

**DEL INDIVIDUO A LA ESPECIE:  
FILOGEOGRAFÍA Y GENÉTICA DEL  
PAISAJE DE LA RATA DE AGUA  
(*Arvicola sapidus*)**

**TESIS DOCTORAL**



**Alejandro Centeno-Cuadros**  
Estación Biológica de Doñana (CSIC)  
(2009)

Editor: Editorial de la Universidad de Granada  
Autor: Alejandro Centeno Cuadros  
D.L.: GR. 3115-2009  
ISBN: 978-84-692-5104-1



CONSEJO SUPERIOR  
DE INVESTIGACIONES CIENTÍFICAS  
Estación Biológica de Doñana



UNIVERSIDAD DE GRANADA  
Comisión de Doctorado

TESIS DOCTORAL

**DEL INDIVIDUO A LA ESPECIE: FILOGEOGRAFÍA Y GENÉTICA DEL  
PAISAJE DE LA RATA DE AGUA (*Arvicola sapidus*)**

Memoria presentada por **Alejandro Centeno Cuadros** para optar al grado de  
Doctor en Biología por la Universidad de Granada

Vº Bº de los directores

**Dr. José Antonio Godoy López**  
Investigador del Consejo Superior de  
Investigaciones Científicas en la  
Estación Biológica de Doñana

**Dr. Miguel Delibes de Castro**  
Profesor de investigación del Consejo  
Superior de Investigaciones Científicas  
en la Estación Biológica de Doñana

Granada, 10 de Junio de 2009



# ÍNDICE

<b>PRÓLOGO</b>	<b>2</b>
<b>INTRODUCCIÓN GENERAL</b>	<b>6</b>
¿POR QUÉ UNA GENÉTICA DE LA CONSERVACIÓN?	7
<b>INTERACCIÓN DE LA GENÉTICA CON OTRAS DISCIPLINAS</b>	<b>13</b>
Genética en auxilio de la taxonomía: filogenias	14
La genética y el estudio de la variabilidad intraespecífica en el espacio y tiempo	17
Genética y Ecología: Ecología Molecular	21
<b>OBJETIVOS Y ORGANIZACIÓN DE ESTA TESIS DOCTORAL</b>	<b>28</b>
<b>BIBLIOGRAFÍA</b>	<b>33</b>
<b>CHAPTER 1: STRUCTURE, ORGANIZATION AND NUCLEOTIDE DIVERSITY OF THE MITOCHONDRIAL CONTROL REGION OF WATER VOLES IN EURASIA (<i>ARVICOLA</i> SPP.)</b>	<b>38</b>
<b>RESUMEN</b>	<b>42</b>
<b>ABSTRACT</b>	<b>44</b>
<b>ACKNOWLEDGMENTS</b>	<b>52</b>
<b>REFERENCES</b>	<b>52</b>
<b>CHAPTER 2: DATING THE DIVERGENCE BETWEEN SOUTHERN AND EUROPEAN WATER VOLES USING MOLECULAR COALESCENT-BASED METHODS</b>	<b>54</b>
<b>RESUMEN</b>	<b>58</b>
<b>ABSTRACT</b>	<b>60</b>
<b>INTRODUCTION</b>	<b>62</b>
<b>MATERIAL AND METHODS</b>	<b>65</b>
Samples and sequence data	65
The IM model	67
<b>RESULTS</b>	<b>69</b>
<b>DISCUSSION</b>	<b>71</b>
Hybridization or gene flow after split?	71
Time of divergence	71
Effective and census population sizes	72
Paleontological and genetic estimates of divergence: a brief history of <i>Arvicola</i>	73
<b>ACKNOWLEDGMENTS</b>	<b>74</b>

<b>REFERENCES</b>	<b>75</b>
<b>CHAPTER 3: PHYLOGEOGRAPHY OF SOUTHERN WATER VOLE (<i>ARVICOLA SAPIDUS</i>): EVIDENCE FOR REFUGIA WITHIN THE IBERIAN GLACIAL REFUGIUM?</b>	<b>78</b>
<b>RESUMEN</b>	<b>82</b>
<b>ABSTRACT</b>	<b>84</b>
<b>INTRODUCTION</b>	<b>86</b>
<b>MATERIAL AND METHODS</b>	<b>89</b>
Sampling and DNA extraction	89
Mitochondrial DNA sequencing	91
Genetic data analyses	92
Genetic diversity	92
Phylogeny	93
Phylogeography	94
Genetic structure	95
Demography	96
<b>RESULTS</b>	<b>97</b>
Genetic diversity	97
Phylogeny	100
Phylogeography	102
Genetic structure	105
Demography	106
<b>DISCUSSION</b>	<b>108</b>
Evolutionary history of <i>A. sapidus</i>	108
Implications for taxonomy and conservation	112
<b>ACKNOWLEDGMENTS</b>	<b>112</b>
<b>REFERENCES</b>	<b>114</b>
<b>CHAPTER 4: THE ROLE OF LANDSCAPE ON GENETIC STRUCTURE AND DISPERSAL ON SOUTHERN WATER VOLE (<i>ARVICOLA SAPIDUS</i>)</b>	<b>118</b>
<b>RESUMEN</b>	<b>122</b>
<b>ABSTRACT</b>	<b>124</b>
<b>INTRODUCTION</b>	<b>126</b>
<b>MATERIAL AND METHODS</b>	<b>130</b>
Study area and sample collection	130
Genotyping and sequencing	132
Genetic data analysis	134
Mitochondrial data	134
Microsatellite data	135
Isolation by distance and landscape genetics	136
Sex-biased dispersal	138
Contemporary gene flow	139
<b>RESULTS</b>	<b>139</b>
Mitochondrial variation	139

Microsatellite diversity and structure	141
Isolation by distance and landscape genetics	145
Contemporary gene flow	147
<b>DISCUSSION</b>	<b>149</b>
Contemporary gene flow in Doñana	151
Effects of landscape on genetic structure	152
Dispersal	155
<b>REFERENCES</b>	<b>157</b>
<b>DISCUSIÓN GENERAL</b>	<b>162</b>
<b>Del individuo a la especie: importancia para la Genética de la Conservación</b>	<b>169</b>
<b>Nuevas preguntas</b>	<b>171</b>
<b>CONCLUSIONES GENERALES</b>	<b>174</b>
<b>AGRADECIMIENTOS</b>	<b>180</b>

# **PRÓLOGO**





El Diccionario de la Real Academia de la Lengua Española define al *hombre* como “Ser animado *racional*, varón o mujer”. Precisamente esta cualidad que nos diferencia del resto de seres animados (el raciocinio) es la que nos ha motivado a intentar comprender cómo *funciona* la naturaleza (qué procesos actúan), una vez percibidos y descritos los patrones originados. Dado que la naturaleza es un sistema complejo (e.g. ecosistema), necesitamos dividirla en subunidades (e.g. biocenosis) y éstas a su vez en divisiones menores, e idealmente conocer cada una de ellas y su interacción con las demás. En el mundo del conocimiento nos comportamos como un niño a quien regalan un puzzle. Sentimos la necesidad de separar las distintas piezas que lo integran. Sin embargo, sólo podremos recomponerlo si identificamos todas y cada una de las piezas y sabemos cómo encajan entre sí. A medida que avanza la ciencia sentimos la necesidad de aumentar la dificultad de nuestro puzzle, descomponiendo la realidad en piezas aún más pequeñas, y adquiriendo así un conocimiento más profundo sobre el entorno que nos rodea. Paradójicamente, la esencia de la complejidad del puzzle de la arquitectura de la vida se reduce a la combinación de sólo cuatro letras (A, T, G, C). Hoy día tenemos la capacidad de transcribir en una combinación de millones de estas cuatro letras una parte importante de lo que ha caracterizado morfológicamente a un individuo, lo que le ha condicionado a dejar o no descendencia, lo que ocurrió a miembros de su misma población y/o especie en el pasado e incluso cuánto tiempo hace que sus ancestros eran comunes con los de otras especies. Tenemos por tanto el dato bruto que nos permite aproximarnos a entender qué ocurre en la naturaleza en sus distintos niveles jerárquicos (desde el individuo hasta la especie) e incluso ir hacia atrás en el tiempo e inferir qué procesos han dado lugar a la actual distribución de los seres vivos sobre la tierra. Nuestra labor está en encontrar la piedra Rossetta que nos permita interpretar correctamente toda esta información. Esta tesis no aspira a ser la tan preciada piedra que sirva de eslabón entre realidad y conocimiento, aunque sí (humildemente) espero que contribuya a dar con ella.



# **INTRODUCCIÓN GENERAL**

## ¿Por qué una Genética de la Conservación?

Hasta los asentamientos de las primeras civilizaciones, el ser humano ha coexistido como una especie más, aprovechando los recursos naturales y, en cierto modo, guardando el equilibrio con su entorno natural. Sin embargo, la aparición de la agricultura y la domesticación del ganado frenaron su búsqueda de los recursos tróficos para favorecer el sedentarismo y consecuente alteración del medio físico y biótico. El ser humano moderno ha ocupado los cinco continentes del planeta, desde zonas de climas templados hasta las más extremas condiciones ambientales, impulsado por un continuo aprendizaje para la explotación de los recursos naturales y adecuación del medio a sus necesidades. El impacto ambiental causado por el hombre en la biosfera podría haber pasado inadvertido hasta los últimos siglos, pero el equilibrio entre la explotación y regeneración de los recursos se rompió drásticamente con la llegada de la Revolución Industrial en el siglo XVIII. Somos por tanto la causa de una de las mayores alteraciones de la biosfera, caracterizada por ser originada por una única especie (*Homo sapiens*), en un breve intervalo de tiempo (relativo a una escala de tiempo geológico) y por reducir tanto la diversidad de ecosistemas existentes en beneficio de tan sólo unos pocos (e.g. agrosistemas) como la diversidad de especies que los habitan. El acelerado ritmo con el que se está reduciendo la biodiversidad del planeta (debido a que la tasa de extinción supera con mucho a la tasa de origen de nuevas especies) ha llevado a algunos autores a nombrarla la *sexta extinción* (Leakey & Lewin, 1995; Wilson, 1992) y equipararla a las cinco previas extinciones masivas ocurridas en los últimos 600 millones de años, con la trágica salvedad del protagonismo que nos atañe. La Unión Internacional para la Conservación de la Naturaleza (IUCN en su acrónimo en inglés) incluyó en el año 1996 al 50% de las especies de vertebrados descritas y a un 12% de las especies de plantas conocidas en un listado de especies amenazadas susceptibles de

desaparecer si no logramos poner freno a esta alteración de la biosfera (Frankham *et al.*, 2002).

La Biología de la Conservación (Soulé, 1985) apareció como respuesta a la alarma que hizo saltar la comunidad científica por la pérdida de biodiversidad. Esta disciplina ha logrado crear un foro multidisciplinar (ecología, genética, biogeografía, veterinaria, ciencias sociales...) para abordar objetivos relacionados con la protección de hábitats (diseño de reservas naturales y manejo), análisis de viabilidad de poblaciones y elaboración de planes de manejo *ex-situ* para los taxones más amenazados (Loeschcke *et al.*, 1994b), etc. Fue precisamente Michael Soulé quien retomó los trabajos de Sir Otto Frankel (Frankel, 1970; Frankel, 1974) y conjuntamente reconocieron el papel que juega la diversidad genética, junto con la diversidad de especies y la riqueza de ecosistemas, como componente fundamental de la biodiversidad (Meffe *et al.*, 1997; Primack, 1998). Los genes se convierten en moneda de cambio en la evolución, de modo que ellos mismos sirven de *espejo* o *memoria* de eventos pasados ocurridos en individuos, poblaciones, especies, etc. La conservación de las especies es por tanto la conservación del resultado de un proceso evolutivo.

Una vez reconocidas la continua alteración de la biosfera, la extinción de las especies y la reducción de la diversidad genética de gran parte de las que sobreviven, uno de los mayores retos propuestos por la Biología de la Conservación consiste en disminuir estas tres tasas y, en la medida de lo posible, elaborar planes de recuperación tanto de las especies como de los hábitats destruidos. Como escribió el padre de la Filogeografía, uno objetivo prioritario de la Biología de la Conservación debe ser evitar *defoliar* aún más el árbol de la vida (Avise, 2005).

La Genética de la Conservación aplica técnicas genéticas para preservar a las especies, consideradas como entidades dinámicas con potencial adaptativo frente a los cambios en el ambiente que las rodea (Frankham *et al.*, 2002). Aparece formalmente como disciplina tras la publicación de *Genetics and Conservation* (Schonewald-Cox *et al.*, 1983) y desde entonces ha sido base de numerosos estudios y publicaciones (e.g. (Avice & Hamrick, 1996; Frankham *et al.*, 2002; Loeschcke *et al.*, 1994a). La Genética de la Conservación debe ser entendida como una aproximación al problema de la conservación de la biodiversidad, donde se dan respuestas a preguntas originadas en un marco ecológico fundamentadas en diferentes bases teóricas (genética de poblaciones, filogenias...) y herramientas (marcadores moleculares). El aporte de la Genética a la Biología de la Conservación en este punto es crucial: a medio y largo plazo, el futuro de una población sometida a cualquier grado de amenaza estará directamente relacionado con su diversidad genética. El análisis comparado de especies amenazadas con especies evolutivamente próximas no catalogadas muestra una reducción de los niveles de diversidad genética de las especies catalogadas en torno a un 35% con respecto a las no amenazadas (Frankham, 1996; Spielman *et al.*, 2004). Por ejemplo, el pingüino de las islas Galápagos (*Spheniscus mendiculus*), es una especie catalogada En Peligro por la IUCN y cuyos niveles de diversidad genética se encuentran reducidos en casi veinte veces los niveles encontrados en otras especies no amenazadas de la misma familia.

La reducción de los tamaños poblacionales causados por pérdida de hábitat incrementa de modo ineludible los riesgos de endogamia o pérdida de diversidad genética. En estos casos, la estocasticidad asociada a procesos ambientales, demográficos y/o genéticos asociados a poblaciones amenazadas determinará el futuro de estas (Evans & Sheldon, 2008; Lande, 1988; Spielman *et al.*, 2004). Gracias a los avances en Ecología y Biogeografía podemos correlacionar ciertas medidas biométricas con la condición física

del individuo, tomar el pulso a la persistencia de una población estudiando su dinámica poblacional y comparar la distribución de la especie en dos momentos distintos para evaluar la expansión o reducción de su área de distribución. Afortunadamente, las herramientas genéticas y los desarrollos teóricos permiten hoy día perfeccionar nuestro conocimiento sobre la ecología, biología e historia evolutiva de las especies. El estudio conjunto de una especie en sus tres niveles jerárquicos (individuo, población y especie) desde el punto de vista genético se convierte por tanto en un paso clave para evaluar su riesgo de extinción y elaborar planes de recuperación si fueran necesarios.

*Para la presente tesis doctoral, hemos elegido a la rata de agua (Arvicola sapidus) como especie modelo para evaluar las causas y procesos que condicionan la distribución espacial de la diversidad genética a tres escalas distintas (especie, población e individuo) y subrayar así la relevancia de la diversidad genética en la distribución de la biodiversidad. Este arvicolino de tamaño medio se encuentra distribuido por toda la Península Ibérica y gran parte de Francia, desde el nivel del mar hasta algo más de dos mil metros de altura en las montañas de Sierra Nevada. Los trabajos sobre esta especie han contribuido al conocimiento de su historia natural en diferentes hábitats del Noreste (Ventura, 1988) y Suroeste (Román, 2007b) de la Península Ibérica, además de caracterizarlo morfológica (e.g. (Garde et al., 1993)) y ecológicamente (e.g. (Fedriani et al., 2002; Garde & Escala, 2000; Román, 2007b; Ventura & Gosálbez, 1987)). Tras haber sido catalogada como “Vulnerable” en el Libro Rojo de los Mamíferos Terrestres de España (Román, 2007a) y ser propuesta para ser igualmente catalogada en Francia (Société Française pour l’Etude et la Protection des Mammifères, octubre 2007) la IUCN incluyó al año siguiente a la rata de agua entre el listado de especies Vulnerables en toda su área de distribución (Rigaux et al., 2008). De entre las causas que han conducido a la rata de agua a esta situación destacan la canalización y contaminación de aguas y la introducción de especies exóticas como el visón americano (Mustela vison). Pocos trabajos genético-moleculares se han publicado hasta la fecha sobre la especie y estos no han ido más allá de una*



*caracterización del cariotipo (Díaz de la Guardia & Pretel, 1978; Díaz de la Guardia & Pretel, 1979; Megias-Nogales et al., 2003). La repentina entrada de la rata de agua en el libro rojo de los mamíferos es un reconocimiento del delicado estado de sus poblaciones en nuestro país. En la actualidad, desconocemos cómo se distribuye la variabilidad genética de la especie a lo largo de su área de distribución. Para abordar nuestro objetivo, hemos seleccionado una batería de marcadores moleculares que nos permitan identificar a los individuos para estudiarlos a nivel poblacional y en toda su área de distribución. De este modo pretendemos entender qué factores condicionan la estructura genética actual de las poblaciones a escala de hábitat o paisajística, reconocer qué procesos históricos han originado la variabilidad genética actual de la especie y evaluar si los factores de amenaza descritos para la especie pueden estar mermando dicha diversidad genética.*

El futuro de la biodiversidad depende por tanto de la diversidad genética de las especies. En un mundo donde la biodiversidad se encuentra amenazada, debemos preservar la mayor variabilidad genética posible a nivel intraespecífico, puesto que de ella dependerá la capacidad de adaptación de las especies frente a las nuevas presiones selectivas (incluidas las de origen antrópico). El diseño y aplicación de marcadores moleculares que aseguren la correcta caracterización genética de las especies se convierte por tanto en un paso obligatorio para mejorar nuestra manera de entender y gestionar la naturaleza y sus recursos.

Hasta bien entrada la década de los años 60, la caracterización molecular de las poblaciones y especies se realizaba fundamentalmente mediante separación de las variantes de proteínas de acuerdo a su movilidad sobre una matriz sólida cuando eran sometidas a un campo electromagnético (alozimas e isoenzimas) o bien mediante el patrón de bandas generado cuando el genoma de una especie era digerido por una combinación de enzimas de restricción (RFLPs). Al igual que el microscopio

revolucionó la microbiología o el telescopio la astronomía, el descubrimiento de la estructura del ácido desoxirribonucleico (ADN) por Watson y Crick abrió, entre otras muchas, nuevas puertas al estudio de la biodiversidad, como casi parecían vaticinar los mismos autores en la segunda frase de su artículo de 1953: “*This structure has novel features which are of considerable biological interest*” (Watson & Crick, 1953). Desde entonces, la caracterización molecular de genomas enteros de algunas especies y el aislamiento de numerosos marcadores moleculares polimórficos permitieron discriminar a los organismos a nivel de individuo, siendo el único limitante la disponibilidad de ADN molde con el que trabajar. La invención de la reacción en cadena de la polimerasa (o PCR por su acrónimo en inglés) en 1986 por el premio Nobel Kary Mullis rompió esta barrera y desde entonces la genética ha formado parte de numerosas disciplinas.

*La ausencia de trabajos de caracterización molecular para A. sapidus se debe principalmente a la falta de atención sobre el estado de conservación de la especie, aunque igualmente puede justificarse por la ausencia de marcadores moleculares específicos para la rata de agua. Por tanto, para el estudio “multiescala” de este arvicolino hemos comenzado por el diseño de marcadores moleculares que permitan caracterizar a la especie en los tres niveles jerárquicos de mayor interés (individuo, población y especie). Para los dos primeros niveles seleccionaremos marcadores de repeticiones de secuencia simple o repeticiones cortas en tándem (responden a los acrónimos en inglés de SSR y STR, respectivamente), más conocidos como microsatélites, compuestos por repeticiones encadenadas de una a seis pares de bases. Estos marcadores moleculares codominantes se localizan en el ADN nuclear y presentan altas tasas de mutación (alrededor de  $10^{-3}$  a  $10^{-4}$  sustituciones por locus por replicación en roedores) que generan un elevado grado de polimorfismo. Al trabajar con microsatélites, interesa medir el tamaño total del fragmento amplificado, puesto que este permite cuantificar el número de veces que se presenta el motivo de repetición de nucleótidos. Este polimorfismo lo emplearemos para identificar a los individuos de rata de agua. Posteriormente, basándonos en la frecuencia*

*con que cada alelo –o variante del marcador por locus- aparece en las poblaciones, describiremos la estructura genética y trataremos de entender qué tipo de procesos la han originado. Las elevadas tasas de mutación reportadas para estos marcadores y la consecuente tendencia a incrementar o reducir los tamaños de los alelos originan que un mismo alelo pueda aparecer por historias evolutivas diferentes (inserción o deleción de un motivo de repetición de nucleótidos). Este fenómeno, conocido como homoplasia, dificultaría la correcta asignación del ancestro común a sus descendientes, principio al fin y al cabo en el que se basa la filogenia (ver más abajo). Por tanto, para abarcar una mayor escala temporal, con objeto de inferir qué procesos históricos han influido en la actual distribución de la diversidad genética de la especie, seleccionaremos marcadores moleculares de tasas mutacionales más lentas localizados en el ADN mitocondrial. A diferencia de los microsátélites, en este caso estamos interesados en la composición y secuencia nucleotídicas de los marcadores mitocondriales. Para ello, diseñaremos cebadores específicos para la rata de agua que nos permitan caracterizar la composición, estructura y organización de dos de los marcadores mitocondriales más usados en Filogeografía y filogenias a nivel intraespecífico: región control y citocromo b. Centraremos especialmente nuestra atención en la región control, dado que las zonas relacionadas con la regulación y replicación de la doble cadena circular de ADN mitocondrial albergan el mayor polimorfismo de todo el orgánulo citoplasmático (Saccone et al., 1993)*

## **INTERACCIÓN DE LA GENÉTICA CON OTRAS DISCIPLINAS**

La genética sirve de cemento de unión entre las escalas jerárquicas (individuo, población y especie) y temporales (genes como *espejos* o *memoria* de los eventos evolutivos pasados de las especies) y su uso debe ser considerado para evaluar el status actual de las poblaciones y elaborar planes de manejo para el futuro de estas. Veremos a

continuación cómo la Genética es utilizada igualmente como herramienta y/o “cemento” entre otras disciplinas de interés en Biología de la Conservación.

### **Genética en auxilio de la taxonomía: filogenias**

La Filogenia es el estudio de la historia evolutiva de las especies o de entidades discretas de rango superior o inferior a éstas. Empleando la alegoría de Darwin y el árbol de la vida, la filogenia trata de seguir cada una de sus ramas desde las *hojas* del árbol hacia el *tronco*. De este modo, logramos estimar no sólo en qué momento se originó cada una de esas ramas, sino también la capacidad que ha tenido cada una de ellas para originar nuevos *brotos*. Dejando las alegorías a un lado y centrándonos en estudio de las filogenias, nos encontramos con dos limitaciones: ¿qué situamos en los extremos de las ramas (i.e. qué unidades taxonómicas empleamos para establecer relaciones filogenéticas entre ellas) y cómo establecemos estas relaciones entre ellas? De nuevo estamos ante una cuestión de escala: podemos asignar desde relaciones de parentesco (pedigrí) hasta inferir qué mecanismos han causado patrones de diferenciación intra e interespecíficos, así como a niveles taxonómicos de mayor rango. Las relaciones filogenéticas entre taxa se pueden establecer en base a los cambios producidos en determinados caracteres morfológicos y biológicos (cladística). El uso de registros fósiles de edad conocida ha permitido asimismo esclarecer la historia evolutiva de las especies tanto a una escala temporal como espacial (Kurtén, 1968). Sin embargo, estos análisis sólo son factibles para aquellas especies bien representadas en los registros fósiles, excluyendo aquellos grupos cuyos reducidos tamaños poblacionales ancestrales y/o estructuras óseas frágiles o pequeñas hayan dificultado el proceso de fosilización (e.g. aves, anfibios). La sinergia entre el ingente desarrollo de las técnicas

moleculares y de modelos teóricos de evolución basados en secuencias de nucleótidos ha logrado salvar este obstáculo, facilitando el establecimiento de relaciones filogenéticas y, por tanto, la reconstrucción de la historia evolutiva (Knowles & Maddison, 2002; Maddison, 1997). La comparación de secuencias de nucleótidos y el controvertido uso de tasas de sustitución nucleotídica mantenidas en el tiempo y similares entre linajes evolutivos cercanos (hipótesis del reloj molecular (Zuckerkandl & Pauling, 1965) ha permitido, además, la datación del tiempo de divergencia entre grupos taxonómicos. Inicialmente se emplearon las divergencias genéticas brutas entre linajes, posteriormente sustituidas por la divergencia neta (i.e. divergencia una vez descontada la varianza encontrada dentro de cada uno de ellos). Poco a poco estos métodos están siendo sustituidos por modelos estadísticos más complejos basados en coalescencia, que permiten además la estima de parámetros demográficos tales como los tamaños efectivos poblacionales, tasas de migración e incluso tasas de crecimiento poblacional (e.g. aislamiento con migración, (Hey & Nielsen, 2004). Estos análisis se basan fundamentalmente en la comparación de las diversidades genéticas contemporáneas y ancestrales. Con el avance de las técnicas moleculares podemos hoy día extraer ADN de muestras fósiles de miles de años de antigüedad (Hoss & Paabo, 1993; Kim *et al.*, 2005; Stuart *et al.*, 2006) así como de especímenes conservados en museos y colecciones científicas (Rohland *et al.*, 2004; Wandeler *et al.*, 2007). La secuenciación a partir de ADN antiguo permite dar un paso más en el estudio de las filogenias, puesto el cálculo de la diversidad genética ancestral pasa de ser una estima teórica a un cálculo derivado de datos empíricos.

En resumen, hemos visto cómo la genética ha influido profundamente en la taxonomía, permitiendo inferir tiempos de divergencia, estimar tasas de intercambio genético entre especies, estimar tamaños poblacionales desde las especies ancestrales a las actuales...e

igualmente constituye un arma de doble filo. Por un lado, la genética plantea soluciones para la conservación de las especies. Por otro lado, ha avivado aún más el debate acerca del controvertido concepto de especie (Hey, 2001), tras proponer una definición basada en la monofilia recíproca encontrada en los árboles filogenéticos (concepto filogenético de especie). La propia definición de especie pone en tela de juicio si los planes de conservación están realmente dirigidos hacia las entidades adecuadas. Por ejemplo, al reconsiderar la clasificación de cerca de 1300 especies bajo el concepto de especie filogenética, el número se incrementó en torno a un 50% con respecto a las inicialmente consideradas (Agapow, 2005). Por este motivo, cobra especial importancia el estudio tanto de la diversidad genética a nivel intraespecífico (y reconsiderar la posibilidad de identificar categorías subespecíficas), como de las divergencias entre especies evolutivamente próximas (y reconsiderar la posibilidad de hibridación entre ellas), siendo más probable esta última posibilidad cuanto menor haya sido el tiempo de divergencia.

*De nuevo, la rata de agua se presenta como una especie modelo para estudiar su relación filogenética con la especie evolutivamente más próxima, la rata topera o europea (*A. terrestris*). Por un lado, el género *Arvicola* es bastante conocido por los paleontólogos (Ruiz Bustos, 1999; Ruiz Bustos, 2002), puesto que los patrones de presencia-ausencia de registros fósiles del género a lo largo de Eurasia desde su origen, hace 1.6 millones de años (Montoya et al., 2001), han permitido trazar los orígenes y sucesiones de especies dentro del mismo género. Por otro lado, ambas especies son monofiléticamente recíprocas y la divergencia nucleotídica en secuencias de genes mitocondriales se ha estimado en torno a un 7.6% (Taberlet et al., 1998). Sin embargo, la datación de divergencias interespecíficas mediante restos fósiles está condicionada tanto por la disponibilidad de éstos como por amplios intervalos de confianza alrededor de la estima de edad de los estratos y siempre bajo la asunción de que la especiación ocurre tras la diferenciación morfológica. Por tanto, especies*

*crípticas podrían dificultar la interpretación de los resultados. Por otro lado, hemos visto cómo las divergencias nucleotídicas pueden sobreestimar el tiempo de divergencia de las especies, hecho de mayor relevancia cuando las especies a considerar han divergido recientemente (Arbogast et al., 2002; Edwards & Beerli, 2000). Además, aunque ambas especies utilizan hábitats diferentes en la única zona donde sus distribuciones coinciden (i.e. Norte de España), su gran plasticidad ecológica (Folitarek, 1971; Román, 2007b) podría facilitar la hibridación entre ambas. Bajo este escenario y empleando los dos marcadores mitocondriales seleccionados en esta tesis doctoral, emplearemos un modelo basado en coalescencia de aislamiento con migración (Hey & Nielsen, 2004). De este modo, podremos situar en un contexto histórico cuándo ocurrió el aislamiento genético de ambas especies, especular sobre qué pudo originar dicha especiación y conocer si ambas especies siguieron estando conectadas por flujo génico en el pasado (tras su diferenciación) e incluso si hibridan en la actualidad en zonas de contacto secundario.*

### **La genética y el estudio de la variabilidad intraespecífica en el espacio y tiempo**

La diversidad genética de una población no puede ser entendido si no es conociendo la historia evolutiva de la especie. La diversidad genética de una especie o población es resultado de procesos contemporáneos (flujo génico contemporáneo, endogamia...) y procesos demográficos pasados (mutación, deriva génica y la selección). Por ejemplo, las poblaciones de guepardo (*Acionyx lubatus*) apenas muestran variación genética en una gran proporción de genes y alozimas estudiados. Los tamaños poblacionales de la especie se han reducido considerablemente debido a las altas tasas de furtivismo y pérdida de hábitat. Esto se ha traducido en la reproducción recurrente entre individuos emparentados (endogamia). Como consecuencia, las tasas de fertilidad se han reducido considerablemente y ha aumentado la susceptibilidad ante enfermedades infecciosas.

Hoy día se estima que este patrón de baja diversidad genética ha sido consecuencia de dos fuertes cuellos de botella que sufrió la especie hace aproximadamente unos 10.000 años y 100 años (O'Brien, 1994).

Hace poco más de veinte años, John C. Avise buscaba en los procesos históricos los factores que condicionaban la diversidad genética contemporánea de las especies, y bautizó a esta disciplina con el nombre de Filogeografía (Avise *et al.*, 1987). Uno de los fenómenos climatológicos probablemente más influyentes sobre la distribución de la fauna y flora actuales han sido las glaciaciones ocurridas desde el Pleistoceno (siendo denominadas las principales, en Eurasia, Günz (600-500 ka), Mindel (450-350 ka.), Riss (200-135 ka.) y Würm (120-18 ka.) (ka = miles de años antes del presente). Durante estos periodos hubo un descenso generalizado de las temperaturas que desencadenó cambios en cascada con repercusión en todo el planeta: se expandieron los hielos desde los polos hacia el Ecuador, lo que originó un descenso del nivel del mar de alrededor de 120 m. (abriendo zonas de paso entre continentes e islas) y un descenso de las cotas de nieve y hielos en todos los sistemas montañosos continentales. Tales modificaciones en la distribución de climas y hábitats sobre la superficie terrestre reafectaron a la flora y fauna. Hoy día conocemos que el avance de los hielos en Europa condujo a la extinción a muchas especies y forzó a otras a desplazarse hacia el sur, convirtiendo a Fennoscandia y zonas centrales de Europa en refugio para las especies asociadas a climas fríos, y a las tres penínsulas mediterráneas (Iberia, Italia y los Balcanes) en refugio para las especies asociadas a climas templados. La alternancia entre periodos glaciares e interglaciares generó aislamiento y posterior diferenciación genética entre poblaciones y/o especies (Hewitt, 2000; Taberlet *et al.*, 1998), en las divergencias profundas observadas actualmente en los árboles filogenéticos (intra e interespecíficos). De nuevo, la genética se presenta como una herramienta imprescindible para detectar el



efecto de las glaciaciones, entendidas como grandes condicionantes del origen y distribución actual de la biodiversidad en el planeta.

Las implicaciones de la Filogeografía en la Biología de la Conservación son inmediatas: frente al objetivo de preservar la mayor parte de la diversidad genética de una especie debemos conocer su distribución en el espacio y los procesos históricos que la han originado. De este modo podemos centrarnos en conservar las poblaciones que han permanecido históricamente aisladas (unidades evolutivamente significativas o ESUs) (Moritz, 1994a; Moritz, 1994b), generalmente representadas como grupos monofiléticamente recíprocos y, por tanto, conservando en cierta manera la historia evolutiva de la especie.

*Como vimos anteriormente, las poblaciones de rata de agua están sufriendo una clara regresión por depredación y por pérdida, contaminación y fragmentación de hábitats. La diversidad genética de sus poblaciones se ve por tanto amenazada debido al aumento de la endogamia y a la estocasticidad (traducida en deriva genética) asociada a este tipo de amenazas. Para establecer medidas de protección de las poblaciones remanentes, hemos de conocer cómo se encuentra actualmente distribuida la diversidad genética de la especie a lo largo de todo su área de distribución y si responde a algún patrón geográfico que facilite la delimitación de ESUs. La distribución de restos fósiles de *A. sapidus* sitúan la Península Ibérica como escenario de la diferenciación de su antecesor (*A. cantiana*), a diferencia del resto de poblaciones europeas que posteriormente divergieron hacia las distintas subespecies descritas de *A. terrestres* (Taberlet et al., 1998; Wust Saucy, 1998). Por otro lado, se han descrito diferencias morfológicas en la especie meridional que podrían corresponder con subgrupos taxonómicos. Casi desde la descripción de la especie (Miller, 1908), se ha hecho notar una leve diferenciación atendiendo a los patrones de coloración del pelaje, que correspondería con una controvertida división subespecífica. De acuerdo a ésta, la variante*

más oscura designaría a *A. sapidus tenebricus* y estaría emplazada en el Norte de la península ibérica, mientras que *A. sapidus sapidus* presentaría un pelaje marrón algo más claro y estaría ampliamente distribuida (Ventura, 2007). Por otro lado, las poblaciones del entorno natural de Doñana (Suroeste de la península) muestran una reducción en torno a un 20-25% en el tamaño corporal (Román, 2007b), que podría asimismo corresponder con una divergencia profunda en la filogenia de la especie. Por todo esto, el estudio de la filogeografía de la rata de agua cobra un doble interés. Por una parte, es una especie originada en un refugio glacial (Iberia) y probablemente colonizó las zonas libres de hielo durante los periodos interglaciales, dejando una “huella genética” reconocible. Por otra parte, los análisis genéticos podrían soportar la clasificación subespecífica propuesta en base a alguno de los criterios morfológicos mencionados anteriormente. De ser así, deberían ser consideradas como ESUs y se deberían elaborar y ejecutar planes de conservación que minimizaran las amenazas vigentes. Otra hipótesis a considerar mantiene que la distribución de la diversidad genética de la rata de agua en Iberia podría explicarse por la complejidad fisiográfica de la península y el potencial papel de los refugios dentro del refugio glacial, similar a lo que ha sido observado en un amplio rango de especies y grupos (refugio dentro de refugio (Gómez & Lunt, 2007)).

En estudios donde la escala espacial juega un papel relevante, se debe prestar especial atención a los esquemas de muestreo, tratando de maximizar la obtención de muestras a lo largo de todo el área de distribución de la variable a evaluar. En nuestro caso, es imprescindible obtener un muestreo representativo de la distribución de la variabilidad genética en todo el rango de la especie (parte de Francia e Iberia). Para ello, recurrimos a una estrategia de muestreo no invasivo raramente empleada hasta ahora en estudios filogeográfico: el empleo para la obtención de ADN de huesos de *A. sapidus* obtenidos de egagrópilas de rapaces (Taberlet & Fumagalli, 1996). El trabajo con este tipo de material conlleva unos protocolos de esterilidad en el laboratorio para evitar (o minimizar) la contaminación con ADN que pueda competir por la

*amplificación durante las PCRs. Todo ello para la extracción de escasas concentraciones de un ADN molde degradado que dificulta la amplificación y secuenciación de fragmentos de gran tamaño, especialmente si proceden del material genético del núcleo (Morin et al., 2007). En nuestro trabajo demostraremos que en estudios de Filogeografía se puede encontrar solución a los muestreos en este tipo de material, condicionado al diseño previo de cebadores específicos para las regiones más polimórficas de marcadores mitocondriales.*

### **Genética y Ecología: Ecología Molecular**

La Ecología Molecular apareció poco antes de la década de los noventa, y hace referencia a la aplicación de métodos moleculares y genéticos para contestar a problemas surgidos de un contexto ecológico. Se centra principalmente en describir los procesos ecológicos que moldean a escala de hábitat la distribución de la diversidad genética de las poblaciones.

La fragmentación y destrucción de hábitat, contaminación, introducción de especies exóticas y sobreexplotación de especies son algunas de las amenazas a la biodiversidad asociadas a la actividad del hombre. El tiempo de respuesta a estos factores varía en función de las características propias de cada población pero, si no desaparece o disminuye el grado de amenaza, el resultado será una disminución del tamaño poblacional hasta conducir a la extinción. El estudio de la diversidad y estructura genética de las poblaciones y de la tasa de flujo génico que las conecta es una de las aplicaciones más comunes de los marcadores moleculares (Bohonak, 1999; Gaggiotti *et al.*, 1999; Hudson *et al.*, 1992; Mardulyn & Milinkovitch, 2005). La mayoría de estos estudios buscan descartar la panmixia (i.e. reproducción al azar) invocando modelos teóricos donde se asume un equilibrio entre la pérdida de diversidad causada por la

deriva génica y la adquisición de nuevos alelos asociados a la migración de individuos entre poblaciones (flujo génico). Por todo esto, la descripción de la estructura genética de las poblaciones suele basarse en las frecuencias alélicas de marcadores neutrales relacionados mediante los índices de fijación o diferenciación genética (estadísticos  $F$ ) propuestos por (Wright, 1969). Estos índices de diferenciación son extremadamente dependientes de la asignación *a priori* de las poblaciones. Al igual que una población continua en el espacio puede presentar estructura genética por la presencia de barreras al flujo génico e incluso por las propias limitaciones asociadas a las capacidades dispersivas de las especies (Irwin, 2002), poblaciones geográficamente aisladas pueden permanecer conectadas por migrantes y eliminar la diferenciación genética que se esperaría por la evolución independiente de cada una de estas unidades. Por ejemplo, los individuos nacidos en las mismas colonias de cría de nóctalo común (*Nyctalus noctula*) (y, por tanto, *a priori*, con elevado grado de parentesco y similitud genética) migran en invierno hacia lugares distanciados por miles de kilómetros (Petit *et al.*, 2001). Si individuos pertenecientes a la misma generación y misma colonia fundaran poblaciones geográficamente separadas, apenas se diferenciarían genéticamente. En los últimos años, los algoritmos de asignación de individuos a poblaciones están emergiendo como una herramienta básica para estudios genético-poblacionales. Disponemos en la actualidad de métodos basados en distancias genéticas entre individuos (representados en forma de árboles filogenéticos) y análisis multivariantes donde se muestra la diversidad genética global y el aporte a ésta de cada individuo, facilitando la asignación de los individuos más parecidos genéticamente a grupos discretos. La comunión entre nuevos modelos teóricos y tecnológicos ha favorecido el desarrollo de métodos de agrupación basados en estadística Bayesiana y aproximaciones por Cadenas de Markov-Monte Carlo (MCMC) (Beaumont & Rannala, 2004). Los métodos Bayesianos emplean

la información genética de cada individuo (su genotipo) para su asignación (o proporción de su genoma) a la población de origen, con la novedad de asignar una probabilidad a esta asignación. De nuevo, la genética ayuda a la Biología de la Conservación para definir qué es una población y vuelve a replantear la definición de conceptos clásicamente usados en Ecología pero, esta vez, evaluados con nuevas herramientas (ver (Waples & Gaggiotti, 2006) para una revisión del concepto de población).

La heterogeneidad ambiental explica en gran medida la estructura genética y espacial de las poblaciones. Por ello, recientemente ha surgido una disciplina derivada de la sinergia entre genética de poblaciones y ecología del paisaje: la Genética del Paisaje (Holderegger & Wagner, 2008; Manel *et al.*, 2003). En ella se relacionan los efectos del paisaje (considerado como una matriz que facilita o dificulta el movimiento) sobre las relaciones de parentesco (nivel de individuo) y estructura genética poblacional (Aars *et al.*, 1998; Berthier *et al.*, 2005; Coulon *et al.*, 2004; Piertney *et al.*, 1998; Vos *et al.*, 2001). En un marco conservacionista, estos trabajos permiten evaluar el efecto de posibles barreras a la dispersión de individuos, así como optimizar el diseño de corredores ecológicos que aumenten la conectividad entre poblaciones aisladas o unidades de manejo (Epps *et al.*, 2007).

Hemos visto que la dispersión juega un papel de gran importancia en el estudio de la dinámica poblacional, estructura genética poblacional y conservación de las especies, tanto por sus implicaciones a corto plazo (dispersión propiamente dicha) como por el resultado promedio para tiempos evolutivos recientes (flujo génico). Para las estimas de flujo génico hemos logrado salvar hoy día las asunciones del clásico modelo de migración-deriva invocando modelos más complejos (teoría de la coalescencia) basados

en la información genealógica recogida por los datos moleculares obtenidos (Beerli & Felsenstein, 2001)

Para las estimas de los movimientos dispersivos de los individuos, los métodos genético-moleculares no requieren sucesivas capturas del mismo individuo (a diferencia de los tradicionales métodos de captura-marcaje-recaptura) y permiten el estudio de especies que abarquen grandes territorios o sean difíciles de muestrear (e.g. dificultad de acceso al área de estudio). Gracias al empleo de la información genética, hoy día podemos corregir sesgos a la hora de inferir las distancias de dispersión (Koenig *et al.*, 1996). Por último, gracias a la genética de poblaciones y al uso de marcadores moleculares de herencia uniparental (i.e. ligados a cromosomas femenino o masculino) podemos detectar diferencias ligadas al sexo, tanto en las tasas como en la distancia media de dispersión (Goudet *et al.*, 2002; Prugnolle & de Meeus, 2002)

*Aunque el nombre común de A. sapidus lo presente como un animal de hábitos acuáticos, sus características morfológicas se encuentran más relacionadas con un tipo de vida semiacuático (ausencia de membranas interdigitales o cola aplanada para nadar, longitud del radio característico de subespecies del género de hábitos terrestres o cavadores, longitud de extremidades intermedio entre hábitos acuáticos y terrestres...). Todo ello dota a la rata de agua de una gran versatilidad ecológica, claramente manifiesta en el entorno natural de Doñana. Situada en el extremo suroccidental de la Península Ibérica, Doñana se emplaza dentro de la región bioclimática mediterránea. El clima de la región es cálido y húmedo (precipitaciones anuales promedio cercanas a 600 mm). El único parámetro meteorológico predecible en Doñana es un fuerte periodo de sequía estival, mientras que la precipitación anual es muy variable. Durante el Cuaternario, se cerró el estuario del río Guadalquivir y provocó su colmatación y el asentamiento de la marisma en las zonas hundidas. Desde su origen, se han superpuesto en Doñana cinco sistemas eólicos que corresponden con cinco eventos de depósitos de arena procedentes del mar,*

*siendo de más antigüedad aquellos localizados en el interior de la comarca. En relación a su antigüedad y a la posición relativa del nivel freático, esta disposición de capas de arenas superpuestas condiciona la disposición de las cubetas lagunares y, con esto, la distribución de los hábitats potenciales para las colonias de rata de agua. A escala fina, estas colonias ocupan parches discretos de hábitat que llegan a secarse y a mantenerse sin agua durante años (Fedriani et al., 2002). Cinco años después de estas primeras prospecciones, Jacinto Román (Román, 2007b) concluía la primera monografía sobre historia natural de este arvicolino en la comarca de Doñana y cerraba con un epílogo sentenciando que “todo proyecto de investigación debería abrir tantas nuevas cuestiones como las viejas que responde”. Gracias a este excelente trabajo, sabemos hoy día cómo trampear a la especie, conocemos sus hábitos de alimentación, las características del hábitat donde se establecen las colonias, los factores de los que depende su supervivencia, los parámetros reproductivos...y un largo etcétera que, en respuesta a su epílogo, ha abierto algunas de las nuevas cuestiones abordadas en esta tesis doctoral. Por ejemplo, sabemos que lagunas, navazos (temporales y permanentes), arroyos, junqueras...constituyen los parches de hábitat en los que se establecen las colonias, por disponer del agua y/o la vegetación asociada que emplean como refugio frente a sus predadores. Debido a los periodos de sequía en la comarca, que condicionan la distribución de los parches de hábitat, la estructura genética de las poblaciones se encontrará fuertemente ligada a ciertos elementos paisajísticos. Por otro lado, las distancias de dispersión obtenidas con trabajos de captura-marcaje-recaptura no diferían significativamente entre machos y hembras (838m y 695m, respectivamente) pero los tamaños de parcela seleccionados en este estudio (3kms de diámetro) resultaron ser insuficientes para las distancias a las que parece se dispersan las ratas de agua en Doñana. En un sistema metapoblacional como el descrito en Doñana (Fedriani et al., 2002), la dinámica de extinción/colonización de las lagunas estará condicionada por los movimientos dispersivos de los individuos. Estos movimientos, a su vez, condicionan el flujo génico entre poblaciones, y dependerán de la estructura del paisaje. Una vez más, este arvicolino se presenta como*

*especie modelo para analizar los factores que a escala de paisaje explican la distribución espacial de la diversidad genética. Mediante el empleo de marcadores moleculares y el uso de Sistemas de Información Geográfica, evaluaremos los efectos que algunos elementos del hábitat pueden tener sobre la estructura genética en Doñana. Asimismo, los genotipos individuales y sus frecuencias en la comarca de Doñana serán de utilidad para estudiar la dispersión de la especie sin necesidad de reducir la escala a los 3 kms de las parcelas de muestreo durante las campañas de captura y recaptura. Con el empleo de técnicas genéticas, tendremos la posibilidad de detectar movimientos dispersivos de largas distancias, al abarcar toda la región de Doñana (aprox. 50 kms de Norte a Sur)*





**OBJETIVOS**  
**Y ORGANIZACIÓN DE ESTA TESIS DOCTORAL**

El objetivo último de esta tesis doctoral es poner de manifiesto la relevancia de los procesos históricos y contemporáneos que han determinado (y determinan) la distribución actual de la diversidad genética de las distintas especies, enfatizando su importancia en el estudio y manejo de la biodiversidad. Para llevar a cabo este objetivo hemos elegido como especie modelo a la rata de agua meridional (*Arvicola sapidus*), un roedor endémico de la Península Ibérica y Francia amenazado por la contaminación, fragmentación y destrucción de hábitat, así como por la introducción de especies exóticas (e.g. visón americano, *Mustela vison*). Todo ello ha causado un declive poblacional de la especie que ha obligado a catalogarla como Vulnerable en el Libro Rojo de las especies amenazadas de la IUCN. Entender cómo se distribuye la diversidad genética de la rata de agua meridional es por tanto en un paso clave para la elaboración de planes de conservación, manejo y recuperación de la especie. Los estudios realizados hasta la fecha sobre esta especie han centrado sus esfuerzos en la ecología e historia natural y ninguno ha afrontado el punto de vista genético-molecular.

En nuestro conocimiento, esta tesis doctoral es el primer estudio genético sobre la rata de agua realizado a distintos niveles jerárquicos y escalas temporales. Para ello, hemos caracterizado molecularmente a los individuos y poblaciones de rata de agua a partir de los marcadores moleculares diseñados en esta tesis doctoral (**capítulos 1 y 4**). A nivel de especie y a escala histórica, estimaremos el tiempo transcurrido desde que la rata de agua meridional (*A. sapidus*) se diferenció de la especie evolutivamente más próxima (*A. terrestris*) a partir del antecesor común (**capítulo 2**). A nivel intraespecífico inferiremos qué procesos históricos han condicionado la distribución de la diversidad genética de la especie a lo largo de toda su área de distribución y evaluaremos la controvertida clasificación subespecífica vigente mediante la combinación de un muestreo exhaustivo no invasivo y análisis filogeográficos (**capítulo 3**). A nivel

poblacional y escala contemporánea, emplearemos la estructura metapoblacional y el ambiente heterogéneo en que se encuentran las poblaciones de rata de agua en la región de Doñana. En este escenario evaluaremos los efectos sobre la estructura genética de las poblaciones de ciertas variables paisajísticas fundamentales para la distribución de la especie (**capítulo 4**). A nivel de individuo, emplearemos este mismo escenario en Doñana para incrementar nuestro conocimiento sobre los patrones de dispersión de los individuos en hábitats heterogéneos y las posibles causas que los condicionan (**capítulo 4**).

En el **Capítulo 1** analizamos la estructura y organización de la región control mitocondrial de las dos únicas especies del género *Arvicola* en Eurasia y reportamos la primera caracterización de la diversidad genética de *A. sapidus* y *A. terrestris* para los dos marcadores mitocondriales más empleados en filogeografía: región control y citocromo b. En base a la distribución de las diversidades nucleotídicas a lo largo de las secuencias completas de ambos marcadores moleculares, proponemos cebadores específicos de *A.sapidus* para la amplificación de los fragmentos más cortos que maximicen el contenido polimórfico.

En el **Capítulo 2** comparamos los métodos clásicos filogenéticos con modelos más recientes de aislamiento con migración basados en coalescencia para estimar el tiempo transcurrido desde la divergencia de las dos especies de rata de agua en Europa. Posteriormente contrastaremos estos resultados con la distribución espacial y temporal de los registros fósiles donde se hayan encontrado fósiles de *A. sapidus*. A pesar de que las especies segregan sus nichos ecológicos en zonas donde viven en simpatria, ambos arvicolinos muestran una gran plasticidad ecológica, que les confiere un gran potencial

de adaptación a ambientes heterogéneos, no descartando por tanto la posibilidad de hibridación entre especies. Con el segundo objetivo de este capítulo testaremos la posibilidad de flujo génico o hibridación entre ambas especies tras su divergencia. Finalmente, proporcionaremos la primera estima de los tamaños efectivos poblacionales de *A. sapidus* y *A. terrestris* y su posible equivalencia a tamaños censales.

En el **Capítulo 3** evaluaremos la influencia que ha tenido la heterogeneidad fisiográfica de la Península Ibérica sobre la distribución geográfica de la variación mitocondrial en la rata de agua, basándonos en los fragmentos polimórficos seleccionados en el Capítulo 1. Para abarcar todo el área de distribución de la especie, obtuvimos el material genético de huesos de rata de agua procedentes de egagrópilas de rapaces y de pieles conservadas en museos y colecciones científicas. En este capítulo emplearemos métodos filogenéticos y análisis demográficos para conocer cómo han condicionado las glaciaciones la actual distribución de la diversidad genética de la especie. Por último, relacionaremos la correspondencia entre la estructura intraespecífica mitocondrial y la localización de refugios glaciares recientemente descritos dentro de Iberia, así como con las diferencias morfológicas y de coloración del pelaje sugeridas con anterioridad como criterios diagnósticos subespecíficos.

En el **Capítulo 4** describiremos la estructura genética de las poblaciones de rata de agua en el entorno de Doñana y la relacionaremos con la distribución espacial de las variables de hábitat que se han descrito como condicionantes para el asentamiento de las colonias. En este mismo capítulo estimaremos las tasas de flujo génico contemporáneo entre poblaciones mediante modelos que no asumen equilibrio de Hardy-Weinberg ni

equilibrio mutación-deriva, dos asunciones comúnmente aceptadas pero rara vez justificadas en especies de ciclo de vida corto y generaciones que se solapan. La estructura genética la determinaremos con el genotipo de los individuos trampeados en Doñana en siete poblaciones geográficas con diez marcadores microsatélites obtenidos tras la amplificación cruzada y rediseño de cebadores de microsatélites diseñados para otras especies de roedores. La caracterización de hábitat y el análisis del efecto del paisaje sobre la estructura genética de las poblaciones la llevamos a cabo empleando Sistemas de Información Geográfica (SIG) y trazando rutas de mínimo coste. El uso combinado de marcadores nucleares (microsatélites) y mitocondriales de herencia uniparental (región control) y de distancias genéticas basadas en el individuo junto a los análisis de genética del paisaje nos permitirán conocer con más profundidad los patrones de dispersión de esta especie en hábitats heterogéneos.

## BIBLIOGRAFÍA

- Aars J, Ims RA, Liu HP, Mulvey M, Smith MH (1998) Bank voles in linear habitats show restricted gene flow as revealed by mitochondrial DNA (mtDNA). *Molecular Ecology* **7**, 1383-1389.
- Agapow P-M (2005) Species: demarcation and diversity. In: *Phylogeny and Conservation* (eds. Purvis A, Gittleman JL, Brooks T). The Zoological Society of London, Cambridge.
- Arbogast BS, Edwards SV, Wakeley J, Beerli P, Slowinski JB (2002) Estimating divergence times from molecular data on phylogenetic and population genetic timescales. *Annual Review of Ecology and Systematics* **33**, 707-740.
- Avise JA (2005) Phylogenetic units and currencies above and below the species level. In: *Phylogeny and Conservation* (eds. Purvis A, Gittleman JL, Brooks T). The Zoological Society of London, Cambridge.
- Avise JA, Hamrick JL (1996) *Conservation Genetics: case histories from Nature* Chapman and Hall, New York.
- Avise JC, Arnold J, Ball RM, *et al.* (1987) Intraspecific Phylogeography: the mitochondrial-DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics* **18**, 489-522.
- Beaumont MA, Rannala B (2004) The bayesian revolution in Genetics. *Nature Reviews Genetics* **5**, 251-261.
- Berli P, Felsenstein J (2001) Maximum likelihood estimation of a migration matrix and effective population sizes in n subpopulations by using a coalescent approach. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 4563-4568.
- Berthier K, Galan M, Foltête JC, Charbonnel N, Cosson JF (2005) Genetic structure of the cyclic fossorial water vole (*Arvicola terrestris*): landscape and demographic influences. *Molecular Ecology* **14**, 2861-2871.
- Bohonak AJ (1999) Dispersal, gene flow and population structure. *The Quarterly Review of Biology* **74**, 21-45.
- Coulon A, Cosson JF, Angibault JM, *et al.* (2004) Landscape connectivity influences gene flow in a roe deer population inhabiting a fragmented landscape: an individual-based approach. *Molecular Ecology* **13**, 2841-2850.
- Díaz de la Guardia R, Pretel A (1978) Karyotype and centric dissociation in water vole *Arvicola sapidus* spp *sapidus* Miller 1908 (Rodentia, Muridae). *Experientia* **34**, 706-708.
- Díaz de la Guardia R, Pretel A (1979) Comparative study of the karyotypes of two species of water vole: *Arvicola sapidus* and *Arvicola terrestris* (Rodentia, Microtinae). *Caryologia* **32**, 183-188.
- Edwards SV, Beerli P (2000) Perspective: Gene divergence, population divergence, and the variance in coalescence time in phylogeographic studies. *Evolution* **54**, 1839-1854.
- Epps CW, Wehausen JD, Bleich VC, Torres SG, Brashares JS (2007) Optimizing dispersal and corridor models using landscape genetics. *Journal of Applied Ecology* **44**, 714-724.
- Evans SR, Sheldon BC (2008) Interspecific patterns of genetic diversity in birds: Correlations with extinction risk. *Conservation Biology* **22**, 1016-1025.

- Fedriani JM, Delibes M, Ferreras P, Roman J (2002) Local and landscape habitat determinants of water vole distribution in a patchy Mediterranean environment. *Ecoscience* **9**, 12-19.
- Folitarek SS (1971) *Ecology of the water vole and its control in West Siberia* The U.S. Department of the Interior and the National Science Foundation, Washington D.C., by the Al Ahram Center for Scientific Translations.
- Frankel OH (1970) Variation, the essence of life. *Proceedings of the Linnean Society of New South Wales* **95**, 158-169.
- Frankel OH (1974) Genetic conservation: our evolutionary responsibility. *Genetics* **78**, 53-65.
- Frankham R (1996) Relationship of genetic variation to population size in wildlife. *Conservation Biology* **10**, 1500-1508.
- Frankham R, Ballou JD, Briscoe DA (2002) *Introduction to Conservation Genetics* University of Cambridge, Cambridge.
- Gaggiotti OE, Lange O, Rassmann K, Gliddon C (1999) A comparison of two indirect methods for estimating average levels of gene flow using microsatellite data. *Molecular Ecology* **8**, 1513-1520.
- Garde J, Escala M, Ventura J (1993) Determinación de la edad relativa en la rata de agua meridional, *Arvicola sapidus* (Miller, 1908) (Rodentia, Arvicolidae). *Doñana Acta Vertebrata* **20**, 266-276.
- Garde JM, Escala MC (2000) The diet of the Southern water vole, *Arvicola sapidus*, in southern Navarra (Spain). *Folia Zoologica* **49**, 287-293.
- Gómez Á, Lunt DH (2007) Refugia within refugia: patterns of phylogeographic concordance in the Iberian Peninsula. In: *Phylogeography of Southern European Refugia. Evolutionary perspectives on the origins and conservation of European biodiversity* (eds. Weiss S, Ferrand N). Springer.
- Goudet J, Perrin N, Waser P (2002) Tests for sex-biased dispersal using biparentally inherited genetic markers. *Molecular Ecology* **11**, 1103-1114.
- Hewitt G (2000) The genetic legacy of the Quaternary ice ages. *Nature* **405**, 907-913.
- Hey J (2001) The mind of the species problem. *Trends in Ecology & Evolution* **16**, 326-329.
- Hey J, Nielsen R (2004) Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. *Genetics* **167**, 747-760.
- Holderegger R, Wagner HH (2008) Landscape genetics. *Bioscience* **58**, 199-207.
- Hoss M, Paabo S (1993) DNA extraction from Pleistocene bones by a silica-based purification method. *Nucleic Acids Research* **21**, 3913-3914.
- Hudson RR, Slatkin M, Maddison WP (1992) Estimation of Levels of Gene Flow from DNA-Sequence Data. *Genetics* **132**, 583-589.
- Irwin DE (2002) Phylogeographic breaks without geographic barriers to gene flow. *Evolution* **56**, 2383-2394.
- Kim JH, Oh JH, Song JH, *et al.* (2005) Molecular genetic analysis of ancient cattle bones excavated from archaeological sites in Jeju, Korea. *Molecules and Cells* **20**, 325-330.
- Knowles LL, Maddison WP (2002) Statistical phylogeography. *Molecular Ecology* **11**, 2623-2635.



- Koenig WD, Van Vuren D, Hooge PN (1996) Detectability, philopatry and the distribution of dispersal distances in vertebrates. *Trends in Ecology and Evolution* **11**, 514-517.
- Kurtén B (1968) *Pleistocene Mammals of Europe* Weidenfeld and Nicolson, London.
- Lande R (1988) Genetics and Demography in Biological Conservation. *Science* **241**, 1455-1460.
- Leakey R, Lewin R (1995) *The sixth extinction: biodiversity and its survival* Doubleday, New York.
- Loeschcke V, Tomiuk J, Jain SK (1994a) *Conservation Genetics* Birkhäuser Verlag Basel.
- Loeschcke V, Tomiuk J, Jain SK (1994b) Introductory remarks: Genetics and Conservation Biology. In: *Conservation Genetics* (eds. Loeschcke V, Tomiuk J, Jain SK). Birkhäuser Verlag Basel.
- Maddison WP (1997) Gene trees in species trees. *Systematic Biology* **46**, 523-536.
- Manel S, Schwartz MK, Luikart G, Taberlet P (2003) Landscape genetics: combining landscape ecology and population genetics. *Trends in Ecology and Evolution* **18**, 189-197.
- Mardulyn P, Milinkovitch MC (2005) Inferring contemporary levels of gene flow and demographic history in a local population of the leaf beetle *Gonioctena olivacea* from mitochondrial DNA sequence variation. *Molecular Ecology* **14**, 1641-1653.
- Meffe GK, Carroll CR, Pimm SL (1997) Community and ecosystem-level conservation, species interactions, disturbance regimes and invading species. In: *Principles of Conservation Biology* (eds. Meffe GK, Carroll CR), pp. 235-267. Sinauer Associates, New York.
- Megias-Nogales B, Marchal JA, Acosta MJ, *et al.* (2003) Sex chromosomes pairing in two Arvicolidae species: *Microtus nivalis* and *Arvicola sapidus*. *Hereditas* **138**, 114-121.
- Miller GS (1908) Eighteen new European Voles. *Annals and Magazine of Natural History* **1**, 100-107.
- Montoya P, Alberdi MT, Barbadillo LJ, *et al.* (2001) Une faune très diversifiée du Pléistocène inférieur de la Sierra de Quibas (province de Murcia, Espagne). *C.R. Acad Sci Paris, Sciences de la Terre et des planètes* **332**, 387-393.
- Morin PA, Hedrick NM, Robertson KM, Leduc CA (2007) Comparative mitochondrial and nuclear quantitative PCR of historical marine mammal tissue, bone, baleen, and tooth samples. *Molecular Ecology Notes* **7**, 404-411.
- Moritz C (1994a) Applications of Mitochondrial-DNA Analysis in Conservation - a Critical-Review. *Molecular Ecology* **3**, 401-411.
- Moritz C (1994b) Defining Evolutionarily-Significant-Units for Conservation. *Trends in Ecology & Evolution* **9**, 373-375.
- O'Brien SJ (1994) Genetic and phylogenetic analyses of endangered species. *Annual Review of Genetics* **28**, 467-489.
- Petit E, Balloux F, Goudet J (2001) Sex-biased dispersal in a migratory bat: A characterization using sex-specific demographic parameters. *Evolution* **55**, 635-640.
- Piertney SB, MacColl ADC, Bacon PJ, Dallas JF (1998) Local genetic structure in red grouse (*Lagopus lagopus scoticus*): evidence from microsatellite DNA markers. *Molecular Ecology* **7**, 1645-1654.
- Primack RB (1998) *Essentials of Conservation Biology* Sinauer, Sunderland MA.

- Prugnolle F, de Meeus T (2002) Inferring sex-biased dispersal from population genetic tools: a review. *Heredity* **88**, 161-165.
- Rigaux P, Vaslin M, Noblet JF, Amori G, Muñoz LJP (2008) *Arvicola sapidus*. In: *2008 IUCN Red List of Threatened Species*. <[www.iucnredlist.org](http://www.iucnredlist.org)>.
- Rohland N, Siedel H, Hofreiter M (2004) Nondestructive DNA extraction method for mitochondrial DNA analyses of museum specimens. *Biotechniques* **36**, 814.
- Román J (2007a) *Arvicola sapidus* Miller, 1908. In: *Atlas y Libro Rojo de los mamíferos terrestres de España* (eds. Palomo LJ, Gisbert J, Blanco JC). Dirección General para la Biodiversidad. Secretaría Territorial para el Territorio y la Biodiversidad. Ministerio de Medio Ambiente, Madrid.
- Román J (2007b) *Historia natural de la rata de agua (Arvicola sapidus) en Doñana*, Universidad Autónoma de Madrid.
- Ruiz Bustos A (1999) The role of the water voles (*Arvicola*, Rodentia) in the Quaternary. *Estudios Geológicos* **55**, 89-95.
- Ruiz Bustos A (2002) Enamel line molar analysis in Arvicolid rodent and its potential use in biostratigraphy and palaeoecology.
- Saccone C, Lanave C, Pesole G, Sbisà E (1993) Peculiar features and evolution of mitochondrial genome in mammals. In: *Mitochondrial DNA in human pathology* (eds. DiMauro S, Wallace DC), pp. 27-37. Raven Press, New York.
- Schonewald-Cox CM, Chambers SM, MacBryde B, Thomas L (1983) *Genetics and Conservation Benjamin / Cumming*, Menlo Park.
- Soulé ME (1985) What is Conservation Biology? *BioScience* **35**, 727-734.
- Spielman D, Brook BW, Frankham R (2004) Most species are not driven to extinction before genetic factors impact them. *Proceedings of the National Academic of Science* **101**, 15261-15264.
- Stuart BL, Dugan KA, Allard MW, Kearney M (2006) Extraction of nuclear DNA from bone of skeletonized and fluid-preserved museum specimens. *Systematics and Biodiversity* **4**, 133-136.
- Taberlet P, Fumagalli L (1996) Owl pellets as a source of DNA for genetic studies of small mammals. *Molecular Ecology* **5**, 301-305.
- Taberlet P, Fumagalli L, Wust-Saucy A-G, Cosson J-F (1998) Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology* **7**, 453-464.
- Ventura J (1988) *Contribución al conocimiento del género Arvicola (Lacepede, 1799) en el Nordeste de la Península Ibérica*, Universidad de Barcelona.
- Ventura J (2007) *Arvicola sapidus* Miller, 1908. In: *Atlas y Libro Rojo de los mamíferos terrestres de España* (eds. Palomo LJ, Gisbert J, Blanco JC). Dirección General para la Biodiversidad. Secretaría Territorial para el Territorio y la Biodiversidad. Ministerio de Medio Ambiente, Madrid.
- Ventura J, Gosálbez J (1987) Reproductive biology of *Arvicola sapidus* (Rodentia, Arvicolidae) in the Ebro delta (Spain). *Zeitschrift für Säugetierkunde* **52**, 364-371.
- Vos CC, Antonisse-De Jong AG, Goedhart PW, Smulders MJM (2001) Genetic similarity as a measure for connectivity between fragmented populations of the moor frog (*Rana arvalis*). *Heredity* **86**, 598-608.
- Wandeler P, Hoeck PEA, Keller LF (2007) Back to the future: museum specimens in population genetics. *Trends in Ecology & Evolution* **22**, 634-642.

- Waples RS, Gaggiotti O (2006) What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. *Molecular Ecology* **15**, 1419-1439.
- Watson JD, Crick FHC (1953) Molecular Structure of Nucleic Acids - a Structure for Deoxyribose Nucleic Acid. *Nature* **171**, 737-738.
- Wilson EO (1992) *The diversity of life* W. W. Norton & Co., New York.
- Wright S (1969) *Evolution and the Genetics of Populations, vol.2* University of Chicago Press, Chicago.
- Wust Saucy AG (1998) *Polymorphisme genetique et phylogeographie du campagnol terrestre Arvicola terrestris*, Université de Lausanne.
- Zuckerlandl E, Pauling L (1965) Evolutionary divergence and convergence in proteins. In: *Evolving Genes and Proteins* (eds. Bryson V, HJ V), New York.

# **CHAPTER 1**



## CHAPTER 1

### **Structure, organization and nucleotide diversity of the mitochondrial control region of water voles in Eurasia (*Arvicola* spp.)**

Alejandro Centeno-Cuadros <sup>1</sup> and José Antonio Godoy<sup>2</sup>

<sup>1</sup> Department of Conservation Biology, Estación Biológica de Doñana, CSIC. Avda. Américo Vespucio, s/n. C.P. 41092. Sevilla, Spain.

<sup>2</sup> Department of Integrative Ecology, Estación Biológica de Doñana, CSIC. Avda. Américo Vespucio, s/n. C.P. 41092. Sevilla, Spain.

**Keywords:** CSB, cytochrome b, D-Loop, ETAS, Muridae, Rodentia.

**Running title:** Structure and organization of CR in *Arvicola*



## RESUMEN

En este trabajo hemos analizado la estructura y organización de la región control mitocondrial entera en las dos únicas especies existentes del género *Arvicola* en Eurasia. Nuestros resultados muestran que la región control de estas especies se caracteriza por presentar un bajo contenido GC, la ausencia de cualquier motivo de repetición dentro de las dos regiones hipervariables y por la presencia de dos secuencias extendidas asociadas a la terminación (ETAS, por su acrónimo en inglés) y bloques de secuencias conservados (CSB). Las comparaciones entre diversidades nucleotídicas dentro de esta región en las dos especies de rata de agua mostró diferencias en la distribución de la variación genética. En este trabajo sugerimos cebadores para la amplificación de un fragmento de 246 pb que incluye la mayor parte de la variación nucleotídica potencialmente útil para el estudio de la variación intraespecífica de la rata de agua meridional (*Arvicola sapidus*). Del mismo modo, reportamos cebadores para la amplificación de la región control completa y fragmentos cortos que maximizan el polimorfismo para el citocromo b. El uso en conjunto de estos marcadores ofrece unas herramientas moleculares para la elaboración de planes de conservación y manejo, así como facilitar los estudios a diferentes escalas espaciales y evolutivas para esta especie.





## ABSTRACT

We analyse the structure and organization of the entire mitochondrial control region in the only two extant species of the genus *Arvicola* in Eurasia. Our results show that the control region of these species is characterized by a low GC content, the absence of any repetition motif within the two hypervariable regions, and the presence of the two extended termination associated sequences (ETAS) and conserved sequence blocks (CSB). Comparisons between nucleotide diversities within this region in the two water voles species revealed differences on the distribution of genetic variation. We suggest primers for the amplification of a 246bp-fragment that includes most of the nucleotide variation to be used in the study of intraspecific variation of the threatened Southern water vole (*Arvicola sapidus*). We also report primers to amplify both the entire coding region and short maximally polymorphic fragments of cytochrome b. The combined application of these markers offers molecular tools to establish future conservation and management guidelines and also will facilitate studies at different spatial and evolutionary scales of this species.



The vertebrate mitochondrial DNA (mtDNA) is a closed, circular and maternally inherited molecule widely used in molecular ecology and phylogeographic studies for the last 25 years (Ballard and Whitlock, 2004). Among 37 genes encoded within mtDNA, Control Region (CR) and cytochrome b (cytb) have been most frequently used in molecular studies at intra and interspecific levels. CR is a non-coding mitochondrial fragment whose length varies depending on the number of repeated sequences within the region. The CR of mammals is flanked by tRNA<sup>Pro</sup> and tRNA<sup>Phe</sup> and is divided into two hypervariable regions separated by a central conserved region. This central region is characterized by three conserved sequence boxes (CSB) involved in regulatory signals for the processing of the RNAs that prime replication. There are also two hypervariable regions composed by extended termination associated sequences (ETAS), the most rapidly evolving part of the mitochondrial genome (Saccone et al., 1993) which are presumably related to the regulation of replication and transcription of mtDNA. Hypervariable fragments within the CR usually meet the polymorphism required to address issues related to population genetic structure or phylogeography. However, due to the variable and heterogeneous distribution of sequence variation along the region, and to the possible occurrence of homopolymer tracts and tandem repeats that might greatly hamper the direct sequencing of PCR products, a thorough characterization of the structure of this region becomes necessary previous to its routine use in molecular ecology and evolutionary studies. Furthermore, finding short fragments that include as much polymorphism as possible becomes a need for studies based on degraded materials like non-invasive samples and museum specimens.

The only two species of the genus *Arvicola* are distributed over Eurasia. The European water vole (*A. terrestris*) range throughout Eastern Asia up to Northern regions of Iberia and it has been used as subject of numerous molecular ecology studies (Berthier et al., 2005; Oliver and

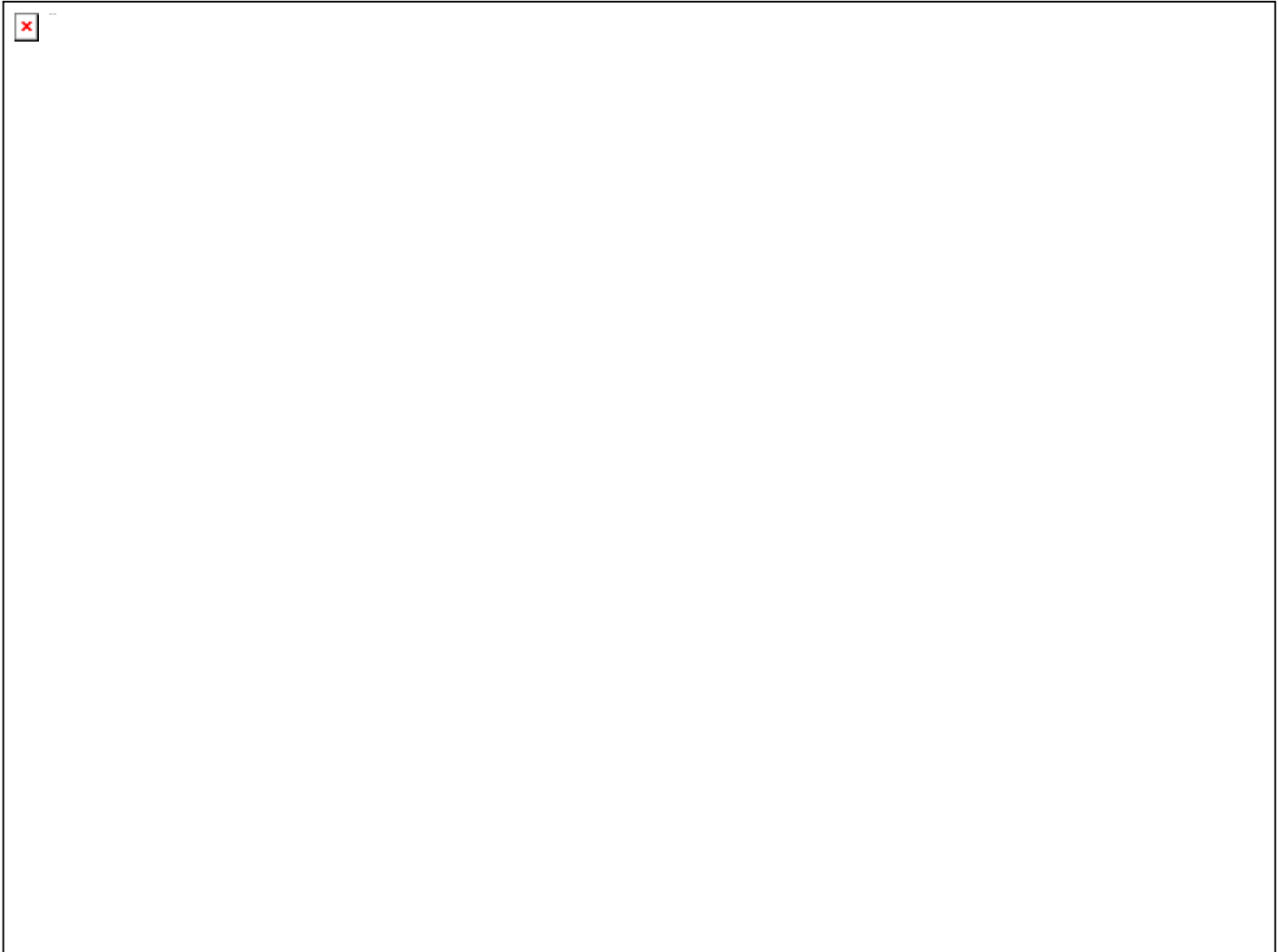
Piertney, 2006; Piertney et al., 2005; Stewart et al., 1999; Stewart et al., 1998). On the other hand, Southern water vole (*A. sapidus*) is distributed through Iberia and France and, despite being endangered by habitat fragmentation, contamination of water bodies and the introduction of the American mink (*Mustela vison*), molecular studies are scarce and only focused on its karyotype (Díaz de la Guardia and Pretel, 1978; Díaz de la Guardia and Pretel, 1979; Megías-Nogales et al., 2003). Divergences on *cytb* have dated the split of European water voles from its evolutionary sister species over Pleistocene glaciations in Iberia (Taberlet et al., 1998). There are no studies on genetic variability in CR in water voles in Eurasia, partially due to homopolymer sequences that hamper the sequencing reaction for the entire region (Piertney et al., 2005) (S. Piertney, personal communication)

In this study we i) characterize the organization and variability of the entire mitochondrial control region sequence in *A. sapidus*, ii) provide entire sequences of control region for *A. sapidus* by designing internal specific primers, iii) provide primers for the amplification of short and highly polymorphic fragments of CR and *cytb* for the study of intraspecific variation and iv) apply these primers on water voles species for a description of the distribution of nucleotide diversities within CR.

Samples from 47 individuals of Southern water voles were used to describe nucleotide variation in the mitochondrial CR. DNA was extracted with a “salting-out” protocol (Müllenbach et al., 1989). We solved the homopolymer sequencing reaction problem in CR by using primers F15708 and R92 (Piertney et al., 2005), primer 5'-TCCCCACCATCAGCACCCAAAGC-3' designed by Stacy et al. (1997) (hereafter, F15374), and four specifically designed internal primers whose combination yield partially overlapping fragments (F15816, 5'-ATGTTTTATCGTCCATACGTTCC-3'; F15872, 5'-AATCAGCCCATGCCTAACAT-3'; R15946, 5'-TAGCCGTCAAGGCATGAAG-3';

RCRasa 5'-AAAAACAACCTCAAAATTCCAAAA-3'). Cytb was amplified and sequenced using primers H15288 and L14115 (Martin et al., 2000). PCR amplifications were performed as follows: 94°C for 5min, 40 cycles at 92°C for 30s, 62°C (CR) or 60°C (cytb) for 30s and 72°C for 30s, finishing with 72°C for 5min. 5µL of mitochondrial PCR products were purified with 2µL ExoSAP-IT enzyme (USB Corporation), incubated during 15min/37°C and inactivated 15min/80°C. Sequencing reactions were performed using the Applied Biosystems BigDye® Terminator Cycle Sequencing Kit v.1.1 following the manufacturer's instructions and the same primers used for the amplifications. Sequencing products were run on an Applied Biosystems 3130xl Genetic Analyzer. Forward and reverse sequences were edited and aligned using Sequencher 4.6 (Gene Codes Corporation, Ann Arbor, MI). We described the distribution of nucleotide diversity along CR using a sliding window of 100 bases (step size of 25 base pairs) as implemented in the software DNAsp 4.50.3 (Rozas et al., 2003)

Herein we report the first analysis of the structure of the whole *Arvicola* CR and of the distribution of nucleotide diversity along the region (figure 1). By using four partially overlapping fragments we were able to obtain sequences of the whole CR from 47 individuals of SWV. We identified a central region that includes two conserved sequence boxes. While CSB3 was completely conserved in *A. sapidus*, CSB2 differs in two indels with respect to corresponding sequences in *M. musculus* mitochondrial genome. We could not allocate *Arvicola* CSB1, indicating large divergences between *M. musculus* and *A. sapidus* sequences. However, the degree of conservation and functionality of this sequence block has been controversial. For example, whereas Sbisa et al. (1997) reported the CBS1 as the least conserved sequence block and suggested it as the most important in functionality, studies in *Clethrionomys* (Matson and Baker, 2001) showed CSB1 as a sequence block with 40% of variability. These authors suggest that only one of the CSB and ETAS elements are involved



**Figure 1** Structure and organization of the complete cytochrome b and mitochondrial control region of *Arvicola sapidus*. Forward and reverse primers are indicated with an arrow above their respective position in the mitochondrial genome (a, b) and as underlined sequences (c). Numbers on the upper line specify the initial position on the consensus sequences of *A. sapidus* obtained in this study. Numbers below the lower line shows the relative position on *Mus musculus* mitochondrial DNA. The graph above the control region map (b) reveals the distribution of nucleotide diversity along the sequence. An asterisk on a nucleotide position highlights a polymorphic site among 47 individual sequences. Arrows on c) delimit conserved blocks in D-Loop sequences. ETAS: Extended Termination Associated Sequences. CSB: Conserved Sequence Blocks.

in mitochondrial replication and, consequently, any variation within any additional element should not have any detrimental effect. Base composition and distances between CSB might be variable among species but are usually conserved at the intraspecific level (Sbisa et al., 1997). *Arvicola* global nucleotide composition is AT-rich and follows the common biased base content found in all organisms [*A. sapidus* average percentages: A (32.4), T (30.7), C (23.8) and G (13.1)] (Saccone et al., 1987). However, Southern water voles showed slight deviations from the expected A>T>C>G pattern in mammals'

ETAS and CSBs [ETAS: T (32.3), A (31.5), C (25) and G (11.3); CSB: C (33.3), A (29.3), T (25.8), and G (11.6)]. The distribution of nucleotide diversity along the CR reveals the allocation of most of the overall polymorphism ( $\pi = 0.0406$ , nucleotide midpoint position 150) at the hypervariable domain I (sliding window analysis, figure 1), whereas the maximum peak of nucleotide diversity in the second hypervariable domain was four times lower ( $\pi = 0.0103$ , midpoint position 725). 14 out of 19 total polymorphic sites were located within the hypervariable domain I. The distribution of nucleotide diversity in *A. sapidus* is biased toward the first hypervariable domain as in other species of rodents [e.g. *Spalax galili* (Reyes et al., 2003)] and differs from most commonly observed patterns where the polymorphism is divided up between both hypervariable regions or even greatly biased toward the second hypervariable domain (Baker and Marshall, 1997; Matson and Baker, 2001; Roques et al., 2004). Considering the observed distribution of polymorphisms and that the CSBs are flanked by several homopolymer tracts (presumably involved in forming a stable hairpin structure in CR), we propose to target a 246bp CR fragment in the first hypervariable region that contains most of the total polymorphism in *A. sapidus* using forward F15468 (5'-GCATTAAATTATATTCCCATGC-3') and reverse R15713 (5'-TTGTTGGTTTCACGGAGGAT-3') primers. Genebank accession number, FJ502319.

We also characterized the complete 1143 bp coding region of *cytb* from six individuals from different geographic locations in Spain (Genebank accession numbers: FJ539341 to FJ539346). We found five different haplotypes defined by ten polymorphic nucleotide positions of which nine corresponded to non-synonymous and one to synonymous mutations. In order to maximize the polymorphism amplified in a single short fragment we suggest to target a 248 bp fragment using specific primers F14559



(5'-TCCTTTTGAGGGGCTACAGT-3') and R14806 (5'-TGGAAGGGAATTTTGTCTGC-3').

Finally we applied primers F15468 and R15713 (CR) and F14559 and R14806 (cytb) on *A. sapidus* (n= 26) and *A. terrestris* (n= 22) to compare nucleotide diversities between fragments and species. Nucleotide diversity in cytb was lower in *A. sapidus* ( $\pi = 0.00374$ ) than in *A. terrestris* ( $\pi = 0.00912$ ), whereas in CR it was higher in *A. sapidus* ( $\pi = 0.03623$ ) than in *A. terrestris* ( $\pi = 0.00770$ ). These nucleotide diversities in CR describe the polymorphism in the 246 bp fragment within the first hypervariable region in both species of water voles. Although sequence data from the 5' end of the hypervariable domain I of *A. terrestris* are lacking (mainly due to the presence of homopolymer tracts) we compare the information obtained in this study with that described in water voles from the United Kingdom (Piertney et al., 2005). In contrast to Southern water voles, European water voles apparently assemble higher nucleotide diversity in the second hypervariable domain, a region that showed scarce variation in Southern water voles.

Following a thorough characterization of the structure and nucleotide variation in *A. sapidus* mitochondrial CR, we propose the first hypervariable domain in *A. sapidus* as the most informative mitochondrial marker for addressing studies at both intra and interspecific scales. We report primers for the amplification of maximally variable but short fragments in both CR and cytb (246 and 248 bp, respectively), which should proof suitable for the characterization of the mitochondrial variation of the species even from highly degraded material.

## Acknowledgements

We are especially grateful to J.M. Llanes, M. Gutiérrez, M. González, F. Alda, J. Román and M. Delibes. This work has been funded by the Dirección General de Investigación (project BOS2001-2391-C02-01). ACC benefited from a grant from the Ministerio de Educación y Ciencia.

## REFERENCES

- Baker, A. J., and Marshall, H. D. (1997) Mitochondrial control region sequences as tools for understanding evolution. In: Avian molecular evolution and systematics (ed. D.P. Mindell), pp. 51-82. Academic Press, Michigan.
- Ballard, J. W. O., and Whitlock, M. C. (2004) The incomplete natural history of mitochondria. *Mol. Ecol.* **13**, 729-744.
- Berthier, K., Galan, M., Foltête, J. C., Charbonnel, N., and Cosson, J. F. (2005) Genetic structure of the cyclic fossorial water vole (*Arvicola terrestris*): landscape and demographic influences. *Mol. Ecol.* **14**, 2861-2871.
- Díaz de la Guardia, R., and Pretel, A. (1978) Karyotype and centric dissociation in water vole *Arvicola sapidus* spp *sapidus* Miller 1908 (Rodentia, Muridae). *Experientia* **34**, 706-708.
- Díaz de la Guardia, R., and Pretel, A. (1979) Comparative study of the karyotypes of two species of water vole: *Arvicola sapidus* and *Arvicola terrestris* (Rodentia, Microtinae). *Caryologia* **32**, 183-188.
- Martin, Y., Gerlach, G., Schlotterer, C., and Meyer A. (2000) Molecular phylogeny of European muroid rodents based on complete cytochrome b sequences. *Mol. Phylogenet. Evol.* **16**, 37-47.
- Matson, C. W., and Baker, R. J. (2001) DNA sequence variation in the mitochondrial control region of red-backed voles (*Clethrionomys*). *Mol. Biol. Evol.* **18**, 1494-1501.
- Megías-Nogales, B., Marchal, J. A., Acosta M. J., Bullejos M., Díaz de la Guardia R., and Sánchez A. (2003) Sex chromosome pairing in two Arvicolidae species: *Microtus nivalis* and *Arvicola sapidus*. *Hereditas* **138**, 114-121.
- Müllenbach, R., Lagoda, P. J. L., and Welter C. (1989) An efficient salt-chloroform extraction of DNA from blood and tissues. *Trends Genet.* **5**, 391-391.
- Oliver, M. K., and Piertney, S. B. (2006) Isolation and characterization of a MHC class II DRB locus in the European water vole (*Arvicola terrestris*). *Immunogenetics* **58**, 390-395.
- Piertney, S. B., Stewart, W. A., Lambin, X., Telfer, S., Aars, J., and Dallas, J. F. (2005) Phylogeographic structure and postglacial evolutionary history of water voles (*Arvicola terrestris*) in the United Kingdom. *Mol. Ecol.* **14**, 1435-1444.

- Reyes, A., Nevo, E., and Saccone, C. (2003) DNA sequence variation in the mitochondrial control region of subterranean mole rats, *Spalax ehrenbergi* superspecies, in Israel. *Mol. Biol. Evol.* **20**, 622-632.
- Roques, S., Godoy, J. A., Negro, J. J., and Hiraldo F. (2004) Organization and variation of the mitochondrial control region in two vulture species, *Gypaetus barbatus* and *Neophron percnopterus*. *J. Hered.* **95**, 332-337.
- Rozas, J., Sanchez-Del Barrio, J. C., Messeguer, X., and Rozas, R. (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* **19**, 2496-2497.
- Saccone, C., Attimonelli, M., and Sbisà, E. (1987) Structural elements highly preserved during the evolution of the D-Loop-containing region in vertebrate mitochondrial DNA. *J. Mol. Evol.* **26**, 205-211.
- Saccone, C., Lanave, C., Pesole, G., and Sbisà, E. (1993) Peculiar features and evolution of mitochondrial genome in mammals. In: *Mitochondrial DNA in human pathology* (eds. S. DiMauro and D. C. Wallace), pp. 27-37. Raven Press, New York.
- Sbisà, E., Tanzariello, F., Reyes, A., Pesole, G., and Saccone, C. (1997) Mammalian mitochondrial D-loop region structural analysis: identification of new conserved sequences and their functional and evolutionary implications. *Gene* **205**, 125-140.
- Stacy, J. E., Jorde, P. E., Steen, H., Ims, R. A., Purvis, A. and Jakobsen, K. S. (1997) Lack of concordance between mtDNA gene flow and population density fluctuations in the bank vole. *Mol. Ecol.* **6**, 751-759.
- Stewart, W. A., Dallas, J. F., Piertney, S. B., Marshall, F., Lambin, X., and Telfer, S. (1999) Metapopulation genetic structure in the water vole, *Arvicola terrestris*, in NE Scotland. *Biol. J. Linn. Soc.* **68**, 159-171.
- Stewart, W. A., Piertney, S. B., and Dallas, J. F. (1998) Isolation and characterization of highly polymorphic microsatellites in the water vole, *Arvicola terrestris*. *Mol. Ecol.* **7**, 1258-1259.
- Taberlet, P., Fumagalli, L., Wust-Saucy, A.-G., and Cosson, J.-F. (1998) Comparative phylogeography and postglacial colonization routes in Europe. *Mol. Ecol.* **7**, 453-464.

## **CHAPTER 2**



## CHAPTER 2

### **Dating the divergence between Southern and European water voles using molecular coalescent-based methods**

Alejandro Centeno-Cuadros<sup>1</sup>, Miguel Delibes<sup>1</sup> and José Antonio Godoy<sup>2</sup>

<sup>1</sup> Department of Conservation Biology, Estación Biológica de Doñana, CSIC. C/ Américo Vespucio, s/n. CP 41092. Sevilla, Spain.

<sup>2</sup> Department of Integrative Ecology, Estación Biológica de Doñana, CSIC. C/ Américo Vespucio, s/n. CP 41092. Sevilla, Spain.

Keywords: *Arvicola sapidus*, *Arvicola terrestris*, coalescence, divergence, isolation-with-migration model, Pleistocene.

Running title: Evolutionary history of *Arvicola*



## RESUMEN

Para comprender la evolución de la biodiversidad y asegurar su conservación es crítico conocer la historia evolutiva de las especies. Las aproximaciones moleculares aparecen encaminadas a alcanzar esta meta, volviéndose más poderosas con la incorporación de avances teóricos y analíticos recientes. En este trabajo hemos aplicado un modelo de aislamiento con migración basado en la teoría del coalescente sobre secuencias de la región control y el citocromo b mitocondrial para investigar la historia evolutiva de la rata de agua europea (*Arvicola terrestris*) y meridional (*A. sapidus*) por todo Eurasia. Hemos datado el tiempo de divergencia del ancestro común a ambas especies de rata de agua hacia el cuarto glacial de Mindel (aprox. 250.000 años). A pesar de que la expansión postglacial de las poblaciones ha hecho entrar a estas especies en contacto secundario en Iberia y Francia, nuestros resultados no contienen evidencias de hibridación reciente o flujo génico entre los dos grupos descendientes. Hemos estimado los tamaños efectivos poblacionales de *A. sapidus* y *A. terrestris* en torno a los 199.000 y 52.0000 individuos (respectivamente). Este estudio muestra la utilidad de las herramientas moleculares basadas en coalescencia para investigar las divergencias recientes entre especies y poblaciones y contribuye al conocimiento de la evolución del género *Arvicola* en Eurasia.





## ABSTRACT

Investigating the evolutionary history of the species is critical for the understanding the evolution of biodiversity and for its conservation. Towards this goal molecular approaches are becoming increasingly powerful with the incorporation of recent theoretical and analytical advances. Here, we apply an isolation-with-migration coalescent-based model to control region and cytochrome b mitochondrial sequences to investigate the evolutionary history of European (*Arvicola terrestris*) and Southern (*A. sapidus*) water voles throughout Eurasia. We date the time of the split of the common ancestor of European and Southern water voles around the fourth Mindel glacial (*ca.* 252000 ybp), a date similar to that estimated by classical phylogenetic methods (*ca.* 241000 ybp). Despite postglacial expansions brought these species into secondary contact in Iberia and France, our results do not contain evidence of subsequent gene flow between descendant groups or current hybridization. We also estimate the effective population sizes of *A. sapidus* and *A. terrestris* on *ca.* 199000 and 52000 individuals, respectively. This study illustrates the utility of modern coalescent-based genetic tools to investigate recent species and population divergences and add some light on the evolution of the genus *Arvicola* in Eurasia.



## INTRODUCTION

The knowledge about the evolutionary history of species is an important topic in biology. Besides, zoologists and wildlife managers need to identify conservation and management units (Moritz, 1994), which should be based on our best knowledge of the evolutionary history of the populations and taxa. Multidisciplinary approaches are thus strongly recommended to integrate, among others, ecological, behavioural and genetic aspects into the optimal design of these operational units (Fraser & Bernatchez, 2001). However, time of divergence between allopatric groups is often used as a shortcut for the harder to evaluate degree of reproductive isolation, adaptive divergences or “ecological exchangeability”. For many vertebrates, fossil registers in strata of known age have allowed a good knowledge on species history and distribution at different spatial and temporal scales (Kurtén, 1968). Nevertheless, low ancestral population sizes, small body sizes and/or fragile and hollow bones which do not easily fossilize, frequently result in scarce fossil registers, hampering the study of the evolutionary paths of many species and groups (e.g. amphibians, birds). The burgeoning advances in genetic techniques and theoretical models on evolution of DNA sequences have shown that the evolutionary history of species hidden in individuals’ morphology is nowadays even more readily and clearly revealed by their genes. Consequently, genetic approaches have filled gaps in the knowledge of the evolution of life since its origin, establishing genetic relatedness between taxa (phylogeny) that are used to infer their evolutionary history (Maddison, 1997; Knowles & Maddison, 2002)

In a genetic frame, dating speciation events has been classically addressed by phylogenetic methods since the proposal of the molecular clock hypothesis (Zuckerkandl & Pauling, 1965), by which genes are proposed to evolve at rates roughly constant over time and across evolutionary lineages. Its use is not exempt of controversies, such as high variance on their

calibration and heterogeneity among genomic regions and lineages, among others (Edwards & Beerli, 2000; Arbogast, Edwards, Wakeley *et al.*, 2002). However, molecular clocks are still the only hope to date speciation events in species for which no fossils registers are available. One recently stressed caveat of phylogenetic methods is the overestimation of the time since divergence of species due to the difference between the time of split of species and genes. Genealogies represent the split between lineages, which generally precede species divergence due to polymorphism in the ancestral population (Nei & Li, 1979; Wilson, Cann, Carr *et al.*, 1985). Whereas the ratio between times of genes split and species split is irrelevant for deeply divergent species, it increases and becomes relevant when recently diverged species are considered (Edwards & Beerli, 2000; Arbogast *et al.*, 2002). The use of net instead of absolute divergences for molecular dating tries to correct this bias by equating the ancestral polymorphism to the average of the polymorphism observed in contemporary derived species (Nei & Li, 1979). Recently developed Isolation-with-Migration (IM) coalescent-based models allow not only the estimation of the time of divergence of pair of species while accounting for the polymorphism in the ancestral species, but also of other demographic parameters, such as current and ancestral effective population sizes and post-split migration rates (Nielsen & Wakeley, 2001). Besides, IM models may simultaneously consider multiple loci with different mutation rates, substantially reducing the variance due to the stochasticity of coalescent processes (Hey & Nielsen, 2004).

Molecular analysis of species divergence does not depend on fossil availability. However when applied to cases with well-known paleontological registers, the results of both approaches may be compared and contrasted. Palaeontology and genetics exchange input and complement each other for a better understanding on evolutionary history of species. The evolutionary history of the genus *Arvicola* (Cricetidae, Rodentia) is well known by

palaeontologists (Ruiz Bustos, 1999; Ruiz Bustos, 2002). Fossil registers reveal that the genus *Arvicola* has been restricted to Eurasia since its origin around 1.6 Mybp (Montoya, Alberdi, Barbadillo *et al.*, 2001) At present, there are two species in the genus: the European water vole (*A. terrestris*), which ranges all over Eurasia up to Northern Iberia, and the Southern water vole (*A. sapidus*), restricted to Iberia and France. Both species overlap their ranges in Northern Iberia, Southern France and the Alps. Paleontological studies have also documented the split of both species from their common ancestor during the Middle Pleistocene (Sesé Benito, 1994; Ruiz Bustos, 1999; Escude, Montuire, Desclaux *et al.*, 2008).

Since the fossil record of these species shows no temporal gaps and reveals a continuous change over time from a single lineage into morphologically distinct forms, they have been considered *chronospecies* i.e. species widely used in palaeontology and other geological sciences to estimate the age of strata (Stanley, 1978). The absence of *A. terrestris* fossils in central and Southern Iberia suggests a fragmentation of the range of the common ancestor during Quaternary glaciations (likely *A. cantianus*, previously named *A. cantiana*). Accordingly, one population of this common ancestor would have remained isolated in the Iberian Peninsula southern glacial refugium (Sesé Benito, 1994; Ruiz Bustos, 1999) and evolved there in isolation from Northern populations, giving origin to the current *A. sapidus* (allopatric speciation). Postglacial range expansions brought these species into secondary contact in France and Northern Iberia.

Previous genetic approaches to the evolutionary history of *Arvicola* in Europe have characterized the intraspecific variation of genetic diversity of both species across their distribution range (Wust Saucy, 1998; Wust Saucy, Hausser, Taberlet *et al.*, 1998; Piertney, Stewart, Lambin *et al.*, 2005). Three genetic lineages have been recognized and proposed as subspecies within *A. terrestris* (namely *A. t. terrestris*, *A. t. italicus* and *A. t. sherman*), while

subspecific delimitations within *A. sapidus* are based on pelage colour variation but are not supported by mitochondrial data (Centeno-Cuadros, Delibes & Godoy, *in press*). Phylogenetic molecular analyses showed *A. terrestris* and *A. sapidus* to be reciprocally monophyletic and estimated their net nucleotide divergence as 7.6% in mitochondrial gene sequences (Taberlet, Fumagalli, Wust-Saucy *et al.*, 1998). Until now it has not been tested, however, whether both species have been connected by gene flow in the past (after the split) or are nowadays hybridizing in secondary contact zones.

Here we use a Bayesian coalescent-based method (Hey & Nielsen, 2007) to specifically: i) date the time of divergence of both extant water voles in Europe, ii) test whether *A. sapidus* and *A. terrestris* were/have been connected by gene flow after the split, *iii*) infer ancestral and current effective population sizes for both species of water voles, and *iv*) compare our results with those derived from paleontological and traditional phylogenetic approaches.

## **MATERIAL AND METHODS**

### **Samples and sequence data**

One of the main assumptions of the IM model is that there should not be any intraspecific genetic structure within each species' data set. Accordingly, although *A. sapidus* shows a shallow phylogeographic structure, we only chose sequences from the most widespread and statistically supported group that expands throughout Iberia and France (Centeno-Cuadros, Delibes & Godoy, *in press*). All sequences of *A. terrestris* in this study belonged to individuals from the *A. t. sherman* lineage, which did not show any internal genetic structure. DNA was extracted from bones obtained in raptors pellets and tissues from specimens of known origin in scientific collections (Taberlet & Fumagalli, 1996; Waits & Paetkau, 2005)

although fresh tissue samples were also used when available. On a pilot study, this kind of material did not prove its adequacy for obtaining nuclear DNA. However, we obtained encouraging results with mitochondrial DNA amplifications. Accordingly, we used the genetic information contained in two mitochondrial DNA markers: control region (CR) and cytochrome b (cytb). Mitochondrial DNA patterns have nevertheless been shown to be robust indicators of patterns of population and species evolutionary history, and the markers most suited for the analyses of recently isolated populations and species delimitation (Zink & Barrowclough, 2008). The targeted sequences were specifically designed to maximize the polymorphism in the shortest fragment for each marker (Centeno-Cuadros & Godoy, submitted). Most polymorphism available in *ca.* 2200 bp (total length of control region and cytochrome b) was roughly contained in a 413 bp fragment resulting from control region (205 bp) and cytochrome b (208 bp) concatenated sequences.

We performed coalescent-based analyses using one dataset composed by the 413 bp concatenated fragment and three independent runs in order to test the robustness of the results. A total of 15 individuals of *A. sapidus* and 12 *A. terrestris* were used. Genetic diversity per locus was estimated using the software DNAsp v.4.50.3 (Rozas, Sánchez-DelBarrio, Messeguer *et al.*, 2003).

In order to compare coalescent-based and classical phylogenetic estimates of divergence dates we used the software MEGA v. 4.0.2 (Tamura, Dudley, Nei *et al.*, 2007) to calculate average and net nucleotide divergences between species. Standard errors were estimated by the bootstrap method using 10000 replicates. Divergence time was calculated by dividing average and net divergence estimates by  $2\mu$  (divergence rate). Confidence intervals of 95% were estimated as  $\pm 1.96$  standard errors of divergence rates.



## The IM model

We apply the Hey & Nielsen (2007) model to contemporary sequences to investigate the most likely demographic scenario for the split of one species (in this case, *A. cantianus*) into two (*A. sapidus* and *A. terrestris*). This model jointly estimates six demographic parameters: effective population size of the ancestral species (*A. cantianus*) ( $N_A$ ), of the current *A. sapidus* ( $N_1$ ) and *A. terrestris* ( $N_2$ ), the proportion of gene exchange between species after splitting ( $m_1$  and  $m_2$ ) and the time since splitting ( $t$ ). The model uses Metropolis-Hasting Markov Chain Monte Carlo (MCMC) simulations to find the optimal combination of values for these six demographic parameters that best fit our dataset. We use the Isolation with Migration model implemented in IMA (updated version 21/4/2008) (Hey & Nielsen, 2007) as described below.

We chose the Hasegawa-Kishino-Yano (HKY) (Hasegawa, Kishino & Yano, 1985) nucleotide substitution model and started with parameters recommended by the authors. We initially performed exploratory runs to define the minimum range of bounded priors for demographic parameters that ensure the incorporation of full marginal posterior probability densities. These exploratory runs are strongly recommended to obtain a considerable reduction on computational time when using MCMC based analysis. Since all these first runs frequently reported the most visited values of migration rates within the lowest interval (which includes 0 values) we accepted the model without gene flow between species. This modification from the original model reduced the number of parameters to be inferred from six to four and, consequently, decrease the uncertainty in the model and computational time. After exploration of several alternative sampling strategies of the parameter space, we finally used multiple simultaneous searches with 50 chains (-n 50, command line options) and a

geometric heating scheme ( $-f$  g). 1000 chains swaps were attempted ( $-k$  1000) in each of the  $3 \times 10^7$  steps ( $-L$  300000). To avoid taking into account parameters values sampled before stationarity was reached, we discarded the first  $3 \times 10^6$  steps (burn-in) ( $-b$  3000000). Convergence was evaluated with the autocorrelation of the parameters among distanced steps, update rates and Effective Sample Size (*ESS*). Scalars for population sizes for *A. sapidus* and *A. terrestris* ( $-q1$  and  $-q2$ ) were set to 20 whereas scalars for ancestral population size and maximum time for populations splitting ( $-qA$  and  $-t$ ) were set to 40. We performed three independent runs per dataset under the parameters fixed after the exploratory runs and different random seed numbers. The same exploratory parameters were used to assess the joint parameter estimate (LMODE). Estimates of model parameters were scaled to demographic and time units (years) assuming mutation rates previously used for the closely related *Microtus arvalis* (4% and 46% per million years for cytochrome *b* and control region, respectively) (Fink, Excoffier & Heckel 2004, Hamilton, Currat, Ray *et al.* 2005), which are high for vertebrates in general, but that match the high mutation rates usually reported for rodents (Heckel, Burri, Fink *et al.*, 2005; Rajabi-Maham, Orth & Bonhomme, 2008).

All IMA analyses were performed on a cluster server configuration composed of 80 servers DELL 750 with Pentium4 3.2GHz processor 1GB/2GB RAM, 2 servers DELL 1425 with Pentium Xeon64 3.2GHz and 3GHz processor 2GB RAM, 1 server HP Proliant DL145 with 2 AMD Opteron 1.8GHz processors 4GB RAM and 40 servers blade: 36 DELL PE 1955 with double quad-core Intel Xeon 5310 1.6GHz processor 4GB RAM, 4 servers blade DELL PE 1955 with double quad-core Intel Xeon 5355 2.66GHz processor 8GB RAM, all of them interconnected by Gigabit Ethernet at the Centro de Supercomputación de Galicia (CESGA), Spain.

## RESULTS

Our survey throughout Iberia and France allowed us to further characterize the mitochondrial genetic diversity of Southern and European water voles in this area. *Arvicola sapidus* showed higher nucleotide and haplotype diversities in control region than cytochrome b, while the opposite pattern was found in *A. terrestris* (table 1).

		N	S	$\pi$ (SD)	h	H (SD)	k
<i>A. sapidus</i>	Control Region	15	24	0.03707 (0.00297)	15	1.000 (0.024)	7.562
	Cytochrome b	15	4	0.00311 (0.00097)	5	0.562(0.143)	0.648
<i>A. terrestris</i>	Control Region	12	3	0.00784 (0.00113)	2	0.530 (0.076)	1.591
	Cytochrome b	12	10	0.00976 (0.00263)	6	0.803 (0.096)	2.030

**Table 1** Control Region and cytochrome b diversities for *A. sapidus* and *A. terrestris* sequences. *N*, sample size; *S*, segregating sites;  $\pi$ , nucleotide diversity; *SD*, standard deviation; *h*, number of haplotypes; *H*, haplotype diversity; *k*, mean number of nucleotide differences.

We did not detect evidences of gene flow between Southern and European water voles ( $m_1=m_2=0$ ), so this parameter was fixed to zero in subsequent runs. Parameters estimated with the IM model on the three independent replicates are shown in table 2.

	$N_1$	$N_2$	$N_A$	t
<b>Replicate 1</b>	81.833 (47.182, 147.866)	21.2483 (9.262, 48.053)	42 (0.217, 362.421)	26.42 (7.98, 33.9)
<b>Replicate 2</b>	81.833 (47.182, 147.866)	21.2483 (9.262, 48.053)	0.217 (0.217, 362.420)	25.74 (7.58, 33.94)
<b>Replicate 3</b>	81.833 (47.400, 148.084)	21.2483 (9.262, 48.053)	42.061 (0.217, 362.421)	25.86 (7.82, 33.78)

**Table 2** Estimates of model parameters and the 90% Highest Posterior Density (HPD) interval (between parentheses) for each of the three independent replicates for mtDNA.  $N_1$ , population size of *A. sapidus*.  $N_2$ , population size of *A. terrestris*.  $N_A$  population size of the common ancestor (*A.*

*cantianus*).  $t$  time since splitting of *A. cantianus* into both extant species of water voles. See text for translation into demographic and time units.

Each replicate reported slightly different results for the splitting parameter ( $t$ ) and similar estimates of effective population sizes. From  $t = 25.86$ ,  $\theta_1 = 81.8333$ ,  $\theta_2 = 21.2483$  and  $\theta_A = 42.0608$  (replicate 3; the combination showing the highest *ESS* for the parameter  $t$ ), we estimate the time since splitting of *A. sapidus* and *A. terrestris* around 252000 years [(76000 - 329000) estimated 90% Highest Posterior Density (HPD90) interval]. Effective population size was estimated in 199000 individuals (HPD90: 115000 - 361000) and 52000 individuals (HPD90: 23000 - 117000) for *A. sapidus* and *A. terrestris*, respectively. Ancestral population size was difficult to estimate, as expected in single locus estimates under reciprocal monophyly (Edwards & Beerli, 2000). Although we obtained a marked peak at the reported highest likelihood estimate, it was followed by a soft plateau falling down to the right of the parameter range of values (i.e. different combination of parameters would not substantially differ on their likelihoods), what results in wide confidence intervals for this estimate (e.g. HPD90: 531 - 883000).

The average pairwise sequence divergence was 12%, while the net nucleotide divergence (Da) between water vole species was 10.6%. The application of a mutation rate of 24.84% substitutions per million years (arithmetic mean of control region (46%) and cytochrome b (4%) mutation rates per site per million years) to these estimates date back the time since

splitting at *ca.* 241000 years before present (ybp) [95% confidence interval (CI) 182000 - 301000 ybp] and 213000 ybp (95% CI: 158000 – 269000 ybp) for average and net divergences, respectively.

## DISCUSSION

We have used a coalescent-based genetic approach to contribute to the knowledge of the evolutionary history of the genus *Arvicola*, focusing on the estimation of the time of divergence and the testing gene flow after the split of both extant water voles in Eurasia.

### Hybridization or gene flow after split?

The isolation with migration model discarded significant gene flow between species and, consequently, our results do not support the hypothesis of current hybridization nor the connection of species by gene flow after split. Díaz de la Guardia & Pretel (1979) proposed that the ancestral karyotype was modified by Robertsonian translocation and pericentric inversions, resulting in the different current karyotypes for both species (*A. terrestris*,  $2n= 36$ ; *A. sapidus*,  $2n= 40$ ). Interspecific differences on the number of homologous chromosomes might have caused postzygotic reproductive isolation of *A. sapidus* and *A. terrestris* by preventing proper chromosomal segregation during meiotic divisions in the hybrids. Furthermore, current habitat segregation might have also contributed to prezygotic isolation (see below).

### Time of divergence

Classical phylogenetic methods estimated the time of divergence of the common ancestor at 241000 ybp, and net divergence did only slightly correct this estimate

(213000 ybp), while our coalescent-based estimate dated this event *ca.* 252000 ybp. These estimates have largely overlapping confidence intervals, which are wider for IM estimates as this method is considering not only sampling variances, but also the stochasticity of coalescent processes. Molecular methods agree in dating the time of divergence of *Arvicola* species during the forth Mindel glacial (stage 8 in the Marine Isotope Stages) (*ca.* 252000 ybp, Middle Pleistocene) and indicate no subsequent gene flow occurred between these species. Although this suggests the emergence of some kind of reproductive isolation, we cannot distinguish between an allopatric fragmentation (as it would result from ice sheet advances toward Southern latitudes) and sympatric reproductive isolation (which might result from sudden chromosomal duplication/deletions). Moreover, since chromosomal rearrangements have been reported in other rodents during glaciations (e.g. *Sorex araneus*) (Searle, 1984), the speciation within the genus *Arvicola* might have been led by the combined effect of chromosomal incompatibilities and allopatry. However, the absence of *A. terrestris* fossils in central and Southern Iberia (Sesé Benito, 1994; Ruiz Bustos, 1999) and the deep divergences between interspecific lineages (Taberlet *et al.*, 1998) suggest a long-term isolation in allopatry until a more recent secondary contact.

### **Effective and census population sizes**

Estimates of population size of *A. sapidus* ( $\theta_1$ ) and *A. terrestris* ( $\theta_2$ ) slightly varied depending on the dataset. Thus, whereas population size for *A. sapidus* was estimated in 199000 individuals (HPD90: 115000 - 361000), a three to five times lower population size was inferred for *A. terrestris*, (52000 individuals; HPD90: 23000 – 117000). These differences in population size estimates reflect differences in nucleotide diversities reported for both species (table 1). It must be noted that sequences of *A. terrestris* were obtained from individuals captured in Northern regions of Iberia and France

(corresponding to *A. t. sherman*), the westernmost portion of the range of the species. As Eckert, Samis & Loughheed (2008) showed, nucleotide diversities can be reduced in marginal areas of distribution with respect to central areas. Consequently, since the IM model is based on intraspecific nucleotide diversities, the effective population size inferred for *A. terrestris* might be applied to the population from Northern Iberia and France and should not be considered as an estimate of the overall population size for the species throughout Eurasia.

It must be remembered that effective populations size usually differ greatly from census size (Frankham, 1996). According to Borkowska & Ratkiewicz (2008), the average ratio of effective/actual population sizes in another Cricetidae (*Microtus arvalis*) ranges from 0.18 to 0.35 (considering the effects of fluctuations in population sizes, variance in family sizes and unequal sex ratios). Assuming similar ratios for water voles, effective population sizes would translate into current population sizes ranging from 569000 to 1105000 individuals for Southern water vole and 149000 to 289000 for European water voles in Iberia and France.

#### **Paleontological and genetic estimates of divergence: a brief history of *Arvicola***

Evolutionary history of water voles has been classically reconstructed by using morphological traits such as body size and length and the enamel band in the first molar (e.g. see Ruiz-Bustos (1997) and Escude et al. (2008) and references therein). These authors coincide on giving a wide timeframe for the origin of *A. sapidus* around the Middle Pleistocene -i.e. 780-130 ky- (López Martínez, 1980; Sesé, 1986; Sesé Benito, 1994; Arribas, 1995; Ruiz-Bustos, 1997; Sesé Benito & Soto, 2005). Our results estimates fall within this broad interval, and further delimit the Mindel glacial as the period in which some population of the common ancestor to both species became

reproductively and genetically isolated within the Iberian Peninsula and started to diverge from the remaining populations in Europe. Consequently, broad time intervals for species divergence based on body, molar and enamel sizes may be attributed in part to the difficulty in delineating recently diverged taxa within this genus based only on morphology. In fact, lack of concordance among morphological and genetic criteria is well represented within the genus *Arvicola*. Differences in coat colour patterns have been used to delineate subspecies within Southern water vole, but there is no correspondence with genetic divergences (Centeno-Cuadros, Delibes & Godoy, *in press*). Besides, a great phenotypic (Escude *et al.*, 2008) and ecological plasticity (Folitarek, 1971) in *A. terrestris* are also apparently hampering a consensus on the delimitation of subspecific taxa.

This study illustrates the utility of classical phylogenetic and coalescent-based models for the estimation of divergence times in *Arvicola* at interspecific levels. We are deeply confident that the use of integrative approaches at intraspecific levels merging ecological and morphometric analyses and isolation-with-migration models will help to disentangle the evolutionary history of *Arvicola* and other species.

## ACKNOWLEDGMENTS

This work has been performed with the invaluable fieldwork assistance, sample request, bibliography and information provided and help from Jacinto Román. The authors are also indebted to all biologists who provided samples for the study. We are very grateful to CESGA (Centro de Supercomputación de Galicia), where all the computational analyses were performed. This work has been funded by the Dirección General de



Investigación (project BOS2001-2391-C02-01). ACC benefited from a FPU fellowship from the Ministerio de Educación y Ciencia.

## REFERENCES

- Arbogast, B. S., Edwards, S. V., Wakeley, J., Beerli, P. & Slowinski, J. B. (2002) Estimating divergence times from molecular data on phylogenetic and population genetic timescales. *Annu. Rev. Ecol. Syst.*, **33**, 707.
- Arribas, A. (1995) Consideraciones cronológicas, tafonómicas y paleoecológicas del yacimiento cuaternario de Villacastín (Segovia, España). *Boletín Geológico y Minero*, **106**, 3.
- Borkowska, A. & Ratkiewicz, M. (2008) Sex-related spatial structure and effective population size in the common vole, *Microtus arvalis*, as revealed by mtDNA analysis. *Ann. Zool. Fenn.*, **45**, 255.
- Centeno-Cuadros, A., Delibes, M. & Godoy, J. A. (*in press*) Phylogeography of Southern Water Vole (*Arvicola sapidus*): evidence for refugia within the Iberian glacial refugium? *Mol. Ecol.*
- Díaz de la Guardia, R. & Pretel, A. (1979) Comparative study of the karyotypes of two species of water vole: *Arvicola sapidus* and *Arvicola terrestris* (Rodentia, Microtinae). *Caryologia*, **32**, 183.
- Eckert, C. G., Samis, K. E. & Loughheed, S. C. (2008) Genetic variation across species' geographical ranges: the central-marginal hypothesis and beyond. *Mol. Ecol.*, **17**, 1170.
- Edwards, S. V. & Beerli, P. (2000) Perspective: Gene divergence, population divergence, and the variance in coalescence time in phylogeographic studies. *Evolution*, **54**, 1839.
- Escude, E., Montuire, S., Desclaux, E., Quere, J. P., Renvoise, E. & Jeannet, M. (2008) Reappraisal of 'chronospecies' and the use of *Arvicola* (Rodentia, Mammalia) for biochronology. *J. Archaeol. Sci.*, **35**, 1867.
- Folitarek, S. S. (1971) *Ecology of the water vole and its control in West Siberia*. The U.S. Department of the Interior and the National Science Foundation, Washington D.C., by the Al Ahram Center for Scientific Translations.
- Frankham, R. (1996) Relationship of genetic variation to population size in wildlife. *Conserv Biol*, **10**, 1500.
- Fraser, D. J. & Bernatchez, L. (2001) Adaptive evolutionary conservation: towards a unified concept for defining conservation units. *Mol. Ecol.*, **10**, 2741.
- Hasegawa, M., Kishino, H. & Yano, T. (1985) Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *J. Mol. Evol.*, **22**, 160.
- Heckel, G., Burri, R., Fink, S., Desmet, J. F. & Excoffier, L. (2005) Genetic structure and colonization processes in European populations of the common vole, *Microtus arvalis*. *Evolution*, **59**, 2231.
- Hey, J. & Nielsen, R. (2004) Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. *Genetics*, **167**, 747.

- Hey, J. & Nielsen, R. (2007) Integration within the Felsenstein equation for improved Markov chain Monte Carlo methods in population genetics. *Proc. Natl. Acad. Sci. U. S. A.*, **104**, 2785.
- Knowles, L. L. & Maddison, W. P. (2002) Statistical phylogeography. *Mol. Ecol.*, **11**, 2623.
- Kurtén, B. (1968) *Pleistocene Mammals of Europe*. London: Weidenfeld and Nicolson.
- López Martínez, N., (1980) Los micromamíferos (Rodentia, Insectivora, Lagomorpha y Chiroptera) del sitio de ocupación achelense de áridos-1 (Arganda, Madrid). In: *Ocupaciones achelenses en el valle del Jarama*: 161. Diputación Provincial de Madrid.
- Maddison, W. P. (1997) Gene trees in species trees. *Syst. Biol.*, **46**, 523.
- Montoya, P., Alberdi, M. T., Barbadillo, L. J., van der Made, J., Morales, J., Murelaga, X., Peñalver, E., Robles, F., Ruiz-Bustos, A., Sánchez, A., Sanchiz, B., Soria, D. & Szyndlar, Z. (2001) Une faune très diversifiée du Pléistocène inférieur de la Sierra de Quibas (province de Murcia, Espagne). *C.R. Acad. Sci. Paris, Sciences de la Terre et des planètes*, **332**, 387.
- Moritz, C. (1994) Defining Evolutionarily-Significant-Units for Conservation. *Trends Ecol. Evol.*, **9**, 373.
- Nei, M. & Li, W. H. (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proc. Natl. Acad. Sci. U. S. A.*, **76**, 5269.
- Nielsen, R. & Wakeley, J. (2001) Distinguishing migration from isolation: A Markov Chain Monte Carlo approach. *Genetics*, **158**, 885.
- Piertney, S. B., Stewart, W. A., Lambin, X., Telfer, S., Aars, J. & Dallas, J. F. (2005) Phylogeographic structure and postglacial evolutionary history of water voles (*Arvicola terrestris*) in the United Kingdom. *Mol. Ecol.*, **14**, 1435.
- Rajabi-Maham, H., Orth, A. & Bonhomme, F. (2008) Phylogeography and postglacial expansion of *Mus musculus domesticus* inferred from mitochondrial DNA coalescent, from Iran to Europe. *Mol. Ecol.*, **17**, 627.
- Rozas, J., Sánchez-DelBarrio, J. C., Messeguer, X. & Rozas, R. (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Methods of biochemical analysis*, **19**, 2496.
- Ruiz Bustos, A. (1999) The role of the water voles (*Arvicola*, Rodentia) in the Quaternary. *Estudios Geológicos*, **55**, 89.
- Ruiz Bustos, A., (2002) Enamel line molar analysis in Arvicolid rodent and its potencial use in biostratigraphy and palaeoecology.
- Ruiz-Bustos, A., (1997) Características biostratigráficas y paleoecológicas que implican los mamíferos cuaternarios en las cuencas de la cordillera Bética. In: *Cuaternario Ibérico*: 283. J. Rodríguez Vidal (Ed.). Asociación Española para el Estudio del Cuaternario (AEQUA), Huelva.
- Searle, J. B. (1984) Three new karyotypic races of the common shrew *Sorex araneus* (Mammalia, Insectivora) and a phylogeny. *Syst. Zool.*, **33**, 184.
- Sesé Benito, C. (1994) Paleoclimatical interpretation of the Quaternary small mammals of Spain. *Geobios*, **27**, 753.
- Sesé Benito, C. & Soto, E., (2005) Mamíferos del yacimiento del Pleistoceno Medio de Ambrona (Soria, España): Análisis faunístico e interpretación paleoambiental. In: *Los yacimientos paleolíticos de Ambrona y Torralba (Soria). Un siglo de investigaciones arqueológicas*: 258.
- Sesé, C. (1986) Insectívoros, roedores y lagomorfos (Mammalia) del sitio de ocupación achelense de Ambrona (Soria, España). *Estudios geol.*, **42**, 355.

- Stanley, S. M. (1978) Chronospecies Longevities, Origin of Genera, and Punctuational Model of Evolution. *Paleobiology*, **4**, 26.
- Taberlet, P. & Fumagalli, L. (1996) Owl pellets as a source of DNA for genetic studies of small mammals. *Mol. Ecol.*, **5**, 301.
- Taberlet, P., Fumagalli, L., Wust-Saucy, A.-G. & Cosson, J.-F. (1998) Comparative phylogeography and postglacial colonization routes in Europe. *Mol. Ecol.*, **7**, 453.
- Tamura, K., Dudley, J., Nei, M. & Kumar, S. (2007) MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol. Biol. Evol.*, **24**, 1596.
- 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.
- Wilson, A. C., Cann, R. L., Carr, S. M., George, M., Gyllensten, U. B., Helmbychowski, K. M., Higuchi, R. G., Palumbi, S. R., Prager, E. M., Sage, R. D. & Stoneking, M. (1985) Mitochondrial DNA and two perspectives on evolutionary genetics. *Biol. J. Linn. Soc.*, **26**, 375.
- Wust Saucy, A.-G., (1998) Polymorphisme genetique et phylogeographie du campagnol terrestre *Arvicola terrestris*. In: *Faculte des Sciences*. Universite de Lausanne, Lausanne.
- Wust Saucy, A. G., Hausser, J., Taberlet, P. & Saucy, F. (1998) Phylogeography of the vole *Arvicola terrestris* as revealed by mtDNA: The role of historical factors? *Zoologia et Botanica*, 764.
- Zink, R. M. & Barrowclough, G. F. (2008) Mitochondrial DNA under siege in avian phylogeography. *Mol Ecol*, **17**, 2107.
- Zuckerandl, E. & Pauling, L., (1965) Evolutionary divergence and convergence in proteins. In: *Evolving Genes and Proteins*. V. Bryson & V. HJ (Eds.). New York.

## **CHAPTER 3**



## CHAPTER 3

### **Phylogeography of Southern Water Vole (*Arvicola sapidus*): evidence for refugia within the Iberian glacial refugium?**

Alejandro Centeno-Cuadros<sup>1</sup>, Miguel Delibes<sup>1</sup> and José Antonio Godoy<sup>2</sup>

<sup>1</sup> *Conservation Biology Department, Estación Biológica de Doñana (EBD-CSIC). C/ Américo Vespucio, s/n. CP 41092. Sevilla (Spain)*

<sup>2</sup> *Integrative Ecology Department, Estación Biológica de Doñana (EBD-CSIC). C/ Américo Vespucio, s/n. CP 41092. Sevilla (Spain)*

**Keywords:** mtDNA, control region, cytochrome *b*, non-invasive sampling, Pleistocene, postglacial expansion

**Running title:** Phylogeography of *A. sapidus*



## RESUMEN

A pesar de que ya está aceptado el papel de las penínsulas meridionales europeas como refugios glaciares para especies de climas templados, sólo recientemente se están estudiando los patrones filogeográficos dentro de refugios. En este trabajo describimos los patrones filogeográficos para la rata de agua meridional (*Arvicola sapidus*) en todo su área de distribución a lo largo de Iberia y Francia. Hemos obtenido secuencias de región control y citocromo b de 228 muestras distribuidas en 130 localidades por Iberia y Francia. En total, hemos encontrado 85 haplotipos con una elevada diversidad mitocondrial ( $\pi = 0.027$ ;  $H = 0.974$ ). La estructura filogeográfica era relativamente débil (divergencia intraespecífica media 3.1%) con unos pocos clados soportados y redes de máxima parsimonia no conectadas al 95 y 90%, aunque significativas, como queda reflejado en un aumento las divergencias nucleotídicas con la distancia ( $r = 0.197$ ,  $p = 0.03$ ) y una autocorrelación significativa hasta aproximadamente 500 kms. Mediante análisis espaciales de varianza molecular hemos localizado siete grupos geográficos que explican el 43.73% de la variación mitocondrial total. En tres de ellos hemos detectado una señal de expansión demográfica. En nuestro estudio, proponemos una colonización reciente de Francia desde Iberia hace unos 62.000 años, empleando para ello un modelo de aislamiento con migración. Nuestros resultados sugieren la contribución de episodios de aislamiento en subrefugios glaciales dentro de Iberia, aunque parecen excluir un aislamiento continuo a largo plazo durante sucesivos ciclos glaciales. La divergencia filogeográfica se ha visto reducida probablemente por los elevados tamaños poblacionales de la especie y periodos de mezcla rápida y prolongados entre subrefugios durante los periodos interglaciales, borrando la estructura filogeográfica acumulada en los picos glaciales. Las diferencias fenotípicas en *A.sapidus* no reflejan aislamientos históricos intraespecíficos y no garantizan por tanto las delimitaciones subespecíficas propuestas. Nuestros resultados apoyan la existencia de subrefugios glaciales dentro de Iberia y su papel en el fomento de divergencias intraespecíficas.





## ABSTRACT

The role of Southern European peninsulas as glacial refugia for temperate species has been widely established, but phylogeographic patterns within refugia are being only recently addressed. Here we describe the phylogeographic patterns for Southern water vole (*Arvicola sapidus*) in its whole distribution across Iberia and France. Control region and cytochrome b sequences were obtained for 228 samples from 130 localities across Iberia and France. 85 haplotypes were found in total yielding a high overall mitochondrial diversity ( $\pi = 0.027$ ;  $H = 0.974$ ). Phylogeographic structure was relatively shallow (3.1% average intraspecific divergence) with few supported clades and 95 and 90% maximum parsimony unconnected networks, but significant, as reflected in increased pairwise nucleotide divergences with distance ( $r = 0.197$ ,  $p = 0.03$ ) and significant autocorrelation up to *ca.* 500 Kms. SAMOVA analysis detected seven geographic groups explaining 43.73% of the total mitochondrial variation. We detected demographic expansions in three of these groups. A recent colonization of France from Iberia was suggested and estimated around 62.000 ybp by an isolation-with-migration model. Our results suggest the contribution of episodes of isolation in glacial subrefugia in Iberia, but seem to exclude a long-term isolation over successive glacial cycles. Phylogeographic divergence was probably tempered by relatively large population sizes and rapid and extensive mixing among subrefugia during interglacials, that might have eroded the phylogeographic structure accumulated at glacial peaks. Phenotypic differences in *A. sapidus* do not delineate historically isolated intraspecific divisions and do not warrant subspecific delimitations. Our results do support the existence of subrefugia within Iberia and their role in promoting intraspecific divergences.



## INTRODUCTION

The combined analysis of geographical and genealogical information provides insights on the historical processes that have conditioned species current distribution and genetic variation (Avice, 2000). This kind of approach has, for example, revealed the predominant effect of glaciations in shaping contemporary genetic patterns across species in Europe (Hewitt, 2000; Taberlet *et al.* 1998). Three main Southern European ice-free glacial refugia (Iberia, Italy and the Balkans) and also central and Northern (Fennoscandia) refugia (Taberlet *et al.* 1998) have promoted intraspecific divergences and have differentially acted as sources for the recolonization of Europe for different species. For some species, the Pyrenees and the Alps seem to have acted as effective barriers to the recolonization from Iberian and Italian peninsulas, respectively. Most studies, being focused on the divergence among refugia have, however, paid little attention to the distribution of genetic variation within refugia, assuming it to be homogeneous.

However, recent studies suggest that glacial refugia are heterogeneous mosaics of climatic and ecological regions that have shaped complex phylogeographic patterns. Furthermore, concordant distributions of divergent intraspecific lineages across taxa have led to the proposition of the “refugia-within-refugia” concept (Gómez & Lunt, 2007; Weiss & Ferrand, 2007). At the species level, these subrefugia usually harbour significantly differentiated intraspecific lineages, thus making an important contribution to overall species diversity. At a broader scale, glacial refugia become hotspots of biodiversity and sources of endemism, as new taxa are originated within a relative reduced region.

The Iberian Peninsula harbours a high physiographic complexity (East-Western oriented mountainous chains) that, in combination with concurrent Mediterranean and Atlantic influences, results in an appropriate area for the study of the “refugia-within-refugia” concept.

Most recent studies on Iberian phylogeography have involved amphibians (Martinez-Solano *et al.* 2006; Sequeira *et al.* 2008) and reptiles (Godinho *et al.* 2008; Pinho *et al.* 2007) (see Gómez & Lunt (2007) for examples in other organisms). These taxa are usually characterized by low dispersal capacities what have favoured the long-term isolation of their populations through successive glacial cycles (Plio-Pleistocene) and the consequent evolution of deep divergences within Iberia. In contrast, the effect of glacial subrefugia on organisms with more complex population dynamics and higher dispersal capacities have been seldom addressed (e.g. Vilà *et al.* 1999). These studies might not only help to evaluate the relative importance of these areas for the conservation of biodiversity, but might also significantly add to our understanding of the complex interactions between life-history traits, demographic history and genealogical patterns.

Rodents are good phylogeographic models because they often provide an abundant fossil register, inspiring hypothesis on their evolutionary history and adaptative evolution. These hypotheses can now be tested with molecular data (Tougaard & Renvoise, 2008). The genus *Arvicola* includes two species of water voles in Eurasia. *Arvicola terrestris* is the most widespread species, extending from the North of the Iberian Peninsula to the East of Asia, while the Southern water vole (*Arvicola sapidus*) is restricted to the Iberian Peninsula and France. The origin of the latter species has been located in Iberia during the Pleistocene using fossil records (Sesé Benito, 1994). Phylogenetic analysis showed these two species to be reciprocally monophyletic and dated their divergence during the Middle Pleistocene (Taberlet *et al.* 1998; Wust Saucy, 1998; Wust Saucy *et al.* 1998). Phylogeographic (Piertney *et al.* 2005; Wust Saucy *et al.* 1998) and population genetic studies using neutral (Aars *et al.* 2006; Berthier *et al.* 2006; Berthier *et al.* 2005; Stewart *et al.* 1999; Telfer *et al.* 2003) and non-neutral markers (Bryja *et al.* 2007; de Bellocq *et al.* 2008; Oliver *et al.* 2009; Oliver & Piertney, 2006; Tollenaere *et al.* 2008) have contributed to the understanding of historical and

current factors promoting genetic structure of *A. terrestris*. Three less divergent genetic lineages have been found within *A. terrestris*, corresponding to Italian, Balkan and Fennoscandian refugia (Taberlet *et al.* 1998). In contrast, genetic studies with *A. sapidus* have only covered cytogenetics and chromosomal characterization (Diaz de la Guardia & Pretel, 1978; Megías-Nogales *et al.* 2003) and phylogeographic analysis of *A. sapidus* are lacking.

Shallow phylogeographic structure for *A. sapidus* is expected if a single large population survived in Iberia and did not undergo any fragmentation during glacial maxima. However, morphological differences among populations of *A. sapidus* suggest local adaptation or genetic structure and belie a simple model of a single population in a single Iberian refugium. Two main morphological differences have been reported: i) a reduction of 20-25% in body size and weight of individuals from a Southwest population (Natural region of Doñana) (Román, 2007) and ii) differences in coat colour which have prompted a controversial subspecific division into *A. sapidus tenebricus* (darker, in the North of Iberia) and *A. sapidus sapidus* (brown coloured coat, widespread over Iberia and France) (Ventura, 2002).

In this study we analyze the geographic distribution of mitochondrial genetic variation in Southern water voles. We relied mostly on bones sampled in raptor pellets, which proved adequate sources for the characterization of two short but polymorphic mitochondrial markers (control region and cytochrome *b*), but poor materials for additional nuclear DNA analyses. Mitochondrial DNA patterns have nevertheless been shown to be robust indicators of patterns of population history, and the markers most suited for the analyses of recently isolated populations (Zink & Barrowclough, 2008). We test whether an intraspecific mitochondrial structure exists and whether it is consistent with historical isolation in subrefugia within the Iberian Peninsula and evaluate its concordance with morphological and coat colour differences. We also dated major events like the colonization of France from Iberia.

## MATERIAL AND METHODS

### Sampling and DNA extraction

Samples from 299 Southern water voles covering the whole distribution area of the species and from 12 *A. terrestris* (from the North region of Iberia), which are included here as outgroup, were obtained mostly as donations from biologists and scientific collections. These samples consist mainly of bones (jaws, molars or incisors) obtained from owl and eagle owl pellets and fixed tissues from museums, although fresh tissues were used when available. All bones collected were morphologically identified as *A. sapidus* by the authors (Dueñas & Peris Álvarez, 1985). Multiple sampling of the same individual is highly unlikely since only a single tooth or mandible were used per pellet and individual water voles are always predated by a single owl.

Because of the variability in amplification success between individuals and fragments, 273 individuals were included in the data set for D-Loop genetic variation, whereas 254 individuals were used for cytochrome *b* analysis. For phylogeographic analysis we considered only those 228 individuals for which both fragments were sequenced. These individuals were sampled at 130 different localities distributed among 38 counties throughout Iberia and France (Table 1).

**Table 1** Geographic and sample type distribution of the samples used in this study. A total of 228 samples were collected from 130 geographic localities. The number of samples corresponding to fresh tissue, museum and hair samples is marked between parentheses, square brackets or with an asterisk, respectively. The remaining samples are bones obtained from owl-pellets.

	Map. Ref.	County	Localities	Sample size	Sample type	Latitude	Longitude
<b>SPAIN</b>	1	Ávila	2	2	1, (1)	40.4271	-5.3032
	2	Barcelona	2	2	[2]	41.2985	1.9937
	3	Badajoz	1	2	2	38.8666	-6.4274
	4	Burgos	9	9	6, (1), [2]	42.1729	-3.7077
	5	Cádiz	1	1	(1)	36.0616	-5.6363
	6	Cáceres	2	2	2	40.1057	-5.7217
	7	Castellón	4	8	(8)	40.7544	-0.1635
	8	Gerona	2	3	2, [1]	41.8490	2.3902
	9	Granada	1	1	1	36.9728	-3.4531
	10	Guadalajara	1	2	2	41.2064	-2.6023
	11	Huelva	15	15	5, (10)	36.9876	-6.4814
	12	Jaén	1	1	1	37.9108	-3.0024
	13	Lérida	1	1	[1]	41.8950	1.0475
	14	La Rioja	3	3	3	42.3269	-3.0377
	15	Lugo	1	1	1	42.5903	-7.1923
	16	Navarra	4	7	7	42.7539	-1.0998
	17	Pontevedra	4	9	9	42.7961	-8.1686
	18	Santander	2	2	2	42.9892	-3.9710
	19	Sevilla	2	2	1, (1)	37.9337	-5.7621
	20	Soria	2	2	2	41.6100	-3.4016
	21	Tarragona	14	19	18, [1]	41.2312	0.7172
	22	Teruel	3	3	3	40.8541	-1.0960
	23	Valladolid	3	5	5	41.5580	-5.0043
	24	Zaragoza	6	7	7	41.6071	-0.7568
	25	Zamora	1	1	1	41.8509	-5.6129
<b>FRANCE</b>	26	Alpes Maritimes	1	1	(1)	43.7738	6.7857
	27	Auvergne	14	14	(14)	45.7344	2.6273
	28	Drôme	1	2	1, 1*	44.6658	5.7063
	29	Creuse	1	5	5	45.8736	1.6440
	30	Corrèze	2	17	17	45.2635	1.6775
	31	Haute-Vienne	1	9	8, 1*	45.8053	0.9337
	32	Vendée	1	1	1	46.9674	-2.2635
	33	Pyrénées-Orientales	3	5	5	42.8332	2.9191
<b>PORTUGAL</b>	34	Beja	1	4	4	38.2041	-7.2327
	35	Évora	3	12	12	38.7047	-7.4000
	36	Faro	1	4	4	37.4394	-7.7670
	37	Setúbal	13	42	(42)	37.9383	-8.7766
	38	Vila Real	1	2	2	41.2826	-7.9041



Bones were frozen under liquid nitrogen and homogenized to powder with a ball-mill (RETSCH Mod. MM301). DNA was extracted using a guanidinium-silica protocol (method C in Rohland & Hofreiter (2007)). Museum pelts were extracted with a standard phenol-chloroform protocol, following extensive washes in TE9 (500mM Tris pH 9, 20mM EDTA, 10mM NaCl, pH 9) (Goetz *et al.* 1985) to remove inhibiting salts (Shiozawa *et al.* 1992). Both types of material were extracted following ancient DNA guidelines in a dedicated lab. Contamination with modern DNA or PCR products was monitored by including two extraction blanks in every extraction round. Finally, DNA from fresh tissues was extracted with a “salting-out” protocol (Müllenbach *et al.* 1989) including two successive LiCl precipitation and chloroform extractions for protein removal.

### **Mitochondrial DNA sequencing**

D-Loop and cytochrome *b* markers were sequenced in this study. As endogenous DNA available in most samples was limited and degraded, we used forward F15468 (5'-GCATTAAATTATATTCCCCATGC-3') and reverse R15713 (5'-TTGTTGGTTTCACGGAGGAT-3') (D-Loop) and forward F14559 (5'-TCCTTTTGAGGGGCTACAGT-3') and reverse R14806 (5'-TGGAAGGGAATTTTGTCTGC-3') (cytochrome *b*) primers (Centeno-Cuadros and Godoy, unpublished data) to amplify short and maximally informative fragments. Polymerase Chain Reaction (PCR) was performed in a final volume of 25µL and consisted of 5 µL of extract, 2mM MgCl<sub>2</sub>, 0.25mM dNTPs, 0.25mM forward and reverse sequence-specific primers, 1x Taq polymerase buffer and 1 U/ml *Taq* (BIOTAQ™ DNA Polymerase, Bioline). 0.8 mg/ml and 0.1 mg/ml BSA (Bovine Serum Albumine, Roche) was added to prevent inhibition in bones-museum and fresh tissue samples, respectively. PCRs were performed in a MJ

research PT200 thermocycler programmed for an initial denaturation at 94°C for 5min, followed by 40 cycles of denaturation at 92°C for 30 s, annealing at 52°C for 30 s and extension at 72°C for 30 s. All reactions were finished with a final extension at 72°C for 5 min. To control the performance of the process and monitor for contamination, positive (DNA obtained from fresh tissue) and negative (water) DNA controls, respectively, were included with each set of PCR reactions. 5 µL of PCR products were incubated with 2µL ExoSAP-IT enzyme (Exonuclease I and Shrimp Alkaline Phosphatase in buffer) (USB Corporation) for 15 minutes at 37°C and inactivated for 15 minutes at 80°C. Fragments were bidirectionally sequenced using the Applied Biosystems BigDye® Terminator Cycle Sequencing Kit v. 1.1 following manufacturer's instructions and the same primers used for the amplification. Reactions were analysed in an Applied Biosystems 3130xl Genetic Analyzer. Forward and reverse sequences for each PCR product were edited and assembled using Sequencher 4.6 (Gene Codes Corporation, Ann Arbor, MI).

## **Genetic data analyses**

### Genetic diversity

Nucleotide and haplotype diversities were estimated for control region and cytochrome *b* fragments separately (see *Sampling and DNA extraction*) and analysed with DnaSP 4.50.3 (Rozas *et al.* 2003). Since both fragments belong to the same DNA molecule, all subsequent analyses were performed with both sequences per individual concatenated in two partitions: control region from nucleotide position 1 to 205 and cytochrome *b* from nucleotide position 206 to 413.

As a first approach to detect a signal for northward range expansion, we tested the existence of a gradient in nucleotide diversity over geographic localities, a pattern that might arise from a series of founding events in a leading-edge colonization process (Ibrahim *et al.* 1996). We estimated nucleotide diversity in geographic localities with more than 5 sampled individuals and compared average nucleotide diversities in Iberia and France by using the Welch approximate t-test to account for the possible difference in variance among samples.

### Phylogeny

Control region and cytochrome *b* sequences were aligned using Clustal W (Thompson *et al.* 1994) as implemented in MEGA4 (Tamura *et al.* 2007). The optimal nucleotide substitutions model was selected among 56 models tested for control region (K81uf+I+G, with unequal base frequencies (uf), the proportion of invariable sites (I) equalling 0.29 and gamma distribution shape parameter ( $\alpha$ ) of 0.36), cytochrome *b* (HKY+G,  $\alpha=0.33$ ) and both sequences concatenated (K81uf, I= 0.56, G= 0.45), based on the Akaike Information Criterion (Posada & Buckley, 2004) as implemented in MODELTEST (Posada & Crandall, 1998). We employed a Bayesian phylogenetic inference approach using a variant of Markov Chain Monte Carlo implemented in MRBAYES 3.1.2 (Huelsenbeck & Ronquist, 2001), with three favoured partitions (control region: nucleotide positions 1 to 205; cytochrome *b*: nucleotide positions 206 to 413; encoded gaps: sequence positions 414 to 416) and allowing for differences in evolutionary models and mutation rates (control region: nst=6, rates=gamma; cytochrome *b*: nst=2, rates=gamma). Gaps were considered as a fifth state and coded with GAPCODER (Young & Healy, 2003). Three independent runs with 10 chains were performed simultaneously, using a heating scheme (temp=0.30). The length of the run was automatically adjusted to that required to reduce the standard deviation of split frequencies

below 0.005 (stopval = 0.005). Consensus trees with average branch lengths and parameter estimates were obtained from the samples remaining after discarding the first 25% of saved states as burnin.

### Phylogeography

Maximum Parsimony network construction methods establish a statistical criterion (statistical parsimony) by which every haplotype that do not exceed a parsimony limit is joined into a single network, leaving as unconnected networks those haplotypes that differ by more than the statistically determined threshold. We favoured this statistical procedure that divided the overall inferred genetic diversity into significantly different units versus algorithms which tend to join into a single network every haplotype independently of the amount of nucleotide differences among them. For this purpose we used the software TCS v.1.21 (Clement *et al.* 2000) and both 90 and 95% parsimony connection limits (MP90 and MP95, respectively). Thus, while unconnected MP95 networks are usually interpreted as intraspecific genetic lineages resulting from long-term isolation by barriers (allopatric fragmentation (Templeton, 2004)) we used differences in connection limits between MP90 and MP95 networks to test the strength of this genetic structure, hypothesizing that patterns which result from long-term isolation and are strongly differentiated must not be importantly affected by slight changes in parsimony limits.

Patterns of isolation by distance in the data were explored using two complementary approaches. ARLEQUIN 3.0 was used to compute a Kimura 2-Parameter corrected pairwise nucleotide distance matrix among haplotypes and its correlation with a geographic distance matrix was evaluated with Mantel tests. On the other hand, we performed an autocorrelation

analysis using II, an Autocorrelation Index for DNA Analysis (AIDA) analogous to Moran's I, which incorporates the nucleotide distances among sequences (Bertorelle & Barbujani, 1995). Class boundaries to geographic distances were chosen according to an equal frequencies criterion, i.e. number of pairwise distances was divided into classes in order to obtain an equal number of comparisons in each class. A subset of 200 randomly chosen individuals was selected to perform this analysis because of computational limitations.

### Genetic structure

We further explored global genetic structure in the species using Spatial Analysis of Molecular Variance (SAMOVA) (Dupanloup *et al.* 2002). This method defines  $k$  continuously homogenous but genetically differentiated populations where the proportion of total genetic variance ( $\Phi_{CT}$ ) is maximized. Since  $\Phi_{CT}$  increases as  $k$  increases we run SAMOVA with 1000 simulated annealing processes for a range of  $k$  values from 2 to 10. For this and other analysis based on allele frequencies we made geographical groupings of samples based on counties, and excluded those with less than five samples.

We estimated pairwise  $\Phi_{ST}$  among counties using Kimura 2-parameters genetic distance, as implemented in ARLEQUIN 3.0 (Excoffier *et al.* 2005), and tested its correlation with the log of geographic distance with Mantel tests after linearization of  $\Phi_{ST}$  [ $\Phi_{ST}/(1-\Phi_{ST})$ ] as recommended by Rousset (1997). We also tested the role of Pyrenees as a barrier to population postglacial expansion of water voles once discounting the effects of Euclidean distance through Partial Mantel test, using a binary matrix of pairs of spatial points separated by the Pyrenees (1) or not (0).

## Demography

Historical demography was analysed on those partitions suggested by SAMOVA. Tajima's D (Tajima, 1989) measures deviations from the correlation between the number of segregating sites and the average number of nucleotide differences expected under a model of neutral evolution; it was estimated with DNASP 4.50.3 (Rozas *et al.* 2003). Since Tajima's D may be affected by other demographic and evolutionary processes and depends on factors as sample size, length of sequence and time of sampling relative to a demographic event and its strength (Depaulis *et al.* 2003; Pannell, 2003; Simonsen *et al.* 1995), we also tested for historical population expansion with mismatch distribution analysis (Marjoram & Donnelly, 1994). A multimodal distribution of pairwise differences is expected from a stationary population whereas an unimodal distribution would be shown if the population have experienced a demographic expansion (Rogers & Harpending, 1992). We compared the observed frequency distribution of pairwise nucleotide differences with their expected distribution in a simulated population expansion using DNASP 4.50.3. Moreover, we estimated raggedness indexes to measure the smoothness of mismatch distribution. Therefore, non-significant deviations of expected and observed mismatch distributions and non-significant or significant but low values of raggedness indexes would indicate a past population expansion. The Rogers & Harpending (1992) model was used to calculate the time since population expansion by estimating Tau ( $\tau$ ) based on the mismatch distribution outputs from ARLEQUIN 3.0. We converted the parameter  $\tau$  to estimate the time since the expansion ( $t$ ) using the equation  $t = \tau / 2u$ , where  $u$  is the mutation rate per sequence per generation. We used per site substitution rates previously reported for rodents cytochrome b (4%) and control region (46%) (Fink *et al.* 2004; Hamilton *et al.* 2005), and a generation time of six months, what results in a mutation

rate per generation for the concatenated sequence of  $u = 5.13 \times 10^{-5}$  (Rogers, 1995; Schneider & Excoffier, 1999).

The time of colonization of France from Iberia was estimated by fitting an Isolation with Migration model to our genetic data, as implemented in IMA (updated version 21/4/2008) (Hey & Nielsen, 2007). Three independent runs of 20 chains and a geometric heating scheme were used to explore the posterior probability space. Ten millions genealogies were sampled of which the first 10% were discarded as burn-in (Mmode), which were then used to assess the joint parameter estimate (Lmode) using the same exploratory parameters. Run parameters and priors were adjusted until stationarity and convergence were reached, as assessed by the autocorrelation of the parameters among independent points ( $< 0.01$ ), large effective sample sizes ( $ESS > 100$ ), and high update rates ( $> 2\%$ ). Lower updates rates might be acceptable when the other criteria are met and independent runs yield concordant estimates (J. Hey, personal communication). Estimates of model parameters were scaled to demographic and year units by considering substitution rates per site per million years of 4% and 46 % for cytochrome *b* and control region, respectively (Fink *et al.* 2004; Hamilton *et al.* 2005).

## RESULTS

### Genetic diversity

A total of 254 and 273 samples were sequenced for cytochrome *b* and control region, respectively. Of these 158 and 140 were obtained from bones collected in owl-pellets (respectively), which showed a high rate of amplification success for the targeted mitochondrial fragments (cytochrome *b*: 92%; control region: 84%). Reported sequences obtained from this material did not show ambiguities in direct sequencing (control region nor

cytochrome *b*), did not contain any nonsense or frameshift mutations nor showed an excess of nonsynonymous mutations (cytochrome *b*), and were not too divergent to sequences obtained from fresh material, suggesting that they were not severely affected by miscoding lesions or that they did not originate from the amplification of nuclear insertions of mitochondrial fragments.

Variation in control region sequences was high, with 69 variable sites that defined a total of 77 haplotypes [ $\pi$  (SD)= 0.046 (0.001); H (SD)= 0.963 (0.004)]. Most variation was due to transitions (92.8%), with only 7.18% transversions and one indel. Cytochrome *b* sequences included 24 variable sites (91.3% transitions and 8.7% transversions) which defined 24 haplotypes in *A. sapidus* [ $\pi$  (SD)= 0.008 (0.001); H(SD)= 0.718 (0.001)]. Parsimony informative sites were found in 17 out of the total amount of variable sites. Most substitutions were synonymous, with only two non-synonymous mutations resulting in conservative aminoacidic changes from alanine to glycine and valine to isoleucine occurring at nucleotide positions 21 and 74, respectively.

Finally, the 228 concatenated sequences defined 85 *A. sapidus* haplotypes [ $\pi$  (SD)= 0.027 (0.011); H(SD)= 0.974 (0.003)] which were used for phylogenetic and phylogeographic analysis (table 2), together with the six *A. terrestris* haplotypes found in the 12 individuals analysed.

**Table 2** GenBank accession numbers to Control Region and cytochrome *b* sequences that composed the 85 concatenated haplotypes of *A. sapidus* used in this study. *N*, total sampling frequency of haplotypes. Geographic localities codes are specified in Table 1. Numbers between parentheses (*n*) denote the relative sampling frequency per haplotype per geographic location.

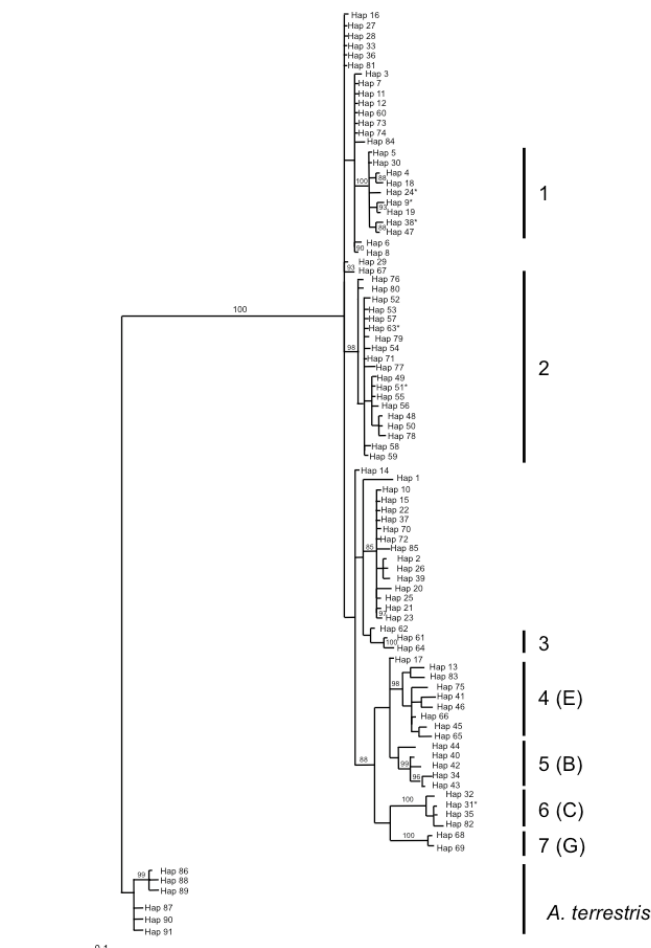


Haplotype	N	Geog. Loc. (n)	CR GenBank #.	Cytb GenBank #.	Haplotype	N	Geog. Loc. (n)	CR GenBank #.	Cytb GenBank #.
Hap_1	7	7 (7)	FJ895493	FJ895408	Hap_44	1	17 (1)	FJ895536	FJ895451
Hap_2	1	7 (1)	FJ895494	FJ895409	Hap_45	1	17 (1)	FJ895537	FJ895452
Hap_3	9	11 (4), 30 (1), 37 (4)	FJ895495	FJ895410	Hap_46	1	15 (1)	FJ895538	FJ895453
Hap_4	3	11 (3)	FJ895496	FJ895411	Hap_47	1	19 (1)	FJ895539	FJ895454
Hap_5	2	11 (1), 35 (1)	FJ895497	FJ895412	Hap_48	2	30 (2)	FJ895540	FJ895455
Hap_6	2	11 (2)	FJ895498	FJ895413	Hap_49	8	27 (1), 29 (1), 30 (6)	FJ895541	FJ895456
Hap_7	1	5 (1)	FJ895499	FJ895414	Hap_50	4	30 (4)	FJ895542	FJ895457
Hap_8	4	19 (1), 36 (1), 37 (2)	FJ895500	FJ895415	Hap_51	5	18 (1), 30 (3), 31 (1)	FJ895543	FJ895458
Hap_9	4	1 (1), 11 (3)	FJ895501	FJ895416	Hap_52	7	31 (7)	FJ895544	FJ895459
Hap_10	3	10 (1), 23 (2)	FJ895502	FJ895417	Hap_53	1	31 (1)	FJ895545	FJ895460
Hap_11	1	3 (1)	FJ895503	FJ895418	Hap_54	1	30 (1)	FJ895546	FJ895461
Hap_12	1	3 (1)	FJ895504	FJ895419	Hap_55	1	29 (1)	FJ895547	FJ895462
Hap_13	1	20 (1)	FJ895505	FJ895420	Hap_56	3	29 (3)	FJ895548	FJ895463
Hap_14	1	20 (1)	FJ895506	FJ895421	Hap_57	1	32 (1)	FJ895549	FJ895464
Hap_15	4	1 (1), 14 (1), 22 (1), 25 (1)	FJ895507	FJ895422	Hap_58	1	26 (1)	FJ895550	FJ895465
Hap_16	2	6 (2)	FJ895508	FJ895423	Hap_59	2	28 (2)	FJ895551	FJ895466
Hap_17	1	18 (1)	FJ895509	FJ895424	Hap_60	1	12 (1)	FJ895552	FJ895467
Hap_18	1	11 (1)	FJ895510	FJ895425	Hap_61	1	21 (1)	FJ895553	FJ895468
Hap_19	1	11 (1)	FJ895511	FJ895426	Hap_62	1	4 (1)	FJ895554	FJ895469
Hap_20	1	14 (1)	FJ895512	FJ895427	Hap_63	1	4 (1)	FJ895555	FJ895470
Hap_21	1	4 (1)	FJ895513	FJ895428	Hap_64	1	2 (1)	FJ895556	FJ895471
Hap_22	1	4 (1)	FJ895514	FJ895429	Hap_65	1	38 (1)	FJ895557	FJ895472
Hap_23	2	4 (2)	FJ895515	FJ895430	Hap_66	1	38 (1)	FJ895558	FJ895473
Hap_24	1	4 (1)	FJ895516	FJ895431	Hap_67	1	14 (1)	FJ895559	FJ895474
Hap_25	1	4 (1)	FJ895517	FJ895432	Hap_68	1	8 (1)	FJ895560	FJ895475
Hap_26	1	10 (1)	FJ895518	FJ895433	Hap_69	5	8 (1), 21 (1), 33 (3)	FJ895561	FJ895476
Hap_27	14	2 (1), 8 (1), 13 (1), 16 (5), 21 (6)	FJ895519	FJ895434	Hap_70	1	21 (1)	FJ895562	FJ895477
Hap_28	1	16 (1)	FJ895520	FJ895435	Hap_71	2	33 (2)	FJ895563	FJ895478
Hap_29	1	16 (1)	FJ895521	FJ895436	Hap_72	3	24 (3)	FJ895564	FJ895479
Hap_30	3	35 (3)	FJ895522	FJ895437	Hap_73	16	37 (16)	FJ895565	FJ895480
Hap_31	13	22 (1), 24 (1), 34 (2), 35 (7), 36 (2)	FJ895523	FJ895438	Hap_74	17	37 (17)	FJ895566	FJ895481
Hap_32	1	34 (1)	FJ895524	FJ895439	Hap_75	1	37 (1)	FJ895567	FJ895482
Hap_33	1	34 (1)	FJ895525	FJ895440	Hap_76	5	27 (5)	FJ895568	FJ895483
Hap_34	1	35 (1)	FJ895526	FJ895441	Hap_77	5	27 (5)	FJ895569	FJ895484
Hap_35	1	36 (1)	FJ895527	FJ895442	Hap_78	1	27 (1)	FJ895570	FJ895485
Hap_36	1	24 (1)	FJ895528	FJ895443	Hap_79	1	27 (1)	FJ895571	FJ895486
Hap_37	1	22 (1)	FJ895529	FJ895444	Hap_80	1	27 (1)	FJ895572	FJ895487
Hap_38	1	24 (1)	FJ895530	FJ895445	Hap_81	10	21 (10)	FJ895573	FJ895488
Hap_39	1	24 (1)	FJ895531	FJ895446	Hap_82	2	37 (2)	FJ895574	FJ895489
Hap_40	2	17 (2)	FJ895532	FJ895447	Hap_83	3	23 (3)	FJ895575	FJ895490
Hap_41	1	17 (1)	FJ895533	FJ895448	Hap_84	1	9 (1)	FJ895576	FJ895491
Hap_42	1	17 (1)	FJ895534	FJ895449	Hap_85	1	4 (1)	FJ895577	FJ895492
Hap_43	3	17 (3)	FJ895535	FJ895450					

As expected, nucleotide diversities were higher for localities where two divergent lineages concur (see below). French localities showed significantly lower nucleotide diversity than Iberian localities, once the admixed population in Pyrénées Orientales is excluded (Welch's approximate t-test:  $t' = 2.047$ ;  $P = 0.032$ ), suggesting a recent colonization of France from a reduced number of founders.

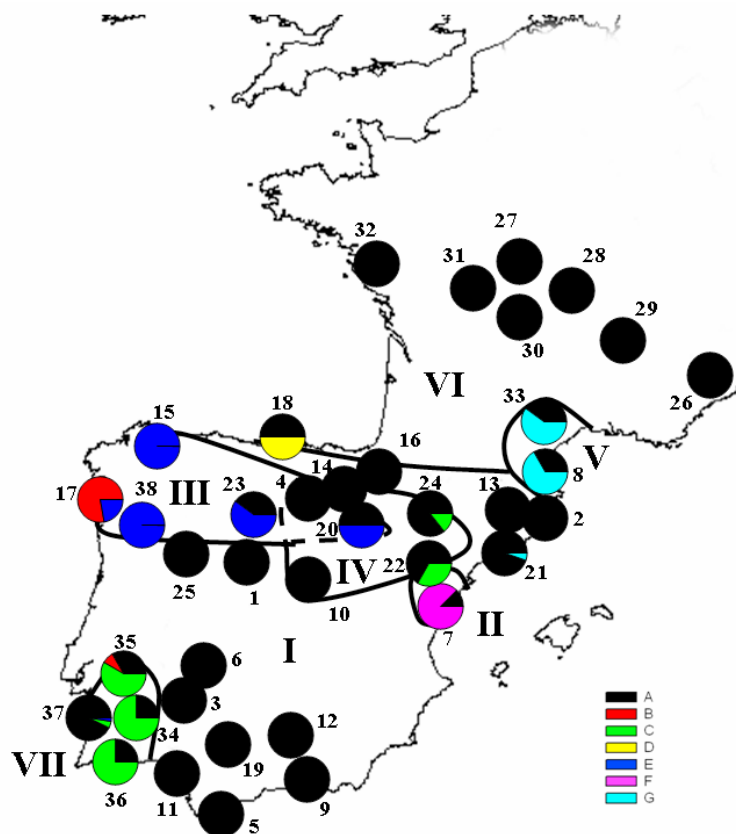
## Phylogeny

The inferred Bayesian phylogenetic tree confirmed the previously reported reciprocal monophyly and deep divergence between *A. terrestris* and *A. sapidus* (average net divergence, 10.1%) (Figure 1)



**Figure 1** Bayesian phylogenetic tree of *Arvicola* mitochondrial haplotypes (CR+cytb). Specified branch numbers refers to Bayesian posterior probability. Numbers on the right specify statistically supported clades (i.e.  $\geq 95$ ). Letters between parentheses refer to unconnected networks obtained with a 95% maximum parsimony limit. Haplotypes marked with and asterisk (\*) were sampled at relatively distant localities.

On the contrary, only a shallow phylogenetic structure was detected among *A. sapidus* haplotypes, as reflected in few statistically supported nodes ( $\geq 95\%$  Bayesian posterior probability) and relatively short internal branches (average intraspecific distances: 3.1%). Supported clades were geographically restricted/coherent, with the exception of haplotypes in three clades sampled in distant areas (marked with an asterisk in Fig. 1). Haplotypes in clade 1 were mostly sampled in SW Iberia but a few clade 1 haplotypes were sampled in Northern regions (Ávila, Burgos and Zaragoza, see codes in table 1 and their geographic position in figure 3).

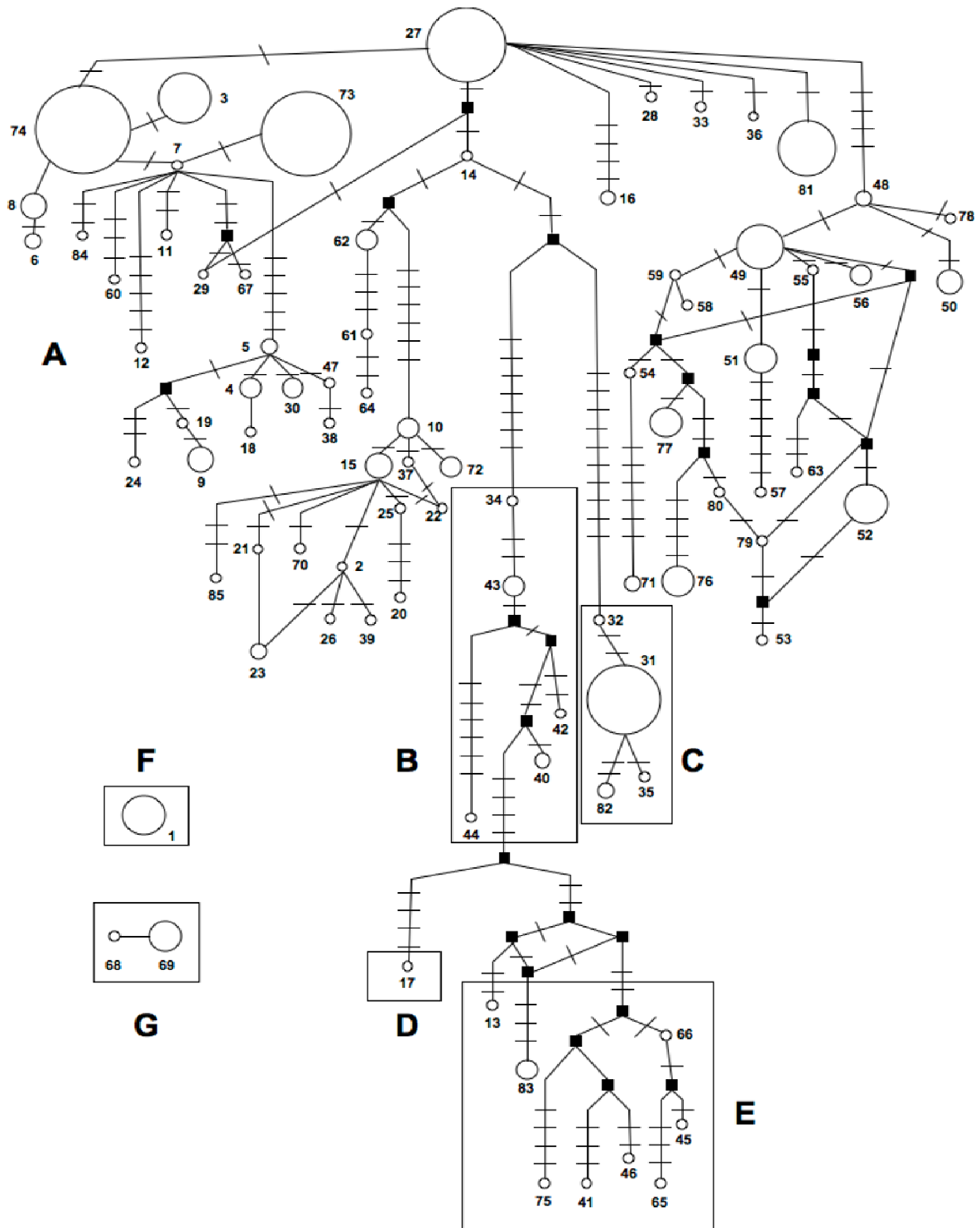


**Figure 3** Geographical distribution of the seven unconnected Networks (A to G, see legend) obtained under a 95% Maximum Parsimony criterion. Bold lines delimit the seven groups obtained with SAMOVA (Latin numbers). Numbers indicate locality codes as listed in Table 1

Clade 2 included all haplotypes found only in France, and also two samples from Santander and Burgos (Northern Iberia). Clade 3 was geographically restricted to the Northeastern quadrant of the Iberian Peninsula (Tarragona and Barcelona). Clade 4 extends over Northwest quadrant of Iberia. Haplotypes from clade 5 were located in NW Spain (Pontevedra) and central Portugal (Évora). Clade 6 was mainly centred over SW Iberia but includes haplotypes located in Northeastern Iberia (Zaragoza and Teruel). Haplotypes sampled on both Spanish and French sides of the Pyrenees (Gerona, Tarragona and Pyrénées-Orientales) composed clade 7.

### **Phylogeography**

Maximum Parsimony networks obtained under a 95% or a 90% parsimony limit were different (Figure 2).



**Figure 2** 90% Maximum Parsimony Networks obtained for the 85 haplotypes (numbers) sampled in *Arvicola sapidus*. Circle sizes are proportional to haplotype frequencies. Capital letters indicate the seven unconnected clades obtained with a 95% parsimony limit. Transversal short lines are mutational steps between haplotypes. Filled squares are unsampled haplotypes.

The dataset was divided into seven unconnected networks (networks A to G, Fig. 2) when using a 95% parsimony limit (connection limit at six mutational steps), while networks A to E were joined in a single network under a 90% parsimony limit (connection limit at ten steps). Networks B, C, E and G coincided with clades 5, 6, 4 and 7 (respectively) in the Bayesian phylogenetic tree. The geographical distribution of these seven different lineages is drawn in Figure 3. A high proportion of the overall haplotype diversity of *A. sapidus* is included in the lineage A, the most frequent and widespread over Iberia and France, and this is surrounded by peripheral, more geographically restricted and partially overlapping lineages (B to G). The only D and F haplotypes were sampled in Northern and Eastern Spain (Santander and Castellón) respectively. The unconnected lineage G was sampled in the French and Spanish sides of Western Pyrenees. All other French haplotypes derive from haplotype 48 (Corrèze), and this is connected to haplotype 27, located at the Iberian side of the Pyrenees, by four mutational steps.

Despite these few geographic inconsistencies, a global phylogeographic pattern was indicated by the existence of a positive global correlation between pairwise nucleotide differences and Euclidean geographic distance between sampled haplotypes ( $r = 0.197$ ,  $p = 0.03$ ), and by the autocorrelogram of II indexes (AIDAs) showing a significant autocorrelation between sequences up to 450-518 kms. The Pyrenees did not significantly contributed to the observed genetic structure once the effects of Euclidean distances were discounted (Partial Mantel Test:  $r = 0.2098$ ,  $p = 0.1097$ )

## Genetic structure

The hierarchical and spatial analysis of genetic structure (SAMOVA) suggested a spatial subdivision of the whole distribution area into seven groups (Figure 3 and Table 3), since  $\Phi_{CT}$  values approached a plateau for  $k = 7$ . The Analysis of Molecular Variance (AMOVA) with these seven groups revealed a more than three times higher significant percentage of genetic variance among groups than among populations within groups (43.7% and 12.8%, respectively), compared to a moderate fraction (43.5%) attributed to differences within populations (fixation indexes:  $\Phi_{SC} = 0.22812$ ,  $\Phi_{ST} = 0.56565$ ,  $\Phi_{CT} = 0.43729$ ; all p-values  $< 0.001$ ).

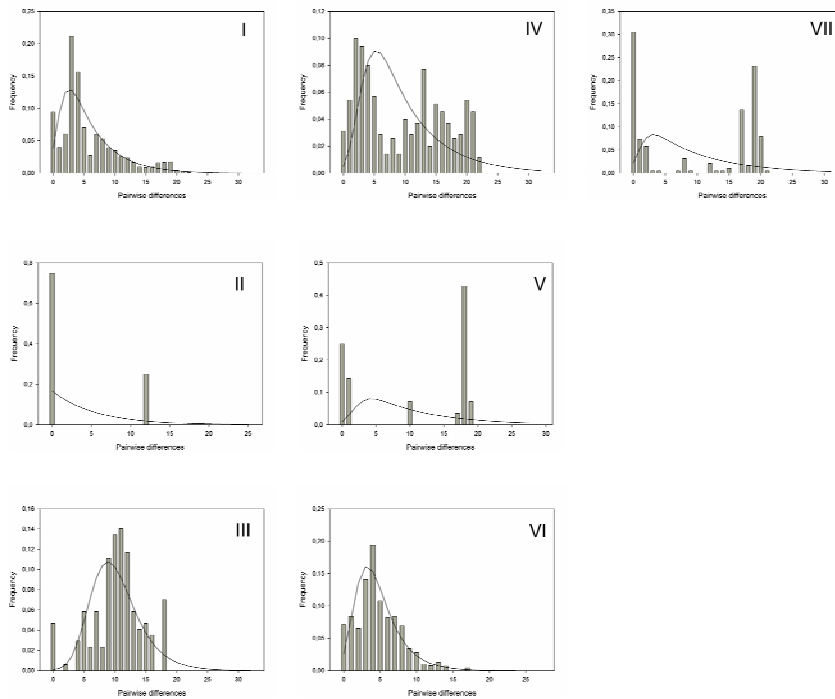
Statistics	I	II	III	IV	V	VI	VII
<b>N</b>	95	8	19	27	8	51	20
<b>S</b>	59	12	35	41	23	32	30
<b>p</b>	0.014	0.007	0.025	0.023	0.025	0.011	0.023
<b>h</b>	26	2	13	20	4	19	8
<b>H</b>	0.905	0.25	0.953	0.969	0.75	0.929	0.695
<b>k</b>	5.766	3.0	10.310	9.601	10.536	4.7	9.816
<b>D</b>	<b>-1.5894</b>	<b>-1.7700</b>	0.1183	-0.3660	0.9842	-1.0799	0.6303
<b>SSD</b>	0.0218	<b>0.0891</b>	0.0191	0.0113	<b>0.2170</b>	0.0052	0.1048
<b>Raggedness</b>	0.0402	<b>0.6875</b>	0.0315	<b>0.0136</b>	<b>0.3304</b>	0.0192	<b>0.1676</b>
<b><math>\tau</math></b>	2.4805	n.e.	11.3262	n.e.	n.e.	4.1055	n.e.

**Table 3** Genetic diversity and demographic statistics for SAMOVA partitions (I to VII). *N*: sample size. *S*: number of segregating sites. *p*: nucleotide diversity. *h*: number of haplotypes; *H*: haplotype diversity; *k*: mean number of nucleotide differences. *SSD*: Sum of Square Deviation (goodness of fit to a simulated population expansion).  $\tau$  (tau) values for the estimates of time since population expansion was estimated based on mismatch distribution (see text). Tajima's D, SSD and Raggedness index values in **bold** show significant tests ( $p < 0.05$ ). n.e. Not estimated (see text).

## Demography

The evidence for historical demographic changes was analysed for each of the partitions suggested by SAMOVA through neutrality tests, goodness of fit to a simulated population expansion, and raggedness index (Table 3) and mismatch distributions (Figure 4). Although only groups I and II showed significant deviations from neutral expectations (Tajima's  $D = -1.5894$  ( $p < 0.05$ ) and  $-1.77$  ( $p < 0.01$ ) respectively), for groups I, III and VI, a signal for demographic expansion was supported by non-significant raggedness indexes and mismatch distributions not deviating significantly from that expected under a sudden expansion model. Tau ( $\tau$ ) estimates for these partitions (group I: 2.4805; group III: 11.3262; group VI: 4.1055) dated population expansion for group III over 110.000 y.b.p, followed by population expansions in groups VI and I around 40.000 and 24.000 y.b.p. respectively.





**Figure 4** Observed mismatch distribution (histogram) and expected distribution under a sudden expansion model (line) for each of the groupings inferred with SAMOVA (latin numbers).

Parameters estimated under the Isolation with Migration model for French and Iberian Southern water voles were largely coincident in three independent runs (Table 4), so estimates obtained in the run with the largest ESS for split time are reported below.

**Table 4** Results of three independent runs evaluating France and Iberia divergence under an IM model.  $N_1$ , effective population size of *A. sapidus* in Iberia.  $N_2$ , effective population size of *A. sapidus* in France.  $N_A$ , effective population size of ancestral *A. sapidus* in Iberia before splitting.  $t$ , time since splitting (years).  $m1$ , migration rate (gene copies per generation) from Iberia to France.  $m2$ , migration rate from France to Iberia.  $HiPt$ , value of the highest count.  $HPD90Lo$  and  $HPD90Hi$  denotes the lower and upper bounds of the estimated 90% highest posterior density (HPD) interval

<b>mtDNA</b>	<b>N<sub>1</sub></b>	<b>N<sub>2</sub></b>	<b>N<sub>A</sub></b>	<b>t</b>	<b>m1</b>	<b>m2</b>
<b>HiPt</b>	496607	106058	41844	61942	0.0016	0.0003
<b>HPD90Lo</b>	310281	58687	790	34452	0.0010	0.0002
<b>HPD90Hi</b>	801361	182905	166588	100499	0.0025	0.0006
<b>HiPt</b>	496607	105533	43423	55515	0.0016	0.0003
<b>HPD90Lo</b>	310281	58687	790	33381	0.0010	0.0002
<b>HPD90Hi</b>	806098	182905	171325	100856	0.0025	0.0006
<b>HiPt</b>	496607	105006	41844	61585	0.0016	0.0003
<b>HPD90Lo</b>	310281	58161	790	33381	0.0010	0.0002
<b>HPD90Hi</b>	798202	182905	169747	101927	0.0025	0.0006

Estimates of effective population size of the Iberian group (Point estimate: 496607; HPD 90% interval: 310281 – 801361) were larger than for the French group (106585; HPD 90% interval: 58687 – 182905), whereas ancestral population size was estimated in 41844 (HPD 90% interval: 790 – 166588). Time since splitting of both groups was estimated in 61942 y. b. p. (HPD 90% interval: 34452 – 100499). Interestingly, the date of split roughly coincides with the date of the demographic expansion of the SAMOVA partition VI, estimated in ca. 40.000 ybp, which suggests that demographic expansion quickly followed France colonization. Average migration rate from Iberia into France was estimated in 0.0016 gene copies per generation (HPD 90% interval: 0.001 – 0.0025), while for the reciprocal migration the posterior probability of the migration parameter peaked at the lowest interval indicating it to be close to zero. Notwithstanding, the occurrence of some gene flow from France to Iberia is indicated by the two observations of French haplotypes in Northern Iberia (haplotype 3 in

Corrèze –France- and Huelva –Spain- and haplotype 69 in Pyrénées Orientales – France- and Gerona and Tarragona –Spain-).

## DISCUSSION

### **Evolutionary history of *A. sapidus***

Taberlet *et al.* (1998) summarized the phylogeny of *Arvicola* spp. as a genus with four lineages, the most basal one corresponding to *A. sapidus* and the other three corresponding to intraspecific lineages within *A. terrestris* (Wust Saucy, 1998; Wust Saucy *et al.* 1998). Our results support this same relatively deep divergence between the two *Arvicola* species as reflected by net divergences of 10.1% in combined control region and cytochrome *b* sequences, which grossly translates into a species split dating back to the mid-Pleistocene (Centeno-Cuadros *et al.* unpublished data).

Our analysis of sequence variation at cytochrome *b* and control region sequences across *A. sapidus* distribution range reveals a high overall mitochondrial diversity for the species ( $\pi = 0.027$ ;  $H = 0.974$ ), and a shallow, but significant phylogeographic structure. The shallowness of Southern water voles structure is reflected in i) few statistically supported clades, ii) relatively short branches on phylogenetic trees (Fig. 1) and iii) a low number of unconnected haplotype networks at 90% parsimony limits, which increased, however, when the parsimony limit was set to 95% (Fig. 2). These phylogenetically shallow intraspecific lineages showed good geographical coherence overall, as reflected in a global pattern of increased pairwise nucleotide divergence with distance and significant autocorrelation up to distances of *ca.* 500

Kms. Lineage sharing across large distances is suggestive of incomplete lineage sorting, although long distance human-and/or-raptors-mediated dispersal cannot be ruled out. When the geographic sampling of haplotypes is taken into account, up to seven spatial partitions can be defined that maximize the proportion of genetic variation distributed among groups of populations (Fig. 3; Table 3).

Theoretical studies have shown that phylogeographic discontinuities can arise in the absence of geographic barriers when dispersal is restricted by distance with respect to the geographic scale considered (Irwin, 2002). The relative shallowness of Southern water voles phylogeographic structure and the apparently short dispersal distances typically reported in rodents, make this a possibility that cannot be discarded. However, some of the results presented suggest the possible contribution of episodes of isolation in glacial subrefugia. Firstly, the geographical distribution of *A. sapidus* lineages show some congruence with other Iberian species, suggesting that their genetic structure have been shaped by isolation in common Pleistocene “refugia within the Iberian refugium”. Interestingly, four of the refugia suggested for Southern water voles (partitions I, III, IV and VII) are coincident with other previously described for lizards (Paulo *et al.* 2008), In particular, *A. sapidus* patterns add evidence for the consideration of the North quadrant of Iberia as a subrefugium, as the spatial distribution of partition III overlaps with that described for divergent lineages in other species [e.g. vertebrates (Paulo *et al.* 2008) and invertebrates (Vialatte *et al.* 2008; Vila *et al.* 2005)]. The Eastern Pyrenees (partition VII) have also been shown to harbour divergent population in other species (Vialatte *et al.* 2008). Interestingly, two Iberian refugia have also been proposed for *Microtus agrestis*, one in central Portugal and one in northern Spain (Jaarola & Searle, 2004). Secondly, a signal of demographic expansion found for some of the partitions

(partitions I, III and VI –Table 3-) would support this scenario by indicating an expansion of *A. sapidus* from their corresponding glacial subrefugia.

In sharp contrast with the deep interlineage divergences found for most of the examples cited above, *A. sapidus* shallow interlineage divergences and evidences for incomplete lineage sorting apparently discard a long-term isolation within these Iberian subrefugia over successive glacial cycles. In the case of *A. sapidus*, divergence was probably tempered by relatively large population sizes. Rapid and extensive mixing among subrefugia during interglacials might have eroded the phylogenetic structure accumulated at glacial peaks (Hofreiter *et al.* 2004).

As observed for many other Southern European refugial species, Southern water voles might have expanded northward following ice retreat, as suggested by a lower nucleotide diversity in France. Furthermore, French Southern water voles form a distinct population segment (partition VI in SAMOVA analysis) and all French haplotypes group in a single clade (clade 2 in Fig. 1), to the only exception of haplotypes belonging to clade 7 sampled in Pyrénées Orientales, suggesting a single event of colonization from Iberia. We estimated the date of the colonization of France from Iberia under a model of isolation with migration (Hey & Nielsen, 2007), in ca. 62 thousands years ago (34452 – 100499, HPD 90% interval), dating it before the end of the last glacial cycle but after its beginning. It must be noted that this estimate heavily depends on assumed mutation rates, a parameter that is admittedly difficult to estimate. We used phylogenetic estimates of mutation rates previously used for closely related *Microtus arvalis* (4% and 46% per million years for cytochrome *b* and control region, respectively; (Fink *et al.* 2004; Hamilton *et al.* 2005)) which are high for vertebrates in general, but that match the high mutation rates usually reported for rodents (Heckel *et al.* 2005; Rajabi-Maham *et al.* 2008). While the use of typical vertebrate rates (2% for

cytochrome *b* and 20% for control region, substitutions per site per million years) that would take this date back to ca. 130000 to the previous interglacial seems unrealistic, the higher mutation rates needed to bring this date forward to the end of the last glacial maximum might be supported by recent evidence indicating accelerated rates of evolution at short time scales (Ho *et al.* 2005; Ho *et al.* 2007; Lambert *et al.* 2002). While the Alps seem to have prevented the expansion of the Italian lineage (*A. t. italicus*) out of the Italian peninsula, the Pyrenees did not act as an absolute barrier for the expansion of Southern water voles from Iberia into France, nor for the colonization of Northern Iberia by *A. terrestris* from its central European refugium. The permeability of the Pyrenees to Southern water voles is further highlighted by group V spanning both sides of the Eastern extreme - a previously described corridor for other species (Hewitt, 1996; Hewitt, 1999) - and by a nonsignificant effect of the Pyrenees on pairwise genetic differentiation once the effect of distance is discounted.

### **Implications for taxonomy and conservation**

Our results do not support the proposed subspecific delimitation based on hair colour pattern upon which the darker pattern in North and West of Iberia would identify *A. sapidus tenebricus* and a lighter brown coloured pattern would characterize *A. sapidus sapidus*. In particular, no association was found between colour pattern and intraspecific mtDNA variation, indicating this phenotypic difference does not delineate historically isolated intraspecific divisions. Lack of concordance was also found between mitochondrial divergences and body size differences, as small-sized Doñana water voles do not match an exclusive divergent lineage, indicating that the former might arise as a plastic response to local differences in environment.

Since the seven spatially continuous groups inferred with SAMOVA (Fig. 3) represent a distinct portion of the total diversity of the species they all have to be preserved. However, these units are not the consequence of long-term isolation, do not seem to harbour any adaptive divergences with respect to the rest of species, and they do maintain extensive gene flow among them, indicating that the species should be treated as a single evolutionary significant unit. Our results, however, support the presence of subrefugia within the Iberian glacial refugium and their role as sources of diversification and endemisms (Gómez & Lunt, 2007)

#### **ACKNOWLEDGEMENTS**

This work has been performed with the invaluable fieldwork assistance, sample request and help from Jacinto Román. Special thanks go to Ana Píriz and Laura Soriano for their help in the lab. We are also grateful with the Centro de Supercomputación de Galicia (CESGA) where all MRBAYES and IMA analyses were performed. Stuart Piertney and four anonymous reviewers provided helpful comments on earlier versions of the manuscript. The authors are also indebted to F. Alvares, R. Balado, J. Barrul, A. Ceña, J.C. Ceña, M.M. Delgado, A. Delibes, C. Escala, N. Indelicato, C. Lemarchand, R. Lorenço, J. Mate, A. Monzón, A. Morcillo, J.F. Noblet, R. Pita, D. Serrano, F. Simonnet, J.V. Turón, J. Ventura and curators from the Natural Science Museum of Barcelona, Aranzadi Society of Sciences and Biological Station of Doñana for providing samples for this study. This work has been funded by the Dirección General de Investigación (project BOS2001-2391-C02-01). ACC benefited from a FPU fellowship from the Ministerio de Educación y Ciencia.

## REFERENCES

- Aars J, Dallas JF, Piertney SB, *et al.* (2006) Widespread gene flow and high genetic variability in populations of water voles *Arvicola terrestris* in patchy habitats. *Molecular Ecology* **15**, 1455-1466.
- Avise JC (2000) *Phylogeography: The History and Formation of Species* Harvard University Press, Cambridge, MA.
- Berthier K, Charbonnel N, Galan M, Chaval Y, Cosson J-F (2006) Migration and recovery of the genetic diversity during the increasing density phase in cyclic vole populations. *Molecular Ecology* **15**, 2665-2676.
- Berthier K, Galan M, Foltête JC, Charbonnel N, Cosson JF (2005) Genetic structure of the cyclic fossorial water vole (*Arvicola terrestris*): landscape and demographic influences. *Molecular Ecology* **14**, 2861-2871.
- Bertorelle G, Barbujani G (1995) Analysis of DNA Diversity by Spatial Autocorrelation. *Genetics* **140**, 811-819.
- Bryja J, Charbonnel N, Berthier K, Galan M, Cosson JF (2007) Density-related changes in selection pattern for major histocompatibility complex genes in fluctuating populations of voles. *Molecular Ecology* **16**, 5084-5097.
- Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology* **9**, 1657-1659.
- de Bellocq JG, Charbonnel N, Morand S (2008) Coevolutionary relationship between helminth diversity and MHC class II polymorphism in rodents. *Journal of Evolutionary Biology* **21**, 1144-1150.
- Depaulis F, Mousset S, Veuille M (2003) Power of neutrality tests to detect bottlenecks and hitchhiking. *Journal of Molecular Evolution* **57**, S190-S200.
- Diaz de la Guardia R, Pretel A (1978) Karyotype and centric dissociation in water vole *Arvicola sapidus* spp *sapidus* Miller 1908 (Rodentia, Muridae). *Experientia* **34**, 706-708.
- Dueñas ME, Peris Álvarez SJ (1985) *Clave para los micromamíferos (Insectivora y Rodentia) del Centro y Sur de la Península Ibérica* Universidad de Salamanca, Salamanca.
- Dupanloup I, Schneider S, Excoffier L (2002) A simulated annealing approach to define the genetic structure of populations. *Molecular Ecology* **11**, 2571-2581.
- Excoffier L, Laval G, Schneider S (2005) Arlequin (version 3.0): An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* **1**, 47-50.
- Fink S, Excoffier L, Heckel G (2004) Mitochondrial gene diversity in the common vole *Microtus arvalis* shaped by historical divergence and local adaptations. *Molecular Ecology* **13**, 3501-3514.
- Godinho R, Crespo EG, Ferrand N (2008) The limits of mtDNA phylogeography: complex patterns of population history in a highly structured Iberian lizard are only revealed by the use of nuclear markers. *Molecular Ecology* **17**, 4670-4683.
- Goetz SE, Hamilton SR, Vogelstein B (1985) Purification of DNA from formaldehyde fixed and paraffin embedded human tissue. *Biochemical and biophysical research communications* **130**, 118-126.
- Gómez Á, Lunt DH (2007) Refugia within refugia: patterns of phylogeographic concordance in the Iberian Peninsula. In: *Phylogeography of Southern European Refugia*.



- Evolutionary perspectives on the origins and conservation of European biodiversity* (eds. Weiss S, Ferrand N). Springer.
- Hamilton G, Currat M, Ray N, *et al.* (2005) Bayesian estimation of recent migration rates after a spatial expansion. *Genetics* **170**, 409-417.
- Heckel G, Burri R, Fink S, Desmet JF, Excoffier L (2005) Genetic structure and colonization processes in European populations of the common vole, *Microtus arvalis*. *Evolution* **59**, 2231-2242.
- Hewitt G (2000) The genetic legacy of the Quaternary ice ages. *Nature* **405**, 907-913.
- Hewitt GM (1996) Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society* **58**, 247-276.
- Hewitt GM (1999) Post-glacial re-colonization of European biota. *Biological Journal of the Linnean Society* **68**, 87-112.
- Hey J, Nielsen R (2007) Integration within the Felsenstein equation for improved Markov chain Monte Carlo methods in population genetics. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 2785-2790.
- Ho SYW, Phillips MJ, Cooper A, Drummond AJ (2005) Time dependency of molecular rate estimates and systematic overestimation of recent divergence times. *Molecular Biology and Evolution* **22**, 1561-1568.
- Ho SYW, Shapiro B, Phillips MJ, Cooper A, Drummond AJ (2007) Evidence for time dependency of molecular rate estimates. *Systematic Biology* **56**, 515-522.
- Hofreiter M, Serre D, Rohland N, *et al.* (2004) Lack of phylogeography in European mammals before the last glaciation. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 12963-12968.
- Huelsenbeck JP, Ronquist F (2001) MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* **17**, 754-755.
- Ibrahim KM, Nichols RA, Hewitt GM (1996) Spatial patterns of genetic variation generated by different forms of dispersal during range expansion. *Heredity* **77**, 282-291.
- Irwin DE (2002) Phylogeographic breaks without geographic barriers to gene flow. *Evolution* **56**, 2383-2394.
- Jaarola M, Searle JB (2004) A highly divergent mitochondrial DNA lineage of *Microtus agrestis* in southern Europe. *Heredity* **92**, 228-234.
- Lambert DM, Ritchie PA, Millar CD, *et al.* (2002) Rates of evolution in ancient DNA from Adelle penguins. *Science* **295**, 2270-2273.
- Marjoram P, Donnelly P (1994) Pairwise comparisons of mitochondrial-DNA sequences in subdivided populations and implications for early human-evolution. *Genetics* **136**, 673-683.
- Martinez-Solano I, Teixeira J, Buckley D, Garcia-Paris M (2006) Mitochondrial DNA phylogeography of *Lissotriton boscai* (Caudata, Salamandridae): evidence for old, multiple refugia in an Iberian endemic. *Molecular Ecology* **15**, 3375-3388.
- Megias-Nogales B, Marchal JA, Acosta MJ, *et al.* (2003) Sex chromosome pairing in two Arvicolidae species: *Microtus nivalis* and *Arvicola sapidus*. *Hereditas* **138**, 114-121.
- Müllenbach R, Lagoda PJJ, Welter C (1989) An efficient salt-chloroform extraction of DNA from blood and tissues. *Trends in Genetics* **5**, 391-391.
- Oliver MK, Lambin X, Thomas C, Piertney SB (2009) Spatio-temporal variation in the strength and mode of selection acting on major histocompatibility complex diversity in water vole (*Arvicola terrestris*) metapopulations. *Molecular Ecology* **18**, 80-92.
- Oliver MK, Piertney SB (2006) Isolation and characterization of a MHC class II DRB locus in the European water vole (*Arvicola terrestris*). *Immunogenetics* **58**, 390-395.

- Pannell JR (2003) Coalescence in a metapopulation with recurrent local extinction and recolonization. *Evolution* **57**, 949-961.
- Paulo OS, Pinheiro J, Miraldo A, *et al.* (2008) The role of vicariance vs. dispersal in shaping genetic patterns in ocellated lizard species in the western Mediterranean. *Molecular Ecology* **17**, 1535-1551.
- Piertney SB, Stewart WA, Lambin X, *et al.* (2005) Phylogeographic structure and postglacial evolutionary history of water voles (*Arvicola terrestris*) in the United Kingdom. *Molecular Ecology* **14**, 1435-1444.
- Pinho C, Harris DJ, Ferrand N (2007) Comparing patterns of nuclear and mitochondrial divergence in a cryptic species complex: the case of Iberian and North African wall lizards (Podarcis, Lacertidae). *Biological Journal of the Linnean Society* **91**, 121-133.
- Posada D, Buckley TR (2004) Model selection and model averaging in phylogenetics: Advantages of Akaike information criterion and Bayesian approaches over likelihood ratio tests. *Systematic Biology* **53**, 793-808.
- Posada D, Crandall KA (1998) MODELTEST: testing the model of DNA substitution. *Bioinformatics* **14**, 817-818.
- Rajabi-Maham H, Orth A, Bonhomme F (2008) Phylogeography and postglacial expansion of *Mus musculus domesticus* inferred from mitochondrial DNA coalescent, from Iran to Europe. *Molecular Ecology* **17**, 627-641.
- Rogers AR (1995) Genetic-evidence for a Pleistocene population explosion. *Evolution* **49**, 608-615.
- Rogers AR, Harpending H (1992) Population-growth makes waves in the distribution of pairwise genetic-differences. *Molecular Biology and Evolution* **9**, 552-569.
- Rohland N, Hofreiter M (2007) Comparison and optimization of ancient DNA extraction. *BioTechniques* **42**, 343-352.
- Román J (2007) *Historia natural de la rata de agua (Arvicola sapidus) en Doñana*. PhD Thesis. Universidad Autónoma de Madrid.
- Rousset F (1997) Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Molecular Ecology* **145**, 1219-1228.
- Rozas J, Sánchez-DelBarrio JC, Messeguer X, Rozas R (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Methods of biochemical analysis* **19**, 2496-2497.
- Schneider S, Excoffier L (1999) Estimation of past demographic parameters from the distribution of pairwise differences when the mutation rates vary among sites: Application to human mitochondrial DNA. *Genetics* **152**, 1079-1089.
- Sequeira F, Alexandrino J, Weiss S, Ferrand N (2008) Documenting the advantages and limitations of different classes of molecular markers in a well-established phylogeographic context: lessons from the Iberian endemic Golden-striped salamander, *Chioglossa lusitanica* (Caudata: Salamandridae). *Biological Journal of the Linnean Society* **95**, 371-387.
- Sesé Benito C (1994) Paleoclimatical interpretation of the Quaternary small mammals of Spain. *Geobios* **27**, 753-767.
- Shiozawa DK, Kudo J, Evans RP, Woodward SR, Williams RN (1992) DNA extraction from preserved trout tissues. *Great Basin Naturalist* **52**, 29-34.
- Simonsen KL, Churchill GA, Aquadro CF (1995) Properties of statistical tests of neutrality for DNA polymorphism data. *Genetics* **141**, 413-429.

- Stewart WA, Dallas JF, Piertney SB, *et al.* (1999) Metapopulation genetic structure in the water vole, *Arvicola terrestris*, in NE Scotland. *Biological Journal of the Linnean Society* **68**, 159-171.
- Taberlet P, Fumagalli L, Wust-Saucy A-G, Cosson J-F (1998) Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology* **7**, 453-464.
- Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**, 585-595.
- Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0. *Molecular Biology and Evolution* **24**, 1596-1599.
- Telfer S, Dallas JF, Aars J, *et al.* (2003) Demographic and genetic structure of fossorial water voles (*Arvicola terrestris*) on Scottish islands. *Journal of Zoology (London)* **259**, 23-29.
- Templeton AR (2004) Statistical phylogeography: methods of evaluating and minimizing inference errors. *Molecular Ecology* **13**, 789-809.
- Thompson JD, Higgins DG, Gibson TJ (1994) Clustal-W - Improving the Sensitivity of Progressive Multiple Sequence Alignment through Sequence Weighting, Position-Specific Gap Penalties and Weight Matrix Choice. *Nucleic Acids Research* **22**, 4673-4680.
- Tollenaere C, Bryja J, Galan M, *et al.* (2008) Multiple parasites mediate balancing selection at two MHC class II genes in the fossorial water vole: insights from multivariate analyses and population genetics. *Journal of Evolutionary Biology* **21**, 1307-1320.
- Tougaard C, Renvoise E (2008) Rodents and palaeogenetics: New perspectives. *Comptes Rendus Palevol* **7**, 125-134.
- Ventura J (2002) *Arvicola sapidus* Miller, 1908. Rata de agua. In: *Atlas de los mamíferos terrestres de España* (eds. Palomo LJ, Gisbert J). Dirección General de Conservación de la Naturaleza-SECEM-SECEMU, Madrid.
- Vialatte A, Guiller A, Bellido A, Madec L (2008) Phylogeography and historical demography of the Lusitanian snail *Elona quimperiana* reveal survival in unexpected separate glacial refugia. *Bmc Evolutionary Biology* **8**, 339.
- Vilà C, Amorim I, Leonard J, *et al.* (1999) Mitochondrial DNA phylogeography and population history of the grey wolf *Canis lupus*. *Molecular Ecology* **8**, 2089-2103.
- Vila M, Vidal-Romani JR, Bjorklund M (2005) The importance of time scale and multiple refugia: Incipient speciation and admixture of lineages in the butterfly *Erebia triaria* (Nymphalidae). *Molecular Phylogenetics and Evolution* **36**, 249-260.
- Weiss S, Ferrand N (2007) *Phylogeography of Southern European Refugia. Evolutionary perspectives on the origins and conservation of European biodiversity*. Springer.
- Wust Saucy AG (1998) *Polymorphisme genetique et phylogeographie du campagnol terrestre Arvicola terrestris*. PhD Thesis. Université de Lausanne.
- Wust Saucy AG, Hausser J, Taberlet P, Saucy F (1998) Phylogeography of the vole *Arvicola terrestris* as revealed by mtDNA: The role of historical factors? *Zoologia et Botanica*, 764-765.
- Young ND, Healy J (2003) GapCoder automates the use of indel characters in phylogenetic analysis. *Bmc Bioinformatics* **4**, 6.
- Zink RM, Barrowclough GF (2008) Mitochondrial DNA under siege in avian phylogeography. *Molecular Ecology* **17**, 2107-2121.

## **CHAPTER 4**



## CHAPTER 4

### The role of landscape on genetic structure and dispersal on Southern water vole

(*Arvicola sapidus*)

(en preparación)

<sup>1</sup> Department of Conservation Biology, Estación Biológica de Doñana, CSIC. C/ Américo Vespucio, s/n. CP 41092. Sevilla, Spain.

<sup>2</sup> Department of Integrative Ecology, Estación Biológica de Doñana, CSIC. C/ Américo Vespucio, s/n. CP 41092. Sevilla, Spain.

**Keywords:** microsatellites, mitochondrial DNA, Doñana, ponds, Least Cost Distance,

**Running title:** Landscape genetics and dispersal of *Arvicola sapidus*



## RESUMEN

El estudio de la influencia del paisaje sobre las distancias y tasas de dispersión y la importancia de estas en la persistencia y evolución de las especies y dinámica poblacional se está estudiando recientemente mediante el uso de marcadores genéticos. En este trabajo hemos utilizado a la rata de agua meridional (*Arvicola sapidus*) como especialista de hábitat y especie modelo para testar la influencia del la matriz paisajística en los patrones de dispersión. Para nuestros análisis hemos analizado el genotipo de 10 marcadores microsátélites y secuencias de toda la región control mitocondrial ( $S = 17$ ,  $\pi = 0.00613$ ,  $H = 0.846$ ) de 142 individuos capturados en siete áreas geográficas distribuidas por la Región Natural de Doñana (Suroeste de España). En todo el área del estudio hemos encontrado una estructura genética mitocondrial significativa y moderada ( $F_{ST} = 0.1272$ ) al igual que con marcadores autosómicos ( $F_{ST} = 0.072$ ). Esta estructura se explicaba parcialmente por la distancia geográfica entre poblaciones (aislamiento por distancia, test de Mantel:  $r = 0.5844$ ,  $p < 0.01$ ), mientras que la cantidad y disposición de las lagunas, drenajes y cobertura vegetal no explicaron significativamente dicho patrón más allá de la propia distancia Euclídea (test parcial de mantel,  $p > 0.05$ ). A diferencia de los métodos directos, las pruebas genéticas para la detección de sesgos de dispersión debido al sexo de los individuos rechazaban la hipótesis nula que mantenía una dispersión similar entre sexos (todos  $p > 0.05$ ), aunque la distancia media estimada entre parentales y su descendencia para machos y hembras ( $\bar{\square}_{machos} = 668$  m,  $\bar{\square}_{hembras} = 661$ m) coincidió con la estimada mediante los métodos de captura-recaptura. Las estimas de migración reciente obtenidas con aproximaciones multilocus bayesianas revelaron elevadas tasas de migración entre las regiones más septentrionales, sugiriendo una dinámica de fuente sumidero moderada dentro del área de estudio. Los resultados presentados en este capítulo apuntan hacia una estrategia dispersiva de la rata de agua propia de especies generalistas de hábitat, probablemente adquirida por especies especialistas de hábitat para facilitar la colonización de parches vacíos localizados a grandes distancias entre colonias situadas en matrices hostiles.





## ABSTRACT

How does the landscape influence on dispersal distance and rates and its importance for species persistence and evolution and population dynamics is being recently addressed by the use of genetic markers. In this study, we use the habitat specialist Southern water vole (*Arvicola sapidus*) as a model species to test the influence of the landscape matrix on dispersal patterns. Genotypes of 10 microsatellites and sequences of the whole mitochondrial Control Region (overall  $S = 17$ ,  $\pi = 0.00613$ ,  $H = 0.846$ ) were obtained from 142 individuals trapped within seven geographical populations throughout the Natural Region of Doñana (Southwest of Spain). There was a significant and moderate mitochondrial genetic structure ( $F_{ST} = 0.1272$ ) as also reported when using autosomal markers ( $F_{ST} = 0.072$ ). Whereas this pattern was partially explained by the geographic distance among populations (Isolation by distance, Mantel Test:  $r=0.5844$ ,  $p<0.01$ ),

ponds, drainages and vegetation cover did not significantly explained the genetic structure further than Euclidean distance did (Partial Mantel Test, all  $p > 0.05$ ). As opposite to direct methods, genetic sex-biased dispersal tests rejected the null hypothesis of equal dispersal between sexes (all  $p > 0.05$ ), although the average axial parent-offspring distances estimated for males and females ( $\bar{\Delta}_{males} = 668$  m,  $\bar{\Delta}_{females} = 661$ m) coincided with that inferred with capture-mark-release methods. Direct estimates of recent migration rates obtained with Bayesian multilocus genotyping procedures revealed higher migration rates between Northern regions and suggested a moderate source-sink dynamic within the studied area. These results pointed out a dispersal habitat generalist species strategy, likely acquired on habitat specialist species to facilitate distanced empty patches colonization between colonies embedded within a hostile (non-breeding habitats) matrix



## INTRODUCTION

Animal dispersal is commonly understood as the one-way movement of individuals beyond their home ranges without any return (or, at least, temporally) (Lidicker Jr & Stenseth, 1992). The importance of dispersal for species persistence and evolution has been addressed by numerous works using different but complementary approaches (e.g. evolutionary, demographic and genetic) (Lambin *et al.*, 2001; Lidicker Jr & Stenseth, 1992). In mammals, for example, dispersal from the natal range to areas where individuals will reproduce (natal dispersal) and movements of adults between reproduction areas (breeding dispersal) are both related to kin interaction, avoidance of inbreeding, resource competition, etc. The three-stages decision of individuals to disperse (leaving, travelling and arriving) has consequences at both individual and population levels. Consequently, dispersal is studied by a number of disciplines each focussing on its consequences on different demographic and evolutionary processes at different scales (gene flow, spread of diseases, alien species invasiveness...). This multidisciplinary approach has established an optimal theoretical and practical framework for the understanding of dispersal. Therefore, movement of individuals may be seen as the glue that holds populations connected (Hansson, 1991) but also as the *glue* that connects different scales and disciplines.

Mammals are distributed through heterogeneous environments (landscapes) which, depending on requirements of species to disperse, might constrain individuals movements. On a genetic perspective, dispersal will not be effective (i.e. it will not influence on the genetic structure of the population) until those movements were followed by the successful reproduction of migrants. This is one of the main differences between dispersal rates estimated by capture-mark-recapture and by

genetic methods. Accordingly, we can estimate gene flow between populations from their genetic structure.

Landscape genetics (Manel et al., 2003) appeared as analytical approach to quantify the effects of geographical and environmental features on individual relatedness and population genetic structure. Early studies included matrix attributes *a posteriori*, that is, landscape features were overlaid on the estimated genetic structure and it was checked if the latter is correlated with the former (Aars et al., 1998; Piertney et al., 1998). More recent studies have tried to quantify how the landscape, considered as a “friction surface” to dispersal, affects individual movements, gene flow and genetic structure beyond the basic Euclidean distance (Berthier et al., 2005; Vos et al., 2001). A better correlation of genetic distances to a landscape modified-distance (i.e. deviations from the straight line as would result of habitat preferences of individuals) than to the Euclidean distance may indicate landscape influences on individual movements during dispersal. Therefore, the spatial distribution of habitat patches and the permeability of the intervening landscape might considerably influence dispersal patterns, especially in species with a highly specialized habitat requirement.

One important caveat of direct methods for measuring dispersal is a known bias towards short-distance dispersal events, due to broad home ranges, difficulties on the accessibility to the study area or the elusive behaviours of animals (Koenig *et al.*, 1996). The use of genetic information may mitigate some biases on the estimation of dispersal distances and rates and yield gene flow estimates that reflect effective (rather than total) dispersal. Deviations from panmixia can be related with dispersal and breeding patterns associated with social organization (Scribner & Chesser, 1993), strong phylopatry in one sex (Chessser, 1991a; Chessser, 1991b) and habitat

requirements (Brouat *et al.*, 2004), independently of the spatial scale considered (Hirota *et al.*, 2004; Verbeylen *et al.*, 2003); Stewart *et al.* 1999). Besides, sex biased dispersal also deviates population structure from panmixia and it might have important consequences in both theoretical and field studies (Pusey, 1987). Apart from field-based direct methods, there are nowadays two indirect (genetic) methods to infer sex-biased dispersal: those based on biparentally inherited markers (e.g. comparing male/female relatedness,  $F_{ST}$  and assignment probabilities) and those relying on uniparentally inherited markers (e.g. mitochondrial markers) (Goudet *et al.*, 2002; Prugnolle & de Meeus, 2002; Slatkin, 1985)

Small mammals have been widely used as model species to test hypothesis on dispersal patterns and individual movements (Lidicker Jr., 1995) through individual labelling (ear tags, passive integrated transponders, photo-identification...) and radiotelemetry techniques. With the advance of recent molecular-based techniques, small mammals are being also used to compare direct and indirect methods for inferring dispersal. In this study we use the habitat specialist Southern water vole (SWV) (*Arvicola sapidus*) (Muridae, Rodentia) as a model species to test the influence of the landscape matrix on dispersal patterns using a genetic approach. During the last years, the acquired knowledge on SWV have mainly focused on its natural history and habitat preferences and little attention have been paid to dispersal habits (Román, 2007; Ventura, 1988). SWV inhabits in Iberia and France, and is closely associated to permanent water bodies in most of its distribution range, although it also uses high vegetation cover as refuges against their common predators when draught periods dry out water bodies (Fedriani *et al.*, 2002; Román, 2007). The Natural region of Doñana (SW of Spain) is a well known area where SWV are restricted to a network of temporary ponds (see *Study area and sample*

*collection* below). Fragmented and unstable habitat patches have favoured here a metapopulation dynamics with frequent extinction and colonization of individual ponds, whose transcendence at larger scales may be hampered by a relatively high density of suitable habitat patches and high recolonization capacity. Previous studies on SWV have suggested sex-biased and long-distance dispersal, although some technical caveats (e.g. inadequacy of the scale of the study area) impede an accurate estimate of dispersal movements.

Here, we address the following four main objectives. First, we describe the genetic diversity and structure of water voles populations in Doñana using both microsatellite markers and mitochondrial sequences. Second, we analyse gene flow patterns through indirect and direct approaches. Our main prediction is that genetic structure will fit a relatively steep isolation-by-distance pattern, suggesting a restricted pattern of gene flow between populations. Third, we evaluate the effects of landscape structure on patterns of individual dispersal using Geographic Information Systems (GIS). An obvious prediction is a strong effect of those landscape variables that determine the species spatial distribution (i.e. water bodies –ponds, drainages- and vegetation cover) on the genetic structure of SWV populations. Fourth, we test sex-biased dispersal by i) analysing isolation by distance patterns using independent datasets per sex and a randomisation procedure to compare sex-specific indices of differentiation, and by ii) comparing genetic structure inferred with maternally inherited markers (mitochondrial control region) with that obtained with biparentally inherited (autosomal microsatellite markers). As in most mammals and field studies suggested, we would expect that females are the philopatric sex (Greenwood, 1980; Román, 2007).

## MATERIAL AND METHODS

### Study area and sample collection

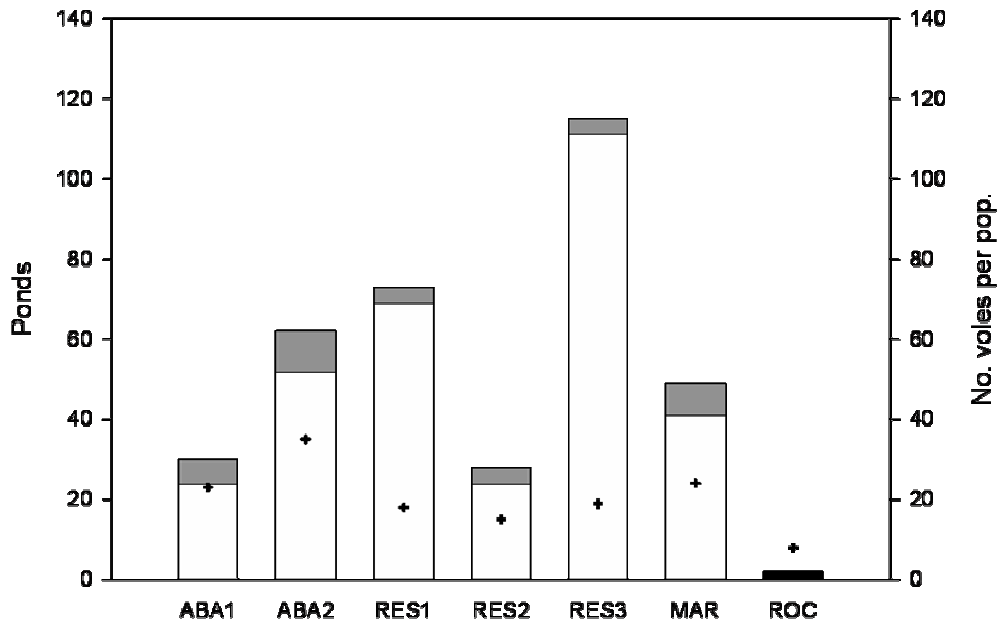
This study was conducted in the Natural Region of Doñana (Southwest of Spain), where climate is typically Mediterranean with hot and dry summers and mild wet winters. The average annual precipitation is about 600 mm, although there is an interannual heterogeneity characterized by cycles of several dry years (i.e. rainfall about 300 mm) interspersed with wet years (i.e. annual rainfall around 900 mm). The study area (600 km<sup>2</sup>) is flat and mostly near sea level. Doñana is a recently formed natural region caused by the replenishment of the old River Guadalquivir gulf by muddy riverine sediments and posterior superposition of five different aeolian systems (sandy marine sediments). These recurrent deposits of sands (older from coastline to inland) and a relatively shallow phreatic level together with winter rains give rise to more than 600 water bodies whose area and shape varies and range from 0.02 to 33 ha (Bravo & Montes, 1992). Smaller ponds tend to dry out during summer, in relation to their size and annual rainfalls. Bank vegetation varies within and all around the perimeter of ponds (open, marshy, canopy, brushes and grasses) and this will condition the amount of food resources and shelter available for SWV. There are three predominant ecosystems in Doñana (from coastline to inland: mobile dunes, fixed dunes and marshes). Fixed dunes are composed by a mixture of different degradation stages of Mediterranean scrubland directly related to the extant vegetation cover. The stream of La Rocina and its tributaries overwhelms the northern region of Doñana and implies a continuous habitat that might be potentially occupied by SWV. Rainfalls, the spatial layout of water bodies, the stream of La Rocina and its tributaries and the distribution of the ecosystems and associated vegetation determine a rich and heterogeneous landscape in



Doñana. On average, less than 2% of the study area is considered as suitable habitat for the presence of SWV.

Colonies of SWV are linked to the vegetation associated to the perimeter of water bodies. During rainfalls, individuals use water and vegetation for predator avoidance. However, Doñana registers drought periods during late-spring and summer seasons that severely influence the distribution and occupancy of these temporary ponds by SWV. During drought periods, movements of SWV are limited to the water bodies and associated vegetation they occupy, and they do not disperse until the next rain season. Previous studies have reported two reproductive maxima (autumn and spring) interrupted by a complete stop in reproduction during summer (Fedriani *et al.*, 2002; Román, 2007).

We used patterns of ponds occupancy by water voles in Doñana described in (Fedriani *et al.*, 2002) to delimit six circular areas of 3 kms in diameter (Abalario1, Abalario2, Reserva1, Reserva2, Reserva3 and Marismillas –ABA1, ABA2, RES1, RES2, RES3 and MAR, respectively-) and also one stream (Rocina -ROC-) running through Doñana (figure 1) that define the seven sampling units used in this study. 321 ponds were visited and screened for SWV trails and finally 142 individuals were trapped in 36 ponds and one stream (ROC) (figure 2). ABA1, ABA2, RES1, RES2 and RES3 were sampled in 2000 and 2001. Individuals from ROC and MAR were trapped in 2004. All trapped individuals were ear-punched and live-released. We also registered their UTM coordinates. Samples were stored at 4°C in 95% ethanol and EDTA 100 µM (Dessauer *et al.*, 1996).



**Figure 2** Overall ponds reviewed in each of the seven geographic areas and number of ponds where Southern water voles were captured (grey areas). Crosses inside each histogram show the sample size per geographic area. ROC is a stream is considered as a continuous water body

### Genotyping and sequencing

We tested 24 pairs of microsatellite primers developed for *A. terrestris* (Stewart *et al.*, 1999; Stewart *et al.*, 1998) or *Microtus oeconomus* (Van De Zande *et al.*, 2000), and eight of these amplified directly and proved to be polymorphic in *A. sapidus*. Two more markers were used after primer redesign based on *A. sapidus* sequences obtained with the original primers. These new primers are AV10 reverse (5'-CAAGGCTTGGAGCTTGGATA-3') and AV14 forward (5'-

TCCTCCCTCCCCAGCAAT-3') and reverse (5'-GCAGAAGGGGGCAGATAAT-3').

PCR products were indirectly labeled using a M13 5' extension (Schuelke, 2000), except AV1, AV10 and AV14 for which 5'- labeled forward primers were used. Amplification reactions were performed in a 25- $\mu$ L volume including 50-100ng template DNA, 1xTaq buffer, 2mM MgCl<sub>2</sub> (0.9mM for AV1), 0.25mM dNTPs, 0.25  $\mu$ M sequence-specific reverse and fluorescently-labeled M13 primers, 0.017  $\mu$ M sequence-specific forward primer with M13 extension, 0.1  $\mu$ g/ $\mu$ l BSA (Roche) and 0.5 U/ $\mu$ l *Taq* DNA polymerase (BIOTAQ™ DNA Polymerase, Bioline). PCRs involved 94°C for 5min, 17 touchdown cycles 92°C for 30s, 66°C for 30s and 72°C for 30s, decreasing one Celsius degree per cycle, and followed by 27 cycles of 92°C for 30s, 50°C for 30s and 72°C for 30s. PCRs for AV1 were as follows: 94°C for 5min, 34 cycles of 92°C for 30s, 60°C for 30s. and 72°C for 30s. All PCR reactions finished with 72°C for 5min. PCR products were checked on a 2% agarose gel, run on an Applied Biosystems 3130xl Genetic Analyzer and scored with GeneMapper software v3.0 (Applied Biosystems).

1024-bp of the mitochondrial control region were sequenced from a subset (n=47) of the total sample size using primers F15708 and R92 (Piertney et al., 2005), primer 5'-TCCCCACCATCAGCACCCAAAGC-3' designed by (Stacy *et al.*, 1997) and four specifically designed internal primers whose combination yield partially overlapping fragments (F15816, 5'-ATGTTTTATCGTCCATACGTTCC-3'; F15872, 5'-AATCAGCCCATGCCTAACAT-3'; R15946, 5'-TAGCCGTCAAGGCATGAAG-3'; RCRasa 5'-AAAAACAACCTCAAATTCCTAAAA-3') (Centeno-Cuadros and Godoy, chapter

1). PCR amplifications were performed as follows: 94°C for 5min, 40 cycles at 92°C for 30s, 62°C for 30s and 72°C for 30s, finishing with 72°C for 5min. We also included with each set of PCR reactions positive and negative (water) DNA controls to monitor the performance of the process and contamination. 5  $\mu$ L of PCR products were purified with 2  $\mu$ L  $\square$  ExoSAP-IT enzyme (Exonuclease I and Shrimp Alkaline Phosphatase in buffer) (USB Corporation), incubated during 15 minutes at 37°C and inactivated 15 minutes at 80°C. Sequencing reactions were performed using the Applied Biosystems BigDye® Terminator Cycle Sequencing Kit v. 1.1 following the manufacturer's instructions, and the same primers used for the amplification. Reactions were analysed in an Applied Biosystems 3130xl Genetic Analyzer. Forward and reverse sequences for each PCR product were edited and assembled using Sequencher 4.6 (Gene Codes Corporation, Ann Arbor, MI).

## **Genetic data analysis**

### *Mitochondrial data*

Nucleotide and haplotype diversities, number of segregating sites and average number of pairwise differences per sampled area and overall samples were estimated with DnaSP 4.5 (Rozas *et al.*, 2003). We calculated population pairwise  $F_{ST}$  (Hudson *et al.*, 1992) using control region sequences as implemented in ARLEQUIN v3.0 (Excoffier *et al.*, 2005). Since  $F$ -statistics derived from mitochondrial and nuclear data assume equilibrium under different effective populations sizes, they need to be calibrated before comparison. For this purpose,

we used the expression  $F_{ST(nuclear)} = F_{ST(mitochondrial)} / (4-3 F_{ST(mitochondrial)})$  derived from (Crochet, 2000)

#### *Microsatellite data*

GENEPOP v3.4 (Raymond & Rousset, 1995) was used to test for deviation of Hardy-Weinberg equilibrium at each locus within each sampling unit. Levels of observed ( $H_O$ ) and unbiased estimates of expected heterozygotes (Nei, 1978) were estimated with GENETIX v4.05 (Belkhir *et al.*, 1996) for each sampling unit by locus and over all loci. FSTAT v 2.9.3 (Goudet, 2001) was used to calculate allelic richness per locus and over all loci per sampled area.

Overall Wright's  $F_{ST}$  and linearized pairwise F-statistics [ $F_{ST} / (1- F_{ST})$ ] were estimated with GENETIX 4.05.2 (Belkhir *et al.*, 1996) (population level). We also computed genetic distances ( $a_i$ ) defined by (Rousset, 2000) between 142 individuals using GENEPOP v3.4. A Factorial Correspondence Analysis (FCA) of individual multilocus genotypes was performed with GENETIX 4.05.2 to visualize genetic variation within and among geographic sampled areas by plotting each individual's genotype in a three-dimensional space. We also performed a Spatial Analysis of Molecular Variance (SAMOVA) (Dupanloup *et al.*, 2002), where individuals trapped in the same pond were grouped into  $k$  continuously homogenous populations and genetically differentiated maximizing the proportion of total genetic variance ( $\Phi_{CT}$ ), avoiding any *a priori* definition of populations. Since  $\Phi_{CT}$  values are expected to increase as  $k$  increases, we run SAMOVA with 1000 simulated annealing processes for  $k$  values ranging from 2 to 10. If our sampling scale were adequate to dispersal patterns of SWV, we would expect  $k = 7$  (i.e. the

number of sampling units used in this study), where each group ( $k$ ) must include those ponds embraced by the 3kms diameter. On the other hand,  $k < 7$  would suggest some degree of clustering of geographic areas.

47 samples from the seven geographic areas were used to compare genetic structures inferred with bi- and uniparentally (maternal) inherited markers (microsatellites and mitochondrial DNA, respectively).

### **Isolation by distance and landscape genetics**

We used the spatial weighted mean coordinate for all individuals trapped as the geographic coordinate of each sampling area (Jenness, 2004) for the calculation of both Euclidean and Least Cost Distances (LCD) (see below). Linearized pairwise  $F_{ST}$  values obtained with nuclear (microsatellites) and mitochondrial markers were plotted against log-transformed Euclidean distances (Rousset, 1997) to test for a negative correlation of gene flow with geographic distance. IBD analyses were performed using the Isolation By Distance Web Service v3.05 (IBDWS 3.05) (Jensen *et al.*, 2005). IBDWS 3.05 estimated the normalized statistic ( $r$ ) and its statistical significance after 30000 permutations using a Mantel Test.

We evaluated the effects of three landscape features on the genetic structure of SWV in Doñana: ponds surface, drainages length and vegetation cover. Here, Least Cost Distance (LCD) is used as the distance covered by an individual on its movement between two localities if one of these landscape attributes were favored (Arnaud, 2003; Coulon *et al.*, 2004; Vignieri, 2005). We divided vegetation cover into five different categories in a decreased order of assumed preference by water voles: water, scrubland cover higher than 50%, scrubland cover between 20-50%, vegetation cover less than 20% and others (farming, buildings...). The whole region

of Doñana was divided into a lattice of 500 m cells and values for each of the three variables under consideration were added within them. We obtained a cost surface where the lowest cost values were assigned to landscape cells that maximized ponds surfaces, drainages or vegetation cover. Least cost paths were then calculated based on the cost of dispersing across each type of landscape cell with the extension PathMatrix (Ray, 2005) for ArcView™3.2 (Environmental Science Research Institute, Redlands, USA). We also performed cost analyses (Schadt *et al.*, 2002) to evaluate the robustness of our results when varying the assigned cost value over a wide range of maximum values. To test for a correlation between LCD and genetic distance and between these two variables once the effect of the Euclidean distances is discounted, we used Mantel and Partial Mantel tests, respectively. Both Mantel and Partial Mantel test were performed in IBDWS 3.05. We also used the delta difference among the respective corrected Akaike's Information Criterion (AICc) as a statistical test to evaluate the goodness of fit of each model. AICc were calculated from the least-squares regressions and adding a second order correction because of the small sample size as follows  $AICc = 2k + n \cdot \ln(RSS/n) + [2k \cdot (k+1)] / (n-k-1)$ , where  $k$  is the number of landscape features in the model,  $n$  is the number of populations and  $RSS$  is the residual sum of squares. Because of the non-independency of the data, no proper method to estimate AIC for pairwise data has been developed. Nevertheless, we used the number of populations as degrees of freedom instead of the number of pairwise comparisons as a better fit of Mantel's test to p-values of ordinary least squares regression has been shown (Hellberg, 1994) (Koizumi *et al.*, 2006).

When delta difference between two models is lower than two, these models do not statistically differ in explaining the variance of genetic distances between populations

and they can be regarded as equally likely (Burnham & Anderson, 2002). The higher the difference of delta values, the better the model with the lowest AICc. According to our null hypothesis, we would expect delta values higher than two between Euclidean line distance models and other landscape modified line models.

### **Sex-biased dispersal**

Sex-biased dispersal promotes differences in genetic structure between sexes. We tested differences in dispersal rates between sexes using two approaches. First, we plotted individual genetic distances (see above) relative to geographic distance using independent datasets for males ( $n = 68$ ) and females ( $n = 74$ ) and compared the slopes of their respective IBD tests. Moreover, we used FSTAT v2.9.3 to calculate  $F_{ST}$  and an assignment index (AI) per sex and population; these two indices have been shown to be most powerful and least sensitive to changes in magnitude of sex biased dispersal (Goudet *et al.*, 2002). Higher  $F_{ST}$  values are expected in the philopatric sex, whereas negative mean assignment indexes (mAIC) or larger variances of the AI (vAIC) would be observed for the dispersing sex. The method proposed by (Goudet *et al.*, 2002) is based on randomisation procedures and eliminate pseudoreplication problems arising from the comparison between sexes. We also performed this analysis using a dataset composed by adults trapped in ABA1 and ABA2 before reproductive dispersal ( $N_{females} = 14$ ,  $N_{males} = 16$ ) in order to avoid the effect caused by the introduction of immigrants into the gene pool.



Assuming IBD in a two-dimensional space, we can estimate the average squared axial parent-offspring distance ( $\sigma^2$ ), which can be interpreted as an average dispersal distance, using the slope ( $b$ ) of the regression of interindividual pairwise genetic distances ( $a_r$ , see above) on the geographical distances (Rousset, 2000). Considering 7.33 ind./km<sup>2</sup> as effective density ( $D$ ) (A. Centeno-Cuadros, unpublished data) we can estimate a sex-specific  $\sigma^2$  using their respective slope obtained in the IBD analyses (equation 1 in (Broquet *et al.*, 2006)). Genepop v3.4 was used to infer  $b$  using the whole (142 individuals) and sex-specific datasets. Higher values of  $b$  would result from lower gene flow and lower average dispersal distance.

### **Contemporary gene flow**

Direct estimates of recent migration rates were obtained with the Bayesian multilocus genotyping procedure implemented in BAYESASS (Wilson & Rannala, 2003). This method does not assume migration-drift nor Hardy-Weinberg equilibrium, two common assumptions that are rarely met in species with high generational overlap and intense population dynamics. We run three replicates of this Markov Chain Monte-Carlo (MCMC) based approach for a total of 3x10<sup>6</sup> iterations and discarded the first 10<sup>6</sup> iterations as burning to assure that chains reached the stationarity. The inference of posterior probability distributions of parameters was performed using samples collected every 2000 iterations.

## **RESULTS**

### **Mitochondrial variation**

Ten mitochondrial haplotypes defined by 17 segregating sites were found after the sequencing a 1024bp D-Loop fragment in 47 SWV individuals, (table 2a). Most variation was due to transitions (96.24%), with only 3.76% transversions and no indels (average number of nucleotide differences,  $k = 6.278$ ). There was a high variability on the distribution of nucleotide diversity ( $\pi$ ) throughout Doñana (table 2b) with observed values ranging from  $\pi=0$  (Marismillas) to  $\pi=0.006$  (Reserva3). Haplotype diversities ( $H$ ) were slightly higher in Northern localities (maximum  $H = 0.929$ , Abalario 2). We found a strong mitochondrial structure (overall  $F_{ST} = 0.3682$ ), which would translate to  $F_{ST} = 0.1272$  for comparison with microsatellites-inferred  $F_{ST}$ , although no haplotype was private to any one area. Pairwise  $F_{ST}$  ranged from -0.0177 to 0.4717 with no clear pattern of isolation by distance (Figure 4; see below).

a)

	1	10	17	Sampled areas						
*	*	*	ABA1	ABA2	RES1	RES2	RES3	ROC	MAR	Hap.freq
Hap_1	GTACTTTCCTTCCTTCT		2	1						3
Hap_2	.CGTC.C.TCCTT..AC		1		1	1	2			5
Hap_3	.....C.....TT.....		4	2				1		7
Hap_4	.CG.CCCTTCC.T.....						2			2
Hap_5	ACG.C.CTTCC.T.....					5	1		8	14

Hap_6	. . . . . C . . . . . TT . CT .	1	1	3		1	3	9
Hap_7	A . . . . . A . . C . T . CT .			1				1
Hap_8	. CG . C . C . TCC . T . . . .		1				1	2
Hap_9	. CG . CCC . TCC . TC . . . .		2				1	3
Hap_10	. CG . CCC . TCC . T . . . .		1					1

---

Sample size    8    8    5    6    6    6    8    47

**b)**

	N	S	$\pi \pm SD$	h	H (SD)	k
Abalarío1	8	13	0.00391 (0.00151)	4	0.750 (0.139)	4.000
Abalarío2	8	13	0.00579 (0.00084)	6	0.929 (0.084)	5.929
Reserva1	5	14	0.00566 (0.00222)	3	0.700 (0.218)	5.800
Reserva2	6	6	0.00195 (0.00126)	2	0.333 (0.215)	2.000
Reserva3	6	14	0.00605 (0.00157)	4	0.867 (0.129)	6.200
Marismillas	8	0	0.00000 (0.00000)	1	0.000 (0.000)	0.000
Rocina		6	11    0.00547 (0.00154)	4	0.800 (0.172)	
	5.600					
Overall	47	17	0.00613(0.00026)	10	0.846 (0.030)	6.278

**Table 2 a)** Haplotypes of Control Region and their geographical distribution all over the Natural Region of Doñana **b)** Control Region nucleotide diversity by sampling area and overall study system. Sample size ( $N$ ), number of segregating sites ( $S$ ), nucleotide diversity ( $\pi$ ), number of haplotypes ( $h$ ), haplotype diversity ( $H$ ) and mean number of nucleotide differences ( $k$ ) are reported. Standard Deviations are given in parenthesis ( $SD$ ).

### Microsatellite diversity and structure

142 individuals were genotyped using 10 polymorphic loci at 7 geographic populations throughout the region of Doñana. Jackknife over loci revealed a major contribution of locus AV14-2 to the within-population deficit of heterocigotes ( $F_{IS} =$

0.052; 95% CI [0.008-0.117]), probably due to null alleles; consequently, overall tests of HWE for each population were performed without this locus. Nevertheless, no significant differences were found when using AV14-2 in estimations of  $F_{ST}$  (one tailed t-paired test  $p = 0.098$ ). Significant heterozygote deficits across loci were detected in the three areas within Biological Reserve of Doñana (RES1, RES2 and RES3), whereas a heterozygote excess was observed at RES1 and ROC (Table 1).

ABALARIO1 (n=23)				ABALARIO2 (n=35)				
	HWE	AR	H(E)n.b	H(O)	HWE	AR	H(E)n.b	H(O)
AV1	0.3032	5.374	0.7333	0.8	0.778	4.384	0.5787	0.6087
AV3	0.3726	5.115	0.7546	0.7391	0.6425	4.729	0.6749	0.7429
AV4	0.5692	7.581	0.8406	0.8261	0.027	10.713	0.9404	0.9143
AV7	0.0618	4.577	0.7705	0.6	0.6057	4.472	0.6792	0.7391
AV8	0.5268	7.529	0.8705	0.8696	0.0966	7.723	0.8654	0.7714
AV9	0.0038	5.548	0.6396	0.4783	0.1723	6.937	0.8211	0.7714
AV10	0.6778	5.367	0.7952	0.7391	0.0354	5.337	0.7805	0.6857
AV15	0.1495	4.273	0.7411	0.8696	0.6962	6.419	0.8340	0.8000
Moe7	0.6998	3.465	0.6232	0.5217	0.4408	5.173	0.7175	0.6970
Overall	0.0513		0.7630	0.7096	0.0819		0.7774	0.7159

RESERVA1 (n=18)				RESERVA2 (n=15)				RESERVA3 (n=19)				
	HWE	AR	H(E)n.b	H(O)	HWE	AR	H(E)n.b	H(O)	HWE	AR	H(E)n.b	H(O)
AV1	0.929	3.423	0.6175	0.7222	0.4704	4.581	0.7632	0.8	0.0209	4.529	0.7508	0.6111
AV3	0.0106	4.785	0.7429	0.6667	0.1416	5.020	0.6782	0.6	0.1432	4.921	0.7454	0.8421
AV4	0.0497	7.427	0.8556	0.8889	0.0003	11.591	0.954	0.9333	0.128	8.902	0.8606	0.7368
AV7	0.1957	3.696	0.7032	0.9444	0.0159	4.889	0.7839	0.6667	0.2892	3.362	0.4921	0.5556
AV8	0.1966	5.647	0.7492	0.7222	0.0372	6.699	0.8207	0.7333	0.0421	7.078	0.8663	0.8947
AV9	0.0779	4.172	0.6048	0.6111	0.192	5.612	0.7724	0.6667	0.303	4.087	0.5462	0.5789
AV10	0.0788	2.889	0.5286	0.6667	0.4263	4.731	0.5839	0.5333	0.3413	3.056	0.3300	0.3158
AV15	0.2523	4.823	0.7175	0.7222	0.682	5.301	0.7356	0.8	0.8774	5.445	0.8137	0.6842
Moe7	0.198	3.616	0.6492	0.6667	0.8585	4.060	0.6759	0.7333	0.6418	3.407	0.6657	0.7895
Overall	0.004		0.6984	0.7222	0.001		0.7632	0.7267	0.0352		0.6860	0.6746

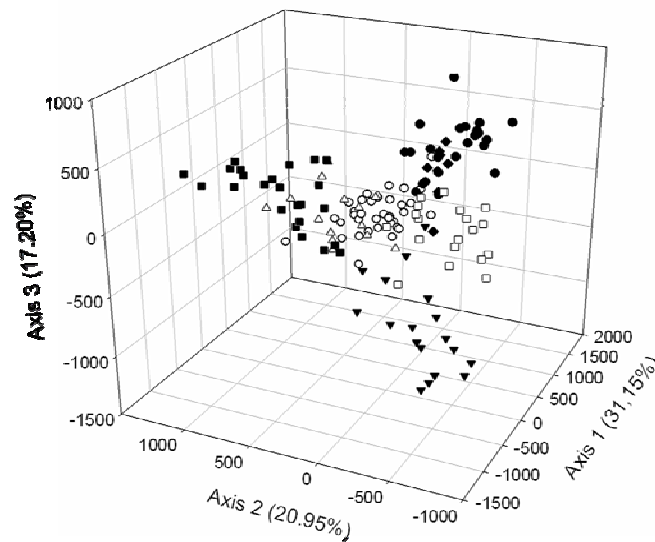
MARISMILLAS (n=24)				ROCINA (n=8)				
	HWE	AR	H(E)n.b	H(O)	HWE	AR	H(E)n.b	H(O)
	0.7119	0.625	0.9145	5	0.775	1		
AV3	0.0677	3.894	0.4805	0.375	0.0689	6	0.8167	0.5
AV4	0.1053	7.208	0.8617	0.75	0.8615	9	0.8917	0.875
AV7	0.643	3.849	0.6729	0.6667	0.8851	4	0.6917	0.75
AV8	0.3588	6.62	0.8085	0.875	0.9729	7	0.75	0.875
AV9	0.3055	4.396	0.727	0.875	0.4095	7	0.775	0.625
AV10	0.1261	3.465	0.4947	0.5	0.9169	4	0.725	0.875
AV15	0.1898	4.939	0.7518	0.75	0.237	6	0.8167	0.75
Moe7	0.6773	4.25	0.7119	0.6667	0.8504	4	0.675	0.75
Overall	0.0518		0.7035	0.6542	0.8817		0.7808	0.7875

	AV1	0.071	4.864
--	-----	-------	-------

**Table 1** Test of deviation of Hardy-Weinberg equilibrium (HWE), allelic richness (AR) and unbiased expected (H(E)) and observed (H(O)) heterozygotes per locus and overall geographic area.

Individuals trapped within the same sampled area tended to cluster together in FCA plots, but with some degree of overlap with other areas (figure 3).



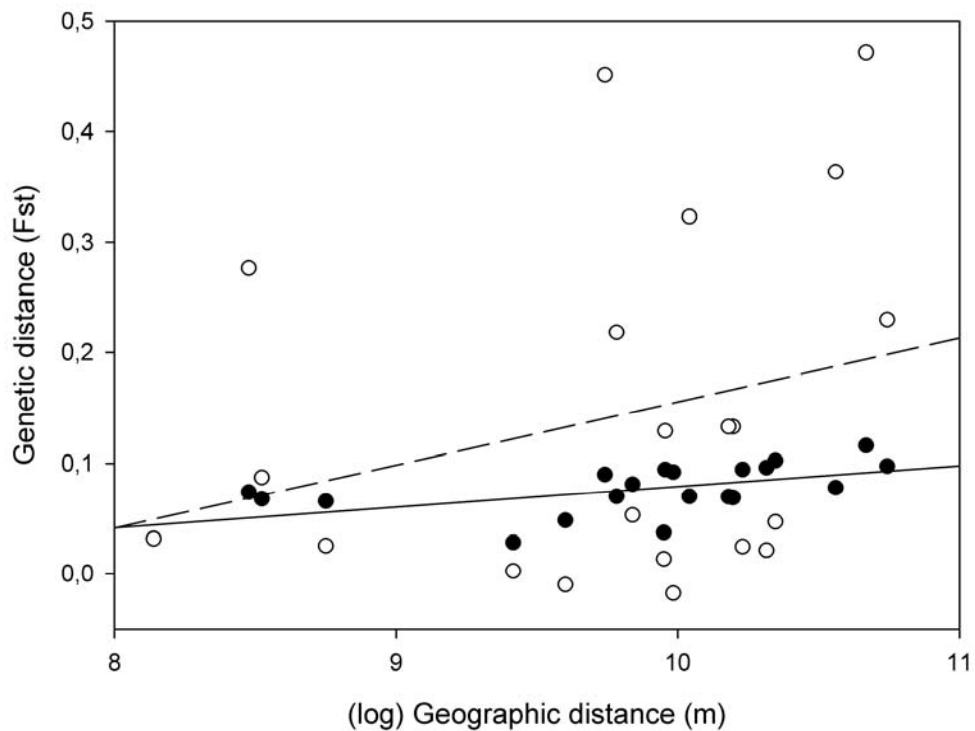
**Figure 3** Factorial Correspondence Analysis performed with 142 individuals trapped in Doñana. ABA1: filled circles; ABA2: empty circles; RES1: filled triangles; RES2: empty triangles; RES3: empty squares; MAR: filled squares; ROC: filled rhombus.

The distribution of each of the sampling units in Doñana within the FCA plot coincided with their geographical location (e.g. ABA1 and MAR –the most geographically distant sampling units- are also the most distant in the FCA). This is reflected in a moderate genetic structure among all seven geographic populations (microsatellites: overall  $F_{ST} = 0.072$ , 95% confidence interval [0.058-0.089], pairwise values ranging from 0.028 to 0.116). However, the hierarchical and spatial analysis of genetic structure (SAMOVA)

does not reveal any statistically significant spatial subdivision on the overall distribution area as shown by a decrease in  $\Phi_{CT}$  as  $k$  increases beyond one.

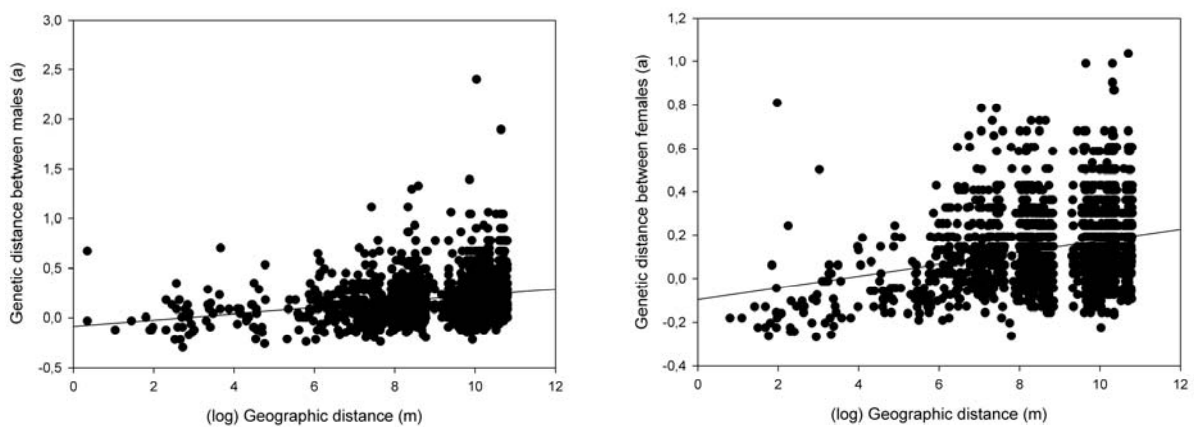
### Isolation by distance and landscape genetics

Geographic (Euclidean) distance between sampled areas explained a 34% of the microsatellite (nuclear) genetic variance (IBD, Mantel Test:  $r=0.5844$ ,  $p<0.01$ ) (figure 4).



**Figure 4** Regression of interpopulation genetic distances ( $F_{ST}$ ) inferred with microsatellites (filled circles, solid line) and mitochondrial Control Region (after calibration, see text) (open circles, dashed line) on the natural logarithm of geographic distance among populations

However, this correlation was not significant when genetic distances inferred with D-Loop sequences were considered ( $r = 0.2501$ ,  $p = 0.1303$ ) (fig.4). At the individual level, a significant pattern of IBD was obtained when considering all individuals ( $r = 0.288$ ,  $p < 0.0001$ ), as well as only males ( $r = 0.2830$ ) or females ( $r = 0.2541$ ) separately (all  $p < 0.0001$ ) (figure 5).



**Figure 5** Regression of interindividual genetic distances  $a$  (Rousset, 2000) between males (a) and females (b) on the natural logarithm of geographic distance among individuals

Least Cost Distance inferred with landscape attributes did not vary when different cost values were assigned ( $CV_{\text{POND}} = 0.096 \pm 0.004$ ;  $CV_{\text{DRE}} = 0.171 \pm 0.007$ ), revealing that LCDs estimated were all similar independently of the assigned costs. Vegetation-, drainages- and ponds- LCDs were all positively correlated with genetic distances in Doñana (regression between genetic and least cost distances:  $LCD_{\text{VEG}}$ ,  $r = 0.5386$ ;  $LCD_{\text{PONDS}}$ :  $r = 0.6078$ ;  $LCD_{\text{DRA}}$ :  $r = 0.6089$ , all  $p < 0.05$ ). However, none of these landscape effects remained significant once the effects of



Euclidean distance on genetic structure were discounted (all Partial Mantel test,  $p > 0.05$ ). Moreover, delta differences of AICc between the model with lowest AICc (Euclidean distance) and landscape modified line models were always less than two.

### **Sex-biased dispersal**

Observed differentiation indices were generally higher for males than for females (females:  $F_{ST} = 0.0652$ ,  $mAIC = -0.01744$ ,  $vAIC = 24.19336$ ; males:  $F_{ST} = 0.0805$ ,  $mAIC = 0.01926$ ,  $vAIC = 14.09704$ ), although none of sex-biased dispersal tests rejected the null hypothesis of equal dispersal between sexes (all  $p > 0.05$ ) at the overall scale of study. The same result was found when only adults trapped in ABA1 and ABA2 after dispersal but before breeding were considered ( $F_{ST}$ ,  $mAIC$  and  $vAIC$  tests, all  $p > 0.05$ ).

Finally, a very similar average axial parent-offspring distances was estimated for males and females ( $\bar{\Delta}_o = 732$ ,  $\bar{\Delta}_m = 668$  m,  $\bar{\Delta}_f = 661$ m) from the slopes of the individual-based IBD analyses ( $b_o = 0.0203$ ,  $b_m = 0.0243$ ,  $b_f = 0.0248$ ) and considering an effective density of 7.33 ind./km<sup>2</sup>.

### **Contemporary gene flow**

Mean posterior probabilities for migration rates between sampling areas are reported in table 3. Most sampling locations in Doñana showed a low proportion of recent immigrants. However, northern regions apparently exchange more frequently individuals per generation, as shown by higher migration rates between the three northernmost sampling areas (migration rates from ABA1 to ABA2: 0.1131, and from ABA1 to ROC: 0.1616). Nonetheless, within the more southern region RES2 received a fairly large proportion of migrants from MAR ( $m = 0.2199$ ), suggesting a moderate source-sink dynamic within the studied area.

To

From	ABA1	ABA2	RES1	RES2	RES3	MAR	ROC
ABA1	<u>0.979</u>	<u>0.113</u>	0.009	0.011	0.006	0.003	<u>0.1616</u>
ABA2	0.004	<u>0.852</u>	0.007	0.009	0.004	0.004	0.017
RES1	0.0031	0.0065	<u>0.9487</u>	0.0186	0.0110	0.0036	0.0317
RES2	0.003	0.006	0.005	<u>0.686</u>	0.004	0.005	0.018
RES3	0.003	0.007	0.008	0.045	<u>0.965</u>	0.003	0.054
MAR	0.003	0.006	0.017	<u>0.219</u>	0.004	<u>0.977</u>	0.016
ROC	0.003	0.007	0.004	0.009	0.004	0.003	<u>0.701</u>

**Table 3** Means of the posterior distributions of  $m$ , the migration rate into each population. The populations from which each individual was sampled are listed in the columns, while the populations from which they migrated are listed in the rows. Values along the diagonal are the proportions of individuals derived from the source populations each generation. Migration rates  $> 0.10$  are underlined.

## DISCUSSION

In this study we have provided new insights on the dispersal biology and genetic structure of Southern water voles in a Mediterranean patchy habitat, increasing the acquired knowledge on SWV already existing more focused on natural history and habitat preferences (Román, 2007; Ventura, 1988).

Overall mitochondrial diversity was moderate to high, but highly variable among populations. None of the ten described haplotypes were locally restricted to our sampling areas, although four out of them were distributed only in the Northern region of Doñana (haplotypes 1, 3, 8 and 9 located in ABA1, ABA2 and ROC)

(table 2a). This distribution of haplotypes in space suggests high connectivity among these three populations, a pattern which is not observed in central (Reserve) or Southern (Marismillas) populations. On the other hand, we cannot relate the absence of diversity in Marismillas ( $\pi=0$ , table 2b) to sampling relatives or whether these low levels of haplotype diversity located in the southernmost population resulted from a very recent founder effect.

A set of 10 microsatellite markers fulfilled the polymorphism required for individual and population characterizations (range of  $H_e$  [0.6860 - 0.7808]). Only one marker (AV14-2, the specific *A. sapidus* homologous marker to *A. terrestris* AV14) showed a consistently significant deficit of heterozygosity across populations, that could be due to the occurrence of null alleles. Population microsatellite diversity was moderate to high. Three populations (RES1, RES2 and RES3) were in Hardy-Weinberg disequilibrium, even when the AV14-2 marker was excluded, showing either excess (RES1) or deficit of heterozygotes (RES2, RES3). Deviations from HWE has been previously reported for some populations of Eurasian water vole (Berthier *et al.*, 2005).

Although the distribution of genetic diversity is not apparently explained by the spatial distribution of sampled populations (SAMOVA,  $k = 1$ ), we observed both variation within and differentiation among these geographic locations and discarded any spatial random distribution of genotypes (FCA, fig.3). This genetic differentiation is also reflected in moderate levels of genetic structure in nuclear and mitochondrial markers ( $F_{ST}$  values, 0.072 and 0.1272, respectively). Sampling of sibs may bias the allele frequencies. However, the influence of temporal variation (i.e. differences on genotypic frequencies between generations) only scarcely

influence the genetic structure in the region Doñana (Centeno-Cuadros, unpublished data). Three out of seven populations (RES1, RES2 and RES3) showed significant deviations from HWE (table 1). Whereas RES1 showed an excess of heterozygosity, RES2 and RES3 showed a deficit of heterozygosity which might be caused by sampling mixed groups of sibling derived from genetically differentiated parental pairs, as have been also reported in a well described metapopulation of *Arvicola terrestris* (Stewart *et al.*, 1999). High levels of genetic structure are commonly reported in other small mammals within similar geographic scales (Brouat *et al.*, 2007; Ehrich & Stenseth, 2001; Gava & de Freitas, 2004). However, when  $F_{ST}$  values are compared between both species of *Arvicola* and similar scales of study, the Eurasian water vole apparently reach lower levels of genetic structure than SWV (Berthier *et al.*, 2006; Berthier *et al.*, 2005; Stewart *et al.*, 1999) (but see (Stewart *et al.*, 1998).

### **Contemporary gene flow in Doñana**

After natal dispersal, SWV establish their home ranges almost exclusively around small vegetation patches along the border of water bodies. At the beginning of the breeding season, reproductive individuals are forced to disperse in order to avoid competition (resource or locale mate competition hypotheses) or inbreeding between young non-dispersed individuals (see e.g. (Lidicker Jr & Stenseth, 1992) for a review). Whereas mitochondrial genetic diversity distribution suggested high migration rates between the Northernmost area of our survey (ABA1, ABA2 and ROC), Abalario1 is apparently a source population of dispersers toward Abalario2 and Rocina (table 3). These high migration rates could happen as result of short

distances between these populations and the presence of small tributaries (drainages) from Rocina which irrigate the whole northern region of Doñana. The highest levels of gene flow connected from Marismillas to Reserva2, two distanced areas (besides being Reserva1 between them) and apparently unconnected by the presence of sand dunes, *a priori* putative barriers for dispersal of SWV. However, shared haplotype 5 (table 2a) was again the most frequent within both RES2 and MAR, which also support this high connectivity between both regions.

### **Effects of landscape on genetic structure**

At the beginning of the present study, we predicted an important role of landscape features and the spatial layout of populations on the genetic structure of *A. sapidus* in Doñana (which would indirectly reflect gene flow rates between populations). Initially, the geographic distance among populations of SWV (Mantel test,  $r=0.5844$ ) explained the genetic structure (all  $p < 0.01$ ). However, although ponds, drainages and vegetation cover apparently affected genetic structure in SWV ( $r$  in Mantel tests; 0.6078, 0.6089 and 0.5386, respectively; all  $p < 0.05$ ), none of the landscape models considering these variables had a significant effect on genetic structure once the effects of the geographic Euclidean distance were discounted (all partial Mantel Test  $p > 0.05$ ; all  $\Delta AICc < 2$ ). At this scale of study, the spatial distribution of sampling areas would suggest random dispersal maximizing the connectivity among populations (spatial analyses, all partial Mantel Tests  $p > 0.05$ ).

It is worth pointing out here two relevant aspects when research is focused on dispersal in a spatially explicit context. First, there is always an anthropocentric way of understanding patterns and processes in wildlife that could bias our

knowledge on dispersal biology. Because rodents rely their spatial orientation on olfaction rather than on vision, their landscape perception could be better described as a network of olfactory paths (Kozakiewicz & Szacki, 1995) and a complete understanding of dispersal would never be feasible unless such kind of variables were included in the analysis. Second, it must be considered the relevant scale to embrace all landscape movements of the species to get a better knowledge about dispersal. Whereas the former is assumed as a common caveat in (almost) all dispersal studies, we are confident in genetic approaches to solve the latter, since widespread samples covering the whole study area and estimates of gene flow ensure the righteous scale of this study.

A plentiful literature distinguish habitat (patches) from non-habitat (matrix) when a classification of land uses for species is intended. This binary division assume the permeability of patches and the reluctance of individuals to cross matrices (barriers). There are, however, other species where habitat preferences might vary according to the reproductive status of individuals. These species apparently differentiate breeding from dispersal habitats and these two from matrix, being the latter considered as a barrier for individual movements. Fieldwork with capture-mark-recapture analyses have shown high population dynamics and interpatch movements of individuals (Román, 2007). For SWV, direct observation of individuals moving among patches becomes virtually impossible, mainly because of the unpredictability of these movements. In this study, least cost distance analyses together with a genetic approach has thrown light on dispersal habitat preferences of SWV in Doñana. On their dispersal pathways, individuals of *A. sapidus* do not apparently look for specific landscape attributes (such as ponds, drainages and vegetation cover) that are maximized for the establishment of breeding habitats.

Habitat is hence minimally classified as optimal for breeding (< 2 %, see *study system*), mostly considered as potential areas for dispersal and scarcely avoided (matrix) for individual movements. Is it an evolutionary stable strategy of habitat specialist species to ensure habitat colonization when patches are isolated and (apparently) beyond their dispersal availabilities? Several habitat specialist species embedded in networks of scarce and isolated patches show similar dispersal behaviours. (Figuerola & Green, 2002). Habitat specialist species are, therefore, apparently acquiring both generalist dispersal behaviours and/or specific long-distance dispersal strategies. Small mammals in a patchy but stable environment have shown the same opportunities to disperse as in continuous habitats (Cockburn, 1992), as a strategy to increase access to breeding pairs and new resources and decrease intrasexual competition (Lambin *et al.*, 2001). There are also evidences in other small mammals where adults dispersal and colonization of new patches in variable environments increase offspring survival (Cockburn *et al.*, 1981), even in suboptimal patches (Gliwicz, 1990). This consequence of high dispersal capacities has been also described in *A.terrestris* (Jannett, 1980; Stenseth, 1978) gaining importance when population densities decrease (Lambin, 1994). SWV at Doñana are enforced to survive at temporally hostile habitats because of recurrent drought seasons that create temporal variation in genetic diversity and population size. (Berthier *et al.*, 2005) studied the genetic consequences on a cyclic population of *A.terrestris* in France and they observed no relationship between voles abundance and genetic diversity. As in their study, if *A.sapidus* would lose genetic diversity as a consequence of draught periods, this would be rapidly recovered in Doñana because these periods of population decrease would be short enough (one-two months per year) and coincide with stop in reproduction (summer). Furthermore, it



would be followed by a fast genetic rescue of populations, since breeding period immediately starts after first rains and drift effects may be weak.

## **Dispersal**

Likely one of the most surprising results presented in this study is a non-significant sex-biased dispersal in SWV in Doñana. The genetic structure inferred with mitochondrial CR (maternally inherited marker, corrected  $F_{ST} = 0.1272$ ) exceeded that found with microsatellite (bi-parentally inherited) markers ( $F_{ST} = 0.072$ ). This difference between F-statistics using different inherited molecular markers would indicate sex-biased dispersal. However, other (and more powerful) sex-biased dispersal analyses presented here do not reject differences on dispersal rates. This may result from small sample sizes per population as a consequence of technical caveats during sequencing reactions (Centeno-Cuadros and Godoy, chapter 1). On the other hand, comparisons of sex-related  $F_{ST}$  using a randomisation procedure (all  $p > 0.05$ ) and comparisons of different slopes obtained in IBD tests using independent sex-specific datasets ( $b_m = 0.0243$ ,  $b_f = 0.0248$ ) do not reveal any signal of sex-biased dispersal in this arvicolid.

In the light of field work in Doñana it has previously suggested sex-biased dispersal (dispersal rates per sex: males, 14.4%; females, 3.5%) (Román, 2007). However, high mortality rates and extremely complex population dynamics might bias these results and must be considered cautiously and only as lower interval values of dispersal. The genetic approach performed here has helped to disentangle this prediction. High dispersal rates in males detected in fieldworks are not apparently followed by high breeding success of male individuals. Since males have been

proposed as the sex prone to disperse in SWV and these differences have been measured, our results suggest that these high dispersal rates are not followed by high breeding success in males. This difference between dispersal (field movements) and effective dispersal (gene flow) would be explained by high variance in reproductive success in one sex (males) or high mortalities in the dispersing sex (males).

The indirect estimates of dispersal using the slope of the regression of genetic vs. logarithm of geographic distances do not reveal either difference on dispersal distances between sexes ( $\bar{\lambda}_m = 668$  m,  $\bar{\lambda}_f = 661$ m). However these values should be taken with caution since their estimation is based on numerous assumptions (Rousset, 2000; Vekemans & Hardy, 2004) and also on an effective density value whose precision is still being solved. Despite both direct and indirect estimates of dispersal distance are known to be biased our estimates of average dispersal distance do not differ from the average estimates per sex using capture-mark-recapture analyses (males, 838m; females, 695m). Consequently, although long distance dispersal has been detected (Román, 2007) these movements might not suppose high proportion of dispersal movements (where the ratio direct-indirect estimates would increase as have previously shown as response to demographic dynamics (Rousset, 1997)

## BIBLIOGRAPHY

- Arnaud JF (2003) Metapopulation genetic structure and migration pathways in the land snail *Helix aspersa*: influence of landscape heterogeneity. *Landscape Ecology* **18**, 333-346.
- Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (1996) GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier II, Montpellier (France).
- Berthier K, Charbonnel N, Galan M, Chaval Y, Cosson J-F (2006) Migration and recovery of the genetic diversity during the increasing density phase in cyclic vole populations. *Molecular Ecology* **15**, 2665-2676.
- Berthier K, Galan M, Foltête JC, Charbonnel N, Cosson JF (2005) Genetic structure of the cyclic fossorial water vole (*Arvicola terrestris*): landscape and demographic influences. *Molecular Ecology* **14**, 2861-2871.
- Bravo MA, Montes C (1992) Inventario de las formaciones palustres del manto eólico del Parque Nacional de Doñana (SO España). In: *IV Spanish Workshop of Limnology*, Granada.
- Broquet T, Johnson CA, Petit E, *et al.* (2006) Dispersal and genetic structure in the American marten, *Martes americana*. *Molecular Ecology* **15**, 1689-1697.
- Brouat C, Chevallier H, Meusnier S, Noblecourt T, Rasplus J-Y (2004) Specialization and habitat: spatial and environmental effects on abundance and genetic diversity of forest generalist and specialist Carabus species. *Molecular Ecology* **13**, 1815-1826.
- Brouat C, Loiseau A, Kane M, Ba K, Duplantier JM (2007) Population genetic structure of two ecologically distinct multimammate rats: the commensal *Mastomys natalensis* and the wild *Mastomys erythroleucus* in southeastern Senegal. *Molecular Ecology* **16**, 2985-2997.
- Burnham K, Anderson D (2002) *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*, 2nd edn. edn. Springer, New York.
- Chesser RK (1991a) Gene diversity and female philopatry. *Genetics* **127**, 437-447.
- Chesser RK (1991b) Influence of gene flow and breeding tactics on gene diversity within populations. *Genetics* **129**, 573-583.
- Cockburn A (1992) Habitat heterogeneity and dispersal: environmental and genetic patchiness. In: *Animal dispersal: small mammals as a model* (eds. Stenseth NC, Lidicker Jr. WZ). Champan and Hall, London.
- Cockburn A, Braithwaite RW, Lee AK (1981) The Response of the Heath Rat, *Pseudomys-Shortridgei*, to Pyric Succession - a Temporally Dynamic Life-History Strategy. *Journal of Animal Ecology* **50**, 649-666.

- Coulon A, Cosson JF, Angibault JM, *et al.* (2004) Landscape connectivity influences gene flow in a roe deer population inhabiting a fragmented landscape: an individual-based approach. *Molecular Ecology* **13**, 2841-2850.
- Crochet PA (2000) Genetic structure of avian populations - allozymes revisited. *Molecular Ecology* **9**, 1463-1469.
- Dessauer HC, Cole CJ, Hafner MS (1996) Collection and storage of tissues. In: *Molecular Systematics* (eds. Hillis DM, Moritz C, Mable BK), pp. 29-41. Sinauer Associates, Inc., Sunderland, Massachusetts (U.S.A.).
- Dupanloup I, Schneider S, Excoffier L (2002) A simulated annealing approach to define the genetic structure of populations. *Molecular Ecology* **11**, 2571-2581.
- Ehrich D, Stenseth NC (2001) Genetic structure of Siberian lemmings (*Lemmus sibiricus*) in a continuous habitat: large patches rather than isolation by distance. *Heredity* **86**, 716-730.
- Excoffier L, Laval G, Schneider S (2005) Arlequin (version 3.0): An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* **1**, 47-50.
- Fedriani JM, Delibes M, Ferreras P, Roman J (2002) Local and landscape habitat determinants of water vole distribution in a patchy Mediterranean environment. *Ecoscience* **9**, 12-19.
- Figuerola J, Green AJ (2002) Dispersal of aquatic organisms by waterbirds: a review of past research and priorities for future studies. *Freshwater Biology* **47**, 483-494.
- Gava A, de Freitas TRO (2004) Microsatellite analysis of a hybrid zone between chromosomally divergent populations of *Ctenomys minutus* from southern Brazil (Rodentia : Ctenomyidae). *Journal of Mammalogy* **85**, 1201-1206.
- Gliwicz J (1990) Habitat-dependent reproductive success in bank voles. In: *Social systems and population cycles in voles* (eds. Tamarin RH, *al e*). Birkhauser Verlag, Basel.
- Goudet J (2001) FSTAT, a program to estimate and test gene diversities and fixation indices, version 2.9.3.
- Goudet J, Perrin N, Waser P (2002) Tests for sex-biased dispersal using biparentally inherited genetic markers. *Molecular Ecology* **11**, 1103-1114.
- Greenwood PJ (1980) Mating Systems, Philopatry and Dispersal in Birds and Mammals. *Animal Behaviour* **28**, 1140-1162.
- Hansson L (1991) Dispersal and connectivity in metapopulations. *Biological Journal of the Linnean Society* **42**, 89-103.
- Hellberg ME (1994) Relationships Between Inferred Levels of Gene Flow and Geographic Distance in a Philopatric Coral, *Balanophyllia elegans*. *Evolution* **48**, 1829-1854.

- Hirota T, Hirohata T, Mashima H, Satoh T, Obara Y (2004) Population structure of the large Japanese field mouse, *Apodemus speciosus* (Rodentia: Muridae), in suburban landscape, based on mitochondrial D-loop sequences. *Molecular Ecology* **13**, 3275-3282.
- Hudson RR, Slatkin M, Maddison WP (1992) Estimation of Levels of Gene Flow from DNA-Sequence Data. *Genetics* **132**, 583-589.
- Jannett FJ (1980) Social dynamics of the montane vole, *Microtus montanus*, as a paradigm. *The Biologist* **62**, 3-19.
- Jenness J (2004) Weighted mean of points (weightmean.avx) extension for ArcView 3.x, v. 1.2c. Jenness Enterprises. Available at: [http://www.jennessent.com/arcview/weighted\\_mean.htm](http://www.jennessent.com/arcview/weighted_mean.htm).
- Jensen JL, Bohonak AJ, Kelley ST (2005) Isolation by distance, web service. *BMC Genetics* **6**, 13 v.13.05 <http://ibdws.sdsu.edu/>.
- Koenig WD, Van Vuren D, Hooge PN (1996) Detectability, philopatry and the distribution of dispersal distances in vertebrates. *Trends in Ecology and Evolution* **11**, 514-517.
- Koizumi I, Yamamoto S, Maekawa K (2006) Decomposed pairwise regression analysis of genetic and geographic distances reveals a metapopulation structure of stream-dwelling Dolly Varden charr. *Molecular Ecology* **15**, 3175-3189.
- Kozakiewicz M, Szacki J (1995) Movements of small mammals in a landscape: patch restriction or nomadism. In: *Landscape approaches in mammalian ecology and conservation* (ed. Lidicker Jr. WZ). University of Minnesota Press, Minneapolis.
- Lambin X (1994) Natal Philopatry, Competition for Resources, and Inbreeding Avoidance in Townsends Voles (*Microtus-Townsendii*). *Ecology* **75**, 224-235.
- Lambin X, Aars J, Piertney SB (2001) Dispersal, intraspecific competition, kin competition and kin facilitation: a review of the empirical evidence. In: *Dispersal* (eds. Clobert J, Danchin E, Dhondt AA, Nichols JD). Oxford University Press, New York.
- Lidicker Jr WZ, Stenseth NC (1992) To disperse or not to disperse: who does it and why? In: *Animal dispersal: small mammals as a model* (eds. Stenseth NC, Lidicker Jr. WZ). Chapman and Hall, London.
- Lidicker Jr. WZ (1995) *Landscape approaches in mammalian ecology and conservation* University of Minnesota Press, London.
- Nei M (1978) Estimation of Average Heterozygosity and Genetic Distance from a Small Number of Individuals. *Genetics* **89**, 583-590.
- Piertney SB, Stewart WA, Lambin X, *et al.* (2005) Phylogeographic structure and postglacial evolutionary history of water voles (*Arvicola terrestris*) in the United Kingdom. *Molecular Ecology* **14**, 1435-1444.

- Prugnolle F, de Meeus T (2002) Inferring sex-biased dispersal from population genetic tools: a review. *Heredity* **88**, 161-165.
- Pusey AE (1987) Sex-Biased Dispersal and Inbreeding Avoidance in Birds and Mammals. *Trends in Ecology & Evolution* **2**, 295-299.
- Ray N (2005) PATHMATRIX: a geographical information system tool to compute effective distances among samples. *Molecular Ecology Notes* **5**, 177-180.
- Raymond M, Rousset F (1995) GENEPOP version 1.2: population genetic software for exact test and ecumenism. *Journal of Heredity* **86**, 248-249.
- Román J (2007) *Historia natural de la rata de agua (Arvicola sapidus) en Doñana*, Universidad Autónoma de Madrid.
- Rousset F (1997) Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Molecular Ecology* **145**, 1219-1228.
- Rousset F (2000) Genetic differentiation between individuals. *Journal of Evolutionary Biology* **13**, 58-62.
- Rozas J, Sánchez-DelBarrio JC, Messeguer X, Rozas R (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Methods of biochemical analysis* **19**, 2496-2497.
- Schadt S, Knauer F, Kaczensky P, *et al.* (2002) Rule-based assessment of suitable habitat and patch connectivity for the eurasian lynx. *Ecological Applications* **12**, 1469-1483.
- Schuelke M (2000) An economic method for the fluorescent labeling of PCR fragments. *Nature Biotechnology* **18**, 233-234.
- Scribner KT, Chesser RK (1993) Environmental and demographic correlates of spatial and seasonal genetic structure in the eastern cottontail (*Sylvilagus floridanus*). *Journal of Mammalogy* **74**, 1026-1045.
- Slatkin M (1985) Gene Flow in Natural Populations. *Annual Review of Ecology and Systematics* **16**, 393-430.
- Stacy JE, Jorde PE, Steen H, *et al.* (1997) Lack of concordance between mtDNA gene flow and population density fluctuations in the bank vole. *Molecular Ecology* **6**, 751-759.
- Stenseth NC (1978) Demographic strategies in fluctuating populations of small rodents. *Oecologia* **33**, 149-172.
- Stewart WA, Dallas JF, Piortney SB, *et al.* (1999) Metapopulation genetic structure in the water vole, *Arvicola terrestris*, in NE Scotland. *Biological Journal of the Linnean Society* **68**, 159-171.

- Stewart WA, Piertney SB, Dallas JF (1998) Isolation and characterization of highly polymorphic microsatellites in the water vole, *Arvicola terrestris*. *Molecular Ecology* **7**, 1258-1259.
- Van De Zande L, Van Apeldoorn RC, Blijdenstein AF, *et al.* (2000) Microsatellite analysis of population structure and genetic differentiation within and between populations of the root vole, *Microtus oeconomus* in the Netherlands. *Molecular Ecology* **9**, 1651-1656.
- Vekemans X, Hardy OJ (2004) New insights from fine-scale spatial genetic structure analyses in plant populations. *Molecular Ecology* **13**, 921-935.
- Ventura J (1988) *Contribución al conocimiento del género Arvicola (Lacepede, 1799) en el Nordeste de la Península Ibérica*, Universidad de Barcelona.
- Verbeylen G, De Bruyn L, Adriaensen F, Matthysen E (2003) Does matrix resistance influence Red squirrel (*Sciurus vulgaris* L. 1758) distribution in an urban landscape? *Landscape Ecology* **18**, 791-805.
- Vignieri SN (2005) Streams over mountains: influence of riparian connectivity on gene flow in the Pacific jumping mouse (*Zapus trinotatus*). *Molecular Ecology* **14**, 1925-1937.
- Wilson GA, Rannala B (2003) Bayesian inference of recent migration rates using multilocus genotypes. *Genetics* **163**, 1177-1191.

# **DISCUSIÓN GENERAL**





En la introducción general de esta tesis doctoral nos marcábamos dos objetivos principales. El primero era intentar probar la relevancia del estudio de la diversidad genética en tres niveles jerárquicos diferentes (individuo, población y especie) para entender qué factores han influido en la distribución espacial de dicha diversidad. El segundo objetivo se desprendía del primero como caso práctico, y planteaba profundizar en el conocimiento de la historia natural y evolutiva de la rata de agua mediante el empleo de marcadores moleculares. En los distintos capítulos de esta tesis hemos puesto de manifiesto cómo el diseño y evaluación de marcadores moleculares con distinto nivel de polimorfismo y tipo de herencia y específicos para la rata de agua (**capítulos 1 y 4**) ha proporcionado las herramientas necesarias para caracterizar a la especie en los tres niveles jerárquicos mencionados. Con esta información hemos podido cuantificar el tiempo transcurrido desde que las poblaciones del antecesor común a *Arvicola sapidus* y *A. terrestris* (verosíblemente *A. cantiana*) se fragmentaran y diferenciaron en aislamiento hasta originar las dos especies del género en Eurasia (**capítulo 2**), así como describir el reparto de la diversidad genética de la rata de agua meridional a lo largo de toda su área de distribución, y los procesos que han dado lugar a este reparto (**capítulo 3**). Finalmente, hemos descrito la estructura genética poblacional mediante el genotipado de individuos distribuidos en un ambiente heterogéneo. Incorporando estos datos a los sistemas de información geográfica disponibles para el área, hemos evaluado los efectos de la estructura del paisaje sobre los patrones de flujo génico contemporáneo, relacionándolos finalmente con la biología de la dispersión de la especie (**capítulo 4**).

Todos los datos moleculares analizados en esta tesis doctoral han sido obtenidos a partir del diseño de cebadores específicos de marcadores moleculares para la rata de agua. Mediante el genotipado de los individuos con los diez marcadores microsatélites

(**capítulo 4**), hemos estudiado los efectos del paisaje sobre la dispersión de los individuos entre las áreas muestreadas en Doñana, así como descrito la estructura genética poblacional de la especie en la región y sugerido las causas que han podido originarla. Los marcadores mitocondriales diseñados en el **capítulo 1** han servido de eslabón entre los distintos niveles jerárquicos y escalas temporales. Así, por ejemplo, la secuenciación de la región control completa nos ha permitido detectar en el **capítulo 4** diferencias entre sexos en las tasas de dispersión en las poblaciones de Doñana. Una vez conocida la distribución de la diversidad nucleotídica en toda la región control y citocromo b, hemos diseñado cebadores específicos para cada marcador con objeto de albergar el mayor polimorfismo en el fragmento más corto posible. De este modo, hemos evaluado los patrones de distribución de la diversidad genética de la especie a lo largo de todo su área de distribución (**capítulo 3**) y datado el evento de aislamiento y divergencia del antecesor común de la rata de agua (**capítulo 2**). Probablemente uno de los aspectos más relevantes de esta tesis doctoral para la Genética de la Conservación es el diseño de marcadores moleculares polimórficos y su amplificación en material obtenido de muestreos no invasivos. Para obtener la información que se analiza y discute en los **capítulos 2 y 3**, hemos extraído el material genético de huesos obtenidos de egagrópilas de rapaces, un método propuesto desde hace tiempo tanto para genética de poblaciones como para filogeografía, pero poco usado en la práctica. Los muestreos no invasivos se han usado frecuentemente en genética de poblaciones y para la monitorización de poblaciones, mediante la amplificación y genotipado de marcadores nucleares microsatélites. Paradójicamente, casi no se emplean los muestreos no invasivos en filogeografía, a pesar de que (como hemos demostrado en esta tesis doctoral) facilitan enormemente la extensión del muestreo a todo el área de distribución de las especies. Tras revisar los trabajos publicados hasta la fecha, hemos visto que el

nuestro es uno de los primeros artículos sobre filogenias y filogeografía donde el material genético se extrae principalmente de huesos obtenidos de egagrópilas de rapaces y, en una menor proporción, de pieles de ejemplares conservados en colecciones científicas. Como resultado de nuestro trabajo, es posible apostar por el uso de métodos no invasivos en filogeografía como alternativa a aquellos casos en los que no hay muestras de buena calidad disponibles; esa alternativa se aplica no sólo a la extracción, sino también a amplificación y secuenciación de secuencias polimórficas de ADN mitocondrial. Por extensión, confiamos en que la aplicación a muestras no invasivas del conjunto de microsatélites propuestos en el **capítulo 4** pueda suponer un gran salto adelante en la monitorización de las poblaciones de rata de agua (e.g. series de ocupación de lagunas en función del genotipado de excrementos frescos).

La primera aplicación de las técnicas genético-moleculares en esta tesis doctoral se orienta a la estima de las fechas de divergencia del género *Arvicola* (**capítulo 2**). Hemos empleado las secuencias de los marcadores mitocondriales (región control y citocromo b) obtenidas de *A. sapidus* y *A. terrestris* para comparar los más avanzados y recientes modelos estadísticos de aislamiento con migración basados en coalescencia con las estimas clásicas filogenéticas de los tiempos de divergencia. Precisamente una de las grandes ventajas de los nuevos métodos para la estima de los tiempos de divergencia entre especies radica en la aplicación de las cadenas de Markov Monte-Carlo (MCMC) y la teoría de la coalescencia a la estima de parámetros demográficos. Los resultados incluyen la estimación de los tamaños poblacionales efectivos de *A. sapidus* y *A. terrestris*, el descarte de la potencial hibridación entre ellas y la datación, partiendo de sus diversidades genéticas actuales, del tiempo de divergencia de su antecesor común más reciente. Gracias a la presencia de las especies de *Arvicola* en los registros fósiles y a nuestra estima de los tiempos de divergencia hemos podido contribuir a la

reconstrucción de la historia evolutiva del género (adelantando hacia finales del Pleistoceno Medio la divergencia genética de las dos especies) y a la evaluación del estatus de *A. sapidus* en su área de distribución (realizando la primera estima del tamaño poblacional de la rata de agua). Para aquellas especies que no se encuentran habitualmente en los registros fósiles, trabajos como el nuestro plantean una destacada ventaja: la posibilidad de reconstruir la historia evolutiva sin disponer muestras fósiles y tan sólo a partir de unas cuantas muestras contemporáneas.

Como vimos en la introducción general, los planes de conservación deben incluir entre sus objetivos prioritarios la conservación de la diversidad genética de los taxones considerados. Sin embargo, la incorporación de las herramientas genéticas en la Biología de la Conservación ha puesto en tela de juicio la definición de las entidades que requieren prioridad para su conservación. La delimitación taxonómica de los dos arvicólidos euroasiáticos queda corroborada por la monofilia recíproca y la ausencia de hibridación (**capítulo 2**). En el caso de *A. terrestris* se han delimitado además categorías taxonómicas subespecíficas gracias a criterios morfológicos y genéticos. Por el contrario, a pesar de las controvertidas propuestas de rangos subespecíficos de *A. sapidus*, atendiendo a variaciones morfológicas, no había hasta la fecha ningún estudio genético que tratara de contrastar dichas variaciones con divergencias genéticas profundas a nivel intraespecífico. En el **capítulo 3** no hemos encontrado relación entre la variabilidad genética y los patrones de coloración del pelaje o tamaños corporales, de forma que nuestros resultados no apoyan la controvertida delimitación entre las subespecies *A. sapidus sapidus* y *A. sapidus tenebricus*. En ese mismo capítulo hemos estudiado el reparto de la diversidad genética de la especie a lo largo de todo su área de distribución. Atendiendo a nuestros resultados, los ciclos glaciares explican el origen y variación genética de la especie en Iberia y Francia. La rata de agua meridional surgió

en Iberia en torno a la glaciación de Mindel (ca. 250.000 años) y traspasó los Pirineos para llegar a Francia, estableciendo el límite más septentrional de su área de distribución en el norte del país (**capítulos 2 y 3**). La filogeografía de la rata de agua ha supuesto un avance en el estudio del efecto de las glaciaciones sobre la fauna y flora ibéricas. Nuestro trabajo confirma el efecto de la heterogeneidad fisiográfica de la península ibérica sobre la diversidad genética de esta especie, sugiriendo el papel determinante que han jugado los refugios situados dentro del refugio glacial ibérico. Bajo nuestro punto de vista, estos resultados aportan un ejemplo que faltaba en la literatura relacionada con los patrones de refugio dentro de refugio: a diferencia del resto de las especies estudiadas, la rata de agua presenta una dinámica poblacional y una capacidad dispersiva más complejas que las especies clásicamente estudiadas. *Arvicola sapidus* representa, por tanto, un nuevo sistema de estudio, que se suma a los ejemplos que avalan la influencia de los refugios glaciales situados dentro de Iberia en el reparto de la diversidad genética de las especies. En este caso, una aparente elevada capacidad de recolonización de zonas libre de hielos en los periodos interglaciales parece haber evitado el aislamiento y divergencia profunda de las poblaciones localizadas durante largos periodos de tiempo dentro de estos refugios ibéricos. Por todo ello, no hay motivos para establecer prioridades de conservación entre las poblaciones residentes en los refugios glaciales descritos y proponemos la gestión de todas las poblaciones como una única unidad de manejo para la especie.

El genotipado de individuos de rata de agua de la región de Doñana nos ha permitido resolver algunas dudas respecto a la historia natural y biología de la dispersión de la especie (**capítulo 4**). La rata de agua es una especie esquiva, de hábitos elusivos y difícil de estudiar con radioemisores. En contra de nuestras hipótesis iniciales, no hemos encontrado indicios de que el paisaje ejerza un papel relevante en la distribución de la

estructura genética de las poblaciones en Doñana. Estos resultados contrastan con nuestra percepción de la rata de agua como un especialista de hábitat que restringe sus movimientos a las zonas óptimas como hábitat de cría. Gracias a los análisis genéticos hemos podido abarcar todo el área de estudio (Doñana), a diferencia de los anteriores análisis de captura-marcaje-recaptura que se restringían a las parcelas de las poblaciones (3 kms). No obstante, las estimas de distancias media de dispersión obtenidas con los marcadores moleculares coinciden con las distancias medias de dispersión estimadas mediante los métodos de trampeo. Estos resultados sugerirían que los posibles sesgos debido al sexo de los individuos en las estimas de las tasas de dispersión mediante captura-recaptura y no detectados con la genética podrían quedar explicados por una elevada tasa de mortalidad de los machos dispersantes y/o elevada variabilidad en el éxito reproductor de los individuos migrantes.

### **Del individuo a la especie: importancia para la Genética de la Conservación**

Al alcanzar esta última sección de la tesis doctoral, esperamos haber dejado claro al lector que la distribución de la diversidad genética de una especie es el resultado de los distintos procesos que afectan a las poblaciones y de cómo los individuos pueden servir de nexo de unión entre éstas, condicionando el patrón de distribución de la diversidad genética global de la especie. En este apartado intentaremos interpretar conjuntamente los resultados obtenidos en los tres niveles jerárquicos analizados, para poder extraer conclusiones útiles para la elaboración de planes de conservación de la especie.

Gracias al estudio sobre filogeografía de la especie en todo su área de distribución (**capítulo 3**) hemos sabido que el reparto de la diversidad genética coincide con la localización de refugios previamente descritos dentro de Iberia. En este capítulo

sugerimos el tratamiento de toda la diversidad albergada en los distintos refugios como una única unidad evolutivamente significativa, debido a la débil estructura filogenética de la especie y a que no encontramos divergencias intraespecíficas profundas que pudieran corresponder con las subespecies anteriormente propuestas o con otras nuevas. Esta débil estructura filogeográfica podría deberse a la recurrencia de altas tasas de flujo génico entre refugios durante los periodos interglaciares. Sin embargo, al estudiar el tiempo de divergencia (**capítulo 2**) concluimos que la rata de agua es una especie de origen relativamente reciente (finales del Pleistoceno Medio). La débil estructura filogenética podría deberse tanto a los periodos de intercambio genético en los interglaciares como a que no ha transcurrido aún mucho tiempo (sólo dos glaciaciones) desde la diferenciación de las especies a partir de su antecesor común (especiación) y, por tanto, no se han originado divergencias intraespecíficas profundas. En promedio, la similitud genética entre dos individuos desaparece a partir de los 500 kms de distancia, lo cual podría ser explicado por altas tasas de flujo génico entre las poblaciones y contribuiría a explicar la débil estructura filogenética registrada para la especie. El flujo génico entre poblaciones se puede relacionar con la dispersión de los individuos entre núcleos poblacionales (movimiento de un individuo seguido de su reproducción). En el **capítulo 4** sugerimos que las altas tasas de flujo génico encontradas en un ambiente heterogéneo pueden deberse al escaso efecto que ejerce el paisaje sobre los movimientos de los individuos. A pesar de esta relativa permeabilidad de la matriz al movimiento dispersivo, las distancias medias de dispersión estimadas con métodos de captura-marcaje-recaptura y empleando marcadores genéticos no superan el kilómetro de distancia (**capítulo 4**). Esta distancia media de dispersión no es mayor a la observada para otras especies de micromamíferos, pero la relativa abundancia y amplia distribución por casi toda la península ibérica y Francia, unas dinámicas poblacionales



complejas, un ciclo de vida corto y generaciones que se solapan, podrían explicar el éxito de recolonización de la especie durante los interglaciares.

En esta tesis doctoral hemos intentado demostrar la necesidad de entender los distintos procesos que actúan en los tres niveles jerárquicos estudiados para entender los patrones de distribución de la diversidad genética de una especie. Gracias a este y a otros trabajos publicados sobre la historia natural y ecología de la rata de agua, disponemos ya de la base de conocimiento necesaria para elaborar planes de conservación que aseguren la viabilidad de sus poblaciones.

### **Nuevas preguntas**

Este trabajo resuelve algunas dudas sobre la biología e historia evolutiva de la rata de agua e igualmente plantea nuevos interrogantes y futuras líneas de investigación, tanto a nivel de especie como de género. El éxito de los nuevos modelos de aislamiento con migración basados en coalescencia en las estimas de divergencia de las dos especies del género, y para datar la fecha de colonización de *A. sapidus*, deja la puerta abierta a la aplicación de los mismos métodos con la especie euroasiática del género. La subdivisión de *A. terrestris* en distintas subespecies atendiendo a criterios morfológicos y genéticos plantea un escenario similar aunque de mayor complejidad. Como hemos demostrado, una metodología como la empleada para *A. sapidus* (marcadores moleculares y muestreos no invasivos) ayudaría a responder cuestiones relacionadas con *A. terrestris*, especie con una extensa área de distribución (e.g. ¿por qué la subespecie cavadora se encuentra en áreas discontinuas donde coincide con la rata de agua? ¿es el resultado de una fragmentación de una población ancestral continua?)

Quedan pendientes nuevos análisis de aislamiento con migración entre pares de linajes dentro de los refugios, para poder estimar el tiempo de divergencia de cada uno de ellos. Para estos análisis serían necesarias las aproximaciones multilocus (e.g. secuencias ADN mitocondrial y nuclear, microsatélites...) con marcadores con distintas tasas mutacionales. A nivel de población, los muestreos no invasivos abren nuevas puertas para estudiar dinámicas poblacionales, patrones de ocupación de las colonias...sin necesidad de trampear a los individuos. El genotipado de individuos permitirá describir la organización social de las colonias, el sistema de reproducción mediante análisis de paternidades, realizar análisis de autocorrelación espacial...La combinación del genotipado de individuos junto a las capturas-recapturas realizadas permitirá corregir los posibles sesgos de las estimas de distancias y tasas de dispersión debidos a la propia metodología de cada método. Pero, como escribía Michael Ende en *La Historia Interminable*, “eso es otra historia y deberá ser contado en otra ocasión”



# **CONCLUSIONES GENERALES**



- 1.** El 74% de la diversidad nucleotídica albergada en la Región Control mitocondrial se concentra en su primer dominio hipervariable. El segundo dominio hipervariable apenas presenta variabilidad nucleotídica y presenta monopolímeros que dificultan la amplificación y secuenciación, restando valor a este fragmento como marcador mitocondrial.
- 2.** En este trabajo hemos diseñado cebadores específicos para la rata de agua para amplificar los fragmentos mitocondriales más polimórficos y de menor tamaño de la Región Control y citocromo b, caracterizando en un fragmento de 413 bp gran parte del polimorfismo albergado en ambos fragmentos completos.
- 3.** Los huesos obtenidos de egagrópilas de rapaces suponen una fuente de ADN de gran utilidad en estudios de Filogeografía, por facilitar enormemente los muestreos de pequeños roedores. Sin embargo, su utilidad puede ser limitada para la caracterización de ADN nuclear de copia sencilla.
- 4.** Los diez marcadores microsátélites propuestos para la rata de agua albergan el suficiente polimorfismo para realizar análisis genéticos desde el nivel de individuo al de especie y son, por tanto, de gran utilidad para describir patrones genéticos y conocer los procesos que los originaron a distintos niveles jerárquicos y escalas temporales.
- 5.** La divergencia de las poblaciones del antecesor común a la rata de agua tuvo lugar a finales del Pleistoceno Medio (*ca.*250.000 años), como consecuencia del aislamiento en la Península Ibérica durante la glaciación Mindel.
- 6.** A pesar de que las dos especies de *Arvicola* habitan en simpatria en Francia y el Norte de la Península Ibérica, no hemos encontrado evidencias de flujo génico o hibridación entre ellas

7. La distribución de la diversidad genética mitocondrial de la rata de agua se reparte entre siete áreas geográficas, seis de las cuales coinciden aproximadamente con refugios previamente descritos dentro de la Península Ibérica. La rata de agua se suma a los ejemplos que avalan la influencia de los refugios glaciales situados dentro de Iberia en el reparto de la diversidad genética de las especies.
8. La colonización de Francia se realizó hace unos 62.000 años por las poblaciones del Sur de los Pirineos. Toda la diversidad genética localizada en Francia deriva de un único haplotipo fundador. Se descarta, por tanto, el papel de los Pirineos como barrera a la expansión de las poblaciones de rata de agua hacia zonas libres de hielos en los periodos interglaciares.
9. Un tiempo de divergencia relativamente corto (*ca.* 250.000 años) y una gran capacidad de colonización de la especie pueden haber eliminado cualquier divergencia profunda originada durante los periodos de aislamiento y divergencia en los refugios glaciares y explicar la débil estructura filogenética encontrada.
10. Nuestros resultados no apoyan la controvertida delimitación taxonómica entre las subespecies *A. sapidus sapidus* y *A. sapidus tenebricus*, no habiéndose encontrado relación entre la variación genética y los patrones de coloración del pelaje o tamaños corporales.
11. Para maximizar la conservación de la totalidad de la diversidad genética deberá garantizarse en los planes de conservación y manejo de la especie la protección de cada uno de los refugios descritos. Sin embargo, dado que estas divisiones no han surgido de un aislamiento prolongado, no han originado divergencias adaptativas que diferencie a unos individuos de otros y todas parecen estar conectadas por flujo génico, toda la especie debería ser considerada como una única unidad evolutivamente significativa.

- 12.** Las colonias de rata de agua se encuentran estructuradas genéticamente en la región de Doñana y responden a un patrón de aislamiento por distancia.
- 13.** Las distancias medias de dispersión inferidas con métodos moleculares coinciden con las estimadas mediante métodos de captura-marcaje-recaptura (*ca.* 700m.). A diferencia de los métodos directos, no encontramos evidencias de sesgos de dispersión relacionados con el sexo de los individuos. La comparación de ambos resultados sugieren una elevada tasa de mortalidad y variabilidad en el éxito reproductor de los dispersantes.
- 14.** En contra de lo esperado para una especie descrita como especialista de hábitat, no hemos encontrado evidencias de que el paisaje ejerza una fuerte influencia sobre la estructura genética de las poblaciones de rata de agua. Este comportamiento dispersivo podría tratarse de una estrategia evolutivamente estable de especialistas de hábitat, cuya supervivencia y persistencia de las poblaciones en el espacio y tiempo dependerá de una gran capacidad de colonización de parches de hábitat generalmente separados por matrices hostiles para el asentamiento de colonias.
- 15.** Los procesos históricos y contemporáneos que han determinado la distribución actual de la variación genética de la rata de agua han sido investigados mediante el estudio de dicha variación en tres niveles jerárquicos diferentes (individuo, población y especie). Esta información será clave para el diseño de planes de gestión y conservación que garanticen la conservación de la especie y de su variación genética.





# **AGRADECIMIENTOS**



Lo último que uno sabe es por dónde empezar. Con toda la razón del mundo lo dijo Blaise Pascal, así que he decidido aplicar una norma simple: comenzaré por el principio. Recuerdo mi primera conversación con Rafael Villafuerte cuando llegué a Ciudad Real en septiembre de 2003. Rafa relacionó la vida del investigador en formación durante su tesis doctoral con una manguera retorcida sobre sí misma. Comienzas la tesis doctoral (entras en la manguera) más o menos seguro de cómo irá evolucionando y decidido en tus pasos. Con el tiempo avanzas por la manguera y comienzas a perder de vista la luz de la entrada de la manguera por la que iniciaste el camino. No sabes muy bien dónde estás ni a qué altura de la manguera te encuentras. Pero sigues caminando. Llega un momento en el que te planteas realmente por qué te has metido ahí dentro y si eso llegará a buen puerto, puesto que estás perdiendo mucho tiempo sin encontrar resultados e incluso afectando a tu vida personal. Pero sigues caminando porque sabes que ya dejarlo todo o volver hacia atrás sería peor aún (algo así como la falacia del Concord...) De repente, ves una luz al extremo opuesto por el que habías entrado y continúas tu travesía dirigiéndote a ella...hasta que logras salir con éxito de la manguera y terminas tu tesis doctoral. Siempre he tenido presentes estas palabras en el transcurso de mi tesis doctoral. En cada una de las fases siempre he tenido la esperanza puesta en ver ese pequeño haz de luz en algún momento u otro. Y es en este momento en el que escribo estas líneas de agradecimientos en el que me doy cuenta de que realmente Rafa tenía razón...y no habría podido continuar adelante por esta “manguera” de la tesis doctoral de no ser por la gente a la que a continuación haré referencia. Serán todos los que están, pero no estarán todos los que son. De antemano, disculpas si alguno de los lectores que ahora fijáis vuestra atención en estas líneas no veis vuestros nombres escritos.

Primero debo agradecer a quienes no conozco personalmente pero han ejercido una gran influencia para la logística de esta tesis doctoral. Este trabajo ha sido financiado por la Dirección General de Investigación (proyecto BOS2001-2391-C02-01) y una beca (2+2) de Formación del Profesorado Universitario del Ministerio de Educación y Ciencia. La coordinación y administración de permisos ha corrido a cargo de la Oficina de Coordinación de Proyectos de la Estación Biológica de Doñana. La obtención de muestras no hubiera sido posible de no ser por la inestimable ayuda de material colectado en el campo por F. Alvarez, R. Balado, J. Barrul, A. Ceña, J.C. Ceña, M.M. Delgado, A. Delibes, C. Escala, N. Indelicato, C. Lemarchand, R. Lorenço, J. Mate, A. Monzón, A. Morcillo, J.F. Noblet, R. Pita, D. Serrano, F. Simonnet, J.V. Turón, J.

Ventura y los encargados de las colecciones científicas del Museo de Ciencias Naturales de Barcelona, la sociedad de ciencias Aranzadi y la Estación Biológica de Doñana. Por último quisiera agradecer igualmente a Juan Jiménez, el grupo Nutria de la Sociedad Española para la Conservación y Estudio de los Mamíferos (SECEM) y la Conselleria de Territori i Habitatge (CTH, Generalitat Valenciana) por su ayuda para los trámites que nos favorecieran la captura de ratas de agua en Castellón y al Centro de Supercomputación de Galicia (CESGA) por la “logística computacional” tan competente y pionera prestada para gran parte de los análisis que hemos realizado.

Algo que siempre he tenido siempre claro era por quién iba a comenzar a escribir esta sección de la tesis. No sé si existirá o no, pero el destino me propuso un día la oportunidad de venir a hacer esta tesis doctoral hacia el otoño de 2003 (qué peligro, ya comienzo a hablar de batallitas de guerra de cuando era joven y lozano...). Recuerdo perfectamente mi primera conversación telefónica con Miguel Delibes. Estaba muy nervioso en la terraza de mi casa en Jaén y Miguel estaba dando un paseo por la playa en la Punta, resumiéndome brevemente mi proyecto de tesis doctoral...y creo que este documento es una prueba más que suficiente para demostrar que me convenció. Miguel me tendió la mano para hacer la tesis doctoral en la Estación Biológica de Doñana y ha sido el mejor San Pedro que me podía abrir las puertas de este mundo de la ciencia. Basta una conversación para conocer a Miguel, para saber que es alguien que se desvive por los suyos y por su trabajo, un naturalista como los que apenas quedan. Estoy convencido de que si no se le hubiera adelantado un señor inglés con barbas y pelo canosos, él hubiera acabado proponiendo la teoría de la selección natural. He tenido la suerte de tener a Miguel siempre ahí, e incluso cuando no estaba, seguía los amables consejos de Los Panchos y si yo le decía ven, él lo dejaba todo por venir y echarme una mano. Darte las gracias es lo que corresponde en esta sección, pero espero poder tener por delante aún muchos años más para poder dártelas en persona.

También recuerdo con claridad la primera vez que vi a José Antonio Godoy. Me acuerdo de su despacho en la esquinita de la última planta del Pabellón del Perú, de cuando bajamos al laboratorio del sótano por esas laberínticas escaleras del torreón...Godoy tiene la virtud de ser un gran docente...y paciente con un testarudo y cabezota como yo. A él le debo tanto que creo que tendría que añadir un segundo volumen o un apéndice a los agradecimientos para enumerarlo todo. Godoy ha sido científico, docente, psicólogo y amigo y me ha enseñado todo cuanto sé hoy día de

Ecología Molecular. Me recibió con los brazos abiertos en su laboratorio a pesar de que hasta ese momento pensaba que un microsatélite era un aparatito chiquitito que circulaba alrededor de la tierra. Compañero y amigo de viajes (¡aunque aún tenemos que convencerte para que nos pagues esa “retirada espiritual” a Marruecos!) espero que sigamos compartiendo vagón en este tren por mucho tiempo...

En Enero de 2004 me planté en Sevilla y me fui al Doñana a “ratear” con Miguel, Godoy y otra de las personas que más admiro de la Estación Biológica de Doñana como científico, biólogo y naturalista. Jacinto Román es una enciclopedia andante, un burgalés de pura cepa, conservacionista convencido y poeta urbano del siglo XXI como nadie. El trabajo de campo en Doñana y nuestra visita a Castellón nos unió un poquito más y me permitió conocer a quién pertenecía ese ADN que estaba dentro de un tubito de 1.5 mL. Mil gracias majo.

Para un recién licenciado como era cuando llegué a la EBD ha sido un honor poder compartir el día a día con un grupo de investigación como es el grupo Carnívoros. Me permitís que comience por una de las *alma mater* del grupo, Sofia Conradi. Sofi, se demostraría la existencia del caos si tú no estuvieras a cargo de toda la administración del grupo. Has sido mi “mamá doñanera” y compañera de despacho, confesora de gran parte de mis pecados, amiga con quien celebrar muchísimas alegrías y hombro sobre el que llorar alguna que otra pena. No sé dónde pararé a partir de ahora, pero recuerda siempre que cuando entre por la mañana a cualquier otro despacho, siempre echaré de menos no poder ser el primero en darte los buenos días. A Paco Palomares lo he ido conociendo poco a poco y cada día que pasa encuentra algo más con lo que pueda sorprenderte. Su pasión por la fotografía ¡y las salidas en bicicleta! han hecho que disfrutemos juntos de grandes momentos...¡espero vuelvan a repetirse muchos más!. Eloy Revilla emana conocimiento, alguien con quien siempre he podido entablar conversación de cualquier aspecto relacionado a mi tesis y cuya carrera profesional creo que es un ejemplo a seguir. Con Alejandro Rodríguez he tenido la suerte de poder intercambiar opiniones de cuanto hay ahora sobre el escenario de la ciencia como de lo que hay en el *backstage* de este mundo. Con Séverine Roques he tenido grandes discusiones de pasillo sobre análisis de datos y grandes pros y contras de la Ecología Molecular, y junto con Elia, hemos compartido de los grandes momentos este periodo. Y todos ellos han sabido elegir sin duda a la mejor infantería para el estudio de los carnívoros. Con Marian Pereira he tenido la suerte de poder compartir los cursos de

doctorado en la Universidad de Granada y un inolvidable viaje al bosque de Bielowieza (Polonia). José Vicente (algún día demostraremos la teoría de la rata surfera...¡no desesperes!), María Viota (y su peculiar manera de ver la vida que poco a poco he ido conociendo), Ana Martínez (la galleginha y más bondadosa del grupo) y Carolina Soto (la malagueña más risueña que jamás he conocido). Poquito fue el tiempo que estuvieron pero dieron un pequeño gran toque internacional al grupo: Tullia, Sonja, Joachim y Nuno Negroes. Tullia Ferreri vino, creció y maduró tanto académica como personalmente y pudimos compartir conciertos, música, juergas inesperadas...dio una gran pincelada de vida a una época en la que realmente la necesitaba. ¡Mil gracias Tutu! Sonja Desnica...jamás sabré cómo aprendiste tan rápido el español...y me hiciste aprender croata e inventar y aquel chiste que tanto te gustaba del brindis de los caracoles. Joachim...fotógrafo, comprometido con la conservación, aparentemente tímido y reservado pero nada más lejos de la realidad...el mejor cocinero de bacalao dorado que conozco...¡gracias!...y Nuno, don Nuno Miguel ¿qué decir de ti? Compañero y amigo (también de juergas inesperadas) compartimos tanto en estas fases finales de nuestras criaturitas...mil gracias por estar siempre ahí y dar siempre esa pincelada de honestidad y amistad. Néstor Fernández, sin tu humor...no sé qué habría sido de mí...gracias por apoyarme en esos momentos tan duros en los que nadie entendía mis chistes...salvo tú. Eres todo un colega. Apenas llegué a la EBD tú te fuiste, pero luego creo aprovechamos mucho el tiempo juntos desde que volviste de Leipzig. Tu racionalidad, capacidad crítica y analítica y cualidades docentes me han enseñado más de lo que imaginas. Gracias por todo hombre de hojalata y espantapájaros...uy no, que eso es de otra película, perdón...

Y ahora es el turno de algo que he esperado con impaciencia desde hace mucho tiempo: ¡Gracias Chunguitas!. Gracias Ana, Laura y Maritxu. La borde, la mala y la chungu. No dejan de ser apodos para mí. Un nombre que no eclipsa ninguna de las tantísimas virtudes vuestras. Esto es un agradecimiento público a todo cuanto habéis hecho por mí. Porque por vosotras ahora soy como soy. Me habéis hecho encontrar mi propio Norte en todo momento. Me habéis sacado de una monotonía y rutina de la que me negaba a salir y a hacerme reír de mis propios errores y aprender de cuantos cometí en el pasado. Pensad que entre estos dos corchetes que escribo a continuación vuelvo a escribir cuanto alguna vez hice [...] Ana Píriz ha sido mi faro y guía desde el principio en el laboratorio y tanto dentro como fuera de la EBD me demostró en todo momento su

transparencia y sinceridad. Siempre irradia felicidad y transmite alegría de vivir, seguridad y confianza siempre que le he pedido cualquier consejo. Viajes, cervezas, fiestas, globos de helio...son muchos recuerdos los que me llevo contigo ¡y muchos más los que espero que compartamos! ¡Espero ir algún día con vosotros y Luzdi a algún concierto de la Selva! Carlos Rodríguez. Nunca me perdonaré el no haberte podido aplaudir en el congreso de la British Ecological Society en Glasgow. Porque te lo mereces como amigo, marido, padre e investigador. Siempre he pensado que el lenguaje es el propio embudo para la comunicación. Y cualquier cosa que escriba ahora sobre Laura Soriano sería un claro ejemplo. Sería injusto transcribir en palabras cuánto hemos compartido, cuánto hemos vivido y cuánto deseo que sigamos forjando esta amistad, porque no reflejarían más que una ínfima parte de la realidad. Mil gracias guapa por todo. María Méndez (Mirentxu, Marutxi o Pikatxu), mi vasca favorita, rabo de lagartija, la mujer más risueña y enérgica que he conocido. Llegaste un poquito más tarde que yo y rápidamente te hiciste con las riendas de tu tesis doctoral. Nos hemos ayudado, desahogado, alegrado y celebrado en todos los análisis que ambos hemos realizado. Pensar en ti es recordar gran parte de los mejores momentos vividos en esta etapa de mi vida... ¡mil gracias por compartirlos conmigo! Hace tiempo que la princesa Mon Ica de la dinastía de los Gutiérrez me permitió entrar en su cuento de princesas y castillos. Mónica se desvive igual por el arte y la cultura que por una buena amistad. Afortunadamente puedo compartir hoy día las tres cosas con ella y me he contagiado esa forma de ver la vida como un cuento. Mireia Casas llegó hace relativamente poco al grupo y sé que ha encontrado en la EBD más de lo que jamás habría esperado encontrar. Te lo mereces todo eso y más. Hemos tenido conversaciones antológicas que recordaré con una sonrisa bien grande. Yareli Esquer vino, vio y venció. Es un diamante en bruto para la ciencia y un pedazo de pan y lo dejó bien claro desde el primer día que vino al laboratorio. Que vayan preparando el Arco del Triunfo porque París seguro que te recibirá como mereces. ¡Mil gracias!

Si algo malo tiene esta población de biólogos en el mundo es que la tenemos que considerar “abierta”. Por un lado enriquece la constante inmigración de nuevos conocidos y amigos que han ido pasando por la EBD, pero por otro entristece cuando la tasa de emigración aumenta. Afortunadamente con muchos de ellos sigo manteniendo aún una muy buena amistad (¡y espero que a partir de ahora pueda ir a visitaros!). Begoña Martínez, mi querida Tata, pueden pasar meses sin que nos veamos que cada



vez que volvemos a encontrarnos parece que no ha pasado el tiempo por ti. Investigadora como la copa de un pino, mi revisora de análisis personal y amiga fiel donde las haya. ¡Mil besos guapetona! Y al acordarme de ella, inevitablemente hablo de Esperanza Ursúa y su tan reconocida risa. Probablemente la EBD ya no es lo mismo desde que no nos deleitas con esas risas a carcajadas que se oían hasta en las cámaras frigoríficas del sótano del Pabellón del Perú. Tienes el don de hacer feliz a cuantos te rodean y luchar por tus principios e ideales. Nuestra amistad nos ha brindado momentos inolvidables desde comiendo succulentas ensaladas en nuestros despachos a las tantas de la noche hasta compartir una ruta por el cono Sur de la Patagonia argentina (Fitz Roy...aquel pico que nunca llegamos a ver por la dichosa nubecita...). Y todo eso fue gracia a esa primera infusión rara que me ofreciste una vez y que acabó instaurándose como la tradición de nuestro mate diario con tantas y tantas conversaciones. Mi más sincera reverencia porque pocas personas como tú hay en este mundo...Y Roger Jovani siempre estuvo (y está) ahí. En el mismo despacho compartiendo bombilla y mate, ideas y pensamientos. Roger transforma la realidad en números y estos en un juego de niños con una gran base teórica a sus espaldas. Más o menos por el 2005 conocí a Ainara Cortés y desde ese primer momento hemos estado al pie del cañón, cada uno viento en popa a toda vela, compañeros de alguna que otra penuria pero sobretodo de muchas alegrías...ya mismito la Dra. Cortés-Avizanda llegará a este mundo...¡y que se preparen todos! ¿Y qué hubiera sido de mí si no hubiera compartido tanto con la banda gangrena? Con Sonia Cabezas tuve la inmensa suerte de compartir mis primeros años en la EBD y guardo bajo llave para que no se me escape ningún recuerdo relacionado con ella. ¡Espero de todo corazón poder verte muy pronto! El mundo entero tiene mucho (muchísimo) que aprender de Paola Laiolo. Bondad, amabilidad, honestidad, humildad, simpatía e inteligencia se han mezclado y tomado forma de Paola. Ella me ha ayudado tanto a interpretar algunos de los resultados presentados en los capítulos de esta tesis como revisado versiones preliminares de los manuscritos. Te debo una visita a tu Asturias patria querida...¡y una buena sidra! El aporte exótico de la banda gangrena nos lo dieron Azucena Canto y José Hernán Sarasola, dos personas con quien he tenido la suerte de compartir mucho fuera de la EBD pero con quienes el tiempo nos ha colocado en puntas distintas del mundo...¡espero veros pronto y que nos juntemos de nuevo todos por aquí! Manu Grande e Isa Luque...anda, pófavó...dejad ya esas tierras tan heladas y venid para acá de nuevo, que ni el mueble bar ni las cervecitas de fin de

semana son lo mismo sin vosotros. ¡Ah! y mil gracias por poner la pincelada musical argentina en mi vida...

En Irene Salicini he encontrado a una gran colaboradora científica, confesora, tertuliana y amiga. Con ella he podido pasarme horas y horas seguidas hablando desde los patrones filogeográficos de refugio dentro de refugio, tratar de definir el verdadero significado de la amistad y analizar las letras de las canciones de Raffaella Carrà. No sé cuál será la mejor definición de amistad, pero tú sí que eres un buen ejemplo para demostrar su existencia. Joan Navarro apareció en mi último año de tesis y basta un minuto de conversación con él para conocerle y un rato largo para hojear su larga trayectoria profesional. ¡Llegarás muy lejos “don Juan”! Cuando quieras podemos comenzar ese estudio de las aves miopes y retomar nuestros entrenamientos ¿vale?. Otra compañera inolvidable de carreras es María del Mar Delgado, que se me fue a tierras finlandesas y ahí sí que estás haciendo entrenamientos en condiciones extremas. Recordaré siempre tantas conversaciones mientras salíamos correr y durante esos cigarritos (incluso “virtuales”) en la terraza del Perú. ¡Espero de todo corazón volver a verte muy pronto!

En algún lado habrá que cenar ¿no? Zulima Tablado ha podido decirme esta frase cientos de veces...y en todas y cada una de ellas he picado. Por causas o azares, como diría algún cantautor, apenas éramos más que compañeros de trabajo hasta hace casi unos meses...y desde entonces podemos decir aquella antológica frase del séptimo arte de “esto puede ser el comienzo de una bonita amistad”. Zulima expresa el mejor significado de la unión de los opuestos, del Yin y el Yang, pues reúne tanto a la mejor imitadora de los huevos poetas del mundo como a la más meticulosa e inteligente investigadora de la EBD. (Jefes, no leáis esto): si de algo me alegro de haberme retrasado en la lectura de tesis es porque gracias a este tiempo he conocido a gente como Marcello D’Amico (lo siento, lo nuestro es imposible...pero siempre te agradeceré el chiste del Playmobil), Margarita Florencio (la más risueña, inocente y bondadosa del mundo...), el camarada Simone Santoro (el italiano más fiel a sus principios, honesto y legal que conozco...te dedico esta canción: “sinceritàààà...”), Letizia Campioni (que escondes detrás de alguien aparentemente reservada a una muy buena amiga, compañera de eventos jolgórico-festivos y apasionada de tu trabajo), Chiara Bettega (la escritora y trampeadora de búhos más eficiente del mundo) y Bruno Fant (profeta de la música francesa y futuro de la investigación en Francia junto con la

tímida pero altamente participativa Justine Bisson). Todos vosotros habéis hecho de esta última etapa de mi tesis doctoral una etapa que nunca olvidaré. Cualquier problema desaparecía momentáneamente cuando nos juntábamos en el café de las 18.00h o con una cervecita delante. Me habéis hecho olvidar los posibles malos momentos y regalado de los grandes recuerdos que me llevo de esta etapa que estoy cerrando...por todo ello...¡mil gracias mozos!

What goes on, Chu! Giulia Bastianelli apareció de repente en mi vida y, como en el Mago de Oz, hizo que todo lo que hasta entonces era blanco y negro cobrara vida y color...y música. Me has hecho entender lo que hasta ahora era ininteligible y conocer cuanto hasta ahora me calificaba de completo ignorante. Me has hecho subir a la mesa de Robin Williams en el Club de los Poetas Muertos y verlo todo desde una perspectiva diferente. Hablar de ti es hablar de música, de literatura, de historia, de política, de compromiso social, de honestidad...has revolucionado mi forma de entender la vida (¡incluso el propio concepto de parmigiano!). Tengo tanto que agradecerte que vamos a tener que recorrer aún mucho camino juntos...¿te parece?

En Laboratorio de Ecología Molecular (L.E.M) nada hubiera sido igual de no ser por la impecable coordinación de Ana Píriz y buen funcionamiento de los técnicos de plantilla Mónica Gutiérrez, Juan Canales, Antonio López y Álvaro Ramos. La secuenciación de la Región Control completa pudo llevarse a cabo gracias a José Manuel Llanes Ferro durante el año 2007 y antes de mi llegada estuvieron echando una mano en el proyecto Mónica Gutiérrez y María González Tirante. Con Juan L. García (Juanele), Joaquín Muñoz y Miguel Alcaide he podido hablar de técnicas avanzadas en ecología molecular para pasar a tomar “una” cerveza en el Chile (o en la plaza de algún pueblo malagueño o pista de algún bar con canciones de karaoke...) y disfrutar de alguno de los grandes momentos vividos en estos años de tesis doctoral.

Tantos investigadores en formación y postdoctorales han pasado por la Estación en estos años que recurro a agradecer al colectivo de Precarios (y con ellos, a todos cuantos se me haya olvidado incluir en estas líneas) por haber logrado tantas cosas en relativamente tan pocos años. Promovidos por el único deseo de mejorar la formación del personal investigador hemos logrado en pocos años más de lo que en otros centros jamás hayan siquiera imaginado. Gracias a Grego Toral por ser tan osado representante de Precarios y rey del taburete en los bares, a Amparo Carrillo y Ana Montero por

retomarlos y sacarlos adelante con éxito cuando parecía estar en decadencia (¡y por tantas y tantas juerguecillas!) y David Canales y a Rosa Agudo (alias Sagudo) por defender siempre que hay vida más allá de la Ciencia. A Juan Antonio Galarza le agradezco enormemente ser siempre casi un libro de referencia ante cualquier duda estadística o técnica que haya tenido en el laboratorio. Carlos Rouco llegó hace un año a la EBD y tuve la suerte de conocer (y trabajar) cuando comencé en otoño de 2003 en el Instituto de Investigación en Recursos Cinegéticos (I.R.E.C.) en Ciudad Real. En esta etapa inicial conocí a Ana Pérez (mi siamesina), con quien tuve el placer de comenzar nuestra carrera investigadora. Guardo desde entonces un especial cariño y buenos recuerdos de nuestros recorridos por las fincas de Despeñaperros, serranía de Cuenca, trabajo de laboratorio...En esta misma época tuve lo que considero el gran honor de trabajar junto a Miguel Delibes Mateos, a quien me costaba mucho invitar a tomar una cerveza y no llegaba bien a entender por qué tanto agobio...una vez comencé yo mi etapa predoctoral, llegué a entenderlo perfectamente. Mil gracias Miguel por enseñarme todo cuanto aprendí en esos meses antes de comenzar en la EBD y por estar siempre ahí por aquel entonces ¡y hoy día!. Gracias en general a todo el personal del IREC que me apoyaron en todo momento desde el principio: Pablo Ferreras, Rafael Villafuerte, Diego Villanúa, Óscar Rodríguez, Paqui Castro...y un muy largo etcétera.

No puedo dejar de agradecer los inestimables comentarios y aportaciones a parte de los manuscritos presentados en esta tesis doctoral realizados por Nuno Ferrand y Paulo Célio Alves y personal del Centro de Investigaçao em Biodiversidade e Recursos Geneticos (CIBIO) (Porto, Portugal).

Two unforgettable experiences abroad were lived during my PhD training at the School of Biological Sciences in the University of Aberdeen (Scotland, UK), hosted by Stuart Piertney and his molecular ecology lab. Stuart received me even without knowing anything at all about me and treated me as another student, same privileges, same dedication...I cannot imagine another better first experience as mine with any other host different than you. We have also deeply discussed not only a couple of the chapters included in this also, but also another non-included analyses (but hopefully published soon!). Tons of thanks Stuart...Talking about Aberdeen would not be the same if I do not make some reference to another two “human-beings” called Dr. Edoardo Tedesco and Dr. Matthew K. Oliver (spaguetti and huggies munchers). Folks, friends, colleagues...thank you very much for everything: hard work during four months, such

patience you had with me, not throwing me the “unnamed” (staple...but I used to give it another name...). I really hope to see you all three musketeers soon! Of course, I am deeply glad to all people from the University as Dr. Xavier Lambin, Thomas Cornulier, Nina Baumgartner, Deborah Davidson, Dan Mifsud, Chrysoula Goubili, Isidora Katara, Eva Maria Zetsche, Lillian Sandeman, Simon Goulak, Nicolas Pade...and a neverending list. Además de nativos y foráneos escoceses también tuve el inmenso placer de verme inmerso en la red de “españolitos” en Aberdeen. Gema Hernández me abrió las puertas de su casa (y el salón de su casa) para pasar unos días y al final me quedé de okupa dos meses en un acogedor colchón. Lucía González y su cangrejo Sebastián me acompañaron en esos dos inolvidables primeros meses en Escocia. A mi buena y gran amiga Lucía Martín “Zagalita”...mi lucerillo. Mil gracias por haber elegido ese año Aberdeen para hacer tu Erasmus y por no dejarme sólo ni uno de los viajes que hicimos en cada fin de semana de mi estancia allí. Sin duda, la mejor “copilota” que puede tener nadie. Ese año conocí a Íñigo Martínez y Ruth Fernández, quienes sin dudarle un minuto me acogieron en Jasmine (su casa) en el mismo sofá que un año había dejado en casa de Gema. Tengo tanto que aprender de los dos que invitaría a todo el que lea esta sección a conocerles. No tengo más que agradecimientos y buenas palabras hacia ellos. Espero subir pronto a daros un fuerte abrazo a los dos. En esta última estancia en Aberdeen coincidí con Ilu Pagán, una grandísima persona con quien (por causas o azares...) hemos vuelto a coincidir gracias a su tan estrecha amistad con otros miembros de la EBD...¡y afortunadamente conmigo tras aquellos dos meses en la tierra del whisky y la cerveza! Ya has estado por México...y no me cabe duda de que devorarás el mundo entero.

And you, Philippe Gaubert, my dear friend...prepare yourself because this is truly finishing! We met during your postdoc at the EBD and that was a real beginning of a friendship and maybe for future collaborations. Thank you for your phylogenetic-advice and your (bad) jokes. We will rock together soon, dude! Ya ya ya!

Durante el verano y parte del otoño de 2004 abandoné todo en Doñana para recibir clases de los cursos de doctorado con mención de calidad “Evolución: la base de la Biología” en la Universidad de Granada. Allí nos juntamos un pequeño grupo de gente interesados en evolución y que disfrutábamos discutiendo acerca del lamarckismo y teorías como la del diseño inteligente todos los días en clase y en los previos de los exámenes e incluso nos daba tiempo a desarrollar nuevas teorías como las de los

enzimas inteligentes (¿verdad Marian?). A ellos (y a los profesores de los cursos) les agradezco enormemente habernos aproximado a las teorías evolutivas...y lograr discutir las con conocimiento de causa. Esta época no hubiera sido la misma sin vosotros: la pri-morcilla María Roldán, Miguel Ángel Calero, Javier Pérez, Marian Pereira, Inma Manrique, Carolina Puerta y Cuauhtemoc Chávez... ¡mil gracias mozos!

A quienes me han visto crecer poco a poco en la biología. Mis compañeros de carrera y muy buenos amigos Carlos Mackintosh, Alicia Ginel, María Jesús Serrano, David Macías y Mamen García. Escribir cuánto he vivido con vosotros, cuánto me habéis enseñado, cuánto he aprendido, cuánto hemos disfrutado, reído, sufrido...con todos vosotros necesitaría cuatro o cinco volúmenes más. Sabéis que os estoy eternamente agradecido por estar siempre ahí, porque de alguna manera u otra estoy donde estoy gracias a vosotros. Habéis sido lo mejor de esa etapa universitaria, sin duda, y dado que nunca había dejado constancia pública de cuánto habéis significado para mí (también como grandes psicólogos durante mi etapa predoctoral) no quería dejar pasar la oportunidad de hacerlo en esta tesis doctoral. ¡Mil gracias!

Les debo muchísimo a otros jiennenses que han hecho mucho más de lo que imaginan por mí. Tamara ha estado, está y estará siempre ahí. No tengo duda alguna. Eres probablemente la persona con quien más he caminado en esta vida... ¡y cuánto nos queda aún! Juan, casi un hermano que creo que no ha terminado de entender por qué he sacrificado tanto para la tesis... ¡espero ahora que la termino poder volver a salir al campo contigo tanto como antes!. Manuel, sabes que te idolatro por tu forma de pensar y vivir acorde a tus ideales... Aitor, casi se puede decir contigo también que nos hemos criado y te tendría que agradecer toda una vida. Ya lo celebraremos con una cervecita delante, ¿ok? Rosaida... porque podrían pasar años sin vernos que cada vez que nos reencontramos eres la misma de siempre... mil gracias a todos!!!

Absolutamente nada de esto hubiera sido posible sin el incondicional apoyo de toda mi familia. Es este un momento muy especial para mí, pues nunca os he podido agradecer públicamente todo cuanto habéis hecho por mí...y, aunque considere injusto reducirlo a tan sólo unas palabras, ahora tengo la mejor oportunidad en esta sección que será leída por el 99% de quienes hayan tenido en sus manos esta tesis. A mi madre (mama, sin tilde en la a) María Antonia Cuadros le debo mi personalidad y constancia en el trabajo. Siempre me has ofrecido tu ayuda en todo cuanto he hecho y (como siempre has dicho)

a llevarme la tacita de chocolate en todo momento y en cualquier lugar del mundo, incluso a pesar de que he empleado palabrotas tan feas como “estocasticidad” o “cadenas de Markov”. Supongo que habré sido fuente de preocupaciones de vez en cuando pero siempre me has entendido y respetado. Has tenido la capacidad de hacerme abstraer de mí mismo, ver los problemas desde fuera y racionalizarlos. Me has hecho ver que desde dentro siempre tendemos a magnificar los problemas...y no tienes más que respirar, alejarte un poquito y darte cuenta de que todo tiene solución. Encontrarla, era siempre bastante más fácil de lo que imaginaba. Always look on the bright side of life...Mi padre (papa, sin tilde en la a), Carlos Centeno...mi faro y guía en el mundo del conocimiento, la cultura y las artes y letras. Cada día que pasa justifica un poquito más el tenerte como icono de mi idolatría. Desde luego mi capacidad para trasnochar trabajando y ejercer de ratón de biblioteca sé que la he heredado de ti. Cuando quieras te cuento lo que he escrito en este libro y trataré de hacer que entiendas algo más del “que” que decía Lorca. Y mi hermanita, Marian Centeno (la tosi...) me ha enseñado el saber apasionarme por cuanto hago. Hace poco recordaba cuando era pequeño y veía a mi hermana realmente dedicada al teatro, el cine, la música y los idiomas. Hoy día has canalizado esa pasión y te está dando la oportunidad a disfrutar de tu trabajo en la docencia así como en la compañía de teatro. Sé que de alguna manera u otra, mi dedicación al cine y muy especialmente a la música vienen de ti...y yo aprendí a canalizar esas “otras pasiones” hacia este otro mundo de la Biología. ¡Mil gracias hermanita!. Carlos Centeno (mi toso favorito...) ¿qué hubiera hecho yo sin ti todo este tiempo? Has ejercido de amigo, confesor, consejero...tu trayectoria profesional como periodista siempre me ha parecido un ejemplo de superación y fruto de todo el gran esfuerzo que has empleado para alcanzarlo. Sin duda, también has sido todo un ejemplo de hermano mayor (pero poco más mayor, ¿eh? que año y medio más tampoco es mucho...jeje, perdón, ahora yo ejercía de hermano menor refunfuñón). ¡Y mil gracias Buzo por sacarme de paseo siempre que lo necesitaba! Y también para mis tíos Juan Cuadros, María Cuadros, María José Úbeda y Rosa Flores junto al enano Alex y la futura pequeñaja que está a puntito de llegar y por el lado motrileño Carmen y Adela Centeno...¡¡¡esta tesis va por todos vosotros!!!

