François Gillet; Laboratoire de Biologie Evolutive, Unité de Génétique de la Conservation, Université de Liège, Institut de Botanique B22, Quartier Vallée 1, Chemin de la Vallée 4, 4000 Liège, Belgium; f.gillet@hotmail.com; f.gillet@alumni.ulg.ac.be; Tel : 003243662130

Genetic structure for the Pyrenean desman

Evidence of fine-scale genetic structure for the endangered Pyrenean desman (*Galemys pyrenaicus*) in the French Pyrenees

F. Gillet*, M. T. Cabria Garrido, F. Blanc, C. Fournier-Chambrillon, M. Némoz, E. Sourp, C. Vial-Novella, R. Zardoya, S. Aulagnier, and J. R. Michaux

Laboratoire de Génétique de la Conservation, Université de Liège, Institut de Botanique B22, Quartier Vallée 1, Chemin de la Vallée 4, 4000 Liège, Belgium (FG, MTCG, JRM)

Comportement et Ecologie de la Faune Sauvage (CEFS), Institut National de la Recherche Agronomique, 24 Chemin de Borde Rouge, Auzeville, CS 52627, 31326 Castanet-Tolosan Cedex, France (FG, SA)

Conservatoire d'Espaces Naturels de Midi-Pyrénées, 75 voie du Toec, BP 57611, 31076 Toulouse Cedex 3, France (FB, MN)

Groupe de Recherche et d'Etudes pour la Gestion de l'Environnement (GREGE), Route de Préchac, 33730 Villandraut, France (CF-C)

Parc national des Pyrénées, 2 rue du IV Septembre, 65 007 Tarbes, France (ES)

Département Biologie Vétérinaire, Laboratoires des Pyrénées et des Landes, rue des Ecoles - 64150 Lagor, France (CV-N)

Departamento de Biodiversidad y Biología Evolutiva, Museo Nacional de Ciencias Naturales, CSIC; José Gutiérrez Abascal, 2, 28006 Madrid, Spain (RZ)

CIRAD, Agirs Unit, TA C- 22/E- Campus international de Baillarguet, 34398 Montpellier Cedex 5, France (JRM)

Johan Michaux 5/2/17 05:01
Mis en forme: Justifié

The Pyrenean desman (*Galemys pyrenaicus*) is a small, semi-aquatic mammal endemic to the Pyrenean Mountains and the northern half of the Iberian Peninsula where it lives in cold and well-oxygenated flowing mountain streams. This species is currently classified as vulnerable on the IUCN Red List and has been undergoing habitat loss and fragmentation for decades, inevitably impacting its distribution. A recent genetic study, based on mitochondrial and intronic sequences, showed that the genetic variability of the Pyrenean desman is very low in the Pyrenees. In this study, we investigated the potential existence of genetic structure and gene flow at a smaller scale using 24 polymorphic microsatellite loci. As the Pyrenean desman is a very elusive species, we supplemented our tissue sample collection with samples of feces collected in the French range of this species. We successfully identified 70 individuals based on 355 fecal samples. Bayesian analyses revealed 3 genetic and geographic clusters (1 eastern, 1 central, and 1 western, including 3 genetic sub-clusters), with origins tracing back only 200 years. These clusters were characterized by low levels of genetic diversity and high inbreeding coefficients. Although gene flow among clusters appeared to be limited, populations seem to have exchanged alleles recently. Therefore, connectivity between watersheds should be enhanced to maintain genetic diversity and potentially improve the long-term survival of the Pyrenean desman in France.

Key words: conservation genetics, *Galemys pyrenaicus,* genetic structure, microsatellites

*Correspondent: f.gillet@hotmail.com

Conservation of endangered species is dependent on knowledge of the genetic structure and diversity of individual populations (Frankham 2003). This diversity is often spatially structured because natural habitat is not always continuous or can vary across the species range. Anthropogenic activities, resulting in fragmentation and habitat loss, also play a role in shaping the structure of wildlife populations. At the individual level, fragmentation can alter spatial and dispersal movement patterns or disrupt the social structure and increase inbreeding (Gerlach and Musolf 2000; Coster and Kovach 2012; Mona et al. 2014). At the population level, fragmentation can reduce suitable habitat, restricting gene flow and augmenting genetic drift (Frankham 2005). Knowledge of species' population dynamics along with their genetics and basic ecology is essential to designing and implementing conservation plans and appropriate management measures, especially in the case of habitat specialists with small ranges that may be particularly sensitive to environmental change.

The Pyrenean desman (*Galemys pyrenaicus*, Soricomorpha, Talpidae) is a small, semi-aquatic mammal endemic to the Pyrenean Mountains and to the northern half of the Iberian Peninsula (Queiroz 1999). This species lives in montane rivers with cold and well-oxygenated flowing waters and is well adapted to aquatic life. The Pyrenean desman is characterized by large webbed hindfeet, double-layered fur, a long tail, and a mobile prehensile snout, which make it a specialist in finding and feeding on larvae of benthic macroinvertebrates (Palmeirim and Hoffmann 1983; Richard 1986). The Pyrenean desman is an endangered species. It is currently classified as vulnerable on the IUCN Red List (Fernandes et al. 2008) and is legally protected in the 4 countries encompassing its range (Andorra, France, Portugal, and Spain). It has been undergoing habitat loss and fragmentation for decades, inevitably impacting its distribution (Nores et al. 2007; Némoz and Bertrand 2008).

71 The elusive behavior and nocturnal activity of this species make it hard to study. Some
72 information is known about its ecology and biology (Stone 1987; Bertrand 1994; Melero et al.
73 2012, 2014), but no data on the genetic structure of the Pyrenean desman was available until a
74 recent study based on mitochondrial and nuclear markers (Igea et al. 2013). This study
75 revealed that the Pyrenean desman was characterized by very low genetic diversity compared
76 to other mammals. Its evolutionary history seems to have been highly influenced by
77 Pleistocene glaciations, leading to a phylogeographic structure encompassing 4 mitochondrial
78 lineages with parapatric distributions. More specifically, Igea et al. (2013) obtained evidence
79 that the desman's Pyrenean populations are genetically homogeneous and that they likely
80 originated from a distant refuge, probably located in the Basque Mountains, after a severe
81 bottleneck.

82 However, these hypotheses were based on a small number of specimens as well as on
83 the use of genetic markers with relatively low rates of evolutionary change, so they did not
84 provide fine-scale information concerning the studied populations. In order to gain further
85 insight into the evolutionary history of the French Pyrenean populations, we conducted a
86 genetic analysis of fecal samples collected throughout the region using microsatellite markers
87 developed in our laboratory (Gillet et al. 2015a). Fecal sampling is the easiest way to detect
88 the presence of the Pyrenean desman and obtain DNA for genetic analyses, as previously
89 demonstrated (Gillet et al. 2015b). We sought to determine the spatial distribution of genetic
90 diversity in desman populations and to quantify gene flow across the French Pyrenees. The
91 ultimate goal of this study was to enhance our general knowledge of this endangered and
92 elusive species to better inform its conservation.

93 **Material and methods**

Sampling and DNA extraction.— A total of 38 tissue and 355 fecal samples derived from the entire French range of this species were used in this study. The tissue samples came from specimens found dead and collected by our research team. Samples were collected from 2011to 2014. The license numbers from the French departments used to collect this material are available upon request. Fecal samples from Pyrenean desmans were identified by amplification of a small mitochondrial cytochrome b fragment (Gillet et al. 2015b). Genomic DNA from tissue and fecal samples preserved in ethanol was extracted using the DNeasy Tissue Kit (Qiagen Inc., Hilden, Germany) and the Stool Mini Kit (Qiagen Inc.), respectively, according to the manufacturer's instructions. To avoid cross-contamination, DNA extractions from feces were conducted in a separate room with an UV-sterilized platform where no Pyrenean desman tissue samples had been previously treated. *DNA amplification.*— The 393 samples used in this study were genotyped at 24 variable microsatellite loci using the multiplex sets and PCR conditions reported in Gillet et al. (2015a), with slightly modified conditions for fecal samples, where PCRs were carried out in a 10-µl volume containing 0.15 of each 20-µM primer, 7.5 µl of Multiplex PCR kit (Qiagen Inc.), and 5 µl of DNA. Amplified DNA was analyzed for length variations on an ABI 3700 sequencer using GeneScan 500LIZ® size standard, and alleles were scored with GENEMAPPER 4.0 (Applied Biosystems, Foster City, California). Consensus genotypes were constructed for fecal samples to prevent genotyping errors in our dataset. For this, we used a modified multitube PCR approach (Taberlet et al. 1996) and repeated each PCR 4 times. Allele scores were accepted if they appeared at least 3 times in 4 PCRs.

Statistical analyses.— Each replicate genotype was compared with the consensus genotype to quantify the error rates. The consensus genotype construction and error rate quantification—such as false alleles (FA) and allelic dropouts (ADO)—were both performed

using GIMLET v1.3.3 (Valiere 2002). The probability of identity among siblings (PIDsibs), i.e., the probability that 2 related individuals have the same genotype (Waits et al. 2001), was estimated using GIMLET v1.3.3. We used MICRO-CHECKER 2.2.3 (Van Oosterhout et al. 2004) to estimate the proportion of null alleles (NA). Genetic diversity was quantified by estimating observed (H_O) and expected (H_E) heterozygosities with GENETIX (Belkhir et al. 2004). Hardy–Weinberg (HW) equilibrium was tested using the exact test implemented in GENEPOP 4.1.0 (Rousset 2008) for each locus separately and over all loci for each cluster (see below). Tests for linkage disequilibrium between loci for each cluster were performed using GENEPOP 4.1.0. Allelic richness (AR) was calculated using the rarefaction procedure implemented in FSTAT 2.9.3.2 (Goudet 2001). Multi-locus F_{IS} was calculated for each cluster and adjusted for multiple tests using Bonferroni's correction with FSTAT 2.9.3.2.

Population structure.— We used STRUCTURE 2.3.1 (Pritchard et al. 2000) to detect genetic structure in our dataset. We used the model-based Bayesian clustering method, with no prior identification of populations, to infer the number of genetic clusters (K) and assign individuals to these clusters according to allele frequencies at each locus. For each K-value from 1 to 10, the program was run 10 times using an admixture model with a burn-in of 10^5 and MCMC values of 10^6. The ΔK method (Evanno et al. 2005) was implemented with STRUCTURE HARVESTER (Earl and vonHoldt 2011) to find the most likely K-value present in the dataset and a visual output of the STRUCTURE results was generated using DISTRUCT (Rosenberg 2003) and CLUMPAK (Kopelman et al. 2015).

We used ARLEQUIN (Excoffier et al. 2005) to estimate pairwise genetic differentiation among populations using F_{ST} statistics, and the online SMOGD application (http://www.ngcrawford.com/django/jost/) was used to estimate Dest statistics (Jost 2008) using 1,000 bootstrap replicates.

142 Isolation by distance (IBD) analyses among and within clusters defined with
143 STRUCTURE 2.3.1 were performed with GENEPOP 4.1.2. The signal strength was estimated
144 by calculating $D\sigma^2$ (i.e., product of the population density and axial mean square parent-
145 offspring distance as defined by Rousset (1997)), according to $b = 1/(4\pi D\sigma^2)$. The value
146 obtained ($D\sigma^2$) was inversely correlated with the IBD strength. The logarithm of the
147 Euclidean distance on GPS coordinates was used to calculate the geographic distance and *âr*
148 statistics were used to represent the genetic distance between pairs of individuals (Rousset
149 2000). A Mantel's test with 10,000 permutations was used to test the significance of the
150 correlation.

151 Finally, we used BOTTLENECK 1.2 (Cornuet and Luikart 1996) to perform a
152 Wilcoxon test under a 2-phase model (TPM) to investigate recent demographic bottlenecks
153 with estimations based on 1,000 replications.

154 *Demographic history.*— The evolutionary history of *G. pyrenaicus* in France was
155 investigated using approximate Bayesian computation as implemented in DIYABC
156 1.0.4.45beta software (Cornuet et al. 2010). This coalescent-based approach allows estimation
157 of the effective population size as well as the splitting time in generations for each tested
158 genetic cluster. A number of biogeographic scenarios were tested and compared to determine
159 whether the observed clusters originated from fragmentation of an ancestral common
160 population or if any cluster resulted from an admixture of the others. Specifically, the type 1
161 scenarios explored 2 consecutive divergences with a 1st divergence between 2 of the 3
162 populations, followed by divergence of the 3rd population from 1 of the 2 others (6 alternative
163 scenarios; Fig. 1). The type 2 scenarios displayed an admixture event between 2 populations,
164 leading to formation of the 3rd population (3 alternative scenarios; Fig. 1), while the type 3
165 scenario showed a radiation process, where the 3 populations would have split at the same

time from a common ancestor (Fig. 1). Two runs were performed, in the 1st one, we considered all alternative scenarios, whereas only those having the highest posterior probabilities (PP) were considered in the 2nd run. The distribution and range of priors for the parameters used to describe these alternative scenarios (effective population size, time of splitting or merging events in generations, and admixture rates) are given in Table 1. A total of 10^7 and $2x10^6$ datasets were simulated for each alternative scenario in the 1st and 2nd runs, respectively, in order to build a reference table from a set of prior parameter distributions. To check if the combination of these distributions of prior parameters and alternative scenarios could generate datasets similar to the observed ones, a principal component analysis (PCA) was performed on the first 10^5 simulated datasets in the 1st run and on the first $2x10^5$ datasets in the 2nd run. Inspection of PCAs helped us choose the most adequate timeframe corresponding to our data (maximum of 500 generations backwards in time). We used microsatellite mutation rates generally used for mammalian species (i.e., 10^{-3} to 10^{-5}; Dallas 1992; Weber and Wong 1993; Ellegren 1995). Software default values were chosen for admixture rates. To determine the most likely alternative scenarios, we used normalized Euclidean distances between each simulated dataset and our observed dataset and 1% of the closest simulated datasets were used to estimate the relative posterior probability (with 95% confidence intervals) of each alternative scenario with a logistic regression (Cornuet et al. 2008). The most likely alternative scenario was that with the highest posterior probability with a non-overlapping 95% confidence interval.

To assess the level of confidence of these analyses, new datasets were simulated with each alternative scenario and the same procedure to estimate their respective posterior probabilities was applied, and the proportion of times the right alternative scenario had the highest posterior probability was measured. According to Cornuet et al. (2010), the type-I

error was estimated for 10,000 simulated datasets generated under the best-supported alternative scenario. The type-II error was estimated by simulating 10,000 datasets generated for each alternative scenario and counting decisions in favor of the selected alternative scenario.

We also estimated historical demographic events and genetic parameters including interactions between clusters (i.e., migration rate (*M*)) using MIGRATE 3.4.4 (Beerli and Felsenstein 1999, 2001; Beerli 2004, 2006; Beerli and Palczewski 2010). This software is able to search through genealogies and obtain estimates of theta (Θ) and *M* by employing a Metropolis-Hastings Markov chain Monte Carlo (MCMC) algorithm and a likelihood ratio test, respectively. It assumes a constant Θ for each population but a variable Θ between them (pairwise migration rate estimates).

We first used MIGRATE 3.4.4 with default parameters, with *FST*- based statistics of Θ and *M*, 10 short chains of 10,000 sampled genealogies, and 3 long chains of 100,000 sampled genealogies. The parameter estimates of Θ and *M* from the previous run were used as starting values to perform a second analysis. The formula x*Ne*m = *M**Θ was used to calculate the headcount of immigrants per generation, with x being the inheritance scalar (set at 4 for diploid species), *Ne* the effective population size, and m the mutation rate per generation and per locus.

Results

Microsatellite genotyping.— A total of 70 individuals were identified out of the 355 fecal samples analyzed with a PIDsibs of $2.24e^{-03}$. The mean proportion of positive PCRs was 67%, ranging from 53% to 76% among loci. No significant allelic dropout or false allele errors could be found in our data (all loci <0.001). MICRO-CHECKER did not detect any

214 significant bias in our dataset that could be attributed to null alleles. Therefore, the total
215 number of individuals used in our analyses was 108 (38 tissue samples and 70 individuals
216 identified from fecal samples). It is important to note that the results of the genotyping of
217 fecal samples were highly dependent on the freshness and size of the feces at the time of
218 collection as feces of Pyrenean desmans are generally small (10 to 15 mm long and 4 to 8 mm
219 wide, Bertrand 1993) and their DNA content rapidly degrades due to contact with water and
220 UV radiation (Lindahl 1993). In our study, only 20% of the collected feces could be attributed
221 to distinct individuals. The percentage of detected individuals also was dependent on the
222 threshold rule of the conservative multitube approach that we used, i.e., allele scores were
223 accepted if they appeared at least 3 times in 4 PCRs. This step was nonetheless necessary to
224 ensure reliable results

225 *Population structure and genetic diversity.*— After using the ΔK method on our
226 STRUCTURE results, the highest ΔK-value was found at K=3 (Fig. 2). The 1st cluster (C),
227 which appeared to have the largest geographic distribution (Fig. 2), mainly included samples
228 from the Garonne watershed in the central Pyrenees. The 2nd cluster (E) mainly included
229 samples from the Tet–Tech–Aude watershed in the eastern Pyrenees. Finally, the 3rd cluster
230 (W) mainly included samples from the Adour–Nive watershed in the western Pyrenees.

231 The mean H_O ranged from 0.19 in the eastern cluster to 0.23 in the central cluster,
232 while the mean H_E ranged from 0.26 in the eastern cluster to 0.37 in the western cluster (see
233 Supplementary Data S1, S2, S3). The mean allelic richness ranged from 1.4 to 1.8 (Table 2).
234 Tests for HWE showed significant deviations for the eastern and western clusters. Four pairs
235 of loci (GpyrGS22 versus GpyrGS33, GpoyrGS30 versus GpyrGS74, GpyrGS33 versus
236 GpyrGS18, and GpyrGS11 versus GpyrGS20) showed significant linkage disequilibrium for
237 the central cluster and 1 pair (GpyrGS33 versus GpyrGS82) for the eastern cluster after

Bonferroni correction. The inbreeding coefficient (F_{IS}) was significant for all 3 clusters (Table 2) and all pairwise F_{ST} were significant (Table 3). Moreover, the results of the Wilcoxon test under a 2-phase model (TPM) performed using the BOTTLENECK 1.2 software package indicated that both the central and eastern clusters had undergone a recent bottleneck ($P <$ 0.001).

The high observed F_{IS}-values, especially in the western cluster (Table 2), could also indicate a Wahlund effect in the genetic clusters. To investigate this possibility, we conducted an additional clustering analysis within each cluster, under the same conditions as previously described. The analyses for clusters C and E did not give any evidence of substructure (all individuals admixed beyond K = 1), whereas that for cluster W supported the existence of 3 sub-clusters (Fig. 3). This sub-structuring could explain the higher inbreeding coefficient (0.434) in the western cluster, even though these sub-clusters did not seem to exhibit any clear geographical distribution (Fig. 3).

Demographic history.— After using DIYABC software to investigate 10 distinct demographic alternate scenarios, a 2nd analysis was performed on the 2 most probable alternate scenarios (those with the highest posterior probabilities) obtained in the 1st run. These 2 alternate scenarios (2 and 9, Fig. 4) exhibited 2 different evolutionary demographic patterns. Alternate scenario 2 reflected the separation of cluster W from more eastern Pyrenean desmans, followed by the separation of cluster C from cluster E. In contrast, alternate scenario 9 reflected the origin of the central cluster following a merging event between the other 2 clusters. The logistic regressions performed on 1% of the closest simulated datasets revealed that the most likely of all of the tested alternate scenarios was alternate scenario 2, with a PP of 0.582 and a confidence interval (95% CI) of 0.475–0.689. However, the confidence intervals of alternate scenarios 2 and 9 overlapped (PP of 0.418 and

95% CI of 0.310-0.525). Analysis of confidence in alternate scenario 2 resulted in type I and type II errors of 0.266 and 0.238, respectively, and inversely for alternate scenario 9. Therefore, the 2 alternate scenarios had similar probability of being correct, with alternate scenario 2 having a slight better probability (76.2%) than alternate scenario 9 (73.4%). However, the divergence times (t2a, t2b, t9a, and t9b) and estimates of effective population size (N_C, N_E, and N_W) were all within the same order of magnitude when comparing the posterior distribution of parameters of the 2 alternate scenarios (Fig. 4). Assuming a minimum generation time of 1 year, the median divergence times t2a and t2b of alternate scenario 2 were estimated at 80 (95% CI: 25 – 190) and 230 years ago (95% CI: 80 – 440), respectively. For alternate scenario 9, the median admixture and divergence times t9a and t9b were estimated at 60 (95% CI: 18 – 160) and 240 years ago (95% CI: 80 – 450), respectively (Fig. 4). The effective population size estimates for clusters C, E, and W had a mean of 370 (95% CI: 90 – 490), 100 (95% CI:30 – 220), and 330 (95% CI:140 – 480) individuals, respectively, for alternate scenario 2, and 290 (95% CI:100 – 470), 100 (95% CI:30 – 220) and 320 (95% CI:130 – 480), respectively, for alternate scenario 9 (Fig. 4).

We also conducted another DIYABC analysis on the 3 western sub-clusters with the same alternate scenarios as for the 3 main clusters, but with slightly modified conditions (effective population sizes and time of events set at max. 300). After the 1st run with 10 alternate scenarios, 2 were well-supported and thus a 2nd run with these 2 alternate scenarios was launched. Finally, the software unambiguously chose 1 alternate scenario from the 2, with a PP of 0.585 (95% CI of 0.565 – 0.605). This alternative scenario displayed the same evolutionary demographic pattern as alternate scenario 2 from Fig. 4. The median times of divergence among these clusters (populations N1, N2, and N3; Fig. 3) were quite similar to

285 those found for the 3 main groups (100 and 200 years ago). The mean effective population
286 sizes were estimated at 185, 160, and 120 individuals for N1, N2, and N3, respectively.

287 We used MIGRATE 3.4.4 on the 3 main clusters to calculate the migration rate per
288 generation according to $N_em = (M_{i\ j}*\Theta_j)/4$, with $\Theta_C = 2.86$, $\Theta_E = 1.14$, $\Theta_W = 1.14$ and $M_{E\ C} =$
289 0.83, $M_{W\ C} = 0.46$, $M_{C\ E} = 1.38$, $M_{W\ E} = 0.39$, $M_{C\ W} = 0.84$, $M_{E\ W} = 0.87$. The effective number
290 of migrants per generation among clusters was low, with less than 1 migrant per generation (1
291 year) with $Nm_{E\ C}= 0.59$, $Nm_{W\ C} = 0.32$, $Nm_{C\ E} = 0.39$, $Nm_{W\ E} = 0.11$, $Nm_{C\ W} = 0.24$ and $Nm_{E\ W} = 0.25$.

293 *Isolation by distance.*— The IBD analyses indicated an absence of isolation by
294 distance among clusters and within clusters E and W. In contrast, a significant signal was
295 observed within cluster C (Mantel test $P < 0.05$), as indicated by the relatively low $D\sigma^2$ value
296 (0.44). However, the slight regression slope indicated that the genetic distance between pairs
297 of individuals was weakly correlated with the geographic distance between them ($r^2 = 0.05$).
298 The absence of significant signal among the 3 clusters suggests that their relationships (e.g., a
299 closer link between the cluster E and C as compared to the cluster W as suggested by the
300 DIYABC results) was not driven by geographic distance alone.

301

302 **Discussion**

303 The present study generated new insight into the genetic structure of Pyrenean desman
304 populations in the French Pyrenees. The large area sampled, covering the entire current
305 distributional area of the Pyrenean desman on the French side of the Pyrenees, and the use of
306 24 highly variable microsatellite markers provided a more fine-scale view of genetic structure
307 than previously available.

Population structure and evolutionary history of the French Pyrenean desman.— The results showed evidence of 3 genetically and geographically distinct clusters (Fig. 2) situated in the eastern (E), central (C), and western (W) French Pyrenees. In addition, sub-structuring seemed to emerge within cluster W, where 3 sub-clusters were detected. These sub-clusters did not exhibit any clear geographical distribution but this may be the result of low sampling within each sub-cluster (6-10 individuals). More extensive sampling in this area is needed to better define this substructure.

The existence of eastern, central, and western genetic clusters was not evident in the study of Igea et al. (2013), which suggested the existence of a single Pyrenean population that would have recently colonized the Pyrenean region from a putative refuge situated in the Basque Mountains after a severe bottleneck event. This hypothesis was deduced from the very low levels of genetic diversity found in both mitochondrial and nuclear (intron) marker sequences in this region. Our microsatellite markers, characterized by higher mutation rates than mitochondrial or nuclear intronic sequences (Schlötterer 2000), allowed us to detect finer genetic structure in the Pyrenean desman. Therefore, according to their different evolutionary rates, both mitochondrial and nuclear markers gave complementary information concerning the evolutionary history of the desman in this region. The species probably colonized the Pyrenean region after the last ice age, and subsequent diversification led to the 3 genetic clusters, presently distributed throughout the French Pyrenean region. The only other vertebrate species having a similar reported genetic structure across the French Pyrenees is the rock ptarmigan (*Lagopus muta*— Bech et al. 2009). However, in contrast to the Pyrenean desman, the genetic structure of this bird species was associated with a significant isolation-by-distance effect, likely the result of short dispersal distances, and high natal and breeding philopatry combined with severe habitat fragmentation.

As no isolation by distance was detected, the structure supported by our study is likely the result of concomitant environmental and anthropogenic factors. The DIYABC analysis proposed that about 80 years ago clusters C and E diverged from an ancestral population, which itself had diverged from the W cluster approximately 230 years ago, or that cluster C was a result of an admixture event between clusters E and W that occurred about 60-80 years ago, after clusters E and W diverged about 230-240 years ago. However, these estimations might be underestimated given that the software algorithm does not assume migration within scenario events. Moreover, these estimations were performed while considering a minimum generation time of 1 year even though this information is unknown for the Pyrenean desman and could be higher. Indeed, as *G. pyrenaicus* belongs to the family Talpidae, if we extrapolate the generation time of the Pyrenean desman from that of the European mole (*Talpa europea*— 1.72 years, Niethammer 1990), the estimated divergence times become 400 years ago for the separation of the E and W clusters and 100–140 years ago for the divergence of C and E (or the admixture of E and W in the 2nd alternative scenario).

Human population growth over the last century, and therefore the increased human impact on nature, inevitably led to riverine habitat loss and fragmentation of species' populations inhabiting mountain streams. The construction and functioning of hydroelectric power plants is an example of the human impact on rivers. These can lead to physical and biotic modifications and alter both hydrologic and thermal regimes, thus impacting resources such as benthic macroinvertebrate larvae (Queiroz et al. 1992; Céréghino and Lavandier 1997). Although little information is available on the Pyrenean desman, these detrimental effects have been studied and highlighted by various authors for other mammals and birds (Nilsson and Dynesius 1994; D'Amico et al. 2000). Furthermore, the development of such infrastructures dates back to the beginning of the 20th century in the French Pyrenees, and

most were built between 1930 and 1960. These dates could therefore coincide with the emergence of the 3rd population found in this study (60-80 years ago). However, more focused studies are needed to further understand the influence of this infrastructure on the distribution of Pyrenean desman and gene flow among populations.

Centennial-scale climatic change could also have played a role in the structuring of Pyrenean desman populations, notably during the last 100 years of the Little Ice Age (1750-1850). Cooler temperatures at high elevations could have induced a shift to lower elevations in the range of the Pyrenean desman. After the Little Ice Age, the Pyrenean desman populations could have shifted back to higher elevations, which in turn could have restricted dispersal. Indeed, rivers at high elevations are less connected and this species favors rivers with high water flows, which are now found at high elevations (Nores et al. 1992; Ramalhinho and Boa Vida 1993; Queiroz et al. 1996; Charbonnel et al. 2015).

Genetic diversity of the Pyrenean desman populations.— The 3 main clusters seemed to be characterized by a heterozygote deficiency, as indicated by the relatively low heterozygosity values (around 0.2) and significantly high F_{IS} indices (Table 2). These data would be associated with recent bottleneck events, at least for clusters C and E, as confirmed by the BOTTLENECK 1.2 analysis.

The very high inbreeding coefficient in the western cluster could also be explained by a Wahlund effect as 3 sub-clusters were found in this population. As for the 3 main populations, the impact of both anthropogenic and climatic factors (notably during the Little Ice Age) along with watershed structure could jointly explain differentiation of the 3 sub-clusters. This sub-structure in the western population also could be due to the fact that this region is characterized by a smaller proportion of favorable habitats compared to the more eastern portion of the distribution (Charbonnel 2015). However, the effective population sizes

estimated by ABC were inconsistent with this favorable habitat gradient, with a higher value in the western than in the eastern population. The sub-structure of the western population could have biased the effective population size estimation.

Gene flow seemed limited among the 3 main clusters, as indicated by pairwise F_{ST} values that were significantly higher than zero and ranged between 0.345 (between the W and E cluster) and 0.203 (between the C and W clusters). This trend also was revealed using the Dest index (Table 3).

Despite this apparent low gene flow between clusters, they were not geographically separated and they overlapped in some areas (Fig. 2). Moreover, individual cluster assignments from the Bayesian analysis clearly showed that some individuals shared alleles from different clusters and have admixed genomes. This pattern was observed particularly between the E and C clusters (Fig. 2), where migrations seemed to have occurred every 2 or 3 generations. Although this result has been viewed as a recent expansion process in other species such as the European otter (*Lutra lutra*— Janssens et al. 2008; Pigneur et al. 2014), the general regression of the range of the Pyrenean desman over the last 3 decades did not allow us to retain this hypothesis. However, this overlapping of clusters could confirm that the genetics of the Pyrenean desman were not markedly impacted by the river networks (Igea et al. 2013) and that its dispersal could be complex, as pointed out by Stone (1987a, b) and Melero et al. (2012, 2014). Our estimates of migration rate also tended to confirm recent but limited gene flow among clusters.

Future studies will be needed to place the patterns found here in a broader spatial framework that includes the entire Pyrenean distribution of this species. For example, it is possible that some individuals included in our study were migrants or descendants of migrants from the Spanish side of the Pyrenees, particularly the 2 green-colored individuals in the

404 central cluster shown in Fig. 2, which could have passed through the Val d'Aran. This may
405 suggest that Pyrenean desmans could cross the mountains from one side to the other. In
406 addition, contact zones between both sides of the Pyrenees have been identified for
407 *Corthippus* grasshoppers (Buño et al. 1994), *Phylloscopus* birds (Helbig et al. 2001), and
408 viviparous lizards (*Zootoca vivipara*— Milá et al. 2013). These contact zones, which are
409 situated across the central high Pyrenees and across the southwestern Pyrenees, could also
410 exist for the Pyrenean desman. Igea et al. (2013) suggested genetic homogeneity across the
411 distribution of Pyrenean desman using relatively slowly evolving mitochondrial and nuclear
412 markers, but further analysis of Spanish and French samples using hypervariable markers is
413 needed to gain further insight concerning the broader-scale genetic structure of the Pyrenean
414 desman.

415 *Implications for conservation of the Pyrenean desman in the French Pyrenees.*—
416 Classified as "Vulnerable" on the IUCN Red List (Fernandes et al. 2008), the Pyrenean
417 desman is legally protected in France and is the focus of a LIFE + project
418 (LIFE13NAT/FR/000092) under a National Action Plan (Némoz and Bertrand 2008). This
419 species has been undergoing habitat loss and fragmentation for decades, particularly in France
420 where its range still requires further investigation (Némoz and Bertrand 2008).

421 The low level of genetic diversity observed in the different French Pyrenean desman
422 clusters as well as the heterozygote deficiency highlighted by the high inbreeding coefficient
423 (F_{IS}) values, and relatively low effective population sizes within clusters, could increase the
424 risk of extinction for this species in the future (Frankham 2005).

425 A lack of genetic diversity within the 3 main populations would lead to increased risk
426 of inbreeding depression. Greater connectivity throughout the Pyrenees should therefore be
427 fostered to facilitate individual dispersal and gene flow among the 3 main populations. This

would favor genetic mixing and a better response to future climatic change. Exchanges between neighboring watersheds should be promoted by improving water quality, mainly at lower elevations where rivers merge. Indeed, the Pyrenean desman favors rivers with high water flows, which are now found at high elevations (Nores et al. 1992; Ramalhinho and Boa Vida 1993; Queiroz et al. 1996; Charbonnel et al. 2015). This preference for high water flows at high elevations could contribute to the genetic structure observed in our study as this species is more inclined to live at elevations where rivers are less connected. In addition, enhanced management of hydroelectric infrastructures and of winter tourism at high elevations should be promoted throughout the mountain chain.

Another improvement could be achieved by increasing water flow in rivers with low trophic resources or by restoring suitable habitats for the Pyrenean desman, notably by placing stones and boulders in rivers to re-create adequate water flow and shelter. Connectivity between main and tributary rivers should also be favored as tributary rivers can serve as refugia in case of short and sudden hydrological events (Lake 2000; Charbonnel 2015).

Although this study generated new insight into the fine-scale genetic structure of the Pyrenean desman in France, a larger study, based on sensitive genetic markers such as microsatellite or SNP markers and encompassing the entire range of the species, particularly the Spanish side of the Pyrenees, is necessary to broaden overall knowledge on this threatened species and its worldwide conservation.

Acknowledgments

We thank the following people who collected tissue samples: EDF agents, Pyrenees National Park agents, M. Bayon, A. Bertrand, J-P. Besson, J-P. Quéré, A. Charbonnel, F. Elzear, L.

Fabre, P. Fantin, B. Le Roux, V. Lacaze, M. Lagardère, F. Lassèrre, B. Le Corre, M. Mas, P. Maunas, G. Nogué, F. Prud'Homme, T. Quintilla, B. Salmeron, T. Tico, S. Torreilles and S. Vernet. We also thank representatives of the following organisations who collected feces samples: Association des Naturalistes de l'Ariège, Conservatoire d'Espaces Naturels d'Aquitaine, Conservatoire d'Espaces Naturels de Midi-Pyrénées, Fédération Aude Claire, Fédération des Réserves Naturelles Catalanes, Groupe de Recherche et d'Etude pour la Gestion de l'Environnement, Office National de la Chasse et de la Faune Sauvage, Office National des Forêts, and Parc National des Pyrénées.

This study is part of the "Plan National d'Actions en faveur du Desman des Pyrénées" and the LIFE+ Desman project (LIFE13NAT/FR/000092) which are coordinated by the Conservatoire d'Espaces Naturels de Midi-Pyrénées (CEN-MP) and financially supported by the following structures: European Union Funding Network (ERDF and LIFE+), Agence de l'eau Adour-Garonne, Agence de l'eau Rhône-Méditerranée-Corse, DREAL Aquitaine, Midi-Pyrénées, and Languedoc-Roussillon, Conseil Régional Aquitaine, Midi-Pyrénées and Languedoc-Roussillon, Conseil Général des Pyrénées-Atlantiques, de l'Aude et des Pyrénées-Orientales, EDF, SHEM, Patagonia, Parc National des Pyrénées, and ANRT (Association Nationale de la Recherche et de la Technologie). F. Gillet is supported by a French research fellowship provided by ANRT (CIFRE N° 2011/1571).

Supplementary data

Supplementary data S1. Observed (H_O) and expected heterozygosity (H_E) per locus for the central cluster. Only polymorphic loci are shown.

Supplementary data S2. Observed (H_O) and expected heterozygosity (H_E) per locus for the eastern cluster. Only polymorphic loci are shown.

Supplementary data S3. Observed (H_O) and expected heterozygosity (H_E) per locus for the western cluster. Only polymorphic loci are shown.

482 **Literature cited**

483 Bech, N., Boissier, J., Drovetski, S., Novoa, C., 2009. Population genetic structure of rock
484 ptarmigan in the "sky islands" of French Pyrenees: Implications for conservation.
485 Animal Conservation 12, 138–146. doi:10.1111/j.1469-1795.2008.00233.x

486 Beerli, P., 2006. Comparison of Bayesian and maximum-likelihood inference of population
487 genetic parameters. Bioinformatics 22, 341–5. doi:10.1093/bioinformatics/bti803

488 Beerli, P., 2004. Effect of unsampled populations on the estimation of population sizes and
489 migration rates between sampled populations. Molecular Ecology 13, 827–836.
490 doi:10.1111/j.1365-294X.2004.02101.x

491 Beerli, P., Felsenstein, J., 2001. Maximum likelihood estimation of a migration matrix and
492 effective population sizes in n subpopulations by using a coalescent approach.
493 Proceedings of the National Academy of Sciences of the United States of America, 98,
494 4563–4568.

495 Beerli, P., Felsenstein, J., 1999. Maximum-Likelihood Estimation of Migration Rates and
496 Effective Population Numbers in Two Populations Using a Coalescent Approach.
497 Genetics 152, 763–73.

498 Beerli, P., Palczewski, M., 2010. Unified framework to evaluate panmixia and migration
499 direction among multiple sampling locations. Genetics 185, 313–26.
500 doi:10.1534/genetics.109.112532

501 Belkhir, K., Borsa, P., Chikhi, L., Raufaste, N., Bonhomme, F., 2004. GENETIX 4.05,
502 logiciel sous Windows TM pour la génétique des populations.

503 Bertrand, A., 1994. Répartition géographique et écologie alimentaire de desman des pyrénées
504 G. pyrenaicus dans les Pyrénées françaises. Diplôme Universitaire de Recherche,
505 Toulouse, 217p.

506 Bertrand, A., 1993. Découvrir le Desman des Pyrénées, Ed. Association des Naturalistes
507 d'Ariège, 32 p.

508 Buño I., Torroja E., López-Fernández C., Butlin R.K., Hewitt G.M., Gosálvez J., 1994. A
509 hybrid zone between two subspecies of the grasshopper Chorthippus parallelus along the
510 Pyrenees: the west end. Heredity 73:625–634.

511 Céréghino, R., Lavandier, P., 1997. Influence des éclusées hydroélectriques sur la distribution
512 et le développement larvaire des Diptères Simuliidae d'une rivière de moyenne
513 montagne. Comptes Rendus l'Académie des Sciences. - Ser. III - Sci. la Vie.
514 doi:10.1016/S0764-4469(97)82775-8

515 Charbonnel, A., 2015. Multi-scale influence of environmental factors in the distribution of the
516 Pyrenan desman (Galemys pyrenaicus) in France. PhD thesis, University of Toulouse,
517 245p.

518 Charbonnel, A., Buisson, L., Biffi, M., D'Amico, F., Besnard, A., Aulagnier, S., Blanc, F.,
519 Gillet, F., Lacaze, V., Michaux, J.R., Némoz, M., Pagé, C., Sanchez-Perez, J.M.,
520 Sauvage, S., Laffaille, P., 2015. Integrating hydrological features and genetically
521 validated occurrence data in occupancy modelling of an endemic and endangered semi-
522 aquatic mammal, Galemys pyrenaicus, in a Pyrenean catchment. Biological
523 Conservation 184, 182–192. doi:10.1016/j.biocon.2015.01.019

524 Cornuet, J., Ravigné, V., Estoup, A., 2010. Inference on population history and model
525 checking using DNA sequence and microsatellite data with the software DIYABC (v1 .
526 0). BMC Bioinformatics 11, 401.

527 Cornuet, J.-M., Luikart, G., 1996. Power Analysis of Two Tests for Detecting Recent
528 Population Bottlenecks From Allele Frequency Data. Genetics 144, 2001–14.

529 Cornuet, J.-M., Santos, F., Beaumont, M. a, Robert, C.P., Marin, J.-M., Balding, D.J.,
530 Guillemaud, T., Estoup, A., 2008. Inferring population history with DIY ABC: a user-
531 friendly approach to approximate Bayesian computation. Bioinformatics 24, 2713–9.
532 doi:10.1093/bioinformatics/btn514

533 Coster, S.S., Kovach, A.I., 2012. Anthropogenic influences on the spatial genetic structure of
534 black bears. Conservation Genetics 13, 1247–1257. doi:10.1007/s10592-012-0368-4

535 D'Amico, F., Manel, S., Mouches, C., Ormerod, S.J., 2000. River birds in regulated rivers :
536 cost or benefit ? Verhandlungen des Internationalen Verein Limnologie 27, 167–170.

537 Dallas, J.F., 1992. Estimation of microsatellite mutation rates in recombinant inbred strains of
538 mouse. Mammalian Genome, 3, 452–456.

539 Earl, D. a., vonHoldt, B.M., 2011. STRUCTURE HARVESTER: a website and program for
540 visualizing STRUCTURE output and implementing the Evanno method. Conservation
541 Genetics Resources 4, 359–361. doi:10.1007/s12686-011-9548-7

542 Ellegren, H., 1995. Mutation rates at porcine microsatellite loci. Mammalian Genome, 6,
543 376–377.

544 Evanno, G., Regnaut, S., Goudet, J., 2005. Detecting the number of clusters of individuals
545 using the software STRUCTURE: a simulation study. Molecular Ecology 14, 2611–20.
546 doi:10.1111/j.1365-294X.2005.02553.x

547 Excoffier, L., Laval, G., Schneider, S., 2005. Arlequin (version 3 . 0): An integrated
548 software package for population genetics data analysis 47–50.

549 Fernandes, M., Herrero, J., Aulagnier, S., Amori, G., 2008. Galemys pyrenaicus. IUCN Red
550 List of Threatened Species. Version 2013.2.

551 Fournier, M., Mesquita, J., Mangin, A., 2010. Evaluation scientifique de l'impact de
552 l'hydroélectricité dans le Parc naturel régional des Pyrénées ariégeoises. 163p.

553 Frankham, R., 2005. Genetics and extinction. Biological Conservation 126, 131–140.
554 doi:10.1016/j.biocon.2005.05.002

555 Frankham, R., 2003. Genetics and conservation biology. Comptes Rendus Biologies 326,
556 S22–S29. doi:10.1016/s1631-0691(03)00023-4

557 Gerlach, G., Musolf, K., 2000. Fragmentation of Landscape as a Cause for Genetic
558 Subdivision in Bank Voles 14, 1066–1074.

559 Gillet, F., Cabria, M.T., Blanc, F., Fournier-Chambrillon, C., Némoz, M., Sourp, E., Vial-
560 Novella, C., Aulagnier, S., Michaux, J.R., 2015a. In press. Isolation, characterization and
561 PCR multiplexing of polymorphic microsatellite markers in the endangered Pyrenean
562 desman, Galemys pyrenaicus.Conservation Genetics Resources

563 Gillet, F., Cabria, M.T., Némoz, M., Blanc, F., Fournier-Chambrillon, C., Sourp, E., Vial-
564 Novella, C., Aulagnier, S., Michaux, J.R., 2015b. PCR-RFLP identification of the
565 endangered Pyrenean desman, Galemys pyrenaicus (Soricomorpha, Talpidae), based on
566 faecal DNA. Mammalia. doi:10.1515/mammalia-2014-0093

567 Goudet, J., 2001. FSTAT, a program to estimate and test gene diversities and fixation indices
568 (version 2.9.3). Available from http://www.unil.ch/izea/softwares/fstat.html. Updated
569 from Goudet (1995).

570 Helbig, A.J., Salomon, M., Bensch, S., Seibold, I., 2001. Male-biased gene flow across an
571 avian hybrid zone: evidence from mitochondrial and microsatellite DNA. Journal of
572 Evolutionary Biology 14:277–287.

573 Igea, J., Aymerich, P., Fernández-González, A., González-Esteban, J., Gómez, A., Alonso, R.,
574 Gosálbez, J., Castresana, J., 2013. Phylogeography and postglacial expansion of the
575 endangered semi-aquatic mammal Galemys pyrenaicus. BMC Evolutionary Biology 13,
576 115. doi:10.1186/1471-2148-13-115

577 Janssens, X., Fontaine, M.C., Michaux, J.R., Libois, R., de Kermabon, J., Defourny, P., Baret,
578 P. V, 2008. Genetic pattern of the recent recovery of European otters in southern France.
579 Ecography (Cop.). 31, 176–186. doi:10.1111/j.0906-7590.2008.4936.x

580 Jost, L., 2008. G ST and its relatives do not measure differentiation. Molecular Ecology 17,
581 4015–4026. doi:10.1111/j.1365-294X.2008.03887.x

582 Kopelman, N.M., Mayzel, J., Jakobsson, M., Rosenberg, N. a, Mayrose, I., 2015. Clumpak: a
583 program for identifying clustering modes and packaging population structure inferences
584 across K. Molecular Ecology Resources. doi:10.1111/1755-0998.12387

585 Lake, P.S., 2000. Disturbance, patchiness, and diversity in streams. Journal of the North
586 American Benthological Society 19, 573–592.

587 Lindahl, T., 1993. Instability and decay of the primary structure of DNA. Nature 362, 709–15.
588 doi:10.1038/362709a0

589 Melero, Y., Aymerich, P., Luque-Larena, J.J., Gosàlbez, J., 2012. New insights into social
590 and space use behaviour of the endangered Pyrenean desman (Galemys pyrenaicus).
591 European Journal of Wildlife Research 58, 185–193. doi:10.1007/s10344-011-0561-7

592 Melero, Y., Aymerich, P., Santulli, G., Gosàlbez, J., 2014. Activity and space patterns of
593 Pyrenean desman (Galemys pyrenaicus) suggest non-aggressive and non-territorial
594 behaviour. European Journal of Wildlife Research 60, 707–715. doi:10.1007/s10344-
595 014-0838-8

596 Milá, B., Surget-Groba, Y., Heulin, B., Gosá, A., Fitze, P.S., 2013. Multilocus
597 phylogeography of the common lizard Zootoca vivipara at the Ibero-Pyrenean suture
598 zone reveals lowland barriers and high-elevation introgression. BMC Evolutionary
599 Biology 13, 192. doi:10.1186/1471-2148-13-192

600 Mona, S., Ray, N., Arenas, M., Excoffier, L., 2014. Genetic consequences of habitat
601 fragmentation during a range expansion. Heredity (Edinb). 112, 291–9.
602 doi:10.1038/hdy.2013.105

603 Némoz, M., Bertrand, A., 2008. Plan National d'Actions en faveur du Desman des Pyrénéés
604 (Galemys pyrenaicus), 2009-2014. Société Française pour l'Etude et la Protection des
605 Mammifères / Ministère de l'Ecologie, de l'Energie, du Développement Durable et de
606 l'Aménagement du Territoire, 1.

607 Niethammer, J., 1990. Talpa europaea Linnaeus, 1758 - Maulwurf. in: Handbuch der
608 Säugetiere Europas. Band 3/I: Insektenfresser - Insectivora, Herrentier - Primates.
609 AULA-Verlag, Wiesbaden, p. 99-133.

610 Nilsson, C., Dynesius, M., 1994. Ecological effects of river regulation on mammals and birds:
611 A review. Regul. Rivers Res. Manag. 9, 45–53. doi:10.1002/rrr.3450090105

612 Nores, C., Ojeda, F., Ruano, A., Villate, I., Gonzalez, J., 1992. Aproximación a la
613 metodología y estudio del área de distribución, estatus de población y selección de
614 hábitat del desmán (Galemys pyrenaicus) en la Península Ibérica. Ministerio de Medio
615 Ambiente, Oviedo.

616 Nores, C., Queiroz, A.I., Gisbert, J., 2007. Galemys pyrenaicus (E. Geoffroy Saint-Hilaire,
617 1811)., in: Palomo, L., Gisbert, J., Blanco, J. (Eds.), Atlas Y Libro Rojo de Los
618 Mamíferos Terrestres de España. Madrid: Dirección General para la Biodiversidad -
619 SECEM - SECEMU, pp. 92–98.

620 Palmeirim, J.M., Hoffmann, S.H., 1983. Galemys pyrenaicus. Mamm. Species 207, 1–5.

621 Pritchard, J.K., Stephens, M., Donnelly, P., 2000. Inference of Population Structure Using
622 Multilocus Genotype Data.

623 Queiroz, A.I., 1999. Galemys pyrenaicus (E. Geoffroy, 1811). in: A.J. Mitchell-Jones et al.
624 (eds) : The atlas of European Mammals. Poyser Natural History, Academic Press,
625 London, 78-79.

626 Queiroz, A.I., Alves, M.H., Almada, V., 1992. The small hydroplants : predicted impacts on
627 the Pyrenean desman populations (Galemys pyrenaicus. Geoffroy)., in: Proceeding
628 Meeting on the Pyrenean Desman. Lisboa, pp. 69–77.

629 Queiroz, A.I., Bertrand, A., Khakhin, G., 1996. Status and conservation of Desmaninae in
630 Europe. Council of Europe, Strasbourg.

631 Ramalhinho, M.G., Boa Vida, M.J., 1993. Habitat of the Pyrenean Desman: assessment of
632 running water quality. Monitoring pollution. In: Proceedings of the Meeting on the
633 Pyrenean Desman, Lisbon, pp. 63–67.

634 Richard, P.B., 1986. Les Desman des Pyrénées , un mammifère inconnu à découvrir. Science
635 et Découvertes, Ed. Le Rocher, Monaco : 118 p.

636 Rosenberg, N. a., 2003. Distruct: a Program for the Graphical Display of Population
637 Structure. Molecular Ecology Notes 4, 137–138. doi:10.1046/j.1471-8286.2003.00566.x

638 Rousset, F., 2008. genepop'007: a complete re-implementation of the genepop software for
639 Windows and Linux. Molecular Ecology Resources 8, 103–6. doi:10.1111/j.1471-
640 8286.2007.01931.x

641 Rousset, F., 2000. Genetic differentiation between individuals 13, 58–62.

642 Rousset, F., 1997. Genetic differentiation and estimation of gene flow from F-statistics under
643 isolation by distance. Genetics 145, 1219–1228.

644 Schlötterer, C., 2000. Evolutionary dynamics of microsatellite DNA. Chromosoma 109, 365–
645 371. doi:10.1007/s004120000089

646 Stone, R.D., 1987a. The social ecology of the Pyrenean desman as revealed by radiotracking.
647 Journal of Zoology 212, 117–129.

648 Stone, R.D., 1987b. The activity pattern of the Pyrenean desman (Galemys pyrenaicus)
649 (Insectivora : Talpidae), as determined under natural condition. Journal of Zoology.
650 London 213, 95–106.

651 Taberlet, P., Griffin, S., Goossens, B., Questiau, S., Manceau, V., Escaravage, N., Waits, L.P.,
652 Bouvet, J., 1996. Reliable genotyping of samples with very low DNA quantities using
653 PCR. Nucleic Acids Research 24, 3189–94.

654 Valiere, N., 2002. GIMLET: a computer program for analysing genetic individual
655 identification data. Molecular Ecology Notes 2, 377–379. doi:10.1046/j.1471-
656 8286.2002.00228.x

657 Van Oosterhout, C., Hutchinson, W.F., Wills, D.P.M., Shipley, P., 2004. Micro-Checker:
658 Software for Identifying and Correcting Genotyping Errors in Microsatellite Data.
659 Molecular Ecology Notes 4, 535–538. doi:10.1111/j.1471-8286.2004.00684.x

660 Waits, L.P., Luikart, G., Taberlet, P., 2001. Estimating the probability of identity among
661 genotypes in natural populations: cautions and guidelines. Molecular Ecology 10, 249–
662 56.

663 Weber, J.L., Wong, C., 1993. Mutation of human short tandem repeats. Human Molecular
664 Genetics, 2, 1123–1128.

665

666

667

668

669

670

671

672

673

674

675

676

677

678

679

680 **Figure legends**

681 **Fig. 1.** Schematic representation of 3 scenarios designed to test the origin of the 3 populations
682 of Pyrenean desman (*Galemys pyrenaicus*) in the French Pyrenees found in our study by
683 approximate Bayesian computation (ABC) analysis. Prior parameters are defined in Table 1.
684 Colors correspond to the colors in Fig.2.

685 **Fig. 2.** Population structure of the Pyrenean desman (*Galemys pyrenaicus*) in the French
686 Pyrenees estimated using STRUCTURE (K = 3). Each individual is represented by a vertical
687 line partitioned into K color segments, with the length of each color being proportional to the
688 estimated membership coefficient (inset, lower left). Geographic distribution of the 3 genetic
689 clusters is shown on the map. The 3 main watersheds of the Pyrenees are, from left to right:
690 Adour–Nive, Garonne, and Tet–Tech–Aude.

691 **Fig. 3**. Population structure estimated in the western cluster for the Pyrenean desman
692 (*Galemys pyrenaicus*) in the French Pyrenees using STRUCTURE (K = 3). Each individual is
693 represented by a vertical line partitioned into K color segments, with the length of each color
694 being proportional to the estimated membership coefficient (insert, lower left). Geographic
695 distribution of the 3 Pyrenean desman clusters is shown on the map. Each diagram represents
696 1 individual with its respective cluster assignations from STRUCTURE.

697 **Fig. 4.** Schematic representations of the 2 most likely alternative scenarios regarding
698 structuring of the Pyrenean desman (*Galemys pyrenaicus*) population in the French Pyrenees
699 based on approximate Bayesian computation (ABC) analysis. NC, NE, and NW are the
700 effective population sizes for the central, eastern, and western clusters, respectively. Numbers
701 for NC, NE, and NW correspond to number of individuals included in the analysis.

Table 1. Prior distribution of parameters used in our approximate Bayesian computation (ABC) analysis of the evolutionary history of the Pyrenean desman (*Galemys pyrenaicus*) in the French Pyrenees based on 24 variable microsatellite loci from tissue and fecal samples from 2011– 2014.

Parameter	Distribution	Min	Max
Effective population size			
N1, N2, N3, Na	Uniform	10	500
Time of events (in generations backward in time)			
Time conditions: ta2>ta1, tb2>tb1, tc2>tc1, td2>td1, te2>te1, tf2>tf1, tg2>tg1, th2>th1, ti2>ti1	Uniform	10	500
Admixture rate (ra)	Uniform	0.001	0.999
Microsatellite mutation model parameters			
Mean mutation rate	Uniform	10^{-5}	10^{-3}
Mean coefficient p	Uniform	0.1	0.3
Mean SNI rate	Log-uniform	10^{-8}	10^{-4}

Johan Michaux 5/2/17 05:01
Mis en forme: Justifié

Table 2. Overview of genetic parameters for each main cluster for the Pyrenean desman (*Galemys pyrenaicus*) in the French Pyrenees based on 24 variable microsatellite loci from tissue and fecal samples from 2011– 2014. N: Number of samples, H_O: Mean observed heterozygosity, H_E: Mean expected heterozygosity, HWE: Deviation from Hardy-Weinberg equilibrium (significance level = 0.002), AR: Mean allelic richness, F_{IS}: Mean inbreeding coefficient.

Clusters	N	H_O	H_E	HWE	AR	F_{IS}
Central	45	0.226	0.274	0.04	1.405	0.179
Eastern	38	0.189	0.258	<0.002	1.803	0.271
Western	25	0.216	0.368	<0.002	1.703	0.434

Johan Michaux 5/2/17 05:01
Mis en forme: Justifié

Johan Michaux 5/2/17 05:01
Mis en forme: Justifié

Johan Michaux 5/2/17 05:01
Mis en forme: Justifié

Johan Michaux 5/2/17 05:01
Mis en forme: Justifié

714 **Table3.** F_{ST} (above diagonal- ARLEQUIN software) and D_{EST} (below diagonal- SMOGD
715 software) for the 3 main clusters ($P < 0.05$) for the Pyrenean desman (*Galemys pyrenaicus*) in
716 the French Pyrenees based on 24 variable microsatellite loci from tissue and fecal samples
717 from 2011– 2014.

Clusters	Central	Eastern	Western
Central	-	0.288	0.203
Eastern	0.037	-	0.345
Western	0.013	0.084	-

Johan Michaux 5/2/17 05:01
Mis en forme: Justifié

718