Tracing historical introductions in the Mediterranean Basin: the success story of the common genet (Genetta genetta) in Europe

Philippe Gaubert, Irene del Cerro, Alejandro Centeno-Cuadros, Francisco Palomares, Pascal Fournier, Carlos Fonseca, Jean-Paul Paillat, José Antonio Godoy

Tracing historical introductions in the Mediterranean Basin: the success story of the common genet (*Genetta genetta*) in Europe

Philippe Gaubert[1*], Irene Del Cerro[2,3], Alejandro Centeno-Cuadros[4], Francisco Palomares[2], Pascal Fournier[5], Carlos Fonseca[3], Jean-Paul Paillat[6], José Antonio Godoy[7]

[1] Institut des Sciences de l'Evolution de Montpellier (ISEM) – UM2-CNRS-IRD, Université Montpellier 2, Place Eugène Bataillon – CC 64, 34095 Montpellier cedex 05, France

[2] Department of Conservation Biology, Estación Biológica de Doñana - CSIC, C/ Americo Vespucio s/n 41092 Sevilla, Spain

[3] CESAM & Department of Biology, University of Aveiro, Campus de Santiago, 3810-193 Aveiro, Portugal

[4] Department of Ecology, Evolution and Behavior, Alexander Silberman Institute of Life Sciences, The Hebrew University of Jerusalem, Edmond J. Safra Campus, Jerusalem 9190401, Israel

[5] GREGE, Route de Préchac, 33730 Villandraut, France

[6] Les Naturalistes Vendéens, Centre Beautour, route de Beautour, 85000 La Roche-sur-Yon, France

[7] Department of Integrative Ecology, Estación Biológica Doñana - CSIC, C/ Americo Vespucio s/n 41092 Sevilla, Spain

* Corresponding author: Philippe Gaubert – Tel. +33 467143261 - Fax. +33 467143622.

Email: philippe.gaubert@univ-montp2.fr

Running title: Introduction scenarios of the common genet in Europe

Abstract

The successful introduction of the common genet (*Genetta genetta*) into Europe has been traditionally associated to the Muslim invasion of Iberia, although diverse evidence suggested an earlier arrival. In this study, we assessed genetic variation at 11 microsatellite loci in 199 individuals from the Mediterranean Basin and used approximate Bayesian computation (ABC) combining genotypes and published mitochondrial sequences. Our objectives were to (i) test alternative scenarios of introduction of the species in Europe, (ii) re-assess the mitochondrial signatures of 'introduction hotspots' in Iberia, and (iii) evaluate how post-introduction demographic processes in the invaded range have shaped genetic structure. ABC estimates favored a scenario of independent introductions from Maghreb into the Balearic Isl. and Iberia; the latter was dated between the Upper Palaeolithic and the end of Phoenicians' influence. Patterns of genotypic diversity broadened the Andalusian introduction hotspot to the antique Tartessos Kingdom and suggested multiple introductions and/or long-term genetic drift. The best fit ABC scenario implied a natural spread from Iberia to France, but was in potential conflict with our delimitation of two genetic clusters (France and Iberia) in continental Europe. In fact, southwestern France populations showed a fair proportion of alleles shared with Maghreb and low levels of heterozygosity that may reflect subsequent introduction from Iberia, in line with the high error rates in favor of this alternative scenario. Significant patterns of isolation-by-distance among individuals within both genetic clusters are suggestive of natural dispersal from both Iberian and French introduction sites resulting in a secondary contact zone in northern Iberia. Overall, our study strongly suggests that the common genet was intentionally introduced in southern Iberia at a time antedating the Muslim invasion, possibly via Phoenicians' commercial routes. Subsequent introduction in France, long-term genetic drift and admixture likely shaped the species genetic variation currently observed in continental Europe.

Keywords: Mediterranean Basin, Viverridae, microsatellites, population genetics, Tartessos, approximate Bayesian computation.

Introduction

Introductions represent natural experiments to study non-equilibrium processes in population biology, such as colonization and spread under new environmental conditions (Sakai et al. 2001). The genetic architecture of introduced populations has been considered a major component of adaptations to newly invaded ecological niches, possibly more important than ecological tolerance (Lee 2002; Crawford and Whitney 2010). Multiple introductions involving different genetic lineages can play a significant role in counter-balancing the important loss of allelic richness and heterozygosity following introduction bottlenecks (Kolbe et al. 2004; Roman and Darling 2007; Dlugosch and Parker 2008). Secondary contact and gene flow ("admixture") within the invaded range may occur through the release of selection against admixture in introduced populations (Verhoeven et al. 2011), with the potential consequence of promoting range expansion (Forsman et al. 2008).

Thus, characterizing the introduction pattern of colonizing species is crucial to understand the determinants of successful introductions (Dlugosch and Parker 2008). The accurate traceability of introduction routes has been greatly improved by the use of multi-locus genotyping to detect founder events (Davies et al. 1999). Nevertheless, the detection of multiple introductions may be rendered difficult by subsequent rapid population growth and admixture among invading populations (Khamis et al. 2009). This issue is exacerbated in historically introduced species, where demographic processes occurring over hundreds of generations can blur the genetic signature of early introductions. In this case, organelle genomes such as mitochondrial DNA may prove useful in retaining the genetic imprint of past introductions (Searle 2008; Jones et al. 2013).

The cultural exchanges connecting the borders of the Mediterranean Basin (MB) since prehistoric times represent an outstanding framework to study historical introductions, notably in mammals (Dobson 1998). Human-mediated introductions since the end of the

Würmian glaciations (14-12 kyr ago) have deeply impacted current patterns of biodiversity in MB (Blondel and Vigne 1993; Vigne et al. 2009). These led to dramatic levels of endemic extinction, at the same time counter-balanced by the establishment of various allochthonous taxa, which are now paradoxically considered part of the "cultural heritage" of MB (Masseti 2009). The intensity of introductions has significantly increased since the first millennium B.C., following massive human migrations from eastern to western borders of the Mediterranean (e.g. Cucchi et al. 2005) that opened several dozens of potential routes to the human-mediated dispersal of species across MB (Ciolek 2011).

The common genet (*Genetta genetta*) is an opportunistic meso-predator (Viverridae, Carnivora) naturally distributed across Maghreb (Mauritania, Morocco, Algeria, Tunisia), sub-Saharan Africa and the southern Arabian Peninsula (Delibes and Gaubert 2013). Its establishment in Europe constitutes a unique example of a successful introduction of a wild African carnivore, since the Egyptian mongoose (*Herpestes ichneumon*) —long thought to be a contemporaneously introduced species— might have dispersed naturally into southwestern Europe (Gaubert et al. 2011). The invaded range of the common genet now spreads from Portugal to continental Spain, Mallorca, Cabrera and Ibiza (Balearic Isl.) and France (west, south-west and south-east). A few records are also known from Italy (Gaubert et al. 2008b). The introduction of the species has traditionally been associated to the Muslim invasion of Iberia (Spain and Portugal) starting 711 A.D. (Morales 1994), although a Greek historical source suggested its presence in Europe as early as the 6th century B.C. (Amigues 1999). Recent investigations based on mitochondrial DNA (mtDNA) reassessed these views, and suggested that the species was introduced from an endemic North African lineage into Andalusia (southern Spain), Catalonia (northwestern Spain), Mallorca and Ibiza (Balearic Isl.), and possibly western France (Gaubert et al. 2009; Gaubert et al. 2011). The distribution of the haplogroup at the origin of the European populations suggested an early influence of

Phoenicians / Carthaginians, possibly later relayed by Muslim conquerors (Gaubert et al. 2011). However, such hypotheses were based on a single locus (mtDNA) and were lacking an explicit test of introduction scenarios, including times of introduction.

Following the postulate that genetic imprints in historically introduced populations are expected to be a good proxy of the history of humans' dispersal, we aimed at refining our mtDNA-based scenario of multiple introductions of the common genet in Europe using multi-locus genotyping (microsatellites). More specifically, we tested different scenarios of introduction from Maghreb and reassessed the delineation of the proposed 'introduction hotspots' in Andalusia and Catalonia. In a second step, we questioned how post-introduction events and demographic processes may have shaped the current genetic structure of the common genet in continental Europe.

Methods

Geographic sampling

A total of 199 samples were collected within the Mediterranean rim, covering Maghreb—the geographic source of European populations— and the range invaded by *G. genetta*, including southwestern continental Europe (Spain, Portugal, France and Italy) and the Balearic Isl. (Fig. 1; Table 1). Samples were gathered through a network of collaborators from the study region, resulting in a variety of sample types that included muscle, blood and ear tissue (146 samples), and guard hairs (53 samples) (Gaubert et al. 2009; Gaubert et al. 2011). The samples were preserved in ~90% ethanol at 4ºC before DNA extraction.

Given the opportunistic strategy of sample collection and our large study area, we followed a threshold between geographic proximity and number of individuals to partition our sample set in 20 geographic 'populations' including two populations from the native range

(Table 1). Six populations with N<5 were not included in population genetic analyses based on allelic frequencies.

Microsatellite genotyping

Fresh tissue and hair samples were processed in separated lab areas to avoid cross-contamination. Genomic DNA was extracted from muscle and blood using the phenol-chloroform extraction described by Sambrook et al. (1989). A similar protocol was applied to hair samples, and DNA was recovered by ultrafiltration with Microcon® YM-30 (Higuchi et al. 1988). We used a modified salt-chloroform method (Müllenbach et al. 1989) to extract DNA from ear tissues.

Fresh tissues samples were genotyped in singleplex PCR at 12 polymorphic microsatellite loci following protocols detailed in Gaubert *et al.* (2008). Hair samples were genotyped through single- or multiplex pre-amplification PCR followed by singleplex PCR. The multiplex pre-amplification step included Mix 1 (loci A5, C101, D4, D111, B103 and B105) and Mix 2 (loci A104, A108, A112, A113, B104, A110). Multiplexing reactions were performed using the QIAGEN Multiplex PCR kit (QIAGEN, Hilden, Germany), including 2X QIAGEN Multiplex PCR Master Mix (1X final concentration), 0.1-0.2 mM primer mix (forward and reverse) and 6 µl DNA template for a final volume of 25 µl. PCR cycling conditions followed the manufacturer's recommendations (QIAGEN Multiplex protocol), with annealing temperatures of 54 and 53°C for Mix 1 and Mix 2, respectively. The second PCR run was performed in singleplex for the twelve microsatellite loci following the conditions detailed in Gaubert *et al.* (2008a), adding 4 µl of pre-amplification product to a 20 µl final volume. For hair samples, we used the modified multiple-tube approach detailed in Ferrando et al. (2008) to circumvent genotyping errors due to null alleles, false alleles and allelic drop-out (Taberlet et al. 1996). We performed a minimum of four PCR replicates per

157 locus and sample. PCR products were analyzed on an ABI 3100 DNA sequencer and allele
158 size was scored with GeneMapper v. 4.0. (Applied Biosystems, Foster City, CA).
159
160 *Genetic diversity and structure*
161 The 12 loci were examined for null alleles and miss-scoring in each geographic population
162 with MICRO-CHECKER v.2.2.3. (Van Oosterhout et al. 2004). Locus B103 showed a
163 significant level of null alleles in all European populations ($P < 0.05$), and thus was not
164 considered in subsequent analyses. *In fine*, our data set consists of individuals with at least
165 seven genotyped loci. The mean number of loci genotyped per individual was 10.2.
166 Genetic diversity was measured as the average number of alleles (N_a), effective
167 number of alleles (N_e), observed (H_o) and expected (H_e) heterozygosities for each geographic
168 population using GenAlEx v.6.1 (Peakall and Smouse 2006) and FSTAT v. 2.9.3 (Goudet
169 2001). The proportion of shared alleles, or similarity (ps; Bowcock et al. 1994), between
170 European populations and Maghreb was calculated with Microsatellite analyzer (MSA) 4.05
171 (Dieringer and Schlötterer 2003). We applied the rarefication method implemented in HP-
172 RARE (Kalinowski 2005) to estimate allelic richness (A_R) and private allelic richness (PA_R).
173 We used GENEPOP v. 4.0 (Rousset 2008) to perform exact tests (Guo and Thompson 1992)
174 for deviations from Hardy-Weinberg equilibrium and linkage disequilibrium among loci for
175 all individuals and for each population within the introduced and native ranges. Significance
176 levels were calculated by the Markov-chain method using 10,000 dememorization steps, 500
177 batches and 1,000 subsequent iterations per batch to keep the standard error estimates <0.01
178 (Raymond and Rousset 1995). The inbreeding coefficient (F_{IS}) (Weir and Cockerham 1984)
179 was calculated for each population using FSTAT with 10,000 permutations. The false
180 discovery rate (FDR) technique was used to eliminate false assignment of significance by
181 chance under α=0.05 (Verhoeven *et al.* 2005).

We used STRUCTURE v.2.3.3 (Pritchard et al. 2000) to assess population structure in the Mediterranean rim and admixture events among European populations. We performed 20 independent runs for K=1–10 using $1x10^5$ Markov chain Monte Carlo (MCMC) iterations and a burnin of 10^4, assuming admixture and uncorrelated allele frequencies, with and without a priori information on geographic populations. The most likely value for K was estimated using the ΛK method (Evanno et al. 2005) as implemented in Structure Harvester (Earl and VonHoldt 2012). Population structure was also explored in a geographic context using Geneland v.3.2.2 (Guillot et al. 2005a; Guillot et al. 2005b) as implemented in R (R Development Core Team 2010). We estimated the most likely number of populations (K) according to genotypic and geographic data under the Dirichlet (D) model (Guillot et al. 2005a). The optimal K value was tested from K=1 to 10 under default parameters of number of nuclei in the Poisson Voronoi tessellation. D model was run five times for $5x10^5$ MCMC iterations, with the first $2x10^4$ iterations discarded as burnin. Then, the Falush (F) model (Falush et al. 2003) was run five times under the same conditions with the fixed optimal number of populations (K) as determined above. Maps of posterior probabilities of F model population membership were obtained and the run with the higher probability was selected. In both STRUCTURE and Geneland analyses, most likely K values were estimated including (i) all the geographic populations (Maghreb, Balearic Isl. and continental Europe) and (ii) continental Europe only.

Genetic differentiation among geographic populations was also calculated by pairwise estimates of F_{ST} (identical to the extended θ_{WC}) (Weir and Cockerham 1984) among geographic populations. Significant departure from zero was tested using 10,000 permutations in Arlequin v.3.01 (Excoffier et al. 2005). In addition, we carried out Principal Component Analysis (PCA) (i) among populations, on the basis of F_{ST} estimates using 10,000

206 permutations in PCAgen v. 1.2 (Goudet 1999), and (ii) among individuals, calculating a
207 squared distance matrix (Φ) (Smouse and Peakall 1999) in GenAlEx.
208 We tested for isolation-by-distance (IBD) at individual and population scales within
209 continental Europe by using, respectively, inter-individual genetic distances (a_r) (Rousset
210 2000) and a linearized Fst index (Fst/1-Fst) versus the logarithm of the geographic distance.
211 The $\hat{a}$ estimates of the a parameters described in Rousset (2000) were computed using
212 GENEPOP (Rousset 2008). Mantel tests were used to quantify the correlation between
213 genetic and geographic distances (r) after 10,000 permutations using the *mantel* function in
214 the VEGAN package for R (Oksanen et al. 2013).
215
216 *Test of alternative introduction scenarios*
217 We estimated the relative likelihood of alternative scenarios of introduction of the
218 common genet in Europe using approximate Bayesian computation (ABC; see Beaumont
219 2010) as implemented in DIYABC v.2.0.3 (Cornuet et al. 2014), combining our genotypic
220 data to a previously generated mtDNA dataset (cytochrome b and control region; Gaubert et
221 al. 2009; Gaubert et al. 2011) (Online Resource 1). DIYABC allows the elaboration and
222 comparison of complex scenarios involving bottlenecks, serial or independent introductions
223 and genetic admixture events as they are often suspected in introduced populations (Estoup
224 and Guillemaud 2010). The demographic parameters considered to model the scenarios are
225 the times of split or admixture events (in number of generations), the stable effective
226 population size, the effective number of founders in introduced populations, the duration of
227 the bottleneck during colonization, and the rate of admixture (whenever admixture occurs).
228 We draw six introduction scenarios following hypotheses posited in the literature and
229 our own results obtained with Bayesian clustering methods (Fig. 2; see below). Since
230 STRUCTURE and Geneland identified four groups (see Results) and that an introduction

hotspot was suspected in Catalonia (northeastern Spain) by our previous mtDNA analyses (Gaubert et al. 2009; Gaubert et al. 2011), we delimited a first set of five populations (set1) as follows: France (Fra), Iberia (Ibe; excluding Catalonia), Catalonia (Cat), Balearic Isl. (Bal) and Maghreb (Mag). Since (i) the clustering methods detected highly admixed populations neighboring Catalonia and (ii) there was a discordant genetic pattern between Ibiza and Mallorca+Cabrera that may imply different sources of introduction, we ran a second analysis with slightly different groupings (set2). In this case, Cat also included S_NBP and S_NCAST, and Ibizan samples were moved to Mag. Following various lines of evidence (Morales 1994; Amigues 1999; Gaubert et al. 2009; Gaubert et al. 2011), Maghreb (Pop5) was in each scenario considered as the native population from which European populations were introduced. Scenario 1 considered Fra, Ibe and Bal as originating from three independent introductions, with a secondary introduction of Cat from Bal. Scenario 2 was similar except that Cat was the result of admixture between Fra and Ibe. Scenario 3 fixed a single introduction event for continental Europe in Iberia followed by a natural spread into Cat and Fra, and an independent introduction in the Balearic Isl. Scenario 4 was similar to scenario 1 except that Fra originated from a secondary introduction event from Ibe. Scenario 5 was similar to scenario 4 except that Cat was the result of admixture between Fra and Bal, thus implying an introduction event from Bal (with a bottleneck). Scenario 6 was similar to scenario 4 except that Cat was the result of admixture between Fra and Ibe.

Prior distributions were uniform and set by default (Online Resource 2), with the exception of (i) the mtDNA mutation model fixed to TrN (Gaubert et al. 2009) and (ii) microsatellites and mtDNA mutation rates having their minimum and maximum distributions increased by a factor 10, respectively (so the fit of the observed data with the model simulations were improved). Microsatellite loci followed the generalized stepwise-mutation (GSM) model as implemented by default in DIYABC (Estoup et al. 2002). Priors were also

constrained to set up realistic posterior estimates as concerns times of split (prior range, by default: 10-10,000 generation time). The time of split between a given pair of primary introduced and secondary introduced populations was systematically fixed to be younger than the initial split between the related pair of native and introduced populations. Similarly, stable effective population size of the native population (Mag) was constrained to be higher than the effective number of founders in introduced populations.

The ABC method relies on summary statistics calculated from the dataset to represent the maximum amount of information in the simplest possible form (Sunnåker et al. 2013). DIYABC uses a series of standard, one sample and two sample summary statistics traditionally used in population genetics (see Cornuet et al. 2013). Following Cornuet et al. (2010), we used the largest series of summary statistics available in DIYABC, excluding a subset of one and two sample summary statistics that were used to check the goodness-of-fit of our dataset under the posterior predictive distribution of the model for the best scenario. Overall, our models (six scenarios) represented 26 historical parameters and 145 summary statistics applied to microsatellites and mtDNA sequences. The summary statistics used to assess the goodness-of-fit were mean size variance (one sample) and mean size variance and shared allele distance (two sample) for microsatellite data, and variance of pairwise differences and of numbers of the rarest nucleotide at segregating sites (one sample) and mean of pairwise differences (two sample) for mtDNA.

We simulated 6,000,000 datasets per scenario based on the coalescent model to produce robust ABC results, as recommended by the authors of DIYABC (Cornuet et al. 2013). As a first analytical step, we checked whether our dataset fitted the range of our pre-defined models (scenarios and parameter priors) using Principal Component Analysis representation on the summary statistics of the first 10% simulated datasets. We concluded that our six models were suitable for proceeding to the ABC analyses by evaluating the

position of our observed data relative to the distribution of summary statistics (Online Resource 3). The relative posterior probabilities of the different scenarios were calculated through polychotomous logistic regression from the 0.1% of simulated data sets most closely resembling the observed data using linear discriminant analysis on summary statistics (Estoup et al. 2012). Then, the posterior distributions of parameters were estimated under the most likely scenario by the logit transformation of parameters and linear regression on the 1% of simulated data sets most closely resembling the observed data. The power of our DIYABC analysis to discriminate between alternative scenarios was evaluated by simulating 500 pseudo-observed data sets per scenario with the same number of loci and individuals as our dataset. The relative posterior probabilities of each competing scenario were used to calculate type I and II errors for the most likely scenario.

Results

No significant linkage disequilibrium was detected between pairs of loci among all the individuals. Departures from Hardy-Weinberg equilibrium (HWE) were detected considering all loci and all individuals, but there was no significant departure for any geographic population ($P<0.001$; FDR correction for multiple comparisons, 1% nominal level). All the loci were polymorphic in all populations, except Cabrera and Ibiza (Balearic Isl.) that showed no allelic variability in eight and five loci, respectively. The mean number of alleles per population (*Na*) ranged from 1.4 to 4.8, and the allelic richness (A_R) ranged from 1.3 to 4.8 (Cabrera and western Maghreb as minimum and maximum values; Table 1). Mean expected heterozygosity values were moderately high in continental Europe populations (0.43-0.61) and lower in the Balearic Isl. (0.09-0.25), compared to the native range (0.63-0.71). Populations from southwestern Iberia had the highest heterozygosity levels in continental Europe (0.59-0.61). Northeastern Iberia and southwestern France had the highest inbreeding

coefficient (F_{IS}) values in continental Europe (0.11-0.13), whereas the highest values among all the populations were found in the Balearic Isl. (0.18-0.20). Performing Hardy–Weinberg exact test by locus and population and applying FDR correction for multiple comparisons, only four values of F_{IS} were found significant: at loci A108, C101 and D111 in southwestern France, and at locus A108 in western France sample (data not shown). However, average F_{IS} values across loci for each population were not significant. Private allelic richness (PA_R) was the highest in Maghreb (0.61-1.26), whereas southwestern Spain and southern Portugal had the greatest richness in Europe (0.11-0.24; see Table 1). The highest similarity values (shared alleles) between European populations and Maghreb (*ps*) were found in Ibiza (0.46) and Catalonia + the eastern, French Pyrenean border (0.43). The rest of the similarity values were slightly inferior (0.33-0.41).

The Bayesian clustering analysis with STRUCTURE identified four clusters (K = 4) within the studied species range (Table 2), including Iberia (cluster 1), France (cluster 2), Cabrera Isl. (cluster 3) and Maghreb + Ibiza Isl. (cluster 4). Mallorca Isl. was admixed between clusters 3 and 4 (posterior probabilities of assignment < 0.70). Individuals from Iberia and France were assigned to clusters 1 and 2 (respectively) when restricting our analysis to continental Europe. Admixed populations were found north of Iberia and at the French border (SF_NE, S_NBP, S_NCAST; Fig. 1 and Table 2). The maps of posterior probability obtained with GENELAND supported four similar clusters within the studied species range but only evidenced admixed populations in northwestern Iberia (i.e., lower posterior probability values of assignment to the two clusters in continental Europe; Online Resource 4).

Principal component analysis (PCA) plots showed genetic structure within the species distribution (Fig. 3). Maghreb, Balearic and continental Europe populations separated along PCI (45.05%, P < 0.05). Along PCII (16,20%, P < 0.05), Mallorca + Cabrera and Ibiza

(Balearic Isl.) separated and European populations stretched from south to north. PCA among individuals yielded a similar pattern, although less clear-cut geographically. The overall measure of genetic differentiation among all populations was high (F_{ST} = 0.276, P = 0.0001), with the majority of pairwise comparisons showing a significant differentiation (Online Resource 5).

The two clusters defined by STRUCTURE for continental Europe were used to delineate three groups in the IBD analysis: 'southern group' (cluster 1: most of Iberia), 'northern group' (cluster 2: most of France) and 'admixed' group (populations with global, posterior assignment probabilities < 0.70: northern Iberia and French boundary). Whereas Mantel tests showed significant positive relationships between pairwise a_r values and geographic distance within each continental group (r values: northern = 0. 1572; southern = 0. 2597; admixed = 0. 1406; in all cases, P < 0.01), a higher correlation was found when grouping all the individuals (r = 0.4052; P < 0.0001) (Fig. 4). At the population level, there was a highly significant correlation between geographic and genetic distances across continental Europe (r = 0.7322, P < 0.001). A similar level of correlation was found when removing the admixed group (northern and southern: r = 0.7377, P < 0.01) and the southern group (northern and admixed: r = 0.7239, P < 0.01). There was a lower correlation when the northern group was removed (southern vs admixed: r = 0.5497, P < 0.001). None of the IBD analyses among populations within the three groups were significant (P > 0.05), probably due to small sample size.

ABC simulations based on six alternative introduction scenarios gave strong support to scenario 3 (mean posterior probability: 0.92-0.97; 95% CI = 0.92-0.93 – 0.97-0.97) and similar posterior parameter estimates, independent of the sets (1 and 2) used. Scenario 3 implies two independent introductions from Maghreb into the Balearic Isl. and Iberia, followed by a spread throughout northeastern Iberia and France (see Fig. 2). The five other

scenarios showed much lower posterior probabilities (<0.07) (see Online Resource 6). False positive rates (type I errors) for scenario 3 were moderately low (24.2-23.9%), but false negative rates (type II errors) reached 56.2-55.5%. Scenario 6 contributed to 41.3-40.7 and 33.8-32.6% of the scenario 3 false positives and negatives, respectively. Our observed data were nested within the posterior predictive distribution of the model for scenario 3 based on alternative summary statistics, suggesting a good fit of scenario 3 with the observed data (Online Resource 7). Randomly modifying the choice of summary statistics used to assess the goodness-of-fit of our observed data had no influence on the choice of the best scenario, neither had the use of separated datasets (i.e. mtDNA versus microsatellites; data not shown).

The posterior distribution estimates of effective population sizes, duration of bottlenecks and locus-specific evolutionary model parameters were in general little informative (Online Resource 2). Given that the observed genetic diversity of our dataset is expected to be the product of the effective population size of our source population (Mag) and mutation rates (i.e. the majority of mutations/alleles must have existed before the first introduction events), we could not estimate independently effective population sizes and mutation rates. Nevertheless, estimates of time parameters describing introduction and expansion events presented the only informative posterior distributions (sharper distribution and narrower confidence interval than prior ranges) and should thus be usable (see Fountain et al. 2014). Those latter represented the only informative posterior distributions (more pointy aspect and narrower confidence interval than prior ranges). The median value was t4=3,130 (set2) – 3,320 (set1) generation times (HPD 95=853-7,610 – 1,150-8,270) for the introduction event in Iberia, t3=1,320 (set1) – 1,790 (set2) generation times for the introduction in the Balearic Isl. (HPD 95=372-6,060 – 554-6,730) and tb=717 (set2) – 797 (set1) generation times for the expansion from Iberia into Catalonia and France (HPD 95=288-1,570 – 338-1,670). Adjusting priors to reflect more realistic effective population sizes (N>10,000) and

mtDNA mutation rate (Gaubert et al. 2009) or to restrict introduction times to more recent periods (e.g. 10-2,000 generation times) systematically resulted in the observed data being off the distribution of the simulated datasets.

Discussion

Scenario of introduction of the common genet from Maghreb to Europe

The descriptive analysis of microsatellite diversity confirmed Maghreb as the source population (Gaubert et al. 2009; Gaubert et al. 2011) through generally lower allelic diversity and levels of heterozygosity in European populations (Nei et al. 1975; Tsutsui et al. 2000). Our approach using ABC modelling allowed us to explicitly test, for the first time, different scenarios of introduction of the common genet in Europe combining multilocus data (microsatellites and mtDNA). ABC simulations yielded strong support for scenario 3, implying two independent introductions from Maghreb into the Balearic Isl. and Iberia (Fig. 2). Considering a generation time of two years in the common genet (Delibes and Gaubert 2013), the posterior distribution estimate of the introduction event in Iberia (6,640 [2,300-16,540] ya for set1 and 6,260 [1,706-15,220] ya for set2) largely antedated the invasion of Iberia by Muslim armies. Although this range is rather large, it covers a realistic time frame from the first trans-migration of humans via the Strait of Gibraltar (Upper Palaeolithic; Derricourt 2005) to the end of the Phoenician influence in the Mediterranean Basin (300 BC; Elayi 2013). It is also congruent with the earliest estimate of transportation of a small carnivore (*Mustela nivalis*) into Mediterranean islands c. 10,000 ya (Lebarbenchon et al. 2010).

Our analyses broadened the Andalusia 'introduction hotspot' to southwestern Iberia (i.e., including southwestern Spain and southern Portugal). Indeed, the highest genetic

406 diversity and private allelic richness observed in southern Portugal mirrored the high genetic
407 diversity found in southwestern Spain and added to the peak of unique mtDNA haplotypes
408 found in Andalusia (Gaubert et al. 2009; Gaubert et al. 2011). The introduction hotspot found
409 in southwestern Iberia superimposes well with the Tartessian Kingdom's zone of cultural
410 influence from 1,200 to 550 BC (Fig. 1; Chamorro 1987). Previous studies investigating
411 historiographical sources and mtDNA diversity have invoked the possible role of Antique
412 civilizations in the introduction of the common genet into southern Iberia (Amigues 1999;
413 Gaubert et al. 2009; Gaubert et al. 2011). Indeed, the Greek author Herodotus mentioned the
414 use of the common genet as a bio-control agent against the pullulating of rabbits at the time of
415 the Tartessos Kingdom, 600 BC (Amigues 1999). Tartessos was a rich harbor city located on
416 the western Andalusian coast, which had vibrant trading exchanges with Phoenicians and
417 their nearby commercial harbor in the present-day Cadiz since at least 800 BC (Moscati
418 1996). It is thus possible that Phoenicians (or Greeks, whom shared almost contemporaneous,
419 similar trading routes) introduced the common genet from their North African colonies
420 through their trading activities with Tartessians. Indeed, Phoenicians are considered the
421 earliest trans-Mediterranean colonizers to southern Iberia having spread 'exotic' vertebrate
422 species (Muñiz et al. 1995).
423 The high genetic diversity and private allelic richness found in southern Portugal and
424 southwestern Spain suggest a bottleneck effect followed by a high rate of population growth
425 (Nei et al. 1975) or long-term genetic drift (Allendorf and Lundquist 2003), in agreement with
426 the ancient time of introduction inferred from ABC. Since multiple introductions may
427 overcome founder effect by producing high genetic diversity in non-native populations
428 (Kolbe et al. 2004) and facilitating rapid adaptation and expansion in the invaded range (Lee
429 2002), we posit that the common genet in southern Iberia might have repeatedly —and
430 deliberately— been introduced from different sources in Maghreb. This hypothesis is also

supported by the co-occurrence of two divergent mtDNA lineages in southwestern Iberia (Gaubert et al. 2009). Although we could not rule on the precise period of introduction of the common genet in Iberia, our results altogether reject the hypothesis of the species being introduced by Muslims from the 8th century, and reinforce the potential role of Phoenicians even though earlier transportations cannot be ruled out.

ABC simulations identified the Balearic Isl. as the second site of introduction of the common genet in Europe, whether Ibiza was grouped with Maghreb or the Balearic Isl. The microsatellite data confirmed that the Balearic Isl. constituted multiple sites of introduction for the common genet (Delibes 1977; Clevenger 1993; Gaubert et al. 2009; Gaubert et al. 2011). Island populations were genetically close to Maghreb and were mainly assigned to two distinct groups by clustering analyses, including (i) Ibiza, significantly differentiated from but grouped with Maghreb, and (ii) Cabrera. Mallorca being admixed between the two genetic clusters and having a small sample size, it remains difficult to establish a definitive scenario of introduction among the Balearic Isl. Nevertheless, our results suggested that severe bottlenecks followed by weak population growth occurred after the introduction of common genets on the islands. The Balearic populations had the lowest allelic richness and levels of heterozygosity in Europe: such a pattern is expected after the introduction of low founder population sizes on relatively small islands followed by genetic drift (Frankham 1998; Broders et al. 1999; Allendorf and Lundquist 2003). Given our limited sample set, we could not clarify whether genets were transported independently to Mallorca and Ibiza, or were translocated from Ibiza to Mallorca in a similar way that they were more recently from Mallorca to Cabrera for regulating rabbit populations (Delibes 1977). Neither we could refine the time frame at which the species was introduced on the islands, given the wide confidence interval that we obtained from ABC estimates either including Ibiza 2,640 [744-12,120] ya, set1) or not (3,580 [1,108-13,460] ya, set2). The influence of the Phoenicians remains

conceivable given that (i) the Antique history of the Balearic Isl. was at its early stage under the influence of Carthaginians (Phoenician colony from Tunisia), as reflected in the Baleares peopling and trade items (Tomàs et al. 2006; Segert 2007), and (ii) one of the first records of a small carnivore (*Mustela nivalis*) in the Balearic Isl. was associated to Carthaginians' remains (Masseti 1995).

Post-introduction scenario in continental Europe

ABC estimates selected the sole introduction scenario (scenario 3) that involved the natural spread of the common genet from Iberia to Catalonia (northeastern Spain) and France. Although the posterior probabilities for choosing this scenario were high (0.92-0.97), a fair rate of type I errors (24%) and more importantly, a high rate of type II errors (56%) indicate that other scenarios may well fit our dataset. Noticeably, scenario 6 that involved a secondary introduction from Iberia to France and subsequent admixture in Catalonia between French and Iberian populations, was responsible for 41 and 33-34% of the false positives and negatives, respectively. We thus consider that ABC did not resolve the issue of whether the French populations originated from a geographic spread of introduced Iberian populations (dispersal hypothesis), or were subsequently introduced from Iberia (bridgehead hypothesis).

At first sight, scenario 3 may appear in conflict with our clustering analyses based on microsatellite data in continental Europe, which inferred two genetic groups distributed in western Iberia and France. However, one should keep in mind that the a priori delineation of 'populations' in ABC does not correspond to a test of population structure, but to simulations of population history (Cornuet et al. 2010). Moreover, STRUCTURE and other clustering methods have been shown to be sensitive to the isolation-by-distance (IBD) effect, marked IBD patterns potentially leading to the biased identification of population clusters along the IBD gradient (Frantz et al. 2009; Meirmans 2012). The strong IBD pattern exhibited within

481 the genotyped populations of Europe is actually in line with scenario 3, which involves
482 northern dispersal from Iberia of the common genet. Indeed, a significant IBD signal may
483 reflect a natural process of diffuse dispersal from introduction sites (e.g., Henry et al. 2009),
484 consistent with the capacity of this small carnivore to colonize new areas in Europe from
485 source populations (Gaubert et al. 2008b). Thus, our results could argue for the biased
486 estimate of two population clusters in continental Europe. Nevertheless, the incidence in
487 southwestern France of (i) a fair proportion of alleles shared with Maghreb and (ii) reduced
488 levels of heterozygosity (low *Ho*, high F_{IS}) could be the signature of a secondary introduction
489 from southern Europe (bridgehead hypothesis), in line with scenario 6 and as suggested by the
490 distribution of mtDNA haplotypes (Gaubert et al. 2011). In this case, the IBD pattern also
491 found within the French and Iberian population clusters (among individuals) and the lower
492 genetic diversity observed in the other French populations would support a scenario of
493 dispersal from both southwestern Iberia and western France leading to subsequent admixture
494 in a contact zone somewhere between the north of Iberia and the French border (including
495 Catalonia). Although we cannot estimate at which period the species could have been
496 introduced in southwestern France (posterior estimates for scenario 6 were uninformative),
497 the Middle Age fashion for the common genet at French courts (Delort 1978; Gaubert and
498 Mézan-Muxart 2010; Mézan-Muxart 2010) or older practices of offering live small carnivores
499 as political gifts (Morales Muñiz 2000) constitute potential landmarks.
500 The Catalonia 'introduction hotspot' was not recovered as a probable scenario by the
501 ABC analysis. Instead, DIYABC favored a scenario of ancestral coalescence (scenario 3) or
502 secondary admixture (scenario 6) of Catalonia and northern Spanish populations with French
503 and Iberian populations, and consistently disfavored scenarios fixing Catalonia as a
504 population introduced from the Balearic Isl. Nevertheless, we observed (i) lower levels of
505 heterozygosity than in other parts of Iberia (low *Ho*, high F_{IS}), (ii) a high proportion of alleles

shared with Maghreb and (iii) the assignment of a small fraction of individuals to cluster 3 (distributed in Mallorca and Cabrera) that could be a signal of introduction of common genets from the Balearic Isl. or Maghreb to Catalonia. It is possible that long-term admixture between Iberian and French populations around the latitude of Catalonia complicated the detection of introduced populations in Catalonia using microsatellite data. All the more since organelle genomes such as mtDNA, which are capable of retaining the genetic imprint of old introductions (Searle et al. 2009), gave a clear signal of independent introduction in Catalonia, either directly from Maghreb or via the Balearic Isl. (Gaubert et al. 2009; Gaubert et al. 2011).

Finally, clustering analyses revealed a high level of admixture between Iberian and French populations of common genets in northern Iberia and at the French boundary, in line with the wide diffusion of the mtDNA haplotype "H1" observed across southwestern Europe (Gaubert et al. 2009; Gaubert et al. 2011). Whether such admixture may have influenced the local fitness of the species remains unknown.

Conclusion

ABC estimates and descriptive analyses of genetic diversity congruently pointed to two primary introductions of the common genet from Maghreb to Iberia and the Balearic Isl. Nevertheless, given the level of uncertainty in the choice of scenarios 3 and 6, ABC failed to reject the hypothesis of a secondary introduction from Iberia to western France. ABC also failed to detect the Catalonian population as introduced from the Balearic Isl. as a likely scenario. Causes of incongruence between ABC and descriptive analyses of genetic data have been poorly explored and were not the scope of our investigations. We anticipate that they could have been due to low sample size (and thus low statistical power), misspecification of priors, conflict between microsatellites and mtDNA prior requirements (see Templeton 2009), and the non-consideration by the ABC approach of genetic diversity patterns within

populations. Eventually, we acknowledge that a wider sampling (e.g. more samples from Maghreb and the Balearic Isl.) and genetic coverage (i.e., more loci) will be necessary to improve the accuracy of the descriptive analyses of genetic diversity in the common genet.

Nevertheless, the descriptive evidence (microsatellite and mitochondrial diversity patterns) used within the permissive range of ABC estimates of scenario choice (scenarios 3 and 6) allowed us to propose that, together with introductions in the Balearic Isl., the common genet may have established thrice in southwestern continental Europe (in southern Iberia, southwestern France and Catalonia). The microsatellite data suggested a scenario of post-introduction gene flow and genetic drift as structuring geographical genetic variation in the invaded range, which is not concordant with the hypothesis of an artificial dispersal of the species by Muslims. Our conclusion raises the question of the specific use of the common genet by humans in historical times, a point that remains almost undocumented in the archaeozoological and historiographical records with the exception of the 'stories' of Herodotus. Further investigations covering the fields of archaeozoology and evolutionary genetics will have to be conducted in the Mediterranean Basin to better understand the factors having influenced the successful introduction of the common genet in Europe.

Acknowledgments

We thank the persons and institutions involved in the collecting of the genetic samples (see Gaubert et al. 2009, 2011). Lab work was funded by the European Commission 6th PCRDT ''EPISARS'' (FP6-2003-SSP-2-SARS; no. 51163). Three anonymous reviewers improved an earlier version of the manuscript.

References

556 Allendorf FW, Lundquist LL (2003) Introduction: population biology, evolution, and control
557 of invasive species. Cons Biol 17:24-30
558 Amigues S (1999) Les belettes de Tartessos. Anthropozool 29:55-64
559 Beaumont MA (2010) Approximate Bayesian computation in evolution and ecology. Annu
560 Rev Ecol Evol Syst 41:379-406
561 Blondel J, Vigne JD (1993) Space, time, and man as determinants of diversity of birds and
562 mammals in the Mediterranean region. In: Ricklefs RE and Schluter D (eds) Species
563 diversity in ecological communities: historical and geographical perspectives. The
564 University of Chicago Press, Chicago and London, pp 135-146
565 Bowcock AM, Ruiz-Linares A, Tomfohrde J, et al. (1994) High resolution of human
566 evolutionary trees with polymorphic microsatellites. Nature 368:455-457
567 Broders HG, Mahoney SP, Montevecchi WA, et al. (1999) Population genetic structure and
568 the effect of founder events on the genetic variability of moose, *Alces alces*, in
569 Canada. Mol Ecol 8:1309-1315
570 Chamorro JG (1987) Survey of archaeological research on Tartessos. Am J Archaeol 91:197-
571 232
572 Ciolek TM (2011) 1999 - present. Georeferenced historical transport/travel/communication
573 routes and nodes. OWTRAD Dromographic Digital Data Archives (ODDDA): Old
574 World Trade Routes (OWTRAD) Project. www.ciolek.com - Asia Pacific Research
575 Online, Canberra, Australia,
576 Clevenger A (1993) Status of martens and genets in the Balearic and Pityusic Islands, Spain.
577 Sm Carn Cons 9:18-19
578 Cornuet J-M, Pudlo P, Veyssier J, et al. (2014) DIYABC v2.0: a software to make
579 approximate Bayesian computation inferences about population history using single

580 nucleotide polymorphism, DNA sequence and microsatellite data. Bioinf 30:1187-
581 1189
582 Cornuet J-M, Ravigne V, Estoup A (2010) Inference on population history and model
583 checking using DNA sequence and microsatellite data with the software DIYABC
584 (v1.0). BMC Bioinfo 11:401
585 Cornuet JM, Pudlo P, Veyssier J, et al. (2013) DIYABC version 2.0: A user-friendly software
586 for inferring population history through Approximate Bayesian Computations using
587 microsatellite, DNA sequence and SNP data. Institut National de la Recherche
588 Agronomique, Campus International de Baillarguet, Saint-Gély-du-Fesc, France,
589 Crawford KM, Whitney KD (2010) Population genetic diversity influences colonization
590 success. Mol Ecol 19:1253-1263
591 Cucchi T, Vigne J-D, Auffray J-C (2005) First occurrence of the house mouse (*Mus musculus*
592 *domesticus* Schwarz & Schwarz, 1943) in the Western Mediterranean: a
593 zooarchaeological revision of subfossil occurrences. Biol J Linn Soc 84:429-445
594 Davies N, Villablanca FX, Roderick GK (1999) Determining the source of individuals:
595 multilocus genotyping in nonequilibrium population genetics. TREE 14:17-21
596 Delibes M (1977) Sobre las ginetas de la isla de Ibiza (*Genetta genetta isabelae* n. ssp.).
597 Doñana, Acta Vertebrata 4:139-160
598 Delibes M, Gaubert P (2013) *Genetta genetta* Common Genet (Small-spotted Genet). In:
599 Kingdon JS and Hoffmann M (eds) The Mammals of Africa, Vol. 5. Carnivores,
600 Pangolins, Equids and Rhinoceroses edn. Bloomsburry Publishing, London, UK, pp
601 224-229
602 Delort R (1978) Le commerce des fourrures en Occident à la fin du Moyen-Âge. Ecole
603 Française de Rome, Palais Farnèse, Roma, Italy

Derricourt R (2005) Getting "out of Africa": sea crossings, land crossings and culture in the hominin migrations. JWP 19:119-132

Dieringer D, Schlötterer C (2003) Microsatellite analyser (MSA): a platform independent analysis tool for large microsatellite data sets. Mol Ecol Not 3:167–169

Dlugosch KM, Parker IM (2008) Founding events in species invasions: genetic variation, adaptive evolution, and the role of multiple introductions. Mol Ecol 17:431-449

Dobson M (1998) Mammal distributions in the western Mediterranean: the role of human intervention. Mamm Rev 28:77-88

Earl DA, VonHoldt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Cons Genet Res 4:359-361

Elayi J (2013) Histoire de la Phénicie. Perrin, Paris, France

Estoup A, Guillemaud T (2010) Reconstructing routes of invasion using genetic data: why, how and so what? Mol Ecol 19:4113-4130

Estoup A, Jarne P, Cornuet J-M (2002) Homoplasy and mutation model at microsatellite loci and their consequences for population genetics analysis. Mol Ecol 11:1591–1604

Estoup A, Lombaert E, Marin J-M, et al. (2012) Estimation of demo-genetic model probabilities with Approximate Bayesian Computation using linear discriminant analysis on summary statistics. Mol Ecol Res 12:846-855

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software structure: a simulation study. Mol Ecol Not 14:2611–2620

Excoffier L, Laval G, Schneider S (2005) Arlequin ver. 3.0: An integrated software package for population genetics data analysis. Evolutionary Bioinformatics Online 1:47-50

Falush D, Stephens M, Pritchard J (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics 164:1567–1587

629 Ferrando A, Lecis R, Domingo-Roura X, et al. (2008) Genetic diversity and individual
630 identification of reintroduced otters (*Lutra lutra*) in north-eastern Spain by DNA
631 genotyping of spraints. Cons Biol 9:129-139
632 Forsman A, Ahnesjö J, Caesar S, et al. (2008) A model of ecological and evolutionary
633 consequences of color polymorphism. Ecology 89:34-40
634 Fountain T, Duvaux L, Horsburgh G, et al. (2014) Human-facilitated metapopulation
635 dynamics in an emerging pest species, *Cimex lectularius*. Mol Ecol 23:1071-1084
636 Frankham R (1998) Inbreeding and extinction: island populations. Cons Biol 12:665-675
637 Frantz AC, Cellina S, Krier A, et al. (2009) Using spatial Bayesian methods to determine the
638 genetic structure of a continuously distributed population: clusters or isolation by
639 distance? J Appl Ecol 46:493-505
640 Gaubert P, del Cerro I, Palomares F, et al. (2008a) Development and characterization of 11
641 microsatellite loci in a historically introduced carnivoran, the common genet (*Genetta*
642 *genetta*). Mol Ecol Res 8:1375-1377
643 Gaubert P, Godoy JA, del Cerro I, et al. (2009) Early phases of a successful invasion:
644 mitochondrial phylogeography of the common genet (*Genetta genetta*) within the
645 Mediterranean Basin. Biol Inv 11:523-546
646 Gaubert P, Jiguet F, Bayle P, et al. (2008b) Has the common genet (*Genetta genetta*) spread
647 into south-eastern France and Italy? It J Zool 75:43-57
648 Gaubert P, Machordom A, Morales A, et al. (2011) Comparative phylogeography of two
649 African carnivorans presumably introduced into Europe: disentangling natural versus
650 human-mediated dispersal across the Strait of Gibraltar. J Biog 38:341-358
651 Gaubert P, Mézan-Muxart V (2010) Where have the ‘black genets’ gone? A likely restriction
652 of melanistic cases of the common genet (*Genetta genetta*) to its introduced range.
653 Mamm Biol 75:353-357

654 Goudet J (1999) PCA-GEN. version 1.2. Lausanne, Switzerland.
655 www.unil.ch/izea/softwares/pcagen.html,
656 Goudet J (2001) FSTAT, version 2.9.3. A program to estimate and test gene diversities and
657 fixation indices. Lausanne University, Lausanne, Switzerland.
658 http://www2.unil.ch/popgen/softwares/fstat.htm,
659 Guillot G, Estoup A, Mortier F, et al. (2005a) A spatial statistical model for landscape
660 genetics. Genetics 170:1261-1280
661 Guillot G, Mortier F, Estoup A (2005b) GENELAND: a computer package for landscape
662 genetics. Mol Ecol Not 5:712–715
663 Guo SW, Thompson EA (1992) Performing the exact tests of Hardy–Weinberg proportion for
664 multiple alleles. Biometrics 48:361–372
665 Henry P, Le Lay G, Goudet J, et al. (2009) Reduced genetic diversity, increased isolation and
666 multiple introductions of invasive giant hogweed in the western Swiss Alps. Mol Ecol
667 18:2819-2831
668 Higuchi R, von Beroldingen CH, Sensabaugh GF, et al. (1988) DNA typing form single hairs.
669 Nature 332:543-546
670 Jones EP, Eager HM, Gabriel SI, et al. (2013) Genetic tracking of mice and other bioproxies
671 to infer human history. Trends Genet 29:298-308
672 Kalinowski ST (2005) HP-RARE 1.0: a computer program for performing rarefaction on
673 measures of allelic richness. Mol Ecol Not 5:187-189
674 Khamis FM, Karam N, Ekesi S, et al. (2009) Uncovering the tracks of a recent and rapid
675 invasion: the case of the fruit fly pest *Bactrocera invadens* (Diptera: Tephritidae) in
676 Africa. Mol Ecol 18:4798-4810
677 Kolbe JJ, Glor RE, Rodriguez Schettino L, et al. (2004) Genetic variation increases during
678 biological invasion by a Cuban lizard. Nature 431:177-181

679 Lebarbenchon C, Poitevin F, Arnal V, et al. (2010) Phylogeography of the weasel (*Mustela*
680 *nivalis*) in the western-Palaearctic region: combined effects of glacial events and
681 human movements. Heredity 105:449-462
682 Lee CE (2002) Evolutionary genetics of invasive species. TREE 17:386-391
683 Masseti M (1995) Quaternary biogeography of the Mustelidae family on the Mediterranean
684 islands. Hystrix 7:17-34
685 Masseti M (2009) Mammals of the Mediterranean islands: homogenisation and the loss of
686 biodiversity. Mammalia 73:169-202
687 Meirmans PG (2012) The trouble with isolation by distance. Mol Ecol 21:2839-2846
688 Mézan-Muxart V (2010) Genette et janette: devises de Jeanne de France au XV^{e} siècle.
689 Reinardus 22:104-125
690 Morales A (1994) Earliest genets in Europe. Nature 370:512-513
691 Morales Muñiz DC (2000) La fauna exótica en la Península Ibérica: apuntes para el estudio
692 del coleccionismo animal en el Medievo hispánico. Espacio, Tiempo y Forma, Serie
693 III, H^{a} Medieval 13:233-270
694 Moscati S (1996) La grande Cadice dei Fenici. Accademia nazionale dei Lincei, Roma, Italy
695 Müllenbach R, Lagoda PJ, Welter C (1989) An efficient salt-chloroform extraction of DNA
696 from blood and tissues. Trends Genet 5:391
697 Muñiz AM, Pecharroman MAC, Carrasquilla FH, et al. (1995) Of mice and sparrows:
698 commensal faunas from the Iberian Iron Age in the Duero valley (Central Spain). Int J
699 Osteoarchaeol 5:127-138
700 Nei M, Maruyama T, Chakraborty R (1975) The bottleneck effect and genetic variability in
701 populations. Evolution 29:1-10
702 Oksanen JF, Blanchet G, Kindt R, et al. (2013) vegan: Community Ecology Package. R
703 package version 2.0-8. http://CRAN.R-project.org/package=vegan,

704 Peakall R, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic
705 software for teaching and research. Mol Ecol Not 6:288-295
706 Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using
707 multilocus genotype data. Genetics 155:945-959
708 R Development Core Team (2010) R: A language and environment for statistical computing.
709 R Foundation for Statistical Computing, Vienna, Austria
710 Raymond M, Rousset F (1995) An exact test for population differentiation. Evolution
711 49:1280-1283
712 Roman J, Darling JA (2007) Paradox lost: genetic diversity and the success of aquatic
713 invasions. TREE 22:454-464
714 Rousset F (2000) Genetic differentiation between individuals. J Evol Biol 13:58-62
715 Rousset F (2008) GenePop'007: a complete re-implementation of the GenePop software for
716 Windows and Linux. Mol Ecol Res 8:103-106
717 Sakai AK, Allendorf FW, Holt JS, et al. (2001) The population biology of invasive species.
718 Ann Rev Ecol Syst 32:305-332
719 Sambrook E, Fritsch F, Maniatis T (1989) Molecular cloning. Cold Spring Harbor Press, New
720 York
721 Searle JB (2008) The genetics of mammalian invasions: a review. Wild Res 35:185-192
722 Searle JB, Jamieson PM, Gündüz İ, et al. (2009) The diverse origins of New Zealand house
723 mice. Proc R Soc B: Biol Sci 276:209-217
724 Segert S (2007) Phoenician and Punic morphology. In: Kaye AS (ed) Morphology of Asia
725 and Africa Vol 1. Eisenbrauns, Winona Lake, IN, pp 75-84
726 Smouse PE, Peakall R (1999) Spatial autocorrelation analysis of individual multiallele and
727 multilocus genetic structure. Heredity 82:561-573

728 Sunnåker M, Busetto AG, Numminen E, et al. (2013) Approximate Bayesian Computation.
729 PLoS Comput Biol 9:e1002803
730 Templeton AR (2009) Statistical hypothesis testing in intraspecific phylogeography: nested
731 clade phylogeographical analysis vs. approximate Bayesian computation. Mol Ecol
732 18:319-331
733 Tomàs C, Jiménez G, Picornell A, et al. (2006) Differential maternal and paternal
734 contributions to the genetic pool of Ibiza Island, Balearic Archipelago. Am J Phys
735 Anthrop 129:268-278
736 Tsutsui ND, Suarez AV, Holway DA, et al. (2000) Reduced genetic variation and the success
737 of an invasive species. Proc Natl Acad Sci USA 97:5948-5953
738 Van Oosterhout C, Hutchinson WF, Wills DPM, et al. (2004) Micro-checker: software for
739 identifying and correcting genotyping errors in microsatellite data. Mol Ecol Not
740 4:535-538
741 Verhoeven KJF, Macel M, Wolfe LM, et al. (2011) Population admixture, biological
742 invasions and the balance between local adaptation and inbreeding depression. Proc R
743 Soc B: Biol Sci 278:2-8
744 Vigne J-D, Zazzo A, Saliège J-F, et al. (2009) Pre-Neolithic wild boar management and
745 introduction to Cyprus more than 11,400 years ago. Proc Natl Acad Sci USA
746 106:16135-16138
747 Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population
748 structure. Evolution 38:1358-1370
749

Table 1 Geographic delimitation and genetic diversity of the populations of common genets in Europe and the Mediterranean Basin. *n* = sample size; $N_{a\ =}$ average number of alleles; N_e = effective number of alleles; A_R = allelic richness; H_o = observed heterozygosity; H_e = expected heterozygosity; F_{IS} = inbreeding coefficient; PA_R= private allelic richness; *ps* = proportion of shared alleles (similarity) between European populations and Maghreb. * including Mallorca.

Introduced range		***Genetic diversity indices***							
		n	N_a	N_e	A_R	H_o	H_e	F_{IS}	PA_R	***ps***
	Continental Europe									
F_HN	Haute-Normandie, France	1	—	—	—	—	—	—	—	—
F_NW	Pays-de-la-Loire, Poitou-Charentes, France	30	3.5	2.1	2.6	0.44	0.44	0.02	0.03	0.39
F_CEN	Centre and Massif Central, France	8	2.8	2.1	2.6	0.47	0.43	-0.03	0.00	0.33
F_SW	Gironde, Dordogne, Landes and Lot-et-Garonne, France	28	4.1	2.4	3.0	0.48	0.53	0.11	0.07	0.41
FI_MED	Languedoc-Roussillon, Rhône-Alpes, Provence-Alpes-Côte-d'Azur, France and Piemont, Italy	9	3.1	2.2	2.8	0.48	0.47	0.07	0.06	0.36
F_HP	Hautes-Pyrénées	1	—	—	—	—	—	—	—	—
SF_NE	Catalonia and Pyrénées-Orientales, Spain	18	4.0	2.4	3.0	0.47	0.52	0.13	0.04	0.43
S_NBP	Basque Province and Castilla-Leon, Spain	13	3.2	2.4	2.8	0.58	0.54	-0.02	0.05	0.36
S_NCAST	Cantabria and Asturias, Spain	8	3.3	2.6	3.0	0.61	0.53	-0.07	0.06	0.39
SP_NW	Galicia, Spain and North Portugal	20	4.0	2.3	3.1	0.52	0.50	-0.01	0.05	0.39
P_C	Centre, Portugal	4	—	—	—	—	—	—	—	—
P_S	Alentejo, Portugal	11	4.1	2.7	3.3	0.66	0.59	-0.07	0.24	0.40

S_SW	Western Andalusia and Extremadura, Spain	7	3.8	2.8	3.6	0.68	0.61	-0.04	0.11	0.40
S_CEN	Madrid, Spain	2	—	—	—	—	—	—	—	—
S_SE	Eastern Andalusia and Murcia, Spain	4	—	—	—	—	—	—	—	—
	Balearic Isl.									
S_MAL	Mallorca Isl., Spain	3	—	—	—	—	—	—	—	—
S_CAB	Cabrera Isl., Spain	15	1.4	1.2	1.3	0.08	0.09	0.18	0.10*	0.26
S_IBI	Ibiza Isl., Spain	5	1.7	1.5	1.7	0.25	0.25	0.20	0.04	0.46
Native range	**Geographic area**	***n***	N_a	N_e	A_R	H_o	H_e	F_{IS}	PA_R	***ps***
	Maghreb									
M_MOMA	Morocco, western Algeria and Mauritania	5	4.8	3.8	4.8	0.67	0.71	0.16	1.26	—
M_ALTU	Eastern Algeria and Tunisia	7	4.5	3.1	4.3	0.62	0.63	0.09	0.61	—

755
756

757
758 **Table 2** Bayesian clustering analysis among populations of the common genet in the Mediterranean range (A; K= 4) and continental Europe (B;

759 K = 2). Proportions of membership to predefined geographic populations (see Fig. 1) are given. See Table 1 for population acronyms.

760

	A) Clusters within the Mediterranean Basin					**B) Clusters within continental Europe**		
		1	2	3	4		1	2
Continental	F_HN	0.111	0.694	0.007	0.189	F_HN	0.315	0.725
Europe	F_NW	0.008	0.978	0.001	0.013	F_NW	0.011	0.989
	F_CEN	0.104	0.89	0.003	0.003	F_CEN	0.078	0.922
	F_SW	0.088	0.864	0.001	0.047	F_SW	0.117	0.883
	FI_MED	0.097	0.901	0	0.002	FI_MED	0.11	0.89
	F_HP	0.627	0.353	0.02	0	F_HP	0.248	0.752
	SF_NE	0.398	0.52	0.08	0.002	SF_NE	0.515	0.485
	S_NBP	0.362	0.63	0.006	0.003	S_NBP	0.353	0.647
	S_NCAST	0.362	0.572	0.014	0.052	S_NCAST	0.563	0.437
	SP_NW	0.760	0.239	0	0	SP_NW	0.881	0.119
	P_CEN	0.975	0.016	0.005	0.004	P_CEN	0.963	0.037
	P_S	0.980	0.014	0.002	0.003	P_S	0.985	0.015
	S_SW	0.770	0.171	0.02	0.039	S_SW	0.868	0.132
	S_CEN	0.635	0.24	0.006	0.118	S_CEN	0.908	0.092

	S_SE	0.725	0.259	0.006	0.01	S_SE	0.732	0.268
Balearic Isl.	S_MAL	0.008	0.01	0.646	0.336			
	S_CAB	0.001	0.001	0.998	0.001			
	S_IBI	0.006	0.006	0.035	0.953			
Maghreb	M_MOMA	0.008	0.027	0.023	0.941			
	M_ALTU	0.163	0.016	0.004	0.817			

761

762 **Figure legends.**

763

764 **Fig. 1** Genetic structure of the common genet in the Mediterranean Basin inferred from
765 microsatellites. a- Map of the geographic populations superimposed to the species distribution
766 range (in green) in the Mediterranean Basin. Pie charts represent the proportional membership
767 of individuals to clusters inferred from STRUCTURE (K = 4). Circles are proportional to the
768 number of individuals. b- Plots of the probabilistic assignments inferred by STRUCTURE per
769 individuals and populations. See Table 1 for population acronyms.

770

771 **Fig. 2** DIYABC graphical representation of the six alternative introduction scenarios of the
772 common genet in Europe used for approximate Bayesian computation simulations.
773 Native population is Maghreb. Time is not to scale. See Material and Methods for details on
774 the scenarios and model parameters. t1-t4 = times of introduction events; ta-tc = times of
775 admixture events; N1-N5 = stable effective population sizes; N1b-N4b = effective numbers of
776 founders in introduced populations; da, dh-dl = times of end of bottleneck since introduction
777 or admixture events; ra-rc = rates of admixture.

778

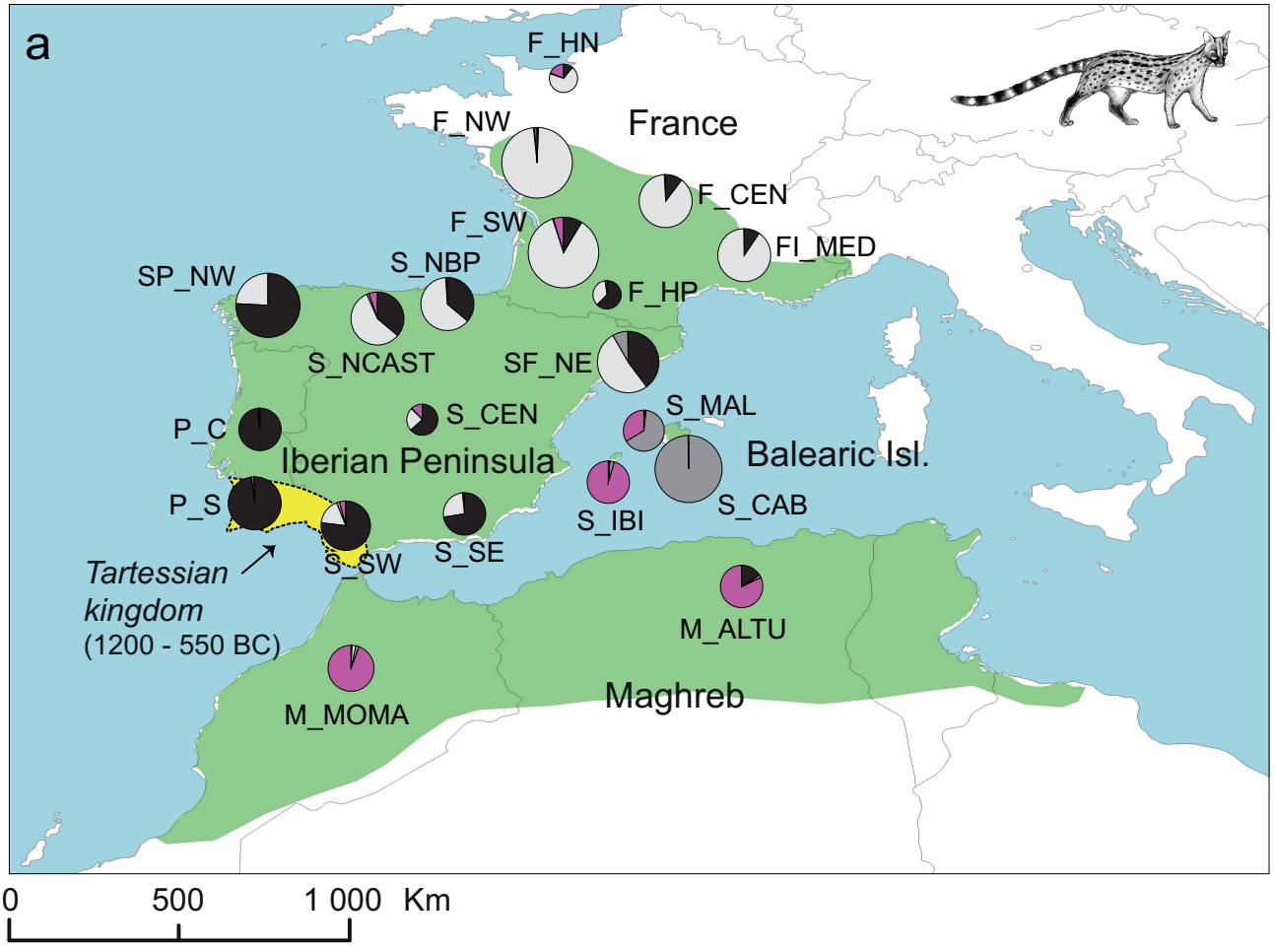
779 **Fig. 3** Principal components analysis (PCA) among a) the geographic populations and b) all
780 the individuals representing the common genet. See Table 1 for population acronyms.

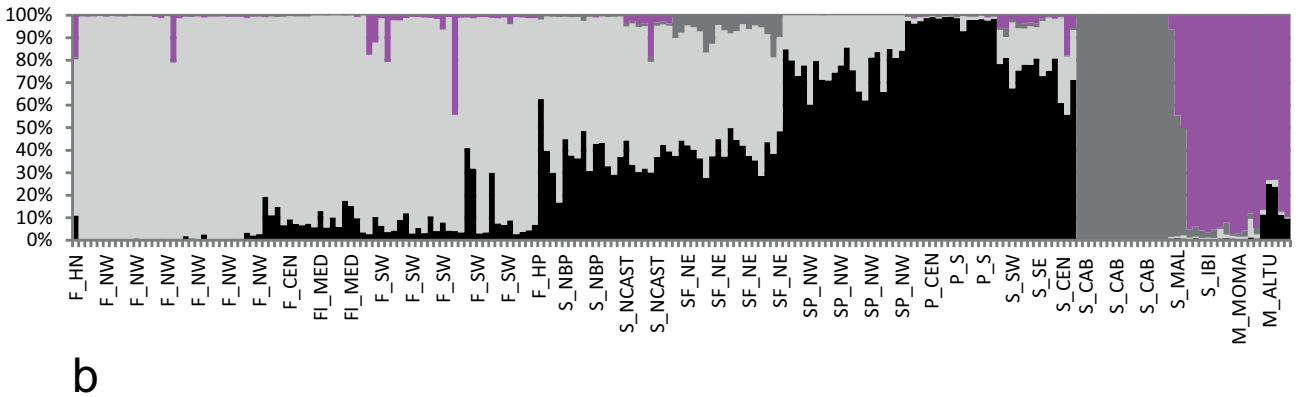
781

782 **Fig. 4** Isolation by distance among (a) individuals and (b) populations of the common genet in
783 continental Europe. Linear regressions are given for the northern (dashed line), southern
784 (dotted line) and admixed (long dashed line) groups. Groups were delineated following results
785 from STRUCTURE. The thick solid lines show the linear regression all over continental
786 Europe (i.e. all individuals and groups together).

Figure 1
Click here to download Figure: Gaubertetal_Fig1.eps

Figure 1





b

Figure 2

Click here to download Figure: Gaubertetal_Fig2.eps

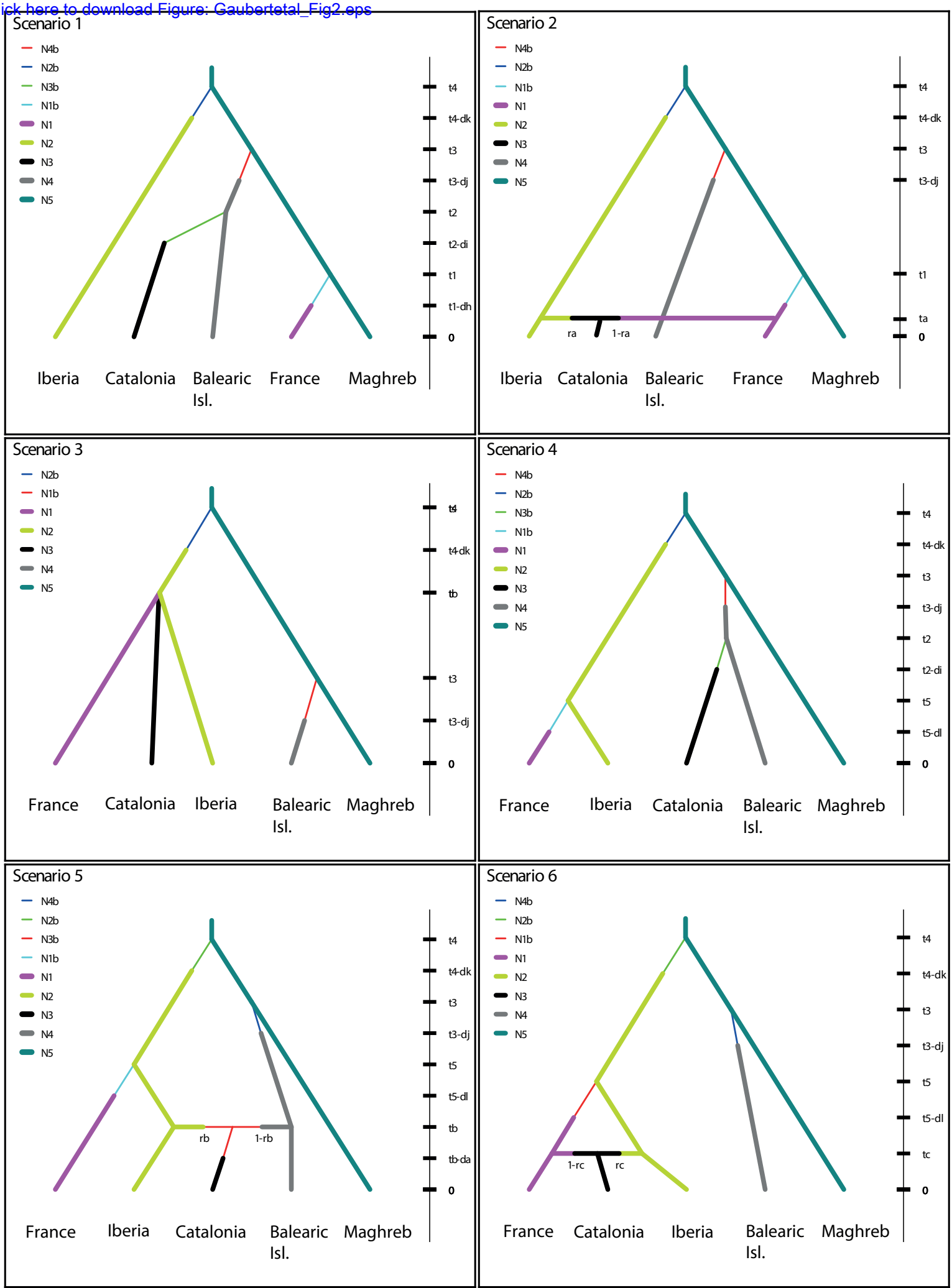


Figure 3
Click here to download Figure: Gaubertetal_Fig3.eps

Figure 3

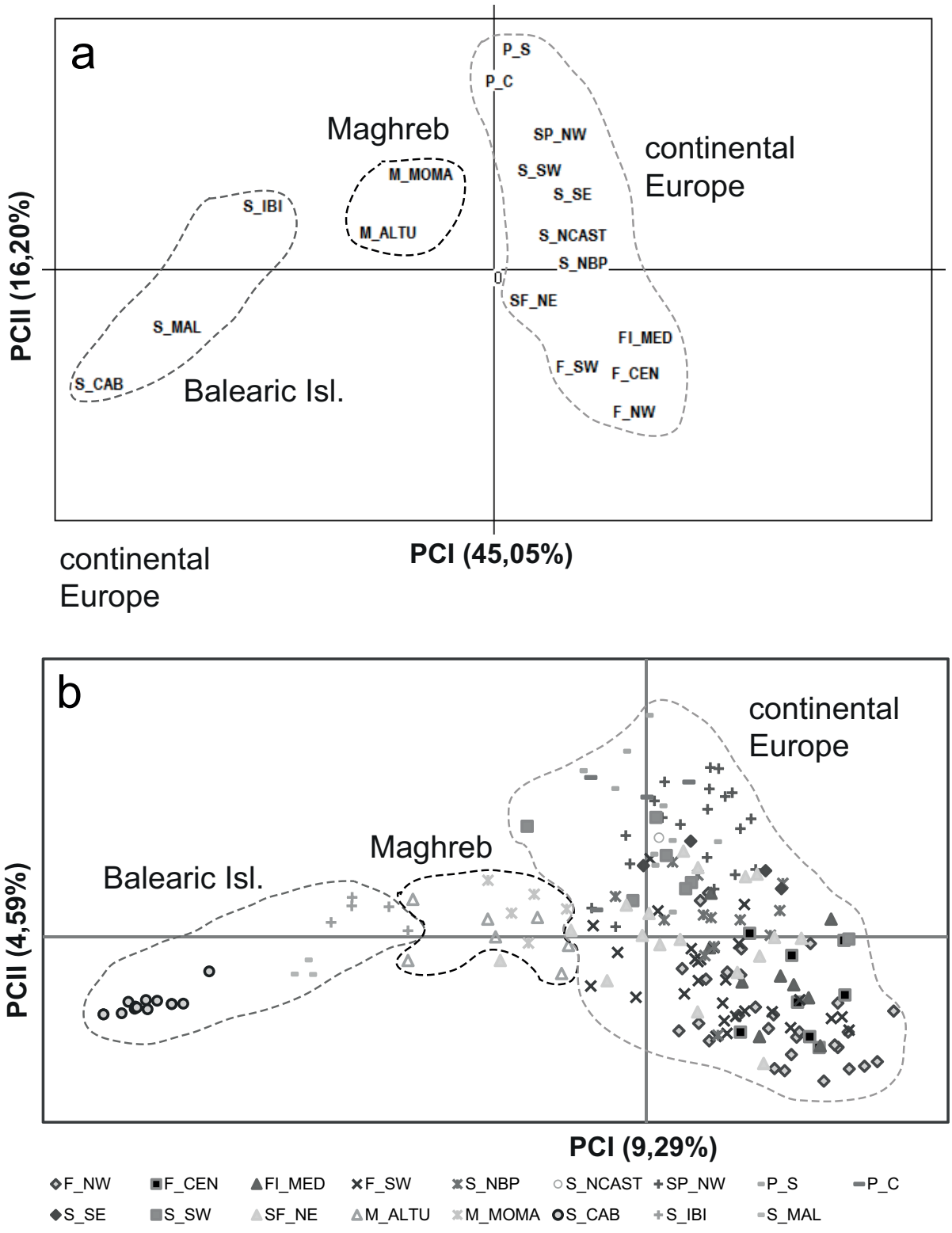


Figure 4
Click here to download Figure: Gaubertetal_Fig4.eps

Figure 4