbeck 2003) was used with default priors and Markov chain settings, and with random starting trees. Each run consisted of four chains of 40 000 000 generations, sampled each 10 000 generations. A plateau was reached after few generations with 25% of the trees resulting from the analyses discarded as burn-in. To assess the relationship between both morphs, an outgroup was included to root the phylogenetic tree. Four *Cataglyphis* species, believed to be among the closest relatives to *C. floricola* (Tinaut 1993*)*, were chosen to for this: *Cataglyphis emmae,* from the *C. emmae* group to which *C. floricola* belongs, and *C. rosenhaueri, C. iberica and C. velox* (Agosti 1990).

The demographic history of both morphs was examined through two separate methods. First, a mismatch distribution analysis of the COI haplotypes executed in ARLEQUIN version 3.5.1.3 (Excoffier & Lischer 2010) by generating 10 000 coalescent simulations. This analysis compares the distribution of nucleotide differences between the haplotypes against that expected for stable populations (at equilibrium with no recombination) generated through coalescent simulations. For populations at equilibrium, the haplotype differences distribution should be multimodal as the coalescent time to the most common recent ancestor is highly variable among haplotypes. On the other hand, for populations that have experienced a recent expansion, such distribution is expected to be unimodal as the time to the most common recent ancestor is similar for most haplotypes (Rogers & Harpending 1992; Harpending *et al.* 1998; Galtier *et al.* 2000). The overall validity of the analyses was tested by obtaining the probability that the observed data fits to the model using the sum of square deviations (SSD) between the observed and expected mismatch distribution as a test statistic. A *P*-value of <0.05 was taken as evidence for significant departure from the simulated distribution model. Second, as a complementary analysis to the mismatch distribution tests, for each colour morph, we estimated effective population size changes through time according to a Bayesian skyline plot method (BSP) implemented in BEAST version 1.7.0 (Drummond *et al.* 2012). We used a relaxed molecular clock and a HKY+ invariable sites distribution heterogeneity model as suggested by JMODELTEST version 0.1.1 (Posada 2008). Three independent runs were performed consisting of 4 million iteration chains, of which the first 25% was discarded as burn-in, sampling genealogies and model parameters every 1000 iterations. Convergence was deemed if the effective sample size (ESS) for the parameters and the analysis generated similar results across the runs with a minimum ESS of 200 for each parameter. The results were analysed and summarized as Bayesian skyline plots in TRACER version 1.5 (available from http://tree.bio.ed.ac.uk/software/tracer/).

Morphological analyses

The male copulatory pieces of both *C. floricola* morphs and *C. emmae* (one of its presumed closest relatives, Tinaut 1993; Tinaut, personal communication) were compared. Fifteen *C. emmae* males were collected 10 km from Amerzgane, Morocco, during May 2009–2010. Fifty-three *C. floricola* males were collected during the months of June–July 2009–2010, 23 of the black morph, collected from the DNP, and another 30 of the bicolour morph, half of which were collected at the DNP and the other half from Alcalá de Guadaíra, Seville (Fig. 1).

Dissections were carried out in Petri dishes with a Nikon SMZ800 stereomicroscope. All measurements were performed with a millimetric ruler, a digital camera Canon 30D with MP-E 65 mm macro lens and the software IMAGEJ (Abramoff *et al.* 2004). The five copulatory piece measurements were the following: (i) subgenital plate width, distance taken from between the sclerotic apical ends of its lobes; (ii) paramere length, distance between the first basal process and the paramere apex; (iii) the volsella (digitus) length, measured from the apex to the orthogonal bend, next to the origin of the cuspis; and maximum (iv) width and (v) length of the penisvalva (Boudinot 2013; Fig. S4, Supporting information). All these variables were used as explanatory variables in a linear discriminant analysis in which the geographical origin of the samples was used as a response variable in Statistica 6.0 (StatSoft).

Chemical analyses

Cuticular hydrocarbons have proven to be useful in conspecific discrimination in social insects (Dani *et al.* 2005; Guerrieri *et al.* 2009). Briefly, chains of about 20–40 hydrocarbons in the ant's cuticle encode social recognition cues, which are perceived by receptor neurons

among conspecifics (Bos *et al.* 2012). To compare the CHC composition of both *C. floricola* morphs with that of *C. emmae*, we collected a total of 15 bicolour workers from one nest near Utrera (Utr), one nest near Alcalá de Guadaíra (Alc) and two nests at the DNP, seven black workers from three nests at the DNP and four workers of *C. emmae* in two different nests near Amerzgane (Morocco). Voucher specimens are conserved at the Doñana Biological Station. Before extraction, the workers were killed by placing them at −18 °C for 20 min. The carcasses were immersed in 2 mL of hexane for at least 24 h. The CHC extracts were then stored at −18 °C until the analyses were conducted. One microlitre of each sample was injected into a gas chromatograph equipped with a mass spectrometer (Turbomass system, Perkin-Elmer, Norwalk, CT, USA, operating at 70 eV) using a nonpolar DB-5HT fused silica capillary column. Samples were run using a temperature program from 100 °C (2 min initial hold) to 320 °C at 6 °C/min with 5 min of final hold. Twenty microgram of eicosane was added to each sample before injection and used as an internal standard.

Only compounds with an average relative quantity higher than 1% of the total CHC composition in at least one species or morph were used in the statistical analyses. To reduce both the number of variables and their collinearity, a principal component analysis was first conducted on the proportion of each CHC. The first six principal components, which explained 96% of the total variance, were then used as explanatory variables in a linear discriminant analysis in which the combination of species and geographical origin was used as a response variable.

Results

Distribution results

We found a very restricted distribution of *Cataglyphis floricola* around the final stretch of the Guadalquivir River, with a parapatric distribution between the two colour morphs, and a tight contact zone between them (Fig. 1). There were differences in the distribution range of each morph and in the environmental conditions in which they occur. The bicolour morph nests appeared distributed mostly inland, in an arc of approximately 200 km long around the Guadalquivir River margins, over fine sand with high temperature and low humidity. The black morph on the other hand, appeared restricted to a narrow coastal strip approximately 35 km long and 7 km wide between two main rivers (Tinto and Guadalquivir), over loose sand and buffered temperatures, and higher humidity due to their proximity to the ocean.

Microsatellite results

None of the six microsatellites showed evidence of null alleles or linkage disequilibrium. However, two loci showed heterozygote deficits, most likely due to cross-amplification issues such as allelic dropout and stutter. A total of 68 alleles were found from both morphs at the transect (Table 1), but few (23.5%) were shared among morphs. Interestingly, none of the 19 alleles recovered from marker Ccur100 were shared between both morphs (Table S1, Supporting information). The number of alleles ranged between 7 and 19 with a mean observed heterozygosity (H_o) of 0.68 vs. 0.56 in the bicolour and black morphs, respectively (range bicolour–black: 0.62 to 0.78–0.10 to 0.88) and a mean expected heterozygosity (H_E) of 0.66 vs. 0.46 in the bicolour and black morphs, respectively (range bicolour and black: 0.55 to 0.80–0.33 to 0.74) (Table 1). Genetic differentiation was higher among nests of the black morph ($F_{ST} = 0.156$) than among nests of the bicolour morph ($F_{ST} = 0.068$). There was a significant differentiation between allelic frequencies between the morphs along the transect area ($F_{ST} = 0.261$, $P < 0.001$).

The results of the Bayesian clustering analysis of multilocus genotypes showed that the log likelihood (ln[$P(D/K)$] values peaked at $K = 17$ (Fig. S5, Supporting information). An additional *ad hoc* statistic, ΔK, which provides an improved predictor of the (K) clusters at the uppermost hierarchical level (Evanno *et al.* 2005), showed a peak at $K = 2$ (Figs 3 and S3, Supporting information), suggesting two independent evolutionary units present in the data set. This was concordant with the transect's MJ network (Fig. 2), with a clear haplotypic differentiation between the morphs.

Table 1 *Cataglyphis floricola* microsatellite descriptive statistics of both morphs

		Bicolour $N_{ind} = 58$		Black $N_{ind} = 60$	
Locus	N_A	H_o	H_E	H_o	H_E
Ccur11	9	0.622	0.551	0.860	0.669
Ccur26	13	0.691	0.764	0.208*	0.472
Ccur51	9	0.672	0.589	0.590	0.685
Ccur61	11	0.784	0.804	0.727	0.720
Ccur99	7	0.645	0.668	0.098*	0.327
Ccur100	19	0.646	0.565	0.879	0.738
Overall	—	0.677	0.657	0.560	0.460

N_A, number of alleles; N_{ind}, number of genotyped individuals; H_E, expected heterozygosity; H_o, observed heterozygosity.
*not in Hardy-Weinberg equilibrium.

Mitochondrial results

The best fitting model for the BI tree was the TIM3+I (−lnL = −2798.489, BIC = 5997.261). The BI consensus phylogram recovered a clear separation between the morphs (BPP: 0.94), with the black morph basal to the bicolour morph, which in turn constituted a monophyletic clade (BPP: 0.99) (Fig. 4). Although a reciprocal monophyly was not observed, both methods based on mtDNA COI sequences, the median-joining network (Fig. 2) and BI tree (Fig. 4) recovered similar results, showing the distinction of the two morphs.

The mean genetic distance (COI sequences, *p*-uncorrected, Table 2) of all black morph haplotypes was 0.80 ± 0.17% (Mean ± SE) and much higher for the bicolour morph haplotypes (1.64 ± 0.03%). The mean genetic distance between both morphs was 1.9 ± 0.04%. The differences in nuclear genotypes at the transect concurred with the median-joining network, with a maximum of one nucleotide difference within morph haplotypes and six between them (Fig. 2). Only one haplotype was found for the bicolour morph and two for the black morph. (Fig. 2).

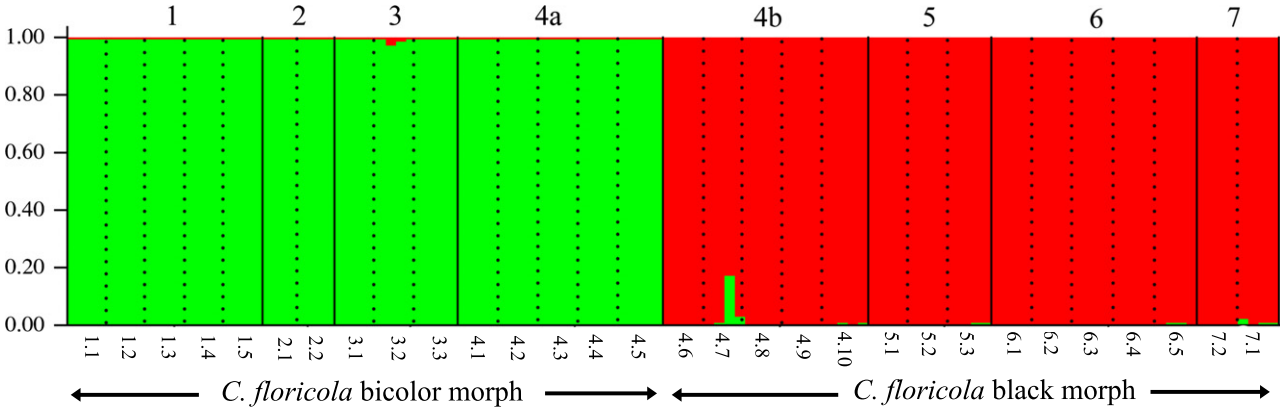

Fig. 3 Population structure of *Cataglyphis floricola* based on microsatellite data assessed by multilocus clustering using STRUCTURE following Evanno *et al*. 2005 method ($K = 2$). Each colony is represented by a bar, and its colour indicates the assigned population.

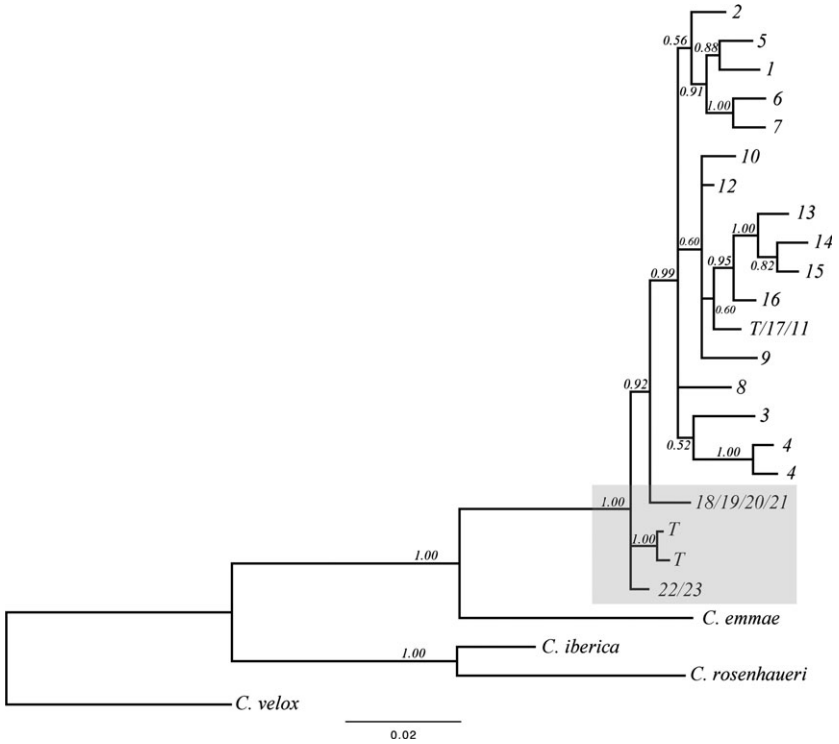

Fig. 4 Bayesian inference tree of both *Cataglyphis floricola* morphs obtained from MtDNA COI (884 bp) data. Node values indicate Bayesian inference posterior probabilities (BPP). Black morph haplotypes are in shaded area.

Table 2 Nucleotide differences (above diagonal) and uncorrected *p*-distances (below diagonal) for each pairwise comparison of all

	24/T	18–21	T	22–23	3	4	2	4	5	6	7	10
24/T	—	11	1	6	16	19	14	20	16	18	18	15
18–21	0.012	—	10	9	18	16	14	17	16	16	18	15
T	0.001	0.011	—	5	15	18	13	19	15	17	17	14
22–23	0.007	0.011	0.006	—	13	17	12	17	14	15	16	13
3	0.018	0.020	0.017	0.015	—	18	14	17	16	16	18	15
4	0.021	0.018	0.020	0.020	0.020	—	17	5	18	18	18	18
2	0.016	0.016	0.015	0.014	0.016	0.020	—	17	11	12	13	11
4	0.023	0.019	0.022	0.020	0.019	0.006	0.020	—	19	17	19	18
5	0.018	0.018	0.017	0.016	0.018	0.020	0.013	0.022	—	12	12	12
6	0.020	0.018	0.019	0.018	0.018	0.020	0.014	0.019	0.014	—	8	13
7	0.020	0.020	0.019	0.019	0.020	0.020	0.015	0.022	0.014	0.009	—	14
10	0.018	0.018	0.017	0.016	0.018	0.022	0.013	0.022	0.014	0.016	0.017	—
12	0.015	0.015	0.014	0.013	0.015	0.017	0.012	0.018	0.012	0.015	0.015	0.006
13	0.024	0.023	0.023	0.022	0.024	0.026	0.021	0.027	0.022	0.022	0.022	0.016
14	0.023	0.021	0.022	0.021	0.020	0.023	0.020	0.023	0.021	0.017	0.021	0.014
15	0.023	0.021	0.022	0.021	0.023	0.023	0.020	0.024	0.021	0.018	0.021	0.014
16	0.020	0.020	0.019	0.015	0.019	0.022	0.017	0.022	0.018	0.019	0.020	0.011
9	0.021	0.021	0.019	0.019	0.019	0.021	0.010	0.017	0.019	0.019	0.023	0.017
8	0.018	0.017	0.017	0.016	0.019	0.019	0.016	0.021	0.015	0.017	0.017	0.016
1	0.017	0.020	0.016	0.015	0.016	0.022	0.012	0.022	0.010	0.014	0.015	0.016
T	0.018	0.018	0.017	0.016	0.018	0.020	0.015	0.022	0.016	0.018	0.018	0.010
17/T	0.020	0.020	0.019	0.019	0.020	0.023	0.017	0.024	0.018	0.020	0.020	0.012
11/T	0.019	0.019	0.018	0.018	0.019	0.021	0.016	0.023	0.017	0.019	0.019	0.011
Cem	0.066	0.065	0.065	0.063	0.069	0.069	0.069	0.070	0.070	0.070	0.068	0.072
Cib	0.101	0.103	0.100	0.102	0.101	0.108	0.106	0.107	0.100	0.108	0.107	0.110
Cro	0.133	0.131	0.131	0.130	0.133	0.141	0.136	0.140	0.129	0.137	0.139	0.137
Cve	0.124	0.124	0.126	0.125	0.120	0.129	0.129	0.130	0.120	0.124	0.126	0.130

The shaded area denotes the black morph. Cem, Cib, Cro and Cve are abbreviations for *C. emmae, C. iberica, C. rosenhaueri* and

The mitochondrial DNA median-joining network for all populations recovered 19 haplotypes for the bicolour morph and four for the black morph (Fig. 2). The mismatch distribution analyses suggested that the bicolour morph is at a demographic equilibrium (multimodal; $P < 0.05$), while the black one appears at demographic expansion (unimodal; $P < 0.01$), as inferred from their respective distribution of nucleotide mismatches between all haplotypes (Fig. 5A,B). Congruently, the BSP supported the recent expansion model in the black morph (i.e. recently diverged haplotypes), and more stable effective population size changes in the bicolour morph (i.e. populations at mutation–drift equilibrium) (Fig. 5C,D). A significant pattern of isolation by distance inferred from the COI data was detected among the bicolour populations ($R = 0.38$, $P = 0.006$).

Morphological results

The first two functions extracted from the linear discriminant analysis explained 75% and 17%, respectively, of the total variance of genitalia and biometry, clearly separating *C. emmae*, *C. floricola* bicolour and *C. floricola* black in three groups (Fig. 6A). The discriminant function was highly significant (Wilks' lambda = 0.006, $F_{18,167} = 46.43$, $P < 10^{-4}$). The variable that contributed the most to the separation among all was the subgenital plate width (Wilks' lambda = 0.013, $F_{3,59} = 20.38$, $P < 10^{-6}$, Fig. S6, Supporting information). Males of *C. floricola* bicolour and black morphs found living in the same area in the DNP clustered apart, while males of the bicolour morph collected in two localities 70 km apart (DNP and Alc) clustered together.

Chemical results

The cuticle of *C. emmae* and both *C. floricola* colour morphs was composed of a mixture of linear, monomethyl and dimethyl alkanes between 25 and 35 carbons (Table S2, Supporting information). One unsaturated CHC was encountered on both morphs of *C. floricola* (C31:1) but not on that of *C. emmae*. Six peaks included three mixtures of heavy C31, C33 and C35 dimethyl alkanes in *C. emmae* but not in *C. floricola*. Among *C. floricola*, both morphs had one unique peak that was absent on the other. Moreover, relative hydrocarbon quantities

Cataglyphis floricola haplotypes recovered

12	13	14	15	16	9	8	1	T	17/T	11/T	Cem	Cib	Cro	Cve
13	21	20	20	17	10	16	15	16	18	17	58	89	93	87
13	20	18	18	17	10	15	17	16	18	17	57	91	92	87
12	20	19	19	16	9	15	14	15	17	16	57	88	92	88
11	19	18	18	13	9	14	13	14	16	15	54	87	91	87
13	21	17	20	16	9	17	14	16	18	17	60	89	93	84
15	23	20	20	19	10	17	19	18	20	19	60	95	99	90
10	18	17	17	14	5	14	10	13	15	14	59	91	95	90
16	24	20	21	19	8	18	19	19	21	20	61	94	98	91
11	19	18	18	15	9	13	9	14	16	15	61	88	90	84
13	19	15	16	16	9	15	12	16	18	17	61	95	96	87
13	19	18	18	17	11	15	13	16	18	17	59	94	97	88
5	13	12	12	9	8	13	13	8	10	9	60	92	94	89
—	10	9	9	6	4	10	12	5	7	6	59	91	94	86
0.011	—	9	7	8	9	16	18	11	13	12	62	93	97	89
0.010	0.010	—	6	9	8	15	19	12	14	13	58	92	95	87
0.010	0.008	0.007	—	9	7	15	17	10	12	11	59	92	96	87
0.007	0.009	0.011	0.011	—	7	14	16	7	9	8	60	89	95	89
0.008	0.019	0.017	0.015	0.014	—	11	8	5	5	5	39	60	73	69
0.011	0.018	0.017	0.017	0.017	0.023	—	14	13	15	14	59	91	95	85
0.014	0.021	0.022	0.020	0.019	0.017	0.016	—	15	17	16	61	92	94	88
0.006	0.012	0.014	0.011	0.008	0.010	0.015	0.017	—	2	1	61	92	96	90
0.008	0.015	0.016	0.014	0.011	0.010	0.017	0.020	0.002	—	1	63	93	97	91
0.007	0.014	0.015	0.013	0.009	0.010	0.016	0.019	0.001	0.001	—	62	92	96	90
0.068	0.071	0.068	0.068	0.071	0.081	0.068	0.071	0.070	0.072	0.071	—	89	91	93
0.103	0.106	0.107	0.106	0.105	0.124	0.104	0.107	0.105	0.106	0.105	0.102	—	35	81
0.134	0.139	0.136	0.138	0.136	0.151	0.136	0.134	0.137	0.139	0.137	0.130	0.050	—	90
0.123	0.127	0.124	0.125	0.127	0.143	0.121	0.126	0.129	0.130	0.129	0.133	0.116	0.129	—

C. velox respectively.

differed greatly between the bicolour and black morph. For instance, the dominant peak on the black morph was a mixture of various C30 monomethyl alkanes, while on the bicolour morph, it was a mixture of C31 monomethyl alkanes. The first two functions extracted from the linear discriminant analysis, which explained 45% and 42% of the total variance in cuticular profiles respectively, showed important differences between *C. emmae* and both *C. floricola* morphs (Fig. 6B). By contrast, among the bicolour morphs, samples collected at three localities between 30 and 70 km apart clustered together.

Discussion

The use of an integrated approach based on chemical, morphological and genetic data shows that the two colour morphs of *Cataglyphis floricola* are two distinct species that have diverged very recently. The low number of shared alleles from the nuclear markers and F_{ST} values between species at the unique contact zone indicate considerable divergence and genetic isolation between them. Further evidence derives from male genital morphology and from qualitative differences of the CHC, which are in agreement with haplotypic differentiation. Based on this strong evidence, we conclude that the two morphs are in fact two different species. The low phylogenetic distance between the two new species suggests a very recent divergence when compared to other species within the genus (Knaden *et al.* 2012). Additionally, the mismatch distribution analysis of both species suggests that they are undergoing different demographic changes, with the bicolour and black species at demographic equilibrium and expansion, respectively. Hence, at first glance, these results are in line with an allopatric speciation scenario as we discuss below.

Although it is difficult to discern whether the observed haplotype diversity is the product of inherited polymorphism from the ancestral population or the result of mutations after speciation, the recent geological changes throughout the area where both species are established today suggest that their differentiation is likely the consequence of rapid divergence between populations through geographical isolation. Based on the assumption of an approximate divergence rate of

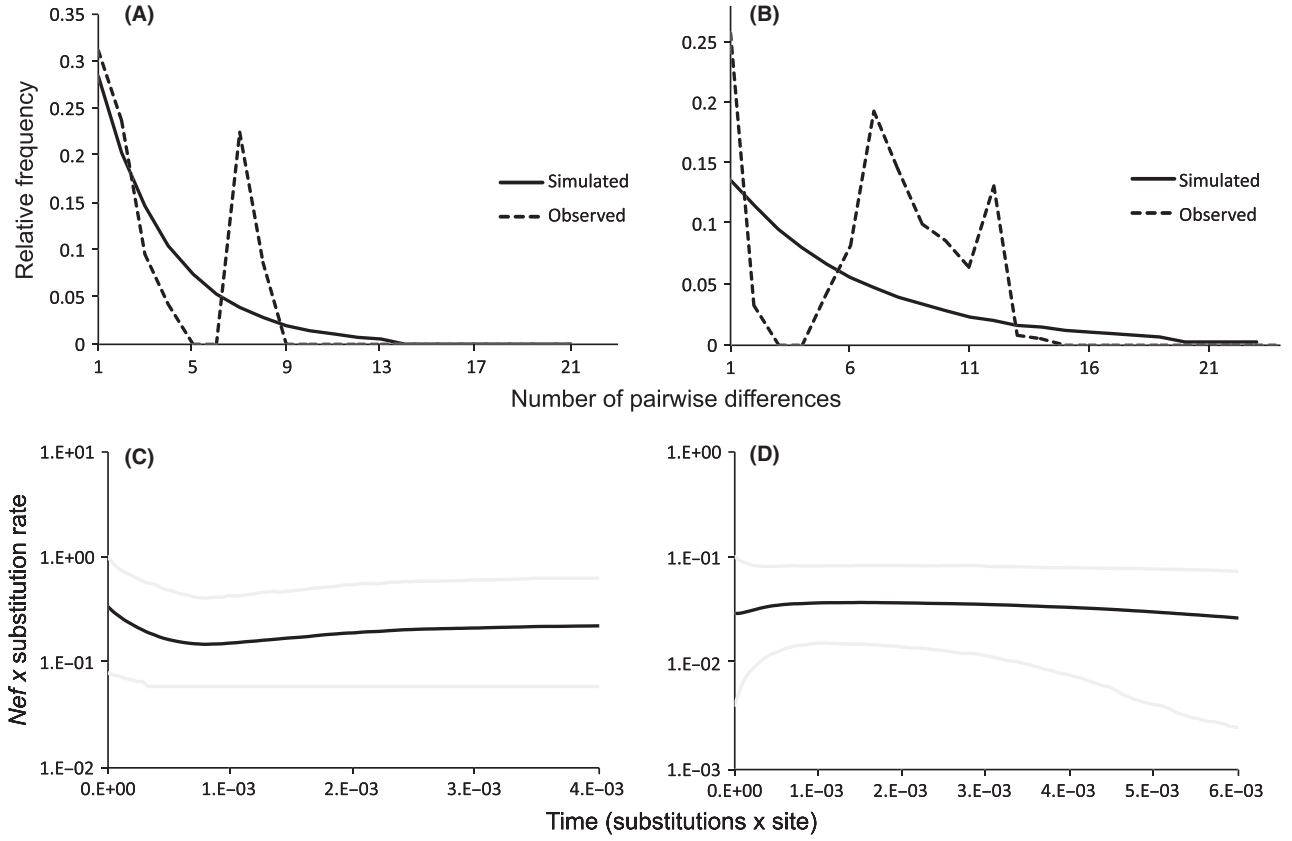

Fig. 5 Demographic expansion models for *Cataglyphis floricola*. Mismatch distribution analysis indicating the number of pairwise differences against their relative frequency for the black (A) and bicolour morphs (B). The observed distributions are compared for their goodness-of-fit to a distribution under a model of sudden expansion. Bayesian skyline plot showing the effective population size changes throughout time for the black (C) and bicolour morphs (D). The centre line indicates the median estimate; upper and lower limits (95% HPD) are indicated by grey lines.

2% per million years (Juan *et al.* 1995; Goropashnaya *et al.* 2004), both species could have diverged about 1 Ma, during the Pleistocene. The Guadalquivir estuary and surrounding areas experienced extensive geological transformations throughout that time (Zazo *et al.* 2005), which could have isolated part of the population, ultimately leading to allopatric speciation. Today's distribution and contact zone of both species at DNP (Figs 1 and S2, Supporting information) could be a consequence of an aridification process since the Holocene, associated with the development of the Doñana split (Rodríguez-Ramírez *et al.* 1996), where the progressive range expansion of both species has led to a secondary contact.

The species' distinct demographic history indicates that the bicolour species is at demographic equilibrium, suggesting that it has been established in the area for a long period of time and is not the consequence of a recent distribution shift or population expansion. In contrast, a population expansion pattern was observed in the black species, which may be a consequence of a recent bottleneck and consecutive founder effect. If so, this may be the result of the last glacial maxima (Würm glacial), when the sea level around the gulf of Cadiz was between 20 and 30 m lower than present, implying that the coast has receded around 10 km since then (Dabrio *et al.* 2000; Zazo *et al.* 2005). Geological events and associated changes in species' genetic divergence are not uncommon in the Iberian Peninsula, which has been reported as a refugia for a number of species (Gómez & Lunt 2006). Thus, today's black species may represent a population that was once much larger and is currently under expansion. This notion is further supported by the genetic divergence found within both morph haplotypes (black 0.8% and bicolour: 1.6%), which also suggests contrasting evolutionary histories since isolation. Taken together, these results are in line with a scenario of ecological isolation leading to allopatric speciation. Therefore, it suggests that these species did not diverge by parapatric speciation, as today's current distribution would imply, with a secondary contact zone and no apparent hybridization. The rapid

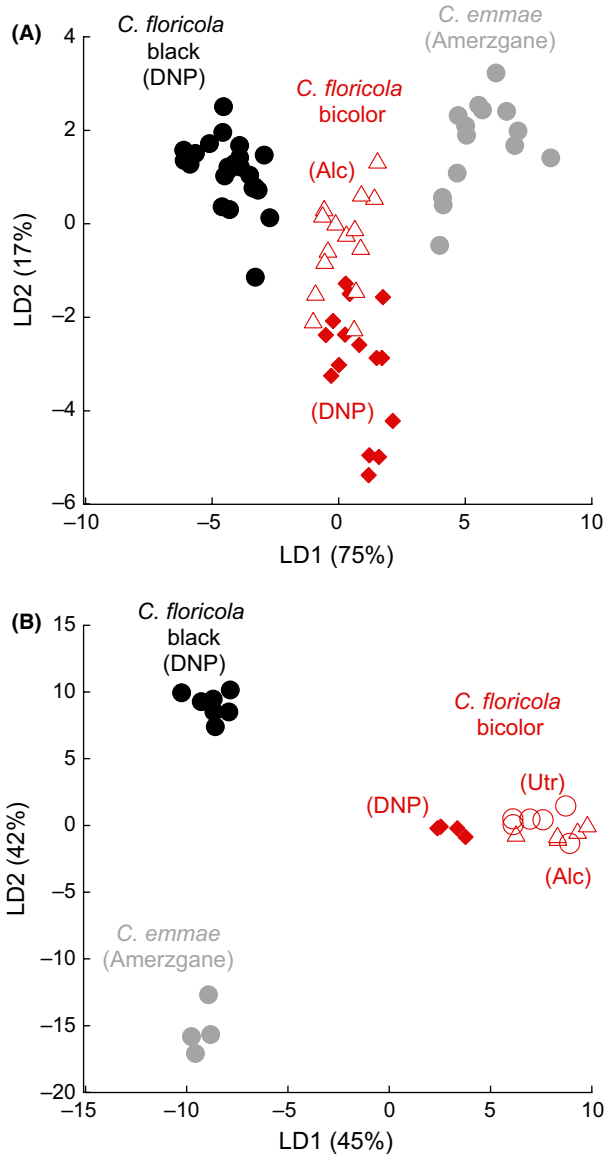

Fig. 6 Discriminant analyses of (A) six biometric measurements of male mating genitalia of *Cataglyphis floricola* (black morph from DNP $n = 23$, bicolour morph from DNP $n = 15$, bicolour morph from Alc $n = 15$) and *C. emmae* ($n = 15$) male mating genitalia and (B) 27 cuticular hydrocarbons of *C. floricola* (black morph from DNP $n = 7$, bicolour morph from DNP $n = 4$, bicolour morph from Alc = 5, bicolour morph from Utr = 6), and *C. emmae* from Amerzgane ($n = 4$). DNP, Doñana National Park; Alc, Alcalá de Guadaíra; Utr, Utrera; Amerzgane, Morocco.

divergence between the species indicates a very rapid evolution of genitalia for which sexual selection is thought to be a primary cause (Hosken & Stockley 2004). Furthermore, in geographically isolated populations, sexual selection may initiate independent episodes of male–female co-evolution, resulting in divergence of both the male trait and female preference (Lande 1981). In addition, reproductive isolation may also evolve via genetic drift in allopatry. When populations experience geographical isolation, genetic drift may change the evolutionary trajectory of sexual traits between the populations (Althoff *et al.* 2014). Finally, small isolated populations are more sensitive to the negative effects of genetic drift (i.e. gene loss, accumulation of mutations), which may contribute to increase the pace of speciation (Turelli *et al.* 2001).

Cataglyphis floricola's dispersal mode (colony fission) could have played an important role in reproductive isolation between the species. As mentioned above, females are philopatric and disperse over short distances (about seven metres), while males are able to fly over longer distances. This colony-founding strategy results in high population structure, slow population expansion, high population viscosity, restricted maternal gene flow and high genetic divergence at a large regional scale (Peeters & Ito 2001). In allopatric speciation, these factors could have had some effect, in conjunction with a geographical barrier, on the speed at which reproductive isolation was achieved. The effect of such a barrier can be inferred from the congruence in both mitochondrial and nuclear markers in differentiating the species, which reflects historical patterns of isolation as well as a lack of contemporary gene flow. The species' range overlap should allow for potential mating encounters, but ecological isolation through habitat or temporal isolation is likely to have prevented successful mating (i.e. prezygotic isolation). Moreover, male-biased gene flow suggests that the evolution of male copulatory structures could have been key for the reproductive isolation of both species. Thus, isolation could have allowed for local selective pressures to act on traits under sexual selection (Peeters & Ito 2001) and for the fixation of different alleles through allopatry and restricted local mating. Moreover, possible differences in female pheromone sensory drive may have accentuated mating preferences with local males, maintaining the level of reproductive isolation between species (Boughman 2002).

CHCs serve many functions in ants, from protection against desiccation to communication among nestmates and non-nestmates (Martin & Drijfhout 2009). Van Oystaeyen *et al.* (2014) recently proposed that their ancestral function is communication between females and males (i.e. in the context of mate choice). Although this hypothesis has not been thoroughly investigated, it is supported by the fact that CHCs serve as sexual pheromones in flies (Ferveur 2005) and solitary wasps (Böröczky *et al.* 2009). Communication among sexual partners and sexual selection are known as important drivers of speciation by reinforcement (selection against

deleterious hybridization; Ritchie 2007). We therefore propose that selection acting on the recognition system of males and females of *C. floricola* may have had consequences on the species-specific hydrocarbon profile that is produced by both sexuals and workers.

Sexual selection can be influenced by environmental change, incrementally increasing divergence among isolated populations and ultimately leading to speciation (Uy & Borgia 2000). Thus, it is worth stating that some phenetic differences such as the colour of both species may be an indication of a cryptic adaptation to microenvironmental differences (e.g. temperature, solar radiation) within preferred habitats. In line with this notion, Keller & Seehausen (2012) recently reviewed the role of divergent selection through thermal exposure due to habitat adaptation, concluding that most initial divergence was likely the result of allopatric speciation and stable coexistence on secondary contact zones. In the case of our study species, this notion warrants further investigation.

Conservation implications

Our intensive survey throughout most of western Andalucía revealed that the *C. floricola* populations east of the Guadalquivir River are small and threatened by agricultural and urban expansions at this margin of the river. These populations have distinct divergent mitochondrial haplotypes and represent unique evolutionary lineages, which deserve a proper conservation management programme. In contrast to the bicolour species, the black species is mostly found in two protected areas, the Doñana Natural Park and Doñana National Park. Nevertheless, because the standard of the natural park conservation is less stringent than that of the national park. Therefore, it should be especially monitored so that specific conservation programmes will guarantee the stability of the populations of both species. The present study brings together all the necessary evidence to rigorously establish the occurrence of two distinct species with the black species having probably the smallest known distribution range of all known *Cataglyphis*. Adequate conservation measures would benefit from a proper taxonomic description. Thus, this protected environment and this endemic species offer a rare and unique model in *Cataglyphis* to study the processes involved in speciation.

Acknowledgements

We are grateful to Ana Carvajal for DNA extractions and assistance in the field. We thank David Aragonés, Isabel Afán and Rubén Solís (LAST-EBD) for their assistance with GIS analyses and to Prof. Alberto Tinaut for comments during the FA thesis committee and also Dr. Antonio Rodríguez-Ramírez for his help with the geology of the area. Cristina Vaquero, from the Electron Microscopy Service of the University of Seville, helped us with the electron photographs. We thank the comments of subject editor Prof. Sean Schoville and five anonymous reviewers, who allowed us to improve the manuscript. This work was funded by grants CGL2009-12472 and CGL2009-09690 MONTES from the Spanish Ministerio de Ciencia e Innovación and FEDER (European Regional Development Fund) and CGL2012-36181 from the Spanish Ministerio de Economía y Competitividad and FEDER. All experiments comply with current Spanish legislation.

References

Abramoff MD, Magalhaes PJ, Ram SJ (2004) Image processing with ImageJ. *Biophotonics International*, **11**, 36–42.

Agosti D (1990) Review and reclassification of *Cataglyphis* (Hymenoptera, Formicidae). *Journal of Natural History*, **24**, 1457–1505.

Althoff DM, Segraves KA, Johnson MTJ (2014) Testing for coevolutionary diversification: linking pattern with process. *Trends in Ecology & Evolution*, **29**, 82–89.

Amor F, Ortega P, Jowers MJ *et al.* (2011) The evolution of worker–queen polymorphism in *Cataglyphis* ants: interplay between individual and colony-level selections. *Behavioral Ecology and Sociobiology*, **65**, 1473–1482.

Bandelt HJ, Forster P, Röhl A (1999) Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution*, **16**, 37–48.

Böröczky K, Crook DJ, Jones TH *et al.* (2009) Monoalkenes as contact sex pheromones components of the woodwasp *Sirex noctilio*. *Journal of Chemical Ecology*, **35**, 1202–1211.

Bos N, Dreier S, Jørgensen CG, Nielsen J, Guerrieri FJ, d'Ettorre P (2012) Learning and perceptual similarity among cuticular hydrocarbons in ants. *Journal of Insect Physiology*, **58**, 138–146.

Boudinot BE (2013) The male genitalia of ants: musculature, homology, and functional morphology (Hymenoptera, Aculeata, Formicidae). *Journal of Hymenoptera Research*, **30**, 29–49.

Boughman JW (2002) How sensory drive can promote speciation. *Trends in Ecology and Evolution*, **17**, 571–577.

Bradshaw HD Jr, Schemske DW (2003) Allele substitution at a flower color locus produces a pollinator shift in monkeyflowers. *Nature*, **426**, 176–178.

Dabrio CJ, Zazo C, Godoy JL *et al.* (2000) Depositional history of estuarine infill during the last postglacial transgression (Gulf of Cadiz, Southern Spain). *Marine Geology*, **164**, 381–404.

Dahbi A, Lenoir A, Tinaut A, Taghizadeh T, Francke W, Hefetz A (1996) Chemistry of the postpharyngeal gland secretion and its implication for the phylogeny of Iberian *Cataglyphis* species (Hymenoptera: Formicidae). *Chemoecology*, **7**, 163–171.

Dani FR, Jones GR, Corsi S, Beard R, Pradella D, Turillazzi S (2005) Nestmate recognition cues in the honey bee: differential importance of cuticular alkanes and alkenes. *Chemical Senses*, **30**, 477–489.

Drummond AJ, Suchard MA, Xie D, Rambaut A (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*, **29**, 1969–1973.

Earl DA (2011) Structure Harvester v0.6.7. Available at http://users.soe.ucsc.edu/~dearl/software/structureHarvester.

Endler JA (1977) *Geographic Variation, Speciation and Clines*. Princeton University Press, Princeton, New Jersey. 262 p.

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

Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, **10**, 564–567.

Ferveur JF (2005) Cuticular hydrocarbons: their evolution and roles in Drosophila pheromonal communication. *Behavior Genetics*, **33**, 279–295.

Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R (1994) DNA primers for amplification of mitochondrial cytochrome *c* oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology*, **3**, 294–299.

Galtier N, Depaulis F, Barton N (2000) Detecting bottlenecks and selective sweeps from DNA sequence polymorphism. *Genetics*, **155**, 981–987.

Gómez A, Lunt DH (2006) Refugia within refugia: Patterns of phylogeographic concordance in the iberian peninsula. In: *Phylogeography of southern european refugia: Evolutionary perspectives on the origins and conservation of european biodiversity* (eds Weiss S, Ferrand N), pp. 155–188. Springer, Dordrecht, The Netherlands.

Goropashnaya AV, Fedorov VB, Pamilo P (2004) Recent speciation in the *Formica rufa* group ants (Hymenoptera, Formicidae): inference from mitochondrial DNA phylogeny. *Molecular Phylogentics and Evolution*, **32**, 198–206.

Goudet J (1995) FSTAT (version 1.2). A computer program to calculate F-statistics. *Journal of Heredity*, **86**, 385–386.

Gouy M, Guindon S, Gascuel O (2010) SeaView version 4. A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Molecular Biology and Evolution*, **27**, 221–224.

Gray SM, McKinnon JS (2007) Linking color polymorphism maintenance and speciation. *Trends in Ecology & Evolution*, **22**, 71–79.

Guerrieri FJ, Nehring V, Jørgensen CG, Nielsen J, Galizia CG, d'Ettorre P (2009) Ants recognize foes and not friends. *Proceedings of the Royal Society B: Biological Sciences*, **276**, 2461–2468.

Harpending H, Batzer M, Gurven M, Jorde L, Rogers A, Sherry S (1998) Genetic traces of ancient demography. *Proceedings of the National Academy of Sciences of the United States of America*, **95**, 1961–1967.

Hosken DJ, Stockley P (2004) Sexual selection and genital evolution. *Trends in Ecology and Evolution*, **19**, 87–93.

Juan C, Oromi P, Hewitt GM (1995) Mitochondrial DNA phylogeny and sequential colonization of Canary Islands by darkling beetles of the genus *Pimelia* (Tenebrionidae). *Proceedings of the Royal Society of London, Series B*, **261**, 173–180.

Keller I, Seehausen O (2012) Thermal adaptation and ecological speciation. *Molecular Ecology*, **21**, 785–799.

Kimura M, Ohta T (1971) Protein polymorphism as a phase of molecular evolution. *Nature*, **229**, 467–469.

Knaden M, Tinaut A, Stökl J, Cerdá X, Wehner R (2012) Molecular phylogeny of the desert ant genus *Cataglyphis* (Hymenoptera: Formicidae). *Myrmecological News*, **16**, 123–132.

Lande R (1981) Models of speciation by sexual selection on polygenic traits. *Proceedings of the National Academy of Sciences of the United States of America*, **78**, 3721–3725.

Larkin MA, Blackshields G, Brown NP *et al.* (2007) Clustal W and Clustal X version 2.0. *Bioinformatics*, **23**, 2947–2948.

Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, **25**, 1451–1452.

Martin S, Drijfhout F (2009) A review of ant cuticular hydrocarbons. *Journal of Chemical Ecology*, **35**, 1151–1161.

Nosil P, Egan SP, Funk DJ (2008) Heterogeneous genomic differentiation between walking-stick ecotypes: 'isolation by adaptation' and multiple roles for divergent selection. *Evolution*, **62**, 316–336.

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

Peeters C, Ito F (2001) Colony dispersal and the evolution of queen morphology in social Hymenoptera. *Annual Review of Entomology*, **46**, 601–630.

Posada D (2008) jModelTest: phylogenetic model averaging. *Molecular Biology and Evolution*, **25**, 1253–1256.

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

de Queiroz K (2007) Species concept and species delimitation. *Systematic Biology*, **56**, 879–886.

Raymond M, Rousset F (1995) Genepop (version 1.2), population genetics software for exact tests and ecumenicism. *Journal of Heredity*, **86**, 248–249.

Rice WR, Hostert EE (1993) Laboratory experiments on speciation: what have we learned in 40 years? *Evolution*, **47**, 1637–1653.

Ritchie MG (2007) Sexual selection and speciation. *Annual Review of Ecology and Systematics*, **38**, 79–102.

Rodríguez-Ramírez A, Rodríguez-Vidal J, Cáceres L *et al.* (1996) Recent coastal evolution of Doñana National Park (SW Spain). *Quaternary Science Reviews*, **15**, 803–809.

Rogers A, Harpending H (1992) Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution*, **9**, 552–569.

Ronquist F, Huelsenbeck JP (2003) Mrbayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, **19**, 1572–1574.

Rundle HD, Nosil P (2005) Ecological speciation. *Ecology Letters*, **8**, 336–352.

Schluter D (2009) Evidence for ecological speciation and its alternative. *Science*, **323**, 737–741.

Seifert B (2009) Cryptic species in ants (Hymenoptera: Formicidae) revisited: we need a change in alpha-taxonomic approach. *Myrmecological News*, **12**, 149–166.

Slatkin M (1993) Isolation by distance in equilibrium and non-equilibrium populations. *Evolution*, **47**, 264–279.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution*, **28**, 2731–2739.

Tinaut A (1993) *Cataglyphis floricola* nov.sp. new species for the genus *Cataglyphis* Förster, 1850 (Hymenoptera, Formicidae) in the Iberian Peninsula. *Bulletin de la Société Entomologique Suisse*, **66**, 123–134.

Truett GE, Heeger P, Mynatt RL, Truett AA, Walker JA, Warman ML (2000) Preparation of PCR-quality mouse genomic DNA with hot sodium hydroxide and Tris (HotSHOT). *BioTechniques*, **29**, 52–54.

Turelli M, Barton NH, Coyne JA (2001) Theory and speciation. *Trends in Ecology and Evolution*, **16**, 330–343.

Uy JA, Borgia G (2000) Sexual selection drives rapid divergence in bowerbird display traits. *Evolution*, **54**, 273–278.

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) Micro-Checker: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecolology Notes*, **4**, 535–538.

Van Oystaeyen A, Oliveira RC, Holman L *et al.* (2014) Conserved class of queen pheromones stops social insect workers from reproducing. *Science*, **343**, 287–290.

Wiley EO (1977) The evolutionary species concept reconsidered. *Systematic Biology*, **27**, 17–26.

Zazo C, Mercier N, Silva PG *et al.* (2005) Landscape evolution and geodynamic controls in the Gulf of Cadiz (Huelva coast, SW Spain) during the Late Quaternary. *Geomorphology*, **68**, 269–290.

F.A., J.A.G., X.C. and R.B. conceived the ideas; F.A., P.O., A.L. collected the data; F.A., M.J., R.B. and J.A.G. analysed the data; M.J. and J.A.G. led the writing and F.A., R.B. contributed.

Data accessibility

DNA sequences: GenBank accessions KJ538563–KJ538585.

Sampling locations including GIS coordinates. Dryad doi: 10.5061/dryad.qk6d9

DNA sequence alignments and tree files. Dryad doi: 10.5061/dryad.qk6d9

Microsatellite genotypes. Dryad doi: 10.5061/dryad.qk6d9

Raw morphological data. Dryad doi: 10.5061/dryad.qk6d9

Raw CHC data files. Dryad: 10.5061/dryad.qk6d9

Supporting additional figures, morph photographs and CHCs composition uploaded online as supporting information.

Supporting information

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

Fig. S1 Picture of workers of *Cataglyphis floricola* black and bicolour morphs.

Fig. S2 Transect at Doñana (DNP). *Cataglyphis floricola* bicolour and black morph distribution join at the central transect point. Red and black points indicate the presence of bicolour and black morph nests, respectively.

Fig. S3 Map showing all locations surveyed in this study with black dots. Main cities are indicated with red dots.

Fig. S4 Views in scanning electron microscopy of the copulatory appendages of *C. floricola* (black morph).

Fig. S5 (A) Maximum change of rate in estimating the number of populations (*K*) as inferred by Structure. (B) Mean logarithmic likelihood values for each *K* tested. (C) Estimated probabilities of population membership assessed by multilocus genotype clustering when $K = 17$. Each colony is represented by a vertical bar. Colors represent the population been assigned. (D) Median joining network of all *C. floricola* haplotypes based on COI partial sequences (884 bp) from the transect. The size of the circles corresponds to the haplotype frequency. Numbers 1, 2, 3, 4a are from the bicolor morph and 4b, 5, 6, 7 from the black morph. fr: haplotypic frequency.

Fig. S6 Pairwise biplots of the relation between the 5 morphometric variables measured on the males genitalia.

Table S1 Microsatellite markers, total number of alleles, alleles shared between both *Cataglyphis floricola* morphs and allelic richness of each morph.

Table S2 Cuticular hydrocarbon composition of *C. floricola* (black and bicolor morphs) and *C. emmae*.

Supplementary Material

Recent speciation and secondary contact in endemic ants
MJ Jowers, F Amor, P Ortega, A Lenoir, RR Boulay, X Cerdá, JA Galarza
Molecular Ecology

Figure S1. - Picture of workers of *Cataglyphis floricola* black and bicolour morphs.

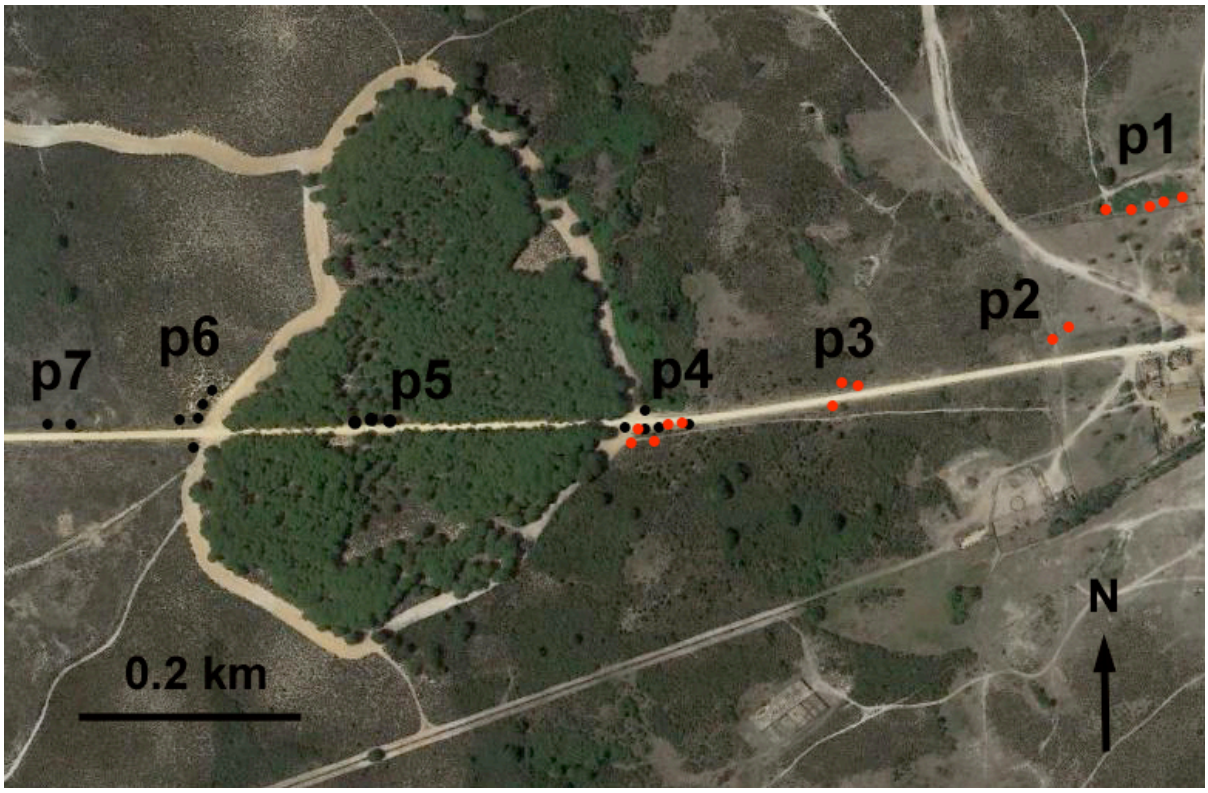

Figure S2. - Transect at Doñana (DNP). *C. floricola* bicolour and black morph distribution join at the central transect point. Red and black points indicate the presence of bicolour and black morph nests respectively.

Table S1. - Microsatellite markers, total number of alleles, alleles shared between both *C. floricola* morphs and allelic richness of each morph.

Loci	Number of alleles	Number of shared alleles	Allelic richness Bicolour morph	Allelic richness Black morph
Ccur21	9	2	3	8
Ccur26	13	2	11	4
Ccur51	9	5	6	8
Ccur61	11	5	8	8
Ccur99	7	2	5	4
Ccur100	19	0	3	16

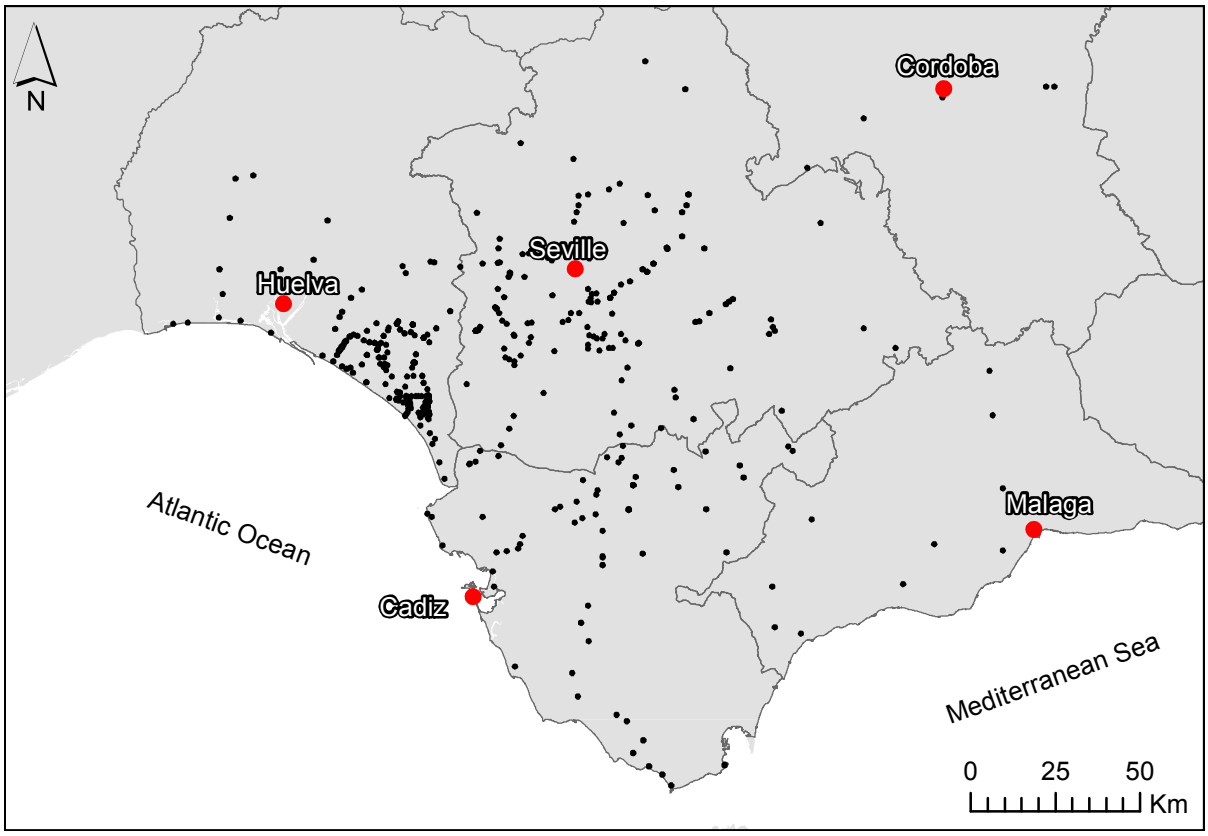

Figure S3. - Map showing all locations surveyed in this study with black dots. Main cities are indicated with red dots.

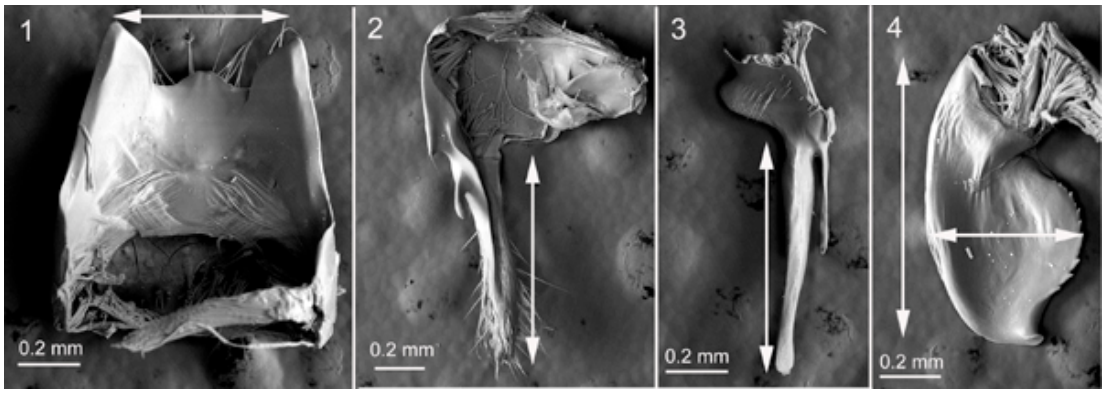

Figure S4.- Views in scanning electron microscopy of the copulatory appendages of *C. floricola* (black morph). (1) subgenital plate; (2) paramere; (3) volsella with digitus (the largest) and cuspis; (4) penisvalva. Measurements are indicated with a double-headed arrows.

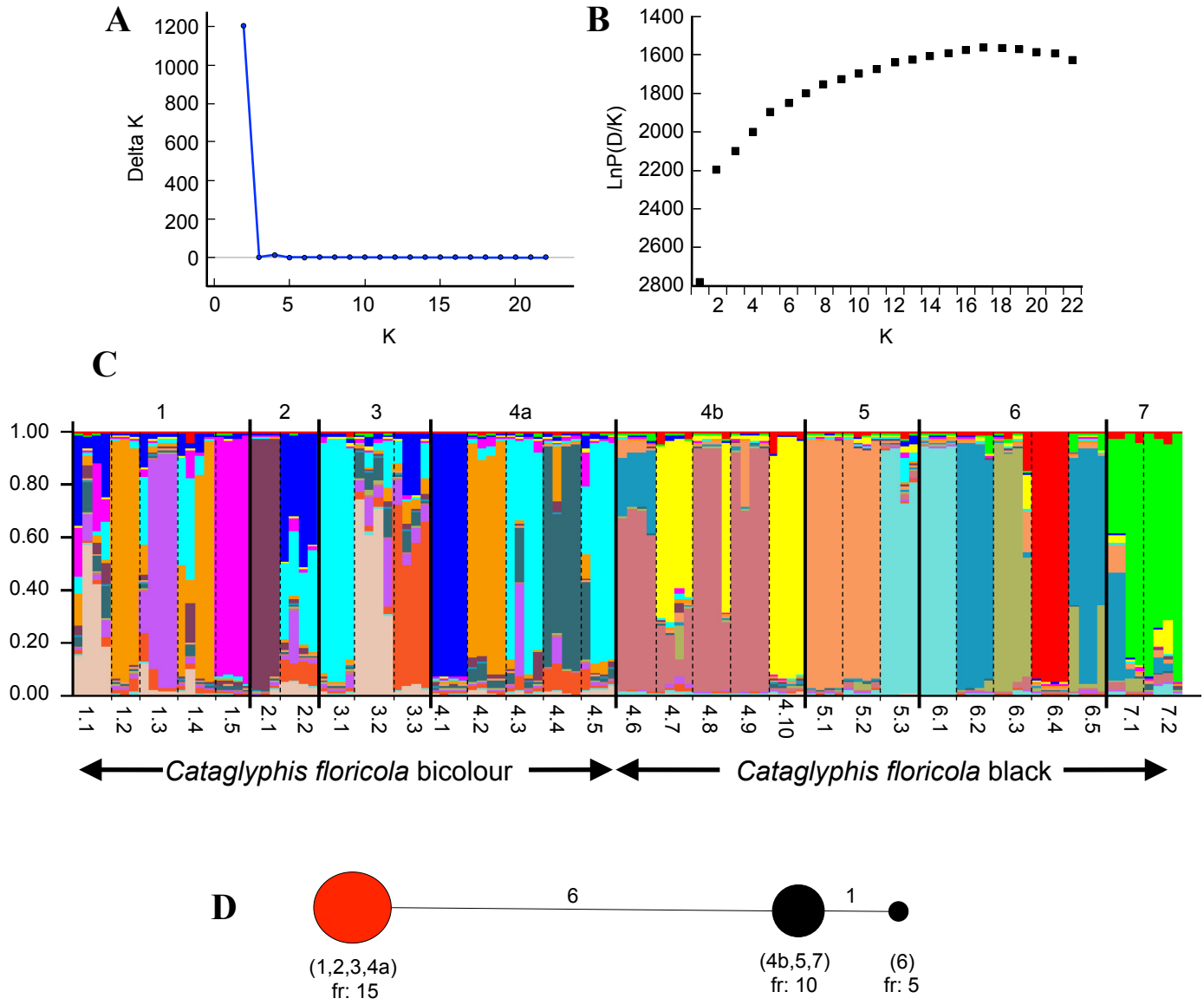

Figure S5. - (A) Maximum change of rate in estimating the number of populations (K) as inferred by Structure. **(B)** Mean logarithmic likelihood values for each K tested. **(C)** Estimated probabilities of population membership assessed by multilocus genotype clustering when K=17. Each colony is represented by a vertical bar. Colours represent the population been assigned. **(D)** Median joining network of all *C.floricola* haplotypes based on COI partial sequences (884 bp) from the transect. The size of the circles corresponds to the haplotype frequency. Numbers 1, 2, 3 ,4a are from the bicolour morph and 4b, 5, 6, 7 from the black morph. fr: haplotypic frequency.

Table S2. - Cuticular hydrocarbon composition of *C. floricola* (black and bicolour morphs) and *C. emmae*. Only the main compounds are shown and ordered according to their retention time. Values are means ± SE. Sample size is given between parentheses after the species names.

Compounds	*C. floricola* black *(7)*	*C. floricola* bicolour *(15)*	*C. emmae (4)*
C25	1.34 ± 0.30	1.22 ± 0.12	2.50 ± 0.40
3MeC25	0.27 ± 0.06	0.01 ± 0.01	6.18 ± 0.32
10 + 11 + 12 + 13MeC26	0.67 ± 0.27	0.27 ± 0.12	13.34 ± 0.76
8, 10DiMeC26	-	-	2.56 ± 0.19
4, 8 + 4, 10DiMeC26	0.02 ± 0.02	-	2.01 ± 0.21
C27	4.49 ± 0.81	5.24 ± 0.40	4.31 ± 0.08
9 + 11 + 13MeC27	10.65 ± 0.92	1.94 ± 0.17	4.58 ± 0.19
3MeC27	3.59 ± 0.57	1.18 ± 0.15	2.11 ± 0.40
C28	-	3.19 ± 0.59	3.47 ± 0.27
10 + 11 + 12 + 14MeC28	5.84 ± 1.19	0.25 ± 0.11	1.40 ± 0.16
10,12 +10,14DiMeC28	6.36 ± 0.58	0.52 ± 0.26	-
6,10 + 6,12 + 8,12DiMeC28	4.00 ± 1.18	2.47 ± 1.26	-
C29	4.84 ± 1.04	12.39 ± 1.12	2.20 ± 0.06
9 + 11 + 13 + 15MeC29	6.40 ± 0.61	9.32 ± 0.53	3.67 ± 0.14
3MeC29	-	1.65 ± 0.46	0.94 ± 0.29
7,15 + 7,17DiMeC29	9.76 ± 1.59	5.43 ± 2.54	-
5,9 + 5,11DiMeC29	3.66 ± 0.9	0.85 ± 0.29	1.70 ± 0.20
8 + 10 + 12 + 13 + 14 + 15MeC30	18.74 ± 4.33	7.19 ± 1.33	1.51 ± 0.52
10,16+12,14DiMeC30	0.16 ± 0.16	2.19 ± 0.44	-
C31:1	5.02 ± 1.88	2.73 ± 0.54	-
C31	-	4.10 ± 0.72	0.73 ± 0.06
9 + 11 + 13 + 15MeC31	0.91 ± 0.91	14.31 ± 1.38	7.03 ± 0.52
11,13 + 11,15 + 13,15 + 13,17DiMeC31	0.48 ± 0.48	5.45 ± 1.11	-
9,13 + 9,17DiMeC31	-	-	6.48 ± 0.82
9 + 11 + 13 + 15 + 17MeC33	0.05 ± 0.05	1.71 ± 0.28	3.59 ± 0.27
11,15 + 13,15DiMeC33	-	-	4.56 ± 1.05
11,15 + 13,15 + 15,17 + 17,19DiMeC35	-	-	2.67 ± 0.36

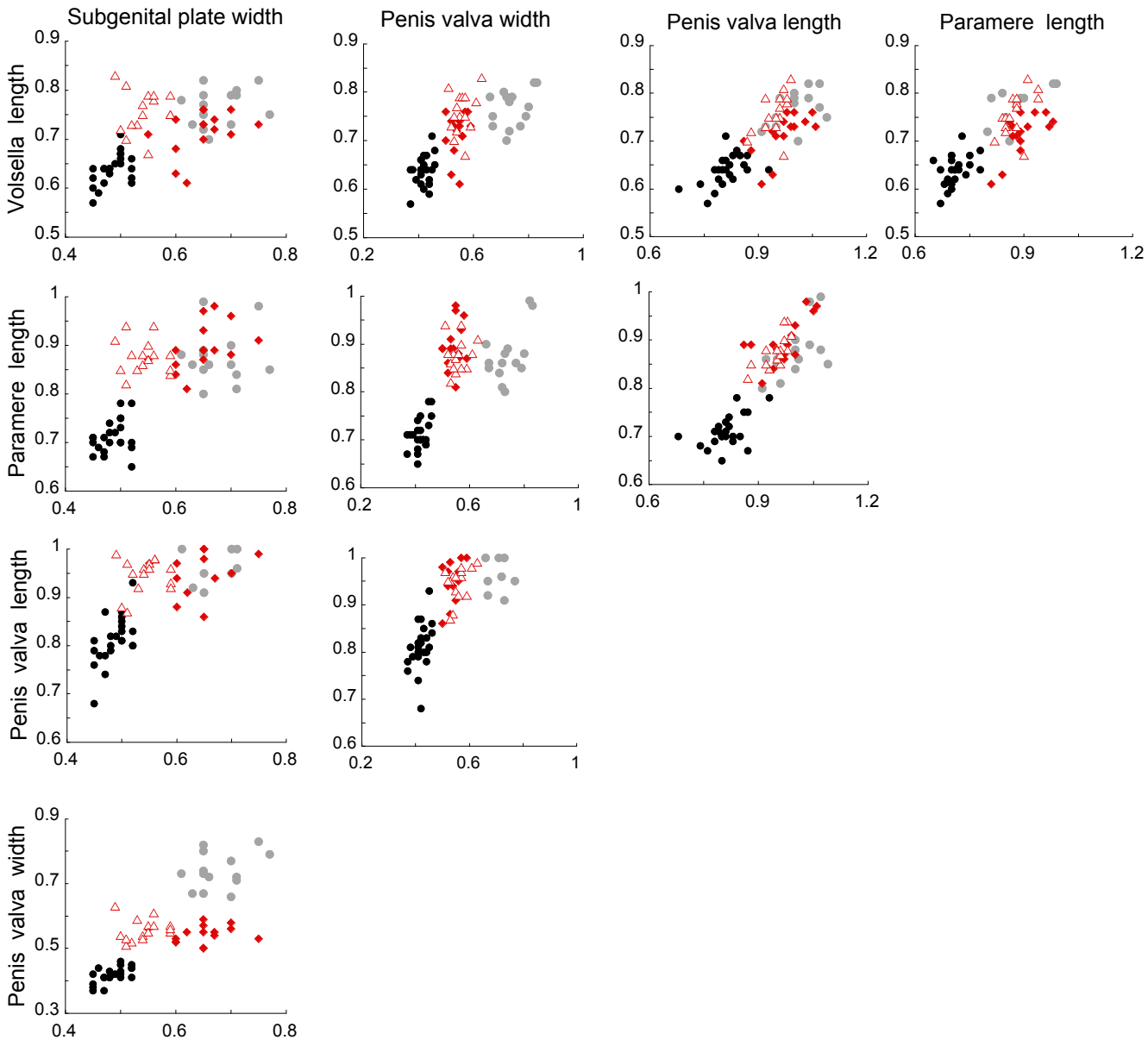

Figure 6. - Pairwise biplots of the relation between the 5 morphometric variables measured on the males genitalia. Grey and black dots represent *Cataglyphis emmae* and *C. floricola* black morph, respectively. Open triangles and red diamonds indicate *Cataglyphis floricola* bicolour morph from Alcala and the Doñana Biological Reserve, respectively.