MOLECULAR ECOLOGY OF EUROPEAN MUSTELIDS:

Unraveling evolutionary and ecological patterns in *Martes* and *Lutra* genera

María Vergara Rosa

PhD Thesis

2015

eman ta zabal zazu

Universidad del País Vasco

Euskal Herriko Unibertsitatea

eman ta zabal zazu
Universidad del País Vasco
Euskal Herriko Unibertsitatea

Department of Zoology and Animal Cell Biology
Faculty of Pharmacy

MOLECULAR ECOLOGY OF EUROPEAN MUSTELIDS:

Unraveling evolutionary and ecological patterns in *Martes* and *Lutra* genera.

María Vergara Rosa

2015

International PhD thesis. Supervisors:

Dr. Aritz Ruiz González and Dr. María José Madeira

We live on an island surrounded by a sea of ignorance.
As our island of knowledge grows, so does its shore

John Archibald Wheeler

Agradecimientos

Cuando escribes una tesis, siempre acabas comparando tu trabajo con el de otros. Yo, y supongo que no seré la única, lo primero que hago es leerme los agradecimientos, porque una tesis es algo propio, pero es evidente que nunca podrías haberla acabado sin la ayuda de mucha de la gente que te rodea.

La primera persona a la que quiero dar las gracias es a **Benjamín** Gómez, mi abuelo de tesis. El término es suyo, no mío. Benja fue el que me ofreció la oportunidad de entrar en el laboratorio y el que me transmitió esas ganas de saber más. Gracias por dejarme formar parte de este gran equipo y por aguantarme y dirigirme durante todos estos años. También quiero agradecer a mis codirectores, Aritz y Marijo todo el tiempo y esfuerzo que han dedicado a esta tesis.

Al Dr. **Aritz Ruiz**, que me introdujo en el apasionante mundo de los mustélidos, la genética, el GIS, las copias de seguridad, los papers y otras muchas cosas imprescindibles para sobrevivir a una tesis de estas características.

A la Dra. **Marijo** Madeira, mi codirectora, gracias por preocuparte por mí, por servirme de ejemplo, por tener dos manos izquierdas, y por tu apoyo incondicional en todas las veces que pensé en tirar la toalla. ¡Ojalá algún día llegue a ser la mitad de buena investigadora que tú!

A **Oihana**, otra de las personas que más ha sufrido el proceso. Mi chófer, amiga e inseparable compañera de viajes, de mesa, de Power-Points, de máster y sobre todo de cafés. 12 horas al día durante seis años (o más) dan para mucho. Te debo tanto… Espero que sepas que siempre estaré ahí si necesitas enterrar un cadáver, o lo que sea.

A **Urtzi**, por tus chistes y tomaduras de pelo, pero sobre todo, por tus estimulantes temas de conversación y canciones. Sin duda alguna, el más tímido joven mastuerzo que nunca llegaré a conocer. A **Luisja**, el próximo en enfrentarse al temido tribunal y que nunca pierde la sonrisa aunque esté, como todos, al borde del ataque de nervios. A **Amaia**, enviada especial a Nottingham, por sus yogures y su diferente y complementaria manera de ver la vida. A **Gurru**, por su inestimable ayuda con el GIS, porque siempre tiene algo nuevo que contar y ¡por ser único en su especie! A **Eder**, la última incorporación, que nos sorprende cada día

con su entusiasmo y ganas de investigar. A **Javi**, otro que ha tenido que aguantar lo suyo, que me ha enseñado todo lo que sé sobre nutrias y visones, y que me llevó a ver mi primera garduña. Muchas gracias por todo.

I am indebted to Dr. Cushman, who introduced me to the amazing world of multiscaling, HSM, landscape genetics and camping in Arizona! Thank you so much **Sam** for your time, patience, fast corrections, help, advice and wonderful company.

To Dr. **Carlos** Fernandes and Dr. **Mafalda** Basto. Thank you both for immersing yourselves in these stone marten papers, and for your time, effort and great work. It was a fantastic (and exhausting :-) collaboration with great scientists!

Cómo olvidar a toda la gente involucrada en la recolección de muestras de excrementos y tejidos sin las que, obviamente, esta tesis no habría podido realizarse. Muchísimas gracias a todos, investigadores de campo, guardas, particulares, encargados de colecciones, administraciones… Espero no haberme dejado ningún nombre por poner en las larguísimas tablas con las miles de muestras utilizadas. Sí, he dicho miles, soy de Bilbao pero esta vez no es una exageración.

La presente tesis ha sido realizada gracias a la financiación procedente del Programa Predoctoral de Formación de Personal Investigador del Dpto. de Educación, Universidades e Investigación durante el periodo 2011-2014 (Ref: RBFI-2012-446) y de la ayuda para la estancia de la tesis (EP-2014-1-184) del Gobierno Vasco.

Rápidamente me gustaría agradecer a toda la gente que ha pasado por el labo, a amigos con los que he pasado grandes ratos y que ahora están lejos, y a los más recientes y que más cerca han estado en esta recta final. Destacar a Mo por sus mensajes de ánimo, las partidas de Quiz y los cotilleos que hemos compartido estos últimos años. A Naiara, por entenderme y escuchar mis quejas, por sus ánimos y por asegurarse de que no se me pasaba ningún plazo para cumplir con el interminable papeleo. A Niamh, mi profe y amiga, que me ha enseñado, entre cafés y Verdejos, lo importante que es ser feliz y lo maravilloso que es estar enamorada (y también a Asier).

También quiero dar las gracias a mis abuelos, mis tíos y mi prima, pero sobretodo a **mis padres** que me han guiado y acompañado a lo largo del duro proceso, me han servido de ejemplo, y me concedieron las exclusivas becas de formación Vergara-Rosa para convertir-

me en lo que soy ahora. Millones de gracias por esforzaros tanto en entender este trabajo, por valorar mi esfuerzo, por no dejar que me rindiera, y por demostrarme lo que me queréis y lo orgullosos que estáis de mí cada día.

Y por último pero no menos importante a **Gus**, mi media mandarina y compañero de viajes. Gracias por tener siempre preparada una sonrisa, un zumo o un café cada mañana. Gracias por animarme, por maquetarme la tesis, por aguantarme, por bromear sobre mis quejas, por hacerme la vida más fácil, por hacerme tan feliz y por quererme tanto.

Table of contents

▹ Abstract 13

▹ Resumen 15

▹ Preface 22

▹ Chapter I: General Introduction 24

1. Fields of study 25

1.1. Molecular ecology
1.2. Genetic monitoring and non-invasive genetic sampling
1.3. Conservation genetics in endangered and common species
1.4. Phylogeography and population genetics
1.5. Landscape genetics: a combination of population genetics and landscape ecology
1.6. Habitat suitability modeling and niche divergence

2. Target species: *Lutra* and *Martes* genera 31

2.1. Molecular studies conducted in genus *Lutra*
2.2. Molecular studies conducted in genus *Martes*
2.3. Comparative studies on the pine and stone marten habitat use and selection

3. Aims of the thesis 35

4. References 37

▹ Chapter II: Non-invasive genetic sampling for species distribution assessment and habitat suitability modeling. 45

Paper 1: Individual identification and distribution assessment of otters (*Lutra lutra*) through non-invasive genetic sampling: Recovery of an endangered species in the Basque Country (Nothern Spain) 45

Paper 2: Shaken but not stirred: Multiscale habitat suitability modeling of sympatric marten species (*Martes martes* and *Martes foina*) in the northern Iberian Peninsula 74

Paper appendix 1: Distribution and habitat use by pine marten *Martes martes* in riparian corridor crossing intensively cultivated lowland 114

▹ Chapter III: Phylogeography, population genetics and landscape genetics of a widely and continuously distributed mustelid, the stone marten. 125

Paper 3: Inferring population genetic structure in widely and continuously distributed carnivores: the stone marten (*Martes foina*) as a case study 125

Paper 4: Phylogeography and population genetic structure of the stone marten (*Martes foina*) in Europe. 159

▹ Chapter IV: Concluding remarks 206

Abstract

A better understanding of how the physical and temporal factors structure populations and shape species ranges is essential to both their long-term conservation and management. The main aim of this thesis is to gain insights into the intraspecific phylogeny, population genetics, and habitat suitability of several elusive European mustelids [the European otter (*Lutra lutra*), the European pine marten (*Martes martes*) and the stone marten (*M. foina*)], through the application of molecular techniques. Firstly, the results of a systematic non-invasive genetic survey were used to determine the current distribution and population size of the endangered European otter in the northern Iberian Peninsula (Basque Country). Otter spraints, confirmed at specific level by means of mtDNA sequencing, were individually identified through microsatellite and sex-genotyping, which enabled us to record the contemporary range expansion of this elusive mustelid at a regional scale, and to perform the first population and density estimations, essential to redefine population status in order to design management and conservation programs. Further, non-invasive genetic sampling was also successfully applied to build a reliable and representative *Martes spp.* presence dataset (i.e. pine and stone marten) where locations from a long-term non-invasive genetic monitoring and unequivocal species records (i.e. road-killed, live-trapped and camera-trapped individuals) were combined. This information was then used to investigate how different environmental variables, measured at varying spatial scales, shape the distribution of these closely-related and sympatric species with similar ecological requirements. Interestingly, once corrected for sampling bias, the optimized habitat suitability models were compared, detecting a significant niche divergence among the pine and the stone marten in a sympatric area of northern Spain. Lastly, the species identification from scats was successfully employed to measure the variation in the pine marten marking intensity, indicative of habitat use and selection in NW Italy. The relative abundance of scats was found to be dependent on the structure and fragmentation of the residual forests, stressing the role of riparian woods as corridors for the pine marten colonization in highly-agricultural landscapes. Then, we filled current knowledge gaps in the phylogeography and population genetics of one of the least studied marten, the stone marten, at different spatial scales (Iberian Peninsula and Europe). Both studies were conducted combining mtDNA sequencing and microsatellite genotyping from tissue and hair samples. In the first study, within Iberia, microsatellite analyses led to the detection of a marked NE-SW cline in genetic diversity, a strong and pervasive isolation by distance pattern and three genetic breaks corresponding, largely, to the presence of the three main rivers (i.e. Ebro, Tagus and Guadiana). Besides, the lack of phylogeographic

structure in stone marten mtDNA data was consistent with a single colonization event. Regarding the study conducted across European *M. foina* populations, results were congruent with an scenario where several regions acted as potential glacial refugia (i.e. Iberian, Italian, and Balkans peninsulas, and the Pannonian basin) during Pleistocene glaciations, and recurrent secondary contacts of haplotypes from distinct origins during consecutive interglacial periods. Thus, a contraction-expansion model during Pleistocene climatic oscillations was detected which corresponded with to that of an ecological flexible mammal occupying wide range of habitats. The early divergence estimated for the stone marten East-West split (0.55 Mya), categorically contradicts the prevailing hypothesis, according to which the stone marten´s first entrance and posterior dispersal followed the Neolithic farming cultures from the Near East to Southwestern Europe during the Holocene. Overall, this PhD thesis highlighted how the combination of different molecular techniques can provide important contributions to the knowledge on the phylogeography, population genetics, and habitat use and suitability for the otter, the pine and the stone marten. This information is central in guiding the effective conservation and management action in these species, and in providing a methodological framework for other carnivores.

Keywords: *Martes foina*, *Martes martes*, *Lutra lutra*, non-invasive genetic sampling, mtDNA, microsatellites, Habitat suitability modeling, niche divergence, phylogeography, genetic structure.

Resumen

El mantenimiento y la conservación de la diversidad genética de cualquier especie requieren conocer su variabilidad genética y establecer cómo ésta se distribuye a lo largo de las diferentes poblaciones que la componen. En este sentido, los datos moleculares son esenciales para poder obtener una aproximación adecuada y entender los procesos que configuran la estructura genética a nivel intraespecífico. Estos datos proporcionan la información necesaria para entender, tanto la distribución genética actual, como la historia evolutiva de cada especie, información que puede ser utilizada como referencia en cualquier programa de conservación.

El objetivo principal de esta tesis es obtener información acerca de la filogenia intraespecífica, la genética de poblaciones y la idoneidad de hábitat ("habitat suitability") de varios mustélidos europeos (Carnivora, Mammalia) mediante la aplicación de diferentes técnicas moleculares. Esta tesis combina aproximaciones y metodologías de diferentes disciplinas -tales como filogenias moleculares, genética de poblaciones, genética del paisaje o modelización de la idoneidad del hábitat- para explorar algunos de los principales patrones y procesos evolutivos y ecológicos de 3 especies de la familia Mustelidae: la nutria Europea (*Lutra lutra,* Linnaeus, 1758), la marta Europea (*Martes martes,* Linnaeus, 1758) y la garduña (*Martes foina,* Erxleben, 1777). Esta información pretende contribuir a obtener un mejor conocimiento de la ecología y la historia evolutiva de estas especies, y proporcionar datos relevantes para la conservación de su diversidad genética.

Capítulo I

En el capítulo introductorio se realiza una descripción de las distintas disciplinas abordadas a lo largo del resto de los capítulos para contextualizar los cinco artículos incluidos en esta tesis. Además, se incluye una revisión de los estudios moleculares más significativos en el ámbito de la ecología molecular de los géneros *Lutra* y *Martes,* así como de los trabajos en los que se establecen comparaciones interespecíficas sobre el uso y selección del hábitat de la marta y la garduña.

Capítulo II

Este capítulo engloba tres de los cinco trabajos incluidos en esta tesis (Artículos 1, 2 y Apéndice 1). Todos ellos se basan en la aplicación de técnicas moleculares en muestras no invasivas, en las que el material genético se obtiene sin necesidad de capturar o interferir con

la especie objeto de estudio. La identificación y seguimiento genético a través de muestreos genéticos no invasivos, en particular a partir de muestras de excrementos, se considera un método fiable y efectivo para el seguimiento de carnívoros elusivos, crípticos o en peligro de extinción. En este caso, el ADN extraído a partir de las muestras fecales se utilizó para determinar de manera no invasiva el origen específico de las muestras (nutria, marta o garduña) y determinar de manera precisa la distribución de las diferentes especies con el fin de abordar diferentes cuestiones ecológicas.

Artículo 1: "**Individual identification and distribution assessment of otters (*Lutra lutra*) through non-invasive genetic sampling: Recovery of an endangered species in the Basque Country (Northern Spain)" -"Identificación individual y distribución de la nutria (*Lutra lutra*) mediante técnicas moleculares no invasivas: recuperación de una especie amenazada en el País Vasco."** *Mammalian Biology* 79 (2014) 259–267.

En este trabajo se emplearon técnicas genéticas no invasivas para realizar un seguimiento de las poblaciones de nutria en el País Vasco. Esta metodología resulta muy idónea para la vigilancia y el estudio de las poblaciones de nutria debido, en gran parte, a la linealidad de sus desplazamientos y sus características de marcaje. En España, como en muchos países europeos, las poblaciones de nutria se han visto reducidas progresivamente desde mediados del siglo XX y en particular desde 1960. Sin embargo, los tres sondeos nacionales de nutria llevados a cabo en 1985, 1996 y 2006 registraron un aumento significativo en la distribución de la especie, que pasó de estar presente únicamente en un 33% del territorio a ocupar el 50% y el 70%. No obstante, la nutria sigue ausente, o presente en muy bajas densidades, en algunas regiones españolas como el País Vasco, donde se encuentra incluida en el Catálogo Vasco de Especies Amenazadas como "En Peligro de Extinción" y, desde 2004, cuenta con un Plan de Gestión en Álava, el único territorio de la Comunidad Autónoma del País Vasco con presencia estable de la especie. Durante el periodo 2007-2010 se realizó un muestreo sistemático para la obtención de muestras fecales en diferentes ríos, arroyos y humedales de la provincia de Álava y zonas limítrofes con el fin de determinar la distribución espacial de la nutria y estimar su tamaño poblacional. Todas las muestras fueron identificadas a nivel específico, previo al proceso de genotipado, mediante la secuenciación de un fragmento diagnóstico de 226 pb del ADNmt. Del total de las 132 muestras analizadas, un 98,4% pertenecieron a la especie en cuestión, lo que se tradujo en un aumento de la distribución de la nutria del 25%, basado en cuadrículas UTM 10×10 km ocupadas por la especie, y un claro patrón de recolonización aguas arriba de los principales ríos. Las muestras identificadas como nutria fueron posteriormente genotipadas a nivel individual mediante una batería de

11 marcadores microsatélites y sexadas mediante el tipado del gen ZFX/ZFY ligado a ambos cromosomas sexuales. De esta forma, se obtuvo un perfil genético individual completo para un total de 55 muestras (éxito de genotipado del 43%). Tras un proceso de reagrupamiento de muestras se identificaron un total de 20 ejemplares: 6 machos, 11 hembras y 3 ejemplares con sexo indeterminado. La densidad media en los ríos con presencia continuada de nutria ha sido de 0,09 nutrias/km (0.06-0.12). Este trabajo ha permitido poner a punto toda una metodología de seguimiento y censo de las poblaciones de nutria mediante muestras fecales en Álava, aspecto de gran importancia para realizar una adecuada gestión y conservación de la especie.

Artículo 2: "**Shaken but not stirred: Multiscale habitat suitability modeling of sympatric marten species (*Martes martes* and *Martes foina*) in the northern Iberian Peninsula"–"Juntas pero no revueltas: modelización predictiva multiescala de la idoneidad del hábitat de las especies simpátricas (*Martes martes* y *Martes foina*) en el norte de la Península Ibérica".** Journal: *Landscape Ecology* (Invited special issue). Status: *Submitted*.

En este trabajo se combinaron los resultados del seguimiento genético no invasivo realizado durante el periodo 2005-2012 con los registros de las especies obtenidos en el marco de otros estudios de fauna (i.e. ejemplares atropellados, ingresados en centros de fauna, foto-trampeados y/o trampeados en vivo), para crear una amplia base de datos de presencia de marta y garduña (n=1286). Estos datos se emplearon posteriormente para estudiar cómo las diferentes características del paisaje afectan a la distribución de estas especies por medio de la modelización de la idoneidad de hábitat ("Habitat Suitability Modeling, HSM") de cada una de ellas y, posteriormente, evaluar si existen diferencias de nicho ecológico significativas entre ambas especies cuando coexisten. Estudios recientes han demostrado que la heterogeneidad del hábitat influye en la distribución y la abundancia de las especies a diferentes escalas, y que las respuestas ecológicas de los organismos varían también en función de la escala de análisis. Por lo tanto, los modelos de idoneidad de hábitat mutiescala y multiespecíficos resultan de gran utilidad para identificar las variables ambientales que influyen en la selección de hábitat y sus escalas y facilitar la comparación de la distribución de especies estrechamente relacionadas pero que potencialmente presentan diferentes requerimientos ecológicos. En este trabajo, previo a la elaboración de los modelos, se analizaron por separado las 40 variables de usos del suelo, topográficas y climatológicas para cada una de las 9 escalas (usando ventanas con radios de 125 m a 32 km), con el fin de seleccionar aquella escala a la que cada variable se encontraba asociada de manera más significativa a la presencia de cada una de las especies. Posteriormente, el subconjunto de variables se empleó para elabo-

rar 300 modelos de idoneidad de hábitat o HSMs para cada una de las especies y seleccionar aquellos que presentaban los mejores rendimientos.

Una limitación de los modelos de idoneidad de hábitat es que, cuando los datos de presencia de las especies derivan de observaciones oportunistas en lugar de muestreos sistemáticos, hay áreas que se muestrean más que otras, por lo que se introduce un sesgo en el muestreo. Por tanto, en este trabajo se exploraron tres de las correcciones más utilizadas, a cuatro escalas diferentes, con el objetivo de subsanar el sesgo de muestreo y optimizar los HSMs identificados como los más acertados para la marta y para la garduña con el fin último de compararlos. De este modo, se detectó una significativa diferenciación de nichos entre estas especies simpátricas. En el área de estudio, la marta se asocia con las zonas más frías, con alta proporción de bosques naturales y plantaciones forestales, menor grado de perturbación humana y mosaicos agroforestales de mediano tamaño. La garduña, sin embargo, se encuentra relacionada con zonas urbanas, tierras de cultivo, zonas con presencia parcial de matorral y la disponibilidad de pastizales semi-continuos. Por tanto, este estudio pone de manifiesto cómo las diferentes características del paisaje, medidas a diferentes escalas espaciales, condicionan la distribución y preferencias de hábitat de la marta y de la garduña. Además, este trabajo proporciona un marco metodológico para futuras comparativas entre múltiples especies y escalas.

Apéndice 1: "**Distribution and habitat use by pine marten in a riparian corridor crossing intensively cultivated lowlands**"**–"Distribución y uso del hábitat de la marta en un corredor ribereño que cruza zonas bajas intensamente cultivadas".** *Ecological Research* (2015). Doi: 10.1007/s11284-014-1220-8.

En este trabajo, la identificación genética a partir de excrementos se utilizó con éxito para medir la variación en la intensidad de marcaje de la marta. En este estudio, la abundancia relativa de muestras fecales aparece relacionada con la estructura y la fragmentación de los bosques residuales. Como resultado más significativo, se puso de manifiesto el papel fundamental de los bosques de ribera como corredores en la colonización de la marta en paisajes altamente agrícolas del Noroeste de Italia.

Capítulo III

El Capítulo III se centra en el estudio de la filogeografía y genética de poblaciones de la garduña a lo largo de su rango de distribución occidental. La filogeografía es una disciplina

que trata de establecer los patrones y los procesos que rigen la distribución geográfica de los linajes y su diversidad genética. La filogeografía, además, sirve de puente entre la filogenia y la genética de poblaciones, e incorpora información complementaria de otras disciplinas, por lo que ha sido ampliamente utilizada en diferentes campos como por ejemplo la ecología o la conservación.

Los resultados obtenidos a partir de los estudios sobre la filogeografía y la genética de poblaciones de la garduña en la Península Ibérica (Artículo 3) y el estudio posterior desarrollado a escala continental, incluyendo todas las muestras obtenidas en Europa (Artículo 4), se presentan en el Capítulo III. Ambos trabajos comparten los mismos marcadores moleculares (secuencias de ADN mitocondrial y genotipado de microsatélites) y el tipo de muestra (tejido o pelo), para responder a preguntas similares pero a diferentes escalas geográficas, permitiendo obtener conclusiones complementarias.

Artículo 3: "**Inferring population genetic structure in widely and continuously distributed carnivores: the stone marten (*Martes foina*) as a case study**"–**"Estudio de la estructuración genética en carnívoros con distribución amplia y continua: la garduña (*Martes foina*) como caso de estudio".** Journal: *Plos One*. Status: *Under review*.

Este trabajo, que se centró en las poblaciones de la garduña en la península Ibérica, representa el primer estudio filogenético a nivel intraespecífico desarrollado sobre esta especie. La garduña es un mustélido ampliamente distribuido en la región paleártica con preferencias de hábitat variables a lo largo de su área de distribución. Según los restos fósiles, la especie llegó a la península desde del suroeste de Asia siguiendo la expansión humana del Neolítico hace aproximadamente 7.000 años. Sin embargo, tanto la biogeografía como la estructura genética de este carnívoro generalista siguen siendo una incógnita. En este estudio se combinó la información inferida a partir de la secuenciación de un fragmento de ADN mitocondrial (620 pb) y el genotipado de 23 marcadores microsatélites polimórficos para determinar la estructura genética de las poblaciones de garduña en la Península Ibérica. Los datos obtenidos no evidenciaron la presencia de estructuración filogeográfica a nivel mitocondrial, lo que parece indicar una colonización relativamente reciente de la península que coincidiría con los datos fósiles publicados. Sin embargo, en este trabajo se apuntó también la importancia de ampliar el área de estudio con el fin de verificar dicha hipótesis.

Los datos obtenidos mediante marcadores microsatélites se analizaron con a) diferentes métodos de agrupación Bayesiana, espaciales y no espaciales (STRUCTURE, TESS, BAPS y

GENELAND), y b) Métodos multivariantes [Análisis Discriminante de Componentes Principales (DAPC) y Análisis espacial de componentes principales (o SPCA por sus siglas en inglés)]. El aislamiento por distancia es un patrón genético espacial muy común en especies con distribuciones continuas y puede presentar un reto para el desempeño de los métodos anteriores. En este trabajo se detectó un patrón espacial noreste-suroeste en la estructuración genética de la garduña, lo que podría explicar la discrepancia observada entre los diferentes resultados de agrupación obtenidos por los diferentes métodos. La división genética detectada entre las 3 subpoblaciones se correspondió con la presencia de los ríos Ebro, Tajo y Guadiana, lo que sugiere que los principales cursos de agua en la Península Ibérica podrían actuar como barreras semipermeables para el flujo genético de la especie. Este trabajo se abordó por primera vez el análisis de la estructuración genética de la especie en una amplia escala geográfica, recalcando la importancia y los beneficios que reporta la utilización y la comparación de varios métodos de agrupación y de análisis multivariante en los estudios genéticos de especies con áreas de distribución amplias.

Artículo 4: "**Phylogeography and population genetic structure of the stone marten (*Martes foina*) in Europe"–"Filogeografía y estructuración genética poblacional de la garduña (*Martes foina*) en Europa".** *Manuscript.*

Para la realización de este trabajo, se llevó a cabo una importante labor de búsqueda de muestras de pelo y tejido de garduña (en museos, colecciones privadas, centros de fauna y otras enviadas por investigadores de campo), lo que nos permitió obtener y analizar 751 muestras recogidas en 29 países europeos y de Oriente Próximo (Israel e Irán). Sobre estas muestras, se aplicaron los mismos marcadores mitocondriales y microsatélites empleados en el artículo 3 con el fin de: i) estudiar los principales patrones filogeográficos de la garduña para conocer la historia biogeográfica de la especie en el continente Europeo, ii) estimar el periodo de entrada y posterior diversificación de los linajes de la garduña en Europa, iii) testar la teoría postulada por varios autores, según la cual la especie habría colonizado Europa siguiendo la expansión humana del Neolítico, y iv) obtener una primera aproximación acerca de la estructuración poblacional actual de la garduña en Europa. Los resultados obtenidos indicaron una importante diferenciación de las poblaciones de garduña en Europa en dos linajes principales (Oeste y Este) que tuvo lugar hace aproximadamente 0.55 millones de años (Pleistoceno Medio), lo que contradice categóricamente la hipótesis preestablecida que relacionaba la expansión de la garduña con la entrada de la cultura Neolítica en Europa. El clado del Oeste lo componen principalmente los individuos muestreados en Europa Occidental, desde la Península Ibérica a Dinamarca. Dentro del clado del Este, que presenta ma-

yor diversidad genética, se diferencian tres clusters designados como 'Balcanes', 'Italia-Norte de Europa ' y ' Centro-Norte de Europa'. Así, la diferenciación genética y geográfica observada entre estos linajes parece ser el resultado de la fragmentación del área de distribución de la garduña y el posterior aislamiento y divergencia de sus poblaciones durante el Cuaternario, confinadas en distintos refugios glaciares. Se han postulado tres refugios Mediterráneos ubicados en las penínsulas Ibérica, Itálica y Balcánica, y un cuarto refugio críptico potencialmente localizado en la Llanura Panónica (Los Cárpatos). Además, atendiendo a la red de haplotipos, la zona de Grecia pudo actuar como un refugio independiente dentro de la zona de los Balcanes. La distribución y estructuración genética actual de la especie sugiere que, durante los recurrentes períodos interglaciares, se habrían sucedido diferentes contactos secundarios entre poblaciones de orígenes distintos, que habrían sido facilitados por la gran capacidad dispersiva de la especie, o bien debido a procesos de introducción-translocación humana, como se infiere de otros estudios filogeográficos en carnívoros. En este caso, los linajes parecen converger en Europa Central que arroja los valores más elevados de diversidad, tanto a nivel mitocondrial como nuclear. Este trabajo nos ha permitido profundizar acerca de la filogeografía y la estructuración genética de la garduña en Europa y conocer una historia evolutiva profundamente ampliamente influenciada por las glaciaciones del Pleistoceno y mucho más antigua de la que se postulaba en estudios previos basados en evidencias fósiles.

Capítulo IV

En este último capítulo se presenta un resumen de las principales conclusiones contenidas en los capítulos anteriores y las contribuciones más relevantes de este trabajo al conocimiento y conservación de las especies objeto de estudio y a las disciplinas bajo las que se enmarca.

Preface

This thesis is divided into 4 chapters. The general framework is given by the introduction (**Chapter I**), where I briefly describe the main subdisciplines in molecular ecology and habitat suitability modeling addressed in the dissertation. Then, I reviewed the main and most significant and comparable molecular studies conducted on genera *Lutra* and *Martes* and those focused on the interspecific habitat use and suitability comparisons among the sympatric pine marten (*Martes martes*) and stone marten (*Martes foina*), to contextualize each of the 5 manuscripts reported in the present work.

Three papers are included under **Chapter II**, all based on the genetic species identification of the target species [i.e. Eurasian otter (*Lutra lutra*), pine marten and stone marten)] from non-invasively collected faecal samples, to address different ecological questions:

First, in ***Paper 1***, the results of a systematic non-invasive genetic survey were used to determine the current distribution and population size of the endangered European otter population in the northern Iberian Peninsula (Basque Country). Otter spraints, genetically confirmed at specific level by means of mtDNA sequencing, were individualy identified through microsatellite and sex-genotyping, which enabled to record the contemporary expansion of this elusive mustelid at regional scale and to perform the first population and density estimations, essential to redefine population status and to design successful and effective management and conservation programs.

In ***Paper 2***, to build a representative marten´s presence dataset, locations from a long-term non-invasive genetic monitoring and from unequivocal species records (i.e. road-killed, live-trapped and camera-trapped individuals) were combined. This information was then used to investigate how different landscape characteristics, measured at varying spatial scales, shape the distribution of each species in the northern Iberian Peninsula (Basque Country and Navarre). Once corrected for potential sampling bias, the resulting optimized habitat suitability models were compared, detecting a significant niche divergence among the pine and the stone marten.

In the third paper, included as an appendix (***Paper Appendix 1***), the species identification from scats was successfully employed to measure the variation in the pine marten marking intensity, and to explore habitat use and selection at microgeographic scale in NW Italy. The relative abundance of scats was found to be dependent on the structure and fragmentation

of the residual forests, stressing the role of riparian woods as corridors for the pine marten colonization in highly-agricultural landscapes.

Both the phylogeography and population genetics of the stone marten in the Iberian Peninsula (***Paper 3***), and the extended research conducted across Europe (***Paper 4***) are reported in **Chapter III**. These studies share the same genetic markers (mtDNA sequences and microsatellite markers, n = 23) and sample type (tissue and hair) to answer similar questions about the instraspeciphic phylogenetics and the current population genetic structure at different spatial scales. In Iberia, microsatellite analyses led to the detection of a marked NE-SW cline in genetic diversity, a strong and pervasive isolation by distance pattern, and genetic breaks largely corresponding to the presence of the three main rivers (i.e. Ebro,Tagus and Guadiana). In this case, the lack of phylogeographic structure in the stone marten mtDNA data, was consistent with a single colonization event. Concerning the study conducted across European *M. foina* populations, results were congruent with a scenario with several regions acting as potential glacial refugia during Pleistocene glaciations (i.e. Iberian, Italian, and Balkan peninsulas, as well as the Pannonian basin), and recurrent secondary contacts of haplotypes from distinct origins during consecutive interglacial periods, favored by the great dispersal ability of the species and/or human mediated introductions/translocations. The early divergence estimated for the stone marten East-West split (0.55 Mya), categorically contradicts the prevailing hypothesis, according to which the stone marten first entrance into Europe and posterior dispersal followed the Neolithic farming cultures from the Near East to Southwestern Europe during the Holocene.

Finally, in **Chapter IV**, I summarized the major findings from the previous chapters and the most important contributions of this work to the field and to the knowledge and conservation of the otter, the pine marten and the stone marten.

Chapter I

General Introduction

1. Fields of study

The multidisciplinary nature of this thesis makes necessary not only to explain each of the main disciplines addressed, but also to briefly describe how they interlace with each other. Thus, the subdisciplines that set the frame for this work are listed below.

1.1. Molecular Ecology

Ecology is a branch of biology focused on the study on how organisms interact with each other as well as in the interactions among organisms and their environment (Freeland et al 2011). These relationships were historically explored by direct observations in the field or experimental manipulations using phenotypic data alone (i.e. the expression of genes in an observable way) contributing, significantly, to the knowledge of processes that maintain ecosystems (Freeland et al 2011). Used in isolation, however, phenotypic data has some limitations as, under different environmental conditions, a single genotype can develop into multiple alternative phenotypes (Ouborg et al 2010). Precisely, this complex interaction among genotypes, phenotypes and environmental conditions, is what makes the development of cost-effective molecular methods necessary (Freeland et al 2011).

The discovery, nearly 50 years ago, of how individual genetic variation could be quantified based on the variance in the structure of proteins (Harris 1966) was considered one of the first step of the discipline of molecular ecology, which is, essentially, the application of different molecular genetic techniques to answer traditional ecological and evolutionary questions (Freeland et al 2011). Molecular ecology encompasses a wide range of aspects such as: a) the study of the relatedness, behavior and dispersal of individuals across landscapes, b) the interaction between different taxa, hybridization, and formation of new species, or c) the evolution and genetics of ecologically important traits (Andrew et al 2013). Since its origins, this discipline has continuously evolved, incorporating the latest analytical approaches and methodologies to address previously unanswerable ecological questions and to evaluate hypothesis that remained untested (Andrew et al 2013). Definitely, the development of analytical methods that facilitate the rapid acquisition of large amounts of genotypic data, the availability of user-friendly software to analyze massive data sets and the continuous increase in theories to interpret those results have been essential drivers of molecular ecology (Beebee and Rowe 2008; Andrew et al 2013).

1.2. Genetic monitoring and non-invasive genetic sampling

Genetic monitoring has become an important tool for wildlife conservation and management, offerring insights into evolutionary and demographic processes in either captive or natural populations impossible to obtain using traditional methods (Allendorf and Luikart 2007; Schwartz et al 2007).

In genetic monitoring, diagnostic genetic markers are used to identify genera, species, subspecies and genetically differentiated populations in order to monitor the temporal changes in populations genetic metrics (i.e. genetic diversity, population size or Ne) and the consequences of anthropogenic impact on wild species (Wayne and Morin 2004; Schwartz et al 2007). Besides, the identification of individuals allow tracking their movements over time, identifying migrants, estimating sex-ratios and abundances and assesing species distribution (Allendorf and Luikart 2007). Depending on the sample type, the genetic monitoring can be invasive (e.g. tissue, hair or blood samples from live trapped animals) or non-invasive, where the DNA is obtained from different sources such as feces, urine, saliva, or shed hair and feathers without capturing or even observing the animal and thus, avoiding disturbance to wild populations (Allendorf and Luikart 2007; Beja-Pereira et al 2009). Thus, species and individual identification via non-invasive genetic sampling, particularly from scats samples, is considered a reliable and cost-effective way for monitoring elusive, cryptic or endangered carnivores (Taberlet and Luikart 1999; Waits and Paetkau 2005; Schwartz and Monfort 2008; Bonesi et al 2012; Ruiz-González et al 2013b).

1.3. Conservation genetics in endangered and common species

Genetic diversity is recognized as a fundamental component of biodiversity and has become a recurrent objective for conservation biology because of its implications for understanding the biology of species and populations (Noos 1990). Genetic diversity, also called genetic variability, is critical for populations to cope with environmental change and both the short and long-term persistence of populations and species (Ouborg et al 2010), thus, a central issue in conservation genetics is the amount of genetic variation present (Pertoldi et al 2007). In this context, the increasingly available DNA analysis methods have favored a rise to conservation studies on the genetic diversity and the causes of spatio-temporal population dynamics in different taxa (Allendorf and Luikart 2007; Schwartz et al 2007).

Since its foundation in the 80´s (Soulé and Wilcox 1980; Frankel and Soulé 1981), the role of conservation genetics has steadily increased, shifting from a branch of conservation biol-

ogy to a full-grown empirical discipline (Frankham et al 2002; Allendorf and Luikart 2007; Pertoldi et al 2007; Ouborg et al 2010). Certainly, the technical breakthroughs on DNA analysis have been important drivers of the expanded role of genetics in conservation (De Salle and Amato 2004). Conservation genetics is an interdisciplinary science aiming to minimize the risk of extinction from genetic factors. Briefly, this discipline encompasses a) the genetic management of small populations to maximize retention of genetic diversity and minimize inbreeding, b) the resolution of taxonomic uncertainties and delineation of management units, and c) the use of molecular genetic analyses in forensics and to understand species´ biology (Frankham et al 2002).

In practice, although not in principle, conservation is a crisis discipline that prioritizes rare species, including those in risk of extinction (e.g. the European otter), constrained into small populations within restricted geographic ranges and thus, common species occurring in large numbers (e.g. the stone marten) receive little of none attention (Gaston and Fuller 2008). There are, however, various sound cases of how once widespread species are currently in danger of disappearing (e.g. the European bison, *Bison bonasus*) or have already became extinct (e.g. The Rocky Mountain locust, *Melanoplus spretus*), supporting that the risk of extinction in common species is more likely that may seem (reviewed in Gaston and Fuller 2008). Thus, to prevent widespread species from disappearing, we need to be pro-active, as the early recognition of threats, decreases in abundance and distribution, and the identification of isolated and fragmented populations, can guide management actions for their long-term conservation (Lindenmayer et al 2011).

1.4. Phylogeography and population genetics

To preserve the genetic diversity of a species it is vital to find out its genetic variability and how it is distributed along the different populations (Avise et al 1987; Avise 2000). Intraspecific phylogenies differentiate between populations from different parts of the distribution of a given taxon (Schwartz et al 2012). In phylogeographic studies, intraspecific phylogenies are analyzed jointly with the spatial distribution of the species to infer the influence of geographic features (e.g. mountains, rivers, roads) and historical processes for structuring populations in evolutionary time scales (Avise et al 1987; Avise 2000). Hence, phylogeography comprises aspects of both time (evolutionary relationships) and space (geographic distributions). In mustelids, for instance, phylogenetic and phylogeographic studies have been useful to investigate historical changes in genetic diversity (e.g. Pertoldi et al 2008), to delimit conservation units (Sato et al 2009), to clarify taxonomic uncertainties (e.g. Drew et

al 2003), to identify the origin of species traslocations (e.g. Drew et al 2003) or to reconstruct postglacial colonization histories (e.g. Ruiz-González et al 2013).

There are several techniques for quantifying the genetic variation within and among populations but the choice of molecular markers depends highly on the population or evolutionary genetic questions raised, as no marker appropriate for all applications exists (Beebee and Rowe 2008). Yet, as reported by Beheregaray (2008), almost the 70% of the phylogeographic studies conducted on mammals used exclusively mitochondrial DNA (mtDNA) data, while 18% combined this with additional molecular markers. There are several reasons for the mtDNA to be the marker of choice. First, it is haploid and maternally inherited and occurs in thousands of copies per cell, and second, its relatively fast evolutionary rate is well suited to examine events that occurred hundreds of thousands to a few million years ago (Allendorf and Luikart 2007). In phylogeographic studies conducted on genera *Lutra* (e.g. Mucci et al 2010) and *Martes* (reviewed in Schwartz et al 2012) the D-loop and cytochrome-b (cyt-b) regions were the most commonly used mtDNA markers.

For the study of relatively recent events (ca. ten thousand years ago) markers with higher mutation rate, such as microsatellites, are needed (Hewitt 2004). These codominant and neutral nuclear markers allow discriminating different individuals and populations, being the most popular type of marker in population genetics (Freeland et al 2011). They can be used to evaluate the genetic structuring and diversity of populations and subpopulations (e.g. F-statistics and genetic distances), to asses demographic histories (e.g. population bottlenecks), to evaluate inbreeding and relatedness between individuals, to assess effective population sizes (Ne), or to measure the strength and direction of gene flow among populations (Schwartz et al 2012).

1.5. Landscape genetics: a combination of population genetics and landscape ecology

Landscape genetics is an interdisciplinary field that combines population genetics, landscape ecology and spatial statistics (Manel et al 2003; Holderegger and Wagner 2008; Manel and Holderegger 2013). The central issue is to quantify the effect that landscape composition, configuration and quality have on micro-evolutionary processes and on the spatial patterns of genetic variability (Balkenhol et al 2009).

In landscape genetic analyses, the individual is used as the operational unit, avoiding potential biases associated with the delimitation of pre-defined populations and leading to studies at higher resolution (Manel et al 2003). The individual genetic data along with each

exact geographical location is then used to correlate the genetic discontinuities found to the landscape features (e.g. rivers, roads) and environmental variables (e.g. temperature, elevation) of the study area (Manel et al 2003; Holderegger and Wagner 2008; Wagner and Fortin 2013). Thus, the main difference between landscape genetics and population genetic studies is that, for the first, the quantity and quality of the matrix separating individuals is considered a key determinant of the underlying biological and ecological processes while the latter is simply concerned with its extent (Storfer et al 2010). Further, landscape genetics (concerned with contemporary processes) differs from phylogeography (dealing with historical processes) in the temporal scale of the processes under study (Wang 2010).

Storfer et al (2010) identified the following major research topics in landscape genetics: a) the quantification of the influence of landscape variables and their configuration on genetic variation, b) the detection of barriers to gene flow, source-sink dynamics and corridors, c) the inference of the spatio-temporal scales of ecological processes and d) the evaluation of species-specific ecological theories.

1.6. Habitat suitability modeling and niche divergence

Habitat Suitability Models (HSMs) (Franklin 1995; Guisan and Zimmerman 2000; Girvetz and Greco 2009; Bellamy et al 2013) also known as Species Distribution Models (SDMs) (Syfert et al 2013; Kramer-Schadt et al 2013; Fourcade et al 2014) or Environmental (or ecological) niche models (ENMs) (Warren et al 2010; Warren and Seifert 2011), have become a fundamental tool in ecology and biogeography as they correlate the presence of species at multiple locations with relevant environmental covariates to estimate habitat preferences or predict species presences and absences throughout the study area (Franklin 1995; Guisan and Zimmerman 2000; Elith et al 2011).

The identification of the factors constraining species presence is essential to identify the most suitable areas for a particular taxon, to infer occurrence probability in areas where no systematic surveys have been conducted (Fourcade et al 2014), to estimate the species niche (Warren and Seifert 2011), to assess the potential expansion of invasive species (i.e. Elith et al 2010; Jiménez-Valverde et al 2011) or to estimate future ranges under different climate change scenarios (i.e. Khanum et al 2013). However, most species presence records correspond to presence-only datasets (PO), providing actual information regarding the species' presence but no direct data regarding absences (Syfert et al 2013).

Most PO dataset are build with records from museums, herbarium collections or from

open biodiversity databases (e.g. GBIF, the Global Biodiversity Information Facility) (Fourcade et al 2014). Yet, any accurate species location can be used as presence data in HSM. MAXENT (Phillips et al 2006) is one of the best performing and most commonly used software to predict species distributions using PO data, as it is considered to work well with sparse and irregularly sampled locations (Elith et al 2006). MAXENT uses the maximum entropy principle to relate species presence-only records and environmental variables in order to estimate species´niches and their potential spatial distibution (Phillips et al 2006). Even if HSM work under the assumption that the entire region under study has been systematically and randomly sampled, most datasets include spatially biased presence records (Elith et al 2011). The effect of sampling bias in HSM performance has been the focus of interest of several recent papers (Anderson and Gonzalez 2011; Syfert et al 2013; Kramer-Schadt et al 2013; Fourcade et al 2014; Brown 2014) in which a number of corrections methods have been proposed.

Species use habitats differently at widely divergent scales (Johnson 1980; Cushman and McGarigal 2004; Graf et al 2005). Therefore, habitat suitability modeling has shifted from models based exclusively in expert's opinion to increasingly complex multiscale models to reveal the true grain at which species respond to the landscape as there is no *a priori* way to infer the grain and extent at which each environmental predictor is most strongly related to species presence (Shirk et al 2012). In this context, recent studies conducted on mammals (e.g. (Wasserman et al 2012; Mateo-Sánchez et al 2013; Bellamy et al 2013; Shirk et al 2014) have demonstrated the effectiveness of multiscale approaches.

HSM can also be employed to explore the environmental characteristics conditioning several species´ overlapping range as a first step to quantitatively estimate the niche conservatism/divergence in related species and to infer the roles of competitive interaction (e.g. Wellenreuther et al 2012), as species realized niche could be restricted by the competence with their congeners, often found in co-existence (Anderson et al 2002). Niche comparisons may be performed in geographic space by comparing the predicted species distributions obtained by HSM (e.g. Anderson et al 2002) or in the environmental space if measuring niche similarities (e.g. Reutter et al 2003). Further, two conditions have to be met in order to consider a significant niche divergence among two species. First, niche characteristics have to differ between species, and second, the differences have to be necesarily greater than the backgrond environmental difference (McCormack et al 2010; Warren et al 2010).

2. Target species: *Lutra* and *Martes* genera

The diversification of the Mustelidae family (Carnivora, Mammalia) is an outstanding example of adaptive radiation (Schluter 2000; Koepfli et al 2008) as mustelids show marked ecomorphological diversity, reflecting how different species within the family have adapted to both different habits and environments. The species of genus *Lutra*, for instance, are semi-aquatic mammals, and thus well-adapted to water and land, while the forest-dwelling martens are better adapted to their semi-arboreal lifestyle.

Three mustelid species with different characteristics were chosen as subject of this dissertation: a) the European otter, a rare and elusive semi-aquatic species listed as endangered in the study area (Basque Country), b) the stone marten, a generalist common and abundant species with important information gaps on its phylogeography and population and landscape genetics, and c) the pine marten, a forest-dwelling species, closely-related to the stone marten and living sympatrically over large areas but showing different ecological requirements and evolutionary patterns, and thus, enabling interesting comparisons regarding their multiscale habitat suitability and niche divergence.

2.1. Molecular studies conducted in genus Lutra

The genus *Lutra* includes three species with different conservation status; the hairy-nosed otter (*L. sumatrana),* the spotted-necked otter *(L. maculicolis)* and the European otter (*L. lutra*) (IUCN, 2008).

The endangered *L. sumatrana* is probably the least studied and rarest of the world's otters (Hussain et al 2008; Wright et al 2008). However, even if more widespread, none phylogeographic or population genetic studies have been conducted on the least concern African species *L. maculicolis* (Hussain et al 2008). The only sound studies on these two otters are the phylogenies build on morfological similarities (van Zyll de Jong 1987) and those based on *cyt-b* mitochondrial sequences (Koepfli and Wayne 1998) and multigene data sets (Koepfli et al 2008) including all of the 13 otter species. *L. lutra*, on the other hand, has been the focus of a wide range of molecular studies (see below).

The European otter (*Lutra lutra*) has one of the widest distributions of all Palaearctic mammals being present in most of Europe, Asia and North Africa (Ruiz-Olmo et al 2008). The Eurasian otter is currently a species of conservation concern and it is listed as near threatened by the IUCN mainly due to the decline of its populations in large areas of Eu-

rope during the twentieth century (Ruiz-Olmo et al 2008). Yet, after a strong decline in the 1980s, the otter is recovering in many European countries where subtantial natural or human-mediated (i.e. reintroductions programs) populations increases have been observed (McDonald et al 2007; Sulkava 2007; Prigioni et al 2007; Koelewijn et al 2010; Romanowski et al 2013). However, studies conducted on this species highlighted both the low levels of mtDNA variation (e.g. Ferrando et al 2004; Mucci et al 2010) and geographic structure (e.g. Randi et al 2003; Ferrando et al 2007; Mucci et al 2010; Hobbs et al 2011).

The marking behaviour of the otter used for intraspecific communication, makes this semi-aquatic mammal a very suitable species for non-invasive genetic monitoring (Kruuk 1995; Ben-David et al 2005; Kruuk 2006). To date, the non-invasive approach has been fruitfully used to estimate central parameters for otter conservation and management such as population sizes and sex-ratio (Dallas et al 2003; Prigioni et al 2006; Kalz et al 2006; Koelewijn et al 2010; Bonesi et al 2012), kindship (Kalz et al 2006), dispersal (Hobbs et al 2006; Kalz et al 2006; Koelewijn et al 2010) and to evaluate the success of reintroduction projects (Ferrando et al 2007; Koelewijn et al 2010). Thus, the combination of species, sex and individual genetic identification through spraints with spatial data has become a valuable method for recording population trends and range expansions (or contractions) in scarced or endangered otter populations, providing basic information for the design of conservation and management programs (Ruiz-González et al 2008a).

2.2. Molecular studies conducted in genus Martes

Among the 8 species of martens (IUCN, 2015), *M. americana*, *M. martes* and *M. pennanti* are clearly the species with the greatest number of molecular studies (reviewed in Schwartz et al 2012). Three unequally well studied species occur in Europe, the pine marten (*M. martes*), the stone marten (*M. foina)* and the sable *(M. zibellina*) (Proulx et al 2004). The pine marten and the sable are sister species within the subgenus *Martes* that diversified during the Plio-Pleistocene, with overlapping populations in part of European Russia (Koepfli et al 2008).

The pine marten has been the focus of several population genetic (Kyle et al 2003; Mergey 2007; Pertoldi et al 2008; Ruiz-González et al 2013a; Pertoldi et al 2014) and landscape genetic studies (Mergey et al 2012; Ruiz-González et al 2014). Regarding its intraspecific phylogeny, a recent study on the pine marten conducted by Ruiz-González et al (2013a) provided interesting evidences of the complex post-glacial recolonization of Europe and

supported the genetic introgression (i.e. contemporary or historical hybridization) of pine marten and sable reported by Davison et al (2001). To date, several molecular studies have been conducted on the sable (e.g. Hosoda et al 1999; Inoue et al 2010; Malyarchuk et al 2010) including the paper by Li et al (2013) where both the phylogenetic relationships within the sable and its place of origin were explored.

Even if the stone marten is one of the most widely distributed mustelids in Eurasia (Tikhonov et al 2008), so far, no work aiming to explore its intraspecific phylogeny or evolutionary history have been conducted (Schwartz et al 2012). Investigations on the population genetics of the stone marten are also virtually absent from literature, with the exception of a country-level study (Nagai et al 2012) and a dissertation on the comparative population and landscape genetic structure of the stone marten and the red fox in Portugal (Basto 2014).

The stone marten shows different habitat preferences and a varying degree of synanthropic behaviour across its range (Proulx et al 2004; Tikhonov et al 2008). Its more generalist behaviour, compared to other members of the genus, makes the research on the genetic structure of the common stone marten particularly challenging and interesting (Virgós et al 2012). However, its relative abundance and wide distribution range enables the development of comprenhensive continental molecular studies focused on unrevealing evolutionary and ecological patterns but also the comparison with ecologically similar and closely related species within *Martes* genus.

2.3. Comparative studies on the pine and stone marten habitat use and selection

The pine marten is primarily a woodland specialist throughout its range (Proulx et al 2004) but it is not as obligately dependent on forest habitats as previously believed (Virgós et al 2012) having found to colonize agricultural landscapes with highly fragmented woods (Mergey et al 2011; Balestrieri et al 2015). In some European countries, the stone marten is well established in urbanized environments being very common in suburban and urban areas, often denning and resting in buildings and barns and causing damage to roofs, insulation and car engines (Tikhonov et al 2008; Herr et al 2009; Herr et al 2010). However, in the Iberian Peninsula the stone marten is not as synantropic as in other parts of Europe occupying well-preserved forested areas, and expressing niche displacement away from preferred pine marten habitats when co-occurring (Delibes 1983; Rodrigues et al 2011). This pattern has being explained as a product of the interspecific competition favoring the slightly bigger pine marten (Delibes 1983; Virgós and García 2002).

Under the niche-complementarity hypothesis (Schoener 1974; Larroque et al 2015) in order to coexist, two related species occupying the same macrohabitat (i.e. pine and stone martens) have to differ in the requirements of one of the three ecological dimensions (habitat use, activity time or diet). In this context, several comparative studies have been conducted on these species with different purposes such as the redefinition of their fine-scale overlapping patterns (Ruiz-González et al 2008b), interspecific variations in frequency and intensity of habitat use (Posłuszny et al 2007), on the differences in diet (Goszczyński et al 2007) and on the variation in the selection of resting sites (Larroque et al 2015).

Ruiz-González et al (2008b), for instance, used genetically confirmed non-invasively collected fecal samples to accurately record each marten presences, resulting in the delimitation of significant differences in the pine and the stone marten distribution patterns, while in the study by Pilot et al (2007) both species scats ocurred inside forest and thus, the habitat in which the sample was found was not indicative of the species. Genetically identified fecal samples were also used by Posłuszny et al (2007), in this case, for comparing the species trophic niches to find both specific and seasonal differences in their diet. In another comparative paper, Goszczyński et al (2007) found that the pine and stone marten coexistence could be largely attributed to the variation in the intensity with which each species used the three-dimensional space in forested areas and in their differing vulnerability to anthropic impact. Recently, a telemetry based study evidenced how species differed greatly in resting patterns (pine martens rested almost exclusively in forest while stone martens did in open zones in the proximity of urban areas) and thus in habitat use, which will allow them living in the same macrohabitat (i.e. syntopic species) (Larroque et al 2015). However, even if the interaction between an organism and its environment is known to occur at many scales (Wiens 1989), none of the studies addressing the differences in the pine and the stone marten habitat associations has investigated these species multiscale habitat relations (Virgós et al 2012)

3. Aims of the thesis

The main aim of this thesis is to gain insights into the intraspecific phylogeny, population and landscape genetics, and habitat suitability of several elusive European mustelids (i.e. the European otter, the pine marten and the stone marten) through the application of complementary molecular techniques.

The general and specific goals, reported below, are to:

1. Perform a genetic species identification assay for the accurate assessment of the spatial distribution of different European mustelids based on mtDNA sequencing of non-invasively collected fecal samples.

   1.1 Obtain a reliable and updated distribution and record the contemporary expansion of the endangered European otter population in the Basque Country (northern Iberian Peninsula).

   1.2 Confirm the genetic identity of the otter-like scats prior to individual genotyping.

   1.3 Unequivocally differentiate between pine and stone marten´s fecal samples obtained from a long term non-invasive genetic monitoring program (2005-2012) to build a presence-only data set for subsequent habitat use/suitability studies.

2. Individually identify each of the genetically otter scats through microsatellite and sex-genotyping for otter population monitoring.

   2.1 Track individual movements over time (i.e. during the study period, 2007-2010) and space (i.e. along the monitored rivers and wetlands).

   2.2 Determine the minimum population size (i.e. the number of unique genotypes found), fundamental to define the population status.

   2.3 Calculate the mean otter density, infer relatedness and estimate the sex-ratio of the monitored otter population.

3. Explore the multiscale relationships of habitat selection by the pine and the stone marten in sympatric areas and test if there is niche divergence in areas where the species coexist.

   3.1 Identify the environmental variables and scales influencing each species habitat selection using bivariate scaling and maximum entropy modeling.

   3.2 Compare the multiscale habitat suitability models obtained for each marten species to determine the differences in their habitat preferences.

3.3 Explore the existence of a significant niche differentiation among these species when co-occurring.

4. Conduct the first phylogeographic study of the stone marten, using mtDNA sequencing in a comprehensive number of stone marten tissue samples, along its European distribution range.

   4.1 Investigate the phylogenetic relationships of stone marten Iberian populations and its expansion within the Iberian Peninsula.

   4.2 Investigate the evolutionary and biogeographic patterns of the stone marten in Europe.

   4.4 Estimate the time of arrival and posterior lineage diversification of the species in Europe.

   4.5 Test whether, as suggested by previous authors, the European colonization of the stone marten followed the Neolithic expansion from the Near East.

5. Explore the recent population genetic structure and landscape genetics of the stone marten via microsatellite genotyping and the application of individual clustering and multivariate methods.

   5.1 Investigate the current genetic structure of the stone marten population in Iberia by seeking for genetic discontinuities in *a priori* continuous population.

   5.2 Explore the effect of landscape features in the distribution and structure of the stone marten throughout the Iberian Peninsula.

   5.3 Provide the first picture of the current population structure of stone martens in Europe.

References

Allendorf F, Luikart G (2007) Conservation and the genetics of populations. Blackwell Publishing, Malden, Massachusetts,USA.

Anderson R, Peterson A, Gomez-Laverde M (2002) Using niche-based GIS modeling to test geographic predictions of competitive exclusion and competitive release in South American pocket mice. Oikos 98:3–16.

Anderson RP, Gonzalez I (2011) Species-specific tuning increases robustness to sampling bias in models of species distributions: An implementation with Maxent. Ecol Modell 222:2796–2811.

Andrew RL, Bernatchez L, Bonin A, et al (2013) A road map for molecular ecology. Mol Ecol 22:2605–2626.

Balestrieri A, Remonti L, Ruiz-González A, et al (2015) Distribution and habitat use by pine marten *Martes martes* in a riparian corridor crossing intensively cultivated lowlands. Ecol Res 30:153–162.

Balkenhol N, Gugerli F, Cushman SA, et al (2009) Identifying future research needs in landscape genetics: where to from here? Landsc Ecol 24:455–463.

Basto MP (2014) Population and landscape genetics of the stone marten and red fox in Portugal: implications for conservation management of common carnivores. M.Sc. Thesis. University of Lisbon, Portugal.

Beebee T, Rowe G (2008) An Introduction to Molecular Ecology, 2nd edition. Oxford University Press, New York.

Beheregaray LB (2008) Twenty years of phylogeography: The state of the field and the challenges for the Southern Hemisphere. Mol Ecol 17:3754–3774.

Beja-Pereira A, Oliveira R, Alves PC, et al (2009) Advancing ecological understandings through technological transformations in noninvasive genetics. Mol Ecol Resour 9:1279–301.

Bellamy C, Scott C, Altringham J (2013) Multiscale, presence-only habitat suitability models: fine-resolution maps for eight bat species. J Appl Ecol 50:892–901.

Ben-David M, Blundell G., Kern J., et al (2005) Communication in river otters: creation of variable resource sheds for terrestrial communities. Ecology 85:1331–1345.

Bonesi L, Hale M, Macdonald DW (2012) Lessons from the use of non-invasive genetic sampling as a way to estimate Eurasian otter population size and sex ratio. Acta Theriol 58:157–168.

Brown JL (2014) SDMtoolbox: a python-based GIS toolkit for landscape genetic, biogeographic and species distribution model analyses. Methods Ecol Evol 5:694–700.

Cushman S., McGarigal K (2004) Patterns in the species-environment relationship depend on both scale and choice of response variables. Oikos 105:117–124.

Dallas JF, Coxon KE, Sykes T, et al (2003) Similar estimates of population genetic composition and sex ratio derived from carcasses and faeces of Eurasian otter *Lutra lutra*. Mol Ecol 12:275–82.

Davison A, Birks JDS, Brookes RC, et al (2001) Mitochondrial phylogeography and population history of pine martens *Martes martes* compared with polecats *Mustela putorius*. Mol Ecol 10:2479–2488.

De Salle R, Amato G (2004) The expansion of conservation genetics. Nat Rev Genet 5:702–712.

Delibes M (1983) Interspecific competition and the habitat of the stone marten *Martes foina* (Erxleben 1777) in Europe. Acta Zool Fenn 174:229–231.

Drew RE, Hallett JG, Aubry KB, et al (2003) Conservation genetics of the fisher (*Martes pennanti*) based on mitochondrial DNA sequencing. Mol Ecol 12:51–62.

Elith J, Graham CH, Anderson RP, et al (2006) Novel methods improve prediction of species' distributions from occurrence data. Ecography 29:129–151.

Elith J, Kearney M, Phillips S (2010) The art of modelling range-shifting species. Methods Ecol Evol 1:330–342.

Elith J, Phillips SJ, Hastie T, et al (2011) A statistical explanation of MaxEnt for ecologists. Divers Distrib 17:43–57.

Ferrando A, Lecis R, Domingo-Roura X, Ponsà M (2007) Genetic diversity and individual identification of reintroduced otters (*Lutra lutra*) in north-eastern Spain by DNA genotyping of spraints. Conserv Genet 9:129–139.

Ferrando A, Ponsà M, Marmi J, Domingo-Roura X (2004) Eurasian otters, *Lutra lutra*, have a dominant mtDNA haplotype from the Iberian Peninsula to Scandinavia. J Hered 95:430–435.

Fourcade Y, Engler JO, Rödder D, Secondi J (2014) Mapping species distributions with MAXENT using a geographically biased sample of presence data: A performance assessment of methods for correcting sampling bias. PLoS One 9:1–13.

Frankel O, Soulé M (1981) Conservation and evolution. Cambridge University Press, Cambridge, UK.

Frankham R, Briscoe D, Ballou J (2002) Introduction to Conservation Genetics. 617.

Franklin J (1995) Predictive vegetation mapping: geographic modelling of biospatial patterns in relation to environmental gradients. Prog Phys Geogr 19:474–499.

Freeland J., Kirk H, Petersen S. (2011) Molecular Ecology, 2nd edition. John Wiley & Sons, Ltd. New York.

Gaston KJ, Fuller R a. (2008) Commonness, population depletion and conservation biology. Trends Ecol Evol 23:14–19.

Girvetz EH, Greco SE (2009) Multi-scale predictive habitat suitability modeling based on hierarchically delineated patches: An example for yellow-billed cuckoos nesting in riparian forests, California, USA. Landsc Ecol 24:1315–1329.

Goszczyński J, Posłuszny M, Pilot M, Gralak B (2007) Patterns of winter locomotion and foraging in two sympatric marten species: *Martes martes* and *Martes foina*. Can J Zool 85:239–249.

Graf RF, Bollmann K, Suter W, Bugmann H (2005) The importance of spatial scale in habitat models: Capercaillie in the Swiss Alps. Landsc Ecol 20:703–717.

Guisan A, Zimmerman NE (2000) Predictive habitat distribution models in ecology. Ecol Modell 135:147–186.

Hall LA, Beissinger SR (2014) A practical toolbox for design and analysis of landscape genetics studies. Landsc Ecol 29:1487–1504.

Harris H (1966) Enzyme polymorphisms in man. Proc R Soc London Ser B - Biol Sci 164:298–310.

Herr J, Schley L, Engel E, Roper TJ (2010) Den preferences and denning behaviour in urban stone martens (*Martes foina*). Mamm Biol.75:138–145.

Herr J, Schley L, Roper TJ (2009) Socio-spatial organization of urban stone martens. J Zool 277:54–62.

Hewitt GM (2004) The structure of biodiversity – insights from molecular phylogeography. Front Zool 1:4.

Hobbs GI, Chadwick EA, Slater FM, Bruford MIW (2006) Landscape Genetics Applied To a Recovering Otter (Lutra Lutra) Population in the Uk: Preliminary Results and Potential Methodologies. 17:47–63.

Hobbs GI, Chadwick EA, Bruford MW, Slater FM (2011) Bayesian clustering techniques and progressive partitioning to identify population structuring within a recovering otter population in the UK. J Appl Ecol 48:1206–1217.

Holderegger R, Wagner HH (2008) Landscape Genetics. Bioscience 58:199–207.

Hosoda T, Suzuki H, Iwata M., et al (1999) Genetic relationship within and between the Japanese marten *Martes melampus* and the sable *M. zibellina*, based on variation of mitochondrial DNA and nuclear ribosomal DNA. Mammal Study 24:25–33.

Hussain S., Kanchanasakha B, de Silva P., Olson A (2008) Lutra sumatrana. The IUCN Red List of Threatened Species. Version 2014.3. <www.iucnredlist.org>. Downloaded on 26 March 2015.

Inoue T, Murakami T, Abramov A., Masuda R (2010) Mitochondrial DNA control region variations in the sable *Martes zibellina* of Hokkaido Island and the Eurasian continent, compared with the Japanese marten *M. melampus*. Mammal Study 35:145–155.

Jiménez-Valverde A, Peterson A., Soberón J, et al (2011) Use of niche models in invasive species risk assessments. Biol Invasions 13:2785–2797.

Johnson D (1980) The comparison of usage and availability measurements for evaluating resource preference. Ecology 61:65–71.

Kalz B, Jewgenow K, Fickel J (2006) Structure of an otter (*Lutra lutra*) population in Germany – results of DNA and hormone analyses from faecal samples. Mamm Biol. 71:321–335.

Khanum R, Mumtaz a. S, Kumar S (2013) Predicting impacts of climate change on medicinal asclepiads of Pakistan using Maxent modeling. Acta Oecologica 49:23–31.

Koelewijn HP, Pérez-Haro M, Jansman H a. H, et al (2010) The reintroduction of the Eurasian otter (*Lutra lutra*) into the Netherlands: hidden life revealed by noninvasive genetic monitoring. Conserv Genet 11:601–614.

Koepfli KP, Wayne R (1998) Phylogenetic relationships of otters (Carnivora: Mustelidae) based on mitochondrial cytochrome b sequences. J Zool 246:401–416.

Koepfli KP, Deere KA, Slater GJ, et al (2008) Multigene phylogeny of the Mustelidae: resolving relationships, tempo and biogeographic history of a mammalian adaptive radiation. BMC Biol 6:22.

Kramer-Schadt S, Niedballa J, Pilgrim JD, et al (2013) The importance of correcting for sampling bias in MaxEnt species distribution models. Divers Distrib 19:1366–1379.

Kruuk H (1995) Wild Otters Predation and Populations. Oxford University Press, Oxford

Kruuk H (2006) Otters. Ecology Behavior and Conservation. Oxford University Press, Oxford

Kyle CJ, Davison A, Strobeck C (2003) Genetic structure of European pine martens (*Martes martes*), and evidence for introgression with *M. americana* in England. Conserv Genet 4:179–188.

Larroque J, Ruette S, Vandel J-M, Devillard S (2015) Where to sleep in a rural landscape? A comparative study of resting sites pattern in two syntopic Martes species. Ecography 38:1-12.

Li B, Malyarchuk B, Ma Z, et al (2013) Phylogeography of sable (*Martes zibellina* L. 1758) in the southeast portion of its range based on mitochondrial DNA variation: Highlighting the evolutionary history of the sable. Acta Theriol (Warsz) 58:139–148.

Lindenmayer DB, Wood JT, McBurney L, et al (2011) How to make a common species rare: A case against conservation complacency. Biol Conserv 144:1663–1672.

Malyarchuk B., Petrovskaya A., Derenko M. (2010) Intraspecific structure of sable Martes zibellina L. inferred from nucleotide variation of the mitochondrial DNA cytochrome b gene. Russ J Genet 46:64–68.

Manel S, Holderegger R (2013) Ten years of landscape genetics. Trends Ecol Evol 28:614–621.

Manel S, Schwartz MK, Luikart G, Taberlet P (2003) Landscape genetics: combining landscape ecology and population genetics. Trends Ecol Evol 18:189–197.

Mateo-Sánchez M, Cushman S., Saura S (2013) Scale dependence in habitat selection: The case of the endangered brown bear (*Ursus arctos*) in the Cantabrian Range (NW Spain). Int J Geogr Inf Sci 1–16.

McCormack JE, Zellmer AJ, Knowles LL (2010) Does niche divergence accompany allopatric divergence in *Aphelocoma jays* as predicted under ecological speciation?: Insights from tests with niche models. Evolution (NY) 64:1231–1244.

McDonald RA, Hara KO, Morrish DJ (2007) Decline of invasive alien mink (*Mustela vison*) is concurrent with recovery of native otters (*Lutra lutra*). Divers Distrib 92–98.

Mergey M (2007) Mergey M (2007). Réponses des populations de martres d'Europe (*Martes martes*) à la fragmentation de l'habitat: mécanismes comportementaux et consequences. University of Reims Campagne-Ardenne, France.

Mergey M, Helder R, Roeder JJ (2011) Effect of forest fragmentation on space use patterns in the European pine marten (*Martes martes*). J Mammal 92:328–335.

Mergey M, Larroque J, Ruette S, et al (2012) Linking habitat characteristics with genetic diversity of the European pine marten (*Martes martes*) in France. Eur J Wildl Res 909–922.

Mucci N, Arrendal J, Ansorge H, et al (2010) Genetic diversity and landscape genetic structure of otter (*Lutra lutra*) populations in Europe. Conserv Genet 11:583–599.

Nagai T, Raichev E, Tsunoda H (2012) Preliminary study on microsatellite and mitochondrial DNA variation of the stone marten *Martes foina* in Bulgaria. Mammal Study 358:353–358.

Noos R (1990) Can we maintain biological and ecological integrity. Conserv Biol 4:241–243.

Ouborg NJ, Pertoldi C, Loeschcke V, et al (2010) Conservation genetics in transition to conservation genomics. Trends Genet 26:177–87.

Pertoldi C, Barker S, Madsen A, et al (2008) Spatio-temporal population genetics of the Danish pine marten (*Martes martes*). Biol J Linn Soc 93:457–464.

Pertoldi C, Bijlsma R, Loeschcke V (2007) Conservation genetics in a globally changing environment: Present problems, paradoxes and future challenges. Biodivers Conserv 16:4147–4163.

Pertoldi C, Elschot K, Ruiz-Gonzalez A, et al (2014) Genetic variability of central–western European pine marten (*Martes martes*) populations. Acta Theriol 59:503–510.

Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species geographic distributions. Ecol Modell 190:231–259.

Pilot M, Gralak B, Goszczyński J, Posłuszny M (2007) A method of genetic identification of pine marten (*Martes martes*) and stone marten (*Martes foina*) and its application to faecal samples. J Zool 271:140–147.

Posłuszny M, Pilot M, Goszczyński J, Gralak B (2007) Diet of sympatric pine marten (*Martes martes*) and stone marten (*Martes foina*) identified by genotyping of DNA from faeces. Ann Zool Fennici 269–284.

Prigioni C, Balestrieri A, Remonti L (2007) Decline and recovery in otter Lutra lutra populations in Italy. Mamm Rev 37:71–79.

Prigioni CL, Remonti LU, Balestrieri AL (2006) Otter (*Lutra lutra*) movements assessed by genotyped spraints in southern Italy. Hystrix It J Mamm 17:91–96.

Proulx G, Aubry K, Birks J, et al (2004) World distribution and status of the genus martes in 2000. In: Harrison D, Fuller A, Proulx G (eds) Martins and fishers (Martes) in human-altered environments: an interna- tional perspective. Springer, New York, pp 21–76.

Randi E, Davoli F, Pierpaoli M, et al (2003) Genetic structure in otter (*Lutra lutra*) populations in Europe: implications for conservation. Anim Conserv 6:93–100.

Reutter B., Helfer V, Hirzel A., Vogel P (2003) Modelling habitat suitability using museum collections: an example with three sympatric Apodemus species from the Alps. J Biogeogr 30:581–590.

Rodrigues M, Santos-Reis M, Elmeros M, Fernandes C (2011) Microsatellite markers for genetic studies in the weasel (*Mustela nivalis*). Eur J Wildl Res 58:507–510.

Romanowski J, Brzeziński M, Zmihorski M (2013) Habitat correlates of the Eurasian otter *Lutra lutra* recolonizing Central Poland. Acta Theriol (Warsz) 58:149–155.

Ruiz-González A, Ferrando A, Pérez-Haro M, et al (2008a) Seguimientos moleculares de las poblaciones de nutria *Lutra lutra* (Linnaeus, 1758) a partir de muestreos no invasivos. La Nutr. en España. Veinteaños Seguim. un mamífero amenazado. SECEM, Málaga, pp 207–215

Ruiz-González A, Gurrutxaga M, Cushman SA, et al (2014) Landscape Genetics for the Empirical Assessment of Resistance Surfaces: The European Pine Marten (*Martes martes)* as a Target-Species of a Regional Ecological Network. PLoS One 9:19.

Ruiz-González A, Madeira MJ, Randi E, et al (2013a) Non-invasive genetic sampling of sympatric marten species (*Martes martes* and *Martes foina*): assessing species and individual identification success rates on faecal DNA genotyping. Eur J Wildl Res 59:371–386.

Ruiz-González A, Madeira MJ, Randi E, et al (2013b) Phylogeography of the forest-dwelling European pine marten (*Martes martes*): new insights into cryptic northern glacial refugia. Biol J Linn Soc 109:1–18.

Ruiz-González A, Rubines J, Berdión O, Gómez-Moliner BJ (2008b) A non-invasive genetic method to identify the sympatric mustelids pine marten (*Martes martes*) and stone marten (*Martes foina*): preliminary distribution survey on the northern Iberian Peninsula. Eur J Wildl Res 54:253–261.

Ruiz-Olmo J, Loy A, Cianfrani C, et al (2008) *Lutra lutra*. The IUCN Red List of Threatened Species. Version 2014.3. <www.iucnredlist.org>. Downloaded on 03 April 2015.

Sato J, Yasuda S, Hosoda T (2009) Genetic diversity of the Japanese Marten (*Martes melampus*) and its implications for the conservation unit. Zool Sci 26:457–466.

Schluter D (2000) The Ecology of Adaptive Radiation. Oxford University Press, Oxford

Schoener TW (1974) Resource partitioning in ecological communities. Science (80) 185:27–39.

Schwartz MK, Monfort SL (2008) Genetic and Endocrine Tools for Carnivore Surveys. In: Long R, MacKay P, Ray J, Zielinski W, editors. Noninvasive survey methods for North American carnivores. Washington D.C.: Island Press. pp. 228–250.

Schwartz M, Ruiz-González A, Pertoldi C, Masuda R (2012) Martes conservation genetics: assessing within-species movements, units to conserve and connectivity cross ecological and evolutionary time. In: Aubry KB, Zielinski WJ, Raphael MG, Proulx G, Buskirk SW (eds) Biology and conservation of marten, sables, and fisher: a new synthesis. Cornell Univesity Press, New York

Schwartz MK, Luikart G, Waples RS (2007) Genetic monitoring as a promising tool for conservation and management. Trends Ecol Evol 22:25–33.

Shirk AJ, Raphael MG, Cushman SA (2014) Spatiotemporal variation in resource selection: insights from the American marten (*Martes americana*). Ecol Appl 24:1434–1444.

Sork VL, Waits L (2010) Contributions of landscape genetics - approaches, insights, and future potential. Mol Ecol 19:3489–3495.

Soulé M, Wilcox B (1980) Conservation Biology: an Evolutionary–Ecological Perspective. Sinauer, Sunderland, MA.

Storfer A, Murphy M A, Spear SF, et al (2010) Landscape genetics: Where are we now? Mol Ecol 19:3496–3514.

Sulkava R (2007) Snow tracking: a relevant method for estimating otter *Lutra lutra* populations. Wildlife Biol 13:208–218.

Syfert MM, Smith MJ, Coomes DA (2013) The Effects of Sampling Bias and Model Complexity on the Predictive Performance of MaxEnt Species Distribution Models. PLoS One. doi: 10.1371/journal.pone.0055158

Taberlet P, Luikart G (1999) Non-invasive genetic sampling and individual identification. Biol J Linn Soc 68:41–55.

The IUCN Red List of Threatened Species. Version 2014.3. <www.iucnredlist.org>. Downloaded on 26 March 2015.

Tikhonov A, Cavallini P, Maran T, et al (2008) *Martes foina*. The IUCN Red List of Threatened Species. Version 2014.3.

Van Zyll de Jong C. (1987) A phylogenetic study of the Lutrinae (Carnivora; Mustelidae) using morphological data. Can J Zool 65:2536–2544.

Virgós E, Zalewski A, Rosalino L., Mergey M (2012) Habitat ecology of genus Martes in Europe: a review of the evidences. In: Aubry KB, Zielinski WJ, Raphael MG, Proulx G, Buskirk SW (eds) Biology and conservation of marten, sables, and fisher: a new synthesis. Cornell Univesity Press, New York, pp 255–266

Wagner H, Fortin M (2013) A conceptual framework for the spatial analysis of landscape genetic data. Conserv Genet 14:253–261.

Waits LP, Paetkau D (2005) Noninvasive genetic sampling tools for wildlife biologists: a review of applications and recommendations for accurate data collection. J Wildl Manage 69:1419– 1433.

Wang IJ (2010) Recognizing the temporal distinctions between landscape genetics and phylogeography. Mol Ecol 19:2605–2608.

Warren DL, Glor RE, Turelli M (2010) ENMTools: a toolbox for comparative studies of environmental niche models. Ecography 1:607–611.

Warren DL, Seifert SN (2011) Ecological niche modeling in Maxent: the importance of model complexity and the performance of model selection criteria. Ecol Appl 21:335–342.

Wasserman TN, Cushman SA, Wallin DO, Hayden J (2012) Multi scale habitat relationships of *Martes americana* in northern Idaho, U.S.A. Environmental Science. Paper 20. http://cedar.wwu.edu/esci_facpubs/20

Wayne R., Morin P. (2004) Conservation genetics in the new molecular age. Front Ecol Environ 2:89–97.

Wellenreuther M, Larson KW, Svensson EI (2012) Climatic niche divergence or conservatism? Environmental niches and range limits in ecologically similar damselflies. Ecology 93:1353–1366.

Wiens J. (1989) Spatial scaling in Ecology. Funct Ecol 385–397.

Wright L, Olsson A, Kanchanasaka B (2008) A Working Review of the Hairy-Nosed Otter (*Lutra sumatrana*). IUCN Otter Specialist Group Bulletin 25:38–59.

Chapter II

PAPER 1: Individual identification and distribution assessment of otters (*Lutra lutra*) through non-invasive genetic sampling: Recovery of an endangered species in the Basque Country (Northern Spain)

María Vergara, Aritz Ruiz-González, Javier López de Luzuriaga and Benjamín J. Gomez-Moliner

2014. Mammalian Biology 79: 259–267

Abstract

Non-invasive genetic techniques have proven to be cost-effective for monitoring and studying otter populations, largely due to the linearity of otter territories and the marking behavior of the species. After a severe decline, in the 60, the Eurasian otter is recovering in the Iberian Peninsula. However, the recovery pattern is not homogeneous and the species is still considered "Threatened" in many regions. During 2007–2010 a systematic non-invasive genetic sampling effort was carried out to determine the spatial distribution and to estimate the population size of an endangered otter population in northern Iberian Peninsula (Basque Country). Samples were identified to species level by sequencing a 226 bp mtDNA fragment prior to genotyping. Among the 132 obtained samples, 127 (98.4%) belonged to the study species, one sample was genetically identified as European mink (*Mustela lutreola*) and one as American mink (*Neovison vison*) while genetic species confirmation was not possible in the three remaining samples. These results provided novel and accurate data on the species distribution, highlighting an overall increase of 25% in 10 × 10 UTM grids occupied by otter and a clear pattern of re-colonization upstream of the main rivers. All samples corresponding to otter were subsequently individually genotyped using a novel multiplex panel of 11 microsatellite markers and sexed by typing the sex-chromosome-related gene ZFX/ZFY. We obtained a complete individual genetic profile for 55 samples (genotyping success 43%), corresponding to 20 different individuals (11 females, 6 males, and 3 individuals of unknown gender). The mean otter density in occupied areas estimated to be 0.09 (0.06–0.12) individuals per river kilometer. The present study enabled us to obtain updated and relevant information about this elusive species' distribution and population size, essential to define population status and to design successful and effective management and conservation programs.

Keywords: *Lutra lutra*, non-invasive genetic sampling, molecular tagging, molecular sexing.

Introduction

Molecular methods incorporating non-invasive sampling (NGS) via the collection of spraints are currently widely used for population monitoring of elusive carnivores (Waits and Paetkau, 2005; Schwartz and Monfort, 2008; Beja-Pereira et al., 2009; Canigliaet al., 2014). The Eurasian otter (*Lutra lutra*), a semi-aquatic carnivore, mostly crepuscular, rare and elusive, is a suitable species for non-invasive genetic monitoring, due to the typical marking behavior used for intra-specific communication (Chanin, 1985; Gorman and Trowbridge, 1989; Kruuk, 1992, 1995; Ben-David et al., 2005). DNA-based assays have become reliable tools for species identification, molecular sex typing and individual identification through microsatellite genotyping using spraint DNA (e.g. Dallaset al., 2000; Prigioni et al., 2006; Mucci and Randi, 2007; Hájkováet al., 2009; O'Neill et al., 2013). The combined application of these methods on spraints is a cost-effective way to investigate spatial distribution patterns, as well as dispersal, population turnover, sex-ratio or kinship for otters in the wild, which enables researchers to identify and resolve information gaps and design effective research and management programs (Dallas et al., 2003; Hung et al., 2004; Kalz et al., 2006; Koelewijn et al., 2010) and to evaluate the success of reintroduction and traslocation projects (Ferrando et al., 2008; Koelewijn et al., 2010).

The Eurasian otter is currently a species of conservation concern and has one of the widest distributions of all Palaearctic mammals being present in most of Europe, Asia and North Africa (Ruiz-Olmo et al., 2008). The Eurasian otter populations declined in large areas of Europe during the twentieth century (Macdonald and Mason, 1983, 1994; Lodé, 1993; Kauhala, 1996). However, substantial increases (Sulkava, 2007) and recolonization events are recently been observed in several European countries (Kranz et al., 2001; Mason and Macdonald, 2004; McDonald et al., 2007; Prigioni et al., 2007). Currently, after a strong decline in the 1980s the Eurasian otter is recovering in many European countries and it has been recently downgraded to the "Near Threatened" category in the IUCN Red List (Ruiz-Olmo et al., 2008).

In Spain, situated on the western edge of the species's distribution range, otter populations have progressively declined since the mid-twentieth-century and particularly since 1960, as in many European countries. The first accurate assessment of the species distribution based on the Otter Surveys Methodology dates from 1985 (Delibes, 1990). As revealed by the first national otter sur-vey (NOS), only a third of the sampled sites were positive for otter presence (Delibes, 1990). Ten years later, in 1996, during the second NOS, the species was present

in almost half the country (Ruiz-Olmo and Delibes, 1998). The last NOS in 2006 showed a marked increase in the species' range probably due to recolonization from surrounding areas and was recorded in the 70% of the territory (López-Martín and Jiménez, 2008). Due to this global recovery pattern, the otter has recently been classified as "Least Concern" according to the IUCN Red List criteria at national level (Palomo and Gisbert, 2002). Nevertheless, it is still absent or rare in some Spanish regions, such as the Basque Country, where it is officially listed as "Endangered" (Decree 167/1996).

Similarly to the global patterns observed in other otter populations in the Iberian Peninsula, also in this region the species suffered a strong and generalized decrease until the 1990s: it disappeared from the whole Atlantic basin, and although the species seems to have partially recovered in the Mediterranean basin in the last decade, its presence is still scarce and potentially existing at very low density (Hernando et al., 2005; López de Luzuriaga, 2008). Fulfilling legal requirements associated with the "Endangered" status of the species (Law 16/1994; Decree 167/1996), the Otter Management Plan for Álava province (OMPA) – the only among the three Basque provinces with a permanent otter population – was finally approved in 2004 (Foral Order 880/2004) to actively promote species recovery and protection. Among the main actions outlined in the OMPA, this plan established the identification of the Conservation Concern Areas for the otter (CCAO, 7 of which are considered Sites of Community Importance, SCI) and also the need of implementing an otter monitoring program to improve knowledge on the biology of the species in those aspects that enable a more effective short-term management (i.e. species distribution and population size). Understanding the underlying ecology of recolonization is fundamental for otter management and conservation (Remonti et al., 2008; López-Martín and Jiménez, 2008). For several recovering populations throughout the otter's European range, and especially in the Iberian Peninsula (López-Martín and Jiménez, 2008), accurate data about presence and abundance are still lacking.

Here, we present the results from a systematic non-invasive genetic survey that aimed to determine the current distribution and population size of the low density and endangered otter population in northern Iberian Peninsula derived from the actions outlined in the official management plan. We also assessed distributional change over the last decade by comparisons with historical datasets (López-Martín and Jiménez, 2008), aiming to evaluate the recovery pattern in this portion of the Iberian Peninsula. Moreover, we provide the first population abundance estimates for the species by the use of a novel multiplex microsatellite panel and a DNA sex typing approach. Results of this study provide information regarding

the distribution of an expanding otter population in northern Spain, help to refine otter sign survey techniques by the use of non-invasive genetic sampling, and provide useful information for otter management activities within the OMPA framework.

Material and methods

Study area

The study area comprises the Álava Province, which is located in northern Spain (Iberian Peninsula, Southwest of Europe), being the largest of the three Provinces of the Basque Country and the only one with a permanent otter population during the last 2 decades (Hernando et al., 2005). It covers an area of 3,037 km2 with altitudes ranging from 376 m to 1,482 m. Álava's hydrological network consist of several large rivers and their tributaries, with important variations in biological conditions. The Zadorra River, a tributary of the Ebro River is the longest river in the study area with a total length of 78 km. The surrounding areas of Álava in the provinces of Burgos and La Rioja and the Enclave of Treviño (Castilla y León) were also included in the survey (Fig. 1).

Non-invasive sample collection and DNA extraction

As the European otter is an elusive and rare species (Kruuk, 2006), we obtained DNA from non-invasive sampling of feces. This DNA source is regularly used for non-invasive genetic studies of carnivores (Schwartz and Monfort, 2008; Beja-Pereira et al., 2009) and has been successfully used in previous otter studies (Kalz et al.,2006; Prigioni et al., 2006; Hájková et al., 2009; Koelewijn et al., 2010). We implemented a multistage sampling scheme in which samples from a pilot study were used to verify the suitability of the field and laboratory methods for individual identification and to assess the preliminary otter distribution. Thus, two spraint-based surveys were conducted between 2007 and 2010. The first one, conducted in the 2007–2008 period was used to preliminarily estimate the distribution range of the otter in the study area. The second one, conducted between 2009 and 2010 was used to refine the species distribution assessment and to obtain a higher number of fresh samples of *Lutra lutra* for individual identification purposes.

The study area was divided in 4 main zones (A to D; West-Álava, Central Álava, South Álava and Enclave of Treviño and the Álava Mountain; Fig. 2) that matched with the 9 CCAO, previously considered suitable for the mustelid. For all of the four study areas, a total of 198 suitable survey sites were identified (based on previous surveys, literature and species' observations) and coded as 0 (survey sites with no otter evidences), 1a (positive survey sites for otter based on field signs) or 1b (positive sites with genetically confirmed fecal samples) for each of the visits. These sites were repeatedly surveyed a minimum of 3 times to increase

the probability of detection and thus, decreasing the probability of false negatives during the whole study period. The sampling was carried out systematically along water bodies such as rivers, streams and wetlands. All surveys were conducted by the same experienced surveyor, thus reducing the observer error in the field (Parry et al., 2013; Ruiz-González et al., 2013).

The sample collection was carried out in cold months during dry weather (October–April) during the morning hours in order to collect fresh samples (<24 h after defecation) as DNA in spraints degrades faster when exposed to warm temperatures (Hájková et al., 2006). Universal Transversal Mercator (UTM) coordinates were recorded for all samples collected using a global positioning system (Garmin eTtrex). Fecal samples, with two replicates when possible, were stored in sterile 1.5 ml tubes containing 96% vol. ethanol and frozen at −20°C until processed (Gómez-Moliner et al., 2004; Ruiz-González et al., 2008a). DNA was then extracted using the DNA Stool MiniKit (Qiagen, Hombrechtikon, Switzerland) following the manufacturer's protocol for DNA extraction from stool samples. Additionally, fresh tissues from otter carcasses were included in the study. DNA was isolated from tissues using the Qiagen DNeasy Tissue DNA extraction kit (Qiagen, Hombrechtikon, Switzerland) according to the manufacturer's instructions.

Species identification and spatial distribution

In order to reliably verify that the visually identified samples correspond to otter we conducted a species identification step consisting of a touchdown PCR amplification of the mitochondrial D-loop region (226 bp). A 2 uL volume of the DNA extraction mixture was added to 23 uL of the PCR mixture containing 0.5 uL of forward primer Ma-2F (Ruiz-González et al., 2008a) and 0.5 uL of reverse primer CR-R (Ruiz-González et al., 2008a) (20 pmol/uL), 2.5 uL 10× reaction buffer, 0.75 uL MgCl2 (50 mM), 1.4 uL deoxynucleotide triphosphates (2.5 mM), 1 uL of bovine serum albumin (10 mg/mL), 16.15 uL of sterile water and 0.2 uL Taq polymerase (5 U/uL), adapted from the protocol used by Ruiz-González et al. (2008) for martens and closely related mustelid species.

After incubation for 5 min at 95°C, samples were subjected to eight cycles of a first denaturation step at 95°C for 1 min, 40 s at 72°C (reducing the temperature by 0.5°C per cycle) and 45 s at 72°C to enhance the enrichment of the specific product over any non-specific one. 40 cycles consisted of denaturation for 1 min at 95°C, annealing for 40 s at 63.5°C, 45 s at 72°C and a final extension step at 72°C for 10 min were performed in a BIO-RAD iCycler (Version 3.021, BIO-RAD Laboratories). 4 uL of the amplified product were run in

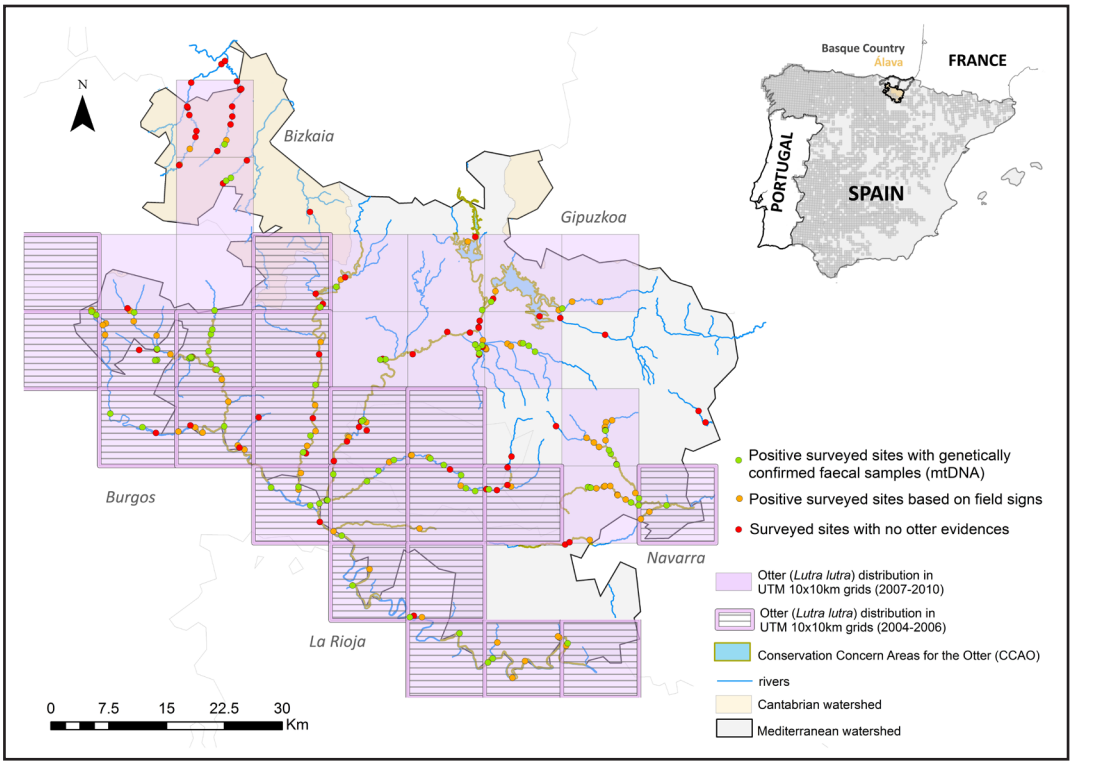


Figure 1. Otter (*Lutra lutra*) distribution in Álava and surroundings in 10x10 UTM km grids for the period 2004-2006 (III. NOS, López-Martín and Jiménez, 2008) and 2007-2010 (present study).

2% agarose gels stained with GelRed, to visualize successful amplifications. Amplicons were purified using EXO-SAP IT (USB, Cleveland, OH, USA) and sequenced using the BigDye® Terminator Kit v1.1 cycle sequencing kit (Applied Biosystems Foster City, USA) according to the manufacturer's instructions. Results were analyzed on an ABI 3130XL automated sequencer (Applied Biosys-tems). Negative controls were used to check for contamination.

To assess the spatial distribution of the otter in the study area the UTM coordinates corresponding to each sample were projected onto a GIS (Arcview 9.0 ESRI) along with the species identification data. Subsequently, the results were displayed in 10 × 10 km UTM grids to infer the species distribution during the whole survey (2007–2010). These results were compared with previous systematic field otter surveys conducted from 2004 to 2006 period in the framework of the III. NOS (López de Luzuriaga et al., 2008; López-Martín and Jiménez, 2008) to assess changes in otter distribution.

Development of a microsatellite multiplex protocol for Lutra lutra *useful for fecal DNA genotyping*

19 microsatellite loci identified in the genomic DNA of the otter (Dallas and Piertney, 1998: Lut733, Lut782, Lut475 and Lut717, Lut833, Lut832, Lut701, Lut615, Lut818, Lut715, Lut453, Lut435,Lut604 and Lut902; Huang et al., 2005, OT-04, OT-07, OT-14, OT-17and OT-19) and used in previous otter genetic studies (e.g. Ferrando et al., 2008; Ruiz-González et al., 2008b; Hájková et al., 2009; Bonesi et al., 2013) were initially considered for designing a novel multiplex panel suitable for non-invasive genetic sampling of our target population.

PCR conditions were firstly optimized for each primer separately and the final annealing temperature per locus was selected based on the results of a temperature gradient PCR run on tissue-derived DNA. Single locus amplifications were performed in a total volume of 15 uL with 2 uL DNA, 0.2 uL of each primers (20 pmol/uL), 1.7 uL 10× reaction buffer, 0.64 uL MgCl2 (50 mM), 0.5 uL dNTP (2.5 mM), 0.1 uL of BSA (10 mg/mL), 9.56 uL of sterile water and 0.1 uL Taq polymerase (5 U/uL). The following PCR conditions were used for all amplifications: after incubation for 5 min at 94°C, the samples were subjected to 40 amplification cycles in a BIO-RAD iCycler (Version 3.021, BIO-RAD Laboratories) consisting of denaturation for 1 min at 94°C, annealing for 1 min at 54–59°C and a final extension stage of 1 min at 72°C.

Loci with a low rate of amplification and with a high quantity of unspecific amplification were discarded. This first screening was evaluated with the only existing 2 tissues of the species obtained during the last 10 years, and 20 fecal otter samples from the study area (unreported data). From the 19 screened loci we selected a suite of 11 markers to design a Multiplex protocol with special emphasis on demonstrated variability in otter samples from the study region and amplification strength with small quantities of DNA (fecal DNA) and powerful enough to allow individual identification.

Microsatellite genotyping

All samples corresponding to otter were individually genotyped using a novel multiplex panel of 11 autosomal microsatellite markers: Lut435, Lut453, Lut701, Lut715, Lut818, Lut833, Lut902 (Dallas and Piertney, 1998; Dallas et al., 1999), OT-04, OT-14, OT-17 and OT-19 (Huang et al., 2005).

In addition to the negative controls for extraction, negative PCR controls were included as

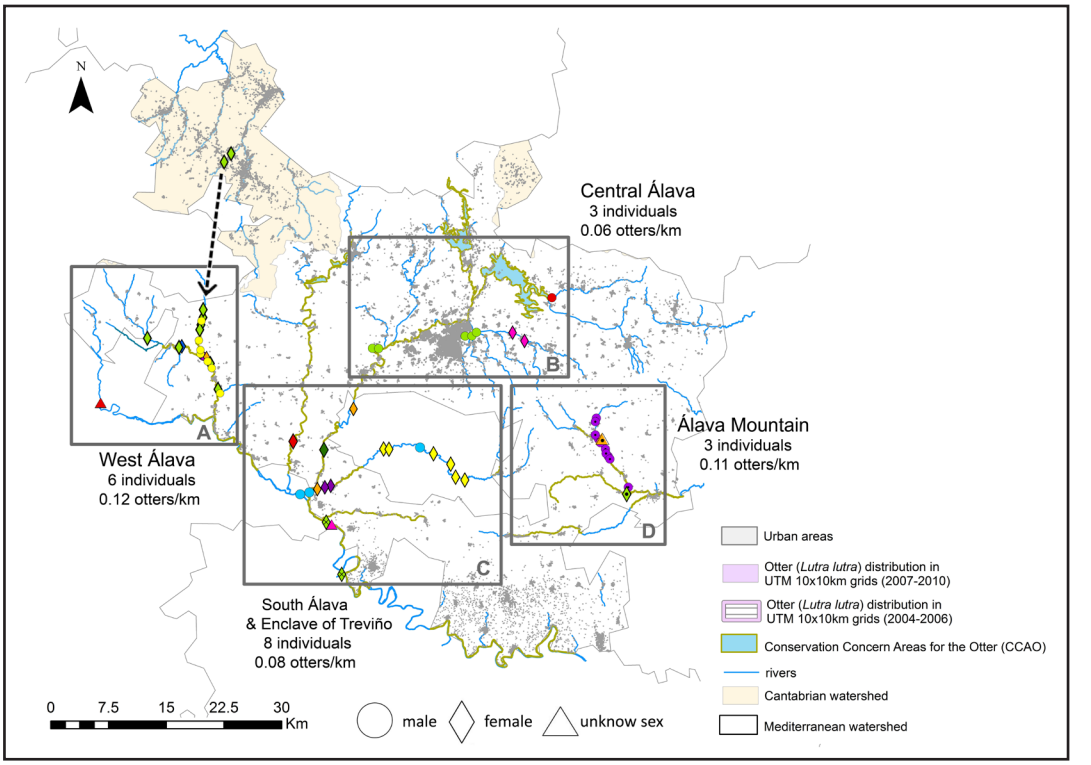


Figure 2. Spatial distribution of the identified individuals with a full multilocus genotype and mean otter density (ind/km river) per zone (A, B, C and D).

proposed by Pompanon et al. (2005). Additionally, we amplified a reference sample as a positive control and test if the electrophoretic mobility of the fragments was consistent across runs (Davison and Chiba, 2003).

In non-invasive genetic sampling, genotyping errors occur due to increased rates of null alleles, allelic drop-out (ADO) and false alleles (FA) (Taberlet and Luikart, 1999; Pompanon et al., 2005). Therefore, we followed a multiple-tube approach from Taberlet and Luikart (1999), amplifying each DNA extract in four independent replicates and in separate rooms dedicated to low DNA content samples.

The primers labeled with the dyes 6-FAM, NED, PET and VIC, were amplified in 4 PCR multiplex reactions (multiplex A, B, C and D, Table 1). PCR multiplex amplifications were carried out with QIA-GEN Multiplex PCR kits using the manufacturer's protocol in a total volume of 10 uL with 2 uL of DNA, 1xMultiplex Kit Master Mix and 2 pmol of each fluorescent labeled forward and unlabelled reverse primer. We applied a hotstart thermocycling protocol. The initial polymerase activation (HotStart PCR) was done at 95ºC for 15 min,

followed by 42 cycles (35 for tissue samples) of denaturation at 94ºC for 30 s, primer annealing at 57ºC for 130 s, and sequence extension at 72ºC for 1 min, and a final extension step at 60ºC for 30 min. DNA quality was initially screened by PCR-amplifying each DNA sample four times at three loci (Multiplex C: LUT453, LUT818 and OT-19). Only positive samples (samples showing >50% positive PCRs) were further amplified four times at the remaining 8 loci. Samples with inconsistent results after four amplifications per locus or with ≤50% successful amplifications across loci were removed from further analysis as they were not considered reliable genotypes.

Multiplex PCR products were run on an ABI (Foster City, CA) 3130XL automated sequencer (Applied Biosystems), with the internal size standard GS500 LIZ™ (Applied Biosystems). Allele sizing was conducted using the ABI software GENEMAPPER 4.0.

Genotype checking, probability of identity and individual identification

Consensus genotypes from four replicates of fecal DNA samples were reconstructed using GIMLET v 1.3.4 (Valière, 2002), accepting heterozygotes if the two alleles were seen at least in two replicates and homozygotes if a single allele was seen at least in three replicates. Afterwards, we grouped identical consensus genotypes using the same software, in order to identify the number of different individuals in our data set. GIMLET was also used to estimate PCR success and genotyping errors: ADO (when a heterozygote from the consensus genotype is typed as a homozygote from repeated genotypes) and FA (when a homozygote from the consensus genotype is typed as a heterozygote) (Pompanon et al., 2005). To test the discrimination power of the 11 microsatellites set, we computed the probability of pairs of individuals bearing an identical multilocus genotype [i.e. probability of identity (PID)] with the software GIMLET v 1.3.2 (Valière, 2002). PID calculations were performed with both the unbiased equation for small sample size and the equation for siblings. The more conservative PID for full-sibs (PID-Sib) was estimated as an upper limit to the probability that pairs of individuals would share the same genotype. GIMLET was also used to estimate PCR success and errors (ADO and FA). Once samples were correctly genotyped and assigned to a known individual, the location and date of the sampling were projected into a GIS (Arcview9.0 ESRI) to evaluate the plausibility of the sample belonging to the same individual.

Sex Identification

Samples were sexed by typing a fragment of the Zink-finger protein genes ZFX/ZFY that can be amplified simultaneously from both male and female chromosomes using the same

MULTIPLEX	LOCUS	DYE	SIZE RANGE	NA	He	Ho	PCR+	ADO	FA
Multiplex A	LUT701	6-FAM	202-210	3	0.6	0.63	0.94	0.218	0.000
	LUT833	NED	154-162	3	0.27	0.32	0.94	0.042	0.000
	Sex	VIC	121-180	2	-	-	-	-	-
Multiplex B	LUT435	NED	119-139	4	0.57	0.63	0.94	0.104	0.013
	LUT715	VIC	200-212	4	0.58	0.63	0.97	0.180	0.033
	OT-17	6-FAM	155-179	3	0.39	0.32	0.96	0.125	0.000
Multiplex C	LUT453	6-FAM	121-131	4	0.52	0.58	0.96	0.147	0.026
	LUT818	PET	160-180	2	0.19	0.21	0.9	0.000	0.010
	OT-19	VIC	218-234	5	0.68	0.53	0.89	0.196	0.061
Multiplex D	LUT902	PET	148-156	3	0.43	0.21	0.97	0.308	0.025
	OT-04	6-FAM	209-217	3	0.59	0.53	0.92	0.164	0.019
	OT-14	VIC	125-149	4	0.58	0.63	0.96	0.089	0.063
Mean	-	-	-	3.45	0.49	0.47	0.94	0.143	0.023

Table 1. Properties of the 11 microsatellite multiplexed loci used in this study and summary of the genetic variability assessed per locus. The table includes: number of alleles (NA), observed (HO) and expected (HE) heterozygosities, rates of positive PCR (PCR+), allelic dropout (ADO) and false allele (FA) for each locus.

PCR primer pair following the protocol used by Mucci and Randi (2007) and following the multi-tube procedure (four replicates). The digestion product was then analyzed together with multiplex A. In order to test the effect of the otter's gender on the number of individual re-samplings a Mann–Whitney test was conducted.

Population size estimates and kinship

Population size estimates were performed by two methods: with software GIMLET (Valière, 2002) and with software CAPWIRE (Miller et al., 2005). GIMLET estimates the population size as the asymptote of the function between the cumulative number of unique genotypes and the number of samples typed in a rarefaction curve and allows fitting the curve using three different equations (Kohn et al., 1999; Eggert et al., 2003 and Chessel's equation (Frantz and Roper, 2006). Program R v.2.8 (Ihaka and Gentleman, 1996) was used to analyze the accumulation curves using a script generated by GIMLET. CAPWIRE is a capture-mark-recapture-based program for non-invasive genetic sampling that appears to work especially well in small populations of less than 100 individuals and when dealing with considerable capture heterogeneity (Miller et al., 2005; Hájková et al., 2009). The mean otter density was calculated independently for each sampling zone. The genetic relatedness and

sibling analyses were calculated with two softwares. The program ML-RELATE (Kalinowski et al., 2006) uses a maximum likelihood method to compute pair-wise genetic relatedness (R_{xy}), representing the proportion of allelic composition shared between each pair of individuals. Sibship analysis was conducted using COLONY version 2.0.4 (Jones and Wang, 2010), with a typing error rate set at 0.01. This approach considers the likelihood of the entire pedigree, as opposed to relatedness of individuals on a pair-wise basis. To infer relatedness among individuals, we used all of the genotyped samples (20 individuals).

Results

Sample collection, species identification and otter distribution assessment

During this study, we have reported 924 sampling occasions at 198 surveyed sites (Fig. 1). At 63 of them (31.8%) at least one sample (from 1 to 6) was collected. 77 (38.9%) were positive sites for signs of the species, but no samples could be obtained for DNA studies. In the remaining 58 sites (29.3%) no otter evidences were detected. Out of 132 fecal samples collected from the entire study area, 26 were collected in 2007–2008 period and 106 between 2009 and 2010 (Fig. 1). Additionally, we obtained 2 otter tissue samples from road-killed animals.

After sequencing a 226 bp mtDNA fragment from fecal DNA, we effectively identified 127 samples as otter, one as European mink (*Mustela lutreola*) and another one as American mink (*Neovison vison*). Thus, unequivocal species identification was possible in 98.4% of the samples with 3 non identified feces. All samples corresponding to otter shared the same haplotype, Lut-1 the most common haplotype in Europe (Effenberger and Suchentrunk, 1999; Mucci et al., 1999; Cassens et al., 2000; Ferrando et al., 2004; Pérez-Haro et al., 2005).

The 129 fecal samples were displayed in a map along with the resulting otter distribution data for the study period. The results found otter presence in 34 of the 48 UTM grids, in comparison with 22 grids in the 2004–2006 period (III.NOS, López de Luzuriaga, 2008) (Fig. 1). This indicates a 25% increase in otter distribution and reflects recolonization from the main rivers, where the otter was previously and solely identified in the southernmost areas of the Mediterranean basin.

Microsatellite genotyping for individual identification and sex typing from otter spraints

The average proportion of positive PCRs (calculated from correctly and fully genotyped samples) was 94% and varied among loci from 89 to 97% and among samples from 82 to 100%. The observed average error rates across loci were: ADO = 0.143 and FA = 0.023.All the 11 loci were polymorphic and the number of alleles per locus ranged from 2 to 5 (mean 3.45). Marker OT-19 (5 alleles) was the most polymorphic one followed by LUT435, LUT715 and OT-14 (Table 1). The average non-biased observed and expected heterozygosity (H_o and H_e) were moderated, 0.49 and 0.47 respectively (Table 1). PID analysis showed that the set of 11 loci would produce an identical genotype with a probability of 3.48×10^{-7}, and with a probability of 2.41×10^{-3} for a full-sib. According to the values, the 5 most informative loci could distinguish two full-siblings with a 99% of probability (Mills et al., 2000).

From the 127 samples identified as otter, we obtained a complete multi-locus profile for 55 samples (43.3% of the analyzed samples). The 96.4% of genotyped samples were fresh (less than 24 h) while only two >24 h samples got a complete genetic profile. After a rejoining process (i.e. pairwise comparison of the different genotypes obtained) 18 different individuals were identified. In addition, two carcasses where found during the study period that were identified as different individuals. Therefore, the minimum population size (i.e. the number of unique genotypes found) for the study period was set at 20 individuals. The average number of detections (re-sampling) per individual was 2.85, with a total number of 39. Spatial distribution of the identified individuals with a full multilocus genotype and mean otter density (individuals/km river) per zone (A, B, C and D) is shown in Fig. 2. Otter density recorded varied between 0.06 and 0.12 otters per shoreline kilometer, with a mean otter density of 0.09 otters/km (Fig. 2). Population size estimated according to Eggert et al. (2003) was 24 individuals (18–56) while the software CAPWIRE estimated 29 (CI 95% = 20–37 otters, CI_{width}= 17).

Sexual typing revealed the presence of 11 females, 6 males and 3 individuals that couldn't be reliably sexed due to the existence of unspecific products, providing for a high sex identification success of 85%. Sex identification results indicate a male/female ratio of ~0.55.

Geographical, temporal distribution and kinship of genetically identified otters

Detailed information about geographical and temporal distribution of the 20 individuals is depicted in Appendix 1a to 1d corresponding to zones A to D, respectively. Otter 2 female was first detected in the Cantabrian watershed and re-sampled in the Mediterranean side for three years covering a minimum of 23 river km and overlapping territory with otter 1 male (Appendix 1a). The largest number of individuals was detected at the confluence of the major rivers and along the river Ebro (1 male and 7 females) where, unfortunately, several samples (n = 8) were genetically confirmed as otter but did not yield any valid genotype which could have increased the re-samples or the number of identified individuals coming from the nearest provinces (Appendix 1c). Individuals 19 and 20 male (Appendix 1d), detected only once, where identified as first order relatives (ML-RELATE, parent/offspring, R_{xy}= 0.724; COLONY, full-sib dyad, r = 0.615). According to COLONY all otters were related to each other (36 half-sib pairs), which prevents the reconstruction of detailed genealogical relationships among individuals.

Discussion

Reliability of the NGS methods for individual and sex identification of otter

We developed a reliable 11 microsatellite multiplex protocol for individual otter identification (Pid sib = 2.41×10^{-3}), proving a means of efficient application of non-invasive genetic monitoring of otters from spraints. The genotyping success from fecal samples (43.3%) was similar to other studies carried out with this species (19%, Bonesi et al., 2013; 20%, Dallas et al., 2003; 21%, Ferrando et al., 2008; 24%, Kalz et al., 2006; 41%, Prigioni et al., 2006; 44%, Arrendal et al., 2007; 48%, Hung et al., 2004; 55–63%, Hájková et al., 2009) and in other non-invasive genetic studies of carnivores (e.g. Ruiz-González et al., 2013, for martens; Caniglia et al., 2014, for wolves). As previously reported in the literature (reviewed in Hájková et al., 2009), we found that genotyping success improves when working with fresh samples, where the 96.4% of the completely genotyped samples were collected within 24 hours of deposition.

The frequency of allelic dropout and false alleles (ADO = 14.3%, FA = 2.3%) was similar to other otter non-invasive genetic studies (Hájková et al., 2006; Arrendal et al., 2007; Ferrando et al., 2008). Both observed and expected heterozygosity values (0.49 and 0.47) were lower than values reported for Catalonia (0.61–0.68 and 0.59–0.67, Ferrando et al., 2008) and Spain (0.64 and 0.58, Mucci et al., 2010). Nonetheless, the results have to be cautiously analyzed due to the low number of identified individuals and the differences between the selected markers among studies.

The success of the sex identification method (85%) was high and similar to other otter studies (79%, Mucci and Randi, 2007; 87.5%, Park et al., 2011). Sex identification results (6 males and 11 females) indicated a male/female ratio of ~0.55, in contrast to the male-biased sex-ratio found in studies with larger sample size (1.4, Dallas et al., 2003; 1.2, Kalz et al., 2006; 1.0, Ferrando et al., 2008; 0.55, Bonesi et al., 2013). According to Kruuk (2006) several female otters may live together in a group territory, each female having its own core-area to rear their young, to defend against other females. Moreover, as reported by Koelewijn et al. (2010), juvenile females stay close to the territories of their mothers while the juvenile males are forced to move by dominant males. Thus, by prioritizing sampling in areas with a higher spraint density, we may have biased the number of female detected and could be one of the possible explanations for the considerable sex-ratio deviation. Overall, we obtained a high capture heterogeneity among individuals

where males were re-sampled more than females (Mann–Whitney U test; P = 0.007). Certainly, dominant male otters with large territories tend to have a higher number of spraints in non-invasive surveys (Bonesi et al., 2013).

Assessing the recovery of otter distribution through non-invasive genetic sampling

Surveys have been conducted for decades to monitor otter distribution in Spain and throughout Europe (Lodé, 1993; Kauhala, 1996; Kranz et al., 2001; McDonald et al., 2007; Sulkava, 2007; Prigioni et al., 2007). However, they have often relied on locating field signs and/or feces with identification from morphological characteristics although this approach could be partially biased due to potential error rates in spraint identification (Kalz et al., 2006; O'Neill et al., 2013). Increasingly, studies that aim to determine otter distribution based on identifying spraints are incorporating DNA analysis to confirm species identity, as a first screening before further genotyping (e.g. Hung et al., 2004; Kalz et al., 2006; Ferrando et al., 2008; Park et al., 2011). In this study, 98% of the visually assigned otter spraints were genetically confirmed as otter spraints. Even though the misidentification rate seems low and similar to that obtained in other studies (e.g. O'Neill et al., 2013; approx. 2.5%) that monitored threatened populations, is worth considering a protocol in which species identification is first conducted to avoid misidentification with other sympatric mustelids such as minks and polecats (Hansen and Jacobsen, 1999). Even if other previous otter studies found nearly perfect species identification (Kalz et al., 2006), it could highly depend on surveyors experience (Ruiz-González et al., 2013).

In addition, incorporating a species identification step using sequencing allowed us to not only verify the specific identity of each spraint but also to identify the shared haplotype of all samples (i.e. Lut 1). These results confirm previous results about mtDNA homogeneity of Spanish otter populations exclusively represented by a unique mtDNA lineage, suggesting that the growing otter population it is unlikely to be derived from surrounding European countries.

The otter is currently widely distributed in the surveyed area being present in the major Mediterranean rivers and their tributaries (34 out of the 48 UTM 10 × 10 km grids). Otter presence was confirmed in all of the 22 grids recorded in the III. NOS (López de Luzuriaga et al., 2008) and evidences were found in 12 new ones, indicating a range increase in the last few years. These results suggest that population expansion is occurring in Álava, in parallel to the observed recovery in other regions within the Iberian Peninsula (López-Martín and

Jiménez, 2008).This is further indicated by recent observations from 2011 in which the otter was detected in 3 new grids (López de Luzuriaga, unpublished data). Nevertheless, although this mustelid seems to follow a clear pattern of re-colonization upstream of the main rivers of the Mediterranean basin, it is still scarce in the Basque Atlantic watershed (López de Luzuriaga et al., 2008), with the first confirmed data about its presence after 50 years being reported in this work.

The European otter is considered a food-limited species and fluctuations in their breeding success are related to food availability (Kruuk, 2006; Ruiz-Olmo et al., 2011). Thus, otters tend to select habitat structures with abundant and predictable food availability, especially fish and crayfish (Ruiz-Olmo et al., 2011). In the study area, the majority of the recently occupied grids correspond to areas with fishing reserves for trout (*Salmo trutta*) and signal crayfish (*Pacifastacus leniusculus*) which might have attracted dispersing otters, as the expansion of the signal and red swamp crayfish (*Procambarus clarkii*) has been suggested to be favorable for the species (López de Luzuriaga et al., 2008).

Population size and density estimations

The Eurasian otter is an "Endangered" species in the Basque Country; therefore, estimates of the number of individuals in the population is a essential information to guide short, medium and long term management of the species. In this study, a total of 20 different otters were identified using a panel of 11 microsatellite markers, so the minimum population size for the study period was set at 20 individuals where at least two were dead. According to the estimations made for the period 2007–2010, the population would consist of 24 to 29 individuals even though this value has to be carefully interpreted due to the extensive study period. Sibship analyses revealed that all the genotyped individuals were closely-related, which is consistent with a recently colonized population. We hypothesize that the Ebro basin, hosting founding populations, acted as a potential source of young animals that dispersed upstream of the main river. Nonetheless, COLONY and ML-RELATE analyses exclusively allowed the reconstruction of potential genealogical relationships between a few individuals (e.g. ind 19 and 20) and therefore no family group reconstruction was possible.

The otter density recorded (0.06 and 0.12 otters per shoreline km), was similar to densities obtained in studies carried out in northeastern Spain (0.04–0.11 otters/km) but lower than densities found in Italy (0.18–0.20 otters/km), Sweden (0.20 otters/km), United Kingdom (0.012–0.33 and 0.27–0.34 otters/km), and eastern Europe (0.11–0.37 and 0.18–0.57 otters/

km) (reviewed in Prigioni et al., 2006). Within the study area, the Ebro River is the only river that has hosted a continuously extant otter population during last two decades (López-Martín and Jiménez, 2008). Accordingly, it seems that the Ebro has played an important role as an axis for otter colonization, allowing the colonization of individuals from adjacent regions, such as Burgos and La Rioja that have maintained stable populations of otter during the most drastic otter decline in Spain (Aguilar and Gómez-Gayubo, 2008).

Conclusions

In this study we combined non-invasive molecular techniques with geographic information systems to record the contemporary expansion of the otter at regional scale, confirming the recovery pattern previously highlighted in surrounding otter populations in the Iberian Peninsula (López-Martín and Jiménez, 2008). More over, microsatellite and sex genotyping from spraints were reliable and effective in identifying the individuals inhabiting the study area and enabled us to perform the first population and density estimations in the Northern Iberian Peninsula provided by non-invasive genetic techniques. Overall, this study provided valuable information regarding the spatial distribution and population size of an expanding otter population in northern Spain. Moreover, it has helped refine otter sign survey techniques by the use of non-invasive molecular techniques, and provide useful information to design successful otter management activities within the OMPA framework.

Acknowledgments

This study was financially supported by Department of Environment of the Diputación Foral de Álava, Department of Environment of the Basque Country Government and SAIOTEK program (Ref: S-PE11UN028). MV (Ref: RBFI-2012-446) and ARG (Ref: DKR-2012-64) were supported by a PhD and post-doctoral fellowships awarded by the Dept. of Education, Universities and Research of the Basque Government. This study has been also partially funded by the Basque Government through the Research group on "Systematics, Biogeography and Population Dynamics" (Ref: IT317-10; GIC10/76).

References

Aguilar, C.M., Gómez-Gayubo, A., 2008. La nutria en La Rioja. In: López-Martín, J.M., Jiménez, J. (Eds). La nutria en España. Veinte años de seguimiento de un mamífero amenazado. SECEM, Málaga, pp. 217-225.

Arrendal, J., 2007. Conservation Genetics of the Eurasian Otter in Sweden. Dissertation, University of Uppsala, Sweden.

Ben-David, M., Blundell, G.M., Kern, J.W., Maier, J.A.K., Brown, E.D., Jewett, S.C., 2005. Communication in river otters: creation of variable resource sheds for terrestrial communities. Ecology. 86(5), 1331-1345.

Beja-Pereira, A., Oliveira, R., Alves, P. C., Schwartz, M. K., Luikart, G., 2009. Advancing ecological understandings through technological transformations in noninvasive genetics. Mol Ecol Resour. 9(5), 1279–301.

Bonesi, L., Hale, M., Macdonald, D.W., 2013. Lessons for the use of non-invasive genetic sampling as a way to estimate Eurasian otter population size and sex-ratio. Acta Theriol. 58, 157-168.

Chanin, P., 1985. The Natural History of Otters. Croom Helm, Kent, 179 pp.

Dallas, J.F., Bacon, P.J., Carss, D.N., Conroy, J.W.H., Green, R., Jefferies, D.J., Kruuk, H., Marshall, F., Piertney, S.B., Racey, P.A., 1999. Genetic diversity in the Eurasian otter, *Lutra lutra*, in Scotland. Evidence from microsatellite polymorphism. Biol J Linnean Soc. 68, 73–86.

Dallas, J. F., Carss, D.N., Marshall, F., Koepfli, K., Kruuk, H., Piertney, B., Bacon, P.J., 2000. Sex identification of the Eurasian otter *Lutra lutra* by PCR typing of spraints. Conserv Genet. 1, 181–183.

Dallas, J.F., Coxon, K.E., Sykes, T., Chanin, P.R.F., Marshall, F., Carss, D.N., Bacon, P., Piertney, S.B. Racey., P.A., 2003. Similar estimates of population genetic composition and sex ratio derived from carcasses and faeces of Eurasian otter *Lutra lutra*. Mol Ecol Notes. 12, 275-282.

Dallas, J.F., Piertney, S.B., 1998. Microsatellite primers for the Eurasian otter. Mol Ecol. 7, 1248-1251.

Delibes, M., 1990. La Nutria (*Lutra lutra*) en España. ICONA. Serie Técnica, Madrid.

Diputación Foral de Álava, 2004. Orden Foral 880/2004, de 27 de octubre por la que se aprueba el Plan de Gestión de la Nutria, *Lutra lutra* (Linnaeus, 1758) en el Territorio Histórico de Álava. BOTHA nº 136.

Eggert, L.S., Eggert, J.A., Woodruff, D.S., 2003. Estimating population sizes for elusive animals: the forest elephants of Kakum National Park, Ghana. Mol Ecol Notes. 12, 1389-1402.

Frantz, A., Roper, T.J., 2006. Simulations to assess the performance of different rarefaction methods in estimating population size using small datasets. Conserv Genet. 7, 315-318.

Ferrando, A., Lecis, R., Domingo-Roura, X., Ponsà, M., 2008. Genetic diversity and individual identification of reintroduced otters (*Lutra lutra*) in north-eastern Spain by DNA genotyping of spraints. Conserv Genet. 9, 129-139.

Gómez-Moliner, B. J., Cabria, M.T., Rubines, J., Garin, I., Madeira, M. J., Elejalde, A., Palazón, S., 2004. PCR-RFLP identification of mustelid species: European mink (*Mustela lutreola*), American mink (*M. vison*) and polecat (*M. putorius*) by analysis of excremental DNA. J Zool, 262(3), 311–316.

Gorman, M.L., Trowbridge, B.J., 1989. The role of odor in the social lives of carnivores. pp 57-88. In J. L. Gittleman (Ed). Carnivore behavior, ecology, and evolution. Cornell University Press, Ithaca, New York, USA.

Hájková, P., Zemanová, B., Bryja, J., Hájek, B., Roche, K., Tkadlec, E., Zima, J., 2006. Factors affecting success of PCR amplification of microsatellite loci from otter faeces. Mol Ecol Notes. 6, 559-562.

Hájková, P., Zemanová, B., Roche, K., Hájek, B., 2009. An evaluation of field and non invasive genetic methods for estimating Eurasian otter population size. Conserv Genet. 10, 1667-1681.

Hernando, A., Martínez de Lecea, F., Illana, A., Bayona, J., Echegaray, J., 2005. Sondeo y evolución de la distribución de la nutria paleártica (*Lutra lutra*, Linnaeus, 1758) en el País Vasco (N España). Galemys, 1, 25–46.

Huang, C-C., Hsu, Y-C., Lee, L-L., Li, S-H., 2005. Isolation and characterization of tetramicrosatellite DNA markers in the Eurasian otter (*Lutra lutra*). Mol Ecol Notes. 5, 314-316.

Hung, C-M., Li, S-H., Lee, L-L., 2004. Faecal DNA typing determines the abundance and spacial organization of otters (*Lutra lutra*) along two stream systems in Kinmen. Anim Conserv. 7, 301-311.

Ihaka, R., Gentleman, R., 1996. R: a language for data analysis and graphics. J. Comp. Graph. Stat. 5, 299-314.

Jones, O.R., Wang, J., 2010. COLONY: a program for parentage and sibship inference from multilocus genotype data. Mol Ecol Resour. 10(3), 551–5.

Kalinowski, S.T., Wagner, A.P., Taper, M.L., 2006. ML-Relate: a computer program for maximum likelihood estimation of relatedness and relationship. Mol Ecol Notes. 6,576-579.

Kalz, B., Jewgenow, K., Fickel, J., 2006. Structure of an otter (*Lutra lutra*) population in Germany – results of DNA and hormone analyses from faecal samples. Mammal Biol. 71(6), 321–335.

Kauhala, K., 1996. Distributional history of the American mink (*Mustela vison*) in Finland with special reference to the trends in otter (*Lutra lutra*) populations. Annales Zoologici Fennici. 33, 283–291.

Koelewijn, H.P., Pérez-Haro, M., Jansman, H.A.H., Boerwinkel, M.C., Bovernschen, D.R., 2010. The reintroduction of the Eurasian otter (*Lutra lutra*) into the Netherlands: hidden life revealed by noninvasive genetic monitoring. Conserv Genet. 11, 601-614.

Kohn, M.H., York, E.C., Kamradt, D.A., Haught, G., Sauvajot, R.M., Wayen, R.K., 1999. Estimating population size by genotyping faeces. Proc. R. Soc. B. 266(1420), 657–663.

Kranz, A., Polednik, L., Pinter, V., Parz-gollner, R., Fischotter, S., 2001. Distribution, status and conservation of otters in Lower Austria. Wiss. Mitt. Niederösterr. Landesmuseum, 14, 39–50.

Kruuk, H., 1992. Scent marking by otters (*Lutra lutra*): signaling the use of resources. Behav Ecol. 3, 133-140.

Kruuk, H., 1995. Wild Otters, Predation and Populations. Oxford Univ. Press. Oxford.

Kruuk, H. 2006. Otters. Ecology, behavior and conservation. Oxford University press. Oxford.

Lodé, T., 1993. The decline of otter *Lutra lutra* populations in the region of the pays de Loire, Western France. Biol Conserv. 65(1), 9–13.

López de Luzuriaga, J., Zuberogoitia, I., Zabala, J., 2008. La nutria en el País Vasco. In: López-Martín, J.M., Jiménez, J. (Eds). La nutria en España. Veinte años de seguimiento de un mamífero amenazado. SECEM, Málaga, pp. 207-215.

López-Martín, J.M., Jiménez, J., 2008. La nutria en España. Veinte años de un mamífero amenazado. SECEM. Málaga.

Macdonald, S.M., Mason, C.F., 1983. The Otter *Lutra lutra* in Southern Italy. Biol Conserv. 25 (2), 95-101.

Macdonald, S.M., Mason, C.F., 1994. Status and conservation need of the otter (*Lutra lutra*) in the western Paleartic. Council of Europe Press, Nat Environ. 67, 1-54.

Mason, C.F., Macdonald, S.M., 2004. Growth in otter (*Lutra lutra*) populations in the UK as shown by long-term monitoring. Ambio. 33, 148–152.

McDonald, R. A., Hara, K. O., Morrish, D. J., 2007. Decline of invasive alien mink (*Mustela vison*) is concurrent with recovery of native otters (*Lutra lutra*). Divers. Distrib. 13, 92–98.

Miller, C.R., Joyce, P., Waits, L.P., 2005. A new method for estimating the size of small populations from genetic mark-recapture data. Mol Ecol Notes. 14, 1991-2005.

Mills, L., Citta, J., Lair, K., Schwartz, M., Tallmon, D., 2000. Estimating animal abundance using non-invasive DNA sampling: promise and pitfalls. J. Appl. Ecol. 10, 283-294.

Mucci, N., Arrendal, J., Ansorge, H., Bailey, M., Bodner, M., Delibes, M., Randi, E., 2010. Genetic diversity and landscape genetic structure of otter (*Lutra lutra*) populations in Europe. Conserv Genet. 11, 583–599.

Mucci, N., Randi, E., 2007. Sex identification of Eurasian otter (*Lutra lutra*) non-invasive DNA samples using ZFX/ZFY sequences. Conserv Genet. 8, 1479-1482.

Palomo, L.J., Gisbert, J., 2002. Atlas de los mamíferos terrestres de España. Dirección General de Conservación de la Naturaleza-SECEM-SECEMU, Madrid, 564 pp.

Park, H., Han, T., Kim, D., Min, M., Han, S., Kim, K., Lee, H., 2011. Individual identification and sex determination of Eurasian otters (*Luta lutra*) in Daegu city based on genetic analysis of otter spraints. Genes & Genomics. 33(6), 653-657.

Parry, G. S., Bodger, O., McDonald, R.A., Forman, D.W., 2013. A systematic re-sampling approach to assess the probability of detecting otters *Lutra lutra* using spraint surveys on small lowland rivers. Ecol. Inform. 14, 64–70.

Pérez-Haro, M., Viñas, J., Mañnas, F., Batet, A., Ruiz-Olmo, J., Pla, C., 2005. Genetic variability in the complete mitochondrial control region of the Eurasian otter (*Lutra lutra*) in the Iberian Peninsula. Biol J Linn Soc. 86(4), 397-403.

Pompanon, F., Bonin, A., Bellemain, E., Taberlet, P., 2005. Genotyping errors: causes, consequences and solutions. Nat. Rev. Genet. 6(11), 847-849.

Prigioni, C., Remonti, L., Balestrieri, A., Sgrosso, S., Priore, G., Mucci, N., Randi, E., 2006. Estimation of european otter (*Lutra lutra*) population size by fecal DNA typing in southern Italy. J Mammal. 87(5), 855-858.

Prigioni, C., Balestrieri, A., Remonti, L., 2007. Decline and recovery in otter *Lutra lutra* populations in Italy. Mammal Review. 37(1), 71–79.

Remonti L., Prigioni C., Balestrieri A., Sgrosso S., Priore G., 2008. Distribution of a recolonising species may not reflect habitat suitability alone: the case of the Eurasian otter (*Lutra lutra*) in southern Italy. Wildl Res. 35, 798-805.

Ruiz-González, A., Rubines, J., Berdión, O., Gómez-Moliner, B., 2008a. A non-invasive genetic method to identify the sympatric mustelids pine marten (*Martes martes*) and stone marten (*Martes foina*): preliminary distribution survey on the northern Iberian Peninsula. Eur J Wildl Res. 54(2), 253-261.

Ruiz-González, A., Ferrando, A., Pérez-Haro, M., Rubines, J., Madeira, M.J., Ponsá, M., Ruiz-Olmo., J., Romero, R., Gómez-Moliner, B.J., 2008b. Seguimientos moleculares de las poblaciones de nutria *Lutra lutra* (Linnaeus, 1758) a partir de muestreos no invasivos. In: López-Martín, J.M., Jiménez, J. (Eds). La nutria en España. Veinte años de seguimiento de un mamífero amenazado. SECEM, Málaga, pp. 207-215.

Ruiz-González, A., Madeira, M. J., Randi, E., Urra, F., Gómez-Moliner, B. J., 2013. Non-invasive genetic sampling of sympatric marten species (*Martes martes* and *Martes foina*): assessing species and individual identification success rates on faecal DNA genotyping. Eur J Wildl Res, 59(3), 371–386.

Ruiz-Olmo, J., Delibes. M., 1998. La nutria en España ante el horizonte del año 2000. SECEM, Málaga.

Ruiz-Olmo, J., Loy, A., Cianfrani, C., Yoxon, P., Yoxon, G., De Silva, P.K., Roos, A., Bisther, M., Hájkova, P., Zemanová, B., 2008. *Lutra lutra*. In: IUCN 2013. IUCN Red List of Threatened Species. Version 2013.2. <www.iucnredlist.org>. Downloaded on 10 February 2014.

Ruiz-Olmo, J., Batet, A., Mañas, F., Martínez-Vidal, R., 2011. Factors affecting otter (*Lutra lutra*) abundance and breeding success in freshwater habitats of the northeastern Iberian Peninsula. Eur J Wildl Res. 57, 827-842.

Schwartz M.K., Monfort, S.L., 2008. Genetic and endocrine tools for carnivore surveys. R. Lomg, P MacKay, J. Ray and W. Zielinski (eds.) Noninvasive survey methods for North American Carnivores. Island, Wahington, DC.

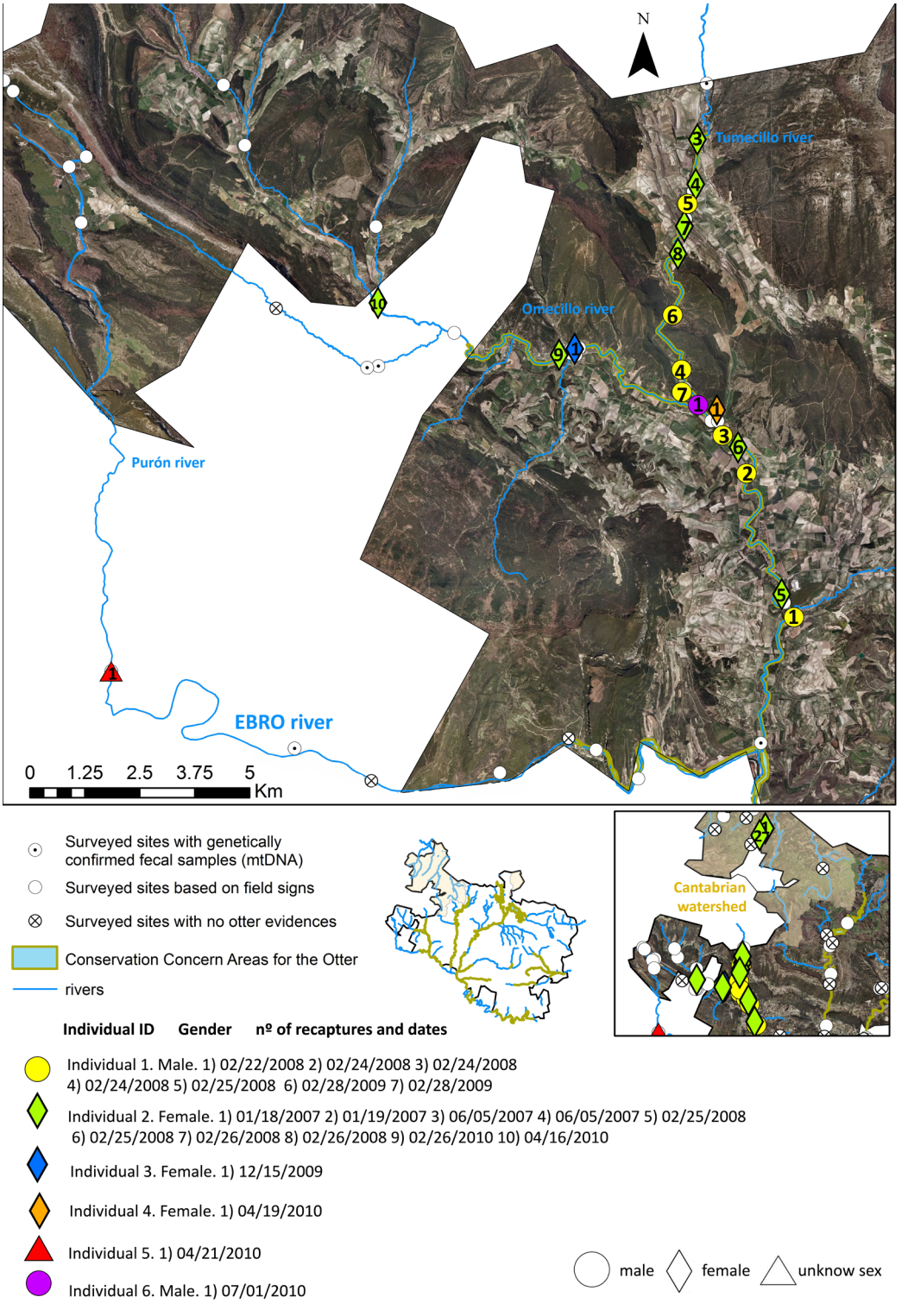
Sulkava, R., 2007. Snow tracking: a relevant method for estimating otter *Lutra lutra* populations. Wildlife Biol. 13(2), 208–218.

Taberlet, P., Luikart, G., 1999. Non-invasive genetic sampling and individual identification. Biol J Linn Soc. 68, 41–55.

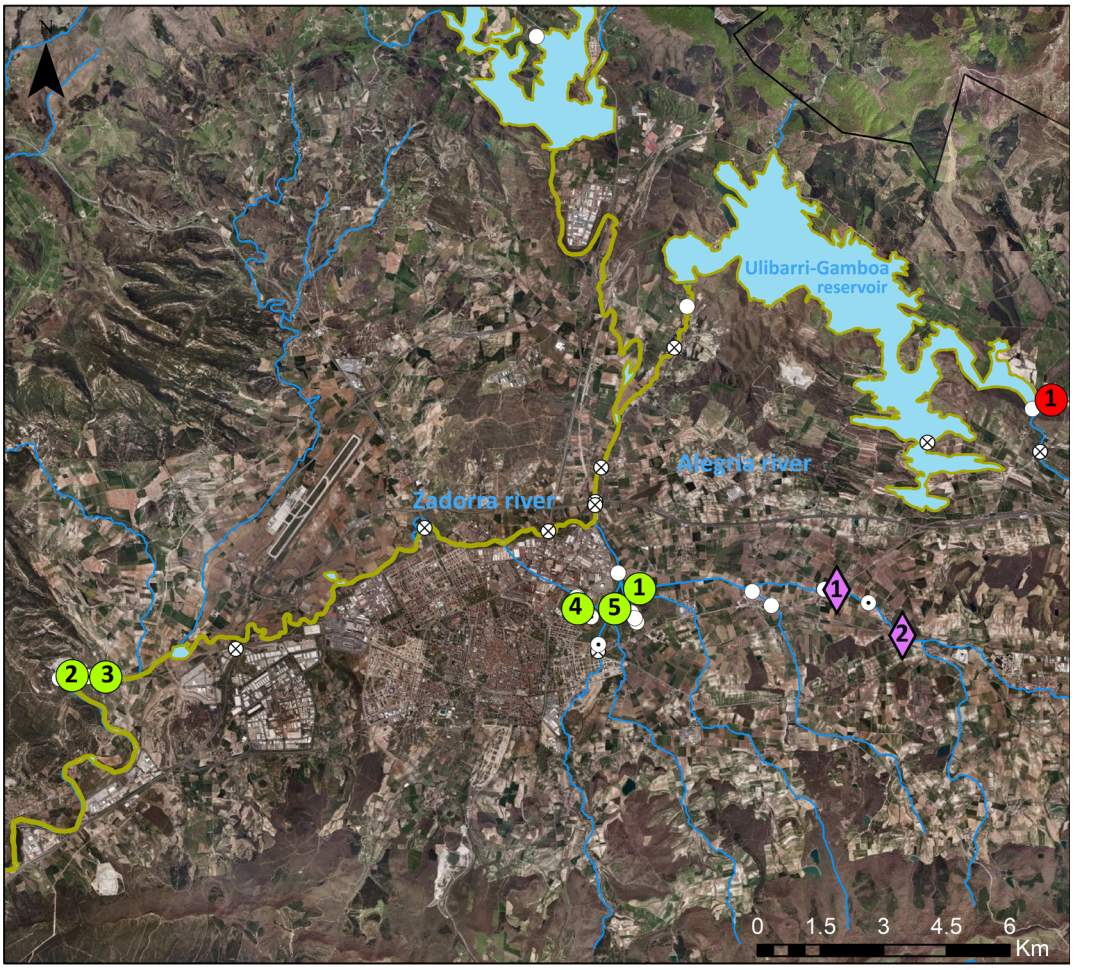
Valière, N., 2002., GIMLET: a computer program for analysing genetic individual identification data. Mol Ecol Notes. 2, 377-379.

Waits, L.P., Paetkau, D., 2005. Noninvasive genetic sampling tools for wildlife biologists: A review of applications and recommendations for accurate data collection. J. Wildl. Manag. 69, 1419-1433.

Appendix 1a


N
Tuñecillo river
Omecillo river
Purón river
EBRO river
0 1.25 2.5 3.75 5
Km
Surveyed sites with genetically confirmed fecal samples (mtDNA)
Surveyed sites based on field signs
Surveyed sites with no otter evidences
Conservation Concern Areas for the Otter
rivers
Cantabrian watershed
Individual ID Gender nº of recaptures and dates
Individual 1. Male. 1) 02/22/2008 2) 02/24/2008 3) 02/24/2008 4) 02/24/2008 5) 02/25/2008 6) 02/28/2009 7) 02/28/2009
Individual 2. Female. 1) 01/18/2007 2) 01/19/2007 3) 06/05/2007 4) 06/05/2007 5) 02/25/2008 6) 02/25/2008 7) 02/26/2008 8) 02/26/2008 9) 02/26/2010 10) 04/16/2010
Individual 3. Female. 1) 12/15/2009
Individual 4. Female. 1) 04/19/2010
Individual 5. 1) 04/21/2010
Individual 6. Male. 1) 07/01/2010
male
female
unknow sex

Appendix 1b

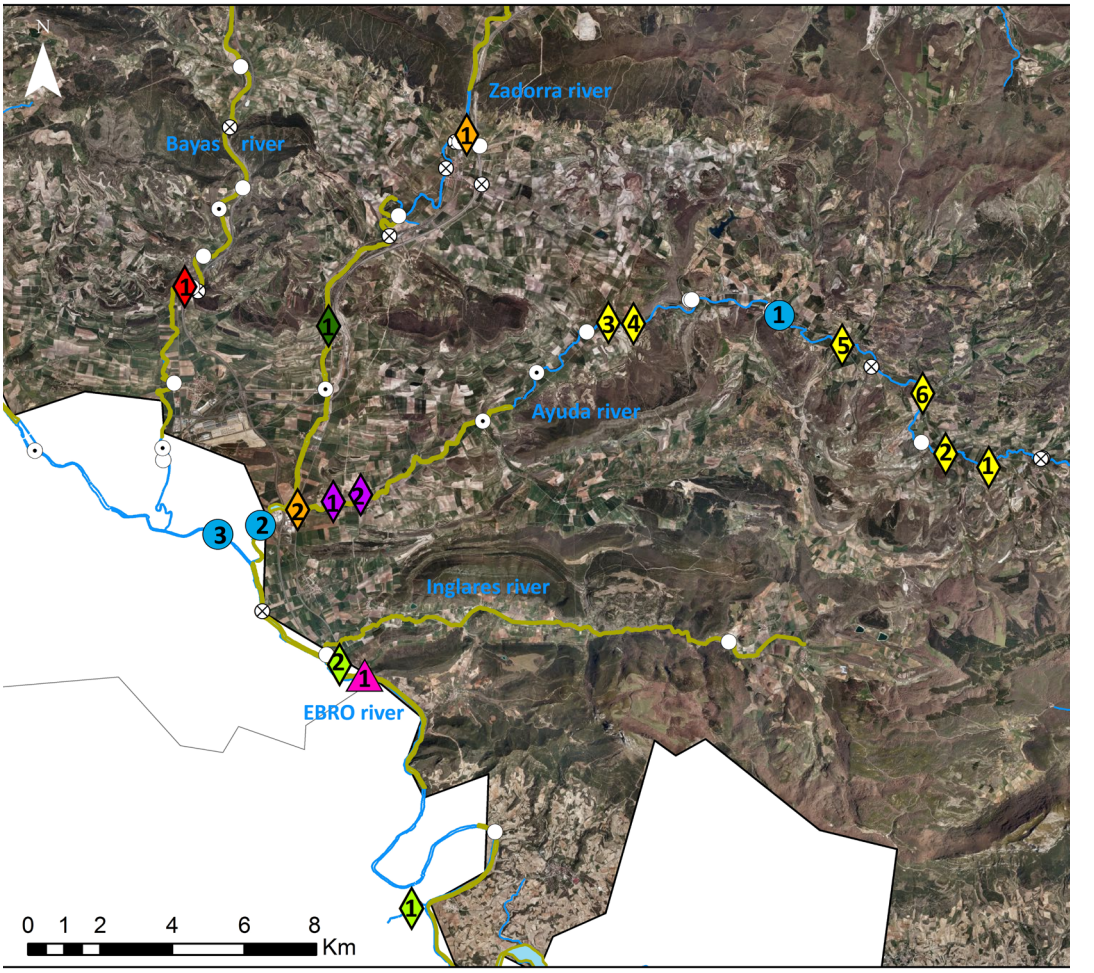


Surveyed sites with genetically confirmed fecal samples (mtDNA)
Surveyed sites based on field signs
Surveyed sites with no otter evidences
Conservation Concern Areas for the Otter
rivers

male female unknow se

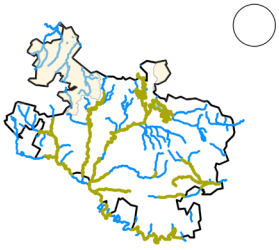
Individual ID	Gender	nº of recaptures and dates
Individual 7	Male	1) 09/24/2007 2) 11/24/2007 3) 11/24/2007 4) 03/04/2009 5) 03/06/2009
Individual 8	Female	1) 02/19/2009 2) 03/13/2009
Individual 9	Male	1) 02/20/2009

Appendix 1c



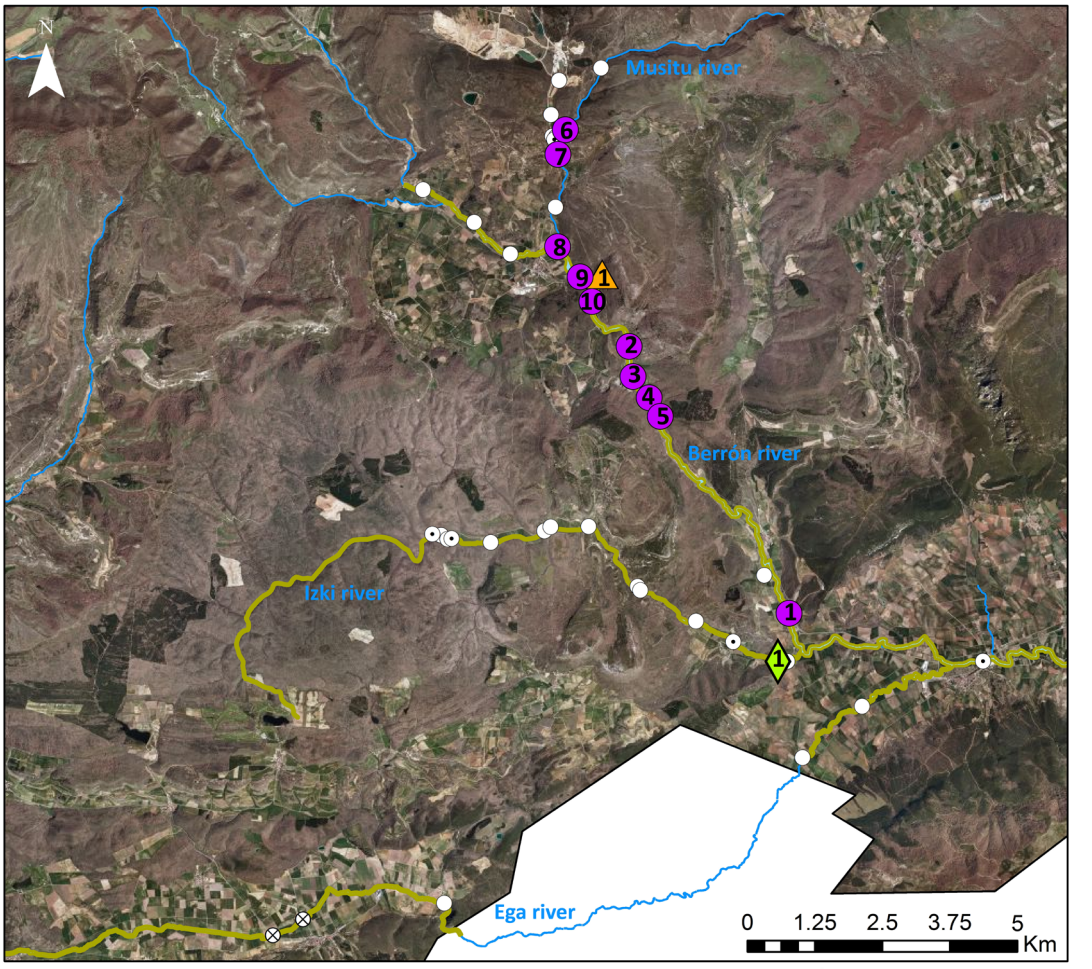
Surveyed sites with genetically confirmed fecal samples (mtDNA)
Surveyed sites based on field signs
Surveyed sites with no otter evidences
Conservation Concern Areas for the Otter
rivers

male female unknow sex



Individual ID	Gender	nº of recaptures and dates
Individual 10.	Female.	1) 02/28/2009 2) 02/28/2009 3) 05/19/2010 4) 03/30/2010 5) 03/30/2010 6) 03/31/2010
Individual 11.	Female.	1) 04/22/2010 2) 09/17/2010
Individual 12.	Female.	1) 05/19/2009 2) 05/20/2009
Individual 13.	Female.	1) 09/08/2010
Individual 14.	Male.	1) 02/28/2009 2) 05/10/2010 3) 05/12/2010
Individual 15.	Female.	1) 04/04/2010 2)04/06/2010
Individual 16.	Female.	1) 03/16/2010
Individual 17.	Female.	1) 05/16/2009

Appendix 1d


N
Musitu river
Berrón river
Izki river
Ega river
0 1.25 2.5 3.75 5
Km
male
female
unknow sex
Surveyed sites with genetically confirmed fecal samples (mtDNA)
Surveyed sites based on field signs
Surveyed sites with no otter evidences
Conservation Concern Areas for the Otter
rivers
Individual ID Gender nº of recaptures and dates
Individual 18. Male. 1) 08/05/2009 2) 12/10/2009 3) 02/02/2010 4) 05/13/2010 5) 05/13/2010 6) 09/29/2010 7) 09/29/2010 8) 10/01/2010 9) 10/01/2010 10) 10/29/2010
Individual 19. 1)10/29/2010
Individual 20. Female. 1)12/10/2009

Chapter II

PAPER 2: Shaken but not stirred: Multiscale habitat suitability modeling of sympatric marten species (*Martes martes* and *Martes foina*) in the northern Iberian Peninsula

María Vergara, Samuel A. Cushman and Aritz Ruiz-González

Manuscript. Special issue of Landscape Ecology "Multi-scale habitat modeling"

Abstract

Context. Habitat heterogeneity at several scales may simultaneously drive species distributions and abundance, and species likely differ in the scales of their ecological responses. Therefore, multispecies and multiscale habitat suitability models are important to identify the environmental variables and scales influencing habitat selection and facilitate the comparison of closely related species with different ecological requirements.

Objectives. This study explores the multiscale relationships of habitat suitability by the pine (*Martes martes*) and stone marten (*M. foina*) in northern Spain to evaluate differences in habitat selection and scaling and to determine if there is habitat niche displacement in areas where both species coexist.

Methods. We combined bivariate scaling and maximum entropy modeling to compare the multiscale habitat selection of the pine and the stone marten. Then, we explored the performance of three sampling bias correction methods at four spatial scales to optimize each marten habitat suitability models (HSM). Finally, the corrected HSMs were compared to explore the presence of a significant niche differentiation between the species.

Results. The comparison among high-resolution HSMs resulted in the detection of a significant niche divergence between species. The pine marten was associated with cooler mountainous areas with low levels of human disturbance, high proportion of natural forests and forestry plantations and medium-extent agroforestry mosaics. The stone marten, in contrast, was most related to the density of urban areas, the proportion and extensiveness of croplands, the existence of some scrub cover and the availability of semi-continuous grasslands.

Conclusions. This study outlines both the influence of the spatial scale and the importance of the sampling bias corrections in HSM. To our knowledge, this is the first study comparing multiscale habitat selection and niche divergence of two related marten species. Further, this study provides a useful methodological framework for multispecies and multiscale comparatives.

Keywords: HSM, scale dependency, sampling bias, niche divergence, Maxent, pine marten, stone marten.

Introduction

Habitat Suitability Models (HSM, Girvetz and Greco 2009; Koreň et al 2011; Bellamy et al 2013), also referred as Species Distribution Models (SDM, Syfert et al 2013; Kramer-Schadt et al 2013; Fourcade et al 2014) or Environmental (or ecological) Niche Models (ENMs, Warren et al 2010; Warren and Seifert 2011), have become a fundamental tool in ecology and biogeography as they correlate the presence of species at multiple locations with relevant environmental covariates to estimate habitat preferences and/or predict species distributions (Elith et al 2011). The identification of the factors constraining species presence is central to identify the most suitable areas for a particular taxa, to infer occurrence probability in areas where no systematic surveys have been conducted (Fourcade et al 2014), to assess the potential expansion of invasive species (i.e. Elith et al 2010; Jiménez-Valverde et al 2011) or to estimate future ranges under different climate change scenarios (i.e. Khanum et al 2013). HSMs can also be particularly useful to explore the environmental characteristics conditioning several species´ overlapping range in order to quantitatively estimate the niche divergence in related species and to infer the roles of competitive interaction (e.g. Wellenreuther et al 2012). This application is particularly interesting as a species' realized niche could be restricted by competition with sympatric congeners (Anderson et al 2002).

Most species records are only available in the form of presence-only datasets (PO), which provide solid information regarding the species' presence but no direct data regarding absences (Pearson et al 2007). Thus, in recent decades, there has been an increasing focus on developing methods to work with this information (reviewed in Yackulic et al 2013). MAXENT (Phillips et al 2006) is one of the most commonly used HSM techniques dealing with PO data (Elith et al 2006; Elith et al 2011; Syfert et al 2013). All PO based HSM methods work under the assumption that the entire area under study has been systematically and randomly sampled (Elith et al 2011). Yet, most datasets include spatially biased presence records. Lately, the effect of sampling bias in model performance has increasingly been acknowledged and several correction methods have been proposed to improve model accuracy (Anderson and Gonzalez 2011; Syfert et al 2013; Kramer-Schadt et al 2013; Brown 2014). Consequently, Fourcade et al (2014) recommended evaluating several types of corrections before choosing the final correction method based on a) their effect in classical model evaluation metrics (e.g. AUC) and b) the adequacy of the produced habitat suitability map to *a priori* knowledge of a given species distribution.

To produce a realistic HSM, habitat features must be measured at spatial resolutions

that are relevant to the organism being modeled based on its ecological adaptations and life-history strategy, as species use habitats differently at widely divergent scales (Johnson 1980; Cushman and McGarigal 2004; Graf et al 2005). Even when the correct variables are employed, the incorrect specification of the scale at which scale-dependant characteristics operate could lead to a dramatically different interpretation of which factors are actually influencing the occurrence of a given species (e.g. Girvetz and Greco 2009; Bellamy et al 2013; Shirk et al 2014). Nevertheless, there is no *a priori* way to infer the grain and extent at which each environmental predictor is most strongly related to species presence (Shirk et al 2012). Therefore, habitat suitability modeling has shifted from models based exclusively in expert's opinion to increasingly complex multiscale models to reveal the true grain at which species respond to the landscape. Multiscale HSMs allow more accurate predictions of species occurrence based on the systematic variation of the scale of analysis of each variable to find the dominant scale at which they operate to build the models (e.g. Shirk et al 2012). In this context, recent studies conducted on mammals (e.g. Wasserman et al 2012; Shirk et al 2014; Mateo-Sánchez et al 2013, Bellamy et al 2013) have demonstrated the effectiveness of multiscale approaches.

The European pine marten (*Martes martes*) and the stone marten (*Martes foina*) are two closely related mustelids living sympatrically over a wide area of Europe (Proulx et al 2004), and the northern Iberian Peninsula represents both the southern limit of the pine marten distribution and the southernmost area of sympatry in southwestern Europe. These species share similar biological characteristics (e.g. body-size, throphic niche or activity patterns) making it challenging to reliably determine differences in distribution based on observational data (Proulx et al 2004). The application of molecular methods on non-invasively collected faecal samples has proven to be a cost-effective way to reliably verify and monitor these elusive species´ presence (e.g. Ruiz-González et al 2013) and thus, also useful to build PO dataset for HSM purposes.

Recent studies have revealed that the pine marten is not as obligately dependent on forest habitats as previously believed (Virgós et al 2012) and may be capable of colonizing agricultural landscapes containing highly fragmented woodland and forest patches (Mergey et al 2011; Balestrieri et al 2015). However, the pine marten is either threatened or rare in many countries, while the stone marten´s geographic range has expanded due to its behavioral plasticity, broader habitat niche and less vulnerability to anthropogenic impacts (Proulx et al 2004; Goszczyński et al 2007; Herr et al 2009). Despite their extensive overlapping range, few studies addressing the differences in the pine and the stone marten habitat associations

have been conducted (e.g. Posłuszny et al 2007; Goszczyński et al 2007; Larroque et al 2015) and currently no in-depth studies have been published on their multiscale habitat selection (Virgós et al 2012).

The goal of this paper is to compare and contrast pine and stone marten multiscale habitat selection in a sympatric area of northern Spain by investigating how different environmental characteristics, measured at varying spatial scales, shape the distribution of each species. Subsequently, we compared the performance of three methods of sampling bias correction (Fourcade et al 2014) to improve the predictions of the final habitat suitability maps. Finally, the resulting HSMs were compared to explore niche differentiation among the species. To our knowledge, this is the first study comparing multiscale habitat selection of two sympatric marten species to clarify their habitat preferences and niche divergence when coexisting.

Material and methods

Study area

The study area comprises the regions of the Basque Country and Navarre and the surrounding territories of Cantabria, Castille and Leon and La Rioja (northern Spain, Fig. 1). The area is located in the southern range limit of the pine marten where it occurs sympatrically with the stone marten. The study area is 31,500 Km^2 with forest covering 31.3%, non-forested mountains 27.3%, cultivated lands 26.8% and urban, infrastructure and remaining land cover types 14.6% of the territory, respectively. Elevation ranges from sea level to 2,017 m (Mount Ori, Pyrenees). Three biogeographic regions converge in the area, including Atlantic, Mediterranean and Alpine, roughly corresponding to three sections with varying landscape compositions: the northern Basque Country, mainly covered by forestry plantations and highly fragmented natural forests; the north-eastern Navarre, dominated by continuous natural forest systems; and extensive cultivated lands and urban areas located primarily in the southern areas of both provinces (Fig. 1).

Marten presence data: combining long-term non-invasive genetic sampling and species records from different field data sources

Two main sources of PO data were used in this study to accurately assess the spatial distribution of both species. First, we used faecal sampling to collect non-invasive genetic samples from both species across the study area between 2005 and 2012. As *Martes* sp. faeces cannot be distinguished from each other visually and can also be easily confused with those of other carnivores (Davison et al 2002; Ruiz-Gonzalez et al 2008), we applied a mtDNA molecular method that effectively identifies the species (Ruiz-González et al. 2008). Additionally, we included unequivocal species records from road-killed, live-trapped, hair-trapped and/or camera-trapped individuals obtained in the framework of different carnivore surveys funded by regional or national administrations (Spanish Ministry of Environment, Regional Governments of Navarre and Basque Country and Alava and Bizkaia Provincial Councils; Table S1). The locations of the genetically identified faeces and those from different sources were combined to build a single PO dataset for each study species (Table S1).

Environmental layers

The environmental variable set comprised a total of 40 variables belonging to 5 categories (Table 1). All variables were resampled to a UTM projection (ETRS89) with a 30 m cell size.

Nine land cover types were derived from the land use information obtained in vector format from the Spanish Forest Map at scale 1:50.000 (Spanish Ministry of Agriculture, Food and Environment, 2006). FRAGSTATS software v 4.2 (McGarigal et al 2012) was used to calculate a) five landscape level metrics characterizing landscape composition, configuration and edge contrast, and b) four class composition and configuration level metrics (Table 1). Elevation data was obtained from a 25m resolution Spanish Digital Elevation Model (Spanish Geographical National Institute, CNIG; 2008) to calculate the Focal Mean Elevation (ELEV), Compound Topographic Index (CTI) and Roughness (ROUGH) using the Surface Gradient and Geomorphometric Modeling tool (Evans et al. 2014) in ArcGIS v.10.0 (ESRI). The density of highways (HWS), national (NAT) and autonomic roads (AUT), were assessed separately derived from the 1:25.000 scale spatial products developed by CNIG to incorporate the antrophic disturbance. To test the influence of temperature in each marten distribution the mean annual temperature (Iberian Climate Atlas; AEMET 2011) was included among the analyzed variables.

Each variable (Table 1) was independently tested at 6 scales using circular windows with radii of 1, 2, 4, 8, 16 and 32 km. HWS, NAT, AUT, ELEV, ROUGH and CTI were additionally tested at high-resolution scale (125, 250 and 500 m). These radii where chosen to encompass the diversity of home-range sizes and the variety and length of daily movements of both martens previously recorded in the literature (e.g. *Martes martes*, Zalewski et al 2004; *Martes foina*, Herrmann 1994).

Marten HSMs

The main methodological steps followed to build the marten HSMs, and to explore the ecological niche divergence among them, were summarized in a workflow (Fig. 2), while details of the steps are reported bellow.

Step 1) Variable pre-selection: bivariate scaling and variable pruning.

Habitat selection is hierarchical and occurs at multiple spatial scales (Johnson 1980; Wiens et al 1987). Therefore, we conducted an initial bivariate scaling step to test at which scale each variable was most strongly related to the species presence (Figure 2, Step 1). In this step, each variable (Table 1) at each scale and species was independently run in MAXENT (Phillips et al 2006), the most widely used HSM software for predicting species distributions from PO data and a set of environmental variables (Fourcade et al 2014). According to Elith et al (2010) MAXENT produces robust results with sparse, irregularly collected occurrence

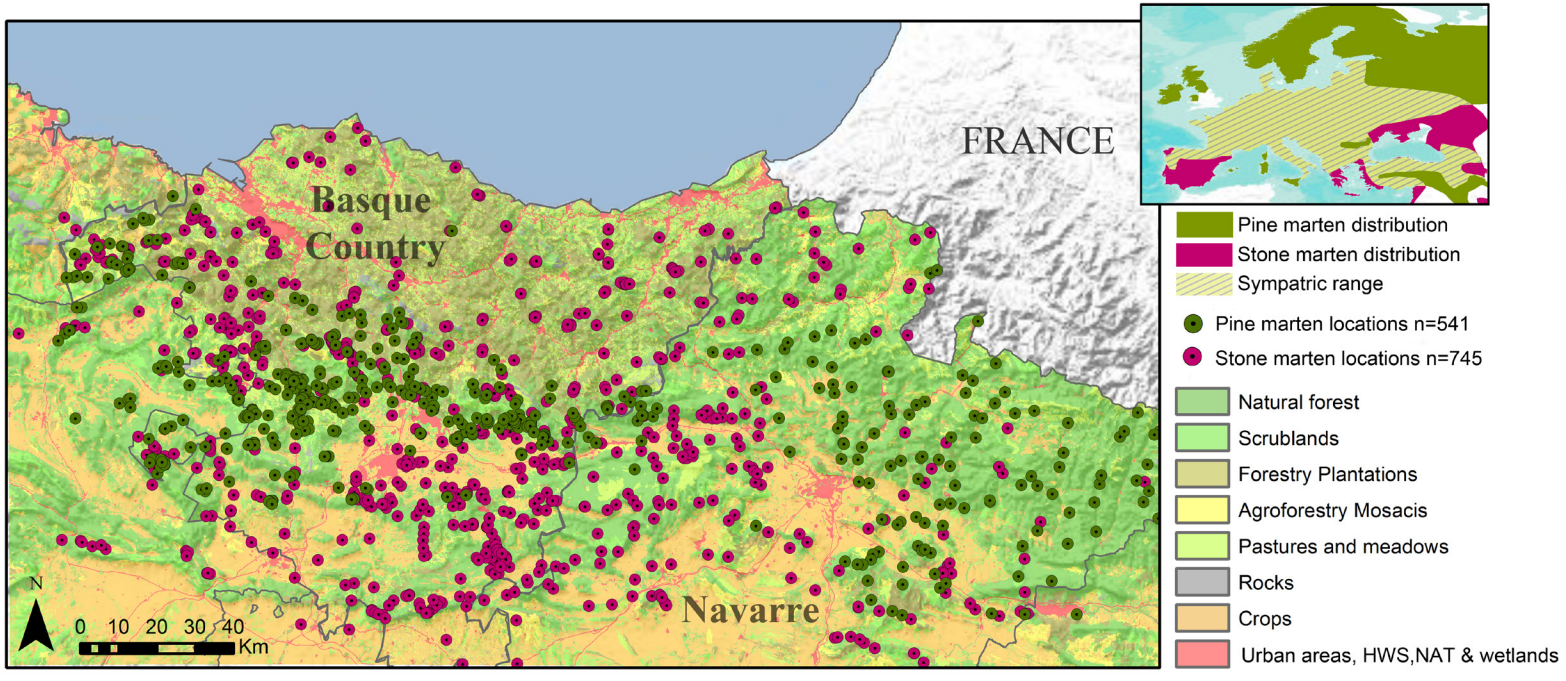


Figure 1. Map of the study area showing the distribution of the pine and stone marten records used to build the models. Each land cover type is shaded in a different color. The inset presents the species ranges and overlap in Europe.

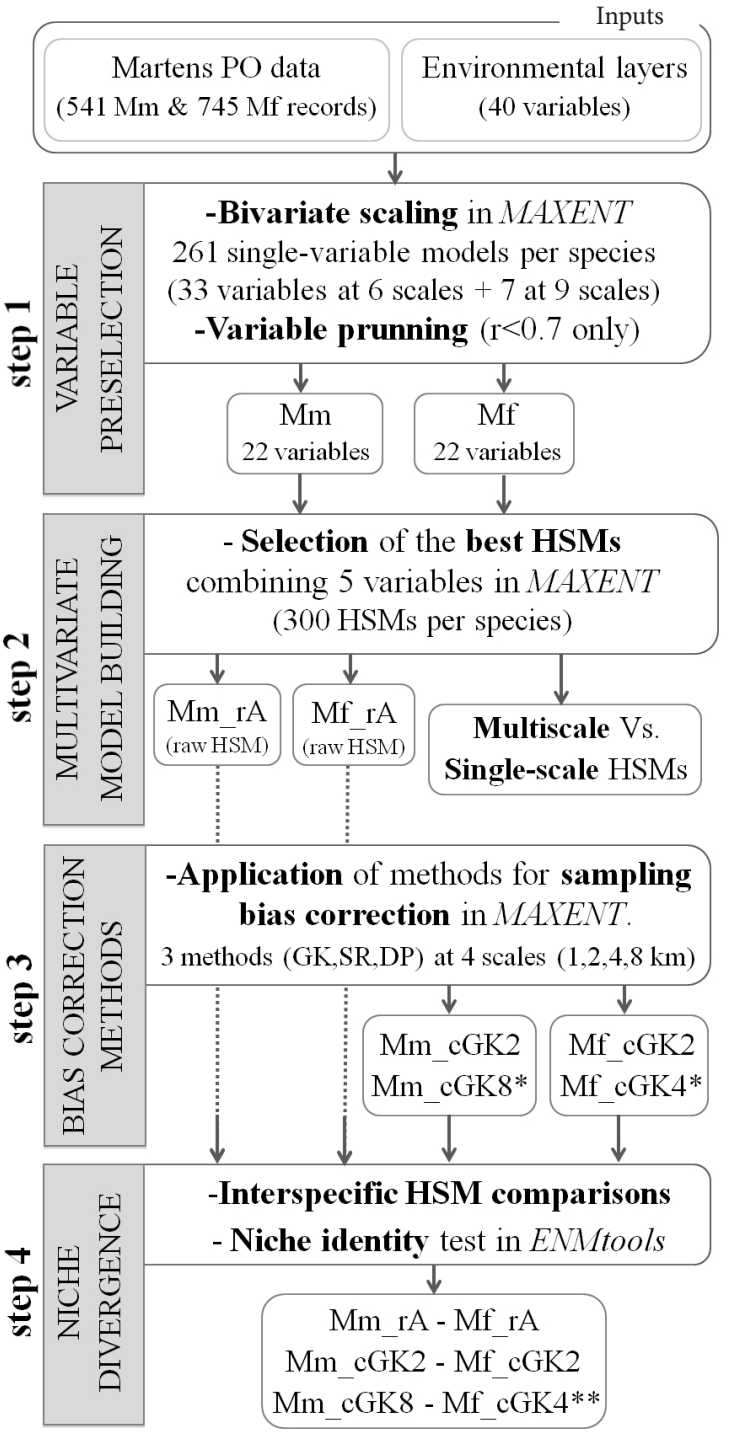


Figure 2. Study workflow including the input files used and the consecutive steps done to build the HSMs and to test martens´ niche divergence.

VARIABLE TYPE	METRIC	ABBREVATION
HUMAN PRESSURE	Highways density	HWS
	National road density	NAT
	Autonomic roads density	AUT
TOPOGRAPHIC	Focal mean of elevation	ELEV
	Roughness	ROUGH
	Compound Topographic Index	CTI
CLIMATIC	Mean Anual Temperature	TEMP
LANDCOVER (Landscape-level metrics)	Aggregation index	AI
	Contrast-weighted Edge Effect	CWED
	Edge density	ED
	Patch density	PD
	Shannon Diversity Index	SHDI
LANDCOVER (Class-level metrics)	Patch density	PD_ (Nat,Fp,Ag,Pa,Scr,Cr,Urb)
	Area-weighted mean	AREAam_(Nat,Fp,Ag,Pa,Scr,Cr,Urb)
	Gyrate_am	GYR_(Nat,Fp,Ag,Pa,Scr,Cr,Urb)
	Percentage of landscape	PLAND_(Nat,Fp,Ag,Pa,Scr,Cr,Urb)

Table 1. The set of 40 independent variables considered for analysing multiscale martens habitat suitability. Variables are grouped into 5 categories (Human pressure, topographic, climatic, and landcover landscape-level and class-level metrics). The 4 class level variables (PD, AREAam, GYR, PLAND were calculated for each of the 7 land cover types considered. All layers were produced at 6 different spatial scales (1, 2, 4, 8, 16 and 32 km) except for HWS, NAT, AUT, ELEV, ROUGH, CTI that were additionally analyzed at high-resolution scale (125, 250 and 500 m). In landcover class-level metrics: Nat: natural forest; Fp: forestry plantations; Ag: agroforestry mosaics; Pa: pastures; Scr: scrublands; Cr: crops and Urb: urban areas.

records, and minimal location errors (Elith et al 2006). Thus, all models were computed in MAXENT v.3.3 using the following parameters as described in Mateo-Sánchez et al (2013): 20,000 background points, a maximum of 5,000 interactions linear and quadratic functional relations (LQ). We selected LQ only as they prevent from locally idiosyncratic responses and are easier to interpret from an ecological perspective (Syfert et al 2013). A random subsample of 75% of the species occurrence points were used to fit the model ("train") and the remaining 25% to assess model performance ("test"). We used the AUC values (Area Under

the receiver operating characteristic Curve, Fielding and Bell 1997) to compare the performance of the single variable models, selecting the scale at which the variable showed the highest AUC value and discarding the rest (e.g. Mateo-Sánchez et al 2013). We used AUC (Phillips et al 2009; Aguirre-Gutiérrez et al 2013) as the scaling criterion as it is the most commonly used metric for model quality assessment (Kramer-Schadt et al 2013). MAXENT performs well in the presence of correlated variables; however, it remains desirable to remove highly inter-correlated variables in multivariate analysis to avoid multicollinearity (Mateo-Sánchez et al 2013) Hence, the candidate set of variables was pruned, discarding the variable (at the best performing scale) with the lowest AUC of each variable pair presenting a Pearson correlation coefficient ≥ 0.7 (Bellamy et al 2013).

Step 2) Multivariate model building

Only those variables remaining after the bivariate scaling and variable pruning steps were included in the multivariate models. 300 HSMs per species were constructed, each containing a different set of 5 randomly selected predictors, including variables belonging exclusively to the same category and combinations of two or more categories (Table 1). Multivariate HSMs, built combining 5 variables were run in MAXENT with the same parameters used in the bivariate scaling step.

To evaluate model performance, each species' multivariate models were ranked according to their AUC values while the relative predictive power of each variable was assessed based on the jackknife measure of test AUC. Then, we explored the 10 top performing models for each species (i.e. those showing the highest AUC values) to compare the incidence of the environmental predictors and select the best raw multivariate HSM for each species (i.e. Mm_rA and Mf_rA). Finally, to estimate how the scale optimization affected the predictive performance of the HSMs, we compared the best performing multiscale models (Mm_rA and Mf_rA) with those built with the same variables but measured at each single scale (i.e. 1, 2, 4, 8, 16 and 32 km).

Step 3) Sampling bias correction methods

One limitation of HSMs, especially when occurrence localities are derived from opportunistic observations rather than representative surveys, is the presence of sampling bias, where some areas in the landscape are sampled more intensively than others (Phillips et al 2009). When information quantifying sampling effort exists, it can be used to correct for sampling bias. However, in empirical studies this information is often unknown (Fourcade

et al 2014). Thus, we implemented three of the correction methods applied in Fourcade et al (2014), to address the most plausible sources of sampling bias in our dataset. We used SDM toolbox v.1.1 (Brown 2014) for ArcGIS v.10.0 (ESRI, 2014) to: a) rarefy the species occurrence data (SR); b) to produce sampling probability maps (GK); and c) to restrict the background area (DP) at four spatial scales.

a) SR: With the aim of eliminating the influence of spatial clusters of localities, the *spatially rarefy occurrence data tool* was used to reduce the locality records to a single point within the specified Euclidian distance. The resulting datasets for each marten, with filtered locations at 1, 2, 4 and 8 km radii, were later used as each species' presence records. Then, we ran MAXENT to produce 4 HSMs per species combining the spatially rarefied locations and the environmental predictors included in the best HSMs.

b) GK: We produced a bias grid that up-weighted PO data points with fewer neighbors in the geographic landscape using the *Gaussian Kernel Density of sampling localities tool*. This sampling probability surface (showing values of 1 to reflect none and higher values representing increasing bias) was computed including both marten species' locations to focus on the widespread spatial sampling biases as the probability of detecting either species is, *a priori,* the same. Each derived Gaussian Kernel map (at 1, 2, 4 and 8 km) was implemented in the bias file option in MAXENT together with the environmental predictors included in each marten best HSMs.

c) DP: The selection of a large rectilinear region, can led to the selection of a higher proportion of less informative background points (pseudo-absences), and therefore, models tend to be over-fitted (Anderson and Raza 2010; Barbet-Massin et al 2012). To circumvent this problem, we used the *sample by distance from observation point tool* which tells MAXENT to sample background points within a maximum buffer size of 1, 2, 4 and 8 km from known occurrences. Each background restriction mask was implemented in the bias file option in MAXENT.

Ten replicates per model (both raw and 12 corrected HSMs per species) were built following a resampling method, randomly selecting a subsample of 25% of observation records for model validation in each replicate.

To compare and evaluate the corrected HSM´s fit, we employed two threshold-independent and one threshold-dependent metric. First, to assess differences in the general model fit we used the AUC_{TEST} which, although criticized (Lobo et al 2008; Warren and Seifert 2011),

is appropriate for comparison of models produced with different settings but for the same species in the same study region, as in this case (Lobo et al 2008; Anderson and Gonzalez 2011). Then, to quantify overfitting, we employed AUC_{DIFF}, the difference between AUC values based on training (AUC_{TRAIN}) and test (AUC_{TEST}) localities (Warren and Seifert 2011). Complementarily, we assessed the spatial patterns of actual presences that are correctly predicted (sensitivity) *versus* actual absences that are correctly predicted (specificity). For this purpose, we chose as probability threshold MAXENT´s maximum training sensitivity plus specificity logistic threshold (MTR), which has been shown to produce accurate predictions (Jiménez-Valverde and Lobo 2007; Svenning et al 2008). Thus, to investigate the spatial pattern of true presences (TP) and false absences (FA) models were converted to binary maps (presence/absence) selecting the MTR of each model as a cut-off (Syfert et al 2013). Additionally, raw and corrected HSMs were visually examined and evaluated based on expert knowledge on the distribution of each species and the habitat types in with they are known to occur to ensure reliability and to identify the best performing final HSMs (e.g. Radosavljevic and Anderson 2014; Brown 2014).

Step 4) Marten niche divergence

First, to explore spatial niche separation, we calculated and visually examined the spatial difference between occurrence probability values of the two species' for the different HSMs: a) raw HSMs, and b) the corrected HSMs according to two different criteria i) those outperforming the raw in all aforementioned criteria and, ii) those HSMs with the lowest percentage of FA (Figure 2, step 4)

Further, to statistically test if models produced for each species were more different than expected by chance, we computed an identity test in ENMtools v.1.4 (Warren et al 2008; Warren et al 2010), for the aforementioned three comparisons, by comparing the observed measure of niche overlap (Schoener's D and Hellinger`s I metrics) to a null distribution calculated with 100 replicates, in which the null hypothesis of niche identity is rejected when the empirically observed value is significantly different from the pseudo-replicate data set (Warren et al 2010).

Results

Marten's presence data

From the 972 faecal samples collected, we selected 899 samples that were the freshest and highest quality for genetic analysis. 754 were genetically identified as one of the two marten species (Table S1). Thus, unequivocal species identification was possible in 83.87% of the samples. In the remaining 16.12%, the DNA extracted was not amplified by the primers used. The results of non-invasive genetic sampling together with the combination of unequivocal species records obtained in the framework of different carnivore surveys (n=532), resulted in a PO dataset with 541 pine marten and 745 stone marten locations to build the HSMs.

Bivariate scaling and variable pruning

The predicted relative habitat suitability, for each environmental variable at each scale and species, revealed high sensitivity of habitat relationships to the scale at which habitat variables were measured (Table S2). Comparisons among different scales revealed that most metrics were most strongly related to marten's habitat suitability at broad-scales (16-32 km, 48.75%), including all three categories of road density and the majority of landcover class-levels metrics across cover types. A significant difference between the stone and the pine marten was found regarding landcover landscape-level metrics (AI, CWED, ED and PD). While variables measured at broad scale performed best for the pine marten, variables at medium-scales (4-8 km) did better for the stone marten. Patch density (PD) for 6 of the 7 landcover classes (except for patch density of natural forest, PD_Nat) had best performance at fine-scales (1-2 km) in both species. The CTI was the only metric selected at a high-resolution scale (125-500 m). After pruning, 22 variables remained from the initial set of 40, 18 of which were shared by both species at diverse scales and 4 were unique to each species (Table S2).

Performance of multivariate HSMs

Among the 10 best performing multivariate models for each species (Table 2) the HSMs produced for the pine marten showed significantly higher discrimination ability, as measured by AUC, (AUC= 0.818-0.835) than those for the stone marten (AUC= 0.720-0.724). The best predictions of both species´ habitat suitability were produced using variables measuring

SP	MODEL	V. CAT	VARIABLES					AUC
PINE MARTEN	Mm_rA	HP,CL,LC	PLAND_Urb 32km	PLAND_NAT 2km	TEMP 4km	Gyr_Ag 32km	PD_Fp 2km	0.835
	Mm_rB	HP,CL,LC	PLAND_Urb 32km	PLAND_Cr 16km	TEMP 4km	Gyr_Ag 32km	PD_Fp 2km	0.833
	Mm_rC	HP,CL,LC	PLAND_Urb 32km	TEMP 4km	Gyr_Ag 32km	PD_Fp 2km	AUT 8km	0.832
	Mm_rD	HP,CL,LC	PLAND_Urb 32km	PLAND_NAT 2km	TEMP 4km	Gyr_Ag 32km	AUT 8km	0.828
	Mm_rE	HP,CL,LC	PLAND_Urb 32km	PLAND_Cr 16km	TEMP 4km	Gyr_Ag 32km	PD_Pa 1km	0.828
	Mm_rF	HP,CL,LC	PLAND_Urb 32km	TEMP 4km	Gyr_Ag 32km	PD_Fp 2km	HWS 32km	0.828
	Mm_rG	HP,CL,LC	PLAND_Urb 32km	PLAND_NAT 2km	TEMP 4km	Gyr_Ag 32km	PD_Urb 2km	0.826
	Mm_rH	HP,CL,LC	PLAND_Urb 32km	PLAND_NAT 2km	TEMP 4km	Gyr_Ag 32km	PD_Ag 1km	0.825
	Mm_rI	HP,CL,LC	PLAND_Urb 32km	PLAND_Cr 16km	TEMP 4km	Gyr_Ag 32km	PD_Cr 1km	0.825
	Mm_rJ	HP,CL,LC	PLAND_Urb 32km	PLAND_Cr 16km	TEMP 4km	Gyr_Scr 16km	PD_Fp 2km	0.818
STONE MARTEN	Mf_rA	HP, LC	PLAND_Urb 32km	PLAND_Cr 4km	PLAND_Scr 32km	AREAam_Cr 32km	PD_Pa 2km	0.724
	Mf_rB	HP, LC	PLAND_Urb 32km	PLAND_Cr 4km	PD_Urb 2km	PD_Pa 2km	PLAND_Scr 32km	0.723
	Mf_rC	HP,CL,LC	PLAND_Urb 32km	PLAND_Cr 4km	PD_Pa 2km	PLAND_Scr 32km	TEMP 16km	0.723
	Mf_rD	HP,LC	PLAND_Urb 32km	PLAND_Cr 4km	PD_Urb 2km	AREAam_Cr 32km	PLAND_Scr 32km	0.723
	Mf_rE	HP,LC	PLAND_Urb 32km	NAT 16km	PD_Pa 2km	PLAND_Scr 32km	PLAND_Cr 4km	0.723
	Mf_rF	HP,CL,LC	PLAND_Urb 32km	PLAND_Cr 4km	AREAam_Cr 32km	PD_Pa 2km	TEMP 16km	0.722
	Mf_rG	HP,LC	PLAND_Urb 32km	NAT 16km	PD_Urb 2km	PLAND_Scr 32km	PLAND_Cr 4km	0.721
	Mf_rH	HP, LC	PLAND_Urb 32km	PLAND_Cr 4km	PD_Pa 2km	PLAND_Scr 32km	SHDI 2km	0.721
	Mf_rI	HP,CL,LC	PLAND_Urb 32km	PLAND_Cr 4km	PD_Urb 2km	PLAND_Scr 32km	TEMP 16km	0.721
	Mf_rJ	HP,CL,LC	PLAND_Urb 32km	PLAND_Cr 4km	AREAam_Cr 32km	PLAND_Scr 32km	TEMP 16km	0.72

Table 2. Ten best performing raw multivariate models for the pine marten (Mm_r) and the stone marten (Mf_r) ranked by AUC values (A-J). Variables are ordered according to their relative contribution to the model. The scale at which each predictor showed the best performance is reported immediately after each variable. V.CAT: Variable categories: HP: Human-pressure, CL: Climate, LC: Landcover.

human-pressure (NAT, AUT, HWS), landcover metrics (SHDI, PLAND_Nat,Scr,Cr,Urb; PD_Fp,Ag,Pa,Cr,Urb; AREAam_Cr) and climatic (TEMP) metrics. Topographic variables (Rough and CTI) were not included in any high-performing models.

The most important predictor of both species' occurrence was the percentage of urban areas (PLAND_Urb). The density of urban patches (PD_Urb) was also an important predictor. The influence of the human impact was further evidenced by the inclusion of both the percentage and the density of crops patches in all the stone marten models (PLAND_Cr, PD_Cr) and the correlation length of agroforesty mosaics (GYR_Ag) in 9 out of 10 pine marten models. Although to a lesser extent, the three types of road density metrics (HWS,

AUT and NAT) were also included. The percentage of natural forested areas (PLAND_Nat) and patch density of forestry plantations (PD_Fp) were highly related to pine marten habitat suitability. Temperature was present in each of the pine marten`s best performing models, whereas the area covered by scrublands (PLAND_Scr) was associated with the presence of the stone marten.

We found significant differences in predictive performance between the multiscale HSMs with the highest AUC (Mm_rA and Mf_rA, Table 2) and the corresponding single-scale models at 1, 2, 4, 8, 16 and 32 km (Fig. S1). Multiscale models showed higher discrimination ability, as evidenced by the AUC values. Among the single-scale models, those including all variables measured at 1 km showed the weakest performance in the pine (AUC_{MM}=0.790) and the stone (AUC_{MF}=0.653) marten. The maximum performance for a single scaled model was archived at 8 km for the pine marten (AUC_{MM}=0.810) and at 32 km for the stone marten (AUC_{MF}=0.702) but far weaker than the optimized multiscale HSMs (AUC_{MM}=0.835; AUC_{MF}=0.724, Fig. S1).

Bias correction methods in marten HSMs

No correction method (GK, SR, DP) or scale (1, 2, 4, 8 km) ranked best based on the three evaluation criteria ($AUC_{TEST,}$ AUC_{DIFF}, % false absences; Table 3) for both martens. For the pine marten, only corrected models run with the Gaussian Kernel bias files (Mm_cGK) presented higher AUC values than the raw model. Mm_cGK and Mm_cDP presented lower overfitting (AUC_{DIFF}) in contrast to what was observed for the Mm_cSR models. Regarding model sensitivity (% of actual presences correctly predicted) the best results were obtained for Mm_cGK and Mm_cSR models, while no Mm_cDP model showed higher values than the raw model. Only Mm_cGK2 outperformed the raw model in all pine marten comparisons, presenting higher AUC, smaller AUC_{DIFF} values and slighter FA rate. Within the stone marten models, only corrections with Mf_cGKs improved the predictive performance and half reduced the overfitting. The proportion of FA was reduced in almost all Mf_cDP and Mf_cGK models but not in any Mf_cSR model. Mf_cGK1 and Mf_cGK2 were the only outperforming the raw model in all comparisons.

The distributions of TP and FA of a) raw HSMs (Mm_rA and Mf_rA), b) corrected HSMs i) outperforming the raw HSMs in all aforementioned criteria (Mm_cGK2 and Mf_cGK2) and ii) corrected HSMs with the lowest percentage of FA (Mm_cGK8 and Mf_cGK4) were visually explored (Fig. S2). Overall, correcting for geographic sampling bias led to a drop in

the actual locations that were predicted as absences (%FA, Table 3) and an increase in the area available for each species (%AREA, Table 3). In the pine marten models, no significant increase in %FA was found in the Basque Country area (Fig. S2). In Navarre (n=134), however, FA were reduced by 21.6% (Mm_cGK2) and 74.7% (Mm_cGK8). Correcting for sampling bias also resulted in a smaller decease in FA in the stone marten (3.6%, Mf_cGK2 and 17.8% Mf_cGK4).

Raw and corrected multiscale HSMs

Three final habitat suitability maps are presented for the pine marten (Fig. 3). The first is the raw Mm_rA, built with PLAND_Urb, PLAND_Nat, Temperature, GYR_Agr and PD_Fp variables at the best performing scale (Table 2) without spatially rarefying occurrence records or applying any bias file. Based on this model, we can clearly distinguish a region of mostly continuous natural forested area and forestry plantations from 800 to 1500 m showing the highest probability of presence for the pine marten. A smaller patch on the west was also identified as an area of elevated probability. The optimal area of the pine marten was well-delimited with the probability decreasing sharply (Fig. 3). The second and third maps correspond to the corrected Mm_cGK2 and Mm_cGK8 models, built using the same variable set, but including a bias file with sampling probabilities based on Gaussian Kernels at 2 and 8 km, respectively. Accounting for sampling bias resulted in an increase in the percentage of optimum area available for the species (Mm_cGK2: 24.65%; Mm_cGK8: 39.32%) and a homogenization of the differences in probability. This resulted in the recognition of the Pyrenean range (near the French border) as a very suitable area for the pine marten, accurately reflecting the known pattern of occurrence. Thus, the Mm_cGK8 model was selected as the optimal pine marten HSM due to its more realistic predictions.

Stone marten HSMs showed, on average, lower predictive performance than the pine marten and no areas of AUC>0.8 were identified (Table 3). Stone marten models were built with PLAND_Urb, PLAND_Cr, PLAND_Bus, AREAam_Cr and PD_Pa variables at the best performing scale. Habitat suitability values decreased gradually showing a wide gradient from very optimal (0.7-0.8) to optimal (0.5-0.7) and suboptimal conditions (<0.35). Although to a lesser extent, accounting for sampling bias also resulted in an increased in the suitable range for the stone marten (Table 3). As a result, based on the Mf_cGK4 model, the species could be more likely found in an extensive and continuous central region (Basque Country and north-western Navarre). Only few localities bordering the study extent are predicted to be non-suitable for the stone marten. The Mf_cGK4 model, which presented the lowest FA

SP	MODEL	N	AUC_{TEST}	AUC_{DIFF}	MTS	% FA	% TP	% AREA
PINE MARTEN	Mm_cDP1	541	0.698	0.012	0.4759	34.38	65.62	35
	Mm_cDP2	541	0.735	0.013	0.4444	34.01	65.99	31.49
	Mm_cDP4	541	0.77	**0.007**	0.438	33.83	66.17	24.55
	Mm_cDP8	541	0.798	**0.005**	0.4542	35.49	64.51	20.96
	Mm_cGK1	541	**0.813**	**0.004**	0.4832	32.9	67.1	19.37
	Mm_cGK2	541	**0.807**	**0.004**	0.4937	**29.21**	70.79	24.65
	Mm_cGK4	541	0.801	**0.005**	0.4931	**26.06**	73.94	29.2
	Mm_cGK8	541	0.772	**0.006**	0.4339	**19.22**	80.78	39.32
	Mm_cSR1	322	0.755	**0.011**	0.4447	32.61	67.39	29.54
	Mm_cSR2	241	0.743	0.013	0.4177	**19.92**	80.08	36.55
	Mm_cSR4	138	0.702	0.032	0.3942	**23.19**	76.81	43.26
	Mm_cSR8	62	0.67	0.055	0.4542	33.87	66.13	38.44
	Mm_rA	541	0.804	0.011	0.4234	31.42	68.58	19.04
STONE MARTEN	Mf_cDP1	745	0.586	**0.004**	0.484	35.03	64.97	53.15
	Mf_cDP2	745	0.601	0.013	0.4546	**22.42**	77.58	60.09
	Mf_cDP4	745	0.645	0.022	0.4372	**23.89**	76.11	50.06
	Mf_cDP8	745	0.706	**0.005**	0.4324	**23.76**	76.24	43.61
	Mf_cGK1	745	**0.722**	**0.002**	0.4517	**23.36**	76.64	39.84
	Mf_cGK2	745	**0.72**	**0.008**	0.47	**24.7**	75.3	42.67
	Mf_cGK4	745	0.682	0.022	0.4879	**21.07**	78.93	45.9
	Mf_cGK8	745	0.641	0.011	0.4922	30.6	69.4	47.01
	Mf_cSR1	466	0.694	**0.003**	0.4502	30.47	69.53	40.37
	Mf_cSR2	355	0.663	0.012	0.4551	28.45	71.55	44.39
	Mf_cSR4	232	0.652	0.014	0.4595	33.19	66.81	44.57
	Mf_cSR8	99	0.583	0.081	0.4909	49.49	50.51	36.37
	Mf_rA	745	0.718	0.009	0.4294	25.64	74.36	38.77

Table 3. Performance of three correction methods (DP: distance to points, GK, Gaussian kernel, SR: spatial rarefaction at 4 scales for the pine (Mm_c) and the stone marten (Mf_c). "SP" refers to the species, "N" to the number of samples included in the model and "MTRS" to the Maximum training sensitivity plus specificity logistic threshold from MAXENT, "%FA" to the percentage of false absences, "%TF" to the percentage of true presences and "%AREA" to the percentage of area above the threshold. (e.g. Mm_cDP1 refers to the corrected HSM for the pine marten using distance to points method at 1km scale). Values outperforming those for the raw models are reported in bold.

rate and which showed the highest consistency with existing knowledge of the stone marten distribution, was selected as the stone marten optimized HSM.

Environmental predictors of marten's occurrence

Based on the top performing corrected HSM for the pine marten (Mm_cGK8) the most important variable (31.2% contribution) was the proportion of the landscape covered by urban areas measured at 32 km, showing a unimodal relationship peaking at 13% cover by urban land uses (Fig. 4). Other important predictors were the percentage of natural forests and temperature (26.1% and 20.3%, respectively). Occurrence probability increased until the area occupied by natural forest reached its maximum (100%, based on an area with a radius of 2 km) revealing a clear preference for this cover type. The temperature curve showed the opposite, depicting a dramatic drop in probability of occurrence as temperature rose. GYR_Agr provided a measure of landscape continuity of the agroforestry mosaics, and was positively correlated to the species presence at finers scales (200-800 m) and contributed 18.6% to the model. The contribution of the patch density of forestry plantations (PD_Fp) at wide scale (32 km) was slight (3.8%) but positively related to the pine marten ocurrence up to intermediate values.

The most important variable (contribution of 39.2%) of the best corrected HSM for the stone marten, (i.e. Mf_GK4), was the percentage of landscape covered by urban areas measured at 32 km, with the response curve showing suitable areas between 7 and 24% and with maximum probability at 15% (Fig. 4). The second variable in terms of importance (18.6%) was the percentage of crops at 4 km (PLAND_Cr) with probability maximal at 20%. The stone marten was more likely to be found in areas with a very low density of scrublands (<6%; PLAND_Scr) and in areas dominated by moderately large cropland patches (with an optimum size around 0.035 km^2; AREAam_Cr) both at the broadest scale evaluated (32 km; Fig. 4). Stone marten occurrence was additionally influenced (7.5%) by the patch density of pastures (PD_Pa), with occurrence probability decreasing as the density of pasture patches increased at fine-scale (2 km).

Martens niche divergence

The spatial differences in marten's occurrence probability values revealed their divergence in habitat selection (Fig. 3). All three comparative maps revealed similar patterns, with the greener areas indicating a higher probability of finding pine martens than stone martens, and declining gradually with the growth in the likelihood of stone marten presence, shown

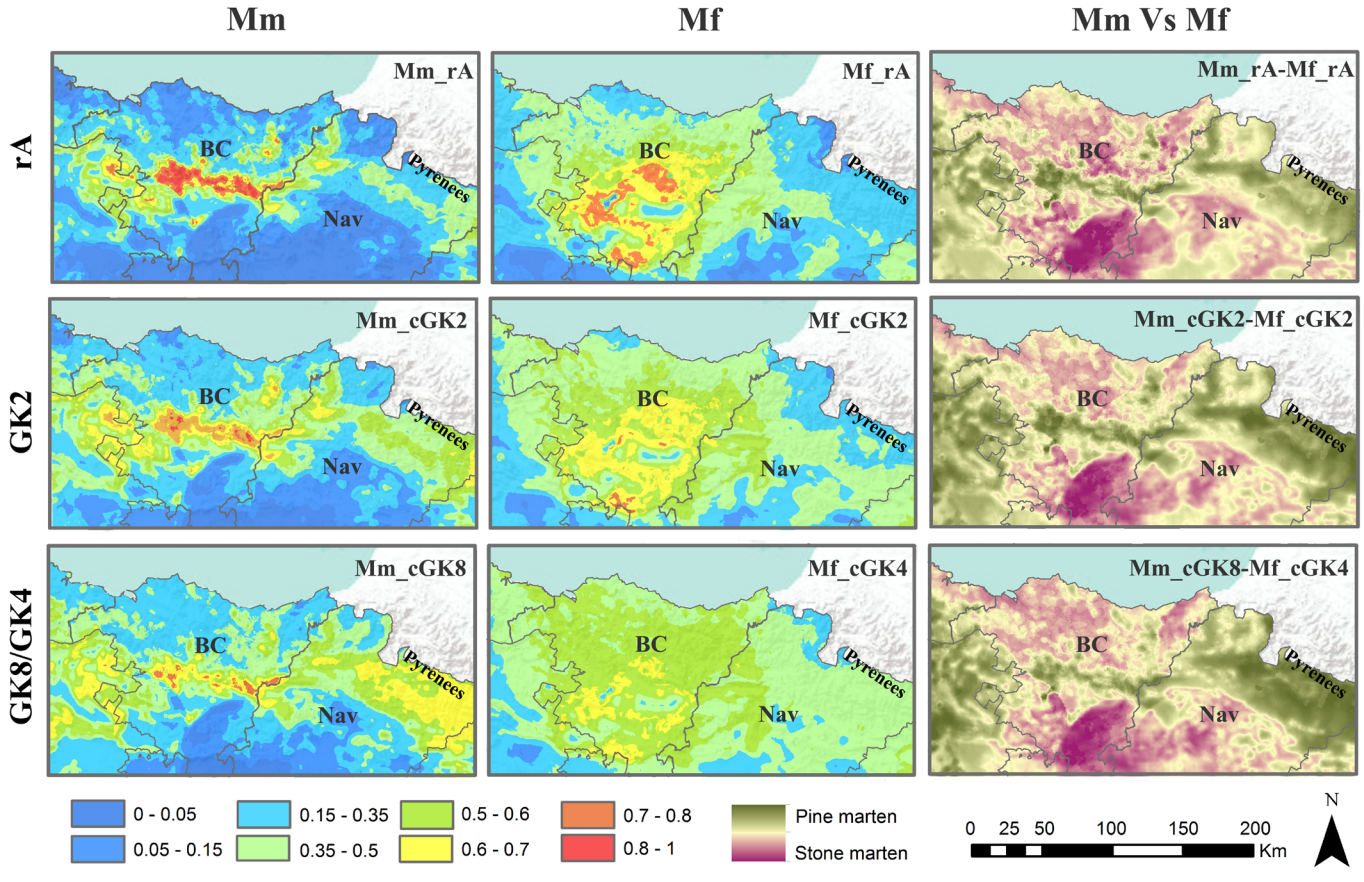


Figure 3. Habitat suitability maps produced by MAXENT for the pine and the stone marten for three model combinations (raw models: Mm_rA-Mf_rA, and corrected models: Mm_cGK2-Mf_cGK2 and Mm_cGK8-Mf_cGK4) averaged across 10 replicates. The Mm-Mf comparisons display the relative probability of each marten presence calculated from the subtraction of the stone marten probability of occurrence to the corresponding pine marten's model. These maps are represented in a gradient from the maximum probability for the pine marten (in green) to the maximum probability for the stone marten presence (in purple).

in purple (Fig. 3). The pine marten occurrence is predicted to be concentrated in the forests and scrublands of the western part of the study area, in the large and continuous natural forested area of northern Navarre and in the forested mountain ranges along the Cantabric-Mediterranean watershed boundary. On the other hand, the stone marten primarily selects the forestry plantations and valleys in the northern Basque Country and the extensive croplands in the south. The species was also found close to the cities where the probability of the pine marten was negligible. However, the stone marten is the only marten recorded in the natural forested areas of the more temperate southern region.

The Identity test further highlighted that the pine and stone marten niches were significantly different ($p < 0.01$) as indicated by the disparity between the null distribution (Fig. 5) and niche overlap values of both metrics observed for the comparative of the optimized HSMs (Mm_cGK8-Mf_cGK4; D=0.786 and I=0.859). The niche dissimilarity between martens was also detected for the Mm_rA-Mf_rA (D=0.709 and I=0.809) and Mm_cGK2-Mf_cGK2 (D=0.770 and I=0.851; Fig. 5) combinations.

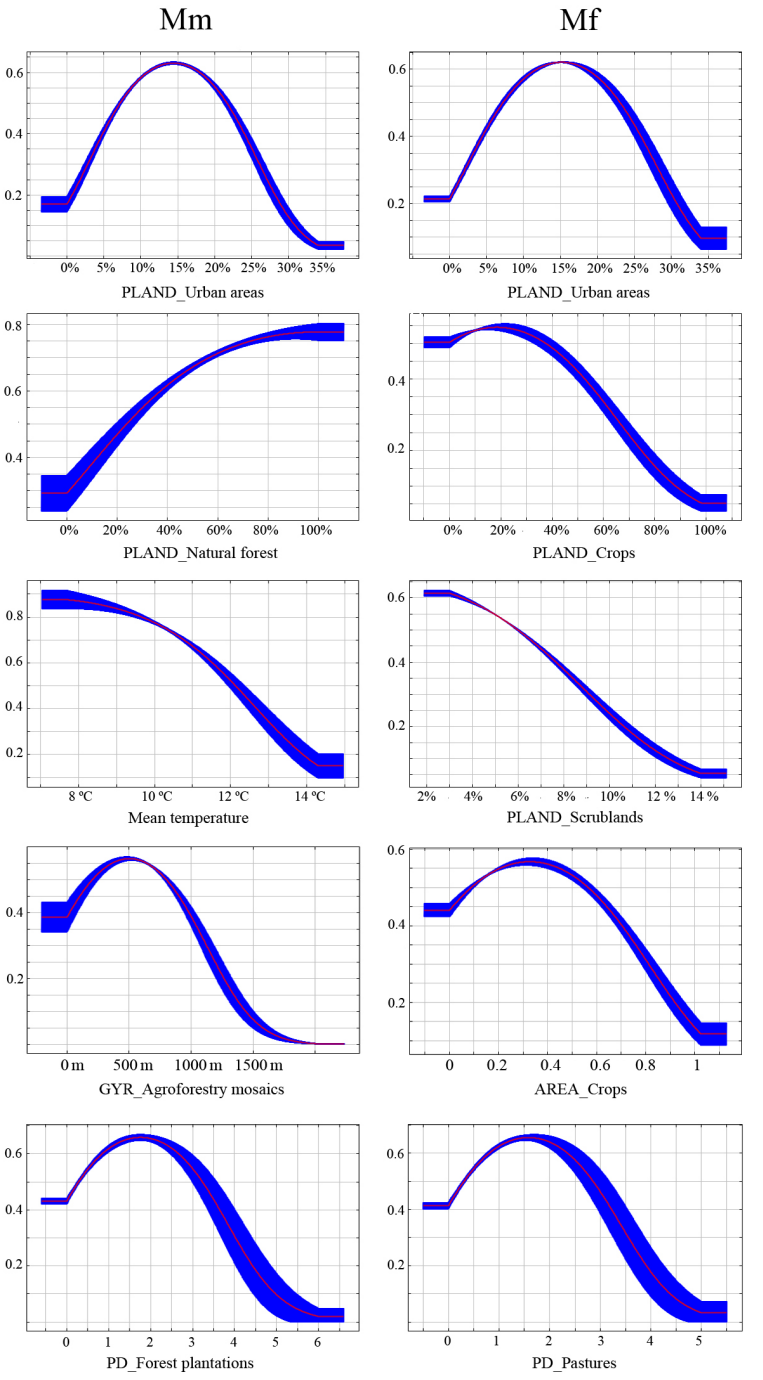


Figure 4. Response curves as estimated from MAXENT showing log response of each marten to five environmental predictors. Probability of presence (logistic output) is shown on the y-axis while the range of the environmental predictor is shown in the x-axis. To prevent that the interaction of variables affect the relationship modeled, the response curves are based on univariate models. Mean response of the 10 replicates is shown in red while the standard deviation is shaded in blue.

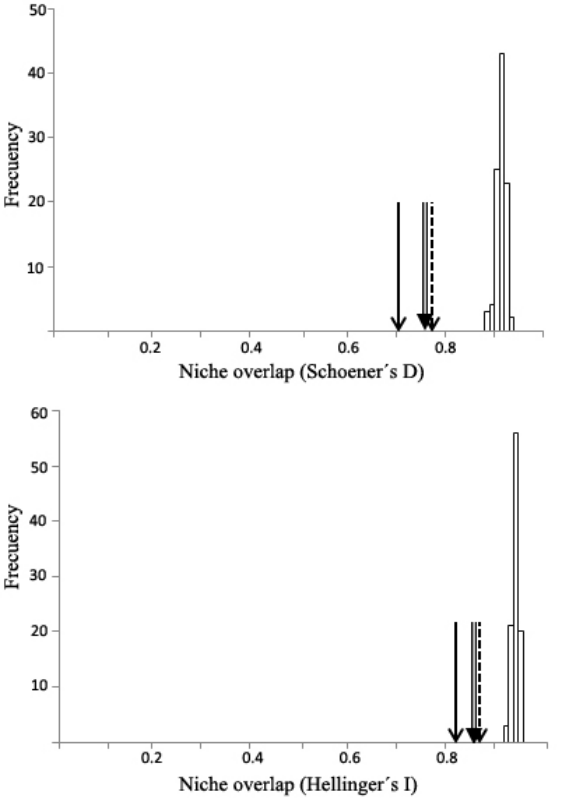


Figure 5. Result for the niche identity test according to a) Schoener's D metric. b) Hellinger`s I metric.. Martens measured niche overlap between species for the Mm_rA-Mf_rA, Mm_cGK2-Mf_cGK2 and Mm_cGK4-Mf_cGK8 HSMs combinations are reported with a simple, double and a dashed arrow respectively while the histogram in white illustrates the distribution of overlaps from the pseudoreplicates.

Discussion

Effect of scale in marten's habitat selection

Habitat selection of both pine and stone marten were determined by habitat predictors at divergent spatial scales. Fisher et al (2013) found that the spatial scale of habitat selection of a species was related to body size in 12 terrestrial mammals. The bigger the species, the larger the distance at which an animal would perceive landscape elements (i.e. its perceptual range). In this study, all six scales analyzed (1 to 32 km) were selected as the most predictive, depending on the habitat variable. Most variables were selected at 32 km (32.5%), indicating that martens respond to landscape features and structure at larger scales than predicted based on body size alone.

Interestingly, road density metrics (evaluated at 9 scales) improved their predictive performance at medium and high scales (8-32 km) despite what could have been expected if the inclusion of road casualties was significantly biasing the predictive performance of the species HSMs.

The multiscale approach outperformed the single-scale multivariate HSMs for both species, further supporting the scale dependence of the pine and stone marten habitat selection and corroborating previous studies conducted on martens (*Martes* spp. Bissonette and Broekhuizen 1995; *Martes americana,* Shirk et al 2012; Wasserman et al 2012; Shirk et al 2014) and other mammals, such as the brown bear (*Ursus arctos*, Mateo-Sánchez et al 2013) and several bat species (Bellamy et al 2013).

Environmental predictors shaping marten´s occurrence

The optimized HSM for the pine marten (i.e. MmcGK8), which showed that the species was positively associated with cooler areas with a small degree of human disturbance, a high proportion of natural forests, well-connected forestry plantations and medium-extent of agroforesty mosaics, is highly consistent with ecological knowledge about the species. For example, forested habitats are key features for all *Martes* species (Buskirk and Powell 1994; Zalewski and Włodzimierz 2006). Forest offer the best combination of abundant food resources, low risk of predation and well insulated denning sites (reviewed in Virgós et al 2012). In a recent study conducted in Italy, the distribution and density of pine marten scats (confirmed by DNA analysis) revealed their preference for woodlands (Balestrieri et al 2015)

while its abundance was found to be related to the structure and degree of fragmentation of residual woods.

In our study, as expected, both the percentage of natural forests (PLAND_Nat) and the presence of continuous forestry plantations (PD_Fp) were among the best predictors of the pine marten occurrence. The species was closely associated with natural forests, selecting the best preserved and connected mountainous forested regions. Brainerd and Rolstad (2002) found that forest structure was more important determinant of the pine marten habitat selection than forest composition and age. Concordantly, presence was also related to connectivity between woodlands (in this case, composed by plantation forests), with occurrence probability higher in unbroken patches than in fragmented forestry landscapes.

The pine marten has been persistently considered to be a forest specialist (Mergey et al 2012). Yet, recent studies indicate greater habitat flexibility, and it has been found in agricultural and fragmented landscapes (Balestrieri et al 2010; Virgós et al 2012; Balestrieri et al 2015). In this study, the extent of agroforestry mosaics was among the variables most related to the species´ occurrence. This finding is in agreement with Ruiz-González et al (2014) where, in addition to woodlands, pine marten gene flow was facilitated by agroforestry mosaics, which represent a transition from forest to agriculture. One possible explanation is that, in fragmented landscapes of temperate regions, the diversity of a wide range of habitat patches promotes higher diversity of food resources than continuous forests (Rosalino and Santos-Reis 2009). This "habitat-diversity" hypothesis has been shown to be related to elevated bird species diversity in moderately fragmented landscapes (e.g. Cushman and McGarigal 2004).

The optimized HSM for the stone marten (i.e. Mf_cGK4) clearly showed that the species' presence was conditioned by the density of urban areas, the proportion and extensiveness of croplands, the existence of some scrub cover and the availability of semi-continuous grasslands. In recent studies, the stone marten was more often detected in rural areas (74.8%) than in forested habitats (25.2%; Santos and Santos-Reis 2010). We also found a tendency of the species to select human associated environments, mostly villages surrounded by extensive agricultural areas. However, in the Iberian Peninsula the stone marten is not as synantropic (Delibes 1983; Rodrigues et al 2011) as in other parts of Europe (e.g. Germany, Luxemburg or Poland), where it is very common in suburban and urban areas, often denning and resting in buildings and barns and causing damage to roofs, insulation and car engines (Proulx et al 2004; Tikhonov et al 2008; Herr et al 2009). In addition to urban envi-

ronments, medium-sized croplands and small scattered pastures together with some shrub cover, which provide food, shelter and areas of lower predation risk (Buskirk and Powell 1994; Herrmann 1994), were the main determinants of stone marten habitat selection. Thus, its presence was better explained by the simultaneous use of different human dominated landcover types. This result is in concordance with those reported by Barrientos and Virgós (2006) and Santos and Santos-Reis (2010), where individuals followed a complementation/supplementation strategy with a temporal segregation of food resources, corroborating the importance of mosaic habitats for stone martens compared to strict forest habitats (Virgós et al 2000; Virgós and García 2002; Santos and Santos-Reis 2010).

Importance of bias correction methods in HSMs building

All datasets derived from opportunistic samplings are likely to suffer from geographic bias, which can strongly affect the predictive performance of the HSMs (Fourcade et al 2014). In addition, when dealing with common and widespread species, observations are frequently under-reported (Fourcade et al 2014). However, such data are often the only data available for many species, and must be used with caution, employing the most appropriate sampling bias correction.

To date, no consensus exists regarding the most appropriate metrics and thresholds for selecting from a candidate set of models (Lobo et al 2008; Warren and Seifert 2011; Kramer-Schadt et al 2013). In the raw HSMs, the areas predicted as the most suitable for each species tightly matched those with the highest density of species records, indicating effects of sampling bias. The application of GKs significantly smoothed the overfitting due to the clumped locations, increased the AUC, halved the FA in the pine marten HSMs, decreased error rates up to 20% in the stone marten, and improved the predictions of the raw models in the Pyrenean region (NE Navarre), tightly matching the areas where the pine marten has been documented to occur. Thus, GKs led to the most realistic potential ranges for these mustelids based on the knowledge of their distribution and the climatic and landcover variables in the area.

Unexpectedly, SR, identified as the most effective sampling bias correction method in recent papers (Kramer-Schadt et al 2013; Boria et al 2014; Fourcade et al 2014), performed poorly across all criteria and spatial resolutions and, in most cases, produced poorer HSMs than the raw ones. DP corrections were also among the less efficient correction methods, in congruence to results reported by Fourcade et al (2014). Based on the evidence, we strongly

recommend evaluating several correction methods and choosing the one that suits each particular species, sampling scheme and objective best.

Niche divergence in sympatric martens

Even if both mustelids are widespread through the area and coexist locally in some forested regions, the optimized HSM comparison (Mm_cGK8-Mf_cGK4) showed a clear spatial segregation and niche divergence, which was further supported by the results of the niche identity test. In agreement with Delibes (1983), Reig (1992) and Virgós et al (2000), we observed that the pine marten in northern Spain was most frequently found in forested landscapes whereas the stone marten was often associated with agricultural and suburban landscapes. The observed pattern could be a product of the interspecific competition among martens, favoring the slightly bigger pine marten, with the stone marten occupying higher quality forests when pine marten is absent and expressing niche displacement away from preferred pine marten habitats when co-occurring (Delibes 1983; Virgós and García 2002).

The smaller extent of high probability areas of the pine marten occurrence has been explained as a consequence of both the direct persecution during recent decades and the increasing effects of forest fragmentation and loss (López-Martín 2007). The stone marten, which shows a greater behavioral plasticity, would likely have colonized areas previously inhabited by the pine marten, increasing the isolation of its subpopulations.

In this study, conducted on the south-western edge of a Eurosiberian species range, temperature played an important role in driving pine marten distribution, which showed a clear preference for cooler and higher elevation environments. Hence, temperature can be considered a limiting factor determining the pine marten distribution across the area, constraining the species southward expansion and shaping the differences in habitat selection of the two species. Consequently, the actual distribution of the pine marten is likely to be further reduced under a climate change scenario, while an increase in temperature is expected to benefit the more thermophilic stone marten.

A comparison conducted in central Poland showed how these closely related species can coexist by differentiating in the use of three-dimensional space in forests and their response to open areas and transformed habitats (Goszczyński et al 2007). Recently, a telemetry based study revealed how the species differed in resting patterns (pine martens rested almost exclusively in forest while stone martens rested in open zones in the proximity of urban areas) and thus in habitat use, which enables coexistence in the same macrohabitat (Larroque et al

2015). In this study we also described how these species can co-occur in certain areas but how each species is influenced by a different set of variables explaining their distribution. Certainly, a fine resolution (pixel size) of the environmental variables is essential to explore the preferences of each marten species in a landscape of highly intermixed small patches and to detect the significant niche divergence among martens that otherwise may remain unknown.

Conclusions

Our results emphasize the importance of analyses conducted at appropriate spatial scales, providing additional support for the need of habitat suitability models that account for the scale at which each environmental characteristic is measured. In addition, the incorporation of the Gaussian Kernel method for sampling bias correction led to a robust prediction of each species distribution, and was more accurate than raw models and the two alternative bias correction methods. Based on these results we recommend that all PO based HSMs should account for sampling bias.

Overall, the pine marten was positively associated with cooler areas with a small degree of human disturbance, a high proportion of natural forests, well-connected forestry plantations and medium-extent of agroforesty mosaics. On the other hand, the stone marten presence was conditioned by the density of urban areas, the proportion and extensiveness of croplands, the existence of some scrub cover and the availability of semi-continuous grasslands.

The high resolution sampling grain selected resulted in the inclusion of small but important landscape elements and patches, particularly decisive in a fragmented and heterogeneous area, which allowed the detection of the significant niche divergence found in the closely related marten species co-occurring in northern Spain. To our knowledge, this is the first study comparing multiscale habitat selection and niche divergence of two related marten species. Further, this study provides a useful methodological framework for multispecies and multiscale comparisons.

Acknowledgments

This study has been partially funded by the Basque Government through the Research group on "Systematics, Biogeography and Population Dynamics" (Ref. IT317-10; GIC10/76). MV (Ref: RBFI-2012-446) and ARG (Ref: DKR-2012-64) were supported by a PhD and post-doctoral fellowships awarded by the Dept. of Education, Universities and Research of the Basque Government. U.S. Forest Service Rocky Mountain Research Station supported Cushman's work on this project.

The authors wish to thank all the people directly involved in the collection of non-invasive genetic sampling and those reporting the species locations used in this study, including rangers, vets and field researchers and their institutions (Table S1).

We are also very grateful to Dr. Jason Brown and Dr. Warren for their useful comments regarding SDMtools and ENMtools, respectively.

References

AEMET (2011) Atlas Climático Ibérico / Iberian Climate Atlas. Agencia Estatal de Meteorología, Ministerio de Medio Ambiente, Rural y Marino, Madrid e Instituto de Meteorologia de Portugal

Aguirre-Gutiérrez J, Carvalheiro LG, Polce C, et al (2013) Fit-for-purpose: species distribution model performance depends on evaluation criteria - Dutch Hoverflies as a case study. PLoS One. doi: 10.1371/journal.pone.0063708

Anderson R., Peterson A., Gomez-Laverde M (2002) Using niche-based GIS modeling to test geographic predictions of competitive exclusion and competitive release in South American pocket mice. Oikos 98:3–16.

Anderson RP, Gonzalez I (2011) Species-specific tuning increases robustness to sampling bias in models of species distributions: An implementation with Maxent. Ecol Modell 222:2796–2811.

Anderson RP, Raza A (2010) The effect of the extent of the study region on GIS models of species geographic distributions and estimates of niche evolution: Preliminary tests with montane rodents (genus *Nephelomys*) in Venezuela. J Biogeogr 37:1378–1393.

Balestrieri A, Remonti L, Ruiz-González A, et al (2015) Distribution and habitat use by pine marten *Martes martes* in a riparian corridor crossing intensively cultivated lowlands. Ecol Res 30:153–162.

Balestrieri A, Remonti L, Ruiz-González A, et al (2010) Food habits of genetically identified pine marten (*Martes martes*) expanding in agricultural lowlands (NW Italy). Acta Theriol (Warsz) 56:199–207.

Barbet-Massin M, Jiguet F, Albert CH, Thuiller W (2012) Selecting pseudo-absences for species distribution models: How, where and how many? Methods Ecol Evol 3:327–338.

Barrientos R, Virgós E (2006) Reduction of potential food interference in two sympatric carnivores by sequential use of shared resources. Acta Oecologica 30:107–116.

Bellamy C, Scott C, Altringham J (2013) Multiscale, presence-only habitat suitability models: fine-resolution maps for eight bat species. J Appl Ecol 50:892–901.

Bissonette J., Broekhuizen S (1995) Martes populations as indicators of habitat spatial patterns: the need for a multiscale approach. In: Lidicker WZJ (ed) Landsc. approaches Mamm. Ecol. Conserv. Minnesota. University of Minnesota Press, pp 95–121

Boria RA, Olson LE, Goodman SM, Anderson RP (2014) Spatial filtering to reduce sampling bias can improve the performance of ecological niche models. Ecol Modell 275:73–77.

Brainerd S., Rolstad J (2002) Habitat selection by Eurasian pine martens *Martes martes* in managed forests of southern boreal Scandinavia. Wildlife Biol 8:289–297.

Brown JL (2014) SDMtoolbox: a python-based GIS toolkit for landscape genetic, biogeographic and species distribution model analyses. Methods Ecol Evol 5:694–700.

Buskirk S., Powell RA (1994) Habitat ecology of fishers and American martens. In: Buskirk S., Harestad A., Raphael M., Powell R. (eds) Martens, sables, and fishers: biology and conservation. Cornell University Press, Ithaca, N.Y, pp 283–296.

Cushman S., McGarigal K (2004) Patterns in the species-environment relationship depend on both scale and choice of response variables. Oikos 105:117–124.

Davison A, Birks JDS, Brookes RC, et al (2002) On the origin of faeces: morphological versus molecular methods for surveying rare carnivores from their scats. J Zool 257:141–143.

Delibes M (1983) Interspecific competition and the habitat of the stone marten *Martes foina* (Erxleben 1777) in Europe. Acta Zool Fenn 174:229–231.

Elith J, Graham CH, Anderson RP, et al (2006) Novel methods improve prediction of species' distributions from occurrence data. Ecography 29:129–151.

Elith J, Kearney M, Phillips S (2010) The art of modelling range-shifting species. Methods Ecol Evol 1:330–342.

Elith J, Phillips SJ, Hastie T, et al (2011) A statistical explanation of MaxEnt for ecologists. Divers Distrib 17:43–57.

Evans J, Oakleaf J, Cushman S, Theobald D (2014) An ArcGIS Toolbox for Surface Gradient and Geomorphometric Modeling, version 2.0-0.

Fielding AH, Bell JF (1997) A review of methods for the assessment of prediction errors in conservation presence/absence models. Environ Conserv 24:38–49.

Fisher JT, Anholt B, Bradbury S, et al (2013) Spatial segregation of sympatric marten and fishers: the influence of landscapes and species-scapes. Ecography 36:240–248.

Fourcade Y, Engler JO, Rödder D, Secondi J (2014) Mapping species distributions with MAXENT using a geographically biased sample of presence data: A performance assessment of methods for correcting sampling bias. PLoS One 9:1–13.

Girvetz EH, Greco SE (2009) Multi-scale predictive habitat suitability modeling based on hierarchically delineated patches: An example for yellow-billed cuckoos nesting in riparian forests, California, USA. Landsc Ecol 24:1315–1329.

Goszczyński J, Posłuszny M, Pilot M, Gralak B (2007) Patterns of winter locomotion and foraging in two sympatric marten species: *Martes martes* and *Martes foina*. Can J Zool 85:239–249.

Graf RF, Bollmann K, Suter W, Bugmann H (2005) The importance of spatial scale in habitat models: Capercaillie in the Swiss Alps. Landsc Ecol 20:703–717.

Herr J, Schley L, Roper TJ (2009) Socio-spatial organization of urban stone martens. J Zool 277:54–62.

Herrmann M (1994) Habitat use and spatial organization by the stone marten. In: Buskirk W, Harestad A., Raphael M., Powell R. (eds) Martens, sables, and fishers: biology and conservation. Cornell University Press, Ithaca, N.Y, pp 122–136.

Jiménez-Valverde A, Lobo JM (2007) Threshold criteria for conversion of probability of species presence to either-or presence-absence. Acta Oecologica 31:361–369.

Jiménez-Valverde A, Peterson A., Soberón J, et al (2011) Use of niche models in invasive species risk assessments. Biol Invasions 13:2785–2797.

Johnson D (1980) The comparison of usage and availability measurements for evaluating resource preference. Ecology 61:65–71.

Khanum R, Mumtaz a. S, Kumar S (2013) Predicting impacts of climate change on medicinal asclepiads of Pakistan using Maxent modeling. Acta Oecologica 49:23–31.

Koreň M, Finďo S, Skuban M, Kajba M (2011) Habitat suitability modelling from non-point data. Ecol Inform 6:296–302.

Kramer-Schadt S, Niedballa J, Pilgrim JD, et al (2013) The importance of correcting for sampling bias in MaxEnt species distribution models. Divers Distrib 19:1366–1379.

Larroque J, Ruette S, Vandel J-M, Devillard S (2015) Where to sleep in a rural landscape? A comparative study of resting sites pattern in two syntopic Martes species. Ecography. doi: 10.1111/ecog.01133

Lobo JM, Jiménez-valverde A, Real R (2008) AUC: A misleading measure of the performance of predictive distribution models. Glob Ecol Biogeogr 17:145–151.

López-Martín J. (2007) Martes martes (Linnaeus, 1758). In: Palomo LJ, Gisbert J, Blanco J. (eds) Atlas y Libro Rojo los Mamíferos Terrestres de España. Dirección General de Biodiversidad- SECEM-SECEMU. Madrid, pp 302–304

Mateo-Sánchez M, Cushman S., Saura S (2013) Scale dependence in habitat selection: The case of the endangered brown bear (*Ursus arctos*) in the Cantabrian Range (NW Spain). Int J Geogr Inf Sci 1–16.

McGarigal K, Cushman SA, Ene E (2012) FRAGSTATS v4: Spatial Pattern Analysis Program for Categorical and Continuous Maps. Computer Software Program Produced by authors University of Massachusetts, Amherst. Available Follow. web site http//www.umass.edu/landeco/research/fragstats/fragstats.html

Mergey M, Helder R, Roeder J-J (2011) Effect of forest fragmentation on space use patterns in the European pine marten (*Martes martes*). J Mammal 92:328–335.

Mergey M, Larroque J, Ruette S, et al (2012) Linking habitat characteristics with genetic diversity of the European pine marten (*Martes martes*) in France. Eur J Wildl Res 909–922.

Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species geographic distributions. Ecol Modell 190:231–259.

Phillips SJ, Dudik M, Elith J, et al (2009) Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. Ecol Appl 19:181–197.

Posłuszny M, Pilot M, Goszczyński J, Gralak B (2007) Diet of sympatric pine marten (*Martes martes*) and stone marten (*Martes foina*) identified by genotyping of DNA from faeces. Ann Zool Fennici 269–284.

Proulx G, Aubry K, Birks J, et al (2004) World distribution and status of the genus martes in 2000. In: Harrison DJ, A.K. F, Proulx G (eds) Martens and Fishers in Human-altered Environments. Springer. New York, pp 21–76

R Development Core Team (2014) R: A language and environment for statistical computing.

Radosavljevic A, Anderson RP (2014) Making better Maxent models of species distributions: Complexity, overfitting and evaluation. J Biogeogr 41:629–643.

Reig S (1992) Geographic variation in pine marten (*Martes martes*) and beech marten (*M. foina*) in Europe. J Mammal 73:744–769.

Rodrigues M, Santos-Reis M, Elmeros M, Fernandes C. (2011) Microsatellite markers for genetic studies in the weasel (*Mustela nivalis*). Eur J Wildl Res 58:507–510.

Rosalino LM, Santos-Reis M (2009) Fruit consumption by carnivores in Mediterranean Europe. Mamm Rev 39:67–78.

Ruiz-González A, Gurrutxaga M, Cushman SA, et al (2014) Landscape Genetics for the Empirical Assessment of Resistance Surfaces: The European Pine Marten (*Martes martes*) as a Target-Species of a Regional Ecological Network. PLoS One 9:19. doi: 10.1371/journal.pone.0110552

Ruiz-Gonzalez A, Jose Madeira M, Randi E, et al (2013) Non-invasive genetic sampling of sympatric marten species (*Martes martes* and *Martes foina*): assessing species and individual identification success rates on faecal DNA genotyping. Eur J Wildl Res 59:371–386.

Ruiz-Gonzalez A, Rubines J, Berdion O, Gomez-Moliner BJ (2008) A non-invasive genetic method to identify the sympatric mustelids pine marten (*Martes martes*) and stone marten (*Martes foina*): preliminary distribution survey on the northern Iberian Peninsula. Eur J Wildl Res 54:253–261.

Santos MJ, Santos-Reis M (2010) Stone marten (*Martes foina*) habitat in a Mediterranean ecosystem: effects of scale, sex, and interspecific interactions. Eur J Wildl Res 56:275–286.

Shirk AJ, Wasserman TN, Cushman SA, Raphael MG (2012) Scale dependency of American marten (*Martes americana*) habitat relationships. In: Aubry KB, Zielinski WJ, Proulx G, Buskirk SW (eds) Biology and conservation of martens, sables, and fishers: a new synthesis. Cornell University Press, New York, pp 269–283.

Shirk AJ, Raphael MG, Cushman SA (2014) Spatiotemporal variation in resource selection: insights from the American marten (*Martes americana*). Ecol Appl 24:1434–1444.

Svenning JC, Normand S, Kageyama M (2008) Glacial refugia of temperate trees in Europe: Insights from species distribution modelling. J Ecol 96:1117–1127.

Syfert MM, Smith MJ, Coomes DA (2013) The Effects of Sampling Bias and Model Complexity on the Predictive Performance of MaxEnt Species Distribution Models. PLoS One. doi: 10.1371/journal.pone.0055158

Tikhonov A, Cavallini P, Maran T, et al (2008) *Martes foina*. The IUCN Red List of Threatened Species. Version 2014.3.

Virgós E, García FJ (2002) Patch occupancy by stone martens *Martes foina* in fragmented landscapes of central Spain: the role of fragment size, isolation and habitat structure. Acta Oecologica 23:231–237.

Virgós E, Recio M., Cortés Y (2000) Stone marten (*Martes foina*) use of different landscape types in the mountains of central Spain. Z Säugetierkd 65:375–379.

Virgós E, Zalewski A, Rosalino L, Mergey M (2012) Habitat ecology of genus Martes in Europe: a review of the evidences. In: Aubry KB, Zielinski WJ, Proulx G, Buskirk SW (eds) Biology and conservation of martens, sables, and fishers: a new synthesis. Cornell University Press, New York, pp 255–266.

Warren DL, Glor RE, Turelli M (2008) Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution. Evolution 62:2868–83.

Warren DL, Glor RE, Turelli M (2010) ENMTools: a toolbox for comparative studies of environmental niche models. Ecography 1:607–611.

Warren DL, Seifert SN (2011) Ecological niche modeling in Maxent: the importance of model complexity and the performance of model selection criteria. Ecol Appl 21:335–342.

Wasserman TN, Cushman SA, Wallin DO, Hayden J (2012) Multi scale habitat relationships of Martes americana in northern Idaho, U.S.A. USDA Forest Service RMRS Research Paper RMRS-RP-94.

Wellenreuther M, Larson KW, Svensson EI (2012) Climatic niche divergence or conservatism? Environmental niches and range limits in ecologically similar damselflies. Ecology 93:1353–1366.

Wiens J, Rotenberry JT, van Horne B (1987) Habitat occupancy patterns of North American shrubsteppe birds: the effects of spatial scale. Oikos 48:132–147.

Yackulic CB, Chandler R, Zipkin EF et al (2013) Presence-only modelling using MAXENT: when can we trust the inferences?. Methods Ecol. Evol 4:236–243.

Zalewski A, Jedrzejewski W, Jedrzejewska B (2004) Mobility and home range use by pine martens (*Martes martes*) in a Polish primeval forest. Ecoscience 11:113–122.

Zalewski A, Włodzimierz JW (2006) Spatial organisation and dynamics of the pine marten *Martes martes* population in Białowieza Forest (E Poland) compared with other European woodlands. Ecography 29:31–43.

Supporting Information

Table S1. Marten locations used to build the habitat suitability models (n=1286).

The table is too big to print in less than 15 pages. Thus, before the papers are published the table will be available online (dropbox folder) in the following link:

https://www.dropbox.com/sh/y4lw39hvnfrepow/AAD72y908pLs-3Zp2G4YUrCGa?dl=0

		HUMAN PRESSURE			TOPOGRAPHIC			CLIMATIC	LANDCOVER-Landscape metrics				
Sp.	Scale	*HWS*	*NAT*	*AUT*	ELEV	*ROUGH*	*CTI*	*TEMP*	AI	CWED	ED	PD	*SHDI*
Pine marten	125m	0,500	0,499	0,495	0,598	0,521	0,469	0,624	-	-	-	-	-
	250m	0,510	0,459	0,492	0,599	0,501	0,459	0,638	-	-	-	-	-
	500m	0,514	0,485	0,484	0,599	0,534	**0,551**	0,639	-	-	-	-	-
	1km	0,502	0,502	0,528	0,600	0,541	0,548	0,646	0,532	0,523	0,534	0,586	0,529
	2km	0,486	0,448	0,563	0,612	0,537	0,526	0,652	0,560	0,545	0,588	0,628	**0,560**
	4km	0,497	0,526	0,524	0,631	0,613	0,524	**0,661**	**0,613**	0,588	0,620	**0,667**	0,534
	8km	0,529	0,545	**0,604**	0,641	**0,640**	0,514	0,661	0,605	**0,600**	**0,622**	0,597	0,512
	16km	0,514	**0,575**	0,534	0,649	0,626	0,514	0,631	0,575	0,580	0,585	0,500	0,505
	32km	**0,597**	0,553	0,533	**0,673**	0,609	0,518	0,642	0,583	0,585	0,591	0,500	0,501
Sp.	Scale	*HWS*	*NAT*	*AUT*	ELEV	*ROUGH*	*CTI*	*TEMP*	AI	CWED	ED	PD	*SHDI*
Stone marten	125m	0,498	0,505	0,542	0,514	0,511	0,467	0,525	-	-	-	-	-
	250m	0,504	0,513	0,563	0,514	0,535	0,485	0,537	-	-	-	-	-
	500m	0,500	0,510	0,561	0,516	0,541	0,504	0,540	-	-	-	-	-
	1km	0,500	0,495	0,462	0,529	0,559	0,507	0,552	0,588	0,572	0,589	0,634	0,536
	2km	0,503	0,494	0,519	0,552	**0,567**	0,527	0,575	0,592	0,591	0,598	**0,665**	**0,586**
	4km	0,502	0,488	0,543	0,555	0,559	**0,543**	0,591	0,587	0,600	0,588	0,664	0,555
	8km	0,533	0,522	0,556	0,531	0,550	0,525	0,589	0,576	0,596	0,578	0,600	0,547
	16km	0,541	**0,567**	**0,593**	0,551	0,540	0,531	**0,596**	0,586	0,602	0,586	0,500	0,530
	32km	**0,593**	0,556	0,586	**0,577**	0,500	0,522	0,586	**0,645**	**0,627**	**0,640**	0,500	0,507

Table S2. Results of the bivariate scaling showing the AUC values of the single variable models across scales and species. The scale with the highest AUC value is represented in bold. Variables underlined (shared by both mustelids) and in italics (species-specific) are those which remained after pruning.

		LANDCOVER-Class metrics												
Sp.	Scale	area_Nat	area_Fp	area_Ag	area_Pa	area_Scr	*area_Urb*	gyr_Nat	gyr_Fp	*gyr_Ag*	*gyr_Pa*	*gyr_Scr*	gyr_Cr	gyr_Urb
Pine marten	125m	-	-	-	-	-	-	-	-	-	-	-	-	-
	250m	-	-	-	-	-	-	-	-	-	-	-	-	-
	500m	-	-	-	-	-	-	-	-	-	-	-	-	-
	1km	0,684	0,580	0,553	0,477	0,487	0,552	0,671	0,604	0,550	0,597	0,567	0,673	0,553
	2km	**0,694**	0,566	0,587	0,436	0,49	0,55	**0,674**	0,555	0,588	**0,624**	0,504	0,673	0,545
	4km	0,674	0,508	0,623	0,445	0,465	0,607	0,668	0,519	0,626	0,608	0,475	0,641	0,607
	8km	0,607	0,542	**0,630**	0,404	**0,614**	0,6	0,662	0,532	0,635	0,603	0,598	0,648	0,630
	16km	0,559	0,620	0,606	**0,601**	0,608	0,597	0,551	0,619	0,607	0,565	**0,626**	0,649	0,618
	32km	0,521	**0,724**	0,597	0,591	0,553	**0,665**	0,550	**0,716**	**0,650**	0,597	0,575	**0,739**	**0,715**
Sp.	Scale	area_Nat	area_Fp	area_Ag	area_Pa	area_Scr	*area_Urb*	gyr_Nat	gyr_Fp	*gyr_Ag*	*gyr_Pa*	*gyr_Scr*	gyr_Cr	gyr_Urb
Stone marten	125m	-	-	-	-	-	-	-	-	-	-	-	-	-
	250m	-	-	-	-	-	-	-	-	-	-	-	-	-
	500m	-	-	-	-	-	-	-	-	-	-	-	-	-
	1km	0,572	0,580	0,537	**0,576**	0,56	0,579	0,583	0,565	0,541	**0,606**	0,567	0,573	0,580
	2km	0,582	0,567	0,551	0,419	0,578	**0,589**	0,561	0,556	0,552	0,592	0,585	0,566	**0,584**
	4km	0,592	0,539	0,587	0,533	0,602	0,576	0,569	0,530	0,589	0,572	0,615	0,546	0,582
	8km	0,583	0,521	**0,626**	0,456	0,577	0,551	0,549	0,519	**0,629**	0,557	0,580	0,515	0,554
	16km	**0,594**	0,539	0,559	0,539	**0,617**	0,572	**0,589**	0,538	0,565	0,515	**0,628**	0,542	0,564
	32km	0,585	**0,603**	0,546	0,501	0,598	0,553	0,585	**0,590**	0,585	0,511	0,579	**0,589**	0,563

Table S2. Continued.

		LANDCOVER-Class metrics													
Sp.	Scale	pland_Nat	pland_Fp	pland_Ag	pland_Pa	pland_Scr	*pland_Cr*	*pland_Urb*	pd_Nat	*pd_Fp*	*pd_Ag*	*pd_Pa*	pd_Scr	*pd_Cr*	*pd_Urb*
Pine marten	125m	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	250m	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	500m	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	1km	0,690	0,601	0,548	0,559	0,467	0,674	0,566	0,548	0,564	**0,508**	**0,589**	0,620	**0,562**	0,570
	2km	**0,706**	0,604	0,583	0,595	0,457	0,675	0,582	0,599	**0,575**	0,500	0,551	**0,655**	0,537	**0,574**
	4km	0,697	0,593	0,593	0,608	0,525	0,648	0,626	**0,608**	0,521	0,500	0,501	0,575	0,507	0,509
	8km	0,665	0,606	0,613	0,632	**0,618**	0,715	0,648	0,500	0,500	0,500	0,500	0,500	0,500	0,500
	16km	0,605	0,643	0,615	0,612	0,537	**0,746**	0,670	0,500	0,500	0,500	0,500	0,500	0,500	0,500
	32km	0,549	**0,711**	**0,633**	**0,702**	0,603	0,737	**0,689**	0,500	0,500	0,500	0,500	0,500	0,500	0,500
Sp.	Scale	pland_Nat	pland_Fp	pland_Ag	pland_Pa	pland_Scr	*pland_Cr*	*pland_Urb*	pd_Nat	*pd_Fp*	*pd_Ag*	*pd_Pa*	pd_Scr	*pd_Cr*	*pd_Urb*
Stone marten	125m	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	250m	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	500m	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	1km	0,592	0,566	0,532	0,561	0,530	0,564	0,594	0,594	0,598	**0,509**	0,598	0,599	0,482	0,625
	2km	0,593	0,580	0,548	**0,587**	0,534	0,571	0,630	**0,595**	**0,562**	0,500	**0,603**	**0,634**	**0,517**	**0,640**
	4km	0,598	0,554	0,573	0,573	0,553	**0,582**	0,630	0,556	0,502	0,500	0,519	0,569	0,508	0,573
	8km	0,594	0,552	0,601	0,585	0,575	0,544	0,630	0,500	0,500	0,500	0,500	0,500	0,500	0,500
	16km	0,613	0,541	**0,604**	0,577	0,568	0,555	0,628	0,500	0,500	0,500	0,500	0,500	0,500	0,500
	32km	**0,624**	**0,614**	0,598	0,577	**0,593**	0,550	**0,655**	0,500	0,500	0,499	0,500	0,500	0,500	0,500

Table S2. Continued.

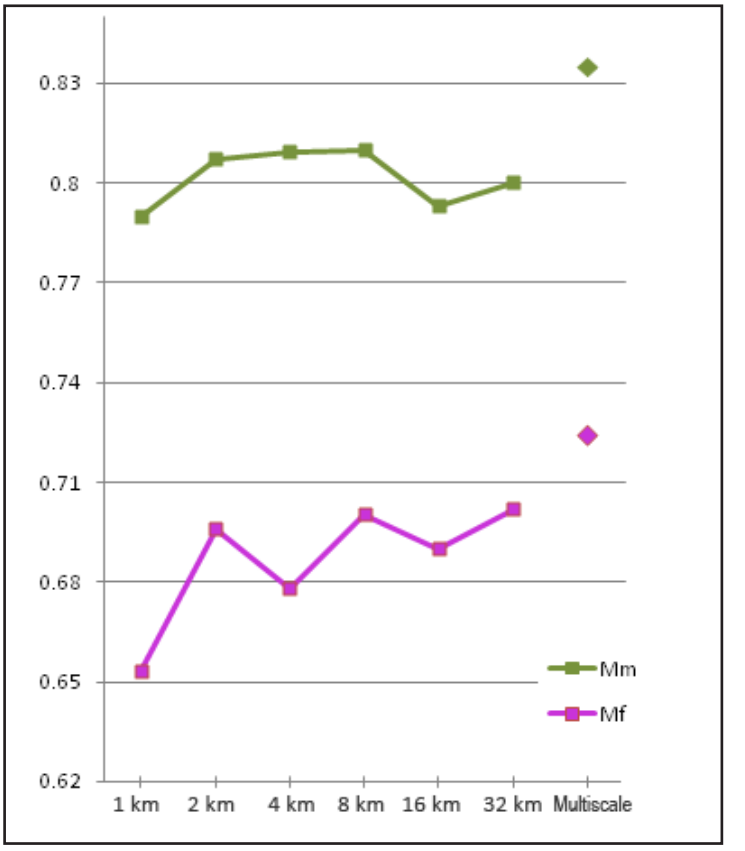


Figure S1. Comparative of the performance of the multiscale and the single-scale models. The AUC values obtained for the pine marten across scales are shown in green and those for the stone marten are shown in purple. The AUC values obtained for the multiscale models are reported last.

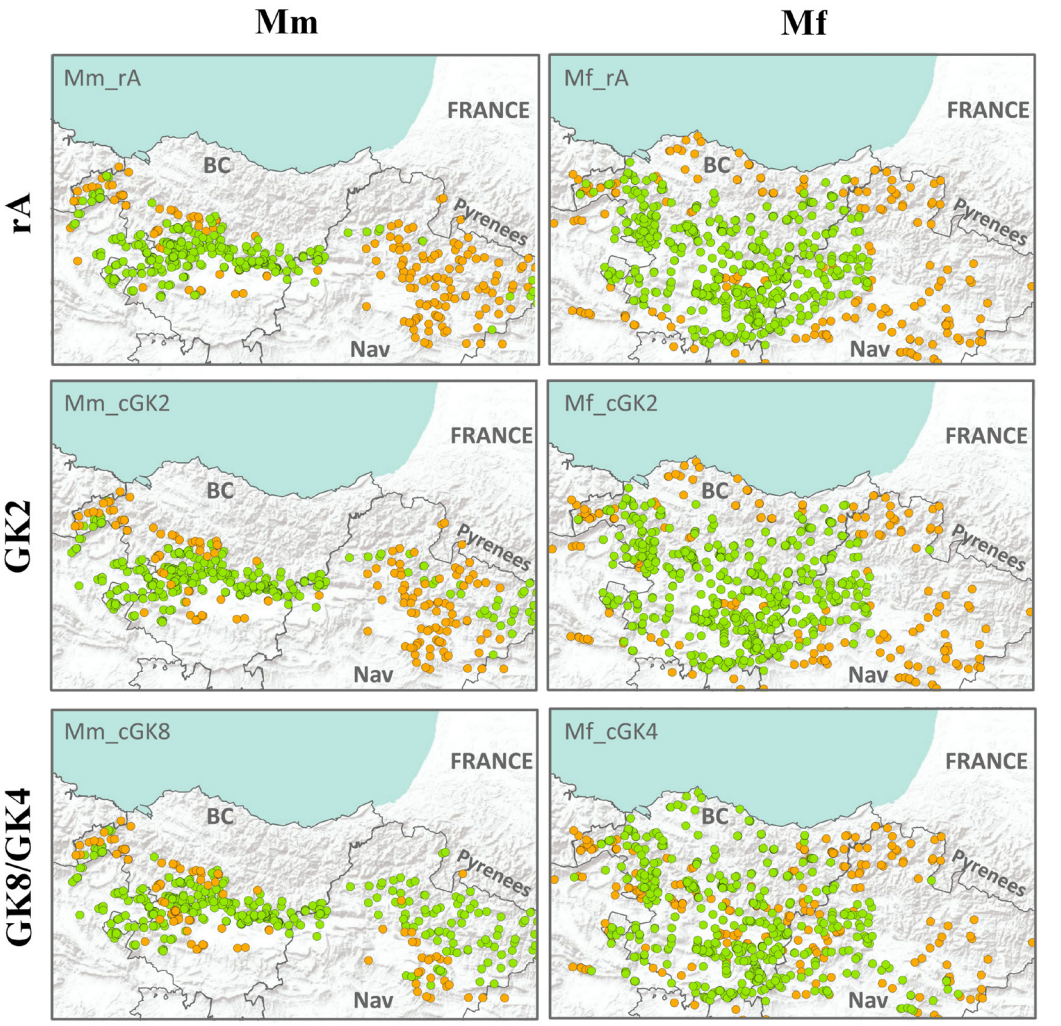


Figure S2. Distribution of false absences (FA) and true presences (TP) for the raw and two corrected HSMs for Martes spp. FA are colored in orange while TP are shown in green. The raw Mm_rA and the corrected Mm_cGK2 and Mm_cGK8 are reported in the first column while Mf_rA, Mf_cGK2 and Mf_cGK4 are represented in the second.

Chapter II

PAPER APPENDIX 1: Distribution and habitat use by pine marten *Martes martes* in riparian corridor crossing intensively cultivated lowland

Alessandro Balestrieri, Luigi Remonti, Aritz Ruiz-González, Michele Zenato, Andrea Gazzola, **María Vergara**, Ettore E. Dettori, Nicola Saino, Enrica Capelli, Benjamín J. Gómez-Moliner, Franca Guidali and Claudio Prigioni

Ecologial Research, 2015, 30: 153-162

Ecol Res
DOI 10.1007/s11284-014-1220-8

ORIGINAL ARTICLE

**Alessandro Balestrieri · Luigi Remonti · Aritz Ruiz-González · Michele Zenato
Andrea Gazzola · Maria Vergara · Ettore E. Dettori · Nicola Saino
Enrica Capelli · Benjamín J. Gómez-Moliner · Franca Guidali · Claudio Prigioni**

Distribution and habitat use by pine marten *Martes martes* in a riparian corridor crossing intensively cultivated lowlands

Received: 16 June 2014 / Accepted: 17 November 2014

Abstract The location of pine marten records in northern Italy suggests that main rivers may play the role of natural corridors favouring this species' colonisation of cultivated lowlands. We assessed the distribution and habitat use by the pine marten on a 35 km long stretch of the River Ticino. Surveys were carried out between October 2011 and June 2012 along linear transects in a 2×2 km grid. Using the variation in marking intensity as an indicator of habitat use, habitat selection was assessed at two landscape levels—at transect-scale by the χ^2 test with Bonferroni's confidence intervals for the proportion of use, and at grid-scale by multiple linear regression. By a polymerase chain reaction–restriction fragment length polymorphism method, 91 faecal samples were assigned to the pine marten. Faeces were mainly located in wooded areas, while fields were avoided. At the grid-scale of analysis, marking intensity was positively related to the mean area of wooded patches and negatively to their mean perimeter-area ratio. This suggests that pine marten relative abundance may partially depend on the degree of fragmentation and structure of residual woods. The survey protocol allowed to assess the probability of detection. Occupancy models outlined that heterogeneity in detection probability may arise as a result of variation in marking intensity, i.e. the number of marking individuals. Our results suggest that the availability of both woodland corridors and wood patches are major factors shaping pine marten distribution in intensively cultivated plains and that non-invasive genetic surveys are a cost-effective method for future studies at a broader scale.

Keywords Non-invasive genetic sampling · Detectability · Faecal DNA · Stone marten · Northern Italy

A. Balestrieri (✉) · N. Saino · F. Guidali
Department of Biosciences, University of Milan, Via Celoria 26, 20133 Milan, Italy
E-mail: alebls@libero.it

L. Remonti · M. Zenato · A. Gazzola · E. E. Dettori · E. Capelli · C. Prigioni
Department of Earth and Environmental Sciences, University of Pavia, Via Taramelli 22, 27100 Pavia, Italy

A. Ruiz-González · M. Vergara · B. J. Gómez-Moliner
Department of Zoology and Animal Cell Biology, University of the Basque Country (UPV/EHU), C/Paseo de la Universidad 7, 01006 Vitoria-Gasteiz, Spain

A. Ruiz-González · M. Vergara
Systematics, Biogeography and Population Dynamics Research Group, Lascaray Research Center, University of the Basque Country (UPV/EHU), Avda. Miguel de Unamuno, 3, 01006 Vitoria-Gasteiz, Spain

A. Ruiz-González
Conservation Genetics Laboratory, National Institute for Environmental Protection and Research (ISPRA), Via Cà Fornacetta 9, 40064 Ozzano dell'Emilia, BO, Italy

Introduction

In the Mediterranean region, human population growth, agricultural intensification and the consequent loss of natural habitat have led to a general decline of biodiversity in both plain and coastal areas (Matson et al. 1997; Benton et al. 2003; Lepers et al. 2005). Plains show the lowest species richness, with the exception of some small wet areas and residual riparian woods, which support a more diverse fauna than surrounding habitats (e.g. Warkentin et al. 1995; Bentley and Catterall 1997; Hilty and Merenlender 2004).

A major biological effect of habitat fragmentation caused by anthropogenic modification is the decline of species that need large areas of connected natural habitats to meet their ecological requirements (Beier 1993; Mortelliti et al. 2010). Mammal distribution is particularly affected by the increased isolation and reduction in area of habitat patches (Bright 1993; Waldron et al. 2006) and habitat fragmentation is considered a major cause of the decline of forest-dwelling species (Reed 2004).

Published online: 29 November 2014

The pine marten (*Martes martes*) has been long considered a forest-specialist and its generalized decline has been imputed to the combined effects of large individual home ranges and deforestation (Buskirk 1992; Buskirk and Zielinski 2003). Nonetheless, recent studies have shown that the pine marten can colonise agricultural landscapes with highly fragmented woods (Balestrieri et al. 2010; Mergey et al. 2011; Caryl et al. 2012), suggesting that, as already observed in Mediterranean Italy (De Marinis and Masseti 1993), it is not such a strict forest-dweller as previously believed (see Virgós et al. 2012).

In Italy, the pine marten occurs sympatrically with the closely related stone marten (*Martes foina*) in mountainous areas, while in plain areas only the stone marten has been reported (Genovesi and De Marinis 2003a, b). Currently, the pine marten is colonising the western sector of the intensively cultivated Po plain (N Italy), where it probably went extinct at the end of the 1960s (Mantovani 2010), and pine marten expansion seems to coincide with the contraction of the stone marten range (Remonti et al. 2012).

The pattern of pine marten expansion seems to suggest that water systems may play the role of natural corridors favouring the dispersal of the pine marten from subalpine areas and the colonization of agricultural lowland (Fig. 1).

Riparian zones are important for maintaining landscape connectivity (Naiman et al. 1993; Taylor et al. 1993) and biological connections for wildlife (Clerici and Vogt 2013). The use by dispersing mammals of both natural and man-made linkages has been reported for several species (see Bennett 2003 for a comprehensive review). Corridors can assist the range expansion of both scarcely and highly mobile mammals, from meadow vole *Microtus pennsylvanicus* (Getz et al. 1978) to mink *Neovison vison*, which, following its introduction, has spread in Great Britain through river corridors (Harris and Woollard 1990).

As in most European lowlands (Coles et al. 1989; Bennett 2003), in northern Italy main rivers form linear habitats covered with remnant riparian vegetation, crossing clearly distinct, heavily disturbed farmland and urban areas. Lowland woods cover only 1,750 km^2, of which about 70 % are located in the western and central plain of the River Po (Camerano et al. 2010) as small residual fragments (mean area 4.5 ha; Lassini et al. 2007). The plain of the River Ticino includes the largest and best conserved riparian woods of northern Italy and represents the less altered river corridor of the western basin of the River Po, a critical European region from a conservation perspective (Clerici and Vogt 2013).

Records of road-killed pine marten suggested that the valley of the River Ticino may represent a suitable dispersal route for the pine marten (Balestrieri et al. 2010). Accordingly, since 2005 a stable pine marten population has dwelt in a small agricultural area, 5 km to the west of the River Ticino, about 20 km upstream its confluence in the River Po (Remonti et al. 2012). Thus, we aimed to draw the actual distribution of both pine and stone marten in the downstream, 35 km long section of the River Ticino valley and assess the pattern of pine marten habitat use in a fragmented landscape. To account for martens elusiveness and the difficulty of distinguishing its tracks from those of other medium-sizes carnivores, we applied a non-invasive genetic sampling method based on mtDNA extracted from faecal samples (Ruiz-González et al. 2008).

As the survey protocol involved repeated surveys, we also modelled the effects of both habitat variables and marking intensity on the probability of pine marten detection, in order to evaluate the cost-effectiveness of the survey method a major factor to consider in the design of occupancy surveys for carnivore mammals (Slauson et al. 2009; Long et al. 2011).

Study area

The Italian stretch of the River Ticino flows southwards from the southern edge of Lake Maggiore to the median course of the River Po, forming a 110 km long and, on average, 7 km wide valley.

The valley is partly protected by two Regional Parks: the Park of the Ticino Valley (Lombardy), covering 906.4 km^2 and the Natural Park of the Ticino Valley (Piedmont), 62.5 km^2.

The river crosses an intensively cultivated and urbanized plain, where mesophilous—*Fraxino-carpinion*—and hygrophilous—*Alno-Ulmion*, *Alnion glutinoso-incanae*, *Salicion albae*—woods are still widespread inside the weave of meanders, streams, canals and oxbow lakes which characterise the downstream stretch of the river. On the whole, water-bodies cover an area of about 48 km^2, while wet woods account for 87 km^2 (Prigioni 1995).

Pine marten monitoring focused on the lower part of the valley, from the towns of Vigevano and Abbiategrasso (Milan, Lombardy) in the north to Gropello Cairoli village (Pavia, Lombardy) in the south (Fig. 2). This ca. 35 km long stretch of the river has a mean annual discharge of about 300 m^3s^{-1} and a catchment area of more than 7,000 km^2. Mean percent riparian vegetation cover, as assessed in a 100 m large belt on both river banks for seven 5 km long river stretches, is 47.8 % (min–max: 12–86 %; Prigioni and Balestrieri 2011).

The climate is temperate, mean annual values being 13 °C for air temperature and ca. 700 mm for rainfall.

Methods

Collection and genetic identification of faecal samples

Surveys were carried out within a 2 × 2 km grid (Fig. 2), superimposed on the kilometric grid of digitalized,

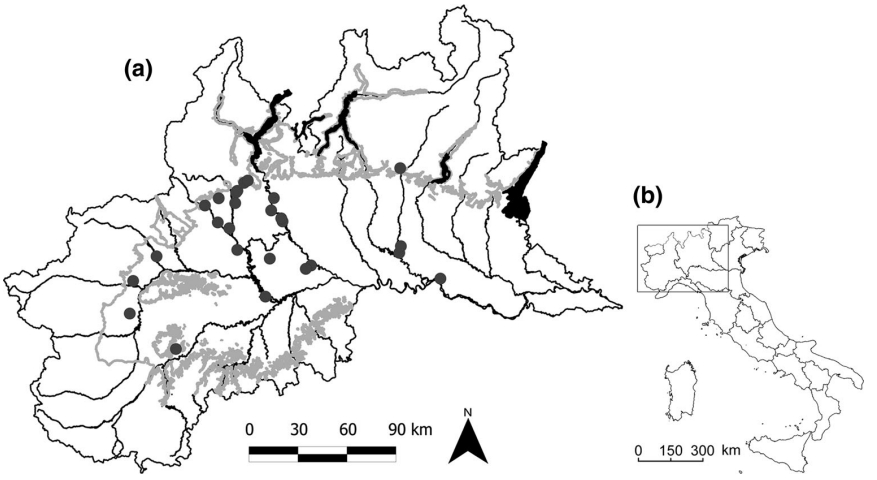


Fig. 1 Distribution of pine marten road kill records (1988–2012*; grey dots*) in the western and central River Po plain (**a**) (Balestrieri et al. 2010; Mantovani 2010) and with respect to northern Italy (**b**). Main rivers and lakes are shown in *dark grey*, while the *light grey line* is the 300 m a.s.l. *contour line*, which broadly marks the upper limit of the plain

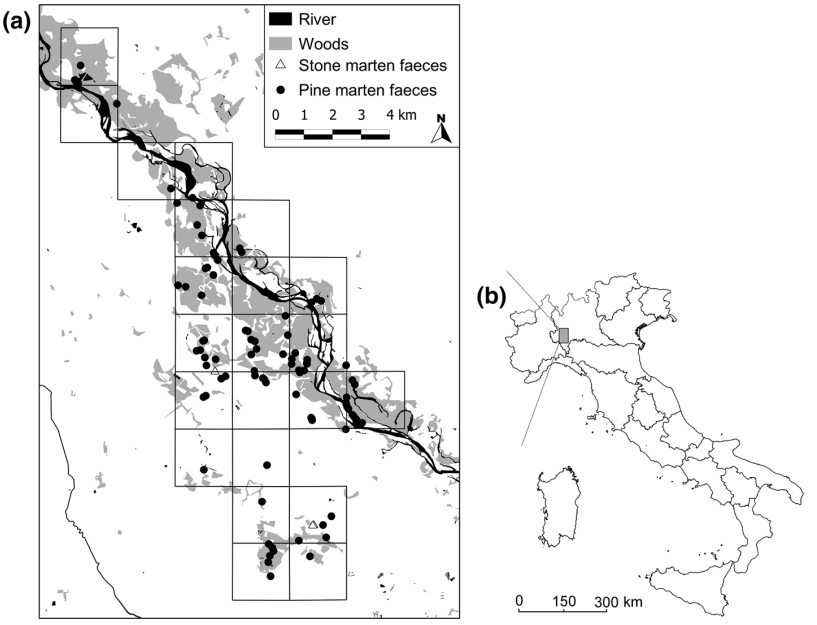


Fig. 2 Genotyped faecal samples into 21, 2 × 2 km grid squares surveyed in the valley of the River Ticino (**a**), and location of the study area in Italy (**b**)

1:10,000 Regional Technical Maps. Grid size was chosen as to broadly correspond to pine marten mean home range in Tuscany (370 ha; Del Fante 2012).

Sampling was conducted between October 2011 and June 2012 along linear transects drawn along wood/field margins, paths and country roads to cover both open and forested habitats. Transects were surveyed 1–3 times each (mean ± SD = 2.44 ± 0.87).

A portion (ca. 1 cm) of each "marten-like" faeces (i.e. < 10 mm large and then suspected to belong to the pine marten; for more details on faeces identification, see Remonti et al. 2012) was picked up using sticks stored in autoclaved tubes containing ethanol 96 % and frozen at −20 °C until processed (Ruiz-González et al. 2008).

Bi-monthly variation in marking intensity was checked by comparing the observed number of faecal samples to that expected based on the overall transect length covered during each period (N = 4).

DNA was isolated using the QIAamp DNA Stool Mini Kit (Qiagen) according to the manufacturer's instructions. The specific identification of faecal samples was accomplished by a polymerase chain reaction–restriction fragment length polymorphism (PCR–RFLP) method, according to Ruiz-González et al. (2008). Two primers amplify the mtDNA from *Martes martes*, *M. foina* and four *Mustela* species, of which only the European polecat *Mustela putorius* had been previously reported for our study area, whilst red fox (*Vulpes vul-*

pes) faeces, which can be easily mistaken for those of martens (Davison et al. 2002), render no amplicons. The simultaneous digestion of amplified mtDNA by two restriction enzymes (*Rsa*I and *Hae*III), generates different restriction patterns for each mustelid species (i.e. DNA fragments differing in both number and length), allowing the unambiguous identification of faecal samples (see Ruiz-González et al. 2008 for further details).

Habitat use

Based on CORINE Land-Cover maps, satellite images and ground surveys, a land-cover, digitalized, 1:10,000 map was produced by a Geographic Information System (QGis®).

Habitat selection was assessed by a use-availability design (Garshelis 2000) at two landscape levels, using faeces relative distribution as an indicator of habitat use.

Faecal counts have been widely used for assessing the abundance and habitat preferences of many mammal species (Putman 1984; Kohn and Wayne 1997; Gese 2001). Their reliability has been long debated (e.g. Kruuk et al. 1986; Messenger and Birks 2000), the interpretation of survey results involving several assumptions about marking activity and dropping identification (Sadlier et al. 2004). Nonetheless, faecal counts have been recently reported to be effective estimators of both relative abundance and habitat use by other carnivores (Clavero et al. 2006; Guter et al. 2008; Lanszki et al. 2008; Rosalino et al. 2008; Balestrieri et al. 2009; Kauhala and Salonen 2012).

To assess habitat use at transect-scale, faecal samples were georeferenced and assigned to a habitat type. Three main habitat types were considered: woods, poplar plantations and cultivated fields (mainly maize and rice). Samples located within a 5-m wide strip on each side of the border between two different habitat types were assigned to both habitats with a 0.5 score. The Chi squared (χ^2) test was performed to test for the goodness of fit of used habitat to available habitat types (White and Garrott 1990). Expected frequencies were calculated based on the overall transect length covered for each habitat. To determine whether a habitat was selected or avoided Bonferroni's confidence intervals for the proportion of use of each habitat were checked:

$$P_i - z_{a/2k}\cdot\sqrt{\frac{P_i(1-P_i)}{N}} < P_i < P_i + z_{a/2k}\cdot\sqrt{\frac{P_i(1-P_i)}{N}};$$

where: P_i is the proportion of use of the ith habitat, N is sample size, z is the percentage point of the standard normal distribution corresponding to an upper tail probability of $\alpha/2\ k$, α is the level of significance and k is the number of habitat types (Neu et al. 1974; Byers et al. 1984).

To assess habitat use at grid scale, the following variables of potential importance to pine marten distribution were measured for each square:

a. W: percentage of woods,
b. U: percentage of urbanised areas,
c. C: percentage of cereal crops,
d. P: percentage of poplar plantations,
e. H: Shannon's index of habitat diversity,
f. minDr: minimum distance from the River Ticino (with minDr = 0 for all grid squares including a river stretch),
g. MWA: mean area of wooded patches, as to account for the degree of fragmentation of pine marten potentially preferred habitat,
h. MPAR: mean perimeter-area ratio of wooded patches, representing their compactness.

The influence of the measured variables on pine marten marking intensity, (MI = number of genetically identified faeces found in each grid square/total transect length per square) was tested by means of a multiple linear regression, using two different approaches.

First, we used a stepwise (forward selection) method to identify which variables were related to pine marten marking intensity (SPSS 12.0.1; SPSS, Chicago, IL, USA). Deviation from normality of the residual distribution was tested by Shapiro–Wilk's test.

Then we performed a linear multiple regression for each subset of uncorrelated (Pearson's test, $p > 0.05$) explanatory variables, entering all independent variables simultaneously. The obtained models were ranked by second order Akaike Information Criterion (AICc; Sugiura 1978), selecting those showing ΔAICc values <2 (ΔAICc = $AICc_i$ − min AICc, where min AICc is the AIC value of the 1st-ranked model; Burnham and Anderson 2002; Posada and Buckley 2004; Mazerolle 2006). The analysis was performed by the software SAM 4.0—Spatial Analysis in Macroecology (Rangel et al. 2010).

Before the analyses, all variables were tested for normality. To achieve normality, U and MPAR were rank transformed, C was square-root transformed, while MWA, P and the length of linear transects were log(x + 1) transformed. Being a frequency of occurrences, MI was log(x + 1) transformed and modelled as a linear function of independent variables (Quinn and Keough 2002).

Detection probability

The influence of habitat variables on pine marten probability of site occupancy (Ψ) and detection probability (p) was analysed by the likelihood-based method for modelling occupancy data proposed by MacKenzie et al. (2002). This modelling is analogous to traditional capture-recapture methods, but uses the proportion of sites occupied by the target species as a state variable (rather than individuals). The goal is to assess Ψ knowing the species is not always detected, even when present. As the method requires multiple surveys in a

demographically closed system (i.e. closed to changes of the occupancy state during the sampling interval; MacKenzie et al. 2003), in a sub-sample of 17 grid squares 1 transect (mean length: 3.6 ± 1.38 km) was surveyed three times between November 2011 and April 2012 (i.e. they were suspended before the time the cubs-of-the-year are supposed to start to scent mark; Genovesi and De Marinis 2003b), as to compile replicate observations into a sequence of 1's (detections) and 0's (non-detections). As grid size was almost equal to the size of home ranges and each home range was not occupied by more than two individuals of same-sex because of territorial behavior (Balharry 1993), we could assume detections at different sites to be independent.

As an example, the likelihood for site i with history 010 would be:

$$\Psi_i(1-p_{i,1})p_{i,2}(1-p_{i,3}),$$

with Ψ_i: probability of site i occupancy; and $p_{i,n}$: detection probability for each visit. Assuming independence of sites, the product of all terms (one for each site) constructed in this manner creates the model (MacKenzie et al. 2002). Ψ may be some function of site characteristics and p may vary with certain variables; using a logistic model this information can be introduced to the model as, respectively, site- and survey-covariates, (MacKenzie et al. 2002).

Analyses were run by the software PRESENCE (Hines 2006), using single-season models. As pine marten presence was expected to depend on wood availability and connectivity (Zalewski and Jedrzejewski 2006), 10 habitat covariates were measured: the length of the transect (L), the percent length of the transect covered by wood (W_t), the overall area of the wood patch crossed by each transect (W_a), wood patch perimeter (W_p), the minimum distance between the wood patch covered by the transect and the nearest >2 km^2 large wood patch (D_{nw}), the mean distance between each transect and the river (D_r), the percent cover of woods (W), urban areas (U) and crops (C) into each grid square; the mean distance between each transect and the nearest urban area (D_u) was considered as an index of human disturbance. To avoid multicollinearity, Spearman's correlation (with $\alpha = 0.01$) test was used to check for any relationship between the covariates. Five covariates (L, W_a, D_{nw}, $\%W_t$ and W) were then selected for modelling.

Variation in marking intensity has been reported as a potential cause of heterogeneity in detection probabilities (Balestrieri et al. 2011a). As any non-invasive genetic survey for *Martes* species is a two-step process—the collection of "marten-like" faeces in the field and their later identification by genetic analyses in the laboratory, we included as a survey covariate the abundance of "marten-like" faeces (FA = number of "marten-like" faeces per km), recording each survey as either 0 or 1 according to the results of genetic analyses.

Before the analyses, all covariates were standardised by the Z transformation (Donovan and Hines 2007). Models were first ranked according to Akaike Information Criteria values, excluding those showing delta values ≥2 (Burnham and Anderson 2002). The goodness of fit of the models was then assessed by Mackenzie and Bailey's test using 1,000 simulated bootstrap detection histories (MacKenzie and Bailey 2004). To assess fit, the test use a Chi square approach, while the c-hat statistic is used for adjusting the standard errors for the model parameters (c-hat >1 indicates a lack of fit).

Naïve probabilities (sensu MacKenzie et al. 2003) were calculated as the number of transects or sites positive for pine marten over the total number of transects or sites surveyed.

The minimum number of surveys required to statistically establish the occurrence of the pine marten was assessed by the probability model: $N_{min} = \log(\alpha)/\log(1-p)$, where $\alpha = 0.05$ sets the confidence level (McArdle 1990; Reed 1996). To account for modelling results, mean p was calculated for three classes of faeces abundance—FA < 0.5 (N = 5), $0.51 <$ FA < 1 (N = 6), FA > 1.1 (N = 6)—using top model estimates.

Results

Overall 160 "marten-like" faecal samples were collected inside 21 of the 2×2 km large sampling squares (84 km^2; Fig. 2), on a total length of 273.9 km of transects (mean transect length ± SD = 3.33 ± 1.45 km, min–max = 1.1–8.0 km), corresponding to 1.9 faeces/km^2 and 0.58 faeces/km, respectively. No time-related variation in marking intensity was recorded ($\chi^2 = 3.81$, $d.f. = 3$, $p = 0.28$).

DNA was extracted from 123 (76.9 %) faecal samples, out of which 91 were assigned by our PCR–RFLP method to the pine marten and nine to the stone marten, corresponding to an overall species identification of 81.3 % of analysed samples.

The pine marten occurred in all surveyed squares, averaging 4.3 ± 3.15 samples per square (min–max = 1–10; Fig. 2), while stone marten samples were found in only two squares. Grid squares overlapping the river showed 30–50 % of woodland, while inside neighbouring squares the percentage of woodland decreased down to 0.9 % (Table 3).

Most pine marten faecal samples (78.8 %) were collected inside or at the margins of woods, whilst 16 (17.6 %) were found in poplar plantations or crops. Along surveyed transects, faeces were not distributed according to habitat availability ($\chi^2 = 15.6$, $d.f. = 5$, $p < 0.001$), being mainly sited in wooded areas, whilst fields were avoided (Fig. 3).

At the grid-level of analysis, the stepwise regression of the index MI on the habitat variables yielded a significant model ($R^2 = 0.28$, $F = 7.49$, $d.f. = 20$, $p = 0.013$), which included only the mean area of

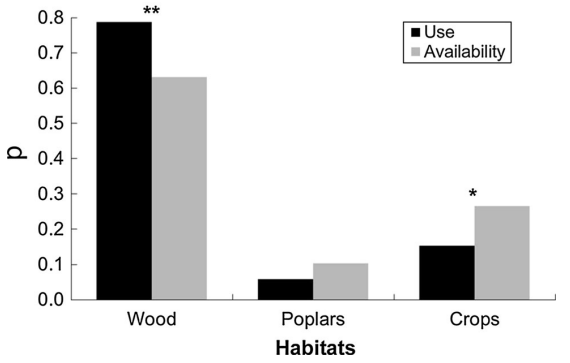


Fig. 3 Habitat selection by pine marten as assessed at transect-scale by the comparison of the observed ("use") and expected ("availability") frequencies (p) of faecal samples (statistical significance according to Bonferroni's confidence intervals; $*p < 0.01$, $**p < 0.001$)

Table 1 Models resulting from linear multiple regressions of pine marten marking intensity MI, ranked according to the Akaike Information Criterion (AICc)

Explanatory variables included in the model	R^2	AICc	ΔAICc	AICc weight
MWA	0.282	−52.166	0.000	0.14
MPAR	0.226	−50.573	1.5935	0.07

MWA mean area of wooded patches, *MPAR* mean perimeter-area ratio of wooded patches, *P* percent cover of poplar plantations, *ΔAICc* delta AICc, see "Methods" section

wooded patches (MWA), positively related to pine marten marking intensity ($\beta = 0.57$). Shapiro–Wilk's test shoved a normal distribution of residuals ($W = 0.93$, $p = 0.11$). The AICc selected three models (Table 1). The best model included the mean area of wooded patches (MWA), positively related to pine marten marking intensity, while the second model included the mean perimeter-area ratio of wooded patches, negatively related to marking intensity ($\beta = -0.46$; Fig. 4).

On the first sampling occasion, the pine marten was genetically ascertained for 15 out of the 17 sampling squares which were surveyed three times (Table 4). One new site was identified during the second survey and an additional one in the third survey. On average, 80.4 % of squares yielded genetically identified pine marten samples from a single survey.

Two occupancy models showed to be supported by the data (ΔAIC scores < 2; Table 2). The 2nd-ranked model was the one which best fitted the occupancy framework (MacKenzie and Bailey's test: $\chi^2 = 0.031$, $p = 0.18$; c-hat = 0.26). Both models supported the inclusion of only FA as a covariate, positively related to the probability of detection. According to the best model, pine marten detection probability was similar in the first and third surveys ($p \pm SE = 0.88 \pm 0.034$ and 0.80 ± 0.012, respectively) and decreased in the second one ($p \pm SE = 0.71 \pm 0.041$). These values almost coincided with naïve probabilities (0.88, 0.82 and 0.7, respectively), as so as for the probability of site occupancy ($\Psi = 1$ for both raw data and top model) (Tables 3, 4).

According to the probability model, the minimum number of surveys needed to ascertain pine marten occurrence is 3.9 (CI = 3.0–5.2) where $FA < 0.5$ and 1.67 (1.66–1.69) where $0.51 < FA < 1$, while only one survey is needed where, on average, more than one "marten-like" faecal sample is found per kilometre of transect.

Discussion

Although in the last two decades of the 20th century sampling in plain areas was probably biased towards the stone marten, indirect signs of presence having been assigned to this species due to its well-known anthropophilia (Sacchi and Meriggi 1995; Lanszki 2003; Herr

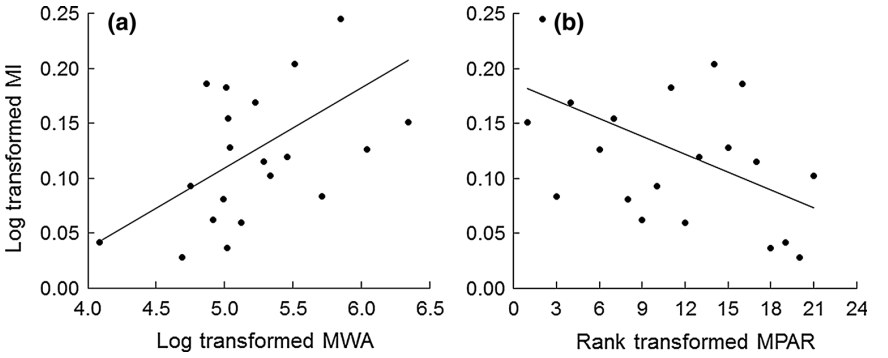


Fig. 4 Relationship between pine marten marking intensity (MI) and both the mean area (MWA) (**a**) and mean perimeter-area ratio (MPAR) (**b**) of wooded patches into 21, 2×2 km wide grid squares. The two linear regressions selected by second order Akaike Information Criterion are shown (see "Methods" section)

Table 2 Top occupancy models obtained by PRESENCE (Hines 2006) and ranked according to Akaike Information Criteria (AIC)

Model	AIC	ΔAIC	AIC weight	Model likelihood	−2Log (likelihood)
Ψ(.), *p*(int) *p*(FA)	14.49	0.00	0.3353	10,000	8.49
Ψ(.), *p*(int) p_1(FA) p_2(FA) p_3(FA)	14.79	0.30	0.2886	0.8607	4.79

$p_{1,2,3}$: detection probability during each survey; int: intercept; Ψ: probability of site occupancy; FA: marten-like faeces km^{-1}; ΔAICc: delta AICc, see "Methods" section

Table 3 Habitat variables of potential importance to pine marten distribution, total length of surveyed linear transects and pine marten marking intensity (MI) in 21, 2 × 2 km wide grid squares

Grid square	W (%)	U (%)	P (%)	C (%)	H	minDr (km)	MWA (km^2)	MPAR	TL (km)	MI
1	17.7	1.7	1.0	79.5	0.29	4.61	0.70	0.009	7.92	0.76
2	8.9	6.8	2.7	81.6	1.24	3.96	0.06	0.031	4.19	0.24
3	13.5	7.9	9.1	69.3	1.57	3.05	0.13	0.032	6.77	0.15
4	11.0	0.7	6.7	81.3	1.18	1.97	0.11	0.027	9.34	0.43
5	5.8	1.5	5.5	86.5	1.22	3.22	0.05	0.094	15.27	0.07
6	0.9	0.5	9.7	88.8	0.71	1.64	0.01	0.062	10.07	0.10
7	13.8	1.7	22.0	62.1	1.72	2.88	0.11	0.036	11.64	0.34
8	3.9	0.8	12.7	82.4	1.17	0.89	0.07	0.038	7.48	0.53
9	42.2	0.3	7.4	42.7	1.96	0.00	0.51	0.021	37.93	0.21
10	52.5	1.6	0.0	36.7	1.81	0.00	0.70	0.024	8.96	1.01
11	22.1	5.0	21.4	51.4	1.88	1.16	0.10	0.031	13.37	0.52
12	54.4	0.8	18.4	26.4	1.67	0.41	1.09	0.024	29.78	0.34
13	47.5	0.1	16.8	21.5	2.53	0.00	0.33	0.034	15.02	0.60
14	53.7	0.0	15.9	28.2	2.29	0.00	0.19	0.046	23.04	0.30
15	26.2	22.2	23.0	20.7	2.87	0.00	0.17	0.021	6.30	0.48
16	57.8	0.0	17.7	8.5	2.22	0.00	0.22	0.098	26.31	0.27
17	12.3	0.7	15.6	70.3	1.44	0.19	0.10	0.030	9.73	0.21
18	20.6	1.7	0.0	68.6	1.99	0.00	0.08	0.031	6.55	0.15
19	36.5	1.2	5.9	31.7	2.73	0.00	0.10	0.048	11.45	0.09
20	30.0	18.4	5.4	29.1	2.49	0.00	0.28	0.034	3.16	0.32
21	59.3	0.0	2.7	35.1	1.23	0.00	2.21	0.007	9.61	0.42

W, U, P, C: percent cover of woods, urban areas, poplar plantations and crops, respectively, into each grid square; H: Shannon's index; minDr: minimum distance to the river; MWA: mean area of wooded patches; MPAR: mean perimeter-area ratio of wooded patches; TL: transect length; MI: marking intensity, expressed as number of faeces km^{-1}; see "Methods" section

et al. 2009), our results draw a quite different picture, with the pine marten as the most widespread *Martes* species on the River Ticino (in terms of the number of both genotyped samples—91 versus 9—, and positive squares—21 versus 2).

As supported by the exponential increase in road-kill reports (Balestrieri et al. 2010), the spreading of the pine marten in the valley may have occurred during the last 15 years. This rapid expansion stresses the role of riparian woods as a suitable corridor for pine marten colonisation of a landscape largely dominated by agriculture and artificial land-cover. Unfortunately, information about the status of the pine marten in Italy is insufficient to determine the causes of pine marten expansion, although, on the Alps, increased forest cover may have favoured the growth of pine marten populations (O'Mahony et al. 2012), as occurred for other forest species (e.g. Jepsen et al. 2004).

In our study area, the distribution of pine marten samples confirmed its preference for woodlands. Although signs were found not only in large woods but also in small, isolated wood plots and on wooded slopes and canal banks surrounded by agricultural land, the relation between marking intensity and mean wood size and perimeter-area ratio suggests that pine marten relative abundance depends on the structure and degree of fragmentation of residual woods (see also Pereboom et al. 2008).

Neither marking intensity nor occupancy rates decreased with the distance of the sampling transects from the river. As the main aim of the study was to assess pine marten distribution and habitat use along a potential river corridor, transect distance from the river never exceeded 5 km. Thus, our survey-scale may have been inadequate to detect the effect of this habitat variable.

Our results are consistent with those obtained by radiotelemetry in rural areas of France (Pereboom et al. 2008) and Scotland (Caryl et al. 2012), suggesting that the pine marten is more generalist in terms of habitat preferences than previously reported (Virgós et al. 2012). Such ecological generalization should increase the likelihood that individuals will find suitable resources in a new area (Mettke-Hofmann et al. 2002; Echeverría et al. 2006; Moritz et al. 2008; Poyry et al. 2009), including man-altered habitats (Musiani et al. 2003). Accordingly,

Table 4 Detection histories and survey- and site-covariates for each sampling station

Site	Detection history			FA (marten like faeces km^{-1})			L	Wa	Wp	Mean Dr	Dnw	Du	Wt	W	U	C
				S 1	S 2	S 3										
1	0	0	1	0.00	0.00	0.25	4.0	0.60	1.70	4.63	2.80	1.23	10.75	6.13	1.63	91.54
2	0	1	1	0.00	0.23	0.69	4.34	2.29	20.56	0.41	0.05	2.50	74.10	28.50	17.53	37.80
3	1	0	0	0.53	0.00	0.00	3.74	6.96	62.47	0.62	0.05	3.45	82.20	54.38	0.80	44.77
4	1	0	0	0.72	0.00	0.00	2.77	0.60	1.70	5.05	3.60	0.34	37.10	6.13	1.63	91.54
5	1	1	1	0.12	0.62	0.25	8.01	6.96	62.47	0.56	0.05	0.99	41.10	38.47	0.30	54.42
6	1	0	1	0.88	0.00	0.88	2.26	0.05	1.90	3.71	0.43	2.28	39.10	13.60	1.65	84.33
7	1	0	1	1.34	0.00	0.69	3.72	0.06	17.58	3.03	2.40	1.20	13.70	0.92	0.52	98.50
8	1	1	1	0.25	1.31	0.65	3.06	2.29	20.56	0.49	0.05	3.39	92.00	48.93	0.00	37.54
9	1	1	1	0.97	0.38	1.12	4.13	6.96	62.47	2.23	0.05	1.99	53.40	54.38	0.80	44.77
10	1	1	1	0.62	1.40	0.83	4.80	2.29	20.56	0.55	0.05	2.20	63.20	52.82	0.00	44.95
11	1	1	0	0.61	2.46	0.00	1.63	6.96	62.47	2.18	0.05	2.34	53.30	3.72	0.75	95.28
12	1	1	1	1.99	0.66	0.66	3.01	0.60	2.90	5.34	4.50	1.02	58.70	17.74	1.72	80.53
13	1	1	1	0.29	0.61	2.45	3.26	1.20	8.12	1.48	0.25	1.20	98.20	52.82	0.00	44.95
14	1	1	1	1.03	0.54	2.15	2.93	0.27	3.96	2.94	0.03	0.61	32.30	23.02	5.17	71.67
15	1	1	1	1.39	2.71	1.16	2.58	6.96	62.47	0.60	0.05	2.39	79.10	40.88	0.09	46.90
16	1	1	1	0.56	3.61	2.72	3.57	2.84	10.82	0.42	0.75	1.33	68.60	52.48	1.59	36.73
17	1	1	1	2.87	1.91	1.59	3.54	0.99	15.22	0.25	0.04	2.18	95.00	12.34	0.69	85.86
Mean				0.83	0.98	0.95	3.61	2.88	25.76	2.03	0.89	1.80	58.34	29.84	2.05	64.24
SD				0.75	1.09	0.84	1.38	2.84	25.32	1.79	1.46	0.90	26.86	20.53	4.17	23.57

(S: survey; L: length, km; Wa: wood area, km^2; Wp: wood perimeter, km; Dr: distance to the river, km; Dnw: minimum distance between wood patches, km; Du: mean distance to urban areas, km; Wt: percent length of wood covered by each transect; W, U, C: percent cover of woods, urban areas and crops, respectively, into each grid square; see the "Methods" for details)

diet analyses suggested that in agricultural habitats the pine marten can cope with the reduced availability of small mammals by relying on introduced Eastern cottontail *Sylvilagus floridanus* (Balestrieri et al. 2011b).

The aims and spatial scale of occupancy surveys drive the choice of survey protocols. For large scales, the need to maximise the cost-effectiveness of the surveys may lead to lower the chance of detecting the target species (Slauson et al. 2009). By the non-invasive genetic method adopted, species identification was obtained for a rather high percentage of faecal samples with respect to other available approaches (e.g. 58 %, Lucentini et al. 2007; 53.4 %, Pilot et al. 2006), confirming the results of previous studies carried out by the same method (Rosellini et al. 2008; Ruiz-González et al. 2008, 2013; Balestrieri et al. 2010, 2011b). Although sample size was small, multiple surveys allowed us to point out that, as previously reported for the Eurasian otter *Lutra lutra* (Balestrieri et al. 2011a), heterogeneity in detection probability and hence the effort (i.e. the minimum number of surveys) needed to ascertain the presence of the target species may arise as a result of variation in the number of marking individuals (see also Kéry 2002; Balestrieri et al. 2011a, b), which should be accounted for by occupancy models (Royle and Nichols 2003).

Variation in p over the three surveys suggests that it may partially depend on covariates different from those tested, such as rain or cloud cover, which can affect the ability of surveyors to detect faecal samples (Olson et al. 2011). We suggest that future survey protocols have to involve multiple surveys, adjusting the overall effort per sample unit on the basis of the number of "marten-like" faeces recorded per km of transect.

In regions exposed to intense anthropogenic pressure, riparian areas are severely threatened, in spite of their crucial role for landscape connectivity (Clerici and Vogt 2013). Although further studies are needed to assess the actual pine marten distribution in the lowlands of northern Italy, on the basis of our results we can hypothesise the availability of both woodland corridors and wood patches spread in the crop matrix to be major factors shaping the distribution of this marten species in the Po-Venetian plain. The PCR–RFLP method adopted in this study, if combined with the search for faecal samples by trained surveyors (see also Ruiz-González et al. 2013), represents a cost-effective tool for future investigations on pine marten distribution at a broader scale.

Acknowledgments This study has been partially supported by the Basque Government through the Research group on "Systematics, Biogeography and Population Dynamics" (Ref. IT317-10; IT575-13). Aritz Ruiz-González (Ref: DKR-2,012-64) and Maria Vergara (Ref: RBFI-2,012-446) were supported by a post-doctoral and PhD fellowships awarded by the Dept. of Education, Universities and Research of the Basque Government. We thank Pietro Tirozzi and Andrea Serioli, who helped with sample collection in the field for their degree theses. Lesley C. Wright kindly revised the paper for English language use. The comments of two anonymous reviewers helped to improve the final version of the manuscript.

References

Balestrieri A, Remonti L, Prigioni C (2009) Habitat selection in a low-density badger *Meles meles* population: a comparison of radio-tracking and latrine surveys. Wildl Biol 15:1–7

Balestrieri A, Remonti L, Ruiz-González A, Gómez-Moliner BJ, Vergara M, Prigioni C (2010) Range expansion of the pine marten (*Martes martes*) in an agricultural landscape matrix (NW Italy). Mamm Biol 75:412–419

Balestrieri A, Remonti L, Prigioni C (2011a) Detectability of the Eurasian otter by standard surveys: an approach using marking intensity to estimate false negative rates. Naturwissenschaften 98:23–31

Balestrieri A, Remonti L, Ruiz-González A, Vergara M, Capelli E, Gómez-Moliner BJ, Prigioni C (2011b) Food habits of genetically identified pine marten (*Martes martes*) expanding in agricultural lowlands (NW Italy). Acta Theriol 56:199–207

Balharry D (1993) Social organization in martens: an inflexible system? Symp Zool Soc Lond 65:321–345

Beier P (1993) Determining minimum habitat areas and habitat corridors for cougars. Conserv Biol 7:94–108

Bennett AF (2003) Linkages in the landscape: the role of corridors and connectivity in wildlife conservation. IUCN, Gland, p 254

Bentley J, Catterall C (1997) The use of bushland, corridors, and linear remnants by birds in southeastern Queensland, Australia. Conserv Biol 11:1173–1189

Benton TG, Vickery JA, Wilson JD (2003) Farmland biodiversity: is habitat heterogeneity the key? Trends Ecol Evol 18:182–188

Bright PW (1993) Habitat fragmentation: problems and predictions for British Mammals. Mammal Rev 23:101–111

Burnham KP, Anderson DR (2002) Model selection and inference—a practical information-theoretic approach, 2nd edn. Springer, New York 488 pp

Buskirk SW (1992) Conserving circumboreal forests for martens and fishers. Conserv Biol 6:318–320

Buskirk SW, Zielinski WJ (2003) Small and mid-sized carnivores. In: Zabel CJ, Anthony RG (eds) Mammalian community dynamics: management and conservation in the coniferous forests of western North America. Cambridge University Press, Cambridge, pp 207–249

Byers CR, Steinhorst RK, Krausman PR (1984) Clarification of a technique for analysis of utilization-availability data. J Wildl Manag 48:1050–1053

Camerano P, Grieco C, Terzuolo PG (2010) I boschi planiziali: conoscenza, conservazione e valorizzazione. Regione Piemonte, Blu edizioni, Torino, p 167

Caryl FM, Quine CP, Park KJ (2012) Martens in the matrix: the importance of nonforested habitats for forest carnivores in fragmented landscapes. J Mamm 93:464–474

Clavero M, Prenda J, Delibes M (2006) Seasonal use of coastal resources by otters: comparing sandy and rocky stretches. Estuar Coast Shelf Sci 66:387–394

Clerici N, Vogt P (2013) Ranking European regions as providers of structural riparian corridors for conservation and management purposes. Int J Appl Earth Obs 21:477–583

Coles TF, Southey JM, Forbes I, Clough T (1989) River wildlife data bases and their value for sensitive environmental management. Regul Rivers 4:179–189

Davison A, Birks JDS, Brookes RC, Braithwaite TC, Messenger JE (2002) On the origin of faeces: morphological versus molecular methods for surveying rare carnivores from their scats. J Zool Lond 257:141–143

De Marinis AM, Masseti M (1993) Pine marten *Martes martes* on the island of Elba. Small Carniv Conserv 8:13

Del Fante S (2012) Comportamento spaziale della martora (*Martes martes*) in ambiente appenninico. Dissertation, Department of Earth and Environmental Sciences, Pavia University

Donovan TM, Hines J (2007) Exercises in occupancy modeling and estimation. http://www.uvm.edu/envnr/vtcfwru/spreadsheets/occupancy/occupancy.htm

Echeverría AI, Vassallo AI, Isacch JP (2006) Experimental analysis of novelty responses in a bird assemblage inhabiting a suburban marsh. Can J Zool 84:974–980

Garshelis DL (2000) Delusions in habitat evaluation: measuring use, selection, and importance. In: Boitani L, Fuller TK (eds) Research techniques in animal ecology: controversies and consequences. Columbia University Press, New York, pp 111–164

Genovesi P, De Marinis AM, (2003a) *Martes foina*. In: Boitani L, Lovari S, Vigna Taglianti A (eds), Fauna d'Italia Mammalia III, Carnivora—Artiodactyla. Calderini, Italy, pp 117–128

Genovesi P, De Marinis AM (2003b) *Martes martes*. In: Boitani L, Lovari S, Vigna Taglianti A (eds), Fauna d'Italia. Mammalia III, Carnivora—Artiodactyla. Ed Calderini, pp 106–117

Gese EM (2001) Monitoring of terrestrial carnivore populations. In: Gittleman JL, Funk SM, Macdonald DW, Wayne RK (eds) Carnivore conservation. Cambridge University Press, Ithaca, pp 372–396

Getz LL, Cole FR, Gates DL (1978) Interstates roadsides as dispersal routes for *Microtus pennsylvanicus*. J Mammal 59:208–212

Guter A, Dolev A, Saltz D, Kronfeld-Schor N (2008) Using videotaping to validate the use of spraints as an index of Eurasian otter (*Lutra lutra*) activity. Ecol Indic 8:462–465

Harris S, Woollard T (1990) The dispersal of mammals in agricultural habitats in Britain. In: Howard DC, Bunce RGH (eds) Species Dispersal In Agricultural Habitats. Belhaven Press, London, pp 159–188

Herr J, Schley L, Roper TJ (2009) Socio-spatial organization of urban stone martens. J Zool Lond 277:54–62

Hilty JA, Merenlender AM (2004) Use of riparian corridors and vineyards by mammalian predators in Northern California. Conserv Biol 18:126–135

Hines JE (2006) PRESENCE2—software to estimate patch occupancy and related parameters. USGS-PWRC. http://www.mbr-pwrc.usgs.gov/software.html

Jepsen JU, Topping CJ (2004) Modelling roe deer (*Capreolus capreolus*) in a gradient of forest fragmentation: behavioural plasticity and choice of cover. Can J Zool 82:1528–1541

Kauhala K, Salonen L (2012) Does a non-invasive method—latrine surveys—reveal habitat preferences of raccoon dogs and badgers? Mamm Biol 77:264–270

Kéry M (2002) Inferring the absence of a species—a case study of snakes. J Wildl Manage 66:330–338

Kohn MH, Wayne RK (1997) Facts from feces revisited. Trends Ecol Evol 12:223–227

Krebs CJ (1989) Ecological methodology. Harper and Row, New York 654 pp

Kruuk H, Conroy JWH, Glimmerveen U, Ouwerkerk EJ (1986) The use of spraints to survey populations of otters *Lutra lutra*. Biol Conserv 35:187–194

Lanszki J (2003) Feeding habits of stone martens in a Hungarian village and its surroundings. Folia Zool 52:367–377

Lanszki J, Hidas A, Szentes K, Révay T, Lehoczky I, Weiss S (2008) Relative spraint density and genetic structure of otter (*Lutra lutra*) along the Drava River in Hungary. Mamm Biol 73:40–47

Lassini P, Monzani F, Pileri P (2007) A green vision for the renewal of the Lombardy landscape. In: Pedroli B, Van Doorn A, De Blust G, Paracchini ML, Wascher D, Bunce F (eds) Europe's living landscapes. Essays on exploring our identity in the countryside, Landscape Europe/KNNV, pp 83–100

Lepers E, Lambin EF, Janetos AC, DeFries R, Achard F, Ramankutty N, Scholes RJ (2005) A synthesis of information on rapid landcover change for the period 1981–2000. Bioscience 55:115–124

Long RA, Donovan TM, MacKay P, Zielinski WJ, Buzas JS (2011) Predicting carnivore occurrence with noninvasive surveys and occupancy modeling. Landsc Ecol 26:327–340

Lucentini L, Vercillo F, Palomba A, Panara F, Ragni B (2007) A PCR-RFLP method on faecal samples to distinguish *Martes martes*, *Martes foina*, *Mustela putorius* and *Vulpes vulpes*. Conserv Genet 8:757–759

MacKenzie DI, Bailey L (2004) Assessing fit of site occupancy models. J Agric Biol Ecol St 9:300–318

MacKenzie DI, Nichols JD, Lachman GB, Droege S, Royle JA, Langtimm CA (2002) Estimating site occupancy rates when detection probabilities are less than one. Ecology 83:2248–2255

MacKenzie DI, Nichols JD, Hines JE, Knutson MG, Franklin AD (2003) Estimating site occupancy, colonization and local extinction probabilities when a species is not detected with certainty. Ecology 84:2200–2207

Mantovani S (2010) Recenti segnalazioni della martora, *Martes martes*, in provincia di Cremona. Pianura 25:95–107
Matson PA, Parton WJ, Power AG, Swift MJ (1997) Agricultural intensification and ecosystem properties. Science 277:504–509
Mazerolle MJ (2006) Improving data analysis in herpetology: using Akaike's information criterion (AIC) to assess the strength of biological hypotheses. Amphibia Reptilia 27:169–180
McArdle BH (1990) When are rare species not there? Oikos 57:276–277
Mergey M, Helder R, Roeder J-J (2011) Effect of forest fragmentation on space-use patterns in the European pine marten (*Martes martes*). J Mammal 92:328–335
Messenger JE, Birks JDS (2000) Monitoring the very rare: pine marten populations in England and Wales. In: Griffiths HI (ed) Management and conservation aspects of small carnivore: human interactions. Backhuys Publishers, Leiden, pp 217–230
Mettke-Hofmann C, Winkler H, Leisler B (2002) The significance of ecological factors for exploration and neophobia in parrots. Ethology 108:249–272
Moritz C, Patton JL, Conroy CJ, Parra JL, White GC, Beissinger SR (2008) Impact of a century of climate change on small-mammal communities in Yosemite National Park, USA. Science 322:261–264
Mortelliti A, Amori G, Capizzi D, Rondinini C, Boitani L (2010) Experimental design and taxonomic scope of fragmentation studies on European mammals: current status and future priorities. Mammal Rev 40:125–154
Musiani M, Mamo C, Boitani L, Callaghan C, Gates CC, Mattei L, Visalberghi E, Breck S, Volpi G (2003) Wolf depredation trends and the use of fladry barriers to protect livestock in Western North America. Conserv Biol 17:1538–1547
Naiman RJ, Décamps H, Pollock M (1993) The role of riparian corridors in maintaining regional biodiversity. Ecol Appl 3:209–212
Neu CW, Byers CR, Peek JM, Boy V (1974) A technique for analysis of utilization-availability data. J Wildl Manag 38:541–545
O'Mahony D, O'Reilly C, Turner P (2012) Pine marten (*Martes martes*) distribution and abundance in Ireland: a cross-jurisdictional analysis using non-invasive genetic survey techniques. Mamm Biol 77:51–357
Olson CA, Beard KH, Koons DN, Pitt WC (2011) Detection probabilities of two introduced frogs in Hawaii: implications for assessing non-native species distributions. Biol Invasions. doi:10.1007/s10530-011-0125-5
Pereboom V, Mergey M, Villerette N, Helder R, Gerard JF, Lodé T (2008) Movement patterns, habitat selection, and corridor use of a typical woodland-dweller species, the European pine marten (*Martes martes*), in fragmented landscape. Can J Zool 86:983–991
Pilot M, Gralak B, Goszczyński J, Posłuszny M (2006) A method of genetic identification of pine marten (*Martes martes*) and stone marten (*Martes foina*) and its application to faecal samples. J Zool Lond 271:140–147
Posada D, Buckley T (2004) Model selection and model averaging in Phylogenetics: advantages of akaike information criterion and bayesian approaches over likelihood ratio tests. Syst Biol 53:793–808
Pőyry J, Luoto M, Heikkinen RK, Kuussaari M, Saarinen K (2009) Species traits explain recent range shifts of Finnish butterflies. Glob Change Biol 15:732–743
Prigioni C (1995) Guidelines for the feasibility study of reintroduction of the otter *Lutra lutra* in Italy: the project of the Ticino Valley (north-western Italy). In: Prigioni C (ed), Proc II It Symp on Carnivores Hystrix (ns) 7(1–2):225–264
Prigioni C, Balestrieri A (2011) Piano di monitoraggio dei Vertebrati terrestri di interesse comunitario in Lombardia: lontra *Lutra lutra*. Università di Pavia, Dipartimento di Biologia Animale 30 pp
Putman RJ (1984) Facts from faeces. Mammal Rev 14:9–97
Quinn GP, Keough MJ (2002) Experimental design and data analysis for biologists. Cambridge University Press, Cambridge
Rangel TF, Diniz-Filho JAF, Bini LM (2010) SAM: a comprehensive application for spatial analysis in macroecology. Ecography 33:1–5
Reed JM (1996) Using statistical probability to increase confidence of inferring species extinction. Conserv Biol 10:1283–1285
Reed DH (2004) Extinction risk in fragmented habitats. Anim Conserv 7:181–191
Remonti L, Balestrieri A, Ruiz-González A, Gómez-Moliner BJ, Capelli E, Prigioni C (2012) Intraguild dietary overlap and its possible relationship to the coexistence of mesocarnivores in intensive agricultural habitats. Popul Ecol 54:521–532
Rosalino LM, Santos MJ, Beier P, Santos-Reis M (2008) Eurasian badger habitat selection in Mediterranean environments: does scale really matter? Mamm Biol 73:189–198
Rosellini S, Osorio E, Ruiz-González A, Pineiro A, Barja I (2008) Monitoring the small-scale distribution of sympatric European pine martens (*Martes martes*) and stone martens (*Martes foina*): a multievidence approach using faecal DNA analysis and cameratraps. Wildl Res 35:434–440
Royle JA, Nichols JD (2003) Estimating abundance from repeated presence-absence data or point counts. Ecology 84:777–790
Ruiz-González A, Rubines J, Berdión O, Gómez-Moliner BJ (2008) A non-invasive genetic method to identify the sympatric mustelids pine marten (*Martes martes*) and stone marten (*Martes foina*): preliminary distribution survey on the northern Iberian peninsula. Eur J Wildl Res 54:253–261
Ruiz-González A, Madeira MJ, Randi E, Urra F, Gómez-Moliner BJ (2013) Non invasive genetic sampling of sympatric marten species (*Martes martes* and *Martes foina*): assessing species and individual identification success rates on faecal DNA genotyping. Eur J Wildl Res 59:371–386
Sacchi O, Meriggi A (1995) Habitat requirements of the stone marten (*Martes foina*) on the Tyrrhenian slopes of the northern Apennines. In: Prigioni C (ed), Proc III Symp on Carnivores Hystrix (ns) 7:99–104
Sadlier MJ, Webbon CC, Baker PJ, Harris S (2004) Methods of monitoring foxes *Vulpes vulpes* and badgers *Meles meles*: are field signs the answer? Mammal Rev 34:75–98
Slauson KM, Baldwin J, Zielinski WJ, Schwartz M (2009) Estimating detection probabilities for fishers using non-invasive methods and implications for survey protocols. U.S. Fish and Wildlife Service, Yreka Field Office and the Fisher Science Team, pp 68
Sugiura N (1978) Further analysis of the data by Akaike's information criterion and the finite corrections. Comm Statist Theory Methods A7:13–26
Taylor PD, Fahrig L, Henein K, Merriam G (1993) Connectivity is a vital element of landscape structure. Oikos 68:571–573
Virgós E, Zalewski A, Rosalino LM, Mergey M (2012) Habitat ecology of genus *Martes* in Europe: a review of the evidences. In: Aubry KB, Zielinski WJ, Raphael MG, Proulx G, Buskirk SW (eds) Biology and conservation of marten, sables, and fisher: a new synthesis. Cornell University Press, New York, pp 255–266
Waldron JL, Bennett SH, Welch SM, Dorcas ME, Lanham JD, Kalinowsky W (2006) Habitat specificity and home-range size as attributes of species vulnerability to extinction: a case study using sympatric rattlesnakes. Anim Conserv 9:414–420
Warkentin IG, Greenberg R, Ortiz JS (1995) Songbird use of gallery woodlands in recently cleared and older settled landscapes of the Selva Lacandona, Chiapas, Mexico. Conserv Biol 9:1095–1106
White GC, Garrott RA (1990) Analysis of wildlife radio-tracking data. Academic Press, USA, p 386
Zalewski A, Jedrzejewski W (2006) Spatial organisation and dynamics of the pine marten *Martes martes* population in Bialowieza Forest (E Poland) compared with other European woodlands. Ecography 29:31–43

Chapter III

PAPER 3: Inferring population genetic structure in widely and continuously distributed carnivores: the stone marten (*Martes foina*) as a case study

María Vergara, Mafalda P. Basto, Mª José Madeira, Benjamín J. Gómez-Moliner, Margarida Santos-Reis, Carlos Fernandes and Aritz Ruiz-González

Under review in PlosOne

Abstract

The stone marten is a widely distributed mustelid in the Palaearctic region that exhibits variable habitat preferences in different parts of its range. The species is a Holocene immigrant from southwest Asia which, according to fossil remains, followed the expansion of the Neolithic farming cultures into Europe and possibly colonized the Iberian Peninsula during the Early Neolithic (ca. 7,000 years BP). However, the population genetic structure and historical biogeography of this generalist carnivore remains essentially unknown. In this study we have combined mitochondrial DNA (mtDNA) sequencing (620 bp) and microsatellite genotyping (23 polymorphic markers) to infer the population genetic structure of the stone marten within the Iberian Peninsula. The mtDNA data revealed low haplotype and nucleotide diversities and a lack of phylogeographic structure, most likely due to a recent colonization of the Iberian Peninsula by a few mtDNA lineages during the Early Neolithic. The microsatellite data set was analysed with a) spatial and non-spatial Bayesian individual-based clustering (IBC) approaches (STRUCTURE, TESS, BAPS and GENELAND), and b) multivariate methods [Discriminant Analysis of Principal Components (DAPC) and spatial principal component analysis (sPCA)]. Additionally, because isolation by distance (IBD) is a common spatial genetic pattern in mobile and continuously distributed species and it may represent a challenge to the performance of the above methods, the microsatellite data set was tested for its presence. Overall, the genetic structure of the stone marten in the Iberian Peninsula was characterized by a NE-SW spatial pattern of IBD, and this may explain the observed disagreement between clustering solutions obtained by the different IBC methods. However, there was significant indication for contemporary genetic structuring, albeit weak, into at least three different subpopulations. The detected subdivision could be attributed to the influence of the rivers Ebro, Tagus and Guadiana, suggesting that main watercourses in the Iberian Peninsula may act as semi-permeable barriers to gene flow in stone martens. To our knowledge, this is the first phylogeographic and population genetic study of the species at a broad regional scale. We also wanted to make the case for the importance and benefits of using and comparing multiple different clustering and multivariate methods in spatial genetic analyses of mobile and continuously distributed species.

Keywords: Genetic structure; Bayesian clustering; Multivariate methods; Isolation by distance; Microsatellites; mtDNA; *Martes foina*; Iberian Peninsula.

Introduction

In continuous populations, genetic structure vary from a single identifiable group of individuals, with gene flow only restricted by the distance among them, to an unknown number of subpopulations separated by barriers to gene exchange [1,2]. Species with high dispersal abilities and weak habitat specificity are expected to show little population genetic differentiation (e.g. [3]). However, recent studies conducted on broadly distributed species have revealed cryptic patterns of genetic structuring in apparently continuous populations due to historical processes and/or the presence of major topographic and landscape features [4–6]. Population and landscape genetic studies often apply individual-based Bayesian clustering (IBC) algorithms to detect genetic discontinuities, which are subsequently correlated with landscape characteristics [1,7–10]. The application of IBC methods to data sets characterized by a pattern of isolation by distance (IBD) may, however, often result in an overestimation of population subdivision, especially when the sampling design is irregular [1,11,12]. Thus, [13] suggested comparing the outputs of several IBC algorithms, in addition to testing for the presence of IBD, to infer a genetic structuring that could be empirically explained. The performance of IBC methods in landscape genetics has been extensively demonstrated in studies and simulations comparing different approaches [1,8,10,11,14–17]. STRUCTURE is one of the most widely used IBC algorithms to infer the number of clusters in a data set and assign individuals to those clusters based exclusively on their multilocus genotypes [18,19]. Conversely, IBC programs such as BAPS [20], GENELAND [14] or TESS [21,22], allow the use of geo-referenced samples in their spatially-explicit models. The spatial prior (i.e. the geographic location of each sample) may provide more support to clustering solutions and be particularly helpful in the case of sparse sampling [23]. Multivariate ordination analyses (e.g. sPCA and DAPC) have been suggested as an alternative to IBC algorithms because they do not make any assumption about the underlying population genetic model [24,25]. Thus, a consensual solution from the combination of different approaches (i.e. Bayesian and multivariate analyses) can be used to obtain reliable inferences on spatial genetic patterns [10,15,17], which may be particularly cryptic and/or complex in mobile and continuously distributed taxa [3].

The stone marten, (*Martes foina,* Erxleben, 1777) is one of the most widely distributed mustelids in Eurasia [26]. It is present in the Middle East, central Asia and in central and southern Europe, with the Iberian Peninsula representing the western limit of its distribution. In parts of its range the stone marten is associated with human-dominated environments (e.g. [27]) but in Iberia the species prefers mosaic habitats [28], forest, scrublands

and rocky areas rather than urbanized environments [29–31]. Nevertheless, its generalist ecological requirements contribute to the species being locally abundant and ubiquitous in the Iberian Peninsula. The limited fossil evidence suggests that stone martens are present in this region since the middle Holocene (*ca.* 7,000 years before present (BP) [32,33]), as a result of a range expansion by the species from southwest Asia into Europe early in the Holocene following the expansion of the Neolithic farming cultures of the Near East [32–34]. However, the details of the historical biogeography of the stone marten in Europe are essentially unknown. In any case, the Iberian Peninsula, because it is only connected to the rest of continental Europe by the Pyrenean isthmus, offers an interesting scenario to investigate how landscape barriers shape the distribution and population structure of the species as it expanded throughout the peninsula after its arrival in Iberia.

Several genetic studies within *Martes* species with a phylogeographic or taxonomic focus have relied on mtDNA sequencing (mainly D-loop and Cytochrome *b* regions) because the relatively fast evolutionary rate of mtDNA is well suited to examine events that occurred hundreds of thousands to a few million years ago (reviewed in [35]). More recently, the number of studies on *Martes* spp. from population or conservation genetic perspectives and using microsatellites has steadily increased [6,36,37]. In contrast, either phylogeographic or population and landscape genetic investigations on the stone marten are virtually absent from literature [35], with the exception of two country-level studies [38,39]. Research on the genetic structure of *Martes foina* should be particularly interesting and challenging because it is more generalist than many other marten species [40].

In this study, we use mtDNA and microsatellite markers to investigate the phylogeography and genetic structure of *Martes foina* in the Iberian Peninsula. This is the first thorough phylogeographic and population genetic survey of this species at a broad regional scale. The results of this work shed light on both the species' history and current population structure in Iberia. We also use this case study to illustrate and argue for a consensual approach comparing complementary methods with distinct assumptions (Bayesian clustering algorithms and multivariate analyses) to accurately characterize genetic structure and to identify processes shaping gene flow.

Material and methods

Ethic Statement

Samples included in the present study (S1 Table) were obtained from different sources; a) specimens from museum collections (n=68), b) samples from Wildlife Rescue Centres (WRC) collected by veterinaries with the permission of the corresponding regional wildlife authority (n=77), c) samples from road-killed individuals collected by wildlife researchers, veterinaries and rangers (n=154) and additionally, d) hair samples from live-trapped individuals obtained in the framework of field studies for other purpose than this study (n=53). Further details for each sample regarding its origin, type of collection, collector and permission obtained (when requested) are listed in S1 Table. No animals were sacrificed or captured for this study. Therefore, a specific formal approval by an Institutional Animal Care and Use Committee was not necessary.

Study area

The study area was the Iberian Peninsula, a region of 582,000 km^2 at the south-western edge of Europe and comprising the mainland territories of Spain and Portugal and the country of Andorra (Fig. 1). On the north, west and southwest it is bordered by the Atlantic Ocean and on the east and southeast by the Mediterranean Sea, while the Pyrenees form the northeast edge of the peninsula, separating it from the rest of Europe. Iberia has a relatively high average altitude (600 m) due to the presence at the centre of a vast plateau, known as the Meseta. A series of other mountain ranges around the plateau and others located on the periphery of the peninsula complete the topographic outline. The major rivers, Ebro, Douro, Tagus, Guadiana and Guadalquivir, flow through the wide valleys between mountain systems. The Iberian Peninsula has the most varied mosaic of climates in Europe and represents the western limit of the stone marten distribution ([26], Fig. 1).

Sample collection and DNA extraction

A total of 599 tissue and hair samples (Spain: n= 509; Portugal: n= 90) were obtained from museum collections and wildlife rescue centres or collected from road-killed or live-trapped stone martens throughout the Iberian Peninsula (Fig. 1). Tissue samples were preserved in ethanol 96% and frozen at -20ºC, while hair samples were stored in dry tubes until DNA extraction. To monitor potential contaminations, we included one negative extraction control

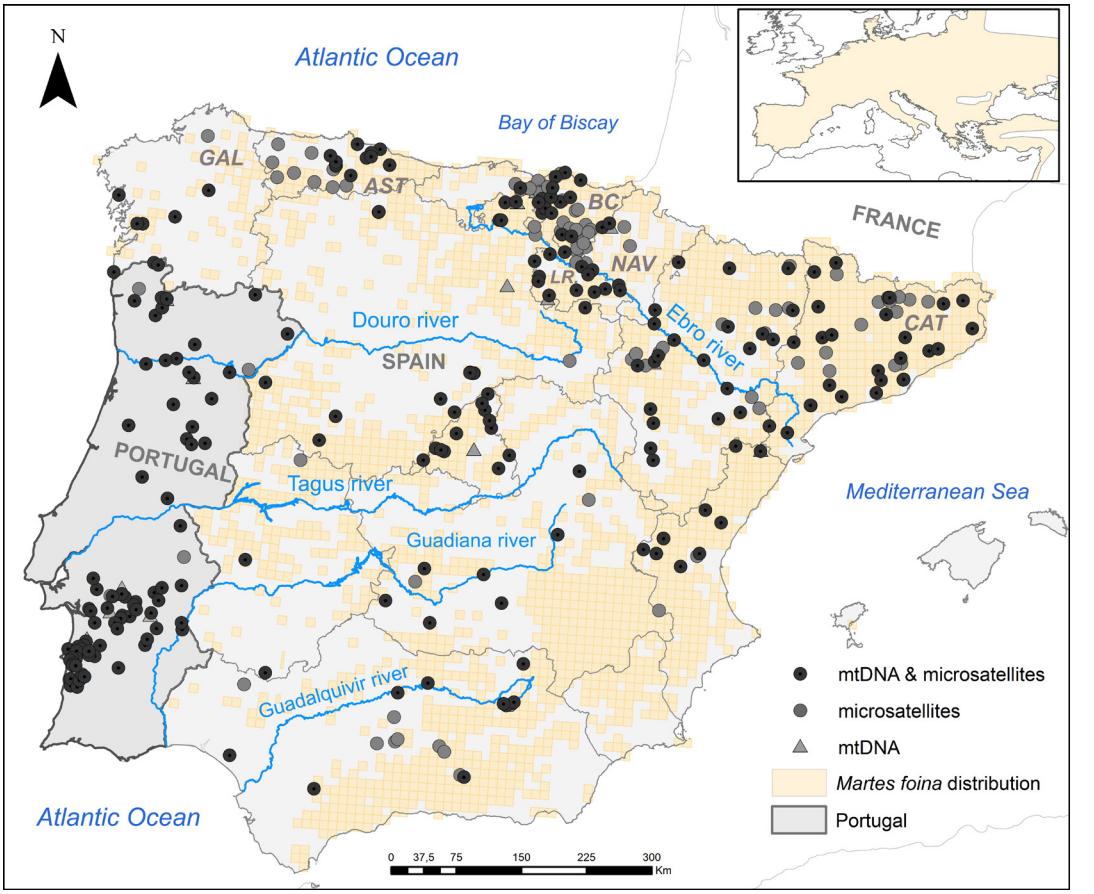


Figure 1. Map of the study area and of the stone marten distribution in Spain [41] with the geographic location of each genotyped (microsatellites) and/or sequenced (mtDNA) sample in the Iberian Peninsula. The inset shows the species range in Europe [26]. Dark grey dots, light grey dots and grey triangles represent individuals with microsatellite and mtDNA data, only microsatellite data and only mtDNA data, respectively. The Autonomous Communities mentioned in the paper are coded as follows: Galicia (GAL), Asturias (AST), Basque Country (BC), La Rioja (LR), Navarre (NAV) and Catalonia (CAT). The distributions maps were specifically generated based on the information provided in [41] and [26] using the country borders and the reference grids layers from the European Environmental Agency as background, available in http://www.eea.europa.eu/data-and-maps/ under a Creative Commons Attribution license, and are therefore for illustrative purposes only.

per extraction session. DNA was extracted with the Qiagen DNeasy Tissue kit following the manufacturer's protocol for tissue and hair samples.

Mitochondrial DNA sequencing and microsatellite genotyping

A fragment of 621 bp of mtDNA encompassing a portion of the cytochrome *b* gene, the threonine and proline tRNAs and the left domain of the control region was amplified with

the primers L15533 [42] and H16437 (5′- GGA GCG AGA AGA GGT ACA-3′). PCRs were carried out in 15 μL comprising 3 μL of DNA extract (20–100 ng), 1X PCR buffer (NZYTech), 2 mM MgCl2, 0.2 mM of each dNTP (Bioline), 1 uM of each primer, 0.17 ug/uL BSA (New England Biolabs) and 1.25U of Supreme NZYTaq DNA polymerase (NZYTech). The reactions were performed with an initial denaturation step at 94 °C for 5 min, followed by 55 cycles at 94 °C for 30 s, 60 °C for 30 s and 72 °C for 1 min. The final extension was at 72 °C for 7 min. PCR products were purified as in [42] and sequenced at Macrogen Inc. Electropherograms were visually inspected using SEQSCAPE 2.5 (Applied Biosystems) and nucleotide sequences were aligned using the default parameters in CLUSTALX 2.0 [43] and manually checked in BIOEDIT 5.0.9 [44].

Samples were individually genotyped with a novel multiplex panel of 23 autosomal microsatellite markers comprising 13 recently described species-specific microsatellites (Mf 1.1, Mf 1.11, Mf 1.18, Mf 1.3, Mf 2.13, Mf 3.2, Mf 3.7, Mf 4.10, Mf 4.17, Mf 6.5, Mf 8.7, Mf 8.8, Mf 8.10; [45] and 10 additional markers described in closely related mustelids: Ma1 and Ma2 [46]; Mel1 [47]; MLUT27 [48]; Mvis072 [49]; Mvi57 [50]; Lut615 [51]; MP0059 [52]; and Lut435 and Lut453 [53]. The forward primers labelled with the dyes 6-FAM, NED, PET and VIC, were used in four PCR multiplex reactions (multiplex A, B, C and D; S2 Table). PCR multiplex amplifications were carried out using the QIAGEN Multiplex PCR kit, following the manufacturer's protocol, at optimized annealing temperatures in a total volume of 9 μL with 1 μL of DNA and 2 pmol of each primer. We applied a hot-start thermocycling protocol with an initial polymerase activation at 95°C for 15 min, followed by 42 cycles of denaturation at 94°C for 30s, primer annealing at 57°C for 130s, and sequence extension at 72°C for 1 min, and a final extension step at 60°C for 30 min. In addition to the negative controls for extraction, negative PCR controls were included. Multiplex PCR products were run on an ABI 3130XL automated sequencer (Applied Biosystems) with the internal size standard GS500 LIZ™ (Applied Biosystems). Fragment analyses were performed using GENEMAPPER v. 4.0 (Applied Biosystems).

The software GIMLET v 1.3.4 [54] was used to calculate the probability of identity with both the unbiased equation for small sample size (PID) and the equation for siblings (PIDsibs), and to estimate genotyping errors [i.e. allelic dropout (ADO) and false alleles (FA)] from replicate genotyping. To check for genotyping errors in hair and tissue samples, we amplified 30 DNA extracts (hair samples: n=20; tissue samples: n=10) in four replicates. As errors were absent in both sample types, the remaining sample set was genotyped only once per locus.

To map the spatial distribution of the analysed samples in the study area, the UTM coordinates corresponding to each sample were projected onto a GIS (ArcGIS 10.0, ESRI) along with the haplotype and microsatellite data.

Mitochondrial DNA analysis

ARLEQUIN v. 3.5 [55] was used to estimate the amount of mtDNA variation through haplotype diversity (H) and nucleotide diversity (π) [56]. Haplotypic richness (Hr), standardized to the lowest sample size, was calculated using the rarefaction method implemented in CONTRIB v. 1.02 [57]. To infer haplotype relationships, a median-joining network was constructed as implemented in NETWORK v. 4.6 [58]. The program DNAsp v. 5 [59] was used to calculate distributions of pairwise nucleotide differences (mismatch distributions), which may be informative about past demographic history: unimodal curves are expected in populations that have undergone a rapid population expansion and multimodal curves are typical of populations with a history of long-term demographic stability [60,61]. We also applied two neutrality tests in ARLEQUIN (Fu's Fs, [62]; Tajima's D, [63]) to obtain further clues concerning demographic history; significance was tested using 1,000 simulations under a model of selective neutrality.

Historical population size changes were also assessed using a coalescent-based Bayesian Skyline Plot (BSP) generated in BEAST v. 1.7.5 [64]. Under a coalescent model it is possible to infer population parameters from genetic sequence data, such as estimates of mutation rate, divergence time, and effective population size through time. A relaxed lognormal clock was selected to construct the BSP. A substitution rate (μ) of 1.95×10^{-8} substitutions per nucleotide per year was estimated using equation $D=2\mu T$ [56], where D is genetic distance and T is the time since divergence. Specifically, as an estimate of D we computed in MEGA v.6 [65] the raw maximum composite likelihood distance [66] between the stone marten haplotypes obtained in this study and the published haplotypes of sable *Martes zibellina* (Genbank Accession Nos. KJ202610, KJ202613-5, KJ202623, KJ202625-7, KJ202633, KJ202636, KJ202640, KJ202644), and we assumed a time to the most recent common ancestor of 2.8 million years [67,68]. The best-fit model of nucleotide substitution for these sequences was HKY+I+G (α=0.33, I=0.780) according to the Akaike Information Criterion in jModelTest2 [69]. Chains were run for 75 million generations; trees were sampled every 1,000 generations and the first 10% of the samples were discarded as burn-in. Four replicate runs were conducted to confirm convergence.

Population structure: Bayesian clustering and multivariate analysis

Population structure was first estimated using the program STRUCTURE v. 2.3.4 [18,19,70] assuming population admixture and correlated allele frequencies within populations. Simulations were run with Markov Chain Monte Carlo (MCMC) of 10^6 iterations after a burn-in of 10^5. *K* was allowed to vary from 1 to 10 and 20 independent simulations were run for each *K* value to check for consistency in the results. To determine the most likely number of clusters we estimated the rate of change in the log probability of data between successive *K* values (ΔK) as described by [71]. For the identified *K* value, we used CLUMPP v. 1.1.2 [72] to determine the population assignment probability of each individual across all simulations. Subsequently, a progressive partitioning approach, forcing K=2 within each inferred cluster, was used to test for finer sub-structuring (e.g. [8]). GENELAND v 4.0.3 [14] was run through an extension of R v.3.0.1 [73] under the correlated allele frequency model without spatial uncertainty in spatial locations. We allowed K to vary between 1 and 10 in 20 independent runs, each with 10^5 iterations, a thinning of 100, a maximum number of nuclei of 1,000, and a maximum rate of Poisson process fixed at 333. The software TESS v.2.3.1 [21,22,74] is designed for seeking genetic discontinuities in continuous populations and estimating spatially varying individual admixture proportions, and can outperform other IBC methods at detecting migrants and recent contact zones between weakly differentiated populations [21]. Following the recommendations in [21], TESS runs were performed under the admixture model for 12,000 sweeps, with a burn-in period of 2,000 sweeps, interaction parameter set to 0.6, 10 independent runs per analysis and the maximum number of clusters fixed to Kmax = 10. Lastly, BAPS v. 6.0 [75] was run for both the microsatellite and mtDNA data sets using spatial clustering of individuals, since the spatial prior may strengthen inferences for sparse molecular data [23]. Then, the membership coefficients of each individual calculated by the different IBC programs were plotted on a map of the study area in ArcGIS 10.0 (ESRI), to assess the relationship between genetic discontinuities and landscape features.

Multivariate analyses have been suggested as an alternative to Bayesian clustering algorithms. Their main asset is that they can summarize the genetic variability without making strong assumptions about the underlying population genetic model, as they do not require populations to be in Hardy-Weinberg equilibrium (HWE) or linkage equilibrium (LE) between loci [24,25]. For instance, the discriminant analysis of principal components (DAPC), [25] identifies genetic clusters through sequential clustering and model selection. The method first transforms the genotype data into principal components and then uses k-means

clustering to define groups of individuals, with the best-supported number of clusters identified by the Bayesian Information Criteria (BIC). The spatial principal component analysis (sPCA, [24]) can explicitly identify cryptic spatial patterns of genetic structuring across the landscape, including clines, accounting for spatial autocorrelation issues associated with neighbour-mating and sample distribution [11].

We tested for the presence of IBD across the study area using both a Mantel test and a spatial autocorrelation analysis. The Mantel test was performed between a matrix of pairwise Edward's genetic distances [76] and a matrix of Euclidian geographic distances, with 9999 permutations to assess significance. The Mantel test, as well as DAPC and sPCA, were implemented in R software v.3.0.1 [73] with the adegenet package [24]. Spatial autocorrelation analysis, which allows examining the correlation of genetic and geographic distance at multiple distance classes, is typically more powerful than Mantel tests for uncovering genetic structure [77]. The spatial autocorrelation analysis was performed in GENALEX v 6.5 [78]. GENALEX generates an autocorrelation coefficient (*r)*, which provides a measurement of the pairwise genetic similarity of individuals whose geographic separation falls within a specified distance class. The significance of the resulting correlogram was determined with the heterogeneity test of [79].

Finally, we used the software ALLELES IN SPACE (AIS, [80]), specifically developed to deal with cases in which individuals are continuously distributed possibly over large spatial scales to obtain a Genetic Landscape Shape interpolation (GLSI) that allows visualizing patterns of genetic diversity across the landscape (a 50x50 grid surface with the distance weighting parameter set at 1). We also used AIS to conduct an allelic aggregation index analysis (AAIA, [80]), a test of non-random patters of spatial genetic diversity. Significance of R_{AVE} (average allelic aggregation index) was tested through the use of 1,000 permutations.

Genetic diversity: over the Iberian Peninsula and within each inferred cluster

Microsatellite genetic variation was characterized, for the whole study area and each inferred cluster, by the number of alleles per locus (N_A) and inbreeding coefficient (Fis) using GENETIX v 4.05 [81]. The software FSTAT v.2.9.3 [82] was used to calculate the observed and expected heterozygosities (H_O and H_E), the allelic richness (A_R), the deviations from HWE, the genotypic disequilibrium among pairs of loci, and pairwise Fst values [83]. Statistical significance was evaluated by running a MCMC consisting of 10,000 batches of 10,000 iterations each, with the first 10,000 iterations discarded before sampling [84]. Significance

levels were adjusted with sequential Bonferroni correction in order to correct for the effect of multiple tests [85]. In addition to Fst, we also estimated genetic differentiation using Jost's D_{EST} [86] and G''_{ST} [87] in GENALEX v 6.5 [78]. MICROCHECKER v.2.2 [88] was used to check for potential scoring errors and the presence of null alleles. Sibship analysis was conducted using COLONY v. 2.0.4 [89], with a typing error rate set at 0.01. This approach considers the likelihood of the entire pedigree, as opposed to relatedness of individuals on a pair-wise basis.

Results

Mitochondrial DNA

After alignment and trimming the ends, a total of 621 base pairs were obtained from the 252 sequenced individuals. Of the 621 aligned nucleotide sites, 12 were variable and 10 were parsimony informative, leading to a nucleotide diversity of 0.00169. We identified 12 haplotypes (H1-H12), 11 excluding sites with gaps, and haplotype diversity was therefore relatively low (0.705). Sequences of haplotypes H1 to H12 were deposited in Genbank (Accession Nos. KM972577-KM972588, S1 Table).

The median-joining network (Fig. 2) showed a star-like topology with most haplotypes differing by a single mutation from two common central haplotypes (H1 and H2), themselves also separated by a single substitution. Haplotype H12, found only in two samples from near the eastern Pyrenees, at eight mutational steps from the centre of the network was the sole exception to this pattern. The network did not suggest any marked phylogeographic structure, with the most common haplotypes (H1: 41.6%; H2: 31.3%) scattered throughout the Iberian Peninsula.

However, haplotype frequencies differed notably between regions, with all of them having private haplotypes with the exception of region SW (Fig. 2). In the NW, H1 was the most common haplotype (68.5%) followed by H2 (20.4%), H7 (5.6%), the private haplotype H11 (3.7%) and H4 (1.8%). In the NE (Fig. 2), H2 was the most frequent haplotype (38.9%), but some private haplotypes were also found in this region (H3, H8 and H10), including the most divergent haplotype H12, rendering the NE region the one with the highest haplotype diversity and haplotypic richness (S3 Table). The SW region, mostly corresponding to western Andalucía and south Portugal, harboured 5 haplotypes, the widespread H1 and H2 (55.9 and 36.8% respectively) plus H5, H7 and H9. Lastly, in the SE region, the analysed samples had haplotypes H1 (38.1%), H5 (33.3%), H9 (9.5%), H2 (4.8%) and the private haplotype H6 (14.3 %).

The tests of Tajima´s D (-1.10943; $p>0.1$) and Fu´s Fs (-3.344; $p>0.02$) did not reject the null hypothesis of demographic stability. However, the BSP indicated a period of constant population size from 6,500 to 1,000 YBP, after which the Iberian population experienced a slow demographic expansion.

BAPS analysis highlighted the mtDNA homogeneity of Iberian stone martens, as it only identified two clusters: one grouping all samples with haplotypes H1-H11 and another exclusively with the two individuals carrying the divergent haplotype H12 that were sampled in NE Iberia close to the French border (Fig. 2).

Microsatellites

We genotyped 333 stone martens from across the Iberian Peninsula (Spain: n=244; Portugal: n= 89) and we obtained full multilocus genotypes at the 23 loci for 257 of the samples (77.17%). All samples were genotyped for more than 20 loci and therefore included in the analyses.

The number of alleles ranged from four (Mf 1.11, Mf 1.18, Mf 2.13 and LUT453) to 13 (Mf 4.17) with a mean number of alleles per locus of 6.52 (S2 Table). The overall PID using all 23 loci was 3.563×10^{-17} and the overall PID-sib was 1.075×10^{-7}. Mean observed heterozygosity over all loci (Ho=0.49) was lower than expected heterozygosity (He=0.57). Twenty-four rare alleles (frequency <0.05) were detected while very rare alleles (frequency <0.01) were not found.

To avoid effects due to the underlying genetic structure, some analyses and tests were performed within the inferred clusters (see next section). Linkage disequilibrium was not apparent for any pair of loci within any of the clusters, before or after Bonferroni correction ($p<0.05$). Fis values were positive and significantly different from zero ($p<0.001$, S4 Table). As recommended by [90], the presence of null alleles was tested in three sets of 25 individuals from South Portugal, Basque Country and Catalonia, based on the results of the clustering analyses, and only three out of the 23 amplified loci showed signs of null alleles (Ma-2, Mf 2.13 and Mf 4.10). COLONY identified 16 full-siblings sampled from less than 1 km away (two road-killed individuals in Burgos, Spain) to almost 70 km apart (two samples collected in Évora district, Portugal), whereas no parent-offspring dyads were detected. Thus, eight individuals (one from each dyad) were excluded from the dataset because the removal of full-siblings is expected to improve estimates of population structure [91].

Population genetic structure

Among the Bayesian clustering approaches, the non-spatial algorithm implemented in STRUCTURE identified two clusters at the uppermost level (K=2, represented by dots and squares in Fig. 3a) following the criterion of Evanno et al [71]. Individuals located at the NE

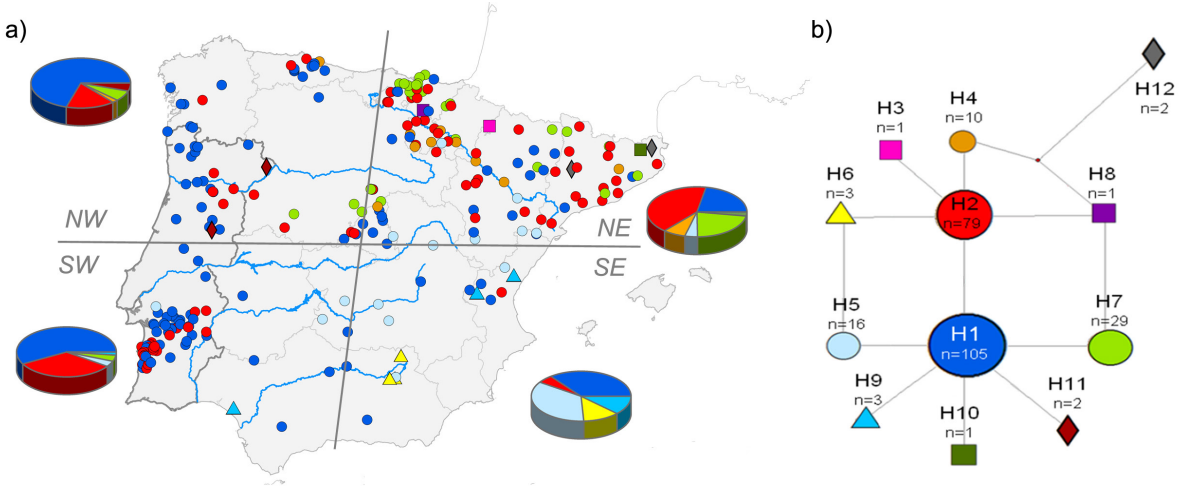


Figure 2. a) Geographic location of the 252 individuals sequenced for mtDNA in the Iberian Peninsula, which was divided into four regions (NE, NW, SE and SW). Pie charts represent the proportion of the samples with each haplotype in each region. b) Median-joining network of the 252 mtDNA sequences. Each haplotype (H1 to H12) is represented by a different coloured symbol (the same in the map). Grey numbers on the lines connecting haplotypes indicate the position of mutations in the alignment. "n" is the number of sequences in each haplotype. Less frequent haplotypes are represented by a square (n=1), diamond (n=2) and triangle (n=3) to facilitate their identification.

and SW extremes had the highest membership coefficients (>0.9) while the other samples elsewhere in the study area showed admixed ancestry (data not shown), a pattern consistent with IBD structure (see below and also Fig. 4a). Progressive partitioning suggested subdivision of the two clusters above, a clustering solution also hinted by the ΔK criterion. Of the four clusters, only one was relatively well delimited (in south Portugal, represented in yellow), but it was possible to roughly associate the others with NE Spain (Basque Country and Catalonia, in green), with an area from east to northwest Spain (in red), and with an area from north Portugal to south Spain (in purple) (Fig. 3a). BAPS and TESS also suggested K=4 but with different cluster compositions among each other and from progressive partitioning in STRUCTURE, especially for the two middle clusters between NE Spain and SW Portugal (Figs. 3b and 3c).

BAPS grouped north Portugal with NW Spain (Galicia and Asturias) (in purple) and inferred a cluster in east and central Spain from south of the River Ebro down to Andalucía (in red) (Fig. 3b), while TESS assigned individuals from Asturias to the latter group (in red; Fig. 3c). GENELAND inferred the clustering solution (K=3) with higher spatial consistency and

more clearly defined cluster boundaries (Fig. 3d). The three identified genetic units are distributed in south Portugal (in yellow, n=57), over a large part of Iberia from NW Spain and north Portugal to east and south Spain (in red, n=126), and in NE Spain (in green, n=150).

Thus, the different Bayesian clustering methods converged in identifying two distinct genetic units in NE Spain and south Portugal (Fig. 3). BAPS and TESS detected two additional clusters but they were incongruent between the two methods, while GENELAND was more conservative and grouped all the individuals elsewhere between NE Spain and south Portugal into a single cluster. The GENELAND solution was the most consistent and meaningful geographically, and so the GENELAND clusters were chosen for further analyses of genetic diversity (S5 Table) including IBD (see below).

The multivariate DAPC identified K=5 as the optimal number of clusters by the Bayesian Information Criterion but, with the exception of the cluster in south Portugal, they were spatially admixed (Fig. 3e). The first principal component differentiated cluster 2 (in yellow) from the others, and the second principal component displayed a slight differentiation between cluster 1 (in green, mainly found in NE Spain) and clusters 3, 4 and 5 (in red, orange and purple respectively, and scattered and overlapping over the majority of the study area). Except for cluster 2, the ellipses delineating the spatial extent of each cluster were substantially overlapping, suggesting weak genetic structuring overall.

Pairwise Fst values between clusters estimated by the different clustering methods ranged from 0.033 to 0.137 (S5 Table). Pairwise values of D_{EST} and G''_{ST}, which correct the dependency of Fst on the amount of within-population variation, were moderate to large and statistically significant (S6 Table). Notably, there was no agreement between the Fst values and the degree of geographic separation or how clearly defined were the boundaries between clusters.

An overall pattern of IBD was detected across the study area by both the Mantel test (0.296, $p < 0.001$; Fig. 4a) and the spatial autocorrelation analysis (heterogeneity test, $p<0.01$). The Mantel regression plot (Fig. 4a) displayed a single high-density nucleus indicating clinal variation, while the autocorrelation correlogram (Fig. 4b) showed a decline in the average relatedness as a function of linear geographic distance, with positive and significant values up to 450 km. The x-intercept of 452 km provided an estimate of the extent of non-random (positive) genetic structure, i.e., of the "genetic neighbourhood" of a population [92] (Fig. 4b). The sPCA ($r=0.633$ $p<0.001$; Fig. 4c) and the Genetic Landscape Shape Interpolation

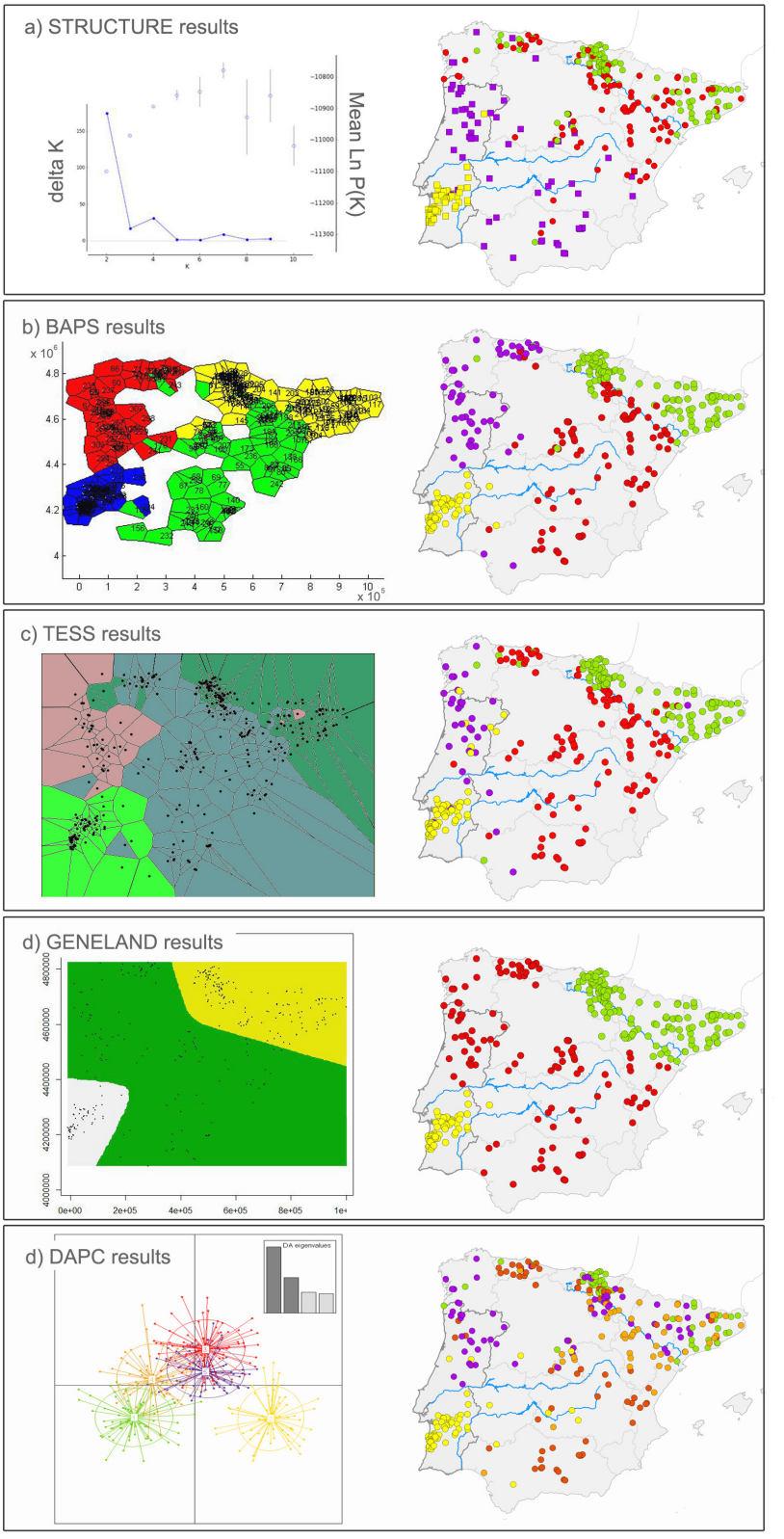


Figure 3. Map of the study area showing the results of the Bayesian clustering algorithms and of the DAPC. Individuals are represented by dots and coloured to reflect the cluster they were assigned to with the highest membership coefficient in each analysis. The Tagus, Ebro and Guadiana rivers are represented in blue. a) Plot from Structure Harvester for STRUCTURE results showing the modal value of K= 2 for the ΔK method (left axis; blue dots connected by line) and K= 4 for the maximum likelihood method (Mean L(K)± SD, right axis; open circles); in the latter case the chosen K is that after which L(K) plateaus or increases slightly and the variance between runs increases b) BAPS results (K= 4) c) TESS results (K= 4) d) GENELAND results (K = 3) e) DAPC identified K=5 as the optimal number of clusters, each indicated by different colours and inertia ellipses, the latter shown on the first two axes.

from the AIS software (Fig. 4d) revealed a cline in allele frequencies from NE to SW Iberia. The AAIA identified significant allelic aggregation (R_{AVE}=0.735, P=0.000), suggesting a non-random distribution of genotypes across the Iberian Peninsula. The strength of the spatial pattern along the NE-SW axis is further illustrated by the results of the clustering analyses (Fig. 3).

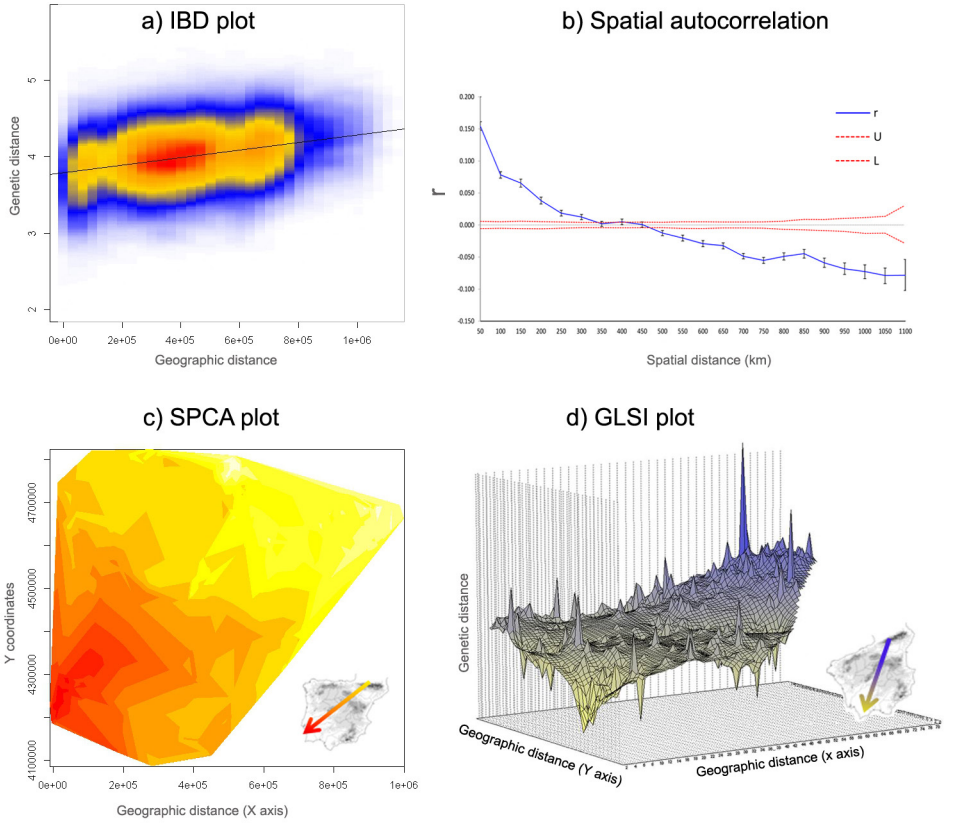


Figure 4. a) Scatterplot showing the results of the Mantel test between the matrix of genetic distances and the matrix of geographic distances to test for the presence of IBD. Colours represent the relative density of points, with warmer colours indicating higher densities, while the line shows the correlation between the two distance matrices. b) Correlogram of the average autocorrelation coefficient (r) as a function of distance classes of 50 km. Error bars bound the 95% confidence interval determined by bootstrap resampling (1,000 replicates) while confidence limits bound the 95% confidence interval about the null hypothesis of no spatial structure as determined by 999 permutations. c) Interpolation using a globally weighted regression of component 1 scores from the sPCA. Contours are component scores representing similarity across the landscape, with the arrow indicating the gradient found within the study area. d) Genetic Landscape Shape Interpolation (GLSI). Surface plot heights reflect genetic distance patterns over the geographical landscape examined and the arrow indicates the gradient found within the study area.

Discussion

Phylogeography of the stone marten in the Iberian Peninsula

The results of several analyses of the mtDNA data support a relatively recent expansion of the stone marten population in the Iberian Peninsula, including the star-like topology of the median-joining network, the unimodal shape of the mismatch distribution, and the BSP estimates. The summary statistics Tajima´s D and Fu´s Fs did not reject the null hypothesis of selective neutrality and demographic stability, but the observed negative values of the parameters indicate an excess of rare mutations that may be due to recent population growth. The median-joining network and the BAPS results for mtDNA, together with the fact that it was found only in two individuals sampled close to the border with France, suggest that the divergent haplotype H12 is the result of a recent dispersal from France, and unrelated to the evolutionary history of the stone marten in Iberia since their first arrival in the peninsula. Other than this episode, the observed mtDNA diversity is consistent with a single colonization event at the origin of the stone marten population in the Iberian Peninsula.

The BSP analysis suggested a period of constant population size from 6,500 to 1,000 YBP, after which the Iberian population experienced a demographic expansion. These estimates, given the absence of suitable calibration points, depend on the estimated substitution rate but agree remarkably well with the age of the oldest fossils of stone marten in the Iberian Peninsula, dating to 3,000-5,000 YBP in the north [32] and 7,000 YBP in the south [32,33]. In this context, it would be important to derive estimates of the time of the first migration of stone martens into Iberia based on their genetic divergence from other populations in Europe, particularly those closest to the Iberian Peninsula (currently underway). Also, because colonization from France across the Pyrenees is not the only possible scenario, since the stone marten could have been introduced from elsewhere in the Mediterranean range of the species [32–34].

In any case, the apparent recent colonization of the Iberian Peninsula could explain the lack of phylogeographic structure found in this study, in contrast to what was observed in the other marten species present in Iberia, the pine marten *Martes martes* [93]. Our results also support a single taxonomic unit in the Iberian Peninsula, in agreement with [41]. The purported differences in size and coloration leading to the recognition of two subspecies in Iberia, the nominal subspecies *Martes foina foina* (Erxleben, 1777) in the northern half of the peninsula and *Martes foina mediterranea* (Barrett-Hamilton, 1898) in the south, likely

fall within the range of continuous morphological variation displayed by the species in the region.

The phylogeographic signal for the stone marten expansion across Iberia, after its putative arrival in NE Spain from France, is still visible in the spatial patterns of genetic diversity. Haplotype richness was highest in NE Spain and there was a clear east-west differentiation in haplotype frequencies (Fig. 2). The Ebro River, the upper courses of the rivers Douro, Tagus and Guadiana, and the Meseta in central Iberia have all likely continuously contributed to these patterns. In line, we observed a marked NE-SW cline of microsatellite allele frequencies in the sPCA and GLSI analyses (Fig. 4).

Population genetic structure

The different Bayesian clustering methods did not converge to the same solution, but they all indicated two distinct genetic units in NE Spain and south Portugal (Fig. 3). Across methods, the cluster in south Portugal (Fig. 3, in yellow) was the most strongly differentiated one, as evidenced by the highest values of pairwise Fst (0.061-0.137, S5 Table), as well as of D_{EST} and G''_{ST} (0.154 and 0.244, respectively; S6 Table). The individuals in this cluster had the highest membership coefficients in all IBC algorithms used (data not shown). The cluster was also supported by the substantial degree of genetic distinctiveness of the southern Portuguese samples revealed by the interpolation plot (Fig. 4d). The cluster in NE Spain was less differentiated from the neighbouring cluster, as hinted by the smaller values of Fst (0.033), D_{EST} (0.049), and G''_{ST} (0.080) (S5 and S6 Tables). Accordingly, the program TESS, which is especially suited for the detection of recent migrants [21], identified putative migrants between these two clusters.

Comparing these findings with the pattern of mtDNA differentiation, it can be hypothesized that the increased microsatellite divergence of the clusters in NE Spain and south Portugal may be the result of the long-term barrier effect of the River Ebro and of the rivers Tagus and Guadiana, respectively, and the high mutation rate of microsatellites. The importance of major rivers as barriers to gene flow in mesocarnivores has been identified in previous studies on other species (e.g. badger, *Meles meles* [94]; wildcat, *Felis silvestris* [9]). Interestingly, [7] in their study of the genetic structure of the Eurasian otter *Lutra lutra*, also using spatial and non-spatial IBC methods, found a genetic break in the Iberian Peninsula roughly corresponding to the course of the Tagus river. Similarly, in our study, GENELAND suggested that the location of the boundary separating the cluster in NE Spain might not

strictly correspond to the River Ebro, at variance with the results from BAPS and TESS (Fig. 3). Although GENELAND is not an edge detection method itself, like for instance those based on Wombling [95] or on the Monmonier algorithm [80], it has been shown to consistently outperform both boundary detection methods and other Bayesian clustering methods in detecting barriers to gene flow [17,96]. The boundary suggested by GENELAND coincides with the convergence of the Ebro Depression and the Iberian System mountain range, which separates the Ebro basin from the watersheds of the major rivers in central Iberia. Hence, even the major rivers in the Iberian Peninsula may only represent moderate barriers to gene flow for the stone marten, but they certainly contribute to shape its genetic structure.

Other than the two consensus clusters in NE Spain and south Portugal, the methods disagreed on the grouping of the individuals elsewhere in between (Fig. 3). This could be the result of IBD, as Bayesian clustering methods are known to be sensitive to IBD and may infer spurious clusters when applied over a large area in which IBD is present [1,97] (see next section).

Of the clusters suggested by the Bayesian methods, the DAPC could only identify the cluster in south Portugal, apparently because it was relatively highly differentiated. This idea is supported by the fact that the clusters from DAPC had higher pairwise Fst values than those observed between the clusters revealed by the Bayesian methods (S4 Table). However, except the cluster in south Portugal, all the other four clusters inferred by DAPC did not have geographic identity.

Inferring genetic structure using clustering methods in the presence of IBD

In continuously distributed species, mating with proximal individuals leads to patterns of close relatedness at finer scales and extensive genetic gradients at larger scales [11]. Accordingly, Mantel tests of IBD were significant for the whole study area ($r = 0.296$, $P < 0.001$; Fig. 4a) and within each of the three clusters identified by GENELAND (S5 Table). The large genetic neighbourhood inferred by the spatial autocorrelation analysis (452 km) further underlined the extent and strength of the spatial structure. Therefore, we conclude that the difficulties and inconsistencies shown by the clustering methods in determining the genetic structure of the Iberian stone marten were due to the presence of strong and pervasive IBD.

This phenomenon has been described in several empirical (e.g. [1,11]) and simulation studies [17,96,98]. The clustering algorithms are known to vary in their ability to accurately

delimit genetic units in the presence of IBD and clines (e.g. [1,13,23], either underestimating [25] or overestimating the true number of populations [1,11]. Besides the widespread IBD in the Iberian stone marten, other characteristics of this study likely increased the challenge. First, the large size of the study area entailed the presence of population processes at different hierarchical levels, which may confound clustering methods [3]. Second, the study design has been shown to influence the performance of spatially-explicit IBC algorithms [11,99,100]. In this study however, even if the sampling was opportunistic, the resulting sampling scheme approximated reasonably well the heterogeneous distribution of the stone marten in Spain [41] (UTM10x10 grids; Fig. 1). Although widespread and locally abundant, the species is irregularly distributed through the Iberian Peninsula and rare in some areas with apparently suitable habitat [41,101], (Fig. 1).

Conclusions and perspectives

The phylogeography of the stone marten in the Iberian Peninsula is compatible with a single recent colonization of the region, possibly in the early Holocene, followed by population expansion sometime later. In turn, its contemporary genetic structure is characterized by a significant pattern of IBD and an apparent differentiation of the populations at the extreme of an NE-SW axis. To our knowledge, this is the first phylogeographic and population genetic study of the stone marten at broad regional scale. This case study also shows the usefulness of comparing and integrating multiple different techniques in the analysis of spatial genetic structure of mobile and continuously distributed species. Comprehensive analyses and tests of isolation by distance and clinal variation are an essential complement to clustering methods aiming to identify distinct genetic units, in particular due to the difficulties of the latter approaches in dealing with populations under IBD.

The genetic breaks detected in the Iberian stone marten population seemed largely related to the major river basins in the peninsula. However, linear barriers are only one of the multiple ways of how landscapes can influence gene flow [17]. Therefore, future studies are needed that focus on the importance of landscape features and variables in determining genetic structure (e.g. [102,103]). These should incorporate explicit tests on the presence of isolation by barriers and isolation by landscape resistance (e.g. land cover, elevation and temperature), especially in the areas where genetic discontinuities were inferred in this study.

Acknowledgments

The authors wish to thank the following institutions for supplying the tissue samples from Spain; CRF Mártioda and CRF Górliz via Departamento de Medio Ambiente y Política Territorial, Gobierno Vasco; CRFS "La Fombera", La Rioja; CRFS La Alfranca via Departamento de Agricultura, Ganadería y Medio Ambiente, Gobierno de Aragón; CR Torreferrusa and CR dels Aiguamolls de l'Empordà via Department d`Àgricultura, Ramaderia, Pesca, Alimentacióí Medi Natural, Generalitat de Catalunya; CRAS Madrid-Viñuelas via Consejería de Medio Ambiente y Ordenación del Territorio de la Comunidad de Madrid; CR El Chaparrilo, via Consejería de Agricultura, Castilla-La Mancha; CR Forn de Vidre via Conselleria d`Infrastructures, Territory i Medí Ambient de la Generalitat Valenciana; CREA Quiebrajano, and CREA Málaga via Consejería de Medio Ambiente, Junta de Andalucía;

National Museum of Natural Sciences, Natural History Museum of Barcelona, Mammal collection-Biological Station of Doñana, IREC and TRAGSATEC. For samples from Portugal, we thank BRISA Autoestradas de Portugal, ICNF, Estradas de Portugal S.A., CERVAS and CERAS. Finally, we wish to thank all the people directly involved in the collection of samples, including rangers, vets and field researchers.

References

1. Frantz AC, Cellina S, Krier A, Schley L, Burke T. Using spatial Bayesian methods to determine the genetic structure of a continuously distributed population: clusters or isolation by distance? J Appl Ecol. 2009 Apr;46(2):493–505.

2. Guillot G, Leblois R, Coulon A, Frantz AC. Statistical methods in spatial genetics. Mol Ecol. 2009 Dec;18(23):4734–56.

3. Latch EK, Reding DM, Heffelfinger JR, Alcalá-Galván CH, Rhodes OE. Range-wide analysis of genetic structure in a widespread, highly mobile species (*Odocoileus hemionus*) reveals the importance of historical biogeography. Mol Ecol. 2014 Jul;23(13):3171–90.

4. Sacks BN, Brown SK, Ernest HB. Population structure of California coyotes corresponds to habitat-specific breaks and illuminates species history. Mol Ecol. 2004 May;13(5):1265–75.

5. McRae BH, Beier P, Dewald LE, Huynh LY, Keim P. Habitat barriers limit gene flow and illuminate historical events in a wide-ranging carnivore, the American puma. Mol Ecol. 2005 Jun;14(7):1965–77.

6. Hapeman P, Latch EK, Fike J, Rhodes OE, Kilpatrick CW. Landscape genetics of fishers (*Martes pennanti*) in the Northeast: dispersal barriers and historical influences. J Hered. 2011;102(3):251–9.

7. Mucci N, Arrendal J, Ansorge H, Bailey M, Bodner M, Delibes M, et al. Genetic diversity and landscape genetic structure of otter (*Lutra lutra*) populations in Europe. Conserv Genet. 2010;11(2):583–99.

8. Hobbs GI, Chadwick EA, Bruford MW, Slater FM. Bayesian clustering techniques and progressive partitioning to identify population structuring within a recovering otter population in the UK. J Appl Ecol. 2011 Oct 6;48(5):1206–17.

9. Hartmann SA, Steyer K, Kraus RHS, Segelbacher G, Nowak C. Potential barriers to gene flow in the endangered European wildcat (*Felis silvestris*). Conserv Genet. 2013 Mar 5;14(2):413–26.

10. Aurelle D, Ledoux J-B. Interplay between isolation by distance and genetic clusters in the red coral *Corallium rubrum*: insights from simulated and empirical data. Conserv Genet. 2013 Feb 20;14(3):705–16.

11. Schwartz MK, McKelvey KS. Why sampling scheme matters: the effect of sampling scheme on landscape genetic results. Conserv Genet. 2009 Jun 17; 10(2):441–52.

12. Tucker JM, Schwartz MK, Truex RL, Wisely SM, Allendorf FW. Sampling affects the detection of genetic subdivision and conservation implications for fisher in the Sierra Nevada. Conserv Genet. 2014 Sep 13;15(1):123–36.

13. Guillot G, Santos F, Estoup A. Analysing georeferenced population genetics data with Geneland: a new algorithm to deal with null alleles and a friendly graphical user interface. Bioinformatics. 2008 Jun 1;24(11):1406–7.

14. Guillot G, Mortier F, Estoup A. Geneland: a computer package for landscape genetics. Mol Ecol Notes. 2005 Sep;5(3):712–5.

15. Rutledge LY, Garroway CJ, Loveless KM, Patterson BR. Genetic differentiation of eastern wolves in Algonquin Park despite bridging gene flow between coyotes and grey wolves. Heredity. Nature Publishing Group; 2010 Dec;105(6):520–31.

16. Tammeleht E, Remm J, Korsten M, Davison J, Tumanov I, Saveljev A, et al. Genetic structure in large, continuous mammal populations: the example of brown bears in northwestern Eurasia. Mol Ecol. 2010 Dec;19(24):5359–70.

17. Blair C, Weigel DE, Balazik M, Keeley ATH, Walker FM, Landguth E, et al. A simulation-based evaluation of methods for inferring linear barriers to gene flow. Mol Ecol Resour. 2012 Sep;12(5):822–33.

18. Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. Genetics. 2000 Jun;155(2):945–59.

19. Falush D, Stephens M, Pritchard JK. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics. 2003 Aug;164(4):1567–87.

20. Corander J, Waldmann P, Sillanpää MJ. Bayesian analysis of genetic differentiation between populations. Genetics. 2003 Jan;163(1):367–74.

21. Chen C, Durand E, Forbes F, François O. Bayesian clustering algorithms ascertaining spatial population structure: a new computer program and a comparison study. Mol Ecol Notes. 2007 Sep;7(5):747–56.

22. Durand E, Jay F, Gaggiotti OE, François O. Spatial inference of admixture proportions and secondary contact zones. Mol Biol Evol. 2009 Sep;26(9):1963–73.

23. Ball MC, Finnegan L, Manseau M, Wilson P. Integrating multiple analytical approaches to spatially delineate and characterize genetic population structure: an application to boreal caribou (*Rangifer tarandus caribou*) in central Canada. Conserv Genet. 2010 Jul 3;11(6):2131–43.

24. Jombart T. Adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics. 2008 Jun 1;24(11):1403–5.

25. Jombart T, Devillard S, Balloux F. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. BMC Genet. BioMed Central Ltd; 2010 Jan;11(1):94.

26. Tikhonov A, Cavallini P, Maran T, Krantz A, Herrero J, Giannatos G, et al. *Martes foina*. IUCN 2013. IUCN Red List of Threatened Species. Version 2013.2. <www.iucnredlist.org>. Downloaded on 05 September 2014. 2008.

27. Herr J. Ecology and Behaviour of Urban Stone Martens (*Martes foina*) in Luxembourg. M.Sc. Thesis, University of Sussex. England; 2008. p. 226.

28. Virgós E, García FJ. Patch occupancy by stone martens *Martes foina* in fragmented landscapes of central Spain: the role of fragment size, isolation and habitat structure. Acta Oecologica. 2002 Aug;23(4):231–7.

29. Santos MJ, Santos-Reis M. Stone marten (*Martes foina*) habitat in a Mediterranean ecosystem: effects of scale, sex, and interspecific interactions. Eur J Wildl Res. 2010 Sep 2;56(3):275–86.

30. Virgós E, Cabezas-Díaz S, Lozano J, Mangas J. Spatial distribution models in a frugivorous carnivore, the stone marten (*Martes foina*): is the fleshy-fruit availability a useful predictor? Anim Biol. 2010 Nov 1;60(4):423–36.

31. Delibes M. Interspecific competition and the habitat of the stone marten *Martes foina* (Erxleben 1777) in Europe. Acta Zool Fenn. 1983;174:229–31.

32. Sommer R, Benecke N. Late- and Post-Glacial history of the Mustelidae in Europe. Mamm Rev. 2004 Dec;34(4):249–84.

33. Llorente L, Cristina R, Careta M. Earliest occurrence of the Beech marten (*Martes foina* Erxleben, 1777) in the Iberian Peninsula. In: Brugal J., Gardeisen A, Zucker A, editors. Prédateurs dans tour leurs états Évoultion, Biodiversité, Interactions, mythes, symboles XXXI rencontres internationales d´archéologie et d´histoire d´Antibes. Éditions A. Antibes; 2011. p. 189–209.

34. Anderson E. Quaternary evolution of the genus *Martes* (Carnivora, Mustelidae). Acta Zool Fenn. 1970;130:1–132.

35. Schwartz M, Ruiz-González A, Pertoldi C, Masuda R. Martes conservation genetics: assessing within-species movements, units to conserve and connectivity cross ecological and evolutionary time. In: Aubry K., Zielinski W., Raphael M., Proulx G, Buskirk S., editors. Biology and conservation of marten, sables, and fisher: a new synthesis. New York: Cornell Univesity Press; 2012.

36. Kyle CJ, Davison A, Strobeck C. Genetic structure of European pine martens (*Martes martes*), and evidence for introgression with M. americana in England. Conserv Genet. Kluwer Academic Publishers; 2003 Mar 1;4(2):179–88.

37. Koen EL, Bowman J, Garroway CJ, Mills SC, Wilson PJ. Landscape resistance and American marten gene flow. Landsc Ecol. 2012 Nov 5;27(1):29–43.

38. Nagai T, Raichev E, Tsunoda H. Preliminary study on microsatellite and mitochondrial DNA variation of the stone marten *Martes foina* in Bulgaria. Mammal Study. 2012;358:353–8.

39. Basto MP. Population and landscape genetics of the stone marten and red fox in Portugal: implications for conservation management of common carnivores. M.Sc. Thesis, University of Lisbon, Portugal. 2014.

40. Virgós E, Zalewski A, Rosalino L., Mergey M. Habitat ecology of genus *Martes* in Europe: a review of the evidences. In: Aubry KB, Zielinski WJ, Proulx G, Buskirk SW, editors. Biology and conservation of Martens, sables, and fishers A new synthesis. New York: Cornell University Press; 2012. p. 255–66.

41. Reig S. *Martes foina* Erxleben, 1777. In: Palomo L., Gisbert J, Blanco J., editors. Atlas y Libro Rojo de los Mamíferos Terrestres de España. Madrid: Dirección General de Biodiversidad-SECEM-SECEMU; 2007. p. 305–7.

42. Fernandes CA, Ginja C, Pereira I, Tenreiro R, Bruford MW, Santos-Reis M. Species-specific mitochondrial DNA markers for identification of non-invasive samples from sympatric carnivores in the Iberian Peninsula. Conserv Genet. 2008 Aug 7;9(3):681–90.

43. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, et al. Clustal W and Clustal X version 2.0. Bioinformatics. 2007 Nov 1;23(21):2947–8.

44. Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/ NT. Nucleic Acids Symp Ser. 1999;(41):95–8.

45. Basto MP, Rodrigues M, Santos-Reis M, Bruford MW, Fernandes CA. Isolation and characterization of 13 tetranucleotide microsatellite loci in the Stone marten (*Martes foina*). Conserv Genet Resour. 2010 Mar 17;2(S1):317–9.

46. Davis C, Strobeck C. Isolation, variability, and cross-species amplification of polymorphic microsatellite loci in the family Mustelidae. Mol Ecol Notes. 1998;7:1776–8.

47. Bijlsma R, Van de Vliet M, Pertoldi C, Van Apeldoorn R, Van de Zande L. Microsatellite primers from the Eurasian badger, *Meles meles*. Mol Ecol. 2000 Dec;9(12):2116–7.

48. Cabria MT, González EG, Gómez-Moliner BJ, Zardoya R. Microsatellite markers for the endangered European mink (*Mustela lutreola*) and closely related mustelids. Mol Ecol Notes. 2007 Nov;7(6):1185–8.

49. Fleming M, Ostrander E, Cook J. Microsatellite markers for American mink (*Mustela vison*) and ermine (*Mustela erminea*). Mol Ecol. 1999;8:1352–3.

50. O'Connell M, Wright JM, Farid A. Development of PCR primers for nine polymorphic American mink *Mustela vison* microsatellite loci. Mol Ecol. 1996 Apr;5(2):311–2.

51. Dallas JF, Piertney SB. Microsatellite primers for the Eurasian otter. Mol Ecol. 1998;7(9):1248–51.

52. Jordan MJ, Higley JM, Matthews SM, Rhodes OE, Schwartz MK, Barrett RH, et al. Development of 22 new microsatellite loci for fishers (*Martes pennanti*) with variability results from across their range. Mol Ecol Notes. 2007 Sep;7(5):797–801.

53. Dallas JF, Coxon KE, Sykes T, Chanin PRF, Marshall F, Carss DN, et al. Similar estimates of population genetic composition and sex ratio derived from carcasses and faeces of Eurasian otter *Lutra lutra*. Mol Ecol. 2003 Jan;12(1):275–82.

54. Valière N. GIMLET: a computer program for analysing genetic individual identification data. Mol Ecol Notes. 2002;(2):377–9.

55. Excoffier L, Lischer HEL. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Resour. 2010 May;10(3):564–7.

56. Nei M. Molecular Evolutionary Genetics. Columbia University Press; 1987.

57. Petit RJ, El Mousadik A, Pons O. Identifying Populations for Conservation on the Basis of Genetic Markers. Conserv Biol. 1998 Jul 7;12(4):844–55.

58. Bandelt H-J, Forster P, Rohl A. Median-Joining Networks for Inferring Intraspecific Phylogenies. Mol Biol Evol. 1994;16(1):37–48.

59. Librado P, Rozas J. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics. 2009 Jun 1;25(11):1451–2.

60. Rogers A, Harpending H. Population growth makes waves in the distribution of pairwise genetic differences. Mol Biol Evol. 1992 May 1;9(3):552–69.

61. Harpending HC. Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. Hum Biol. 1994 Aug;66(4):591–600.

62. Fu YX. Statistical Tests of Neutrality of Mutations Against Population Growth, Hitchhiking and Background Selection. Genetics. 1997 Oct 1;147(2):915–25.

63. Tajima F. Statistical Method for Testing the Neutral Mutation Hypothesis by DNA Polymorphism. Genetics. 1989 Nov 1;123(3):585–95.

64. Drummond AJ, Rambaut A. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol Biol. 2007 Jan;7(1):214.

65. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol Biol Evol. 2013 Dec;30(12):2725–9.

66. Tamura K, Nei M, Kumar S. Prospects for inferring very large phylogenies by using the neighbor-joining method. Proc Natl Acad Sci U S A. 2004;101:11030–5.

67. Koepfli K-P, Deere KA, Slater GJ, Begg C, Begg K, Grassman L, et al. Multigene phylogeny of the Mustelidae: resolving relationships, tempo and biogeographic history of a mammalian adaptive radiation. BMC Biol. 2008 Jan;6(10):22.

68. Li B, Wolsan M, Wu D, Zhang W, Xu Y, Zeng Z. Molecular Phylogenetics and Evolution Mitochondrial genomes reveal the pattern and timing of marten (*Martes*), wolverine (*Gulo*), and fisher (*Pekania*) diversification. 2014;80:156–64.

69. Darriba D, Taboada GD, Doallo R, Posada D. jModelTest 2: more models, new heuristics and parallel computing. Nat Methods. 2012;9(8):772.

70. Falush D, Stephens M, Pritchard JK. Inference of population structure using multilocus genotype data: dominant markers and null alleles. Mol Ecol Notes. 2007 Jul 1;7(4):574–8.

71. Evanno G, Regnaut S, Goudet J. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol. 2005 Jul;14(8):2611–20.

72. Jakobsson M, Rosenberg NA. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics. 2007 Jul 15;23(14):1801–6.

73. R Development Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria; 2014.

74. François O, Ancelet S, Guillot G. Bayesian clustering using hidden Markov random fields in spatial population genetics. Genetics. 2006 Oct;174(2):805–16.

75. Corander J, Marttinen P. Bayesian identification of admixture events using multilocus molecular markers. Mol Ecol. 2006 Sep;15(10):2833–43.

76. Edwards AW. Distance between populations on the basis of gene frequencies. Biometrics. 1971;27:873–81.

77. Peakall R, Ruibal M, Lindenmayer DB. Spatial autocorrelation analysis offers new insights into gene flow in the australian bush rat, *Rattus fuscipes*. Evolution (NY). 2003;57(5):1182–95.

78. Peakall R, Smouse PE. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. Bioinformatics. 2012 Oct 1;28(19):2537–9.

79. Smouse PE, Peakall R, Gonzales E. A heterogeneity test for fine-scale genetic structure. Mol Ecol. 2008;17:3389–400.

80. Miller MP. Alleles in space (AIS): computer software for the joint analysis of interindividual spatial and genetic information. J Hered. 2005;96(6):722–4.

81. Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F. Genetix v. 4.02, logiciel sous Windows pour la génétique des populations. Laboratoire Génome, Populations, Interactions, Université de Montpellier II, Montpellier, France. 2000.

82. Goudet J. FSTAT, a program to estimate and test gene diversitiesand fixation indices version 2.9.3. Updated from Goudet (1995). http://www2.unil.ch/popgen/softwares/fstat.htm

83. Weir B, Cockerham C. Estimating F-statistics for the analysis of population structure. Evolution 1984;38:1358–70.

84. Guo S, Thompson E. Performing the exact test of Hardy– Weinberg proportion for multiple alleles. Biometrics. 1992;48:361–72.

85. Rice W. Analyzing tables of statistical tests. Evolution 1989;43:223–5.

86. Jost L. GST and its relatives do not measure differentiation. Mol Ecol. 2008 Sep;17(18):4015–26.

87. Meirmans PG, Hedrick PW. Assessing population structure: FST and related measures. Mol Ecol Resour. 2011 Jan;11(1):5–18.

88. Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P. Micro-Checker: Software for Identifying and Correcting Genotyping Errors in Microsatellite Data. Mol Ecol Notes. 2004 Sep;4(3):535–8.

89. Jones OR, Wang J. COLONY: a program for parentage and sibship inference from multilocus genotype data. Mol Ecol Resour. 2010 May;10(3):551–5.

90. Hale ML, Burg TM, Steeves TE. Sampling for microsatellite-based population genetic studies: 25 to 30 individuals per population is enough to accurately estimate allele frequencies. PLoS One. 2012 Jan;7(9):1–10.

91. Goldberg CS, Waits LP. Quantification and reduction of bias from sampling larvae to infer population and landscape genetic structure. Mol Ecol Resour. 2010 Mar;10(2):304–13.

92. Wright S. Isolation by distance under diverse systems of mating. Genetics. 1946;31:39–59.

93. Ruiz-González A, Madeira MJ, Randi E, Abramov A V., Davoli F, Gómez-Moliner BJ. Phylogeography of the forest-dwelling European pine marten (*Martes martes*): new insights into cryptic northern glacial refugia. Biol J Linn Soc. 2013 May 26;109(1):1–18.

94. Frantz AC, Pope LC, Etherington TR, Wilson GJ, Burke T. Using isolation-by-distance-based approaches to assess the barrier effect of linear landscape elements on badger (*Meles meles*) dispersal. Mol Ecol. 2010 Apr;19(8):1663–74.

95. Cercueil A, François O, Manel S. The genetical bandwidth mapping: a spatial and graphical representation of population genetic structure based on the Wombling method. Theor Popul Biol. 2007 May;71(3):332–41.

96. Safner T, Miller MP, McRae BH, Fortin M-J, Manel S. Comparison of bayesian clustering and edge detection methods for inferring boundaries in landscape genetics. Int J Mol Sci. 2011 Jan;12(2):865–89.

97. Meirmans PG. The trouble with isolation by distance. Mol Ecol. 2012 Jun;21(12):2839–46.

98. Landguth EL, Cushman SA, Schwartz MK, McKelvey K, Murphy M, Luikart G. Quantifying the lag time to detect barriers in landscape genetics. Mol Ecol. 2010 Sep 6;4179–91.

99. Landguth EL, Cushman SA, Johnson NA. Simulating natural selection in landscape genetics. Mol Ecol Resour. 2012 Mar;12(2):363–8.

100. Landguth EL, Schwartz MK. Evaluating sample allocation and effort in detecting population differentiation for discrete and continuously distributed individuals. Conserv Genet. 2014 Mar 26;15(4):981–92.

101. Reig S. Martes foina Erxleben, 1777. In: Palomo L, Gisbert J, editors. Atlas de los Mamiferos Terrestres de España. Madrid: Dirección General de Conservación de la Naturaleza-SECEM-SECEMU; 2002. p. 270–3.

102. Cushman SA, Wasserman T, Landguth E, Shirk A. Re-Evaluating Causal Modeling with Mantel Tests in Landscape Genetics. Diversity. 2013 Feb 18;5(1):51–72.

103. Balestrieri A, Ruiz-González A, Remonti L, Gómez-Moliner BJ, Genovese S, Gola LA, et al. A non-invasice genetic survey of the pine marten (*Martes martes*) in the western river Po plain (Italy): Preliminary results. Hystrix It J Mamm. 2008;19(1):77–80.

Supporting Information

Table S1. Information on the 352 stone marten specimens analysed in this study. Given are, for each individual, the sample codes, localities, coordinates, sample donor, mtDNA haplotype, and genotypes for the 23 microsatellite loci.

The table is too big to print in less than 15 pages. Thus, before the papers are published the table will be available online (dropbox folder) in the following link:

https://www.dropbox.com/sh/y4lw39hvnfrepow/AAD72y908pLs-3Zp2G4YUrCGa?dl=0

Multiplex	Locus	Repeat	DYE	Size range	Na	Ho	He
Multiplex A	Mlut27	dinucleotide	PET	178-202	8	0.341	0.375
	Mel1	dinucleotide	PET	264-276	6	0.468	0.577
	Mf1.1	tetranucleotide	VIC	152-168	5	0.399	0.446
	Mf4.17	tetranucleotide	6-FAM	194-242	13	0.773	0.854
	Mf8.10	tetranucleotide	NED	128-152	7	0.648	0.738
Multiplex B	Lut453	dinucleotide	6-FAM	104-110	4	0.398	0.415
	Ma1	dinucleotide	PET	205-215	6	0.644	0.689
	Mf1.18	tetranucleotide	PET	153-165	4	0.374	0.391
	Mf3.7	tetranucleotide	VIC	181-201	6	0.398	0.456
	Mf8.8	tetranucleotide	6-FAM	223-255	9	0.681	0.799
	Mp0059	dinucleotide	VIC	142-152	6	0.497	0.557
Multiplex C	Lut615	dinucleotide	6-FAM	115-127	6	0.301	0.47
	Ma2	dinucleotide	PET	174-186	6	0.437	0.491
	Mf1.11	tetranucleotide	NED	205-217	4	0.345	0.391
	Mf2.13	tetranucleotide	6-FAM	285-297	4	0.388	0.588
	Mf3.2	tetranucleotide	PET	146-170	7	0.648	0.677
	Mf6.5	tetranucleotide	VIC	215-247	9	0.642	0.695
	Mvi072	dinucleotide	NED	262-276	8	0.351	0.436
Multiplex D	Lut435	dinucleotide	NED	128-144	6	0.344	0.416
	Mf1.3	tetranucleotide	NED	196-220	7	0.591	0.645
	Mf4.10	tetranucleotide	VIC	311-339	8	0.649	0.769
	Mf8.7	tetranucleotide	PET	151-171	6	0.559	0.632
	Mvi-57	dinucleotide	6-FAM	101-109	5	0.469	0.543
Mean					**6.52**	**0.493**	**0.567**

S2 Table. Properties of the 23 multiplexed microsatellite loci used in this study, including repeat type, dye type, allele size range, number of alleles (Na) and observed (Ho) and expected (He) heterozygosities for each locus.

Region	n	Nh	Ph	π	H	Hr
NW	54	5	1	0.00080	0.4934	2.374
NE	108	9	4	0.00131	0.7542	3.990
SW	69	5	0	0.00099	0.5550	1.997
SE	21	5	1	0.00166	0.7480	3.628

S3 Table. mtDNA summary statistics for the NW, NE, SW and SE regions of Iberia. Number of individuals (n) and haplotypes (Nh), private haplotypes (Ph), nucleotide diversity (π), haplotype diversity (H), and haplotypic richness (Hr) calculated using the rarefaction method to correct for unequal sample sizes (n=21-108; rarefaction to 15).

DAPC	DAPC_red	DAPC_purple	DAPC_orange	DAPC_yellow
DAPC_green (n:62)	0.0887	0.0711	0.0624	0.1372
DAPC_red (n:78)	-	0.0541	0.0611	0.0897
DAPC_purple (n:72)	-	-	0.0604	0.078
DAPC_orange (n:53)	-	-	-	0.1191
DAPC_yellow (n:68)	-	-	-	-

STRUCTURE	STR_red	STR_purple	STR_yellow
STR_green (n:106)	0.0355	0.0625	0.1129
STR_red (n:104)	-	0.0348	0.1043
STR_purple (n:68)	-	-	0.0621
STR_yellow (n:55)	-	-	-

BAPS	BAPS_red	BAPS_purple	BAPS_yellow
BAPS_green (n:121)	0.0555	0.041	0.1045
BAPS_red (n:58)	-	0.0437	0.0846
BAPS_purple (n:96)	-	-	0.0957
BAPS_yellow (n:58)	-	-	-

TESS	TESS_red	TESS_purple	TESS_yellow
TESS_green (n:111)	0.0408	0.0657	0.0977
TESS_red (n:127)	-	0.0545	0.0783
TESS_purple (n:35)	-	-	0.061
TESS_yellow (n:60)	-	-	-

GENELAND	GL_red	GL_yellow
GL_green (n:150)	0.033	0.1024
GL_red (n:126)	-	0.0753
GL_yellow (n:57)	-	-

S4 Table. Pairwise Fst values between the inferred clusters by each of the clustering methods. All values were significantly different from zero at $p<0.001$.

Cluster	n	A	Pa	Ho	He	Fis	HWE	Mantel's r
GL_red	126	5.56	14	0.49	0.57	0.152	<0.001	0.227
GL_green	150	5.78	19	0.53	0.6	0.129	<0.001	0.145
GL_yellow	57	3.82	0	0.47	0.52	0.117	<0.001	0.272
total	333	6.6	33	0.49	0.6	0.167	<0.001	0.295

S5 Table. Genetic diversity indices for the three clusters identified by GENELAND. Number of individuals (n), mean number of alleles (A), private alleles (Pa) observed (Ho) and expected (He) heterozygosities, overall Fis, Hardy-Weinberg equilibrium p-value, and Mantel's correlation (r) of IBD tests, p<0.001).

G''_{ST} \ D_{EST}	GL_green	GL_red	GL_yellow
GL_green (n:150)	-	0.049	0.100
GL_red (n:126)	0.080	-	0.154
GL_yellow (n:57)	0.169	0.244	-

S6 Table. Matrix of pairwise values of Jost's D_{EST} (above diagonal) and Hedrick´s unbiased G''_{ST} (below diagonal) between the clusters inferred by GENELAND. All values were significantly different from zero at p<0.001 (9999 permutations).

Chapter III

PAPER 4: Phylogeography and population genetic structure of the stone marten (*Martes foina*) in Europe.

María Vergara*, Carlos Fernandes*, Mafalda Basto, María José Madeira, Aritz Ruiz-González, Benjamín Gómez-Moliner, Margarida Santos-Reis

*Both authors contributed equally to this work

Manuscript

Abstract

In this study, we examined the phylogeography and genetic structure of *M. foina* through mitochondrial (620 pb) and microsatellite (23 autosomal markers) analyses, based on a large and geographically comprehensive sampling across Europe (n=751). The results revealed high levels of haplotype diversity and a major phylogeographic east-west split within the Western Palearctic, consistently found in the mtDNA analyses, which was dated to the Middle Pleistocene (0.55 Mya). The western clade comprised mostly individuals sampled in Western Europe, from the Iberian Peninsula to Denmark, while three geographic lineages were identified in the eastern clade; the 'Balkan', the 'Italian-North European' and the 'Central-North European'. The genetic and geographic differentiation observed between these lineages is congruent with a range fragmentation, isolation, and posterior divergence within different refugia during Late Quaternary glaciations. The basal position of the Balkanic samples among the extant mtDNA diversity, in conjunction with the Middle Pleistocene age inferred for the European stone martens, supports an early entrance into the Balkans from the Middle East, and subsequent expansion westwards reaching Western Europe. Four refugia were postulated; three traditional Mediterranean refuges located in the Iberian, Italian and Balkan peninsulas and additional cryptic northern refugia in the Pannonian Basin. During the recurring interglacial periods of the Quaternary, the secondary contacts of haplotypes from distinct origins would have been facilitated by the high dispersal ability of the stone marten or human mediated introductions/translocations. The use of both mtDNA and microsatellites markers, together with new fossil evidences, helped to elucidate the complex phylogeography and history of population structure of *M. foina* in Europe. Besides, the early divergence estimate categorically contradicts the prevailing hypothesis based on previous fossil evidences, according to which the stone marten first entrance and posterior dispersal followed the Neolithic farming cultures from the Near East to Southwestern Europe during the Holocene.

Keywords: *Martes foina*; stone marten; European phylogeography; Quaternary; molecular versus fossil data; large-scale genetic structure; mitochondrial DNA; microsatellites.

Introduction

Phylogeography is a branch of biogeography concerned with the study of the geographic distributions of genealogical lineages within and among closely related species, and a major focus is on understanding how intraspecific genealogies have been shaped by geography, history and demography (Avise et al 1987; Avise 2000). In the last two decades, phylogeography has experienced explosive growth fuelled by developments in molecular markers, theoretical concepts and analytical methods, and provided valuable contributions to several areas of biological research, including speciation, taxonomy, conservation genetics, and the inference of general historical processes underlying the origin and distribution of genetic diversity across species (Taberlet et al 1998; Beheregaray 2008).

In Europe, it is generally accepted that phylogeographic patterns and structure in terrestrial animals and plants have been heavily influenced by the recurrent cycles of range contraction/expansion driven by environmental change associated with the Quaternary climatic oscillations (Hewitt 1996; Hewitt 2000; Hewitt 2004). Temperate species are thought to have survived glacial periods in southern refugia, namely Iberia, Italy and the Balkans, and recolonized central and northern Europe following ice retreat during interglacials (Hewitt 1996; Taberlet et al 1998; Hewitt 1999). Such recolonizations often resulted in suture zones, zones of secondary contact between lineages from different refugia, along the Alps and Pyrenees, along different meridians in western and central Europe, in the Alpine-Carparthian belt, and in Scandinavia (Jaarola and Tegelström 1995; Taberlet et al 1998; Hewitt 2000; Hewitt 2004). In addition to the main refugia in the southern peninsulas, areas located further north have been postulated as favourable cryptic refugia during the last glacial period (Stewart and Lister 2001; Provan and Bennett 2008), particularly for cold-tolerant species (Bhagwat and Willis 2008), and the importance of these northern refugia has been supported by accumulating phylogeographic data (Valdiosera et al 2007; McDevitt et al 2012; Schmitt and Varga 2012; Ruiz-González et al 2013).

Among mammals, carnivores have been recently the subjects of several pan-European phylogeographic surveys that provided unique insights into the evolutionary history of the studied species and contributed to our understanding of the historical processes that shaped the distribution and genetic patterns of the biota in Europe. These studies found genetic signals of population isolation in multiple refugia during the Pleistocene (Edwards et al 2012; McDevitt et al 2012; Ruiz-González et al 2013), but also instances where mitochondrial haplotypes were shared between distant regions (Mucci et al 2010; Pilot et al 2010; Frantz et al

2014). Limited phylogeographic structure in carnivores has been attributed to their high dispersal capabilities that allow relatively rapid gene flow across large distances during postglacial recolonization, erasing the signs of Pleistocene separations (Vilà et al 1999). Recurrent population connectivity in the interstadials of the last glaciation may have also contributed to the weak phylogeographic differentiation found in some species (Hofreiter et al 2004).

The stone or beech marten (*Martes foina* Erxleben, 1777) is a medium-sized mustelid carnivore with a wide Palearctic distribution, from the Iberian Peninsula, through central and southern Europe, the Middle East, central Asia, extending as far east as Mongolia, eastern China and northern Myanmar (Tikhonov et al 2008). In the Eastern Mediterranean, the species also found its way to Crete, Rhodes, and many of the Aegean and Ionian islands (Masseti 1995). The stone marten is generally common in Europe but population status and trends in other parts of its range, where hunting law enforcement is weaker and census estimates of the species are lacking, are in need of research (Tikhonov et al 2008). The habitat preferences of *M. foina* vary in different parts of its range, and the species can be found in mountain forests, rocky hillsides, mixed forests, woodlands and mosaic habitats (Delibes 1983; Sacchi and Meriggi 1995; Al Shafee et al 1997; Santos and Santos-Reis 2010). Stone martens can use more open habitats than other martens and its more generalist and flexible habits may allow avoidance of competition with the pine marten (*Martes martes* Linnaeus, 1758), as the two species are sympatric over most of their ranges (Delibes 1983). In fact, *M. foina* is often synanthropic where it coexists with *M. martes* (e.g. Herr et al 2009).

The subfossil record of *M. foina* from the Late Glacial and Holocene in Europe is relatively poor (Sommer and Benecke 2004). The oldest Holocene finds are from 9,000-7,500 years BP in central Europe, France and Italy, with further records in Poland, the Alps, Spain and Greece from the second half of the Hypsithermal, 7,500-5,000 years BP (Sommer and Benecke 2004; Llorente et al 2011). A record from the Upper Palaeolithic site at Duritor in Moldova (Dallas et al 1999) has been postulated by Sommer & Benecke (2004) to be due to a possible mixing of layers, and these authors concur with Anderson (1970) who first suggested that the stone marten only became a member of the European fauna following the expansion of the Neolithic cultures of the Near East to southern Europe.

More recently, however, *M. foina* has been confirmed as present in a layer from the Palaeolithic cave site of Veternica, Croatia, assigned to the end of the Pleniglacial (Miracle et al 2010). This raises the importance of determining the age of other *M. foina* remains ascribed to the Late Pleistocene, like the one from Duritor and those from the Caverna degli Orsi in

the Trieste Karst (NE Italy; Berto and Rubinato 2011) and the Equi Cave in the Apuan Alps (NW Tuscany, Italy; Ghezzo et al 2014), and to the Middle Pleistocene, such as those from Es-Taliens in the French Pyrenees (Bonifay 1973) and from the Apidima Caves in the Greek Peloponnese (Tsoukala 1999), but affected by stratigraphic uncertainties. What is needed is direct radiocarbon dating of these putative Pleistocene fossils (Sommer and Benecke 2004) and their reanalysis, possibly including genetic comparisons with modern stone martens (Consuegra et al 2002), to ascertain whether they belong to *M. foina* (e.g. Berto and Rubinato 2011).

Sommer & Benecke (2004) also noted that stone martens might have been absent or rare across much of the North European Plain until the Middle Ages, when more intense occupation and use of the landscape by humans could have facilitated *M. foina* colonization. Finally, the same authors suggest that the stone marten was introduced into Crete during the Minoan-Mycenaean period in the Late Bronze Age, although some limited evidence supports an earlier introduction (Masseti 1995).

Given the sparse fossil record, the biogeographic and evolutionary history of *M. foina* in Europe is therefore largely unknown. A morphometric analysis of cranial variation found an east-west gradient of decreasing skull size, which was seen as consistent with the hypothesis of European colonization from the Middle East (Reig 1992). Another notable finding of the same study was that samples from Iberia grouped with those from France and northern Italy, while samples from the Alps grouped with those from northern Central Europe. Previous genetic studies focusing on the phylogeography or population structure of *M. foina* are few and at the national or regional level (Bulgaria, Nagai et al 2012); Portugal, Basto et al, in review; Iberia, Vergara et al, in review).

In this study, we examined the phylogeography and genetic structure of *M. foina* in Europe based on mitochondrial and microsatellite DNA analyses of a comprehensive sampling across the continent. The specific aims were to: (i) identify the main phylogeographic patterns, and thus shed light on the biogeographic history of the stone marten in Europe; (ii) estimate the timing of the arrival and subsequent lineage diversification of *M. foina* in Europe, to assess the prevailing current view that the species entered Europe only with the Neolithic expansion from the Near East; (iii) provide the first overall picture of the current population structure of stone martens in Europe.

Material and methods

Sampling and DNA extraction

We collected tissue and hair samples of 751 stone martens from 29 countries in the Western Palearctic and from Iran, and obtained detailed geographic coverage of the European range. Sample location and information are given in Fig. 1 and Table 1; all maps in this paper were generated using QGIS 2.8.1 (QGIS Development Team 2015). Samples originated from road-kills, museum specimens, individuals in wildlife rescue centres, and live-trapped animals (see Table S1 for additional sample details). The Iberian samples have been used in our previous studies on the phylogeography and genetic structure of *M. foina* within the Iberian Peninsula (Basto et al. in review; Vergara et al. in review).

Tissue samples were preserved in ethanol or in a salt-saturated solution of 20% DMSO in water (Seutin et al 1990) and stored at -20ºC; hair samples were kept dry until DNA extraction. We extracted genomic DNA using the DNeasy Blood & Tissue Kit (Qiagen) following the manufacturer's protocols.

Mitochondrial DNA sequencing and microsatellite genotyping

We amplified and sequenced a fragment of 621 bp of mitochondrial DNA (mtDNA) encompassing a portion of the Cyt *b* gene, the threonine and proline tRNAs and the left domain of the control region using the primers L15533 (Fernandes et al 2008) and H16437 (5'- GGAGCGAGAAGAGGTACA-3'). Polymerase chain reactions (PCRs) were carried out in a total volume of 15 µl containing 1.25 U of Supreme NZYTaq DNA polymerase (NZYTech), 1X PCR buffer, 2 mM $MgCl_2$, 0.2 mM of each dNTP (Bioline), 1 µM of each primer, 0.17 µg/µL BSA (New England Biolabs) and 3 µL of DNA extract (20-100 ng). The thermal cycling was as follows: initial denaturation at 94 ºC for 5 min, followed by 55 cycles of 94 ºC for 30 s, 60 ºC for 30 s, and 72 ºC for 1 min, and a final extension at 72 ºC for 7 min. PCR products were analysed by electrophoresis in 2% agarose gels; positive results were purified using ExoSAP (Hanke and Wink 1994) and sequenced at Macrogen Inc. Sequences were assembled, aligned, and edited using Sequencher 4.7 (Gene Codes Corporation).

Samples were genotyped with a multiplex panel of 23 autosomal microsatellite markers designed by Vergara et al (in review), comprising the 13 available species-specific loci (Basto et al 2010) plus 10 selected from panels developed in other mustelid species. PCR conditions were as described in the same paper. Briefly, multiplex amplifications were carried out using

Country	Samples (n)	mtDNA (n)	Microsatellites (n)
Albania	1	1	1
Austria	26	17	26
Belgium	3	3	3
Bulgaria	6	6	6
Croatia	21	21	19
Czech Republic	3	3	3
Denmark	26	26	26
Estonia	2	2	2
France	25	24	25
Germany	29	28	25
Greece	58	54	30
Hungary	22	19	22
Iran	2	1	1
Israel	7	7	7
Italy	31	31	25
Lithuania	4	4	4
Luxembourg	26	25	22
Macedonia	2	2	2
Nederland	21	20	21
Poland	26	25	26
Portugal	93	86	85
Romania	5	4	5
Servia	25	22	24
Slovakia	7	7	7
Slovenia	4	4	4
Spain	257	170	248
Switzerland	13	12	12
Syrien	1	0	1
Turkey	1	1	1
Ukraine	4	4	3
TOTAL	**751**	**629**	**686**

Table 1. Number of stone marten (*Martes foina*) samples collected per country and used for mtDNA sequencing and microsatellite genotyping

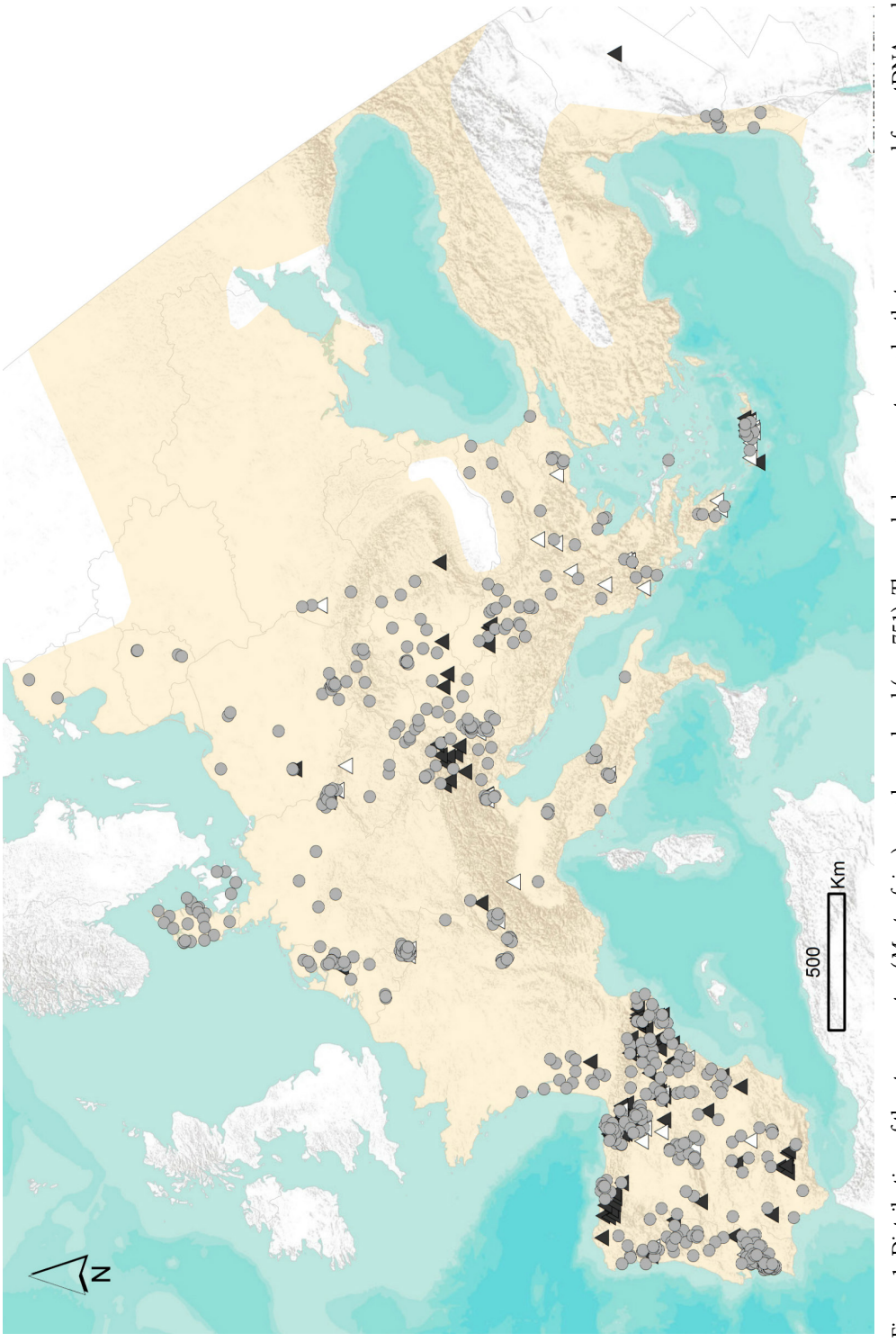


Figure 1. Distribution of the stone marten (*Martes foina*) samples analysed (n = 751). The symbols represent samples that were sequenced for mtDNA and genotyped for microsatellites (grey circles), only sequenced for mtDNA (white triangles), and only genotyped for microsatellites (black triangles). The present stone marten range is shaded in orange (Tikhonov et al. 2008).

the Qiagen Multiplex PCR Kit in a total volume of 9 µl containing 1 µl of DNA and 2 pmol of each primer. Thermal cycling consisted of a hot start protocol as follows: initial polymerase activation at 95 °C for 15 min, followed by 42 cycles at 94 °C for 30 s, 57 °C for 130 s, and 72 °C for 1 min, and a final extension at 60 °C for 30 min. PCR products were run on an ABI 3130xl Genetic Analyzer (Applied Biosystems) with the internal size standard GeneScan-500 LIZ (Applied Biosystems), and fragment analysis was performed using GeneMapper 4.0 (Applied Biosystems). To check for genotyping errors, we amplified 30 samples (20 of hair and 10 of tissue) in four replicates and estimated allelic dropout (ADO) and false allele (FA) rates using GIMLET 1.3.3 (Valière 2002). Since no errors were detected, the remaining samples were genotyped only once per locus.

Mitochondrial DNA analysis

Genealogical relationships among haplotypes

The sequence alignment was analysed with FaBox 1.41 (Villesen 2007) to collapse identical sequences into haplotypes. Phylogenetic analyses were performed using Bayesian inference (BI) in MrBayes 3.2.4 (Ronquist et al 2012), maximum likelihood (ML) in RAxML 7.2.8 (Stamatakis 2006), minimum evolution (ME) in FastME 2.1.3 (Desper and Gascuel 2002), and maximum parsimony (MP) in MEGA 6.0 (Tamura et al 2013). A GenBank sequence of stone marten from China (HM106325) was added to the data set. As outgroups, we used sequences of pine marten, sable *Martes zibellina* (Linnaeus 1758), American marten *Martes americana* (Turton 1806), Japanese marten *Martes melampus* (Wagner 1840) and yellow-throated marten *Martes flavigula* (Boddaert 1785) from GenBank (accession numbers respectively: KC660129, HM106323, HM106324, AB291076 and HM106326). The outgroups were chosen based on previously published phylogenetic hypotheses for the genus *Martes* (e.g. Koepfli et al 2008; Li et al 2014). Since the addition of outgroups introduced a few small indels in the control region segment of the alignment, we used M-COFFEE (Wallace et al 2006), a meta-aligner that combines the solutions of alternative alignment methods, to estimate a consensus alignment. We combined the three top-performing methods (PROBCONS, T-COFFEE and MAFFT) in a recent benchmark study of sequence alignment algorithms (Thompson et al 2011), with MAFFT being previously shown to be highly accurate when dealing with indel-rich sequences (Golubchik et al 2007). Substitution saturation was checked using the index of (Xia et al 2002) in DAMBE 5.3.74 (Xia 2013). Analyses in MrBayes were conducted with two parallel runs, each with four Markov chains (one cold and three heated) and 20 million generations. The first five million generations were

discarded as burn-in and, thereafter, chains were sampled every 500 generations. The entire general time-reversible (GTR) substitution model space was sampled within the analyses (Huelsenbeck et al 2004). Convergence of the two runs was assumed when the average standard deviation of the split frequencies was less than 0.01. Support for tree nodes was determined based on the values of Bayesian posterior probability (BPP) obtained from a majority-rule consensus tree (Holder et al 2008). In RAxML we used a GTR+Γ model of sequence evolution, in FastME the initial tree was constructed with the BIONJ version of the neighbour-joining algorithm (Gascuel 1997) and the subsequent tree search was carried out by subtree-prunning-and-regrafting (SPR), and in MEGA the MP trees were obtained using the tree-bisection-regrafting (TBR) algorithm (Nei and Kumar 2000) on initial trees obtained with 10 replicates of random addition of sequences. The support for each node in the ML, ME, and MP trees was evaluated by 1000 bootstrap replicates. Majority-rule consensus trees were visualized and edited with TreeGraph 2.4.0 (Stöver and Müller 2010).

Genealogical relationships among haplotypes were also estimated using a median-joining network (Bandelt et al 1999) as implemented in PopART 1.7 (http://popart.otago.ac.nz). We also attempted to root the network, using the sister group of *M. foina* as outgroup (*M. martes* + *M. zibellina* + *M. americana* + *M. melampus*, which together with *M. foina* comprise the subgenus *Martes*), to infer ancestral haplotypes and polarity in the network.

mtDNA variation and geographical structure

The geographic structure of mtDNA variation in European stone martens was analysed with BAPS 6.0 (Corander et al 2008). The genetic mixture analysis was performed using the 'clustering with linked loci' model (Corander and Tang 2007), and either the option 'clustering of individuals' or the option 'clustering of groups of individuals' in which individuals were grouped a priori by country, or the 'spatial clustering of individuals' model (Cheng et al 2013).

We used Arlequin 3.5.1.3 (Excoffier and Lischer 2010) and DnaSP 5.10.1 (Librado and Rozas 2009) to compute the number of haplotypes (n_H), haplotype diversity (h), nucleotide diversity (π) and the mean number of pairwise differences (k) (Tajima 1983; Nei 1987) for each of the countries with sample sizes equal or greater than four individuals, and for identified clades and geographical groups of stone martens. Haplotype richness (H_R) was calculated by rarefaction in CONTRIB 1.02 (Petit et al 1998) (Petit et al. 1998). Genetic diversity estimates per country are obviously influenced by sample size and distribution of sampling

locations within each country, but they may still be useful in suggesting areas of glacial refugia, of secondary contact between differentiated lineages, or of impoverished genetic variability.

We also used Arlequin to assess the genetic differentiation between defined populations, as measured by the net average distance (D_a, Nei 1987) and pairwise Φ_{ST} (Excoffier et al 1992) under the Tamura-Nei model (Tamura and Nei 1993) with a Γ shape parameter α of 0.05 estimated using MEGA. The statistical significance of D_a and Φ_{ST} values was tested using 20,000 permutations.

Demographic history

Demographic events can leave characteristic signatures in the distribution of pairwise nucleotide differences between individuals, or mismatch distribution (Slatkin and Hudson 1991; Rogers and Harpending 1992). Multimodal or "ragged" mismatch distributions are assumed to characterize old populations of constant size, whereas unimodal distributions are expected in populations that underwent a recent expansion. We tested the fit of the observed mismatch distributions with each of the two models of population expansion implemented in Arlequin: a model of pure demographic expansion in an undivided population (Schneider and Excoffier 1999) and a model of range expansion in a subdivided population (Ray et al 2003; Excoffier 2004). Tests based on the mismatch distribution use little information available in the data (Felsenstein 1992) and can be very conservative (Ramos-Onsins and Rozas 2002). Therefore, we also employed two test statistics, F_S (Fu 1997) and R_2 (Ramos-Onsins and Rozas 2002), which have been shown to be powerful for detecting population growth (Fu 1997; Ramos-Onsins & Rozas 2002). We used DnaSP to calculate R_2 and Arlequin to calculate F_S and Tajima's D (Tajima 1989); the statistical significance of the estimates was assessed using 20,000 random samples under the hypothesis of selective neutrality and population equilibrium.

Age and timing of diversification of mtDNA lineages

We estimated coalescence times of the main identified clades and geographical groups of haplotypes using Bayesian inference as implemented in BEAST 1.8.2 (Drummond et al 2012). We conducted BEAST analyses on an alignment containing only stone marten sequences, assuming a constant-size coalescent tree prior and applying a strict clock with a substitution rate of 1.675% nucleotide substitutions per million years. This rate was estimated using equation $D = 2\mu T$ (Nei 1987), where D is genetic distance, μ is substitution

rate and T is time since divergence. Specifically, we used the average genetic distance between *M. foina* and *M. zibellina* and the estimates of about 2.8 million years for the time to their most recent common ancestor (TMRCA) obtained by Koepfli et al (2008) and Li et al (2014). The average genetic distance between the stone marten haplotypes in this study and the sable GenBank haplotypes reported by Malyarchuk et al (2014) was estimated using the maximum composite likelihood model (Tamura et al 2004) in MEGA, and equalled 0.0938 nucleotide substitutions per site. In BEAST we specified the HKY+I model, which best fitted the stone marten data according to jModelTest, and the respective parameter estimates for I and κ were set as initial prior values.

Four independent replicate runs were performed and their convergence and effective sample sizes (ESS), as well as the parameter estimates, were monitored in Tracer 1.6 (Rambaut et al 2014). In all runs, the ESS were greater than 500 for all parameters. Results from replicate runs were convergent and log and tree files were combined using LogCombiner. TreeAnnotator was used to annotate the maximum clade credibility tree, which was displayed by FigTree.

Microsatellite analysis

Genetic diversity and structure

The data was first checked using GenAlEx 6.501 (Peakall and Smouse 2012) to identify matching multilocus genotypes likely to represent replicate samples; one of each pair of matching samples collected within the same country was removed prior to subsequent analyses. Microsatellite genetic diversity was characterized by the mean number of alleles (N_A), observed heterozygosity (H_O) and unbiased expected heterozygosity (H_E) across loci for each of the countries with sample sizes equal or greater than six individuals, and for the genetic clusters inferred by the Bayesian clustering algorithms (see below). Mean allelic richness across loci (A_R) was calculated in HP-Rare 1.0 (Kalinowski 2005) using a rarefaction procedure to account for unequal sample size between populations (Kalinowski 2004).

We used the microsatellite data to investigate the large-scale genetic structure of European stone martens using spatially explicit models in BAPS and Geneland 4.0.5 (Guillot et al 2005) and a non-spatial model in Structure 2.3.4 (Pritchard et al 2000). In BAPS we conducted an individual level mixture analysis (Corander et al 2008) and the most likely number of clusters (*K*) was determined by evaluating the 10 best partitions found across runs with different values of tested *K* (10 replicates for each *K*) in terms of marginal likelihood.

In Geneland we used the spatial Dirichlet model, assumed uncorrelated allele frequencies as recommended by Guillot et al (2005), allowed for uncertainty in the positioning of individuals, defined an uniform prior on *K* between 1 and 20, and performed 10 independent runs each with 500,000 iterations and a thinning of 1,000 after a burn-in of 1,000 iterations. The choice of K was based on the histograms obtained in each run and the highest mean values of posterior density for each K. In Structure we implemented the admixture model with correlated allele frequencies (Falush et al 2003) and carried out 10 replicate runs for *K* = 1-20, each with one million iterations after a burn-in of 100,000. We used Structure Harvester 0.6.94 (Earl and vonHoldt 2012) to summarise Structure output and calculate for each *K* the posterior probability of the data (Ln P(D)) and a measure of the second-order rate of change of the likelihood function (Δ*K*, Evanno et al 2005). See Table S2 for a summary of the models and setting used in the Bayesian clustering analyses.

Results

Mitochondrial DNA

Haplotypes and their relationships

We identified 50 haplotypes (H1-H50) among the 629 stone martens sequenced. The GenBank sequence from China (HM106325) corresponded to an additional haplotype (H51). The 621-bp alignment of 630 sequences contained 56 (9%) polymorphic sites, of which 24 (4%) were parsimony informative and two were 1-bp indels; the number of observed substitutions was 55 (50 transitions and 5 transversions). The haplotype diversity was 0.882 ± 0.005, the nucleotide diversity was 0.0094 ± 0.0050, and the mean number of pairwise differences was 5.836 ± 2.793.

No substitution saturation was evident in the alignment with the five added outgroups for phylogenetic analyses, as the index of substitution saturation (I_{ss}, Xia et al. 2003) was significantly lower than the critical value (P = 0.0001).

The Bayesian and maximum likelihood inferences of the phylogenetic relationships of the mtDNA haplotypes are given in Fig. 2; the maximum parsimony and minimum evolution trees were essentially identical to the Bayesian tree and are shown in Fig. S1. Within *M. foina*, two main clades were consistently recovered (Clades I and II) while the remaining haplotypes were ungrouped or clustered into clades of two haplotypes each. The only major difference between the ML tree and the other trees concerned the placement of haplotype H51, from a specimen from China, which falls in the same clade with other *M. foina* haplotypes in the former tree, whereas it is basal to a clade of western Eurasian stone martens in all the other trees.

Clade I included haplotypes mainly found in Western Europe, particularly in the Iberian Peninsula, France, Switzerland, the Low Countries, Germany and Denmark. Haplotypes of this clade were also present in Italy, in the Alpine-Carparthian belt (Slovenia-Hungary-Romania), and in Israel. Clade II was composed of haplotypes distributed in the North European Plain and the Baltics, from Luxembourg and Denmark up to Estonia, and in East-Central Europe across the Alpine-Carpathian range, from Austria and Czech Republic to Croatia, Serbia and Romania.

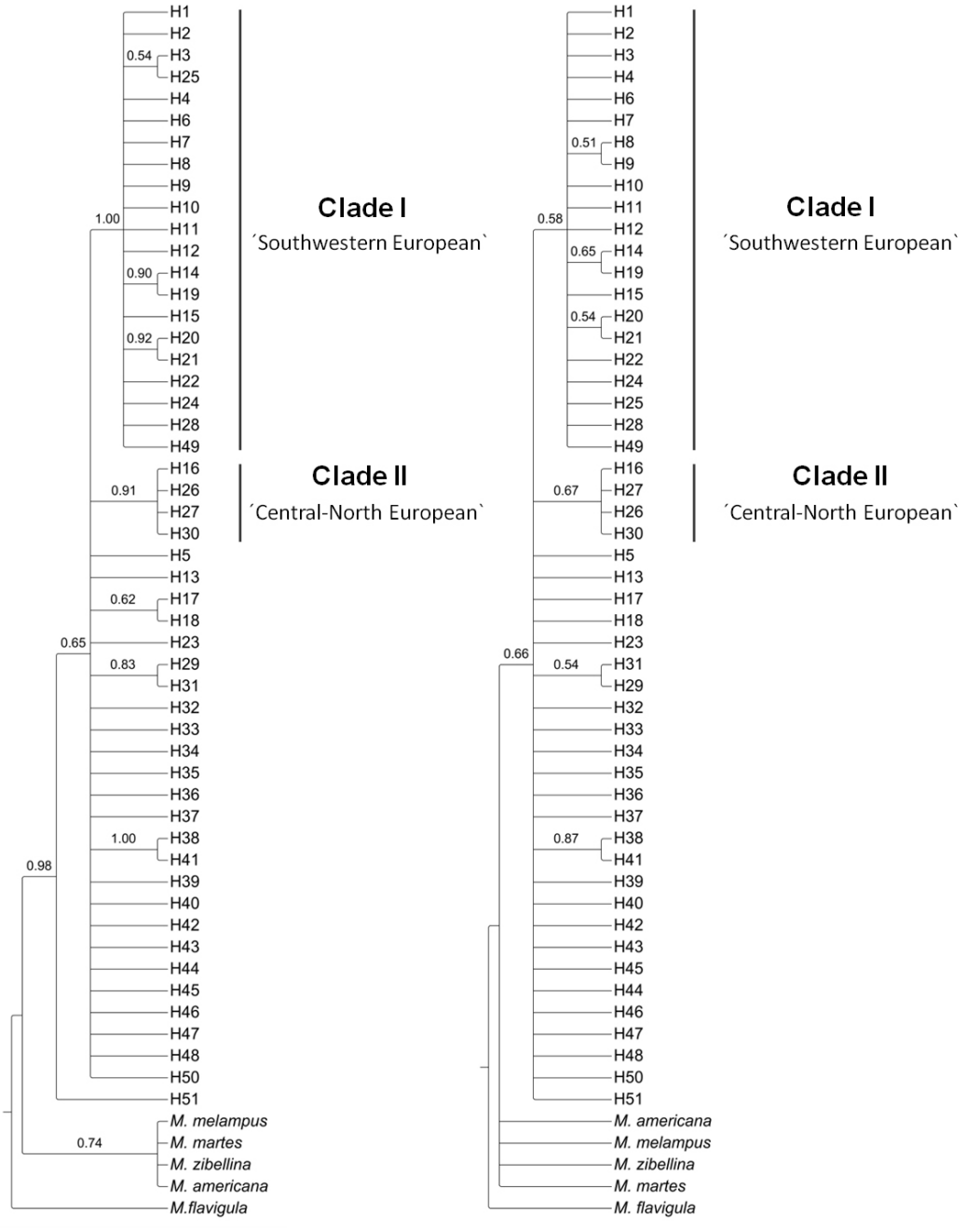


Figure 2. Majority-rule consensus trees of the Bayesian (left) and maximum likelihood (right) analyses of mtDNA haplotypes. Numbers above branches are, respectively, Bayesian posterior probabilities and bootstrap values. Haplotype codes are as in Table S.1. Clades were named (I and II) for discussion purposes.

The haplotype genealogy estimated using the median-joining algorithm was consistent with the phylogenetic trees, while displaying more information on the relationships between haplotypes and between clades (Fig. 3). For instance, among the haplotypes that did not fall into a major cluster in the phylogenetic trees, most were private to the Balkans and Crete, but it was possible to identify a set of haplotypes (H13, H17, H18, and H23) that were rare or absent in southeastern Europe, being instead found from Lithuania across the North European Plain to France and Switzerland and in Italy. Hereafter, for convenience, this group of haplotypes is called 'Italian-North European'. A common haplotype in the Balkans and Crete (H5), occupying a central position among the haplotypes found in those areas, was also detected in the Alpine-Carparthian belt, in the Netherlands, in Italy, and in Spain near the eastern Pyrenees. H16, which appeared to be ancestral in Clade II, was common in central and north Europe (Fig. 3). Star-shaped patterns, with one or two common central haplotypes surrounded by several rarer haplotypes differing by one to a few mutations from the central haplotypes, were observed in Clade I and around H5. Such star-like radiations of haplotypes are consistent with past demographic expansions (Slatkin and Hudson 1991; Merila et al 1997). The network rooted with the sister group of *M. foina* (Fig. S2) suggest that the oldest stone marten lineages in the data set are the one from China and, within Europe, those that occur predominantly in southeastern Europe. Accordingly, the Iranian haplotype (H50) was closely related to haplotypes found in the Balkans. Outgroup rooting is more reliable when ingroup and outgroup are reciprocally monophyletic sister groups and when they are separated by a relatively short branch (Posada and Crandall 2001; Cassens et al 2003), which is the case here.

mtDNA variation and geographical structure

The results of the BAPS analyses of the mtDNA data differed according to the model employed, but there were clear common patterns. The 'spatial clustering of individuals' model identified three clusters in western Eurasian stone martens (Fig. 4a) that corresponded, respectively, to Clade I, Clade II and the set of haplotypes that did not fall into a major clade in the phylogenetic trees. The two latter clusters were also identified using the non-spatial model 'clustering with linked loci' with the option 'clustering of individuals' but, under this model, BAPS detected differentiation among individuals within Clade I haplotypes: between those with haplotypes mostly restricted to the south of Western Europe (Iberian Peninsula, France, Italy), with a few reaching Switzerland and Slovenia, and those with haplotypes also found further to the north and east (Low Countries, Germany, Denmark, Hungary, Romania, Israel) and closely related ones (Fig. 4b, see also Fig. 3). The same partition of individuals

with Clade I haplotypes, as well as a cluster corresponding to Clade II, were recovered using the 'clustering with linked loci' model with the option 'clustering of groups of individuals'. With this model, individuals were grouped *a priori* by country, but this model inferred two additional clusters composed of individuals carrying the haplotypes that did not fall into a major clade in the phylogenetic trees: one of these clusters corresponded to the set of haplotypes also identified in the median-joining network and that we called 'Italian-North European', while the other cluster contained individuals with the remaining haplotypes (Fig. 4c). Members of the latter cluster were predominantly found in the Balkans and Crete, and it was clearly the dominant cluster in southeastern Europe; therefore it is subsequently designated as 'Balkan' group. As implied by their names, spatial segregation of the 'Italian-North European' and 'Balkan' clusters was evident. In contrast, the division of Clade I, consistently obtained with the 'clustering with linked loci' model, was less geographically explicit (Fig. 4c). Indeed, the two central haplotypes in each of these subgroups suggested by BAPS within Clade I, H2 and H7, respectively, were common in Iberia, even if the subgroup containing H2 could be found much further north and east than the subgroup containing H7 (Fig. 3).

Tables 2 and 3 respectively present measures of mtDNA diversity estimated by country and for clades and groups of stone martens identified in the genealogical and Bayesian clustering analyses above. In the analysis by country, haplotype richness and diversity were highest in Slovenia, Switzerland, Serbia and mainland Greece, and lowest in Hungary, Slovakia and Israel. Nucleotide diversity was highest in some countries of Western Europe (Netherlands, Denmark, France and Switzerland) and lowest in Portugal, Crete and Israel (Table 2). Among the clades and geographical clusters of European stone martens, haplotype diversity was highest in Clade I but haplotype richness was highest in the 'Balkan' cluster. In both of these groups, haplotype richness and diversity were moderately high and nucleotide diversity was relatively low. In Clade II and in the 'Italian-North European' cluster, haplotype and nucleotide diversity were similar and very low (Table 3). Given their main geographical distribution (Fig. 4a), and to facilitate discussion in a phylogeographic context, clades I and II are henceforth referred to as 'Southwestern European' and 'Central-North European', respectively.

Pairwise D_a distances between delineated geographic lineages in European stone martens varied from 0.17 to 1.51%, and differentiation measured by Φ_{ST} ranged from 0.60 to 0.99; all comparisons were significant ($P < 0.05$). Using equation $D = 2\mu T$ and the substitution rate of 1.675% nucleotide substitutions per million years, the observed Da distances translate into coarse estimates of the time since divergence between lineages. All the estimates fell within

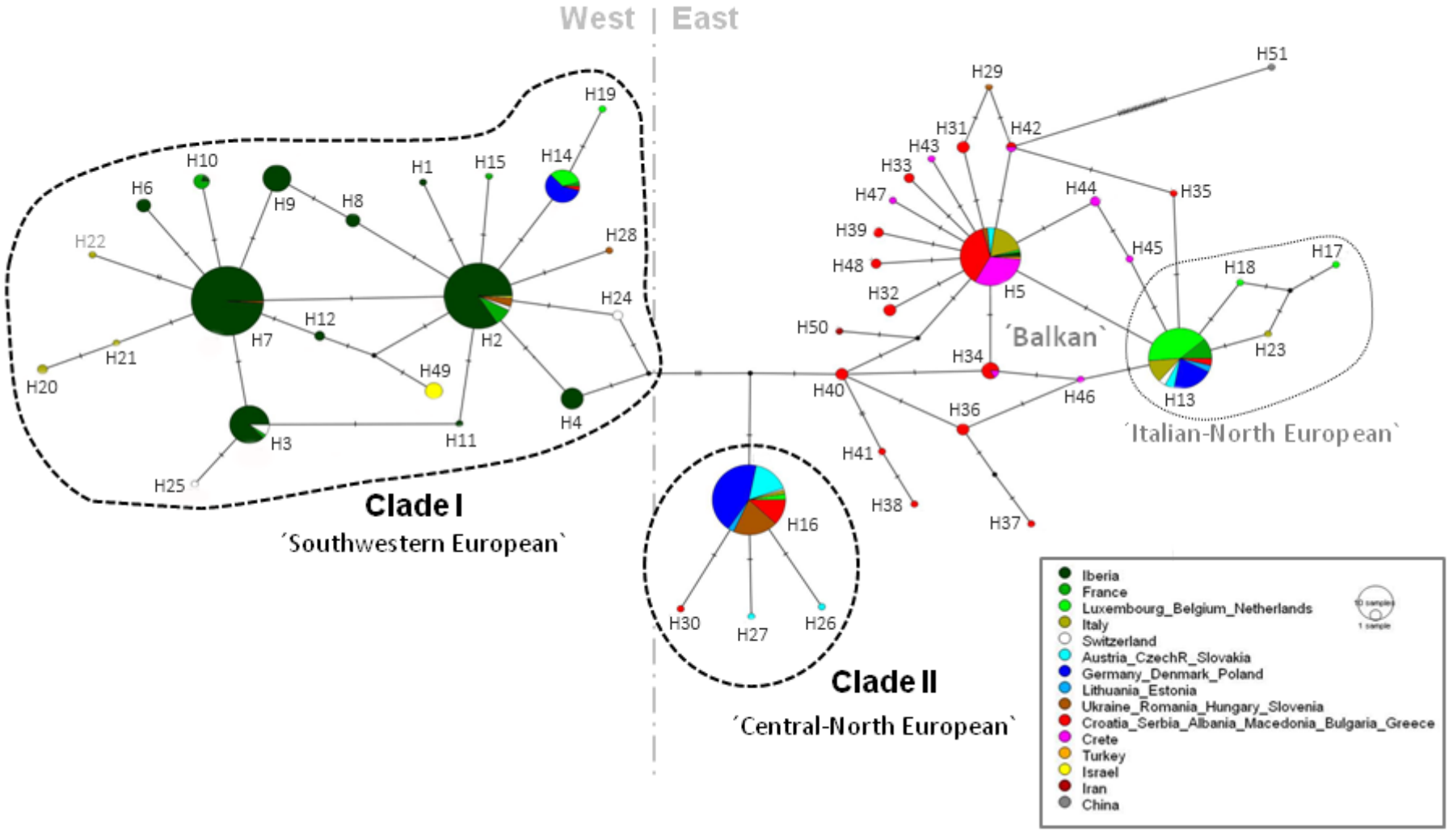


Figure 3. Median-joining network based on mtDNA haplotypes. Circles represent haplotypes and their size is proportional to the frequency observed. Coloured pie charts indicate the frequency of haplotypes in countries or regions. Small black circles represent hypothetical haplotypes. Dashes on lines connecting haplotypes represent the number of nucleotide substitutions separating them. Haplotype codes are given in Table S.1. The clades identified in the phylogenetic analyses are indicated, as well as a group of haplotypes, named 'Italian-North European', discussed in the text.

Country*	Mitochondrial DNA					Microsatellites				
	n	n_H	H_R	h	π	n	N_A	A_R	H_E	H_O
Portugal	86	4	0.92	0.49±0.04	0.0008±0.0001	85	4.87	3.16	0.58±0.03	0.49±0.03
Spain	170	11	1.89	0.78±0.02	0.0021±0.0002	248	7.17	3.46	0.61±0.03	0.50±0.03
France	24	6	1.74	0.75±0.05	0.0071±0.0006	25	4.65	3.74	0.61±0.04	0.51±0.04
Luxembourg	25	5	0.9	0.42±0.12	0.0033±0.0012	22	4.35	3.38	0.59±0.03	0.47±0.05
Netherlands	20	4	1.23	0.60±0.08	0.0075±0.0008	21	4.35	3.32	0.60±0.03	0.44±0.04
Italy	31	7	1.61	0.70±0.06	0.0045±0.0011	25	5.09	3.83	0.66±0.03	0.49±0.04
Switzerland	12	6	2.31	0.88±0.06	0.0068±0.0010	12	4.35	3.65	0.63±0.04	0.52±0.04
Austria	17	5	1.46	0.64±0.12	0.0037±0.0008	26	5.26	3.74	0.65±0.03	0.50±0.04
Germany	28	3	0.86	0.42±0.10	0.0039±0.0010	25	4.48	3.18	0.58±0.04	0.48±0.04
Denmark	26	3	1.48	0.69±0.03	0.0074±0.0005	26	3.65	2.7	0.50±0.04	0.39±0.04
Slovakia	7	2	0.57	0.29±0.20	0.0023± 0.0016	7	3.22	3.04	0.55±0.04	0.51±0.05
Poland	25	3	0.91	0.45±0.10	0.0041±0.0010	26	4.13	3.18	0.60±0.03	0.49±0.04
Lithuania	4	2	1	0.50± 0.27	0.0040±0.0021	4	-	-	-	-
Ukraine	4	2	1	0.50±0.27	0.0032±0.0017	3	-	-	-	-
Romania	4	2	1	0.67±0.20	0.0054±0.0017	5	-	-	-	-
Hungary	19	2	0.39	0.20±0.11	0.0016±0.0009	22	5.22	3.45	0.58±0.05	0.43±0.05
Slovenia	4	4	3	1.00±0.18	0.0100±0.0026	4	-	-	-	-
Croatia	21	4	1.23	0.60±0.07	0.0036±0.0004	19	4.78	3.74	0.62±0.05	0.54±0.06
Serbia	22	8	2.31	0.87±0.04	0.0038±0.0010	24	6.3	4.44	0.70±0.03	0.55±0.03
Greece (mainland)	22	9	2.11	0.82±0.07	0.0035±0.0006	18	6.26	4.18	0.69±0.03	0.53±0.04
Crete	32	8	0.99	0.44±0.11	0.0009±0.0003	12	4.26	3.37	0.59±0.05	0.54±0.05
Bulgaria	6	3	1.6	0.73±0.16	0.0019±0.0007	6	3.96	3.74	0.67±0.03	0.52±0.06
Israel	7	2	0.57	0.29±0.20	0.0005±0.0003	7	3.91	3.72	0.65±0.04	0.50±0.04

Table 2. Measures of mitochondrial DNA and microsatellite diversity by country in western Eurasian stone martens. n, number of samples; nH, number of haplotypes; HR, haplotype richness; h, haplotype diversity; π, nucleotide diversity; NA, mean number of alleles; AR, average allelic richness; HE, mean expected heterozygosity; HO, mean observed heterozygosity. The asterisk signals the exception for Greece, where samples from the mainland and from Crete were analyzed separately. Dashes indicate parameters that were not estimated when the sample size was lower than six individuals

Clade/Group	n	n_H	H_R	h	π	*k*	SSD (sudden)	SSD (spatial)	F_S	R_2	*D*
M. foina	630	51	16.998	0.88±0.01	0.0094±0.0050	5.84±2.79	0.02	0.023	-16.344*	0.00*	-1.192
W. Eurasian	629	50	16.878	0.88±0.01	0.0091±0.0048	5.62±2.70	0.02	0.022	-15.748*	0.00*	-0.39
SW. European	318	21	11.548	0.78±0.01	0.0021±0.0015	1.31±0.82	0.008*	0.008*	-12.835*	0.00*	-1.378
C.-N. European	113	4	2.257	0.05±0.03	0.0001±0.0002	0.06±0.13	0	0	-5.984*	0.00*	-1.591*
Italian-N. European	85	4	3.000	0.07±0.04	0.0002±0.0004	0.13±0.20	0	0	-4.027*	0.00*	-1.489*
Balkan	113	21	17.167	0.56±0.06	0.0018±0.0013	1.12±0.73	0	0	-20.248*	0.00*	-1.932*

Table 3. Measures of mitochondrial DNA diversity and results of mismatch distribution tests of demographic expansion and of tests of neutrality and mutation-drift equilibrium for clades and geographical groups of stone martens. n, number of samples; nH, number of haplotypes; HR, haplotype richness; h, haplotype diversity; π, nucleotide diversity; k, mean number of pairwise differences; SSD, sum of square deviations between the observed and expected mismatch distributions under the sudden and spatial expansion models; FS, test statistic of Fu (1997); R2, test statistic of Ramos-Onsins & Rozas (2002); D, test statistic of Tajima (1989). Asterisks indicate significant results ($P < 0.02$ for FS and $P < 0.05$ for the other tests).

the second half of the Middle Pleistocene, with the exception of the less differentiated 'Italian-North European' and 'Balkan' clusters, whose split was dated to the Pleniglacial (Table 4).

	Southwestern European	**Central-North European**	**Balkans**	**Italian-North European**
Southwestern European		0.87	0.88	0.9
Central-North European	1.065 (317,910)		0.88	0.99
Balkans	1.443 (430,746)	0.672 (200,597)		0.6
Italian-North European	1.512 (451,343)	0.974 (290,746)	0.165 (49,254)	

Table 4. Pairwise mtDNA Da distances in percentage (below diagonal) and ΦST values (above diagonal) between identified geographical lineages of European stone martens. Within parentheses next to the Da values are the corresponding divergence time estimates in years before present, using a substitution rate of 1.675% nucleotide substitutions per million years. All Da and ΦST values were significant ($p < 0.05$).

Demographic history

The mismatch distributions and the results of the three neutrality tests were consistent with the scenario of past demographic expansion for the 'Balkan', 'Italian-North European', and 'Central-North European' lineages. For the 'Southwestern European', the hypothesis of past growth was supported by significant values of F_S and R_2, but Tajima's D and the mismatch distribution did not differ from those expected under demographic equilibrium (Table 3). For both lineages 'Central-North European' and 'Italian-North European', which showed similar levels of mtDNA variation (Table 3), the time since expansion in units of mutational time before present (tau, τ) under the model of pure demographic expansion was estimated at 3.00, with 95% confidence intervals of 0.54 to 3.00. Using equation $\tau = 2ut$ (Rogers & Harpending 1992), where t is the time in generations since expansion and u is the substitution rate per year of the entire nucleotide sequence studied (i.e. $1.675 \times 10^{-8} \times 621$), and assuming a generation time of two years for stone martens (Jones et al 2009), the expansions were estimated to have started around 72,103 years ago (95% CI: 12979-72103). For the 'Balkan' lineage, the inferred expansion was estimated at 0.43 τ units (95% CI: 0.00-3.28), which translated into 10,335 years BP (95% CI: 0-78833). Analyses of the 15 country samples containing more than 10 individuals (in the case of Greece, the mainland and Crete were considered separately), only found strong evidence of population growth for the Cretan sample (Table S3).

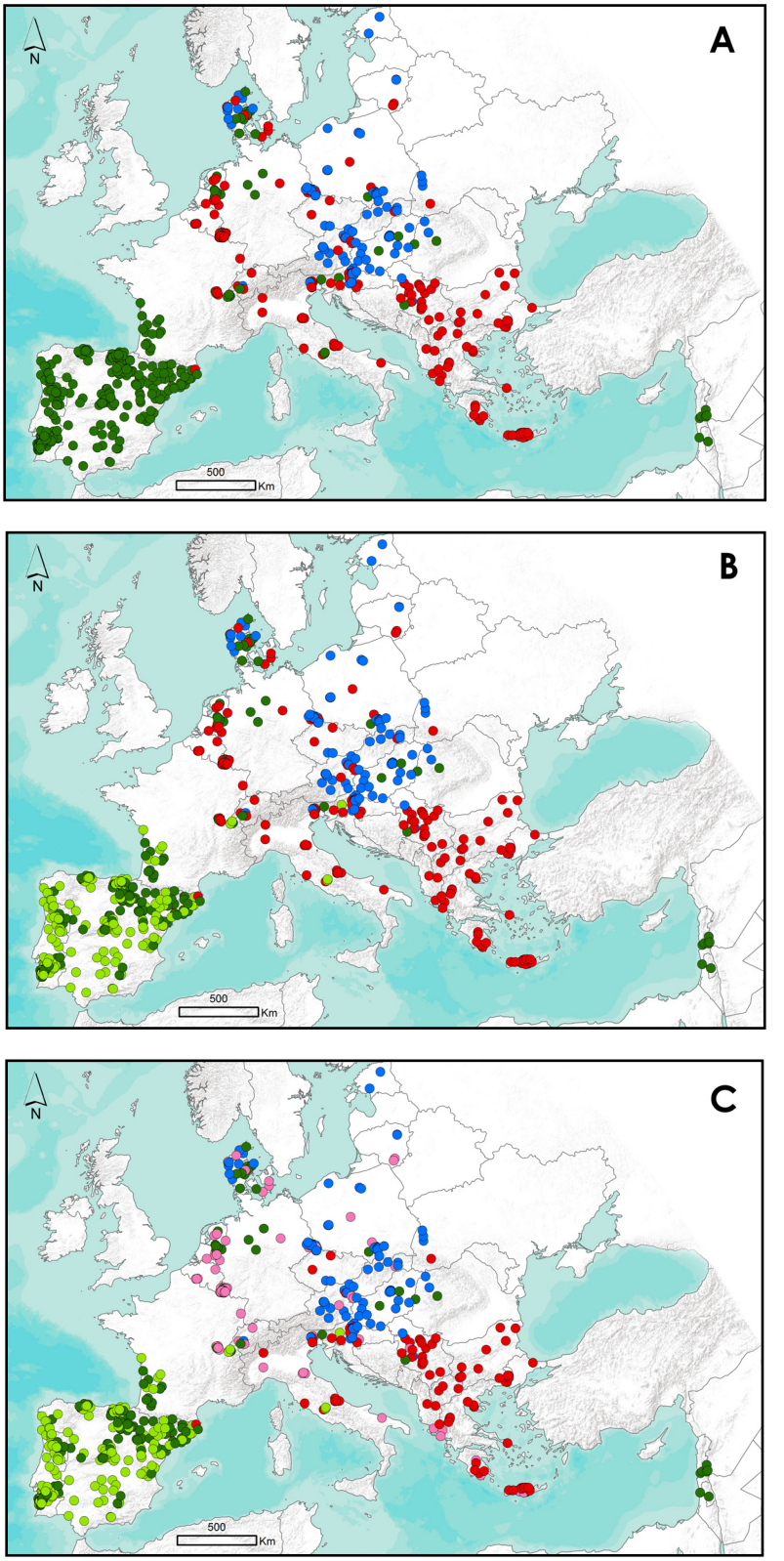


Figure 4. BAPS inferences of the geographic structure of mtDNA variation in European stone martens. Circles represent samples and different colours represent clusters in the optimal partition. a) 'spatial clustering of individuals' model (K=3). The cluster in green corresponds to Clade I in the phylogenetic trees, the cluster in blue corresponds to Clade II, and the cluster in red to haplotypes that did not fall into a major clade in the phylogenetic trees; b) 'clustering with linked loci' model with the option 'clustering of individuals' (K=4). A similar solution to that of the previous model, with two clusters identical (in blue and red), but with a division of Clade I sequences, represented by the two clusters in dark and light green; c) 'clustering with linked loci' model with the option 'clustering of groups of individuals', in which individuals were grouped a priori by country (K=5). A similar solution to that of the previous model, with three clusters identical (in blue and dark and light green), but with a partition of individuals carrying the haplotypes that did not fall into a major clade in the phylogenetic trees (the red cluster in the previous two models) into an 'Italian-North European' cluster (pink) and a 'Balkan' cluster (red).

Estimates of the age of defined mtDNA lineages

Table 5 and Fig. S3 show the coalescence time estimates obtained from BEAST, for clades and geographical groups of stone martens. These results agree with the divergence times estimated from the pairwise D_a distances in suggesting the presence of *M. foina* in Europe since the Middle Pleistocene and an east-west split in the second half of this period. The TMRCAs of the lineages 'Central-North European' and 'Italian-North European' at 100-90 kyr BP are also compatible with the timing for their expansion estimated from the mismatch distribution.

Clade/Lineage	TMRCA (95% HPD)
M. foina	2,043 (1081-3149)
Western Eurasian	554 (313-835)
Southwestern European	229 (103-373)
Central-North European	92 (14-202)
Italian-North European	100 (19-210)
Balkan	404 (188-661)

Table 5. Time to the most recent common ancestor (TMRCA) in thousands of years for clades and geographical lineages of stone martens. Shown are the means and 95% highest posterior density (HPD) intervals obtained using an estimated rate of 1.675 x 10-8 substitutions per site per year on a stone marten-only alignment, with a strict molecular clock and a constant coalescent tree prior.

Microsatellite analysis

Genetic diversity and structure

In total, 686 stone martens from 30 countries across the range of the species in western Eurasia were genotyped for 23 microsatellite loci (Table 1). Allelic richness and expected heterozygosity were highest in southeastern Europe (Serbia, mainland Greece, Bulgaria), Italy, and in some countries of Central and Western Europe (Austria, Switzerland, France). Some of these countries (Serbia, mainland Greece, Switzerland) also showed the highest observed mtDNA diversity (Table 2). Microsatellite diversity was lower in Iberia, in the North European Plain (Luxembourg, Netherlands, Germany, Denmark, Poland), and in Crete.

The non-spatial algorithm used in Structure identified two clusters ($K = 2$; $F_{ST} = 0.096$, $D_{EST} = 0.191$, P = 0.000 for both) as the uppermost hierarchical level of genetic structure in

European stone martens (Fig. 5a). This East-West split separates individuals in southwest Europe from those found elsewhere in the continent. The distinction of the southwest European populations mirrors the pattern observed in the mtDNA data (Figs. 3 and 4a). Spatially explicit analyses in GENELAND with varying K inferred seven clusters ($K = 7$; Fig. 5b) as the mode of the posterior distribution, but subsequent runs with K fixed to 7, to check the consistency of the clustering arrangement and to obtain cluster membership probabilities for each individual, indicated six clusters ($K = 6$; Fig. 5c). The two clustering solutions revealed some similar patterns but also had important differences. Both consistently identified the same clusters in east-central and northern Europe (in dark blue) and in the Pannonian Basin (in brown), and suggested the presence of two clusters in the Iberian Peninsula (S-W and N-E: light and dark green, respectively). However, the two hypotheses differed with respect to the location of the NW-SE boundary in Iberia, and the $K = 6$ solution grouped NE Iberia with SW France. The two other clusters proposed in this solution grouped, respectively, stone martens from Italy, the Alps and the south of the Low Countries (in pink), and from the southeastern Europe and the Middle East (in red) (Fig. 5c). Conversely, the $K = 7$ solution grouped individuals from France, Switzerland and the south of the Low Countries (in light blue), from Italy, eastern Alps and Balkans (in red), and from Crete and the Middle East (in purple) (Fig. 5b). Genetic differentiation between the clusters in each geographic structuring inferred by GENELAND, as measured by F_{ST} and D_{EST} (Table 6), was significant for all pairwise comparisons ($P < 0.05$). Global F_{ST} and D_{EST} values were, respectively, 0.079 and 0.152 for the $K = 6$ solution, and 0.119 and 0.224 for the $K = 7$ solution. There are remarkable resemblances between the geographic patterns of microsatellite variation inferred by GENELAND, particularly as revealed by the $K = 6$ solution, and the identified mtDNA phylogeographic structure (compare Figs. 4c and 5b). BAPS identified partitions with a large number of clusters in both the spatial ($K = 16$; Fig. S4a) and non-spatial ($K = 18$; Fig. S4b) models. The BAPS results were overall comparable with the $K = 7$ solution from GENELAND, and the additional clusters are potentially either a signal of finer-scale genetic structure, for instance within Iberia (Vergara et al, in review), or artefacts.

F_{ST} \ D_{EST}	S & W Iberia	N & E Iberia	France + Switzerland + southern Low Countries	East-central & northern Europe	Pannonian Basin	Balkans + eastern Alps + Italy	Crete + Middle East
S & W Iberia	-	0.073	0.188	0.272	0.295	0.226	0.342
N & E Iberia	0.049	-	0.146	0.257	0.26	0.258	0.364
France + Switzerland + southern Low Countries	0.113	0.086	-	0.148	0.133	0.139	0.365
East-central & northern Europe	0.165	0.15	0.088	-	0.077	0.135	0.354
Pannonian Basin	0.169	0.144	0.073	0.049	-	0.088	0.35
Balkans + eastern Alps + Italy	0.119	0.129	0.068	0.072	0.043	-	0.263
Crete + Middle East	0.18	0.18	0.165	0.177	0.156	0.109	-

F_{ST} \ D_{EST}	West-central Iberia	NE Spain & SW France	Italy + Alps + southern Low Countries	East-central & northern Europe	Pannonian Basin	Southeastern Europe + Middle East
West-central Iberia		0.044	0.174	0.228	0.225	0.221
NE Spain & SW France	0.028		0.151	0.204	0.196	0.235
Italy + Alps + southern Low Countries	0.093	0.078		0.081	0.05	0.106
East-central & northern Europe	0.133	0.115	0.044		0.053	0.187
Pannonian Basin	0.128	0.107	0.026	0.033		0.123
Southeastern Europe + Middle East	0.108	0.107	0.045	0.089	0.054	

Table 6. Pairwise FST (below diagonal) and Jost's DEST (above diagonal) between the clusters inferred by GENELAND from the microsatellite data: a) with varying K; b) with K fixed at the modal value in the runs with varying K.

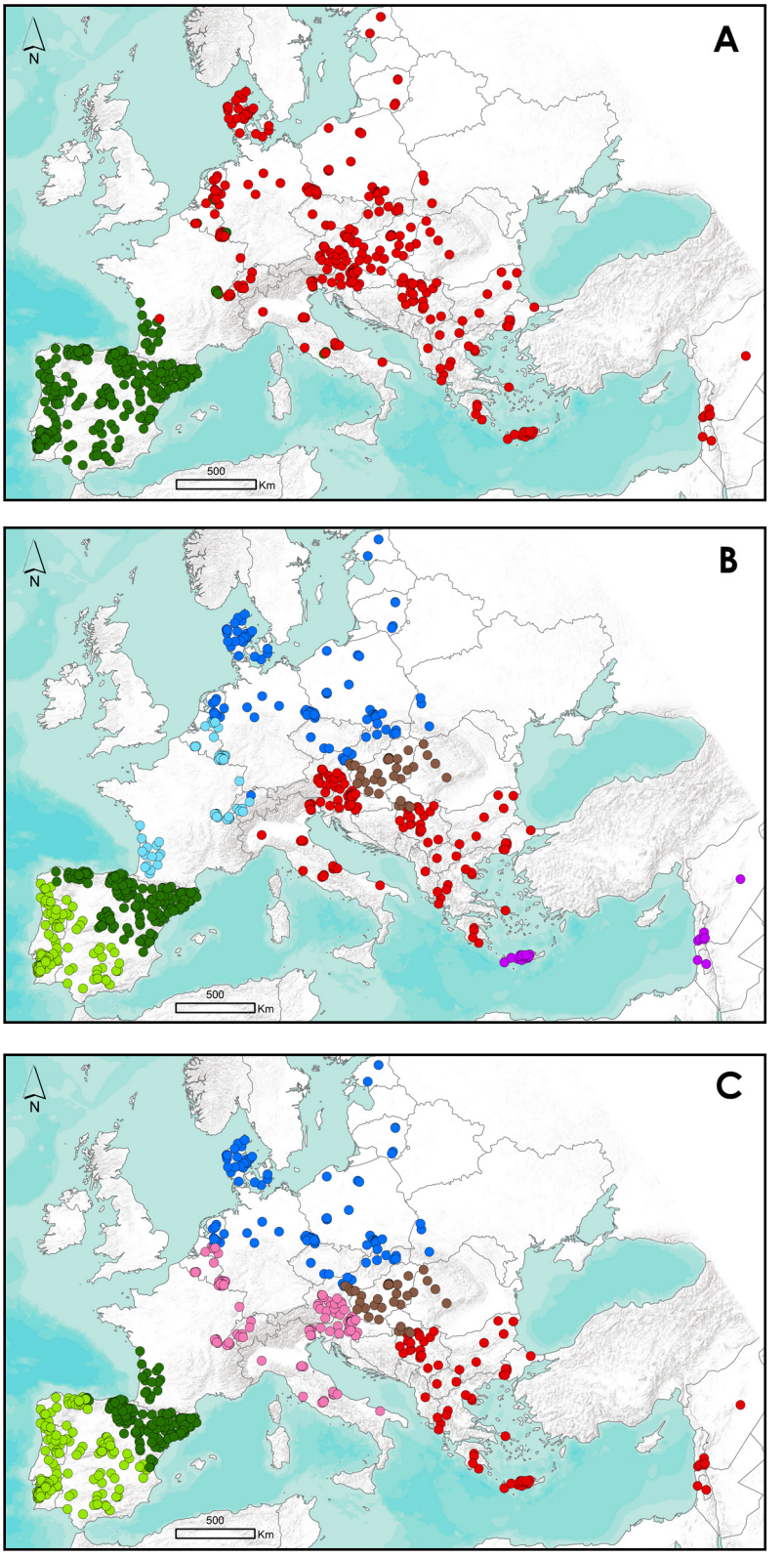


Figure 5. Bayesian clustering analyses of geographic genetic structure in European stone martens based on microsatellite loci. Circles represent samples and different colours represent clusters in the optimal partition. a) STRUCTURE (K=2). A split between southwest Europe (in green) and the rest of the range in the continent (in red); b) GENELAND with varying K (K=7). Clusters in east-central and northern Europe (in dark blue), in the Pannonian Basin (in brown), in the Iberian Peninsula (light and dark green), in France, Switzerland and the south of the Low Countries (in light blue), in Italy, eastern Alps and Balkans (in red), and in Crete and the Middle East (in purple); c) GENELAND with K fixed at the modal value in the runs with varying K (K=6). Clusters in east-central and northern Europe (in dark blue), in the Pannonian Basin (in brown), in south and west Iberia (in light green), in NE Spain and SW France (in dark green), in Italy, the Alps and the south of the Low Countries (in pink), and in the Balkans and the Middle East (in red).

Discussion

The present work is the first study focused on the phylogeography and genetic structure of *M. foina* in Europe and was based on a large and geographically comprehensive sampling across the continent. The results revealed high levels of haplotype diversity and a major phylogeographic east-west split, consistently found in the mtDNA analyses, which was dated to the Middle Pleistocene. The western clade comprised mostly individuals sampled in Western Europe, from the Iberian Peninsula to Denmark, and we identified three geographic lineages among populations occurring further east that we named 'Balkan', 'Italian-North European' and 'Central-North European'.

The geographic distribution of stone marten mtDNA haplotypes, together with the divergence time estimates, are congruent with a scenario where several regions potentially acted as glacial refugia during Late Quaternary glaciations, and secondary contact of lineages from different origins recurrently occurred during interglacial periods. Extensive expansion during these warmer intervals would have been facilitated by the good dispersal ability of the species, as inferred in other phylogeographic studies of mesocarnivores (e.g. Ruiz-González et al 2013; Frantz et al 2014). The use of both mtDNA and microsatellites helped to more clearly elucidate the complex phylogeography and history of population structure of *M. foina* in Europe.

Origin and timing of arrival of the stone marten in Europe

The distribution of the earliest fossils of the putative ancestral form of true martens *Martes wenzensis* from the Pliocene, suggest that martens evolved in Central Europe and then spread to the Middle East (reviewed in Hughes 2012). The subsequent speciation process may have occurred somewhere in Eurasia, with a divergence of the *M. foina* lineage from that of the other true martens at about 2.8 Mya (Koepfli et al 2008; Li et al 2014). According to our results, including a greater similarity to the haplotypes found in Iran and China, the 'Balkan' lineage is the most basal among the extant mtDNA diversity in European stone martens. This, in conjunction with the Middle Pleistocene age inferred for the European stone martens, supports an early entrance into the Balkans from the Middle East and subsequent expansion westwards reaching Western Europe. Thus, the findings of this study strongly contradict the prevailing hypothesis, based on the fossil record, that stone martens only colonized Europe following the expansion of the Neolithic farming cultures from the Middle East (Anderson 1970; Sommer & Benecke 2004). Conversely, it is not possible to reconstruct

from the data how far north in Europe *M. foina* extended in the Pleistocene. However, the lack of known subfossils in the northern European lowlands dated to before the Middle Ages (Sommer and Benecke 2004) may suggest that the species was mostly restricted to southern latitudes.

Impact of the Late Quaternary glaciations and persistence in multiple refugia

After the colonization of Europe, an east-west divergence was estimated to have started ≈ 0.55 Mya and complete isolation was inferred to have occurred at approximately 0.45 Mya, which coincides with the Mindel-Elsterian glaciation (Raymo 1997). Indeed, the 0.60-0.42 Mya interval was a period of predominantly cool-temperate climate, with cold peaks at 0.56-0.53 Mya (MIS 14) and 0.48-42 Mya (MIS 12) (Lisiecki and Raymo 2005). Therefore, as found for other European mammals (Hewitt 2000; Feuda et al 2015), the likely causes of this ancient phylogeographic break among European stone martens are the effects of Middle Pleistocene glaciation. The geographic distribution of haplotypes and their genealogical relationships are consistent with a scenario in which, following range fragmentation and contraction induced by climatic worsening ≈ 0.56 Mya, *M. foina* populations in Europe eventually became restricted to the Iberian and Balkan peninsulas. This would explain the observed east-west division and these two regions are known to have been important Pleistocene glacial refugia for many European taxa (Hewitt, 2000). The primary east-west split was apparent in the haplotype network (Fig. 3) and in the Bayesian clustering analyses of the mtDNA (Fig. 4a), but also in microsatellite (Fig. 5a) data sets. Our estimates of how long the stone marten has been present in Europe lend some credibility to *M. foina* remains ascribed to the Pleistocene but affected by stratigraphic uncertainties: in the French Pyrenees (c. 0.57-0.65 Mya, Bonifay 1973), Greece (c. 0.20-0.45 Mya, Tsoukala 1999) and Czech Republic (c. 0.1-0.8 Mya, von Koenigswald & Heinrich 1999).

Eastern lineages

The low to moderate haplotype diversity and low nucleotide diversity in the three identified 'eastern' lineages ('Balkan', 'Central-North European' and 'Italian-North European'; Table 3) are compatible with rapid demographic expansion of small populations (Avise 2000). Actually, results of the mismatch distribution and neutrality tests all supported the hypothesis of past demographic expansion for the three lineages. We propose that the genetic and geographic differentiation observed between these lineages may be the result of range fragmentation, and isolation and divergence in different refugia during Pleistocene glacial peri-

ods. These refugia would have been in the Balkan Peninsula, the Pannonian Basin, and the Italian Peninsula, respectively.

Balkan lineage

Several of our results, including the observed high levels of haplotype and allelic richness, suggest a long presence of *M. foina* in the Balkans. However, the moderate haplotype diversity and low nucleotide diversity of the 'Balkan' lineage are likely to be signatures of past demographic bottlenecks (Grant and Bowen 1998). Subsequent population growth was indicated by the aforementioned mismatch distribution and neutrality tests and also visible in the star-shaped pattern around H5, with widespread range expansion attested by the presence of H5 for instance in Spain and the Netherlands.

Numerous phylogeographic studies on species present in the Balkans have revealed the existence of independent refugia within the peninsula through the Quaternary (e.g. Schmitt et al 2006; Karaiskou et al 2014) thereby supporting the refugia-within-refugia concept (Gómez and Lunt 2006). For instance, Greece, due to its location at the southern tip of the Balkan Peninsula, has been reported to hold genetically differentiated lineages from those elsewhere in the peninsula in different taxa (Stamatis et al 2009; Karaiskou et al 2014). The possibility that stone marten populations in the Balkans may have been isolated in separate refugia, north and south in the peninsula, is raised by the presence of a distinct set of haplotypes (H36, 37, 38, 40, 41; Fig. 3) exclusive to Greece. Persistence in multiple refugia may have contributed to the maintenance of genetic richness in Balkan stone martens. Moreover, radial range expansion from the northern Balkans could have prevented the spread northwards of the haplotypes private to Greece, as described for the meadow grasshopper (Cooper et al 1995; Hewitt 1999). Future studies with increased sampling in the Balkans are warranted to test these hypotheses.

Italian-North European lineage

The 'Italian-North European' lineage, closely related to the 'Balkan', is distinguishable as a geographic expansion represented by H13 from the Balkans into Italy in the Late Pleistocene. We hypothesize that, by drift, it became rare in the Balkans and common in Italy, from where it subsequently expanded into northern Europe through France and Switzerland, eventually reaching Lithuania. The genetic differentiation between the 'Italian-North European' and 'Balkan' lineages was also found in Bayesian clustering analyses of the mtDNA data in BAPS and of the microsatellite data in Geneland.

Central-North European lineage

Among the Eastern lineages, the 'Central-North European' is the best-supported phylogroup (Clade II). Divergence between the 'Central-North European' and 'Balkan' lineages was estimated to have occurred near the end of the Middle Pleistocene (Table 4).

Recently, evidence has been accumulating for the existence of cryptic refugia for temperate species well to the north of the classic southern refugia (i.e. Iberian, Italian and Balkan peninsulas) (Hewitt 1999; Hewitt 2001; Stewart and Lister 2001; Provan and Bennett 2008). Given the distribution of H16, we suggest that the 'Central-North European' lineage is the result of persistence and isolation in cryptic refugia potentially located in the Carpathian region, possibly in the Pannonian Basin. This basin, surrounded by the Carpathian Mountains, the Eastern Alps, the Dinarides and the Balkan Mountains, has been postulated as a glacial refugium for a wide range of vertebrates, from amphibians (*e.g. Rana arvalis*, Babik et al 2004) to small mammals (e.g. *Clethrionomys glareolus*, Kotlík et al 2006; *Mustela nivalis*, McDevitt et al 2012) and large mammals (e.g. *Ursus arctos*, Saarma et al 2007). The haplotype network indicates a radial postglacial expansion of Clade II both into northern Europe, where it is dominant, and to the Balkans in the south. Clade II was consistently identified in the mtDNA analyses, while the microsatellite data revealed a more recent differentiation between central-northern Europe and the Pannonian Basin.

Western clade

After the east-west split, the western clade (Clade I, 'Southwestern European' lineage) remained rather stable over time. Both the phylogenetic tree and the haplotype network show a complex system with several homoplasies but with a clear geographical structure. The presence of many private haplotypes in Iberia, along with the ancient internal subdivision detected in the network and clustering analyses, stressed the important role of the Iberian Peninsula as potential glacial refuge for *M. foina*, as previously proposed for a wide range of species (Hewitt 1999; Hewitt 2004; Sommer and Benecke 2004; Randi 2007; Ruiz-González et al 2013).

During interglacial stages, the Iberian Peninsula, situated in the western edge of the stone marten distribution, had little influence in the postglacial range repopulation of Central Europe, in contrast to the key contribution reported in numerous studies on mammals (*Ursus arctos*, Valdiosera et al 2007; *Martes martes* Ruiz-González et al 2013; *Meles meles*, Frantz et al 2014). A possible explanation is that the Pyrenees acted as a barrier for the stone marten

dispersal, as did in other species (e.g. *Apodemus sylvaticus*, Michaux et al 2003). Equally plausible, by the time the stone martens found their way out through the Pyrenees, France and Central Europe were already occupied by extant stone martens populations, and thus the low rate of immigration from the Iberian refugia had a relative small effect on the composition of the already established populations (e.g. Cooper et al 1995; Hewitt 2000).

The long-term isolation in refugia, like the one in Iberia, make species more prone to evolutionary divergence, usually leading to phylogeographic structuring or speciation (Stewart et al 2010). However, the specific responses to the Quaternary Climate changes ultimately depend on each species adaptations and environmental tolerances (Steward et al 2010). Limited phylogeographic structure in carnivores has been attributed to their high dispersal capabilities that allow relatively rapid gene flow across large distances, erasing the signs of Pleistocene separations (Vilà et al 1999). In this context, the lack of a clear genetic structure for mtDNA in the Iberian Stone marten populations was already detected by Vergara et al, in review).

Some haplotypes within this Western clade displayed intriguing distributions, as the common western European H2 found in samples from Israel, (dominant haplotype is H49), and the presence of H2, H7 and H28 at low rates in the Carpathian belt. In light of this finding, two alternative explanations have to be contemplated. First, as documented for other species of the genus, stone martens were introduced into Luxemburg, Germany, Denmark, Russia, and into several Mediterranean islands (Powell et al 2012), but the existence of additional unreported translocations across its range should not be discarded. Thus, stone marten translocations and/or human mediated introductions could be behind the irregular distribution of H2, H7 and H28 haplotypes. Another hypothesis for this finding is that these regions maintained ancestral populations carrying haplotypes from before the basal east-west split, which either became extinct elsewhere in Europe or might have gone undetected in this study. The contemporary genetic structuring obtained from microsatellites where; i) all samples from Israel belonged to a private cluster or were closely related to Crete or the Balkans and ii) all software converged in restricting the Sourthwestern cluster to Iberia and France, speaks in favor of the stone marten relictic population hypothesis. Still, additional sampling from Near-East populations are needed to shed light on these haplotype distributions.

Contact zones between stone marten lineages

This study provided evidences to believe that the present stone marten populations were originated from individuals from four distinct glacial refugia; the Balkan Peninsula, the Ital-

ian Peninsula and the Pannonian Basin in the Easter, and the Iberian Peninsula in the Western clade. When divergent expanding lineages met, they formed what is called hybrid or contact zones (Hewitt 2000). In Europe, the main contact zones are located in The Alps and Central Europe, and in a minor degree, in the northern Balkans and the Pyrenees (Hewitt 2000). In this study, regardless of when the diversification occurred, the mitochondrial lineages seem to spatially converge in Southern-Central Europe. The highest observed mtDNA diversity found in Switzerland, Slovenia and Serbia suggest that the region could be the main intermixing zone. Allelic richness and expected heterozygosity were highest in Serbia, mainland Greece and Bulgaria, and in Austria, Italy, Switzerland and France. Thus, the contact zone might well be best considered a continuum along the Alps. Mitochondrial BAPS results also identified Denmark as an important mixing zone with Slovenia and Austria as the main overlapping region, in consonance with the output of the same software BAPS from microsatellites.

The stone marten in the island of Crete

Several mammalian species have been studied to identify the genetic link between insular and mainland populations and to answer how and when colonized and dispersed across the different Mediterranean islands (e.g. Dubey et al 2008; Lebarbenchon et al 2010). As generally agreed, the stone marten, as well as the pine marten, were introduced by humans in ancient times in more than 20 islands, mostly in the eastern Mediterranean (Alcover, 1980; Masseti, 1995). Crete has not been connected to the mainland since the Early Pliocene, 5 Myr BP (Dermitzakis 1990). Therefore, colonization through human intervention (intentionally or not) seems to be the prevailing hypothesis. Sondaar et al (1996) suggested that the process of differentiation of the stone marten in Crete was hidden by multiple introductions, possibly intermixed by local extinctions, as the stone marten fossils were found in the Early Neolithic layers, and reappeared in the Minoan period but were absent from the Middle and Late Neolithic material. The strong evidence of recent population growth and the significant number of private haplotypes detected in Crete reinforced the idea of recurrent reintroductions from multiple source locations, such as mainland Greece or Turkey, but also from other Mediterranean Countries.

Conclusions

In this study we provided the first phylogeographic and population genetic results of the stone marten in Europe through the mtDNA sequencing and microsatellite genotyping of a comprehensive sampling in the Western Paleartic. First, our results refuted the prevailing hypothesis, built on fossil evidences, according to which the stone marten entered Europe only during the Holocene and following the Neolithic cultures. In contrast, phylogenetic and dating analyses of the mitochondrial sequences revealed a strong an ancient genetic east-west split that can be dated back to the Middle Pleistocene. The geographic distribution of stone marten mtDNA haplotypes, together with the divergence time estimates, are congruent with a scenario where several regions potentially acted as glacial refugia. Regardless of when the diversification occurred, the four postulated refugia were located not only in the Mediterranean refuges (i.e. the Iberian, Italian and Balkan peninsulas) but also possibly in cryptic northern refugia (i.e. the Pannonian Basin). Further, both genetic markers agreed in the presence of secondary contacts of lineages among individuals from different origins during interglacial periods, facilitated by the high dispersal ability of the stone marten, and by human mediated introductions and translocations.

Acknowledgments

This study has been partially funded by the Basque Government through the Research group on ´Systematics, Biogeography and Population Dynamic` (Ref. IT317-10; GIC10/76). MV (Ref: RBFI-2012-446) and ARG (Ref: DKR-2012-64) were supported by a PhD and post-doctoral fellowships awarded by the Dept. of Education, Universities and Research of the Basque Government. MPB and CF acknowledge financial support from Fundação para a Ciência e Tecnologia through the PhD fellowship SFRH/BD/38410/2007, the Ciência 2007 contract C2007-UL-342-CBA1, and the research project PTDC/BIA-BEC/101511/2008. Finally, we wish to thank all the institutions and people directly involved in the collection of samples, including rangers, vets and field researchers (See Table S1, Collectors).

References

Al Shafee D., Yuosef M, Al Melhim W., Amr Z. (1997) The status of the Stone Marten, *Martes foina syriaca* (Nehring, 1902), in Jordan. Zool Middle East 15:5–8.

Anderson E (1970) Quaternary evolution of the genus *Martes* (Carnivora, Mustelidae). Acta Zool Fenn 130:1–132.

Avise J (2000) Phylogeography: the History and Formation of Species. Harvard University Press, Cambridge, MA.

Avise J, Arnold J, Ball RM, et al (1987) Intraspecific Phylogeography: The Mitochondrial DNA Bridge Between Population Genetics and Systematics. Annu Rev Ecol Syst 18:489–522.

Babik W, Branicki W, Sandera M, et al (2004) Mitochondrial phylogeography of the moor frog, Rana arvalis. Mol Ecol 13:1469–1480.

Bandelt HJ, Forster P, Röhl a (1999) Median-joining networks for inferring intraspecific phylogenies. Mol Biol Evol 16:37–48.

Basto MP, Rodrigues M, Santos-Reis M, et al (2010) Isolation and characterization of 13 tetranucleotide microsatellite loci in the Stone marten (*Martes foina*). Conserv Genet Resour 2:317–319.

Beheregaray LB (2008) Twenty years of phylogeography: The state of the field and the challenges for the Southern Hemisphere. Mol Ecol 17:3754–3774.

Berto C, Rubinato G (2011) The upper Pleistocene mammal record from Caverna degli Orsi (San Dorligo della Valle - Dolina, Trieste, Italy): A faunal complex between eastern and western Europe. Quat Int. doi: 10.1016/j.quaint.2011.07.025

Bhagwat SA, Willis KJ (2008) Species persistence in northerly glacial refugia of Europe: A matter of chance or biogeographical traits? J Biogeogr 35:464–482.

Bonifay M (1973) Principaux gisements paléontologiques français du Pléistocène moyen: essai de classification. Le Quat 41–49.

Cassens I, Van Waerebeek K, Best PB, et al (2003) The phylogeography of dusky dolphins (*Lagenorhynchus obscurus*): A critical examination of network methods and rooting procedures. Mol Ecol 12:1781–1792.

Cheng L, Connor TR, Sirén J, et al (2013) Hierarchical and spatially explicit clustering of DNA sequences with BAPS software. Mol Biol Evol 30:1224–1228.

Consuegra S, García De Leániz C, Serdio A, et al (2002) Mitochondrial DNA variation in Pleistocene and modern Atlantic salmon from the Iberian glacial refugium. Mol Ecol 11:2037–2048.

Cooper S, Ibrahim K, Hewitt G (1995) Postglacial expansion and genome subdivision in the European grasshopper Chorthippus parallelus. Mol Ecol 4:49–60.

Corander J, Marttinen P, Sirén J, Tang J (2008) Enhanced Bayesian modelling in BAPS software for learning genetic structures of populations. BMC Bioinformatics 9:539.

Corander J, Tang J (2007) Bayesian analysis of population structure based on linked molecular information. Math Biosci 205:19–31.

Dallas JF, Bacon PJ, Carss DN, et al (1999) Genetic diversity in the Eurasian otter, Lutra lutra, in Scotland. Evidence from microsatelliye polymorphism. Biol J Linn Soc 73–86.

Delibes M (1983) Interspecific competition and the habitat of the stone marten *Martes foina* (Erxleben 1777) in Europe. Acta Zool Fenn 174:229–231.

Dermitzakis M (1990) Paleogeography, geodynamic processes and event stratigraphy during the Late Cenozoic of the Aegean area. Atti Convegni Lincei 85:263– 288.

Desper R, Gascuel O (2002) Fast and accurate phylogeny minimum-evolution principle. J Comput Biol 9:687–705.

Drummond AJ, Suchard M a., Xie D, Rambaut A (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol Biol Evol 29:1969–1973.

Dubey S, Koyasu K, Parapanov R, et al (2008) Molecular phylogenetics reveals Messinian, Pliocene, and Pleistocene colonizations of islands by North African shrews. Mol Phylogenet Evol 47:877–882.

Earl DA, vonHoldt BM (2012) Structure harvester: A website and program for visualizing Structure output and implementing the Evanno method. Conserv Genet Resour 4:359–361.

Edwards CJ, Soulsbury CD, Statham MJ, et al (2012) Temporal genetic variation of the red fox, Vulpes vulpes, across western Europe and the British Isles. Quat Sci Rev 57:95–104.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 14:2611–20.

Excoffier L (2004) Patterns of DNA sequence diversity and genetic structure after a range expansion: lessons from the infinite-island model. Mol Ecol 13:853–864.

Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Resour 10:564–7.

Excoffier L, Smouse P, Quattro J (1992) Analysis of molecular variance infered from metric distances among DNA haplotypes: application to human mitochondrial DNA restricyion data. Genetics 131:479–491.

Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics 164:1567–87.

Felsenstein J (1992) Estimating effective population size from samples of sequences: inefficiency of pairwise and segregating sites as compared to phylogenetic estimates. Genet Reseach 59:139–147.

Fernandes CA, Ginja C, Pereira I, et al (2008) Species-specific mitochondrial DNA markers for identification of non-invasive samples from sympatric carnivores in the Iberian Peninsula. Conserv Genet 9:681–690.

Feuda R, Bannikova AA, Zemlemerova ED, et al (2015) Tracing the evolutionary history of the mole, Talpa europaea, through mitochondrial DNA phylogeography and species distribution modelling. Biol J Linn Soc 114:495–512.

Frantz A, McDevitt A, Pope L, et al (2014) Revisiting the phylogeography and demography of European badgers (*Meles meles*) based on broad sampling, multiple markers and simulations. Heredity 1–11.

Fu YX (1997) Statistical Tests of Neutrality of Mutations Against Population Growth, Hitchhiking and Background Selection. Genetics 147:915–925.

Gascuel O (1997) BIONJ: an improved version of the NJ algorithm based on a simple model of sequence data. Mol Biol Evol 14:685–695.

Ghezzo E, Palchetti A, Rook L (2014) Recovering data from historical collections: Stratigraphic and spatial reconstruction of the outstanding carnivoran record from the Late Pleistocene Equi cave (Apuane Alps, Italy). Quat Sci Rev 96:168–179.

Golubchik T, Wise MJ, Easteal S, Jermiin LS (2007) Mind the gaps: Evidence of bias in estimates of multiple sequence alignments. Mol Biol Evol 24:2433–2442.

Gómez A, Lunt DH (2006) Refugia within refugia: patterns of phylogeographic concordance in the Iberian Peninsula. In: Weiss S, Ferrand N (eds). Phylogeography of Southern European Refugia. Springer: The Netherlands. pp 155–188.

Grant WS, Bowen BW (1998) Shallow population histories in deep evolutionary lineages of marine fishes: Insights from sardines and anchovies and lessons for conservation. J Hered 89:415–426.

Guillot G, Mortier F, Estoup A (2005) Geneland: a computer package for landscape genetics. Mol Ecol Notes 5:712–715.

Hanke M, Wink M (1994) Direct DNA sequencing of PCR-amplified vector inserts following enzymatic degradation of primer and dNTPs. Biotechniques 17:858–860.

Herr J, Schley L, Roper TJ (2009) Socio-spatial organization of urban stone martens. J Zool 277:54–62.

Hewitt G (2000) The genetic legacy of the Quaternary ice ages. Nature 405:907–913.

Hewitt G (2001) Speciation, hybrid zones and phylogeography - or seeing genes in space and time. Mol Ecol 537–549.

Hewitt GM (1996) Some genetic consequences of ice ages, and their role, in divergence and speciation. Biol J Linn Soc 58:247–276.

Hewitt GM (2004) The structure of biodiversity – insights from molecular phylogeography. Front Zool 1:4.

Hewitt GM (1999) Post-glacial re-colonization of European biota. Biol J Linn Soc 68:87–112.

Hofreiter M, Serre D, Rohland N, et al (2004) Lack of phylogeography in European mammals before the last glaciation. Proc Natl Acad Sci U S A 101:12963–12968.

Holder MT, Sukumaran J, Lewis PO (2008) A justification for reporting the majority-rule consensus tree in Bayesian phylogenetics. Syst Biol 57:814–821.

Huelsenbeck JP, Larget B, Alfaro ME (2004) Bayesian phylogenetic model selection using reversible jump Markov chain Monte Carlo. Mol Biol Evol 21:1123–1133.

Hughes S (2012) Synthesis of Martes evolutionary history. In: Aubry K, Zielinski W, Raphael M, Proulx G, Buskirk S (ed) Biology and conservation of marten, sables, and fisher: a new synthesis. New York: Cornell University Press. pp. 3–22.

Jaarola M, Tegelström H (1995) Colonization history of north European field voles (*Microtus agrestis*) revealed by mitochondrial DNA. Mol Ecol 4:299–310.

Jones K, Bielby J, Cardillo M, et al (2009) PanTHERIA: a species-level database of life history, ecology, and geography of extant and recently extinct mammals: Ecological Archives E090-184. Ecology 90:2648–2648.

Kalinowski ST (2004) Counting alleles with rarefactions: private alleles and hierachical sampling design. Conserv Genet 5:539–543.

Kalinowski ST (2005) HP-RARE 1.0: A computer program for performing rarefaction on measures of allelic richness. Mol Ecol Notes 5:187–189.

Karaiskou N, Tsakogiannis A, Gkagkavouzis K, et al (2014) Greece: a Balkan subrefuge for a remnant red deer (*Cervus elaphus*) population. J Hered 105:334–44.

Koepfli K-P, Deere KA, Slater GJ, et al (2008) Multigene phylogeny of the Mustelidae: resolving relationships, tempo and biogeographic history of a mammalian adaptive radiation. BMC Biol 6:22.

Kotlík P, Deffontaine V, Mascheretti S, et al (2006) A northern glacial refugium for bank voles (*Clethrionomys glareolus*). Proc Natl Acad Sci U S A 103:14860–14864.

Lebarbenchon C, Poitevin F, Arnal V, Montgelard C (2010) Phylogeography of the weasel (*Mustela nivalis*) in the western-Palaearctic region: combined effects of glacial events and human movements. Heredity 105:449–62.

Li B, Wolsan M, Wu D, et al (2014) Molecular Phylogenetics and Evolution Mitochondrial genomes reveal the pattern and timing of marten (Martes), wolverine (Gulo), and fisher (Pekania) diversification. 80:156–164.

Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25:1451–2.

Lisiecki LE, Raymo ME (2005) A Pliocene-Pleistocene stack of 57 globally distributed benthic of D18 records. Paleoceanography 20:1–17.

Llorente L, Cristina R, Careta M (2011) Earliest occurrence of the Beech marten (Martes foina Erxleben , 1777) in the Iberian Peninsula. In: Brugal J., Gardeisen A, Zucker A (eds) Prédateurs dans tour leurs états. Évoultion, Biodiversité, Interact. Myth. Symb. XXXI rencontres Int. d´archéologie d´histoire d´Antibes, Éditions A. Antibes, pp 189–209.

Malyarchuk B, Derenko M, Denisova G (2014) A mitogenomic phylogeny and genetic history of sable (*Martes zibellina*). Gene 550:56–67.

Masseti M (1995) Quaternary biogepgraphy of the mustelidae family on the Mediterranean islands. Hystrix 7:17–34.

McDevitt AD, Zub K, Kawałko A, et al (2012) Climate and refugial origin influence the mitochondrial lineage distribution of weasels (*Mustela nivalis*) in a phylogeographic suture zone. Biol J Linn Soc 106:57–69.

Merila J, Bjorklund M, BAker A. (1997) Historical demography and present day population structure of the Greenfinch, *Carduelis chloris*-An analysis of mtDNA Control-Region sequences. Evolution 51:945–956.

Michaux JR, Magnanou E, Paradis E, et al (2003) Mitochondrial phylogeography of the woodmouse (*Apodemus sylvaticus*) in the Western Palearctic region. Mol Ecol 12:685–697.

Miracle PT, Mauch Lenardić J, Brajković D (2010) Last glacial climates, "Refugia", and faunal change in Southeastern Europe: Mammalian assemblages from Veternica, Velika pećina, and Vindija caves (Croatia). Quat Int 212:137–148.

Mucci N, Arrendal J, Ansorge H, et al (2010) Genetic diversity and landscape genetic structure of otter (*Lutra lutra*) populations in Europe. Conserv Genet 11:583–599.

Nagai T, Raichev E, Tsunoda H (2012) Preliminary study on microsatellite and mitochondrial DNA variation of the stone marten *Martes foina* in Bulgaria. Mammal Study 358:353–358.

Nei M (1987) Molecular Evolutionary Genetics. Columbia University Press.

Nei M, Kumar S (2000) Molecular evolution and phylogenetics. Oxford University Press, New York. 333 pp.

Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. Bioinformatics 28:2537–9.

Petit RJ, El Mousadik A, Pons O (1998) Identifying Populations for Conservation on the Basis of Genetic Markers. Conserv Biol 12:844–855.

Pilot M, Branicki W, Jedrzejewski W, et al (2010) Phylogeographic history of grey wolves in Europe. BMC Evol Biol 10:104.

Posada D, Crandall K a (2001) Intraspecific phylogenetics: Trees grafting into networks. Trends Ecol Evol 16:37–45.

Powell RA, Lewis J., Slough B., et al (2012) Evaluating translocations of martens, sables, and fishers: Testing model predictions with field data. In: Aubry K, Zielinski W, Raphael M, Proulx G, Buskirk S (ed) Biology and conservation of marten, sables, and fisher: a new synthesis. New York: Cornell University Press, pp. 93–127.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–59.

Provan J, Bennett KD (2008) Phylogeographic insights into cryptic glacial refugia. Trends Ecol Evol 23:564–571.

Rambaut A, Suchard M, Xie D, Drummond A (2014) Tracer v1.6. Available from http://beast.bio.ed.ac.uk/Tracer

Ramos-Onsins SE, Rozas J (2002) Statistical properties of new neutrality tests against population growth. Mol Biol Evol 19:2092–2100.

Randi E (2007) Phylogeography of South European mammals. In: Weiss S, Ferrand N (eds). Phylogeography of Southern European Refugia. Springer: The Netherlands. pp 101–126.

Ray N, Currat M, Excoffier L (2003) Intra-deme molecular diversity in spatially expanding populations. Mol Biol Evol 20:76–86.

Raymo M. (1997) The timing of major climate terminations. Paleoceanography 12:577–585.

Reig S (1992) Geographic variation in pine marten (*Martes martes*) and beech marten (*M. foina*) in Europe. J Mammal 73:744–769.

Rogers A, Harpending H (1992) Population growth makes waves in the distribution of pairwise genetic differences. Mol Biol Evol 9:552–569.

Ronquist F, Teslenko M, Van Der Mark P, et al (2012) Mrbayes 3.2: Efficient bayesian phylogenetic inference and model choice across a large model space. Syst Biol 61:539–542.

Ruiz-González A, Madeira MJ, Randi E, et al (2013) Phylogeography of the forest-dwelling European pine marten (Martes martes): new insights into cryptic northern glacial refugia. Biol J Linn Soc 109:1–18.

Saarma U, Ho S, Pybus O, et al (2007) Mitogenetic structure of brown bears (*Ursus arctos* L.) in north-eastern Europe and a new time frame for the formation of European brown bear lineages. Mol Ecol 16:401–413.

Sacchi O, Meriggi A (1995) Habitat Requirements of the Stone Marten (*Martes Foina*) on the Tyrrhenian Slopes of the Northern Apennines. Hystrix 7:99–104.

Santos MJ, Santos-Reis M (2010) Stone marten (*Martes foina*) habitat in a Mediterranean ecosystem: effects of scale, sex, and interspecific interactions. Eur J Wildl Res 56:275–286.

Schmitt T, Habel JC, Zimmermann M, Müller P (2006) Genetic differentiation of the marbled white butterfly, *Melanargia galathea*, accounts for glacial distribution patterns and postglacial range expansion in southeastern Europe. Mol Ecol 15:1889–1901.

Schmitt T, Varga Z (2012) Extra-Mediterranean refugia: The rule and not the exception? Front Zool 9:22.

Schneider S, Excoffier L (1999) Estimation of demographic parameters from the distribution of pairwise differenced when the mutation rates vary among sites: application to human mitochondrial DNA. Genetics 153:1079–1089.

Seutin G, White B., Boag P. (1990) Oreservation of avian bblood and tissue samples for DNA analyses. Can J Zool 69:82–90.

Slatkin M, Hudson R (1991) Pairwise Comparisons of Mitochondrial. Evolution 562:555–562.

Sommer R, Benecke N (2004) Late- and Post-Glacial history of the Mustelidae in Europe. Mamm Rev 34:249–284.

Sondaar P, Dermitzakis M, de Vos J (1996) The paleogeography and faunal evolution of the land mammals of Crete. In: Reese D (ed) Pleistocene Holocene fauna Crete its first settlers. World Archaeol. No. 28. Madison Prehistory Press, pp 61–67.

Stamatakis A (2006) RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22:2688–2690.

Stamatis C, Suchentrunk F, Moutou K, et al (2009) Phylogeography of the brown hare (*Lepus europaeus*) in Europe: a legacy of south-Eastern Mediterranean refugia? J Biogeogr 36:515–528.

Stewart JR, Lister AM (2001) Cryptic northern refugia and the origins of the modern biota. Trends Ecol Evol 16:608–613.

Stewart JR, Lister AM, Barnes I, Dalén L (2010) Refugia revisited: individualistic responses of species in space and time. Proc Biol Sci 277:661–671.

Stöver BC, Müller KF (2010) TreeGraph 2: combining and visualizing evidence from different phylogenetic analyses. BMC Bioinformatics 11:7. Taberlet P, Fumagalli L, Wust-Saucy A-G, Cossons J-F (1998) Comparative phylogeography and postglacial colonization. Mol Ecol 7:453–464.

Tajima F (1989) Statistical Method for Testing the Neutral Mutation Hypothesis by DNA Polymorphism. Genetics 123:585–595.

Tajima F (1983) Evolutionary Relationship of DNA Sequences in Finite Populations. Genetics 105:437–460.

Tamura K, Nei M (1993) Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol Biol Evol 10:512–526.

Tamura K, Nei M, Kumar S (2004) Prospects for inferring very large phylogenies by using the neighbor-joining method. Proc Natl Acad Sci USA 101:11030–11035.

Tamura K, Stecher G, Peterson D, et al (2013) MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol Biol Evol 30:2725–9.

Thompson JD, Linard B, Lecompte O, Poch O (2011) A comprehensive benchmark study of multiple sequence alignment methods: Current challenges and future perspectives. PLoS One. doi: 10.1371/journal.pone.0018093.

Tikhonov A, Cavallini P, Maran T, et al (2008) *Martes foina*. The IUCN Red List of Threatened Species. Version 2014.3.

Tsoukala E (1999) Quaternary large mammals from the Apidima Caves (Lakonia, S. Peloponnese, greece). Beitraege zur Palaeontol 24:207–229.

Valdiosera C, García N, Anderung C, et al (2007) Staying out in the cold: glacial refugia and mitochondrial DNA phylogeography in ancient European brown bears. Mol Ecol 16:5140–5148.

Valière N (2002) GIMLET: a computer program for analysing genetic individual identification data. Mol Ecol Notes 377–379.

Vilà C, Amorim I, Leonard J, et al (1999) Mitochondrial DNA phylogeography and population history of the grey wolf Canis lupus. Mol Ecol 8:2089–2103.

Villesen P (2007) FaBox: an online toolbox for fasta sequences. Mol Ecol Notes 7:965–968.

Von Koenigswald W, Heinrich W (1999) Mittelpleistozäne Säugetierfaunen aus Mitteleuropa - der Versuch einer biostratigraphischen Zuordnung. Kaupia 9: 53-112.

Wallace IM, O'Sullivan O, Higgins DG, Notredame C (2006) M-Coffee: Combining multiple sequence alignment methods with T-Coffee. Nucleic Acids Res 34:1692–1699.

Xia X (2013) DAMBE5: A comprehensive software package for data analysis in molecular biology and evolution. Mol Biol Evol 30:1720–1728.

Xia X, Xie Z, Salemi M, et al (2002) An index of sustitution saturation and its application. Mol Phylogenet Evol 26:1–7.

Supporting Information

S1 Table. Information on the samples used in this study, including ID number, internal sample codes, country of origin, geographic coordinates, genetic markers analysed, haplotype code and microsatellite genotypes.

The table is too big to print in less than 15 pages. Thus, before the papers are published the table will be available online (dropbox folder) in the following link:

https://www.dropbox.com/sh/y4lw39hvnfrepow/AAD72y908pLs-3Zp2G4YUrCGa?dl=0

Algorithm	Iterations	Models	Tested values of K, number of runs
STRUCTURE	burn-in - 10^5 iterations - 10^6	Admixture and correlated allele frequencies (no prior population info)	K = 1 to 20
BAPS	Not applicable	Clustering at the individual level Spatial and non-spatial models	K = 1,5,10,15,20 10 runs
GENELAND	500,000 – to identify K 200,000 – with identified K 1,000 - thinning Posterior probability of population membership with burn-in of 1,000	Spatial D-model; Uncorrelated; Uncertainty spatial cords – 1,000; Maximum rate of Poisson process = number of samples; Maximum number of nuclei in Poisson-Voronoi tessellation = 3 x sample size; Null alleles – False; Spatial model – True	Priors on K - uniform 1 to 20

S2 Table. Summary of the models and settings used in the Bayesian clustering algorithms for K estimation

S3 Table. Results of mismatch distribution tests of demographic expansion and of Tajima (1989) and Fu (1997) tests for country samples containing more than 10 individuals (in the case of Greece, the mainland and Crete were considered separately).

The table is too big to print in less than 15 pages. Thus, before the papers are published the table will be available online (dropbox folder) in the following link:

https://www.dropbox.com/sh/y4lw39hvnfrepow/AAD72y908pLs-3Zp2G4YUrCGa?dl=0

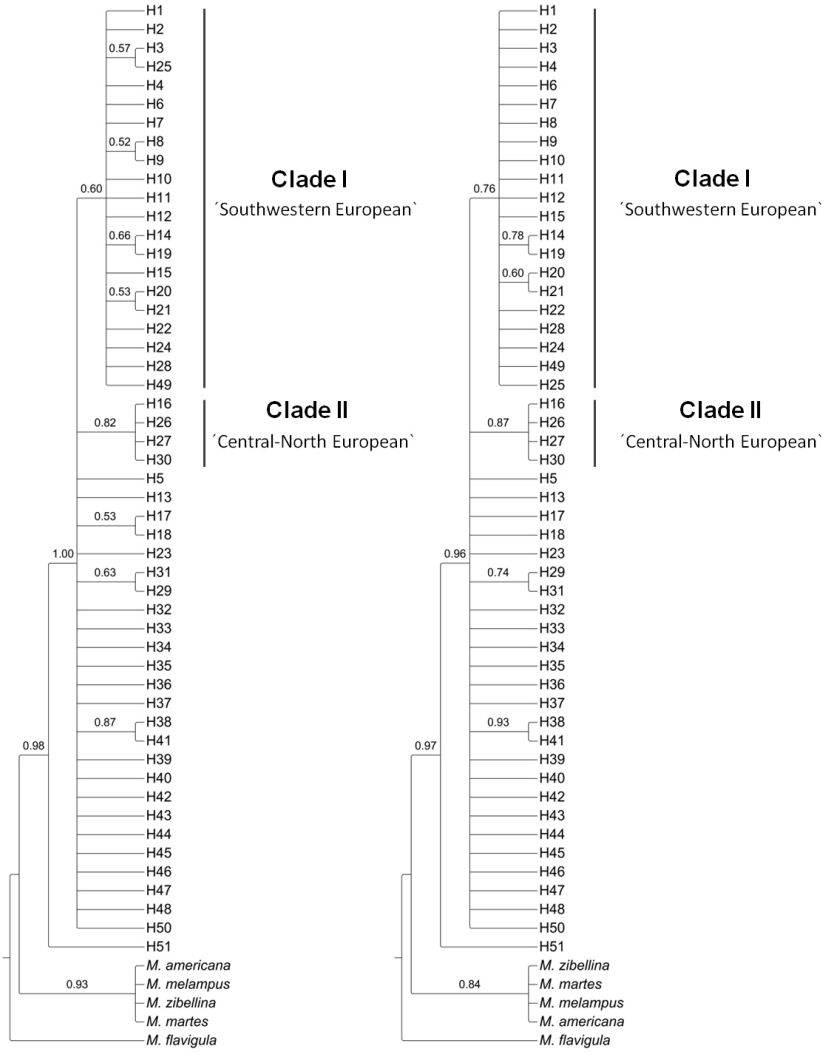


Figure S1. Majority-rule consensus trees of the minimum evolution (left) and maximum parsimony (right) analyses of mtDNA haplotypes. Numbers above branches are bootstrap values. Haplotype codes are as in Table S.1. Clades were named (I and II) for discussion purposes.

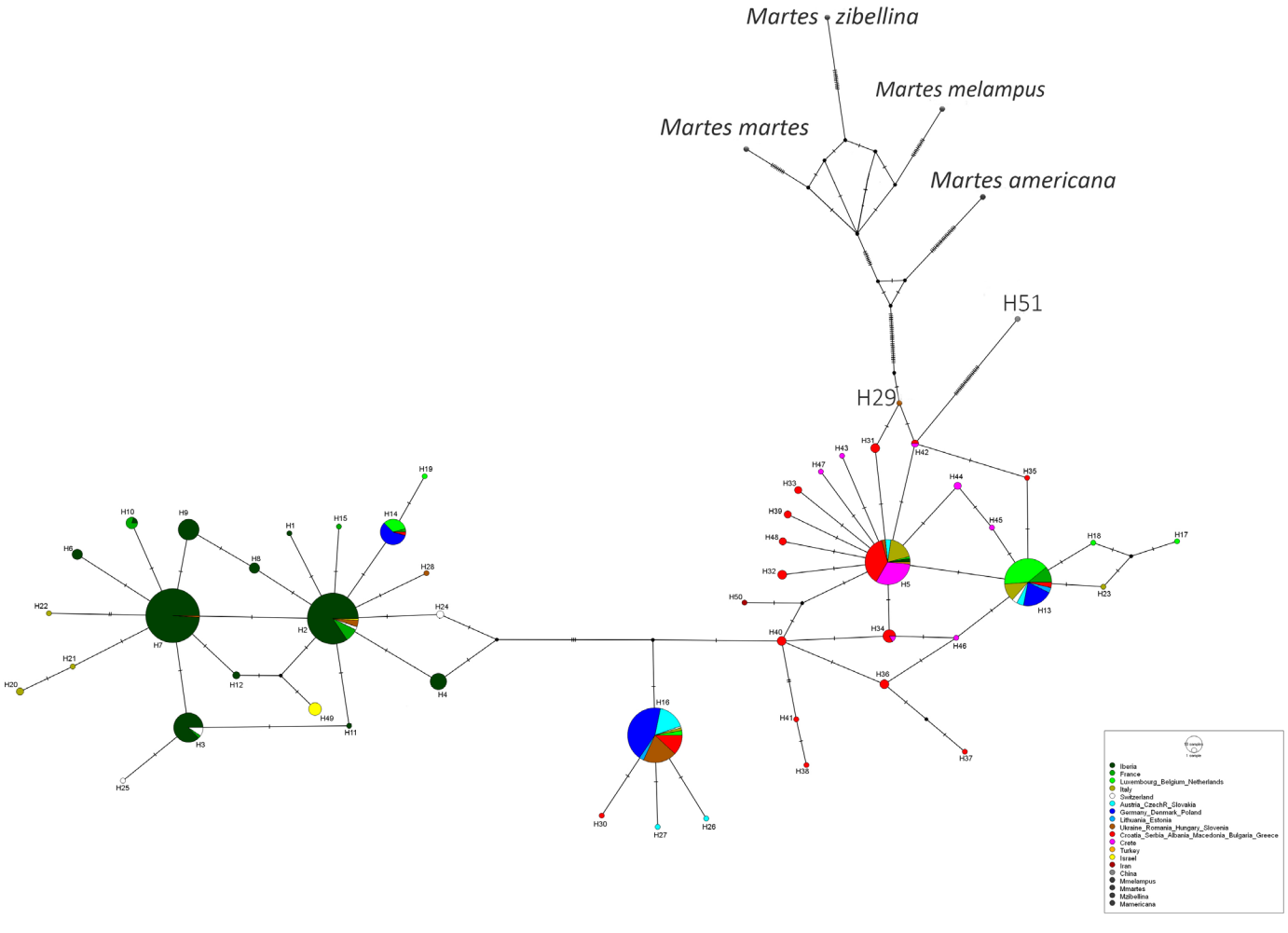


Figure S2. Median-joining network of mtDNA haplotypes rooted with the sister group of *M. foina* (*M. martes* + *M. zibellina* + *M. americana* + *M. melampus*). Circles represent haplotypes and their size is proportional to the frequency observed. Coloured pie charts indicate the frequency of haplotypes in countries or regions. Small black circles represent hypothetical haplotypes. Dashes on lines connecting haplotypes represent the number of nucleotide substitutions separating them. Haplotype codes are given in Table S.1.

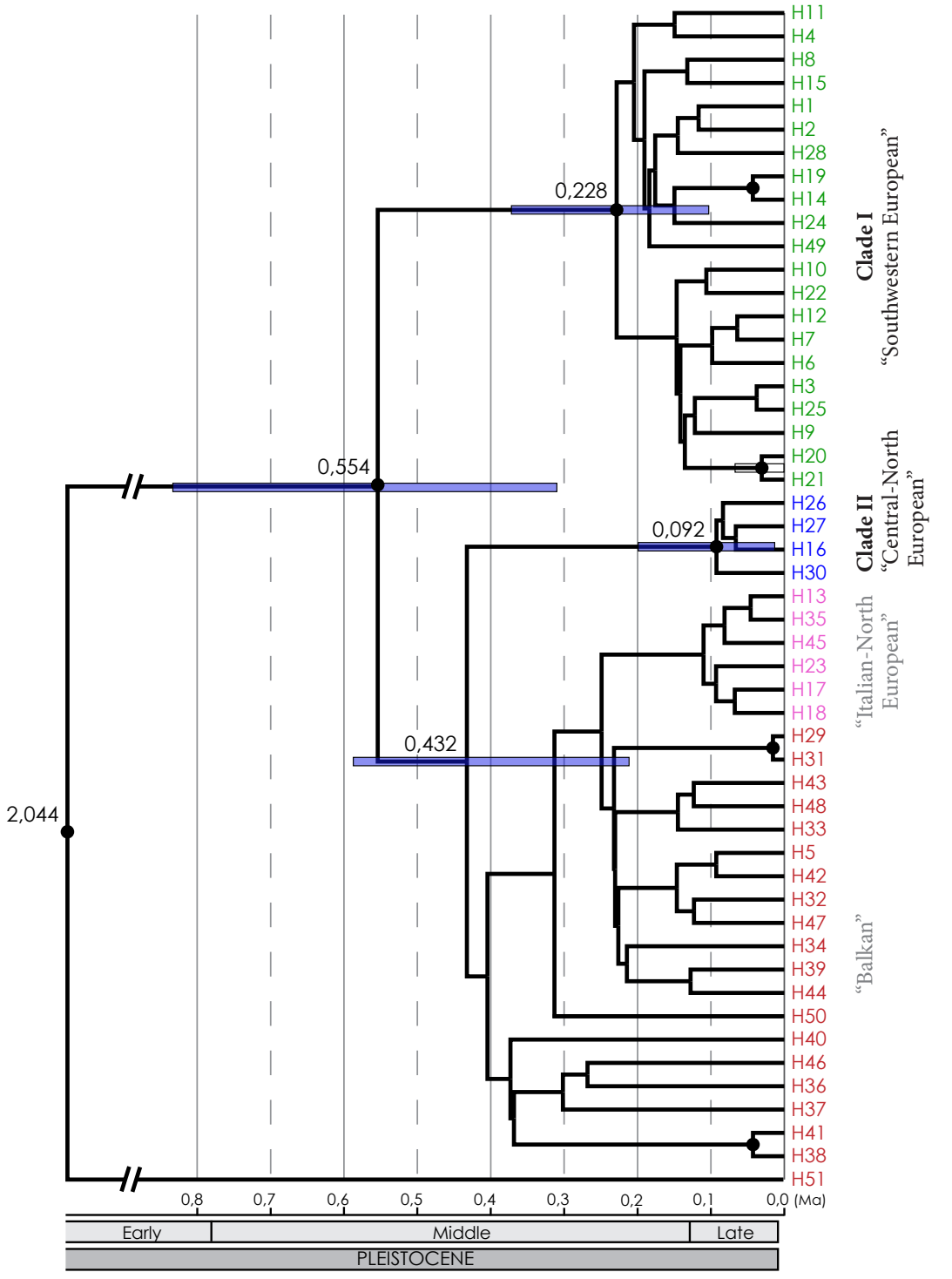


Figure S3. Maximum clade credibility tree with divergence times generated by the BEAST analysis for mitochondrial genes analyzed. Node numbers indicate estimated mean ages and bars correspond to the 95% highest posterior density intervals. Nodes supported with posterior probability values higher than 0.95 are marked with black circles.

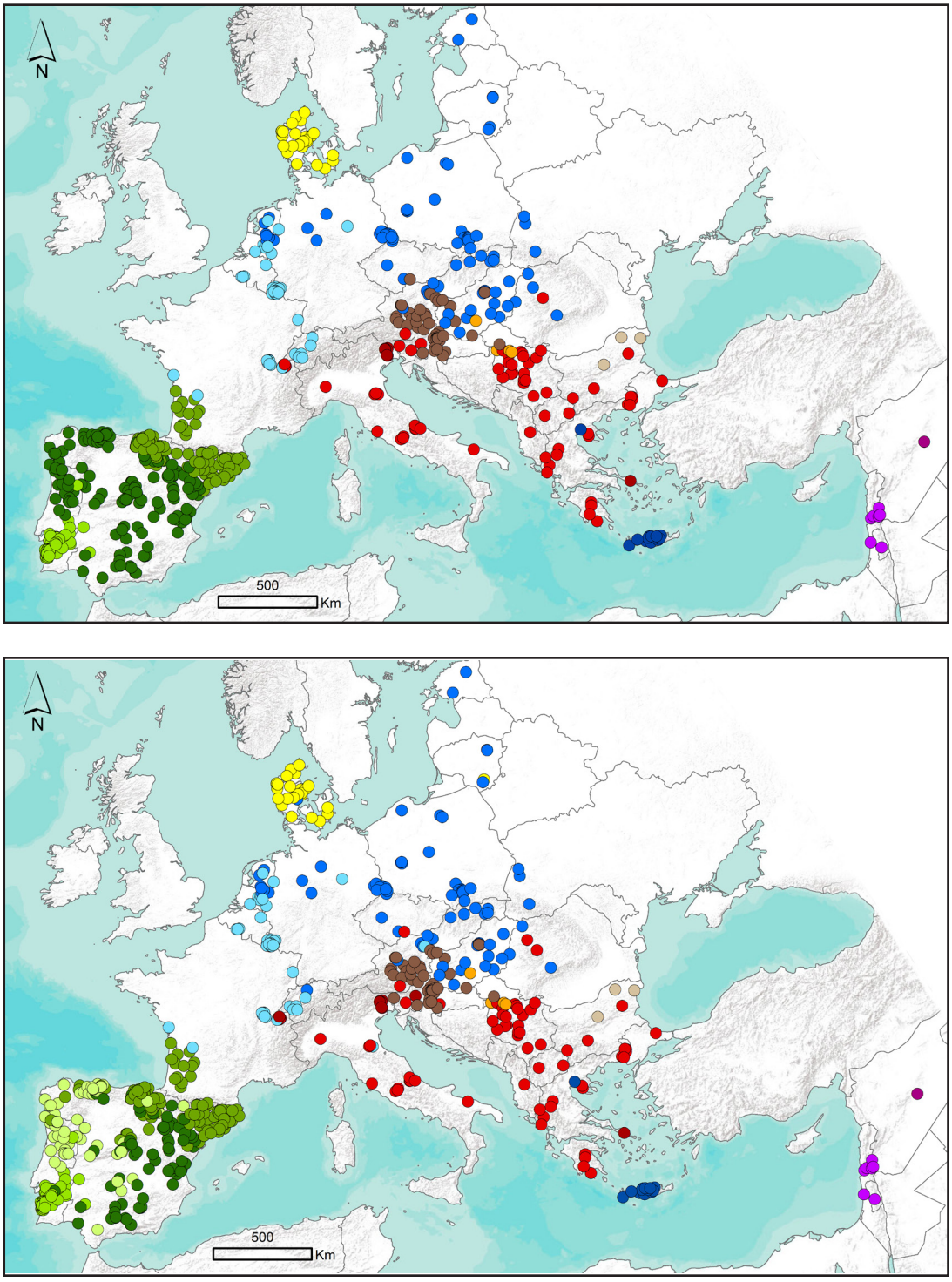


Figure S4. BAPS inferences of the geographic structure of microsatellite variation in European stone martens. Circles represent samples and different colours represent clusters in the optimal partition. a) 'spatial clustering of individuals' model (K=16); b) 'non-spatial clustering of individuals' model (K=18). Compare results with that from GENELAND with varying K (Fig. 5b).

Chapter IV

Concluding remarks

Concluding remarks

From the studies performed in this PhD thesis the following conclusions can be drawn:

A. Non-invasive genetic sampling has proven to be an effective approach to cover a number of research topics in elusive European mustelids. The specific identification from spraints, used to define the otter distribution, led to the detection of a contemporary expansion of this species in northern Spain (Basque Country). Besides, the estimation of the number of individuals inferred through sex and individual characterization, provided essential information to guide short, medium and long-term conservation programs of this endangered species, as contemplated in the only prevailing management plan for the otter in the Basque Country (Álava).

B. The combination of pine and stone marten presence records from different sources (non-invasive genetic sampling and unequivocal species detections from field data) was crucial to obtain a comprehensive high-resolution presence dataset for habitat suitability modeling. Multiscale habitat suitability models (HSM) showed higher discrimination ability than the equivalent single scale models, supporting the scale dependence of the pine and stone marten's habitat selection. The incorporation of sampling bias corrections considerably improved the predictive performance of the HSM, and their adequacy to the known distribution of both species. Further, the comparison between the best performing HSM for each marten led to the detection of a significant niche divergence between the closely-related sympatric mustelids. The pine marten was positively associated with cooler areas, with a small degree of human disturbance, a high proportion of natural forests, well-connected forestry plantations and medium-extent of agroforestry mosaics. On the other hand, the stone marten was conditioned by the density of urban areas, the proportion and extensiveness of croplands, the existence of some scrub cover, and the availability of semi-continuous grasslands. Moreover, this research provided a useful methodological framework for future multispecies and multiscale comparatives.

C. The research presented on the stone marten intraspecific phylogeny, phylogeography and population and landscape genetics has greatly enhanced our knowledge of one of the least studied martens. Results revealed high levels of haplotype diversity and a major phylogeographic east-west split in Europe, which was dated to the Middle Pleistocene (0.55 Mya). Based on mtDNA and microsatellite results, four refugia were postulated for the Western Palaearctic: three traditional Mediterranean refuges located in the Iberian, Italian and Bal-

kan peninsulas and an additional cryptic northern refugia potentially located in the Pannonian Basin (Carpathians). During the recurring interglacial periods of the Quaternary, the secondary contacts of populations from distinct origins would have been facilitated by the high dispersal ability of the stone marten and/or human mediated introductions/translocations. The early divergence estimates categorically contradict the prevailing hypothesis, based on previous fossil evidence, according to which the stone marten´s first entrance and posterior dispersal followed the Neolithic cultures from the Near East in the Holocene. Additionally, the high-resolution genetic study conducted within the Iberian Peninsula, led to the detection of a marked NE-SW cline in genetic diversity, a strong and pervasive isolation by distance pattern, and genetic breaks corresponding, largely, to the presence of the three main rivers in Iberia (i.e. Ebro, Tagus and Guadiana). Overall, the use of both mtDNA and microsatellites markers together with new fossil evidence helped to elucidate the complex phylogeography and history of population structure of *M. foina* both within Iberia and across Europe.