



Filogenia, morfometría y especiación de dos topillos ibéricos: *Microtus duodecimcostatus* y *Microtus lusitanicus*

Ana Rodríguez Prieto

ADVERTIMENT. La consulta d'aquesta tesi queda condicionada a l'acceptació de les següents condicions d'ús: La difusió d'aquesta tesi per mitjà del servei TDX (www.tdx.cat) i a través del Dipòsit Digital de la UB (diposit.ub.edu) ha estat autoritzada pels titulars dels drets de propietat intel·lectual únicament per a usos privats emmarcats en activitats d'investigació i docència. No s'autoritza la seva reproducció amb finalitats de lucre ni la seva difusió i posada a disposició des d'un lloc aliè al servei TDX ni al Dipòsit Digital de la UB. No s'autoritza la presentació del seu contingut en una finestra o marc aliè a TDX o al Dipòsit Digital de la UB (framing). Aquesta reserva de drets afecta tant al resum de presentació de la tesi com als seus continguts. En la utilització o cita de parts de la tesi és obligat indicar el nom de la persona autora.

ADVERTENCIA. La consulta de esta tesis queda condicionada a la aceptación de las siguientes condiciones de uso: La difusión de esta tesis por medio del servicio TDR (www.tdx.cat) y a través del Repositorio Digital de la UB (diposit.ub.edu) ha sido autorizada por los titulares de los derechos de propiedad intelectual únicamente para usos privados enmarcados en actividades de investigación y docencia. No se autoriza su reproducción con finalidades de lucro ni su difusión y puesta a disposición desde un sitio ajeno al servicio TDR o al Repositorio Digital de la UB. No se autoriza la presentación de su contenido en una ventana o marco ajeno a TDR o al Repositorio Digital de la UB (framing). Esta reserva de derechos afecta tanto al resumen de presentación de la tesis como a sus contenidos. En la utilización o cita de partes de la tesis es obligado indicar el nombre de la persona autora.

WARNING. On having consulted this thesis you're accepting the following use conditions: Spreading this thesis by the TDX (www.tdx.cat) service and by the UB Digital Repository (diposit.ub.edu) has been authorized by the titular of the intellectual property rights only for private uses placed in investigation and teaching activities. Reproduction with lucrative aims is not authorized nor its spreading and availability from a site foreign to the TDX service or to the UB Digital Repository. Introducing its content in a window or frame foreign to the TDX service or to the UB Digital Repository is not authorized (framing). Those rights affect to the presentation summary of the thesis as well as to its contents. In the using or citation of parts of the thesis it's obliged to indicate the name of the author.

**FILOGENIA, MORFOMETRÍA Y ESPECIACIÓN
DE DOS TOPILLOS IBÉRICOS:**

Microtus duodecimcostatus y *Microtus lusitanicus*

Ana Rodríguez Prieto



FACULTAD DE BIOLOGÍA
DEPARTAMENTO DE BIOLOGÍA ANIMAL
PROGRAMA DE DOCTORADO DE BIODIVERSIDAD

Filogenia, morfometría y especiación de dos topillos ibéricos:
Microtus duodecimcostatus y Microtus lusitanicus

Memoria presentada por **Ana Rodríguez Prieto** para optar al título de
Doctora por la Universidad de Barcelona

Trabajo realizado en el Instituto de Biología Evolutiva (CSIC-UPF)

Ana Rodríguez Prieto

Barcelona, Julio, 2013

Director

José Castresana Villamor

Investigador Científico

Instituto de Biología Evolutiva (CSIC-UPF)

Tutor

Miquel A. Arnedo Lombarte

Profesor Titular

Universitat de Barcelona

AGRADECIMIENTOS

Durante los cuatro años que ha durado esta tesis doctoral han pasado por mi vida una gran cantidad de personas imprescindibles que me han ayudado a que este proyecto fuera posible. Además de todas las personas, esta tesis ha ido acompañada de muchas horas de música, libros y viajes que han sido necesarios para poder liberar la mente en los momentos de mayor cansancio y agotamiento de fuerza y energía.

Sin todo mi grupo, el p59, esto no hubiera sido posible. Muchas conversaciones y discusiones sobre ciencia, mucho trabajo en equipo, muchas risas, algún que otro mal trago y muchas birras han aderezado este trabajo. Javi, rubito, gracias por haber empezado antes que yo el doctorado. Has sido como un hermano mayor que me ha ido dejando muchas ideas, mucho trabajo ya hecho y que me ha ayudado en todo momento. Espalda contra espalda. Víctor, qué hubiera hecho sin tu ayuda informática. Alex, gracias por la acidez de tus comentarios, imprescindibles para sobrevivir en el mundo de la ciencia. Joan, el hermano adoptado, ha sido genial compartir estos años con un animal de *sangre fría*. No olvidaré tus días de contacto con las pipetas. Marina, el último refuerzo, a pesar de que hemos coincidido poco ha sido muy intenso. Me hubiera encantado compartir más horas contigo, pero ya sabes que el whatsapp hace milagros. Jose, supongo que encontrar el ajuste perfecto entre dos personas que trabajan juntas día a día es un ejercicio de madurez. Me llevo una buena sensación al haber trabajado cuatro años contigo, he aprendido mucho de ciencia, pero también has hecho que me haga un poco más mayor a nivel personal.

A mis dos tutores, Salvador Carranza y Miquel A. Arnedo por soportar el tedio burocrático.

Agradezco mucha la ayuda y compañía de otras personas del IBE con las que he podido trabajar y compartir momentos durante estos cuatro años. No me puedo olvidar de todos los de administración que habéis vivido todos mis momentos de locura cuando los pedidos no llegaban o cuando los aparatos se rompían.

Voglio ringraziare tutta la gente dell'Università di Pavia, che mi ha accolto così bene quando sono andata a lavorare con loro. Roberto, grazie mille per essere così nerd di R, per insegnarmi un sacco di cose di morfometria e di statistica. Però, soprattutto grazie per avermi dato sempre una mano nonostante fossimo lontani.

Ha sido fundamental toda la gente e instituciones que me han ayudado a recoger egagrópilas, Jose, Jesús, Mar, Manolo, Javi, Lorena, Mónica y Luis, o que me ha enviado muestras, Jacint Ventura, Pablo García, el BTAC y el Museo de Ciencias de Nueva York.

Fuera del IBE ha habido un grupo de gente enorme que ha estado en todo momento a mi lado, y que ha sabido animarme en los momentos más difíciles del doctorado.

Colorattos, todos los que habéis pasado por esta casa, Pri, Manuela, Marius, Leo, Patejo y Dea (cagna maledetta). Habéis sido geniales. Me habéis dado las risas, las fiestas, las conversaciones, la tranquilidad y los viajes necesarios para tener una energía infinita. Cargoles (que sois miles), sin unos vecinos así de majos la vida en nuestro Kerido

barrio el Karmelo hubiera sido mucho más aburrida. Bionenas, sois una pieza clave de esta tesis. Vuestra compañía siempre ha servido para seguir adelante, muchas gracias por haberme hecho un hueco en este grupo de locas. Circenses, que venís de un mundo tan diferente al mío y habéis conseguido atraparme. Siempre me acordaré de cuando os decía que hacía un doctorado con vomitados de lechuzas, menudas caras. María, esto viene de hace mucho, pero muchas gracias por haber aparecido ese día y no haberte ido nunca. Pri, qua si campa d'aria. Sei stato un punto d'inflessione nella mia vita a Barcelona, non avrei mai detto che finivamo a vivere qua assieme, ma è successo ed è stato meraviglioso. Manuela, que te voy a decir si ya me lo dices tu todo, gracias por tu energía y puff, al canal!. Patejo, un placer compartir de nuevo una experiencia de nuestras vidas. Haber hecho el doctorado a la vez que tú me ha hecho mucha ilusión, la comprensión ha sido máxima y la ayuda inolvidable.

A todos mis otros amigos y a mis tíos-amigos, a lo mejor no habéis vivido tan de cerca esta tesis, porque estábamos un poco lejos pero es como si hubierais estado. Muchas gracias por los mails, por los skypes y por las veces que nos hemos visto. El apoyo lo he notado siempre, sois muy grandes.

A tutti i miei amici italiani, tutto il tempo che ho passato in Italia in questi quattro anni, siete stati super carini. Tutto il gruppo di fotografi pazzi, indimenticabili tutti i momenti a cercare il posto giusto per questa maledetta foto. Al gruppo trentino, un super grazie per accogliermi nella città. Mic, mi sono piaciute sempre le conversazioni di scienza con te, molto interessanti e utili per la scrittura di questa tesi.

A toda mi familia —la de siempre y la nueva— le agradezco la comprensión acerca del trabajo que he estado haciendo estos cuatro años. Seguro que para muchos es algo surrealista que yo recogiera vómitos de lechuza y mirara que había dentro, pero, siempre me han dado mucho apoyo y muchísimo cariño. Las abuelas han sido unas fieles seguidoras del trabajo, muy atentas y siempre preguntándose si era necesario tener dos carreras para recoger algo que ellas limpiaban de sus casas porque hacía feo.

Padres, sois unos fans definitivamente incondicionales. Nunca lo había dudado, pero ésta ha sido una prueba de fuego. Habéis aguantado todos los temporales de desánimo y frustración dando todavía más ánimos y cariño. Habéis seguido el trabajo, como buenos biólogos que sois, con pasión y con ganas de aprender. No habéis faltado en el momento que he tenido que tomar decisiones importantes, siempre con un buen consejo preparado. Supongo que todo esto significa ser unos buenos padres y vosotros lo habéis hecho a la perfección. Sois magníficos.

Fabio, secondo me tra poco ti faranno santo. Lo sai quanto mi è cambiata la vita da quando ci conosciamo però anche come è cambiato il mio dottorato. Grazie per non permettermi mai di mollarlo. Grazie per donarmi un sacco di consigli di scienza, per l'aiuto quando abbiamo lavorato assieme, per i consigli artistici. Però soprattutto, per la pazienza che hai avuto nei momenti peggiori, e per non farmi mai sentire sola, che sempre eri presente per appoggiarmi e farmi un sacco di coccole, senza dubbio questa è la cosa più importante. Sei un punci ñopez.

Gracias a todos, he disfrutado mucho de esta tesis doctoral, he aprendido y me he hecho un poco más como persona. Siempre será inolvidable.

CONTENIDO

I. INTRODUCCIÓN	1
1. DEL DESORDEN AL ORDEN	3
1.1 ESPECIE Y ESPECIACIÓN	3
1.2 BIOGEOGRAFÍA Y GENÉTICA DE POBLACIONES EN EL PROCESO DE ESPECIACIÓN	5
2. DE LOS GENES A LAS ESPECIES	9
2.1 ÁRBOLES DE GENES	9
2.2 ÁRBOLES DE ESPECIES	11
2.3 MÚLTIPLES MARCADORES MOLECULARES: VARIOS, PERO CUÁLES.	16
3. DESCRIBIR O CUANTIFICAR	18
3.1 PRIMEROS PASOS	18
3.2 MORFOMETRÍA GEOMÉTRICA: SUS PUNTOS FUERTES	19
3.3 APLICACIONES PRÁCTICAS DE LA MORFOMETRÍA GEOMÉTRICA	21
4. PROTAGONISTAS: EL TOPILLO MEDITERRÁNEO Y EL TOPILLO LUSITANO	22
4.1 SUS ORÍGENES	22
4.2 SUS CARACTERÍSTICAS	25
4.2.1 <i>Microtus duodecimcostatus</i>	25
4.2.2 <i>Microtus lusitanicus</i>	26
4.3 SU HÁBITAT	26
4.4 SUS DIFERENCIAS MORFOLÓGICAS TRADICIONALES	28
II. OBJETIVOS	31
1. OBJETIVOS	33
III. MATERIALES Y MÉTODOS	35
1. DESARROLLO DE NUEVOS MARCADORES NUCLEARES PARA EL ORDEN <i>RODENTIA</i>	37
1.1 OBTENCIÓN Y FILTRADO DE INTRONES DE ROEDORES	37
1.1.1 <i>Obtención de los genomas</i>	37
1.1.2 <i>Filtros bioinformáticos</i>	37
1.2 VALIDACIÓN EXPERIMENTAL DE LOS INTRONES SELECCIONADOS	39
1.2.1 <i>Diseño de cebadores</i>	39
1.2.3 <i>Amplificación por PCR</i>	41
1.2.4 <i>Reconstrucción filogenética</i>	42
2. ANÁLISIS MORFOMÉTRICO DEL TOPILLO MEDITERRÁNEO Y DEL TOPILLO LUSITANO EN SUS DIFERENTES ÁREAS DE DISTRIBUCIÓN	43
2.1 OBTENCIÓN DE MUESTRAS Y LOCALIDADES DEL ESTUDIO	43
2.1.1 <i>Obtención y procesado de las muestras</i>	43

2.1.2 Localidades de muestreo	44
2.2 MORFOMETRÍA GEOMÉTRICA	45
2.2.1 Material fotográfico	45
2.2.2 Selección y tratamiento de landmarks	46
2.3 PROTOCOLO PARA LA ASIGNACIÓN DE ESPECIE	47
2.3.1 Asignación de especie a los individuos de las zonas de distribución alopátrica	47
2.3.2 Asignación de especie a los individuos de las zonas de distribución solapada	47
2.4 ANÁLISIS ESTADÍSTICO DE LOS DATOS	49
2.4.1 Análisis estadístico de los datos de las zonas de distribución alopátrica	50
2.4.2 Análisis estadístico de los datos de las zonas de distribución solapada	50
2.4.3 Análisis morfométrico del tercer molar superior (M3)	51
2.5 PLATAFORMAS INFORMÁTICAS	51
3. ANÁLISIS DE LAS RELACIONES FILOGENÉTICAS ENTRE EL TOPILLO MEDITERRÁNEO Y EL TOPILLO LUSITANO MEDIANTE EL USO DE MARCADORES MITOCONDRIALES Y NUCLEARES	52
3.1 OBTENCIÓN DEL MATERIAL BIOLÓGICO	52
3.2 EXTRACCIÓN DE ADN GENÓMICO	52
3.2.1 Extracción de ADN genómico a partir de material óseo	52
3.2.1 Extracción de ADN genómico a partir de muestras de tejido fresco	53
3.3 AMPLIFICACIÓN POR PCR	53
3.3.1 Amplificación de secuencias mitocondriales	54
3.3.2 Amplificación de secuencias nucleares	54
3.3.3 Evaluación de la posible pérdida de alelos	56
3.4 OBTENCIÓN DE LAS FASES HAPLOTÍPICAS	56
3.4.1 Obtención computacional de las fases haplotípicas	56
3.4.2 Obtención experimental de las fases haplotípicas	57
3.5 ANÁLISIS FILOGENÉTICO DE LOS MARCADORES AMPLIFICADOS	58
3.5.1 Análisis filogenético de las secuencias mitocondriales	58
3.5.2 Análisis filogenético de las secuencias nucleares	58
3.6 DISCORDANCIA CITO-NUCLEAR Y MORFOLÓGICA: CLASIFICACIÓN DE LOS INDIVIDUOS	58
3.7 ANÁLISIS DE LA DIVERSIDAD Y DE LA ESTRUCTURA GENÉTICA	60
3.8 SELECCIÓN DE LOS MODELOS EVOLUTIVOS Y DEL RELOJ MOLECULAR MÁS ADECUADO PARA LOS MARCADORES NUCLEARES	61
3.9 CÁLCULO DE LAS TASAS EVOLUTIVAS DE LOS INTRONES SELECCIONADOS PARA EL GÉNERO <i>MICROTUS</i>	61
3.10 CÁLCULO DEL TIEMPO DE ESPECIACIÓN DE LAS DOS ESPECIES DE TOPILLOS Y DE SUS PARÁMETROS DEMOGRÁFICOS MEDIANTE ÁRBOLES DE ESPECIES	64
3.10.1 Estima de parámetros empleando *Beast	64
3.10.2 Estima de parámetros empleando IMA2	65
3.11 IDENTIFICACIÓN DEL PROCESO QUE CAUSA LA DISCREPANCIA CITO-NUCLEAR: INTROGRESIÓN O SEPARACIÓN INCOMPLETA DE LINAJES	66
IV. RESULTADOS Y DISCUSIÓN	69
1. DESARROLLO DE NUEVOS MARCADORES NUCLEARES PARA EL ORDEN <i>RODENTIA</i>	71
1.1 RESULTADOS	71
1.1.1 Conjunto de intrones seleccionados	71
1.1.2 Validación experimental de los intrones para el orden <i>Rodentia</i>	73
1.1.3 Análisis de la reconstrucción filogenética de los intrones amplificados	75
1.2 DISCUSIÓN	76
1.2.1 Características generales del conjunto final de intrones	76
1.2.2 Posibles aplicaciones de los intrones seleccionados	77

2. VARIACIONES MORFOMÉTRICAS EXISTENTES ENTRE <i>M. DUODECIMCOSTATUS</i> Y <i>M. LUSITANICUS</i> RELACIONADAS CON SUS ÁREAS DE DISTRIBUCIÓN	78
2.1 RESULTADOS	78
2.1.1 Identificación y caracterización morfológica de las dos especies de topillos en las zonas de distribución alopatrica.....	78
2.1.1.1 Asignación de especie y correlación con la distribución geográfica	78
2.1.1.2 Variación interespecífica e intraespecífica del tamaño.....	80
2.1.1.3 Variación interespecífica e intraespecífica de la forma	82
2.1.2 Identificación y caracterización morfológica de las dos especies de topillos en las zonas de distribución solapada.....	85
2.1.2.1 Asignación de especie	85
2.1.2.2 Variación interespecífica e intraespecífica del tamaño.....	89
2.1.2.3 Variación interespecífica e intraespecífica de la forma	90
2.2 DISCUSIÓN	95
2.2.1 Individuos de las zonas de distribución alopatrica	95
2.2.2 Individuos de las zonas de distribución solapada	96
3. CARACTERIZACIÓN DE LAS RELACIONES FILOGENÉTICAS ENTRE <i>M. DUODECIMCOSTATUS</i> Y <i>M. LUSITANICUS</i> MEDIANTE EL USO DE SECUENCIAS MITOCONDRIALES Y NUCLEARES	100
3.1 RESULTADOS	100
3.1.1 Variabilidad interespecífica del citocromo b y de los intrones	100
3.1.2 Análisis de la historia evolutiva intraespecífica e interespecífica: *Beast e IMA2.....	103
3.1.2.1 Uso de calibraciones fósiles en la filogenia del orden <i>Rodentia</i>	103
3.1.2.2 Cálculo de las tasas evolutivas de los intrones para el género <i>Microtus</i>	104
3.1.2.3 Estima del tiempo de divergencia entre las dos especies de topillos	105
3.1.2.4 Estima de los parámetros demográficos ancestrales y actuales de los dos topillos ibéricos	107
3.1.2.5 Análisis de la incongruencia cito-nuclear bajo un escenario de introgresión o de separación incompleta de linajes.....	108
3.2 DISCUSIÓN	114
3.2.1 Diferentes escenarios para la historia evolutiva de <i>M. duodecimcostatus</i> y <i>M. lusitanicus</i>	114
V. CONCLUSIONES	121
VI. APÉNDICE.....	127
VII. BIBLIOGRAFÍA	161

ÍNDICE DE FIGURAS

Figura 1. Proceso de especiación	4
Figura 2. Relación espacial entre los rangos de distribución geográficos de dos especies	6
Figura 3. Mecanismos que permiten el mantenimiento de dos o más especies como entidades separadas	7
Figura 4. Árbol de especies reconstruido bajo un modelo de coalescencia	11
Figura 5. Incongruencia en la topología entre un árbol de especies y un árbol de genes.. ..	12
Figura 6. Diferentes procesos que provocan la existencia de incongruencias topológicas entre los árboles de especies y los árboles de genes.....	13
Figura 7. Señal filogenética registrada en un árbol de especies y en un árbol de genes de un proceso de separación incompleta de linajes y de un proceso de introgresión.....	15
Figura 8. Conjunto de procesos básicos en un análisis de morfometría geométrica.....	19
Figura 9. Diferentes eventos de colonización del género <i>Microtus</i>	23
Figura 10. Distribución estratigráfica y posibles relaciones filogenéticas entre el género <i>Allophaiomys</i> y el género <i>Terricola</i> en Europa	25
Figura 11. Características de <i>M. duodecimcostatus</i> y <i>M. lusitanicus</i>	26
Figura 12. Áreas de distribución de <i>M. duodecimcostatus</i> y <i>M. lusitanicus</i> en la Península Ibérica.....	27
Figura 13. Características morfológicas tradicionalmente usadas para la diferenciación entre <i>M. duodecimcostatus</i> y <i>M. lusitanicus</i>	29
Figura 14. Puntos geográficos de las localidades de muestreo	43

Figura 15. Selección de <i>landmarks</i> para los cinco caracteres	45
Figura 16. Uso de calibraciones fósiles para la estima de tasas evolutivas de marcadores moleculares.....	62
Figura 17. Conjunto de filtros bioinformáticos aplicados para la obtención de un conjunto final de 208 intrones	71
Figura 18. Análisis filogenético de los diez intrones amplificados en diez especies del género <i>Rodentia</i>	75
Figura 19. Clasificación de los individuos de las zonas de distribución alopátrica sin información <i>a priori</i> sobre la especie a la que pertenecen	79
Figura 20. Valores de logCS de los individuos de las zonas distribución alopátrica para los cinco caracteres.....	81
Figura 21. Análisis relativo de deformación para los cinco caracteres de los individuos de zonas de distribución alopátrica.....	83
Figura 22. Análisis de la variación del M3 para los individuos de las zonas de distribución alopátrica de las dos especies de topillos	84
Figura 23. Clasificación de los individuos de las zonas de distribución alopátrica y de las zonas de distribución solapada sin información <i>a priori</i> sobre la especie a la que pertenecen.....	86
Figura 24. Clasificación de los individuos de las zonas de distribución alopátrica y de las zonas de distribución solapada de las dos especies para los cinco caracteres mediante algoritmos E-M.....	88
Figura 25. Valores de logCS de los individuos de las zonas distribución alopátrica y de las zonas de distribución solapada para los cinco caracteres.....	89
Figura 26. Análisis relativo de deformación para los cinco caracteres de los individuos de zonas de distribución alopátrica y de las zonas de distribución solapada de las dos especies de topillos	92

Figura 27. Análisis de la variación del M3 para los individuos de las zonas de distribución alopátrica y solapada de las dos especies de topillos	93
Figura 28. Genealogías haplotípicas de los seis intrones y del citocromo <i>b</i> amplificados en 44 individuos	103
Figura 29. Comparación de las tasas evolutivas de los seis intrones amplificados obtenidas mediante el uso de calibraciones fósiles con las distribuciones posteriores de las tasas evolutivas obtenidas con *Beast y IMA2.....	106
Figura 30. Comparación de las estimas del tiempo de especiación de <i>M. duodecimcostatus</i> y <i>M. lusitanicus</i> calculados con *Beast e IMA2.....	110
Figura 31. Árbol de inferencia Bayesiana del citocromo <i>b</i> del total de los 184 individuos	113
Figura 32. Modelización del árbol de especies de <i>M. duodecimcostatus</i> y <i>M. lusitanicus</i>	119

ÍNDICE DE TABLAS

Tabla 1. Diez intrones amplificados en diferentes especies del género <i>Rodentia</i>	40
Tabla 2. Resultados de la amplificación por PCR de diez intrones en las diez especies del género <i>Rodentia</i> seleccionadas.....	74
Tabla 3. Asignación de especie a los individuos de las zonas de distribución alopatrica mediante un análisis discriminante para los cinco caracteres.....	78
Tabla 4. Análisis de la varianza del logCS en los individuos de zonas de distribución alopatrica para los cinco caracteres	80
Tabla 5. Análisis multivariante de la varianza de la forma de los cinco caracteres para los individuos de zonas de distribución alopatrica	82
Tabla 6. Análisis de la varianza del logCS en todos los individuos de las dos especies... ..	90
Tabla 7. Análisis multivariante de la varianza de la forma de los cinco caracteres en todos los individuos de las dos especies.....	91
Tabla 8. Cuantificación de la variación del módulo de los vectores de forma en el espacio para los cinco caracteres	94
Tabla 9. Análisis de la variabilidad genética de las dos especies de topillos	102
Tabla 10. Tasas evolutivas de los seis intrones amplificados calculadas para el género <i>Microtus</i>	104
Tabla 11. Tiempo al antecesor común más cercano de los seis intrones y el citocromo <i>b</i> calculados con *Beast.....	107
Tabla 12. Tamaños efectivos poblacionales ancestral y actuales de las dos especies de topillos obtenidos con *Beast e IMA2.....	108
Tabla 13. Valores del índice GSI.....	111

ÍNDICE DE FIGURAS DEL APÉNDICE

Figura A1 Árbol de máxima verosimilitud reconstruido con el concatenado de 2288 intrones obtenidos a partir de los filtros bioinformáticos	131
Figura A2. Distribución de frecuencias de los valores de <i>scaling factor</i> y de <i>K score</i> de los 2288 intrones.....	132
Figura A3. Árboles de máxima verosimilitud de los diez intrones reconstruidos con las secuencias amplificadas en diez especies de roedores	133
Figura A4. Distribución de frecuencias de los valores de logCS del carácter ventral de los individuos de las zonas de distribución alopátrica.....	148
Figura A5. Árbol de máxima verosimilitud reconstruido con las 184 secuencias mitocondriales de las dos especies de topillos.....	150
Figura A6. Árboles de máxima verosimilitud reconstruidos con las secuencias de los intrones amplificados en <i>M. duodecimcostatus</i> y <i>M. lusitanicus</i>	152
Figura A7. Distribuciones de las edades calculadas para cada nodo calibrado.....	159

ÍNDICE DE TABLAS DEL APÉNDICE

Tabla A1. Especies del orden <i>Rodentia</i> utilizadas para la extracción de ADN genómico	129
Tabla A2. Modelos de sustitución nucleotídica y relojes moleculares calculados para cada uno de los diez intrones	130
Tabla A3. Individuos recolectados de las dos especies de topillos	134
Tabla A4. Descripción de los <i>landmarks</i> seleccionados para los cinco caracteres.	138
Tabla A5. Valores del primer y del segundo componente de deformación para cada uno de los cinco caracteres para todos los individuos.	140
Tabla A6. Asignación de especie a los individuos de las zonas de distribución alopátrica y de las zonas de distribución solapada	144
Tabla A7. Individuos de las dos especies de topillos de los que se extrajo ADN genómico a partir de tejido fresco	148
Tabla A8. Conjunto de cebadores diseñados para la amplificación de las secuencias mitocondriales y de las secuencias nucleares de las dos especies de topillos	149
Tabla A9. Asignación de especie genética y morfológica a los individuos de los zonas de distribución alopátrica y de las zonas de distribución solapada	153
Tabla A10. Modelos de sustitución nucleotídica y relojes moleculares calculados para cada uno de los seis intrones y para el citocromo <i>b</i>	158
Tabla A11. Combinación de individuos y marcadores incluidos para cada uno de los análisis realizados con *Beast e IMA2 para el cálculo de diferentes parámetros	160
Tabla A12. <i>Priors</i> utilizados para cada uno de los análisis realizados bajo un modelo de aislamiento con migración y sin migración.....	160

I. INTRODUCCIÓN

1. DEL DESORDEN AL ORDEN

Para un naturalista probar a definir las especies equivale a “probar a definir lo indefinible” (Charles Darwin, en una carta del 1856 a su amigo Joseph Dalton Hooker).

1.1 ESPECIE Y ESPECIACIÓN

Una de las mayores preocupaciones de la biología contemporánea ha sido encontrar una definición universal del concepto de especie, entendido como una definición que englobe todas las diferentes propuestas que se han realizado. Hay multitud de conceptos de especie postulados —alrededor de 24 (Mayden 1997)— y cada uno de ellos reclama una serie de requisitos que los organismos deben cumplir para constituirse como una especie con entidad propia. Se ha definido la especie desde un punto de vista biológico (Wright 1940; Mayr 1942; Dobzhansky 1950), ecológico (VanValen 1976; Andersson 1990), filogenético (Hennig 1966; Rosen 1979; Donoghue 1985; Mishler 1985; Ridley 1989; Avise & Ball 1990; Baum & Shaw 1995; Meier & Willmann 2000), genotípico (Mallet 1995) y fenético (Michener 1970; Sokal & Crovello 1970; Sneath & Sokal 1973), entre otros. Esta diversidad de conceptos de especies indica que la identificación de los límites que existen entre dos especies ha sido extremadamente difícil de identificar. Esta situación se produce debido a que estos conceptos de especie tradicionales proponen requisitos inquebrantables en la gran mayoría de los casos y, casi siempre, incompatibles entre las diferentes definiciones de especie.

Ya que la especie es una unidad fundamental en Biología, ya sea para la descripción de la biodiversidad o para la elaboración de proyectos de conservación, sería necesario encontrar un concepto unificador. Una de las mejores tentativas ha sido la definición propuesta por Kevin De Queroiz (De Queroiz 1999, 2007), donde una especie es “un segmento de un linaje metapoblacional que evoluciona separadamente de otro”, entendiéndose linaje como una serie de antecesores y descendientes (Simpsons 1961; Hull 1980) y metapoblación como una población inclusiva formada por subpoblaciones conectadas (Levins 1970; Hanski & Gaggiotti 2004).

I. Introducción

En realidad, esta nueva definición extrae la esencia común de todas las anteriores definiciones. Esta definición, siendo menos concreta que las otras, representa de forma mucho más clara lo que realmente sucede en la naturaleza. Es decir, el proceso de especiación no es un fenómeno ni que ocurra rápido en el tiempo ni que siga unos pasos predefinidos, como podría ser por ejemplo que primero se produce un aislamiento reproductivo, luego una diferenciación de nicho ecológico y luego una monofila de las variantes alélicas (De Queiroz 1999). Por todo esto, la especiación se podría definir como un conjunto de procesos evolutivos, como las mutaciones, la deriva genética, la migración (o su ausencia) y la selección natural, que afectarán a diferentes caracteres en un modo diverso y harán que los linajes empiecen a divergir. Durante este continuo se podrán identificar los requisitos imprescindibles recogidos por los conceptos de especie tradicionales (aislamiento reproductivo, aparición de nicho ecológico propio, características morfológicas propias, monofilia de genes, etc.) y que ocurrirán en el tiempo con una determinada velocidad y con un determinado orden dependiendo del proceso de especiación que se este estudiando (figura 1) (De Queiroz 2007). Se tiene que tener en cuenta que cuanto más avanzado esté el proceso más fácil será reconocer los límites entre las especies nuevas que se están formando. Por esto, se debe presentar mucha más atención a los primeros estadios de la especiación para poder detectarla y caracterizarla.

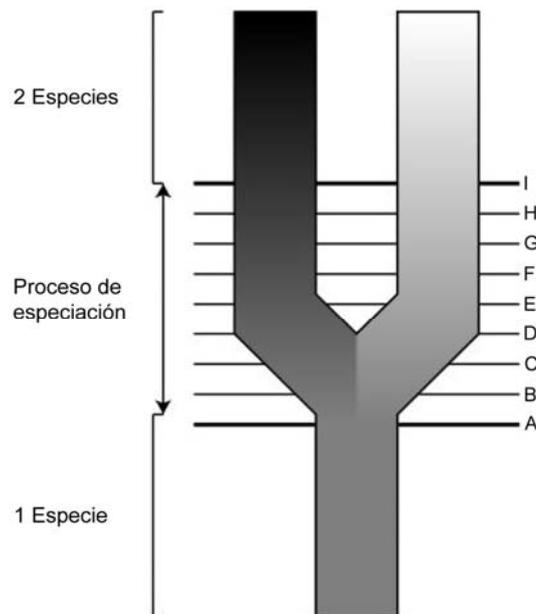


Figura 1. Proceso de especiación. Cada una de las letras de la A a la I indica la adquisición de una característica que hace posible la separación de la especie ancestral en dos especies diferentes. Estas características se corresponden con los requisitos, que cada concepto de especie, retiene necesario para la existencia de una completa separación de dos linajes. Modificado de De Queiroz (2007).

1.2 BIOGEOGRAFÍA Y GENÉTICA DE POBLACIONES EN EL PROCESO DE ESPECIACIÓN

Dos factores que influyen en gran medida en la capacidad reconocer el proceso de especiación y la delimitación de las especies es el tipo de relación espacial que existe entre las especies divergentes y la existencia o no de flujo génico entre ellas. Por tanto, la caracterización del proceso de especiación se puede realizar bajo un marco biogeográfico y de genética de poblaciones. Además de estas dos consideraciones, en el análisis del proceso de especiación se debe incluir un marco temporal. Así, por un lado, se debe distinguir entre cómo y dónde se produjo el proceso de especiación, es decir, cuáles fueron los mecanismos que favorecieron la especiación y cuál era la relación espacial existente de las especies incipientes. Y por otro lado, se debe precisar cuáles son los mecanismos que mantienen las diferentes especies como entidades separadas en la actualidad y cuál es la relación espacial actual de esas especies. El marco temporal es un elemento importante ya que la distribución geográfica actual de las especies no tiene por qué corresponder a la ancestral (Chesser & Zink 1994; Webb & Gaston 2000; Losos & Glor 2003) y, por tanto, los mecanismos impulsores de la especiación pudieron ser diferentes a los que mantienen actualmente las especies como entidades independientes. Estimar los rangos de distribución ancestrales mediante la suma de los rangos actuales de distribución de las especies implicaría asumir que ni unos ni otros hayan podido sufrir eventos de contracción o expansión por diferentes motivos, situación que podría resultar bastante improbable (Losos & Glor 2003). Por tanto, se han utilizado diferentes técnicas de estimación indirecta (registro fósil, correlación de la distribución geográfica y el tiempo de especiación con diferentes eventos geológicos, registros paleoclimáticos y características de las zonas geográficas de distribución actuales [Schliewen 1994]) de estos rangos de distribución ancestrales y tratar así determinar el origen geográfico de la especiación.

Una vez establecido el marco temporal en el que se estudia el proceso de especiación (pasado o actual) los procesos de especiación se pueden incluir en tres diferentes categorías: alopatría, parapatría y simpatría. Debido a la similitud de características geográficas y de genética de poblaciones, estos conceptos se pueden reagrupar, por un lado, en alopatría y, por otro, en simpatría y parapatría.

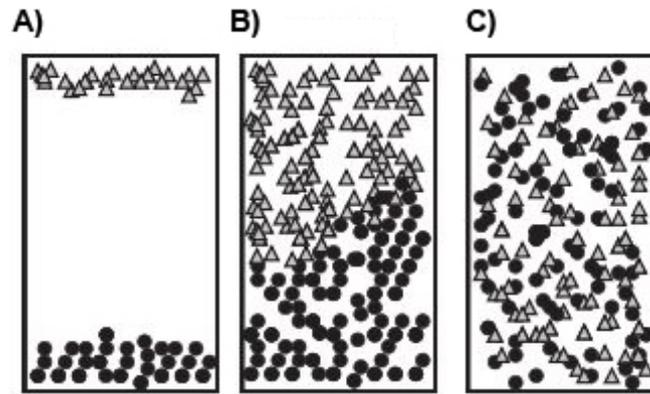


Figura 2. Relación espacial entre los rangos de distribución geográficos de dos especies. **(A)** Alopatría. Las dos especies presentan rangos de distribución totalmente separados debido a la existencia de una barrera geográfica. **(B)** Parapatría. Las dos especies presentan rangos de distribución adyacentes con una pequeña zona de solapada de rangos **(C)** Simpatría. Las dos especies presentan rangos de distribución totalmente solapados. Modificado de Mallet *et al.* (2009).

Alopatría define el estado de estar en diferentes lugares (Mayr 1942) (figura 2A). Por tanto, dos especies han sufrido un proceso de especiación alopátrico cuando las poblaciones de la especie ancestral estaban separadas por una barrera geográfica, provocando que el flujo génico entre ellas fuera inexistente (Futuyma & Mayer 1980; Losos & Glor 2003; Gavrillets 2003, 2004; Coyne & Orr 2004). Dentro de la categoría de especiación alopátrica se pueden diferenciar diferentes subcategorías. La especiación vicariante, en la que poblaciones que presentaban una distribución geográfica continua sufren una división geográfica y cada uno de los grupos originados —con un tamaño considerable cada uno de ellos— dan origen a una nueva especie (Bush 1994; Losos & Glor 2003). La especiación peripátrica se produce debido al aislamiento geográfico de un pequeño grupo periférico que dará lugar a una nueva especie (Losos & Glor 2003). Dos especies presentan rangos de distribución alopátrica, en la actualidad, cuando sus áreas de distribución se encuentran separadas por la existencia de diversas barreras geográficas (figura 2A). Esta separación sigue impidiendo el flujo génico y mantiene la diferenciación entre las especies que, a su vez, se refuerza con una diferenciación en términos ecológicos y sexuales (especiación ecológica y especiación por selección sexual, respectivamente) (Smadja & Butlin 2011) (figura 3). Normalmente, un proceso de especiación alopátrico —en el pasado— entre dos especies cercanas se corresponde con una distribución alopátrica actual (Losos & Glor 2003). Por tanto, este modo de especiación es uno de los más fácilmente identificables.

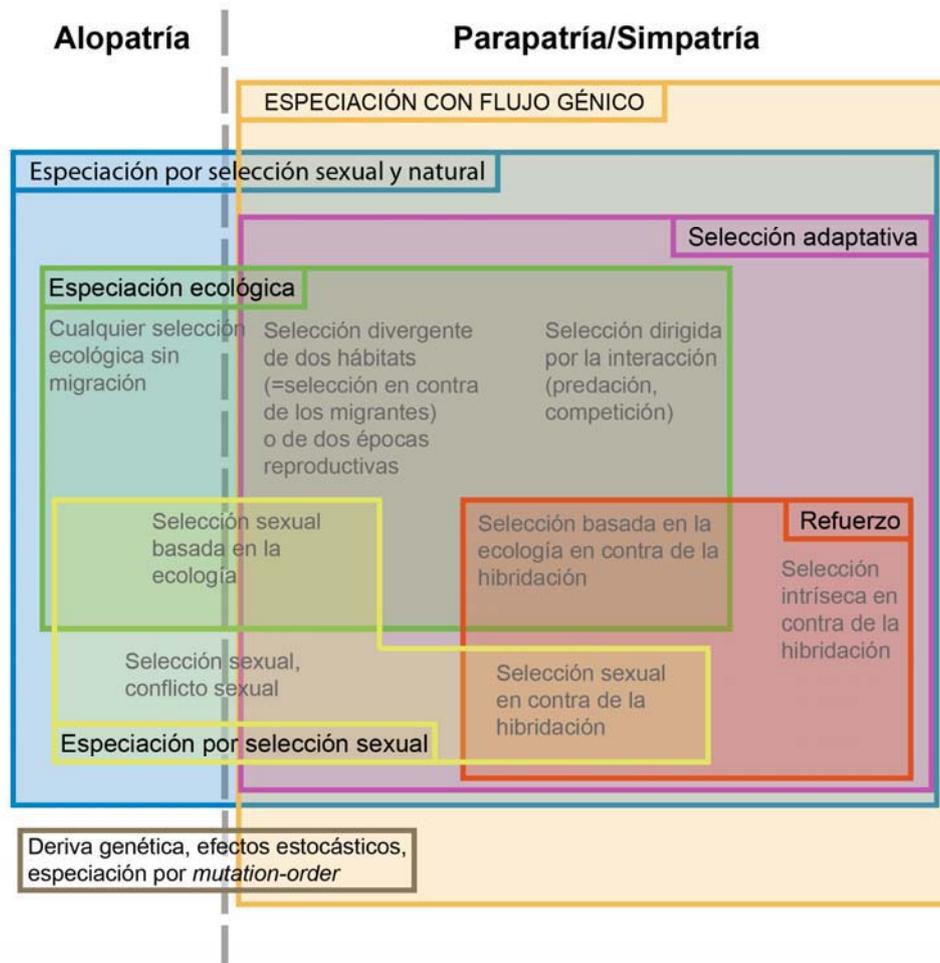


Figura 3. Mecanismos que permiten el mantenimiento de dos o más especies como entidades separadas. Se diferencia si el proceso de especiación se ha producido bajo un escenario en ausencia de flujo génico entre los individuos (alopatría) o bajo un escenario en el que el flujo génico entre individuos de las diferentes especies es posible debido a que comparten rangos de distribución geográfica (parapatría y simpatría). Modificado de Smadja & Butlin (2011).

Simpatría es el estado de estar en el mismo sitio (Poulton 1903) y parapatría es el estado de estar en contacto pero no solapando (Smith 1955) (figura 2B y 2C). Debido a que los dos conceptos representan la existencia de una distribución geográfica en común (más o menos grande, respectivamente) y que el proceso de especiación se verifica bajo la existencia de flujo génico entre las especies incipientes, estos dos conceptos se pueden analizar en conjunto pero sin olvidar que representan dos escenarios de especiación diferentes. Por tanto, la especiación simpátrica ocurre cuando las poblaciones ancestrales que originaron las nuevas especies presentaban el mismo rango de distribución geográfica (Futuyma & Mayer 1980; Bush 1994; Losos & Glor 2003). La especiación parapatrica se verifica cuando las poblacionales ancestrales se encontraban

separadas pero en regiones adyacentes, provocando que una fracción de individuos pudiera contactar (Futuyma & Mayer 1980; Losos & Glor 2003). La dificultad de distinguir entre estos dos procesos de especiación está obviamente relacionada con la distribución espacial de las poblaciones de ancestrales. Es decir, dependiendo del grado de rigurosidad con el que se determine si el territorio de dos especies incipientes estaba totalmente superpuesto o, por el contrario, era solamente adyacente (Bush 1969; Feder 1998; Berlocher & Feder 2002; Coyne & Orr 2004; Savolainen *et al.* 2006; Fitzpatrick *et al.* 2008), nos encontraremos ante un proceso de especiación simpátrico o parapátrico, respectivamente. A pesar de las diferencias en cuanto a la cantidad de extensión geográfica solapada, estos dos tipos de especiación comparten la característica común de que en ambas situaciones la diferenciación entre especies se verificó ante la presencia de flujo génico entre individuos. La interrupción de flujo génico se verificaría de manera progresiva, siendo reforzada por mecanismos de especiación ecológica y sexual (diferenciación de nichos, diferentes épocas de reproducción, diferenciación en los sistemas reproductivos o incompatibilidades genómicas) (Bush 1994; Mallet 2008). Dos especies presentan una distribución simpátrica, en la actualidad, cuando sus rangos de distribución geográfica se encuentran totalmente solapados (figura 2C). Se podría proponer una especiación simpátrica como el origen más plausible de dos especies con distribución actual simpátrica (Losos & Glor 2003). Por el contrario, dos especies con una distribución actual parapátrica (figura 2B) pueden ser el resultado de una especiación parapátrica o de un primer contacto entre especies que han sufrido una especiación alopátrica y una posterior expansión geográfica (Bull 1991).

Tanto en especies con distribución simpátrica (total solapamiento), como para los individuos que viven en las zonas de solapamiento de sus áreas de distribución parapátrica, existe un gran número de mecanismos que permiten la existencia de las especies como entidades separadas (figura 3). Todos estos conceptos se podrían agrupar bajo el concepto de la especiación adaptativa. Dentro de este concepto se identifican mecanismos ecológicos de especiación como la selección divergente de hábitats o las interacciones entre especies (como la competencia o la depredación). Igualmente se identifican procesos de selección sexual y procesos de refuerzo, entendidos como procesos que implican una menor eficacia biológica de los híbridos resultantes de la reproducción de individuos de dos especies diferentes (Smadja & Butlin 2011). Este conjunto de mecanismos que promueven la diferenciación de especies en simpatria y/o

parapatría (en el pasado) y que mantienen las diferencias entre especies con distribución simpátrica y/o parapátrica (en la actualidad) indica que el proceso de especiación en presencia de flujo génico entre individuos no es un proceso único, sino un conjunto de mecanismos que minimizan el flujo génico y refuerzan la barreras reproductivas, y que encaminan el proceso de especiación a su conclusión (Smadja & Butlin 2011) (figura 3).

Identificar los mecanismos concretos que están operando, verificar la completa interrupción de flujo génico, identificar la presencia de individuos híbridos e identificar el papel de estos híbridos en el proceso de especiación son tareas que requieren una gran cantidad información genética, ecológica, filogenética, morfológica y geográfica que permita identificar y describir el punto exacto en el que se encuentra el proceso evolutivo que se está estudiando para poder entender de manera más precisa la relación existente entre dos o más especies establecidas o incipientes (Wiens & Penkrot 2002; Sites & Marshall 2004; Dayrat 2005; Leaché *et al.* 2009; Padial 2010; Schlick-Steiner *et al.* 2010).

2. DE LOS GENES A LAS ESPECIES

2.1 ÁRBOLES DE GENES

Las relaciones filogenéticas existentes entre las especies se representan mediante árboles en los que las longitudes de las ramas y sus diferentes topologías indican cuál ha sido la historia evolutiva de estas especies desde su origen. El material que mejor contiene esta información y que se utiliza para estas reconstrucciones es el ADN, a su vez formado por genes. Cada uno de estos genes presenta una historia evolutiva diferente. Es decir, cada uno de ellos ha cambiado a través de las diferentes especies a un ritmo determinado y en una proporción determinada. Este parámetro se cuantifica en biología evolutiva en número de sustituciones por posición. Los árboles filogenéticos resultantes se denominan árboles de genes. La longitud de las ramas de estos árboles representa mutaciones. Con este tipo de árboles podemos indicar si uno de los linajes que se están estudiando ha sufrido más mutaciones por posición respecto a otros linajes, para un gen determinado. Debido a que un mismo gen no se comporta del mismo modo en cada unas de las especies, o bien por presentar tasas mutacionales distintas (para

especies lejanas) o bien por la aparición de un mayor número de mutaciones por cuestiones estocásticas (para especies cercanas), el resultado serán árboles de genes con diferentes longitudes de ramas.

Si se quiere conocer en cuánto tiempo se ha producido estas mutaciones estos árboles se tienen que transformar en árboles de genes ultramétricos. Para ello es necesario conocer las tasas evolutivas, es decir, incluir una referencia del tiempo en el que se producen las mutaciones por posición. Por tanto, los taxones terminales del árbol estarán a la misma altura temporal. Cada uno de los nodos de estos árboles representa el tiempo al antecesor común más cercano (TMRCA, *time to the most recent ancestor*). La cuestión que se plantea ahora es cómo introducir el tiempo en una reconstrucción filogenética. En 1965 Zuckerkandl y Pauling utilizaron por primera vez el término reloj molecular (Kumar 2005). Calcularon el número de diferencias existentes entre dos proteínas de dos especies de las cuáles conocían su tiempo de divergencia. Con este valor de referencia pudieron estimar los tiempos de divergencia entre otras especies de los que sólo se disponía información relativa a la proteína utilizada en el estudio (Kumar 2005). Se establecía, por tanto, el concepto de que el número de diferencias existentes —en una secuencia aminoacídica o en una secuencia de ADN— eran directamente proporcionales al tiempo que separaba los taxones analizados. El concepto de reloj molecular estricto quedaba instaurado (Zuckerkandl & Pauling 1965). Se asumía que la tasa de mutación (número de cambios) era constante a lo largo de toda la filogenia, es decir, todos los taxones sufrían cambios al mismo ritmo. Posteriormente (Hasegawa *et al.* 1987) se propuso el concepto de relojes moleculares relajados (Kumar 2005). Este concepto refleja la posibilidad de asignar a diferentes partes de la filogenia con la que se trabaja una tasa de mutación diferente, es decir, cada linaje podría evolucionar a un ritmo diferente. Ante esta dualidad de relojes moleculares se presenta la duda de cuál de ellos utilizar para introducir el factor tiempo en una filogenia. Actualmente existen una serie de tests estadísticos que predicen cuál es el tipo de reloj molecular que más se ajusta al marcador utilizado y, gracias a la inferencia Bayesiana, esta información puede ser introducida para las reconstrucciones filogenéticas (Kumar 2005).

A pesar de que existe la posibilidad de obtener estas reconstrucciones filogenéticas (árboles de genes) que recogen una gran cantidad de información acerca de las relaciones entre taxa, no se puede considerar que la reconstrucción filogenética de estos

árboles sea equivalente a la reconstrucción filogenética de una especie como entidad evolutiva, ya que cada uno los genes ha sufrido diferentes procesos evolutivos (Knowles & Kubatko 2010), en especial, cuando se realizan reconstrucciones filogenéticas de especies cercanas.

2.2 ÁRBOLES DE ESPECIES

La historia filogenética de una especie se obtiene, en un modo más completo, mediante los árboles de especies. Los árboles de especies son el resultado de la combinación de los árboles de genes usando la teoría de la coalescencia (Kingman 1982). Por tanto, estos árboles de especies están basados en la información de las genealogías, es decir, en cada una de las filogenias que representa la historia evolutiva de cada gen para cada una de las especies (figura 4).

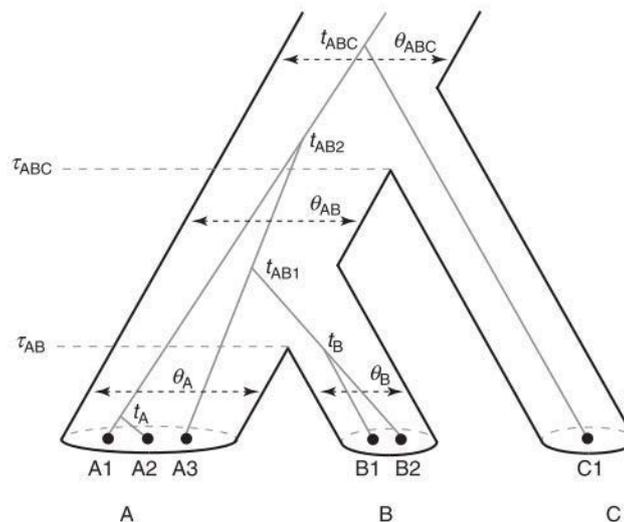


Figura 4. Árbol de especies reconstruido bajo un modelo de coalescencia. Las líneas negras representan las relaciones filogenéticas entre las tres especies (A y B especies hermanas y C como grupo externo). La anchura de las ramas refleja el tamaño poblacional ancestral y actual (θ). Los nodos representan tiempo de especiación (τ). En gris se representa la reconstrucción filogenética de un solo gen dentro del árbol de especies. Cada nodo representa el tiempo de coalescencia de cada alelo en la población, es decir, el TMRCA (t). Modificado de Fujita *et al.* (2012).

Cada uno de los nodos del árbol de especies representa un evento de especiación en unidades de tiempo, τ . La anchura de las ramas representa el tamaño poblacional de cada una de las especies, ancestral y actuales. El tamaño poblacional, θ , viene definido por la ecuación $\theta=4N_e\mu$ para marcadores diploides, donde μ representa la tasa de mutación por posición por tiempo de generación y N_e es el tamaño efectivo poblacional,

que representa el número de individuos reproductores en una población que contribuirán al acervo genético en la siguiente generación (Fujita *et al.* 2012). Las longitudes de las ramas se pueden medir también en unidades de tiempo de coalescencia. En este caso, la longitud de las ramas es directamente proporcional al número de generaciones e inversamente proporcional al tamaño efectivo poblacional. Por tanto, la presencia de ramas cortas en un árbol de especies puede ser debida a un bajo número de generaciones, a un elevado tamaño poblacional, o a ambas situaciones a la vez (Degnan & Rosenberg 2009). Si, por el contrario se quisiera expresar la longitud de las ramas en número de mutaciones esperadas por generación se debería multiplicar el tiempo de especiación por $\theta/4$ ($\theta=4N_e\mu$) (Degnan & Rosenberg 2009). Otro de los parámetros que se pueden calcular en un árbol de especies y sus genealogías es el tiempo de coalescencia de un alelo en una población, t (figura 4), que es el mismo anteriormente descrito como TMRCA, pero en este caso calculado bajo las premisas de la teoría de coalescencia.

Por todas estas características los árboles de especies son una herramienta más adecuada para reconstruir la historia evolutiva entre dos especies, ya que son capaces de integrar la información contenida en los diferentes árboles de genes. Por otro lado, previenen de una reconstrucción filogenética errónea debido a la diferente historia evolutiva que pueden haber sufrido diferentes genes afectados por diversos procesos (figura 5).

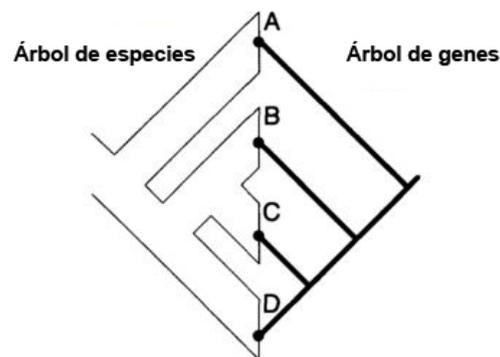


Figura 5. Incongruencia en la topología entre un árbol de especies y un árbol de genes para las especies A, B, C y D. En el árbol de especies las dos especies más cercanas serían B y C mientras que para la reconstrucción filogenética de diferentes genes las especies más cercanas serían C y D. Modificado de Maddison (1997).

Algunos de los fenómenos que pueden provocar esta situación de incongruencia topológica entre árboles de genes y árboles de especies son la duplicación génica, la recombinación, la introgresión y la separación incompleta de linajes (Maddison 1997; Degnan & Rosenberg 2009) (figura 6).

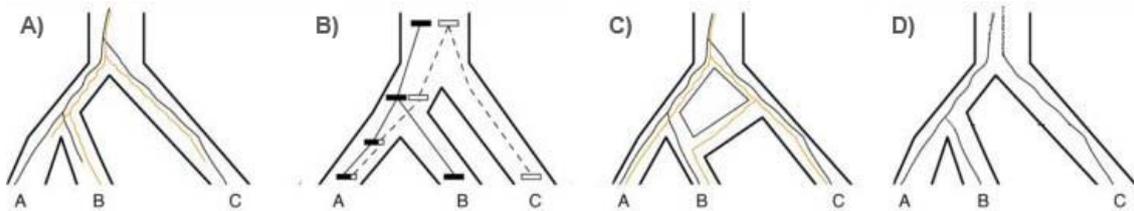


Figura 6. Diferentes procesos que provocan la existencia de incongruencias topológicas entre los árboles de especies y los árboles de genes. **(A)** Duplicación génica. **(B)** Recombinación. **(C)** Introgresión por procesos de hibridación. **(D)** Separación incompleta de linajes. Modificado de Degnan & Rosenberg (2009).

La duplicación génica puede ser una fuente de discordancia ya que después de la duplicación cada gen sufrirá diversos procesos evolutivos. En teoría, estas situaciones se podrían evitar si se realizan controles rigurosos sobre la ortología de los genes que se usan en los análisis (figura 6A).

La recombinación sólo afecta a genes nucleares. Si la información que se recoge en un árbol de genes nucleares incluye en una de las ramas información de un fragmento génico que ha sufrido recombinación (presenta un fragmento paterno y otro materno) podría generar una reconstrucción filogenética diferente a la que presentaría un segmento adyacente del mismo gen que no haya sufrido recombinación (figura 6B) (Degnan & Rosenberg 2009).

La introgresión es entendida como el intercambio génico entre dos especies cercanas después del evento de especiación por procesos de hibridación (figura 6C). En ocasiones este evento es difícil de detectar ya que puede afectar únicamente a una parte del genoma después de varias generaciones (Linnen & Farrell 2007; Peters *et al.* 2007; Jacobsen & Omland 2011; Sequeira *et al.* 2011). Esta incidencia diferencial en diferentes partes del genoma de un proceso de introgresión se ha demostrado que puede ser debida, por ejemplo, a que en las especies de animales con una dispersión diferencial entre sexos será más susceptible de sufrir una introgresión el marcador molecular propio del sexo menos dispersante (Petit & Excoffier 2009). Por este motivo en muchas ocasiones se han detectado procesos de introgresión a nivel del genoma mitocondrial

pero no a nivel nuclear para especies en las que los machos son más dispersantes que las hembras (Good *et al.* 2008; Geraldès *et al.* 2008; Ishida *et al.* 2011; Bastos-Silveira *et al.* 2012; Melo-Ferreira *et al.* 2012). Si en estos casos se hubiera usado únicamente la información mitocondrial se habría podido inferir una historia evolutiva errónea. Por tanto, hay que tener en cuenta cuáles son los marcadores moleculares que son más idóneos en cada situación para estudiar las diferencias entre las especies. En las situaciones donde no se disponga de información *a priori* se deberían realizar análisis combinados tanto de marcadores mitocondriales como nucleares para asegurar una reconstrucción filogenética más correcta.

La separación incompleta de linajes o coalescencia profunda ocurre cuando un polimorfismo ancestral persiste a diferentes eventos de especiación y, como consecuencia, ciertas variantes alélicas se pueden perder en distintos linajes, provocando un árbol de genes diferente al árbol de especies (Maddison 1997) (figura 6D). Las filogenias que presentan mayores tamaños efectivos poblaciones y ramas internas más cortas son más susceptibles a sufrir de retención de polimorfismo ancestrales (Pamilo & Nei 1988).

En muchas ocasiones, la señal filogenética que se registra —mediante la reconstrucción filogenética de árboles de genes— debida a un proceso de introgresión o a un proceso de separación incompleta de linajes es la misma (figura 7). Por tanto, determinar si uno u otro proceso es el responsable de la incongruencia entre los árboles de genes y los árboles de especies ha sido objeto de estudio en muchas ocasiones (Wang *et al.* 1997; Sang & Zhong 2000; Holder *et al.* 2001; Machado *et al.* 2002; Huson *et al.* 2005; Buckley *et al.* 2006; Holland *et al.* 2008). Aparte de una cuestión puramente teórica de poder diferenciar entre los dos procesos, existe además un interés añadido en determinar la existencia real de híbridos entre dos especies.

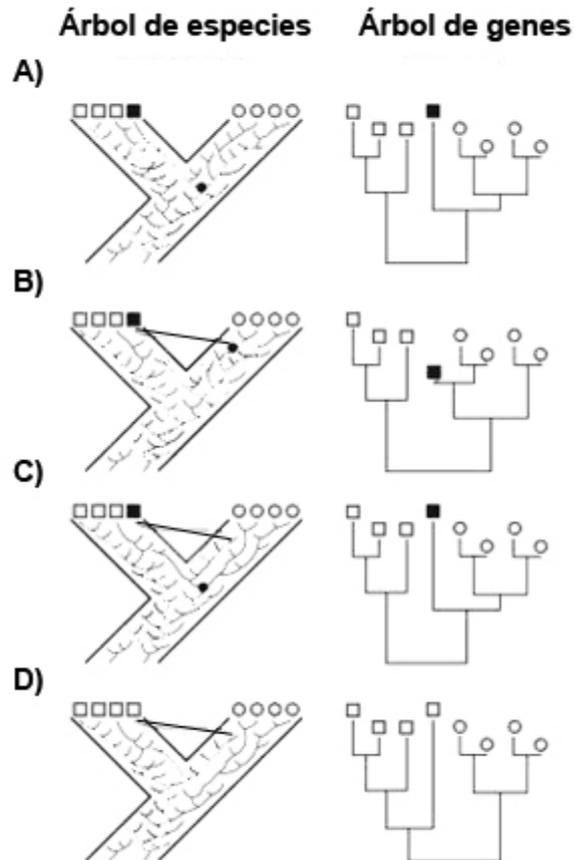


Figura 7. Señal filogenética registrada en un árbol de especies y en un árbol de genes de un proceso de separación incompleta de linajes (A) y de un proceso de introgresión (B, C y D). El cuadrado negro representa una secuencia que es más cercana a la otra especie que a la suya propia. La señal filogenética que se registra en un árbol de genes es la misma para un proceso de separación incompleta de linajes y para un proceso de introgresión (A, B y C). En los árboles de especies para un proceso de separación incompleta de linajes el tiempo de coalescencia de la secuencia incongruente se producirá siempre antes del evento de especiación (A); mientras que en un escenario de introgresión la coalescencia se puede producir después de la especiación (B) o antes de ésta (C). Por último, un proceso de introgresión no origina siempre como resultado una topología parafilética, pudiendo, por tanto, pasar inadvertido (D). Modificado de Joly *et al.* (2009).

El papel desempeñado por los individuos híbridos en el proceso de especiación ha sido siempre muy controvertido. Algunos autores defienden la hibridación como una fuente potencial de variación genética, de novedades funcionales y de nuevas especies. Por el contrario, los detractores de esta teoría opinan que los híbridos son entidades con una menor eficacia biológica que sus conespecíficos parentales y que por tanto son una rama muerta de la evolución (Seehausen 2004). Para el caso concreto de mamíferos se ha demostrado recientemente que en algunas ocasiones la hibridación ha resultado ser una fuente de diversificación y adaptación (Wyner *et al.* 2002; McDevitt *et al.* 2009; Shurtliff 2011). Si la existencia de individuos híbridos puede condicionar el desarrollo de ciertos procesos evolutivos y adaptativos, determinar su verdadera existencia

mediante el análisis de la señal filogenética que haya podido dejar este proceso evolutivo y conseguir diferenciarlo de un escenario de separación incompleta de linajes es un reto que se empieza a abordar con el uso de diferentes metodologías.

En definitiva, desentrañar cuál o cuáles de todos los procesos evolutivos descritos anteriormente son los que han delineado la historia evolutiva de las especies, estimar el tiempo de divergencia y estimar los tamaños poblacionales ancestral y actuales entre dos o varias especies es posible en un modo más preciso y realista gracias a la reconstrucción de árboles de especies, utilizando múltiples marcadores moleculares, bajo los principios de la teoría de la coalescencia.

2.3 MÚLTIPLES MARCADORES MOLECULARES: VARIOS, PERO CUÁLES.

Tanto en el campo de la filogeografía, como en el campo de la filogenia el ADN mitocondrial fue el marcador más empleado durante las dos últimas décadas del siglo pasado (Avice 2000). Su extendido uso quedaba justificado por una serie de características. Principalmente, presenta una elevada tasa de mutación (debido a un sistema más deficitario de reparación de mutaciones [Brown *et al.* 1979]), que se traduce en una gran cantidad de información filogenética. Además, al transmitirse únicamente por vía materna, existe una única copia de cada gen, evitándose de este modo los procesos de recombinación (aunque se han descrito algunas excepciones [Ladoukakis *et al.* 2001]). Por este mismo motivo —herencia materna— y por haber una sola copia de cada gen presenta un menor tamaño efectivo poblacional (N_e), situación que promueve más fácilmente la existencia de monofilia recíproca de los genes entre las especies. Su principal inconveniente es que se reconstruye la historia evolutiva de las especies teniendo en cuenta un único marcador. Es decir, se sobrevaloraba la capacidad de obtener topologías, estimas precisas de tiempo de especiación, de tamaños efectivos poblacionales y de tasas de flujo génico, usando una única fuente de información (Felsenstein 1981; Kuhner *et al.* 1998).

La necesidad de realizar reconstrucciones filogenéticas lo más completas posibles provocó que se empezaran a buscar otro tipo de marcadores aparte de los mitocondriales. Era necesario introducir en estos análisis marcadores que informaran de la historia evolutiva de las especies teniendo en cuenta la herencia genética de los dos

progenitores, es decir, marcadores del genoma nuclear. El problema es decidir cuáles son los mejores marcadores. Los primeros marcadores nucleares que se utilizaron para la reconstrucción de filogenias profundas fueron los exones (Stanhope *et al.* 1992; Porter *et al.* 1996). El mayor inconveniente de estos marcadores es que no son suficientemente informativos como para resolver relaciones filogenéticas en el marco de árboles de especies próximas (Sánchez-Gracia & Castresana 2012).

Otros marcadores que están empezando a utilizarse en estudios filogenéticos son los SNPs (polimorfismos de nucleótido simple). Aunque son muy útiles en genética de poblaciones, estos marcadores presentan dos inconvenientes para árboles de especies. Por un lado, muchas de las técnicas de reconstrucción de árboles de especies basadas en la coalescencia son muy incipientes para estos marcadores (Bryant *et al.* 2012). Por otra parte, no se puede calcular una tasa de mutación para cada uno de ellos. Debido a esto y a que el cálculo de parámetros demográficos es sensible a la presencia de esta información los SNPs no son tampoco, de momento, los mejores candidatos para la reconstrucción de árboles de especies (Maruki *et al.* 2012).

Por tanto, se necesitaría —para la reconstrucción de árboles de especies cercanas— un marcador nuclear que sea lo suficientemente variable como para obtener una reconstrucción fiable de las genealogías, que se pueda calcular su tasa de mutación para obtener estimaciones de parámetros demográficos y temporales más precisas, que se pueda utilizar con las metodologías de reconstrucción de árboles de especies basadas en coalescencia y que se pueda amplificar fácilmente. Los candidatos ideales bajo estos requisitos son los intrones. Los intrones han sido denominados también como EPIC (exon-primed, intron-crossing) (Palumbi & Baker 1994; Creer *et al.* 2005; Li *et al.* 2010), que hace referencia al último de los requisitos, es decir, a que sean fácilmente amplificables. Al encontrarse entre exones —que se caracterizan por estar altamente conservados a lo largo de la filogenia— es posible, por una parte, encontrarlos e identificarlos (usando filtros adecuados) a lo largo del genoma; por otra parte, son fácilmente amplificables por PCR ya que se pueden diseñar cebadores conservados en las regiones exónicas y obtener con relativamente poco esfuerzo secuencias a lo largo de las filogenias.

Aunque estos marcadores sean ideales bajo estas condiciones, no han sido utilizados con gran difusión hasta ahora pero parece que esta tendencia esta empezando a revertir. Así, han sido utilizados para la reconstrucción de árboles de especies (Salicini *et al.* 2011), para la estimación más precisa del tiempo de especiación (Hailer *et al.* 2012), para el análisis de patrones de dispersión (Peters *et al.* 2008) y para discernir entre diferentes escenarios evolutivos posibles (Jacobsen & Omland 2012; Igea *et al.* 2013). Por tanto, una buena estrategia a seguir, aprovechando la disponibilidad de genomas completos de una serie de organismos en las bases de datos, sería buscar y caracterizar conjuntos de intrones que pudieran ser utilizados para resolver preguntas filogenéticas en diferentes grupos de organismos (Igea *et al.* 2010).

3. DESCRIBIR O CUANTIFICAR

3.1 PRIMEROS PASOS

Las primeras clasificaciones taxonómicas mediante la morfología de los seres vivos tenían un carácter meramente descriptivo, es decir, eran medidas de tipo cualitativas. A principios del siglo XX este modo de trabajo cambió radicalmente y se empezó a recoger información de tipo cuantitativa, lo que se conoce como la “Revolución Cuantitativa” (Bookstein 1998). Los primeros estadios de esta revolución promovieron la obtención de diversas medidas de caracteres morfológicos cuantificables y permitieron la comparación entre diferentes individuos mediante el uso de valores medios. El desarrollo de métodos estadísticos como el coeficiente de correlación (Pearson 1895), el análisis de la varianza (Fisher 1935) o el análisis de componentes principales (Pearson 1901; Hotelling 1933) generó una mayor potencia en el uso de las variables cuantitativas.

Pero el verdadero cambio en el estudio de la forma estaba por llegar. Fue a mitad del siglo XX cuando se instauró el concepto de morfometría definido como el estudio de la variación de la forma y su co-variación con otras variables (Bookstein 1991; Dryden & Mardia 1998). Sobre esta base se realizaban medidas sobre la altura, la anchura y la longitud de las estructuras y se aplicaban métodos estadísticos multivariantes para su análisis. El problema de esta metodología residía en que no se podía utilizar toda la

información que estas medidas recogían porque los métodos de análisis no permitían la correlación entre ellas y, por tanto, no se tenía en cuenta la variación de la forma como estructura global. Además, la representación gráfica de la variación en forma era prácticamente imposible debido a la complejidad de representar las matrices numéricas de las variables que se cuantificaban.

A finales de los años ochenta llegó la verdadera innovación en el estudio de la forma: la morfometría geométrica (Rolh & Marcus 1993). La gran ventaja de esta técnica es que se implementa la posibilidad de recoger la geometría de las estructuras morfológicas a analizar y preservar esta información a lo largo de los análisis (Adams *et al.* 2004). En este momento empezó la “Revolución morfométrica” (Rolh & Marcus 1993).

3.2 MORFOMETRÍA GEOMÉTRICA: SUS PUNTOS FUERTES

En los análisis de morfometría geométrica la información de las estructuras se recoge mediante coordenadas (2D o 3D) de *landmarks* (puntos de referencia) morfológicos (Rolh & Marcus 1993). Un buen *landmark* debe ser de tipo homólogo, es decir, debe estar situado en estructuras iguales en diferentes individuos—y, a ser posible, fácilmente reconocibles—. De esta forma, aparte de la información intrínseca sobre la forma se recoge al mismo tiempo información con un sentido biológico. En el caso de que no sea posible identificar este tipo de puntos se utilizarán *pseudolandmarks*, que se situarán en zonas de curvatura, en bordes o en extremos de las estructuras. La presencia de estos *landmarks* nos permite obtener una representación gráfica de la estructura a estudiar fácilmente interpretable (figura 8).

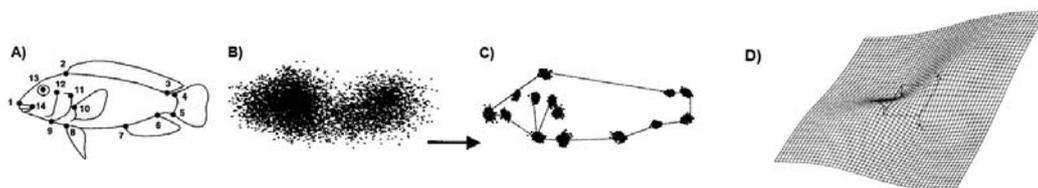


Figura 8. Conjunto de procesos básicos en un análisis de morfometría geométrica. **(A)** Selección de los *landmarks* más adecuados. **(B)** Disposición de los *landmarks* en todos los individuos. **(C)** Escalar, rotar, y trasladar los *landmarks* de todos los individuos mediante un método de superimposición o de *Procrustes* para que se puedan comparar. **(D)** Uso de la función TPS (*thin-plate spline*) para cuantificar las deformaciones, en una malla, de una estructura respecto a otra. Modificado de Bookstein (1991); Adams *et al.* (2004).

Una vez situados los *landmarks* en los lugares seleccionados (para poder comparar las diferencias en forma de los individuos) se tiene que realizar una estandarización de los mismos. Es decir, se tienen que eliminar las diferencias existentes en ubicación, orientación y escala entre cada *landmark* homólogo para cada individuo (figura 8). Esta estandarización de los *landmarks* se realiza mediante los denominados métodos de superimposición o de *Procruster*. Existen diferentes variantes, cada una de ellas con sus particularidades: Coordenadas de forma de Bookstein (Bookstein 1986); GPA, análisis generalizado de *Procruster* (*generalized procruster analysis*) (Rohlf & Slice 1990), GRF, ajuste rotacional generalizado (*generalized rotacional fitted*) (Rohlf & Slice 1990; Slice *et al.* 1996). Una de las más utilizadas es el análisis generalizado de *Procruster* (GPA).

En el momento en el que todos los *landmarks* homólogos están dispuestos en el espacio de una manera estándar la metodología de la morfometría geométrica propone dos opciones. La primera sería describir las diferencias en forma mediante las diferencias existentes entre las coordenadas de cada *landmarks*. Es decir, se evalúa cuánto se desplaza cada *landmark* de una configuración base. Digamos que sería un procedimiento análogo al del ajuste de un conjunto de datos a un modelo de regresión lineal. Las pequeñas diferencias respecto al modelo serían los residuos. Estas diferencias (residuos) pueden ser utilizadas como variables en diferentes análisis multivariantes. La principal ventaja de esta metodología (GPA) radica en el hecho de que es capaz de detectar variaciones en forma cuando éstas están provocadas únicamente por pocos *landmarks*. La segunda opción propondría continuar con el análisis de las coordenadas. Se utilizaría la función *thin-plate spline* (TPS) descrita por Bookstein (Bookstein 1991) para mapear la deformación de una estructura respecto a otra. Es decir, las diferencias en forma se pueden representar como deformaciones de una malla (figura 8D). Los parámetros que describen las deformaciones de esta malla (denominados como *partial warp scores*) son los que representan la variación no uniforme de la forma, es decir, son los que recogen las variaciones locales en forma (Bookstein 1991). Además, la forma está compuesta por las dos componentes uniformes (variación de la forma sobre el eje X y el eje Y [Bookstein 1996; Rohlf & Bookstein 2003]). Tanto la variación uniforme como la no uniforme se pueden utilizar como variables óptimas, para el análisis de la variación de forma mediante métodos multivariantes, formando una matriz global de forma.

Otra de las características a destacar de la técnica de la morfometría geométrica es la capacidad de separar la variación de tamaño de la variación de forma en las estructuras analizadas. Esta implementación es absolutamente necesaria si se quiere realizar una adecuada comparación de la forma. Se podría llegar a conclusiones erróneas si se comparara la forma de dos estructuras cuando una es más grande que la otra. Por tanto y como uno de los primeros pasos descritos anteriormente las estructuras se reescalan antes de compararse. El tamaño en morfometría geométrica de una estructura viene definido por el *centroid size*, CS que es la raíz cuadrada del cuadrado de la suma de la distancia de cada *landmark* al centroide o punto central de la estructura. Éste es el parámetro que se utiliza para reescalar las estructuras. Aparte de este uso, el CS se puede emplear para analizar cómo varía la forma en función del tamaño, por ejemplo, debido a la existencia de procesos alométricos. Por tanto, en morfometría geométrica el tamaño se separa de la forma pero no se abandona.

El último aspecto por el que la morfometría geométrica es una técnica revolucionaria en el estudio de la forma es que no es necesario disponer de información *a priori* sobre cuál será la parte de la estructura más importante a medir o a analizar, ya que sólo es necesaria la selección de los *landmarks* y los análisis posteriores indicarán cuáles son los puntos calientes de variación de forma entre los individuos analizados.

3.3 APLICACIONES PRÁCTICAS DE LA MORFOMETRÍA GEOMÉTRICA

Una de las muchas aplicaciones de esta técnica es la caracterización de las diferencias morfológicas existentes entre diferentes taxa (Wallace 2006; Sidlauska 2008; Serb *et al.* 2011; Kaliontzopoulou *et al.* 2011). De forma más concreta, y debido a que en muchas ocasiones las diferencias morfológicas entre especies cercanas pueden llegar a ser muy sutiles o difícilmente identificables, se ha empezado a usar la morfometría geométrica para identificar los rasgos morfológicos que diferencian dos especies cercanas (Adams & Funk 1997; Martínez & Di Cola 2011). Debido a que en muchas ocasiones existen situaciones donde las especies de estudio comparten áreas de distribución (zonas de simpatria, o en la zona de distribución solapada en distribuciones parapátricas), esta técnica puede ser utilizada para identificar cuáles son los principales procesos que están modelando la forma de las especies en las zonas de contacto y si existen diferencias con las zonas alopátricas. Procesos como el desplazamiento de caracteres morfológicos

(Adams 2004; Rychlik *et al.* 2006), la convergencia morfológica (Adams & Nistri 2010), los procesos de adaptación local a diferentes ecosistemas (Caumul & Polly 2005), la plasticidad fenotípica (Sacchi *et al.* 2007; Piras *et al.* 2010) o la presencia de formas híbridas (Gaubert *et al.* 2005) son procesos que se pueden identificar y cuantificar gracias a la morfometría geométrica, ya que esta técnica nos aporta información sobre la dirección del cambio de forma en el espacio y la magnitud del mismo.

4. PROTAGONISTAS: EL TOPILLO MEDITERRÁNEO Y EL TOPILLO LUSITANO

4.1 SUS ORÍGENES

El género *Microtus* es el género con más especies (unas 70 [Fink *et al.* 2010]) de la subfamilia *Arvicolinae* (familia *Cricetidae*, orden *Rodentia*) y uno de los géneros con más número de especies de todos los mamíferos (Nowak 1999). Es un género que ha sufrido una rapidísima radiación, distribuyéndose por todo el territorio Holártico (Jaarola *et al.* 2004). A diferencia de otros grupos animales que han experimentado radiaciones masivas como es el caso de los cíclidos (Albertson *et al.* 1999; Danley & Kocher 2001) o el de los pinzones de las Galápagos (Sato *et al.* 2001), todas las especies del género *Microtus* presentan una morfología bastante similar (Niethammer & Krapp 1982; Musser & Carleton 1993; Mitchell-Jones *et al.* 1999; Nowak 1999). Muchos de los análisis realizados comparando cráneos y dientes de diferentes especies han corroborado que existe una gran variabilidad intraespecífica, un gran número de evidencias de la existencia de procesos de convergencia adaptativa y un gran número de especies hermanas (Zakrzewski 1985; Chaline & Graf 1988; Nadachowski & Zagorodnyuk 1996; Chaline *et al.* 1999; Jaarola *et al.* 2004). Por todos estos motivos este género se presenta como un modelo ideal para estudiar procesos de especiación (Jaarola *et al.* 2004).

Se postula, basándose en registros estratigráficos, que todo el género *Microtus* se originó en la zona del sur de Asia partir de un ancestro común denominado *Allophaiomys pliocaenicus* (Chaline 1966, 1972; Kretzoi 1969; Van der Meulen 1973,

1978) hace unos 2 millones de años (Ma), aunque según nuevos registros fósiles encontrados en China esta cifra podría elevarse hasta los 2,3-2,4 Ma (Zheng & Zhang 2000). Siempre sobre la base del registro fósil, se postula que la dispersión del género por el territorio Holártico se produjo mediante tres eventos independientes de colonización. Uno hacia Asia Central, otro hacia el Norte de Eurasia y otro alrededor de 1.2-2 Ma a través del estrecho de Bering llegando al Norte del continente Americano (Brunet-Lecomte & Chaline 1991; Chaline *et al.* 1999). Por el contrario, y basándose en el uso de marcadores moleculares (AFLPs), se defiende la existencia de varias ondas de colonización a partir del sur de Asia originándose así la actual distribución de las especies (Fink *et al.* 2010) (figura 9).

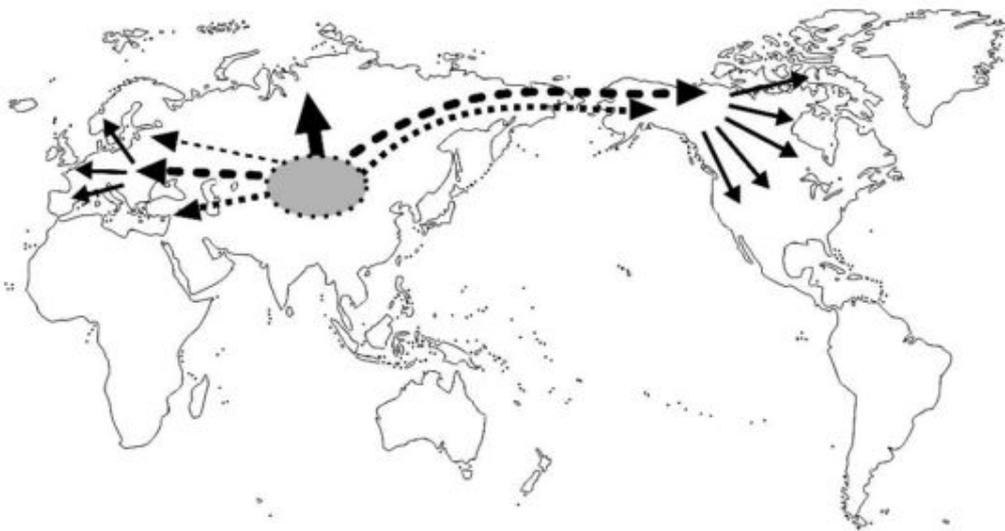


Figura 9. Diferentes eventos de colonización del género *Microtus* hacia Europa y Norteamérica (líneas punteadas) desde un antepasado común situado en el sureste asiático (círculo gris). Las líneas negras representan las sucesivas radiaciones locales. Escenarios reconstruidos a partir de información obtenida de la secuenciación de AFLPs. Modificado Fink *et al.* (2010).

Tanto con evidencias del registro fósil como con evidencias moleculares (citocromo *b* y AFLPs) dentro del clado europeo se ha definido la existencia de dos subgéneros, *Microtus* y *Terricola* (Brunet-Lecomte & Chaline 1990, 1991; Jaarola *et al.* 2004; Jaarola & Searle 2004; Fink *et al.* 2010). El subgénero *Terricola* en un principio fue considerado como un género denominado *Pytimys* que presentaba una distribución tanto Neártica como Paleártica. Las especies que formaban este género se caracterizaban por presentar una estructura en el primer molar inferior, M1, denominada rombo pitimiano. Ésta es una estructura originada por la unión de dos triángulos surgidos de la evolución del complejo anterocoronoide del ancestro común (*Allophaiomys pliocaenicus*) de todo el

género *Microtus* (Chaline 1985). Debido a su distribución disyunta y a que formaba un grupo polifilético a nivel mitocondrial, se decidió separar los dos grupos geográficos incluyendo al clado europeo en el subgénero *Terricola* (Brunet-Lecomte & Chaline 1992; Chaline & Graf 1988; Zagorodnyuk 1989; Jaarola *et al.* 2004).

El subgénero *Terricola* está formado por 15 especies (Chaline *et al.* 1999). A pesar de las variaciones en el cariotipo entre las diferentes especies, existe una gran controversia acerca de la delimitación de las especies dentro del subgénero. A nivel mitocondrial, en el género *Microtus*, se observa que existe un solapamiento entre la divergencia intraespecífica (4-7%) y la interespecífica (4-18%) (Jaarola *et al.* 2004). Es decir, hay especies que presentan una gran variabilidad intraespecífica y especies hermanas con muy pocas diferencias entre sí. Por tanto, el uso de distancias genéticas o la existencia de monofilia recíproca en el ADN mitocondrial no son criterios suficientes para definir los límites entre especies en este género (Bradley & Baker 2001; Nichols 2001; Hudson & Coyne 2002; Rosenberg & Nordborg 2002; Hudson & Turelli 2003).

Microtus duodecimcostatus y *M. lusitanicus*, son los topillos, junto a *M. gerbei*, representantes del subgénero *Terricola* en la Península Ibérica. Dentro de este subgénero, estas dos especies presentan unas características inmejorables para el estudio de procesos de especiación ya que cuentan con un gran número de particularidades. Por un lado, tienen un número cromosómico propio, $2n=62$ (Chaline *et al.* 1999) Además, según los registros estratigráficos parece ser que hayan tenido una evolución casi directa e independiente del resto de los topillos desde el ancestro común del género (Chaline *et al.* 1999) (figura 10) y cuentan con una estructura primitiva en el tercer molar superior, M3, (triángulos 1 y 2 más cerrados y alternantes), muy similar a la del antecesor común (Chaline 1974). Por último, a nivel mitocondrial presentan una de las tasas de divergencia menores de todo el género *Microtus*, entre el 4-5% (Jaarola *et al.* 2004).

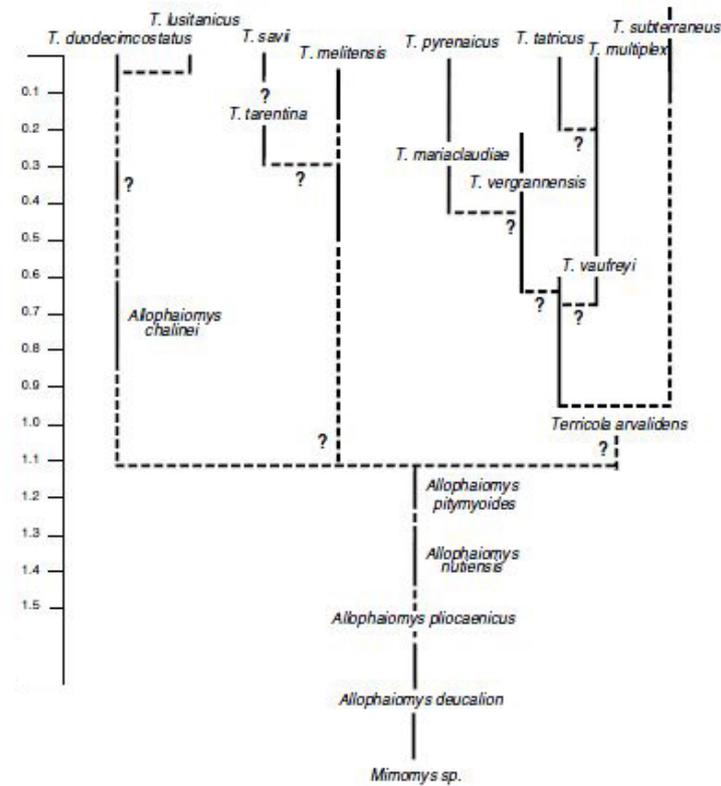


Figura 10. Distribución estratigráfica y posibles relaciones filogenéticas entre el género *Allophaiomys* y el subgénero *Terricola* en Europa. Tiempo representado en millones de años. Modificado de Chaline *et al.* (1999).

4.2 SUS CARACTERÍSTICAS

4.2.1 *Microtus duodecimcostatus* (de Sélys-Longchamps 1839)

El topillo mediterráneo (figura 11A) es un animal de pequeño tamaño, aproximadamente 15 cm, y un peso de entre 15 y 30 gramos (Cotilla & Palomo 2007). Presenta una coloración del pelo parda amarillenta aunque puede variar en función de las localidades. La cola siempre es de tonalidad grisácea (Cotilla & Palomo 2007). Debido a sus hábitos excavadores presenta pequeñas orejas y ojos, pelo corto y flexible. Del mismo modo, su dieta está condicionada por esta actividad excavadora y se alimenta básicamente de bulbos, tubérculos y raíces (Cotilla & Palomo 2007). Presentan actividad sexual casi durante todo el año, aunque existen periodos de mínima actividad (en el norte de la distribución en invierno y en el sur en verano). Su mayor depredador es la lechuza (*Tyto alba*) (Cotilla & Palomo 2007).

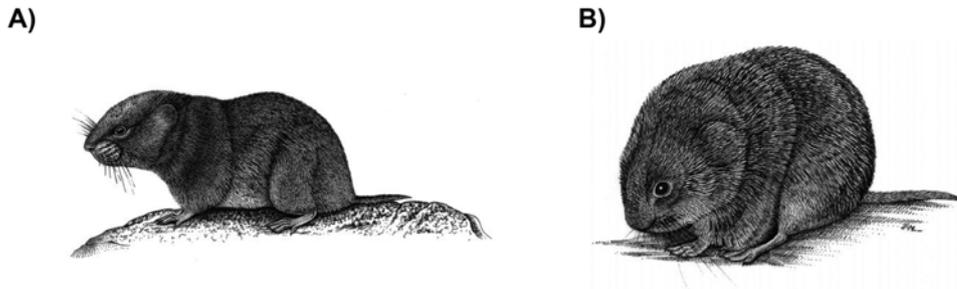


Figura 11. (A) *M. duodecimcostatus*. (B) *M. lusitanicus*. Se observa que *M. duodecimcostatus* presenta una forma craneal más alargada y robusta mientras que *M. lusitanicus* presenta una forma craneal más redondeada y aplastada. Modificado de Cotilla & Palomo (2007); Mira & Mathias (2007).

4.2.2 *Microtus lusitanicus* (Gerbe 1879)

El topillo lusitano (figura 11B) es más pequeño que el topillo mediterráneo, midiendo aproximadamente unos 13 cm y con un peso que oscila entre 14 y 20 gramos (Mira & Mathias 2007). Presenta una coloración corporal predominantemente grisácea y la cola es bicolor, presentando un gris más oscuro en la zona dorsal (Mira & Mathias 2007). Presenta las mismas adaptaciones a hábitos excavadores pero su aspecto todavía es más redondeado en la parte craneal y más cilíndrico en la corporal que el topillo mediterráneo (Mira & Mathias 2007). Su dieta está compuesta de diversos geófitos aunque la pueden alternar con hierbas anuales de pequeño tamaño en ciertas épocas del año (Mira & Mathias 2007). Su actividad reproductora está influenciada por las condiciones climáticas anuales aunque presentan la posibilidad de reproducirse durante todo el año. Sus mayores depredadores son la lechuza y el cárabo (*Strix aluco*) (Mira & Mathias 2007).

4.3 SU HÁBITAT

En la actualidad estas dos especies de topillos se distribuyen por casi todo el territorio de la Península Ibérica y parte de Francia. El topillo lusitano presenta una distribución predominantemente occidental (desde el nivel del mar hasta unos 2000 m), llegando por el sur hasta Setúbal (Portugal), mientras que el topillo mediterráneo se distribuye en la parte oriental de Península y sureste de Francia (desde el nivel del mar hasta unos 3000 m). En la parte centro-septentrional de la Península Ibérica (norte de Navarra, sur del País Vasco, oeste de La Rioja y la mayor parte de Castilla y León, sobre todo, las

provincias de Burgos, Palencia, Valladolid, Segovia, Zamora y Salamanca) las dos especies comparten área de distribución (Cotilla & Palomo 2007; Mira & Mathias 2007) (figura 12).

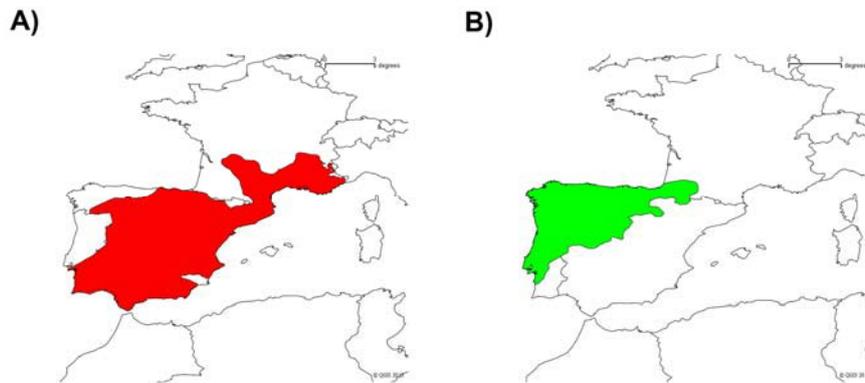


Figura 12. Áreas de distribución de *M. duodecimcostatus* (A) y *M. lusitanicus* (B) en la Península Ibérica. Modificado de Cotilla & Palomo (2007); Mira & Mathias (2007); Aulagnier & Palomo (2008); Aulagnier *et al.* (2008); Santos (2009a).

Los requerimientos ecológicos, en lo que se refiere al hábitat que ocupan, son diferentes para las dos especies tanto a pequeña como a gran escala, aunque son los primeros los que básicamente determinan su distribución (Santos *et al.* 2010). Muchas de las características que conllevan a la elección del hábitat idóneo se basan en la mayor o menor capacidad de excavación, del topillo mediterráneo y del topillo lusitano, respectivamente. Por tanto, el topillo mediterráneo prefiere zonas con cobertura arbórea o zonas de campos de cultivo, un régimen no muy pronunciado de lluvias, alto contenido de arcilla en el suelo que le permita mantener túneles más profundos y suelos luvisoles o litosoles caracterizados por valores intermedios de pH (Santos *et al.* 2009b). El topillo lusitano, debido a su menor capacidad de excavación prefiere zonas a mayor altitud, en los que en los periodos de mayor sequía las condiciones son más favorables —debido a una mayor humedad—, mayor cobertura vegetal pero de formación arbustiva ya que no permite la presencia de predadores potenciales —como la lechuza— y suelos ácidos, con bajo contenido en arcilla para facilitar la excavación de los túneles a menor profundidad (Santos *et al.* 2009b).

Estas dos especies han sido tradicionalmente consideradas como especies con distribución simpátrica, es decir, especies que presentan territorios de distribución solapados de forma completa. Seguramente estas afirmaciones fueron realizadas debido a que la metodología utilizada de forma más habitual para obtener muestras de estas dos

especies ha sido la recolección de egagrópilas de lechuza, (Madureira 1981; Veiga 1986; Brunet-Lecompte *et al.* 1987; Cotilla & Palomo 2007; Rodríguez & Peris 2007) por lo que se ha recogido información de zonas más amplias de lo que se corresponde con el área de distribución de estos topillos (Shawyer 1998; Taylor 2004; Santos *et al.* 2009b). En realidad, sólo se ha descrito la presencia de estas dos especies en la misma área en una zona de los Pirineos españoles con información de presencia recogida mediante el uso de trampas de captura directa (Borghi *et al.* 1994). Por el contrario, otros estudios realizados en la zona de distribución compartida del sur de Portugal han demostrado que la proporción de área de distribución rigurosamente simpátrica es muy reducida, predominando una área de distribución parapátrica. Es decir, estas dos especies presentarían en términos generales rangos de distribución adyacentes pero no solapados (Bull 1991), combinados con una estrecha área de territorio solapado (Santos *et al.* 2009b; Santos *et al.* 2010).

4.4 SUS DIFERENCIAS MORFOLÓGICAS TRADICIONALES

La diferenciación de estas dos especies a nivel morfológico externo es bastante evidente. Como se describe anteriormente, el topillo mediterráneo es de mayor tamaño que el lusitano. La coloración del pelo (más pardo en el topillo mediterráneo y más gris en el lusitano) es otra característica que puede ayudar a la distinción de las dos especies. El problema es que en la mayoría de los casos, la forma más sencilla de obtener muestras de estas dos especies es a partir de egagrópilas de lechuza. De ahí que muchas claves de identificación se basen en diferencias existentes en el cráneo, la mandíbula y los dientes.

En referencia al cráneo, existe un mayor desarrollo del rostro, una mayor diastema (mayor de 7,5 mm) y un mayor palato en el topillo mediterráneo, debido a su mayor actividad excavadora (Madureira 1982). Se presenta una mayor inclinación de los bordes de la constricción post-orbital en los adultos de *M. duodecimcostatus* respecto a *M. lusitanicus* (Madureira 1982) (figura 13A). Los incisivos superiores en el caso de *M. duodecimcostatus* son de tipo prognato, es decir, están inclinados ligeramente hacia delante, mientras que en *M. lusitanicus* son ortognatos, es decir, desde una vista lateral del cráneo se disponen paralelos al rostro (Ellerman & Morris-Scott 1951; Almança 1973; Herrera 1973) (figura 13B).

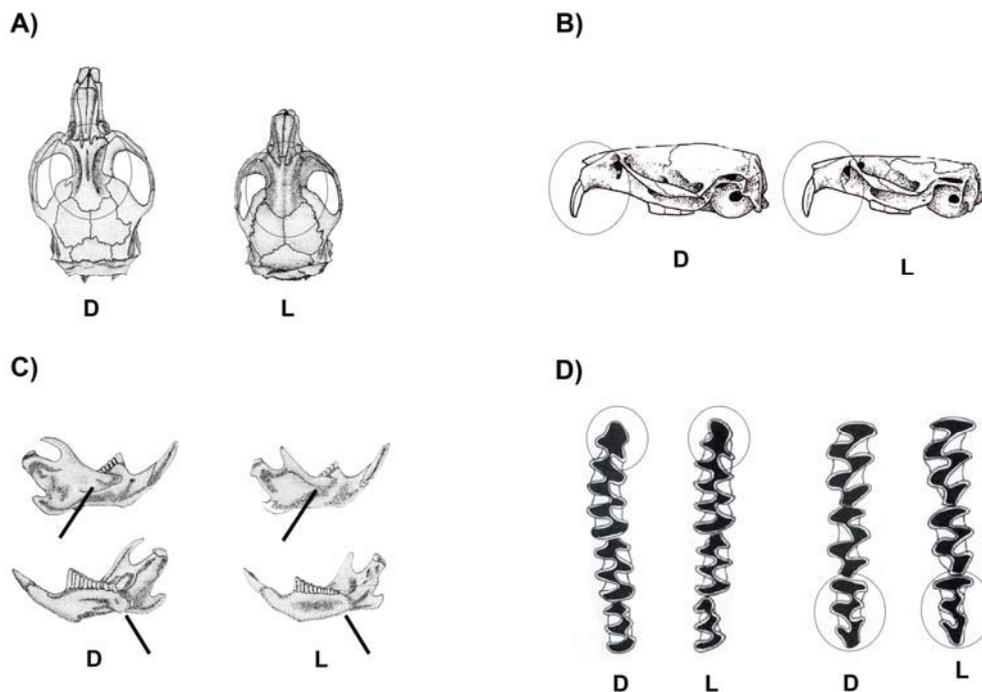


Figura 13. Características morfológicas tradicionalmente usadas para la diferenciación entre *M. duodecimcostatus* (D) y *M. lusitanicus* (L). **(A)** Diferencias en la inclinación de los bordes de la constricción post-orbital. **(B)** Dientes de tipo prognato para *M. duodecimcostatus* y ortognatos para *M. lusitanicus*. **(C)** Un mayor desarrollo de la cresta masetérica (parte superior de la figura) y del proceso articular (parte inferior de la figura) en *M. duodecimcostatus* respecto a *M. lusitanicus*. **(D)** Complejo anterocoronoide más redondeado en *M. duodecimcostatus* (parte izquierda de la figura). Segundo triangulo bucal ligeramente más prominente en *M. duodecimcostatus* (parte derecha de la figura). Modificado de Madureira (1982); Blanco (1998).

Las mandíbulas del topillo mediterráneo son más grandes y robustas en cuanto a la dentición, pero se diferencian respecto a las del topillo lusitano en el mayor desarrollo de la cresta masetérica y del proceso articular (Madureira 1982) (figura 13C).

En numerosas ocasiones la diferenciación entre especies a nivel morfológico en el género *Microtus* se ha estudiado mediante la comparación de las variaciones existentes en los diferentes molares (Ácz *et al.* 2005; Nappi *et al.* 2005; Wallace 2006; Markova *et al.* 2009). Del mismo modo, las diferencias existentes en el primer molar inferior (M1) y en el tercer molar superior (M3) entre el topillo mediterráneo y el topillo lusitano han sido utilizadas como carácter diferenciador entre estas especies. El carácter que discrimina en el M1 es el grado de abertura del complejo anterocoronoide, más cerrado en el topillo lusitano (Brunet-Lecompte *et al.* 1987) (figura 13D). Para el M3 se describe una mayor prominencia —aunque muy sutil— del segundo triángulo bucal

respecto al primero y al tercero, en *M. duodecimcostatus* (Niethammer 1970; Madureira 1982) (figura 13D).

Todos estos ha sido los caracteres utilizados tradicionalmente para la clasificación de estas dos especies de topillos, aunque en todas las guías morfológicas o publicaciones científicas se advierte que el grado de complejidad en la diferenciación aumenta en las áreas de distribución compartidas ya que estos caracteres tienden a igualarse entre las dos especies.

II. OBJETIVOS

1. OBJETIVOS

Establecer las relaciones existentes entre las especies y definir los límites existentes entre las mismas ha sido uno de los objetos principales de estudio de la biología evolutiva durante las últimas décadas. El establecimiento de estos límites se hace más complejo cuanto más cercanas son las especies entre sí debido a que comparten más características, tanto genéticas como morfológicas, que hacen que en algunas ocasiones estos límites sean más difusos, generando una mayor complejidad en la búsqueda de los criterios más precisos y menos sesgados para establecer los límites entre las especies. Bajo este marco conceptual, el objetivo principal de esta tesis doctoral ha sido la búsqueda, sirviéndose de herramientas morfológicas y filogenéticas, de los criterios que explican con un mayor grado de exactitud la relación existente entre dos especies cercanas de topillos, *Microtus duodecimcostatus* y *Microtus lusitanicus*, que se distribuyen a lo largo de la Península Ibérica.

Este objetivo principal se ha desglosado en los siguientes objetivos parciales:

1. Teniendo en cuenta que los roedores componen la mitad de las especies de los mamíferos y que cuentan con un genoma particularmente acelerado respecto al resto de los mamíferos, el primer objetivo fue el desarrollo de un conjunto de marcadores nucleares intrónicos específicos para este grupo y con unas determinadas características (tamaño, copia única, alta divergencia y evolución ajustada a la del resto del genoma) que permitieran usar estos intrones para resolver cuestiones filogenéticas a los niveles taxonómicos más bajos (familias, géneros y especies) dentro del orden *Rodentia*. Una vez obtenido este conjunto de marcadores mediante una serie de filtros bioinformáticos, se escogería un subconjunto de ellos y se comprobaría su validez mediante la amplificación y posterior uso para la reconstrucción filogenética en especies pertenecientes a varias familias representativas del orden *Rodentia*.
2. Uso de la técnica de la morfometría geométrica para determinar las características morfológicas principales que diferencian las dos especies de topillos, *M. duodecimcostatus* y *M. lusitanicus*, en las zonas de distribución

alopátrica. Comparación de estos nuevos caracteres obtenidos con los utilizados tradicionalmente para su diferenciación. Posteriormente se pretendía, mediante estos caracteres, la asignación de especie a los individuos de las zonas de distribución solapada de las dos especies de topillos. Evaluación de si en estas zonas de distribución solapada están ocurriendo procesos de adaptación fenotípica, desplazamiento de caracteres o convergencia de caracteres a nivel de forma, entre las dos especies del estudio.

3. Caracterización de las dos especies de topillos ibéricos a nivel filogenético mediante el uso de un marcador mitocondrial, citocromo *b*, y de seis intrones de los desarrollados en el primer objetivo de la tesis. Datación del tiempo de especiación usando metodologías de reconstrucción de árboles de especies y teniendo en cuenta la posible existencia de procesos de introgresión o separación incompleta de linajes entre las dos especies. Asimismo, se evaluaría cómo los procesos de introgresión o separación incompleta de linajes podrían influir en la correcta descripción de los límites entre dos especies cercanas.

III. MATERIALES Y MÉTODOS

1. DESARROLLO DE NUEVOS MARCADORES NUCLEARES PARA EL ORDEN *Rodentia*

1.1 OBTENCIÓN Y FILTRADO DE INTRONES DE ROEDORES

1.1.1 Obtención de los genomas

Se obtuvieron de la base de datos de Ensembl (Flicek *et al.* 2012) los genomas disponibles de roedores (en formato GenBank), que tuvieran una buena cobertura y que estuvieran bien anotados. En el momento del estudio se contó con tres genomas que cumplieran estas características: ratón común (*Mus musculus* versión NCBIIm37) (Mouse Genome Sequencing Consortium 2002), rata parda (*Rattus norvegicus* versión RGSC 3.4) (Rat Genome Sequencing Project Consortium 2004) y conejillo de indias (*Cavia porcellus* versión cavPor3) (Flicek *et al.* 2012). Asimismo se obtuvo el genoma de humano (*Homo sapiens* versión GRch37 [International Human Genome Sequencing Consortium 2001; Venter *et al.* 2001]) para usarlo como grupo externo en los árboles de roedores. A partir de estos genomas se seleccionaron los intrones (junto con sus exones correspondientes) que tenían una longitud menor de 50.000 pares de bases. Una serie de scripts en Perl fueron diseñados para extraer estas secuencias y para procesarlas en los pasos sucesivos.

1.1.2 Filtros bioinformáticos

Se aplicaron una serie de filtros bioinformáticos con el objetivo de obtener un conjunto de intrones que cumpliera una serie de características. En primer lugar se obtuvieron las relaciones de ortología 1:1 existentes entre los genes de las cuatro especies. Esta información se extrajo de la base de datos BioMart (Smedley *et al.* 2009). Para obtener una lista de genes ortólogos para las cuatro especies, se hicieron listas intermedias de dos especies basadas en la proximidad filogenética. Así, primero se compararon rata y ratón. La lista resultante, usando ratón como referencia, se comparó con la información para el conejillo de indias y, a su vez, este resultado se comparó con la información de humano. De este conjunto de genes se eliminaron todos aquellos que no tenían el mismo

número de intrones para las cuatro especies. A partir de estos genes, se obtuvieron todos los intrones y los exones que los forman.

A continuación, se seleccionaron los intrones con un tamaño razonable para su amplificación por PCR. Para ello se estableció una longitud, basándose en la información del ratón, de entre 200 y 1.600 pares de bases. Además, para evitar la selección de intrones con longitudes muy dispares entre especies y que posteriormente pudieran ocasionar problemas de alineamiento, se controló la diferencia de longitud entre especies. De esta manera, se seleccionaron aquellos intrones que tenían una diferencia en longitud menor del 90% entre humano y el resto de las especies; del 80% en el caso de las diferencias entre el conejillo de indias y el resto de especies; y del 70% entre ratón y rata.

Debido a algunas malas anotaciones del genoma y para evitar problemas posteriores, se eliminaron los intrones con anotaciones redundantes y aquellos que presentaban más de 9 Ns (entendido como una base ambigua) en alguna de las cuatro especies.

Para asegurar que hubiera espacio suficiente para el diseño de los cebadores se seleccionaron aquellos intrones que estaban flanqueados por exones con una longitud mayor de 40 nucleótidos.

Para evitar bandas múltiples en el posterior proceso de amplificación por PCR se eliminaron los intrones pertenecientes a genes con más de una copia en el genoma. Para ello, se utilizaron los exones flanqueantes en una búsqueda mediante el algoritmo Blast (Altschul *et al.* 1997) para cada especie dentro de su propio genoma. El umbral de la búsqueda se estableció en un *e-value* de 10^{-4} .

Los últimos filtros que se aplicaron fueron de carácter filogenético. Primero se alinearon los intrones de cada grupo ortólogo (para las cuatro especies) con Mafft versión 6.708 (Katoh & Toh 2008). Se aplicó a los alineamientos el programa Gblocks versión 0.91 (Castresana 2000) con parámetros relajados (Talavera & Castresana 2007) para eliminar las posiciones mal alineadas. Para cada alineamiento, se reconstruyó un árbol filogenético de máxima verosimilitud usando RaxML versión 7.0.4 con un modelo de substitución nucleotídica GTR y una distribución gamma para las tasas evolutivas

(Stamatakis 2006). Además, se concatenaron todos los intrones alineados y se obtuvo un árbol filogenético de máxima verosimilitud bajo las mismas condiciones anteriores para usar como árbol de referencia. Este árbol de referencia se comparó con cada uno de los árboles individuales para calcular el factor escalador (*scaling factor*) y el índice de K del árbol (*K tree score*), usando el software Ktreedist (Soria-Carrasco *et al.* 2007). El *scaling factor* indica la tasa global relativa que presenta un árbol dado respecto al de referencia, es decir, es el valor por el que se deberían multiplicar las ramas de un árbol dado para igualar lo más posible su longitud a la longitud de las ramas del árbol de referencia. Así, valores de *scaling factor* menores de 1 corresponden a árboles más divergentes que el árbol de referencia. El *K score* refleja diferencias topológicas y diferencias relativas en longitudes de ramas entre un árbol dado y un árbol de referencia (y, así, detecta desviaciones del reloj molecular entre ramas). Un *K score* alto refleja que las tasas evolutivas son muy dispares entre los taxa que se analizan. Esta situación puede causar problemas a la hora de estimar tasas evolutivas (usando calibraciones fósiles) incluso aplicando relojes moleculares relajados no correlacionados (Drummond *et al.* 2006). Teniendo en cuenta lo que cuantifica cada uno de estos índices se seleccionaron aquellos intrones con mayor divergencia (*scaling factor* < 0,8) y con menos variación en el reloj molecular del árbol respecto al árbol de referencia (*K score* < 0,09). Después de la aplicación de todos estos filtros, se consideró que los intrones resultantes eran adecuados para su aplicación en distintas especies de roedores.

1.2 VALIDACIÓN EXPERIMENTAL DE LOS INTRONES SELECCIONADOS

A partir del set final de intrones se seleccionaron intrones con una longitud para ratón de entre 300 y 500 nucleótidos. Después de una inspección visual para controlar que existieran regiones apropiadas en los exones para poder diseñar los cebadores, se escogieron 10 intrones que se utilizaron para su amplificación por PCR usando muestras de diferentes roedores

1.2.1 Diseño de cebadores

Con el objetivo de tener un conjunto de cebadores que reconocieran las regiones exónicas en diferentes especies del orden *Rodentia*, se usaron para el diseño de los cebadores los genomas de las tres especies de roedores usadas hasta este momento

(ratón, rata y conejillo de indias) más otras secuencias de roedores y lagomorfos. Estas nuevas secuencias se obtuvieron de la base datos Ensembl (*Dipodomys ordii*, versión dipOrd1; *Ictidomys tridecemlineatus*, versión speTri1; *Oryctolagus cuniculus*, versión OryCun 2 y *Ochotona princeps*, versión OchPri 2.0). Debido a que estos genomas no están tan bien anotados como los genomas usados inicialmente, no en todos los casos fue posible obtener las secuencias exónicas. Todas las secuencias de exones de cada uno de los intrones que se consiguieron se alinearon y se diseñaron los cebadores en los lugares más conservados. Se intentó que la longitud de los cebadores estuviera entre 20 y 25 nucleótidos, que la temperatura de hibridación fuera parecida para cada pareja de cebadores y que hubiera un número máximo de 48 degeneraciones (tabla 1).

Tabla 1. Diez intrones que fueron amplificados en diferentes especies del género *Rodentia*. El nombre del intrón esta formado por la abreviatura del gen al que pertenece y el número del intrón. La longitud, en pares de bases, se refiere a la que presentan en *Mus musculus*. Los cebadores fueron diseñados específicamente para la amplificación de estos intrones.

Nombre intrón	Código Ensembl	Longitud (ratón)	Secuencia cebadores (forward y reverse)
Abcb9-intrón-2	ENSMUSG00000029408	423	GCATYGTSATCCAGAARAGCAYGGA CTGTGCGRTTCTCRTCRAARAAGCT
Agxt-intrón-10	ENSMUSG00000026272	420	GGCTACAACCTGGAGGGACATC TGCAGGGCCTCCYTCAGGGCCT
Catsper3-intrón-5	ENSMUSG00000021499	382	TGCTKGCMTCTTTCATCTT AGRATYAYYTGCTTCTYCTCC
Dhcr24-intrón-7	ENSMUSG00000034926	357	CAGGACATGCTGGTGCCCATGAA CCTGGCTGGCTGGGCAGGATGAA
Ivd-intrón-8	ENSMUSG00000027332	478	CTGGACCTRGARCGCTGGT CTGRAAKTGSCCRATYTTCT
Nadsyn1-intrón-4	ENSMUSG00000031090	498	GTYCGYTACAAYTGCAGAGT TCCTKSHCCAKGGGGTRAACCA
Rras-intrón-4	ENSMUSG00000038387	488	ACWCAGATCCTCMGRGYAAGGA AGTTTGGCDGAKGCCTCRAAGTA
Smo-intrón-9	ENSMUSG00000001761	344	GCCACCCTGCTCATCTGGAGGCG TTGGCRATCATCTTGCTYTTCTTGA
Trpv4-intrón-8	ENSMUSG00000014158	398	TTACCRBACCACVGYGGACTACCT CTGGAAGGAGCCRTCGAYGAAGA
Wls-intrón-7	ENSMUSG00000028173	364	AAYCACATYGCMGGSTAYTGGA TCYGTKCCAACRTCYGTRGTCCA

1.2.2 Extracción de ADN genómico

Se recopilaron muestras de especies pertenecientes a nueve familias de roedores de diferentes instituciones (Museo de Ciencias Naturales de Nueva York y Banco de tejidos de Cataluña) y de la colección propia (tabla A1). Todas ellas estaban conservadas en etanol a -20° C. Se extrajo el ADN genómico a partir de unos 25 mg de tejido, usando el *kit DNeasy Blood & Tissue* de Qiagen siguiendo las instrucciones del fabricante.

1.2.3 Amplificación por PCR

El volumen final de las reacciones de amplificación fue de 25 µl. Las concentraciones de ADN obtenidas fueron muy diferentes dependiendo de la eficacia de las extracciones pero se intentó disponer de 50 ng. Se usó 1 µM de concentración final para cada cebador, 0,2 mM de dNTPs, 0,75 unidades de Promega *GoTaq DNA* polimerasa y 17,5 µg de BSA (*bovine serum albumin*). La PCR se llevo a cabo con el método *touch-down* (Korbie & Mattick 2008) siguiendo las siguientes condiciones: una desnaturalización inicial a 95° C durante 3 minutos seguida de 15 ciclos formados por una desnaturalización inicial (30 segundos a 95° C), una hibridación (30 segundos empezando a 65° C) y una extensión (30 segundos a 72° C). En cada ciclo la temperatura de hibridación bajaba entre 0,33 y 1° C hasta alcanzar la temperatura de hibridación del segundo conjunto de ciclos. Este segundo conjunto, de 20 ciclos, se realizaba con una temperatura de hibridación de 60, 55 o 50° C. El resto de los pasos se mantuvo igual que el primer conjunto de ciclos. Al final se añadió una extensión durante 10 minutos a 72° C.

Los productos de amplificación obtenidos fueron cargados en geles de agarosa al 1% preteñidos con *SYBR Safe DNA Gel Stain* (Invitrogen) para comprobar si la amplificación había ocurrido con éxito.

Se probó, de forma sistemática, una *touch-down* PCR con temperaturas de hibridación de 65-55° C para todas las muestras de ADN genómico obtenidas y para cada marcador. Si el resultado era la aparición de bandas múltiples, se realizaba una PCR más

restrictiva, a 65-60° C. Por el contrario, si no se había obtenido amplificación, se realizaba una PCR menos restrictiva, a 65-50° C.

Las amplificaciones que resultaron positivas fueron purificadas con ExoSAP-It (Affymetrix) para su posterior envío al servicio de secuenciación de Macrogen Inc. (Seoul, Corea del Sur). Se utilizaron para la secuenciación los mismos cebadores originales de la amplificación. Las secuencias finales fueron inspeccionadas, recortadas y ensambladas con el programa Geneious Pro v4 (Drummond *et al.* 2009) (Biomatters Ltd.).

1.2.4 Reconstrucción filogenética

Las secuencias obtenidas se alinearon junto con las secuencias de ratón, rata, conejillo de indias y humano usando Mafft versión 6.708 (Kato & Toh 2008). Se eliminaron las regiones mal alineadas usando Gblocks con parámetros relajados (Castresana 2000; Talavera & Castresana 2007). Para cada intrón se realizó una reconstrucción filogenética basada en máxima verosimilitud usando RaxML con los parámetros descritos anteriormente (Stamatakis 2006). Se reconstruyó asimismo un árbol de máxima verosimilitud con todos los intrones concatenados.

Además se realizó una reconstrucción filogenética siguiendo un método Bayesiano usando Beast versión 1.6.1 (Drummond & Rambaut 2007). Cada uno de los 10 intrones se introdujo como una partición independiente. Se eligió el modelo de sustitución más adecuado mediante jModeltest (Posada 2008) (tabla A2). Con el objetivo de implementar el reloj molecular más apropiado (estricto o relajado lognormal no correlacionado) se estimó con PAUP* versión 4.0b10 (Swofford 2003) la verosimilitud de una topología de PhyML (Guindon & Gascuel 2003) con y sin reloj molecular. Se escogió el reloj más apropiado con un test de tasas de verosimilitud (Felsenstein 1981) (tabla A2). Como prior de especiación se usó un modelo *Yule*. Se corrió un total de 50 millones de generaciones, se descartaron a modo de *burn-in* el 10% inicial de las muestras generadas y se comprobó la correcta convergencia de todos los parámetros usando la utilidad de Beast Tracer, asegurando que todos los tamaños de muestra efectivos fueran superiores a 200. Mediante la utilidad TreeAnnotator se calculó el árbol de máxima credibilidad de clados usando las medianas de las longitudes de ramas.

2. ANÁLISIS MORFOMÉTRICO DEL TOPILLO MEDITERRÁNEO Y DEL TOPILLO LUSITANO EN SUS DIFERENTES ÁREAS DE DISTRIBUCIÓN

2.1 OBTENCIÓN DE MUESTRAS Y LOCALIDADES DEL ESTUDIO

2.1.1 Obtención y procesado de las muestras

Con el objetivo de obtener cráneos de los topillos ibéricos se realizó una búsqueda de egagrópilas de lechuza (*Tyto alba*) por gran parte del territorio de la Península Ibérica, intentando abarcar en la medida de lo posible la distribución de las dos especies de topillos. Se recolectó un total de 664 egagrópilas procedentes de 15 localidades (figura 14). De cada una de ellas se extrajeron todos los individuos encontrados y se identificaron al menos hasta el nivel de género.

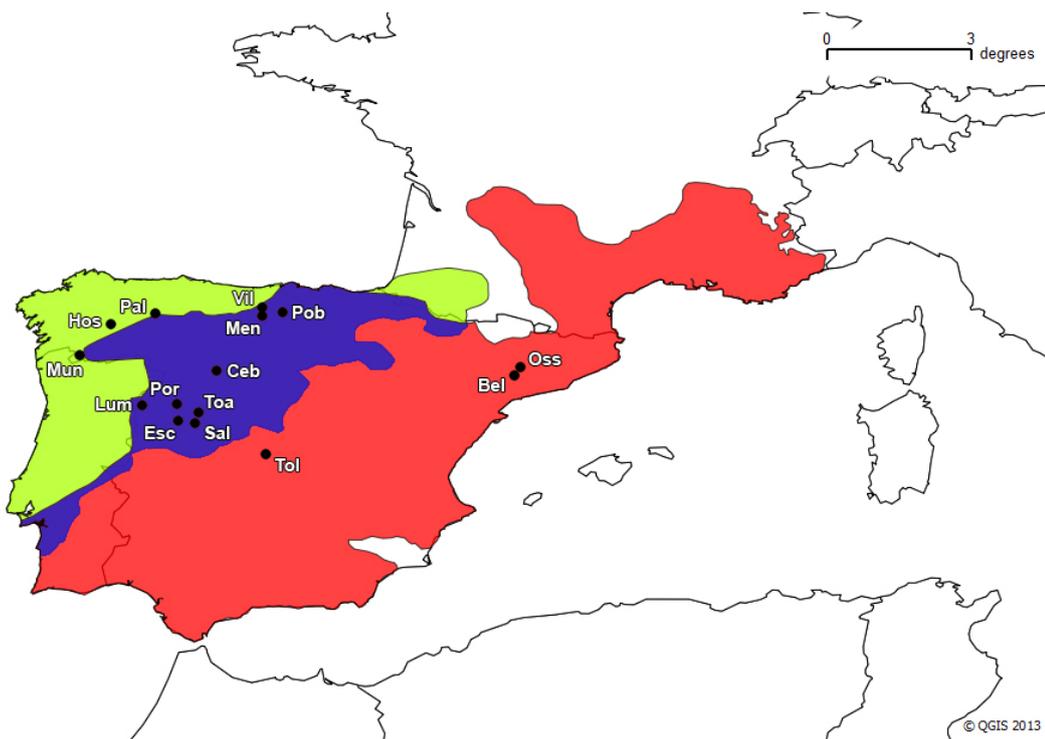


Figura 14. Puntos geográficos de las localidades de muestreo. Áreas de distribución de *M. duodecimcostatus* (rojo), *M. lusitanicus* (verde) y zonas de distribución solapada (azul). **Bel:** Belianes; **Ceb:** San Cebrián de Mazote; **Esc:** Escorial de la Sierra; **Hos:** Hospital de Incio; **Lum:** Lumbrales; **Men:** Menaza; **Mun:** Muñíos; **Oss:** Ossó de Sió; **Pal:** Palacios de Sil; **Pob:** Poblaciones de Arreba; **Por:** Porqueriza; **Sal:** Salvatierra de Tormes; **Toa:** Torrejón de Alba; **Tol:** Toledo y **Vil:** Villacantid. Modificado de Cotilla & Palomo (2007); Mira & Mathias (2007); Aulagnier & Palomo (2008); Aulagnier *et al.* (2008); Santos (2009a).

Para las dos especies de topillos, se identificaron como individuos independientes todos los cráneos encontrados y, en el caso de las mandíbulas, sólo se conservaron aquellas que se encontraron unidas físicamente al cráneo. De este modo se intentó prevenir la mezcla de material biológico de individuos diferentes y la sobre-estimación de individuos.

Se obtuvo un total de 179 individuos, 60 de las zonas de distribución alopátrica y 119 de las zonas de distribución solapada (tabla A3). Se considera que este método de muestreo recoge la variabilidad de individuos real de las dos especies, ya que las lechuzas no presentan preferencias por las presas (Mikkola 1983). Se asumió, por tanto, que se contaba con información tanto de individuos adultos como de juveniles, y tanto de machos como de hembras.

2.1.2 Localidades de muestreo

La distribución de los individuos en las localidades elegidas está sujeta al método de muestreo. Es decir, ya que las muestras de este estudio se obtuvieron a partir de egagrópilas de lechuza, los individuos pueden pertenecer a una zona de entre uno y unos veinticinco kilómetros del punto donde fueron recolectadas. Esta distribución se corresponde con el área media de caza de una lechuza y, además, varía en función de que sea o no la época de cría de la lechuza (Shawyer 1998; Taylor 2004). Por tanto, cada punto de muestreo puede recoger información de varias poblaciones de topillos debido a que los territorios de las dos especies son más pequeños que el área de caza de una lechuza (estimado en unos 50 m² para *M. duodecimcostatus* [Guédon 1992] y 5,5 – 3.063 m², para *M. lusitanicus* [Santos *et al.* 2010]).

Los ambientes de muestreo variaron desde campos abiertos cerealistas hasta pequeños valles un poco boscosos con prados de media montaña. Es decir, se procuró recolectar muestras en los ecosistemas propios de la distribución de las dos especies, haciendo un mayor esfuerzo de muestro en las descritas como zonas de distribución solapada de las dos especies (figura 14). Así, dentro la zona de distribución alopátrica se muestreó en cuatro localidades diferentes y en la zona de distribución solapada en once localidades diferentes (figura 14).

2.2 MORFOMETRÍA GEOMÉTRICA

2.2.1 Material fotográfico

Para cada uno de los individuos se realizaron cinco fotos de diferentes estructuras para obtener información relativa a la forma (figura 15). Así, se obtuvo, del cráneo, una fotografía de la vista lateral izquierda (L) y otra de la parte ventral (V); de la mandíbula, una fotografía de la parte interna de la mandíbula izquierda (J); y de los dientes, una fotografía del primer molar inferior izquierdo (M1) y otra del tercer molar superior izquierdo (M3).

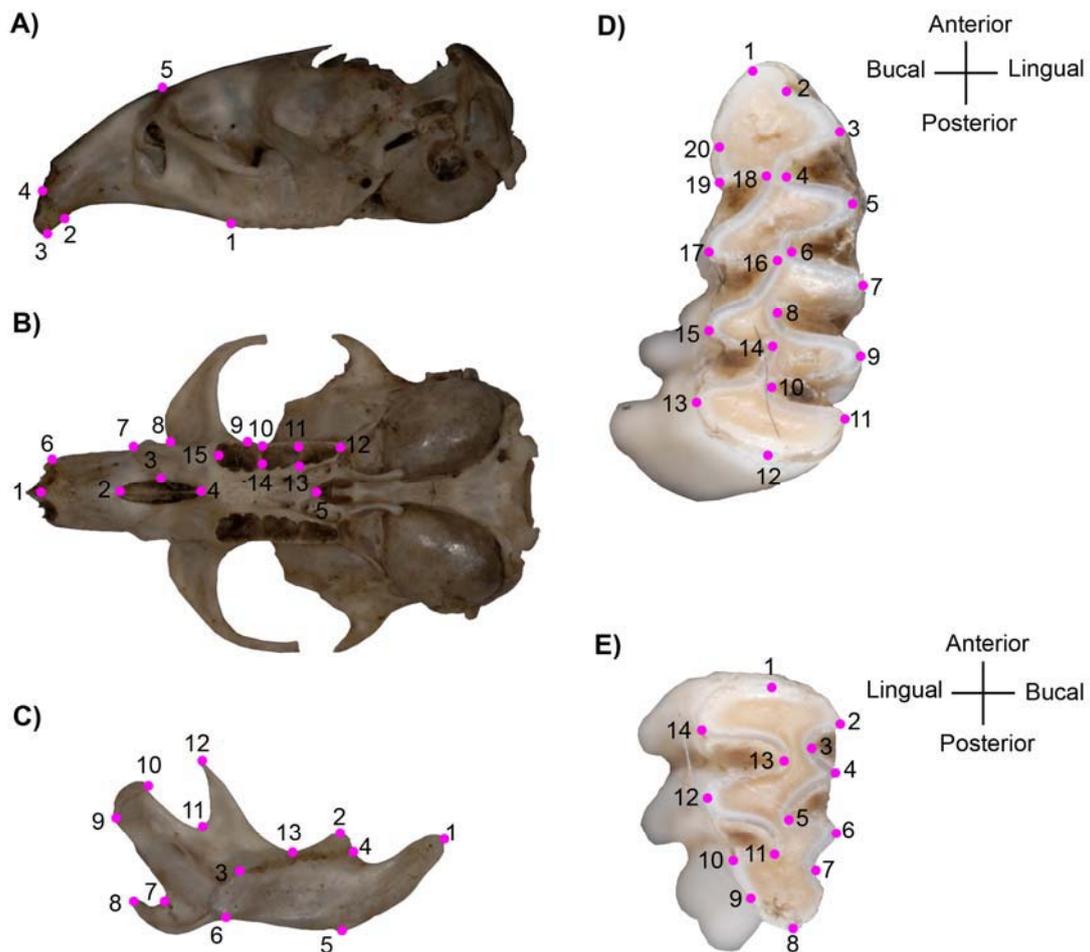


Figura 15. Selección de *landmarks* para cinco caracteres. **(A)** Vista lateral del cráneo (L), 5 *landmarks*. **(B)** Vista ventral del cráneo (V), 15 *landmarks*. **(C)** Vista interna de la mandíbula izquierda (J), 13 *landmarks*. **(D)** Primer molar inferior izquierdo (M1), 20 *landmarks*. **(E)** Tercer molar superior izquierdo (M3), 14 *landmarks*.

Para la mandíbula y los dientes, cuando no fue posible obtener la foto del lado seleccionado (porque estaba roto o porque no se encontró), se realizó la fotografía del otro lado (en el caso de que estuviera disponible) y se obtuvo una imagen espejada para su posterior análisis. No fue posible obtener fotografías de los cinco caracteres en todos los individuos, ya que no se recuperaron todas las estructuras para todos los individuos. Se empleó una cámara fotográfica Nikon D200, un objetivo de 60 mm con una velocidad de apertura de 1/250 segundos y F.22 de apertura de diafragma. La distancia focal para V, L y J fue diferente en cada foto y posteriormente se estandarizó. Para el caso de los dientes la distancia focal era fija.

2.2.2 Selección y tratamiento de landmarks

Entre cinco y veinte *landmarks* para cada carácter (L: n=5; V: n=15; J: n=13; M1: n=20; y M3: n=14) fueron digitalizados en 179 individuos (L: n=175, V: n=179, J: n=137, M1: n=166, y M3: n=177) usando el software tpsDig2 (Rohlf 2005a) (figura 15, tabla A4).

No para todos los caracteres fue posible la identificación de los correspondientes *landmarks*, ya que las estructuras estaban dañadas. Teniendo en cuenta que no se encontraron todas las estructuras para todos los individuos y que algunas estaban dañadas, se obtuvo un conjunto de 128 individuos para los que se contaba con información de todos los caracteres. Para el resto, se utilizó la información de la que se disponía en los análisis sucesivos (tabla A3).

Por medio del software tpsRelw (Rohlf 2005b) las configuraciones de los *landmarks* fueron escaladas, rotadas y trasladadas usando un análisis generalizado de *Procrustes* (GPA, Rohlf & Slice 1990). Se estimó el tamaño de cada estructura usando el centroide (CS), que fue posteriormente transformado con el logaritmo para su normalización (logCS). Se obtuvo igualmente una matriz de forma (Y) incorporando los componentes uniformes (2 para cada estructura) y los no uniformes (L=4, V=24, J=20, M1=34, M3=22). Todos estos componentes se utilizaron como variables de forma. La información que contienen se obtuvo mediante un análisis de componentes principales (en este caso, denominado análisis relativo de deformación, RW).

2.3 PROTOCOLO PARA LA ASIGNACIÓN DE ESPECIE

El análisis de morfometría geométrica se dividió en dos fases según el conjunto de individuos que fueron utilizados. En primer lugar se efectuó un análisis con los 60 individuos que provenían de las zonas de distribución alopatrica. En segundo lugar se analizaron los 119 individuos de las zonas de distribución solapada de las dos especies junto con los 60 individuos de las zonas de distribución alopatrica (figura 14).

2.3.1 Asignación de especie a los individuos de las zonas de distribución alopatrica

Aunque por la distribución geográfica de los individuos se les podría haber asignado una de las dos especies, se realizó un análisis de asignación de especie sin información *a priori*. Se empleó la información del logCS y de la matriz de forma (Y) para estimar la densidad *Kernel* (KDE) (Wand & Jones 1995). Los KDE son estimadores no paramétricos de la distribución de las variables. Si existen diferentes grupos en el conjunto de datos se deberían observar diferentes agrupaciones de densidad. Para el logCS se llevo a cabo un KDE en una dimensión y para la matriz de forma en dos dimensiones.

Con el objetivo de corroborar que la asignación de especie por distribución geográfica era correcta se efectuó un análisis discriminante (LDA) para cada uno de los cinco caracteres. El modelo para cada uno de los caracteres fue $\text{discr.especie-lda}(Y, \text{especie})$, en el que se discrimina la especie en función de la información contenida en la matriz de forma (Y) y se compara con la información *a priori* sobre la especie. Para testar el poder de clasificación de la función discriminante (LDA) y ya que el tamaño muestral no era muy elevado, se repitió el mismo modelo pero con un procedimiento *leave-one-out* (LDA *leave-one out*).

2.3.2 Asignación de especie a los individuos de las zonas de distribución solapada

Del mismo modo que en el apartado anterior se realizó un análisis de asignación de especie sin información *a priori* a cerca de la especie, tanto para el logCS como para la matriz de forma. En este caso se incluyeron los 179 individuos disponibles en el estudio.

Igualmente, se llevo a cabo un análisis discriminante como en el apartado anterior pero con una serie de modificaciones. En este caso, y con el objetivo de utilizar la información de los cinco caracteres al mismo tiempo —para poder tener una mayor potencia en la asignación de la especie—, en vez de utilizar los componentes uniformes y no-uniformes de cada uno de los caracteres, se utilizó la información del primer componente principal y del segundo (RW1 y RW2). Para cada uno de los caracteres se realizó un análisis de deformación relativo y se recogió la información, para cada individuo, del RW1 y del RW2. Con estas nuevas variables se construyó una nueva matriz de forma. Como el análisis discriminante necesita que exista información para todos los individuos y para todos los caracteres, se calculó la media de RW1 y RW2 (de los 5 caracteres) de los individuos que presentaban esta información y de este modo se cubrió la falta de información de algunos individuos (tabla A5). Al asignar la media a un carácter se consigue que este valor no compute en la asignación de una u otra especie pero permite que se pueda usar la información de los otros caracteres (Quinn & Keough 2002). Con toda esta información se construyó el análisis discriminante siguiendo el siguiente modelo $\text{discr.especie-lda}(Y, \text{especie})$, donde se asigna la especie en función de la información recogida en la matriz de forma (Y) y la información *a priori* sobre la especie. Este LDA se realizó con la información correspondiente a 32 individuos de las zonas de distribución alopátrica (tabla A3), de los que se poseía información de los cinco caracteres y de los que se conocía su especie (individuos de referencia). Para el mismo conjunto de individuos se realizó un LDA *leave-one-out*. Posteriormente, con la información del LDA se llevó a cabo la predicción de la especie del resto de individuos siguiendo el siguiente modelo: $\text{predict}(\text{discr.especie}, Y_n)$. El subíndice n indica que la información introducida pertenece, por un lado, a los individuos de las zonas de distribución solapada (que eran el principal objeto de clasificación en este apartado) y por el otro los individuos de las zonas de distribución alopátrica a los que les faltaba información para alguno de los caracteres. Estos últimos individuos fueron incluidos en esta predicción de especie con el propósito de validar, por un lado, el método de clasificación y, por el otro, comprobar la legitimidad de introducir la media en las situaciones de falta de información. A continuación se recopiló el coeficiente de discriminación *a posteriori* (que es un valor presente en el archivo de salida del LDA) para cada individuo, con el fin de comprobar cuál era la robustez en la asignación de la especie para cada caso. Este coeficiente presenta valores de 0 a 1, indicando 1 la máxima robustez en la pertenecía a una de las dos especies analizadas.

Seguidamente al LDA, y para comprobar la validez de la clasificación de especies obtenidas, se llevo a cabo un análisis de agrupación de individuos basados en algoritmos E-M (*expectation-maximization*). Ésta es una metodología que combina el uso de modelos basados en la clasificación jerárquica aglomerativa (es decir, un individuo se irá uniendo a un grupo de individuos ya existente) y el algoritmo E-M para la estimación de la máxima verosimilitud de modelos mixtos multivariantes (Banfield & Raftery 1993). El procedimiento E-M se calcula sobre 6 modelos de distribución diferentes. Las características de estos 6 modelos varían en función de los parámetros de volumen, forma y orientación de los datos. La elección entre modelos se realiza mediante el índice BIC (*Bayesian Information Criterion*) (Kass & Raftery 1995), cuyo resultado indica cuál es el mejor modelo que explica la información introducida y cuál es el número más óptimo de agrupación de los datos introducidos en el modelo. Por tanto, en primer lugar y sin ninguna información acerca de la especie, se calculo cuál era el mejor modelo y cuáles eran los grupos más probables. Como fichero de entrada se utilizó la matriz de forma creada para el LDA (con la información completa para los 179 individuos). Se realizó una representación gráfica de los resultados del análisis para cada pareja de RW (RW1 y RW2) de cada carácter y se recogieron los coeficientes de agrupación posterior para cada uno de los individuos (coeficientes que tienen el mismo significado que los comentados anteriormente para el LDA). Posteriormente se calculó cuáles eran las discrepancias entre la asignación de especie por el LDA y la clasificación en grupos del análisis E-M. Por último, se calcularon cuáles eran las zonas en el espacio forma, para cada uno de los caracteres, donde la probabilidad de realizar una asignación de un individuo a un grupo (en este caso, a una especie) de forma errónea era más probable.

2.4 ANÁLISIS ESTADÍSTICO DE LOS DATOS

Todos los análisis llevados a cabo a partir de este punto se realizaron teniendo en cuenta la asignación de especie para cada individuo obtenida en el apartado anterior mediante el análisis discriminante.

2.4.1 Análisis estadístico de los datos de las zonas de distribución alopátrica

Se evaluaron las variaciones en tamaño y forma de manera interespecífica e intraespecífica en función de diferentes factores mediante diversos tests estadísticos. Se realizó un análisis de la varianza (ANOVA) con el modelo $\log CS \sim sp/pob$, en el que se evalúa la variación del $\log CS$ (variable dependiente) en función de la variación interespecífica (sp) o intraespecífica (sp/pob , población anidado en especie). Se llevó a cabo un análisis multivariante de la varianza (MANOVA, test=Pillai) según el modelo $Y \sim \log CS + sp/pob$, donde se analiza la variación de la forma (Y) en función del $\log CS$ (para controlar los posibles efectos alométricos), de la variación interespecífica (sp) y la variación intraespecífica (sp/pob).

2.4.2 Análisis estadístico de los datos de las zonas de distribución solapada

Se llevaron a cabo los mismos análisis estadísticos que en el apartado anterior con pequeñas variaciones de los modelos. Para el $\log CS$ el modelo testado fue $\log CS \sim sp * geo$, donde geo es el factor que determina si los individuos (de cada especie por separado) pertenecen a la zonas de distribución alopátrica o las zonas de distribución solapada. Se evaluó, por un lado, la variación interespecífica del tamaño (sp) y por otro la variación intraespecífica ($sp * geo$), en relación a la distribución geográfica de los individuos.

El modelo aplicado para analizar la variación de la forma fue $Y \sim \log CS + sp * geo$. Por tanto, se controló la variación interespecífica de la forma (sp) y usando la especie y la distribución geográfica como factores que interaccionan, se pretendía evaluar si existían modificaciones intraespecíficas en forma debido a la coexistencia de las dos especies en la misma aérea geográfica. El $\log CS$, como en el apartado anterior, controlaba la existencia de procesos alométricos.

Paralelamente, se realizó una comparación del cambio de forma medido a través de la cuantificación del desplazamiento de la forma en el espacio. Este cambio de forma se midió a través de vectores en tantas dimensiones como componentes no uniformes de cada uno de los caracteres. Para cada uno de los vectores se cuantificó el cambio en su módulo, en función de las variaciones de forma interespecíficas o intraespecíficas, entre

las zonas de distribución alopátrica y en las zonas de distribución solapada, siguiendo la metodología descrita previamente (Collyer & Admas 2007). Los modelos que se testaron fueron: DA-DS (*M. duodecimcostatus* en zonas de distribución alopátrica, DA, contra *M. duodecimcostatus* en zonas de distribución solapada, DS); LA-LS (el mismo modelo que el anterior pero para *M. lusitanicus*); DA-LA (*M. duodecimcostatus* en zonas de distribución alopátrica contra *M. lusitanicus* en zonas de distribución alopátrica, LA) y DS-LS (*M. duodecimcostatus*, DS, contra *M. lusitanicus*, LS, de las zonas de distribución solapada).

2.4.3 Análisis morfométrico del tercer molar superior (M3)

Para el caso del M3, se cuantificó la distancia existente entre el *landmark* número 4 de M3 y una línea que unía el *landmark* número 2 y 6. Para los individuos de las zonas de distribución alopátrica, se efectuó un análisis de la varianza (ANOVA) con el modelo $d \sim sp$, en el que la distancia (d) del punto a la línea es la variable dependiente en función de la especie (sp). Para el total de los individuos, zonas de distribución alopátrica y solapada, el modelo de la ANOVA fue $d \sim sp * geo$, donde además de la variación interespecífica (sp) se controló el efecto de la distribución geográfica de los individuos en la variación intraespecífica de la forma ($sp * geo$).

2.5 PLATAFORMAS INFORMÁTICAS

Todos los análisis anteriores se realizaron con el software “R” (*R development Core Team* <http://cran.r-project.org/>) versión 2.15 para Windows con diferentes librerías: Mclust (Fraley & Raftery 2002; Fraley *et al.* 2012); geomorph (Adams & Otárola-Castillo 2013); GenKern (Lucy & Aykroyd 2012); gdata (Warnes *et al.* 2013); gtools (Warnes *et al.* 2012); plotrix (Lemon 2006); gridExtA (Auguie 2012).

3. ANÁLISIS DE LAS RELACIONES FILOGENÉTICAS ENTRE EL TOPILLO MEDITERRÁNEO Y EL TOPILLO LUSITANO MEDIANTE EL USO DE MARCADORES MITOCONDRIALES Y NUCLEARES

3.1 OBTENCIÓN DEL MATERIAL BIOLÓGICO

Los 179 individuos (60 de las zonas de distribución alopátrica y 119 de las zonas de distribución solapada), de los que se realizó la caracterización morfológica, fueron también utilizados para el análisis filogenético de estas dos especies de topillos (figura 14, tabla A3). Aparte de las precauciones ya mencionadas para la identificación de individuos a partir de cráneos de egagrópilas (*apartado 2.1.1*) se tomó otra precaución adicional para minimizar las posibles contaminaciones cruzadas a la hora de la extracción de ADN genómico. Para ello, cada una de las estructuras óseas seleccionadas fue cuidadosamente limpiada para eliminar cualquier resto de pelos y/o tejido que pudiera proceder de otros individuos. Una vez finalizado el proceso exhaustivo de limpieza las estructuras óseas se conservaron en etanol a 4° C.

Para completar el muestreo se obtuvieron cinco muestras de tejido fresco de las dos especies de topillos (1 de *M. duodecimcostatus* y 4 de *M. lusitanicus*), de nuestra propia colección o provenientes de donaciones (tabla A7). Estas muestras fueron conservadas en etanol a -20° C.

3.2 EXTRACCIÓN DE ADN GENÓMICO

3.2.1 Extracción de ADN genómico a partir de material óseo

Siempre que fue posible, se usó como material de partida, para la extracción de ADN genómico, los dientes de cada individuo. Se eligieron los dientes ya que presentan más cantidad de ADN que otras estructuras óseas (Casas-Marce *et al.* 2009). Además, este ADN se encuentra más protegido frente a la degradación por la propia estructura de los dientes (esmalte y dentina). En los casos en los que no se disponía de los dientes se utilizó una de las dos mandíbulas.

Se emplearon unos 50 mg como material de partida. Se utilizó el kit de extracción *DNEasy Blood&Tissue* de Qiagen siguiendo las instrucciones del fabricante con las siguientes modificaciones. Las muestras fueron tratadas previamente a la extracción con nitrógeno líquido para provocar una rotura física del hueso y favorecer así la extracción de ADN. Igualmente, el tiempo de incubación con el tampón de extracción y la proteinasa *K* se modificó, ya que se realizó una incubación durante seis horas, consiguiendo así la máxima eficiencia en la extracción. Se llevó a cabo una elución final del ADN en 60 µl de agua *MiliQ*. Se midió la concentración de ADN de cada muestra en un instrumento ND-1000 (Nanodrop Technologies), midiendo la absorbancia a 260 nm. Las concentraciones de ADN genómico que se obtuvieron fueron muy diferentes dependiendo del grado de preservación del tejido óseo y de la eficacia del proceso de extracción.

3.2.1 Extracción de ADN genómico a partir de muestras de tejido fresco

Se emplearon unos 25 mg de tejido fresco. Se utilizó el kit de extracción *DNEasy Blood&Tissue* de Qiagen siguiendo las instrucciones del fabricante. La elución final del ADN se realizó en un volumen final de 50 µl de agua *MiliQ*. Se comprobó la concentración de ADN de cada muestra con el ND-1000 (Nanodrop Technologies) midiendo la absorbancia a 260 nm.

3.3 AMPLIFICACIÓN POR PCR

Debido a que la mayor parte de este estudio se realizó con muestras de carácter no invasivo, las cantidades de ADN con las que se trabajó fueron mucho menores que si se hubiera trabajado con muestras de tejido fresco. Por tanto, se tomó una serie de medidas para asegurar que el material biológico amplificado no estuviera contaminado de forma cruzada con otros materiales genéticos. Todo el proceso de amplificación por PCR se llevó a cabo en una sala especialmente equipada, con irradiación ultravioleta, para su completa esterilización antes y después de cada uso.

3.3.1 Amplificación de secuencias mitocondriales

Se amplificó un total de 184 secuencias de citocromo *b*, 179 a partir de ADN genómico extraído de huesos y 5 a partir de ADN genómico de tejido fresco, de las dos especies de topillos, con una longitud total de 1143 pares de bases. Debido a la fragmentación del ADN genómico que se obtuvo del tejido óseo, se amplificó este gen en tres fragmentos más pequeños, con una longitud de 340, 446 y 464 pares de bases, respectivamente. Para ello, se desarrollaron 6 cebadores, de los cuáles cuatro internos (tabla A8).

El volumen final de las reacciones de amplificación fue de 25 µl. Se procuró que la concentración de partida de ADN genómico estuviera alrededor de 50 ng. Para cada cebador la concentración final fue de 1 µM, 0,2 mM para cada uno de los dNTPs, 0,75 unidades de *Promega GoTaq* ADN polimerasa y 17,5 µg de BSA (*bovine serum albumina*). Las fases de la amplificación fueron: una desnaturalización inicial a 95° C durante 3 minutos, seguida de 35 ciclos formados por: una desnaturalización inicial (30 segundos a 95° C), una fase de hibridación (30 segundos a 50° C) y una extensión (90 segundos a 72° C). Finalmente se añadió una extensión durante 10 minutos a 72° C.

Los productos de amplificación obtenidos fueron cargados en geles de agarosa al 1% pre-teñidos con *SYBR Safe DNA gel Stain* (Invitrogen) para comprobar si la amplificación se produjo con éxito.

Las amplificaciones que resultaron positivas fueron purificadas con *ExoSAP-It* (Affymetrix) para su posterior envío al servicio de secuenciación de Macrogen Inc. (Seoul, Corea del Sur). Se utilizaron para la secuenciación los mismos cebadores originales de la amplificación. Las secuencias finales fueron inspeccionadas, recortadas y ensambladas con el programa Geneious Pro v4 (Drummond *et al.* 2009) (Biomatters Ltd.).

3.3.2 Amplificación de secuencias nucleares

Se seleccionaron seis intrones (Agxt-10, Catsper3-5, Dhcr24-7, Nadsyn1-4, Smo-9 y Wls-7) de los diez utilizados en el primer capítulo. La selección de estos seis intrones se

debió a que fueron con los que se obtuvieron más amplificaciones positivas después de realizar unas pruebas previas de amplificación con muestras de las dos especies de topillos. Los cebadores utilizados para la amplificación fueron los mismos que los diseñados en el primer capítulo, sin que fuera necesaria ninguna modificación para la amplificación específica para estas dos especies de topillos (tabla A8).

Para los individuos en los que la calidad de la secuencia mitocondrial obtenida fue de muy baja calidad, no se intentó la amplificación de las secuencias nucleares, ya que la probabilidad de éxito era muy reducida. Para el resto de individuos se probó, de forma sistemática, la amplificación inicial de los intrones Agxt-10, Dhcr24-7 y Smo-9. Estos tres intrones son los que presentaban un mayor éxito de amplificación en pruebas preliminares, por lo que fueron considerados como indicadores de la calidad del ADN genómico. Por tanto, para aquellos individuos que presentaron una amplificación positiva para estos intrones, se continuó con la amplificación de los otros tres intrones (Catsper3-5, Nadsyn1-4 y Wls-7). Los individuos finales para los que se consiguió la amplificación de secuencias de intrones fueron 44.

Se llevó a cabo la amplificación mediante un protocolo de *touch-down* PCR (*apartado 1.2.3*) o mediante una amplificación por PCR a temperatura fija (*apartado 3.3.1*). Las únicas modificaciones que se realizaron fueron las diferentes temperaturas de hibridación que se utilizaron para obtener amplificaciones de banda única y de mejor calidad (tabla A8).

Los productos de amplificación obtenidos fueron cargados en geles de agarosa al 1% pre-teñidos con *SYBR Safe DNA gel Stain* (Invitrogen) para comprobar si la amplificación se produjo con éxito.

Las amplificaciones que resultaron positivas fueron purificadas con *ExoSAP-It* (Affymetrix) para su posterior envío al servicio de secuenciación de MacroGen Inc. (Seoul, Corea del Sur). Se utilizaron para la secuenciación los mismos cebadores originales de la amplificación. Las secuencias finales fueron inspeccionadas, recortadas y ensambladas con el programa Geneious Pro v4 (Drummond *et al.* 2009) (Biomatters Ltd.).

3.3.3. Evaluación de la posible pérdida de alelos

Otro de los problemas asociados al hecho de trabajar con ADN genómico procedente de muestras biológicas de tipo no invasivo es la posible pérdida de alelos en la amplificación por PCR en el caso de las secuencias nucleares, de modo que se puede obtener un índice de homocigosidad más elevado del real. Este fenómeno es más habitual cuando uno de los alelos es más pequeño que su complementario (Pompanon *et al.* 2005).

Para cuantificar la posible pérdida alélica se testó, mediante un test de Chi-cuadrado, si existían diferencias significativas entre el número de secuencias heterocigotas amplificadas a partir de las muestras de tejido fresco y a partir de las muestras de tejido óseo para todos los intrones para cada especie por separado. Se contabilizó la heterocigosidad mediante el número de polimorfismos presentes en cada secuencia analizada.

3.4 OBTENCIÓN DE LAS FASES HAPLOTÍPICAS

Con el objetivo de obtener la mayor información posible de las secuencias de los intrones amplificados, se procedió a la resolución de las fases haplotípicas de las secuencias heterocigotas (bien por polimorfismos de mutaciones puntuales o por diferencia de longitud de los alelos) empleando dos metodologías diferentes.

3.4.1 Obtención computacional de las fases haplotípicas

Se determinó la fase haplotípica usando el programa PHASE versión 2.1.1 (Stephens *et al.* 2001). El umbral de probabilidad de aceptación de un haplotipo resuelto se fijó en 0,9 (-p 0,9). Este umbral se estableció así de restrictivo para evitar resoluciones de fases haplotípicas no existentes en las poblaciones muestreadas. Se empleó SeqPHASE (Flot *et al.* 2006) para generar los archivos de entrada requeridos por PHASE.

3.4.2 Obtención experimental de las fases haplotípicas

Para las secuencias cuya fase haplotípica no se pudo resolver de forma computacional se empleó un método de clonación tradicional. Los productos de PCR fueron introducidos en vectores plasmídicos *pSTblue-1* (Invitrogen) por un proceso de ligación, siguiendo las instrucciones del fabricante. Posteriormente se realizó una transformación mediante choque térmico de células competentes *Nova Blue* de *Escherichia coli*. Las células se sembraron en placas con medio LB con carbenicilina y se les añadió X-Gal e IPTG. Como la secuencia a clonar se insertó en el marco de lectura del gen *lacZ* (del plásmido) todas aquellas bacterias transformadas pero cuyos plásmidos no presentaban una ligación correcta (de la secuencia a clonar) fueron capaces de sintetizar la enzima β -galactosida y por tanto fueron capaces de degradar el X-gal, ayudadas por el inductor IPTG, originando colonias de color azul. Por otro lado, las bacterias que fueron transformadas con el plásmido correctamente ligado a la secuencia a clonar produjeron colonias de color blanco, ya que eran incapaces de sintetizar la enzima β -galactosida. Por último, aquellas bacterias que no fueron transformadas por el plásmido murieron ya que no presentaban resistencia a la carbemcilina. Por cada placa se seleccionaron seis colonias totalmente blancas.

Seguidamente se procedió a la amplificación de las secuencias clonadas. Se utilizaron los cebadores sugeridos por los fabricantes del kit de clonación (T7 y SP6). Las condiciones de amplificación por PCR fueron las mismas que las descritas en el apartado 3.3.1 pero con una temperatura de hibridación de 54° C.

Aquellas amplificaciones que resultaron positivas fueron purificadas con *ExoSAP-It* (Affymetrix) para su posterior envío al servicio de secuenciación de Macrogen Inc. (Seoul, Corea del Sur). Se utilizaron para la secuenciación los mismos cebadores que para la amplificación por PCR.

Se intentó amplificar y secuenciar correctamente al menos cuatro copias de cada secuencia de referencia para poder así controlar los posibles errores clonación. Para los casos en los que la heterozigosidad era debida a la existencia de uno a varios polimorfismos, si mediante la clonación se obtuvo solamente uno de los alelos, el otro se dedujo a partir del clonado como secuencia complementaria. Sin embargo, en las

situaciones en las que existía una diferencia de longitud entre los alelos de la secuencia y por clonación solamente se obtuvo uno de ellos el otro no se dedujo, ya que para ciertos marcadores había una gran variabilidad en el número de pares de bases que formaban las deleciones o inserciones.

3.5 ANÁLISIS FILOGENÉTICO DE LOS MARCADORES AMPLIFICADOS

3.5.1 Análisis filogenético de las secuencias mitocondriales

Se realizó el alineamiento múltiple con las 184 secuencias mitocondriales correctamente amplificadas y secuenciadas mediante el programa Mafft versión 6.708 (Katoh & Toh 2008). A partir de este alineamiento se reconstruyó un árbol de máxima verosimilitud usando RaxML versión 7.0.4 con un modelo GTR de sustitución nucleotídica y una distribución de tipo gamma para las tasas evolutivas (Stamatakis 2006) (figura A5).

3.5.2 Análisis filogenético de las secuencias nucleares

En total se resolvieron 431 alelos de los seis intrones seleccionados para los 44 individuos en los que se realizaron las amplificaciones. Cada uno de los alelos obtenidos fue tratado como una secuencia independiente. Se realizaron alineamientos múltiples con todos los alelos para cada uno de los seis intrones mediante Mafft versión 6.708 (Katoh & Toh 2008) y posteriormente se reconstruyeron los correspondientes árboles de máxima verosimilitud usando RaxML versión 7.0.4 con un modelo GTR de sustitución nucleotídica y una distribución de tipo gamma para las tasas evolutivas (Stamatakis 2006) (figura A6).

3.6 DISCORDANCIA CITO-NUCLEAR Y MORFOLÓGICA: CLASIFICACIÓN DE LOS INDIVIDUOS

En el capítulo anterior se llevó a cabo un proceso de asignación de especie basándose en características morfológicas de los individuos de las zonas de distribución solapada usando una metodología LDA (tabla A6). Al realizar la reconstrucción filogenética del citocromo *b* se observó que no existía una total correspondencia entre la asignación de especie morfológica y el clado mitocondrial al que pertenecían los individuos. En el

árbol de máxima verosimilitud del citocromo *b* (figura A5) se podían diferenciar dos grandes clados, uno formado por individuos pertenecientes a la especie *M. duodecimcostatus* y otro con individuos de la especie *M. lusitanicus*. La discordancia morfológica-mitocondrial se produjo al observar individuos de las zonas de distribución solapada clasificados como *M. lusitanicus*, basándose en su morfología, dentro del clado mitocondrial perteneciente a *M. duodecimcostatus* (tabla A9A).

Posteriormente, se analizó la información que contenía los intrones de estos individuos con una discordancia morfológica-mitocondrial en los 44 individuos en los que fue posible (*apartado 3.3.2*) (figura A6). Se observó que para los intrones que presentaban haplotipos propios de cada especie existía una discordancia mitocondrial-nuclear pero no una discordancia morfológica-nuclear (tabla A9B).

Por todo esto, se decidió denominar y agrupar a estos individuos en una categoría propia para los análisis sucesivos, debido a sus particularidades. La denominación arbitraria que recibieron fue: grupo discordante cito-morfonuclear (DCMN). Por tanto, el resto de individuos que presentó una plena concordancia entre su clasificación morfológica, mitocondrial y nuclear a nivel de especie, recibió el nombre de grupo concordante cito-morfonuclear (CCMN). En los análisis en los que se requería una asignación específica *a priori* de los individuos, a los del grupo DCMN les fue asignada la especie *M. lusitanicus*, ya que dos de los dos criterios para la asignación de especie (morfológico y nuclear) les colocaban dentro del grupo del topillo lusitano (tabla A9).

Siguiendo los criterios anteriores, y ya que los individuos C1293 y C1948 fueron clasificados como *M. duodecimcostatus* a nivel morfológico (tabla A6) y a nivel mitocondrial se situaron dentro del grupo *M. duodecimcostatus* (figura A5) su clasificación natural hubiera sido ser incluidos dentro del grupo CCMN. Sin embargo, como de estos dos individuos se obtuvo información nuclear que les clasificaba como *M. lusitanicus* (figura A6) y los coeficientes de asignación posterior de especie —a nivel morfológico— eran intermedios entre las dos especies (tabla A6) se decidió incluirlos dentro del grupo DCMN (tabla A9B).

Por último, para los individuos C1296 y C1252, sus secuencias de citocromo *b* se encontraban dentro del grupo *M. duodecimcostatus* (figura A5) pero el análisis

morfométrico no discriminó de manera completa la especie a la que pertenecían (tabla A6) y aunque no fue posible obtener secuencias nucleares que pudieran corroborar la especie a la que pertenecían se decidió incluirlos dentro del grupo DCMN para los análisis posteriores (tabla A9A).

3.7 ANÁLISIS DE LA DIVERSIDAD Y DE LA ESTRUCTURA GENÉTICA

Se estimó la diversidad haplotípica y nucleotídica para los seis marcadores nucleares y para el marcador mitocondrial, por separado, para cada una de las dos especies de topillos. Se cuantificó la desviación de la neutralidad mediante tres tests diferentes: el F_s de Fu (Fu & Li 1993), la D de Tajima (Tajima 1989) y el R_2 (Ramos-Onsins & Rozas 2002). Para obtener la significación estadística del test de Fu y el R_2 se hicieron 1.000 simulaciones usando el algoritmo de coalescencia. Todos estos parámetros se calcularon con DnaSP versión 5.10.01 (Librado & Rozas 2009).

Se evaluó la presencia de eventos de recombinación entre los alelos de las secuencias nucleares mediante dos test diferentes. El test de los cuatro gametos (R_m) implementado en DnaSP v5 calcula el número mínimo de eventos de recombinación para cada locus y para cada especie. El estadístico ϕ_w , que es un índice de homoplasia para todas las comparaciones por pares (*pairwise homoplasy index*) (Bruen *et al.* 2006) se calculó para cada uno de los intrones por separado confrontando todos los alelos disponibles independientemente de la especie a la que pertenecían. Como resultado del test se obtuvo un *p-value* de probabilidad asociada a la existencia de procesos de recombinación en las secuencias analizadas. Además del *p-value* de los datos reales se realizaron 1.000 permutaciones para comprobar la robustez del test.

Se construyeron, de forma paralela, genealogías haplotípicas con el programa Hapviewer versión 1.0 (Salzburger *et al.* 2011). Se construyó una genealogía haplotípica para cada intrón y una para representar la información mitocondrial de los 44 individuos que contaban además con información de secuencias nucleares. Para construir estas genealogías se utilizó la información contenida en los alineamientos y en los árboles de máxima verosimilitud reconstruidos en el apartado 3.5.

Por último, se calculó la distancia genética media del citocromo *b* entre *M. dudodecimcostatus* y *M. lusitanicus* utilizando la información mitocondrial de los 184 individuos con el programa Mega versión 5.2 (Tamura *et al.* 2011).

3.8 SELECCIÓN DE LOS MODELOS EVOLUTIVOS Y DEL RELOJ MOLECULAR MÁS ADECUADOS PARA LOS MARCADORES NUCLEARES

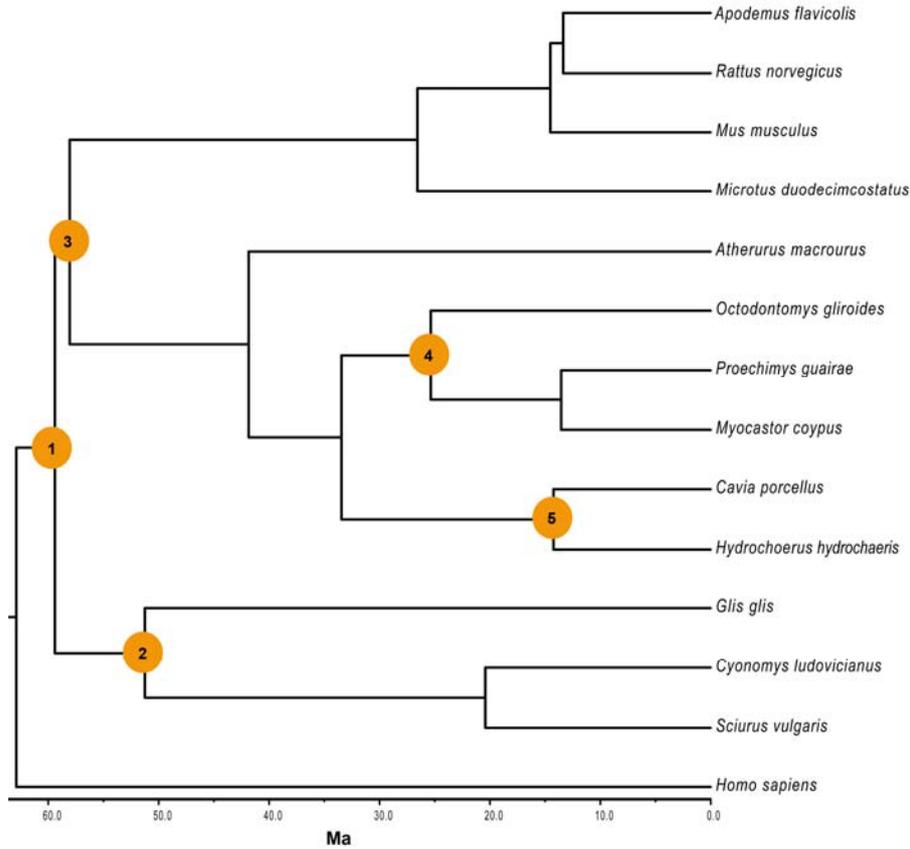
Se eligió el modelo de sustitución más adecuado mediante jModeltest (Posada 2008) (tabla A10). Con el objetivo de implementar el reloj molecular apropiado (estricto o relajado no correlacionado) se estimó con PAUP* versión 4.0b10 (Swofford 2003) la verosimilitud de una topología de PhyML (Guindon & Gascuel 2003) con y sin reloj molecular. Se decidió el modelo más apropiado con un test de índice de verosimilitud (Felsenstein 1981) (tabla A10).

3.9 CÁLCULO DE LAS TASAS EVOLUTIVAS DE LOS INTRONES SELECCIONADOS PARA EL GÉNERO *Microtus*

Para poder realizar estimas filogenéticas más precisas sobre las dos especies de topillos se calcularon las tasas evolutivas de los seis intrones utilizados para el género *Microtus*.

En primer lugar, se seleccionaron las secuencias del orden *Rodentia* obtenidas en el primer capítulo (tabla 2) y de las secuencias obtenidas de las bases de datos (*Mus musculus*, *Rattus norvegicus*, *Cavia porcellus* y *Homo sapiens*) de los seis intrones que fueron utilizados en este capítulo (Agxt-10, Catsper3-5, Dhcr24-7, Nadsyn1-4, Smo-9 y Wls-7). Estas secuencias se alinearon usando Mafft versión 6.708 (Kato & Toh 2008) y se eliminaron las regiones mal alineadas usando Gblocks con parámetros relajados (Castresana 2000; Talavera & Castresana 2007). Estos alineamientos se introdujeron como particiones independientes en el análisis de inferencia Bayesiana realizado con Beast versión 1.6.1 (Drummond & Rambaut 2007).

A)



B)

Clado	Límite mínimo estricto (<i>Hard bound</i>)	Límite máximo difuso (<i>Soft bound</i>)	Parámetros lognormal	
			Media	Offset
<i>Rodentia</i> (B) (M)	55,6	65,8	3,25	55,6
<i>Sciuridae</i> (M)	48,4	61,1	4,04	48,4
<i>Hystricognathi</i> (B)	52,5	58,9	2,04	52,5
<i>Octodontoidea</i> (M)	24,5	37,3	4,07	24,5
<i>Hydrocaeridae-Caviidae</i> (M)	11,8	34	7,06	11,8

Figura 16. Uso de calibraciones fósiles para la estima de tasas evolutivas de marcadores moleculares. **(A)** Árbol de inferencia Bayesiana reconstruido con seis intrones (Agxt, Catsper3, Dhcr24, Nadsyn1, Smo y Wls) para 10 especies del orden *Rodentia*. Los cinco puntos indican las calibraciones fósiles utilizadas. (1) Delimita el clado *Rodentia*. (2) Delimita el clado *Sciuridae*. (3) Delimita el clado *Hystricognathi*. (4) Delimita el clado *Octodontoidea*. (5) Delimita el clado *Hydrocaeridae-Caviidae*. **(B)** Priors de calibración en millones de años. (B, Benton *et al.* 2009; M, Meredith *et al.* 2011).

Paralelamente, se seleccionaron cinco puntos de calibraciones fósiles, descritos en la literatura (Benton *et al.* 2009; Meredith *et al.* 2011) distribuidos a lo largo de la filogenia del orden *Rodentia* (figura 16A). La información de estas calibraciones fósiles fue introducida en el análisis de inferencia Bayesiana del siguiente modo. Se usaron como *priors* distribuciones lognormales definidas por el límite mínimo de divergencia, que determinó el *offset* y por el límite máximo difuso, cuyo valor se hizo coincidir con el percentil 95% de la densidad de probabilidad, determinando de este modo la media. La desviación estándar de la distribución lognormal se fijó en 1 (figura 16B).

Para todos aquellos clados que estaban delimitados por las calibraciones se forzó la monofilia. Se usaron para cada una de las seis particiones (intrones) secuencias de *Homo sapiens* como grupo externo del análisis. A cada intrón se le asignó el modelo evolutivo y el reloj molecular que le correspondía (tabla A2). Se corrió un total de $50 \cdot 10^6$ generaciones. Se descartó a modo de *burn-in* el 10% inicial de las muestras generadas y se comprobó la correcta convergencia de todos los parámetros usando la utilidad de Beast Tracer, asegurando que todos los tamaños de muestra efectivos fueran superiores a 200.

En segundo lugar, a partir del conjunto de árboles muestreados de la distribución posterior obtenidos por Beast se obtuvieron los datos de la tasa evolutiva de la rama correspondiente al género *Microtus* para cada uno de los seis intrones. Cada una de estas tasas evolutivas fue ajustada a una distribución normal, obteniendo así un valor medio y una distribución estándar.

Para comprobar posibles interacciones entre los *priors* de calibración empleados para los distintos nodos del árbol se estimaron las distribuciones efectivas conjuntas de estos *priors* (*effective joint prior distributions*) (Drummond *et al.* 2006) realizando un análisis igual al descrito anteriormente pero sin incluir los datos de las secuencias. Estos resultados se compararon con las distribuciones posteriores de las edades estimadas para cada nodo donde se introdujo una calibración y con la información de los *priors* de calibración introducidos en el análisis inicial (Sanders & Lee 2007) (figura A7).

3.10 CÁLCULO DEL TIEMPO DE ESPECIACIÓN DE LAS DOS ESPECIES DE TOPILOS Y DE SUS PARÁMETROS DEMOGRÁFICOS MEDIANTE ÁRBOLES DE ESPECIES

Para calcular el tiempo de especiación entre *M. duodecimcostatus* y *M. lusitanicus* así como sus respectivos parámetros demográficos (tamaño poblacional ancestral y tamaño de las poblaciones actuales) se emplearon dos metodologías diferentes de árboles de especies basadas en análisis Bayesiano: *Beast verion 1.6.1 (Heled & Drummond 2010) e IMA2 (Hey 2010). De ellas, la última incorpora un modelo de aislamiento con migración.

Para los parámetros estimados con *Beast se recogió la mediana y los intervalos inferior y superior del 95% de densidad de probabilidad (95% HPD). Para el caso de IMA2 se recogió el valor de HPit (*the value of the bin with the highest count*) y de los intervalos inferior y superior del 90% de densidad de probabilidad (HPD90) de cada parámetro estimado.

3.10.1 Estima de parámetros empleando *Beast

Para estos análisis sólo se tuvieron en cuenta los 44 individuos de los que se tenía información mitocondrial y nuclear, con una asignación *a priori* de especie según lo descrito en el apartado 3.6 (tabla A9B). Se plantearon diferentes situaciones de análisis variando el tipo de marcador molecular empleado (mitocondrial, nuclear, o los dos) y los individuos incluidos (sólo CCMN o CCMN junto DCMN) con el objetivo de realizar una estima de los parámetros lo más precisa posible (tabla A11).

Se introdujo, en cada análisis, la información de cada uno de los marcadores moleculares como particiones independientes. Las tasas evolutivas de cada marcador introducidas en el análisis fueron las calculadas en el apartado 3.9. Se introdujeron como *priors* de una distribución normal con la media y desviación típicas calculadas. A cada marcador le fue asignado el modelo evolutivo y el reloj molecular que le correspondía (tabla A10). Cada uno de los análisis se corrió un total de $50 \cdot 10^6$ generaciones. Se descartó a modo de burn-in el 10% inicial de las muestras generadas y se comprobó la correcta convergencia de todos los parámetros usando la utilidad de

Beast Tracer, asegurando que todos los tamaños de muestra efectivos fueran superiores a 200.

Algunos de los análisis de *Beast se corrieron bajo un modelo de coalescencia constante, y otros análisis se corrieron bajo un modelo de coalescencia linear y raíz constante para obtener parámetros más completos (tabla A11). Los análisis de IMA2 se corrieron siempre bajo un modelo de coalescencia constante ya que es el único implementado en el programa.

3.10.2 Estima de parámetros empleando IMA2

Con los modelos IM (*isolation and migration*) e I (*isolation*; con la migración limitada a cero) del programa IMA2 se realizaron los mismo análisis en cuanto a variables introducidas que los realizados con *Beast (tabla A11).

Las tasas evolutivas por gen requeridas por IMA2 para cada uno de los marcadores se calcularon multiplicando las tasas evolutivas obtenidas en el *apartado 3.9* por la longitud de cada molécula. En las situaciones en las que existía una diferencia en la longitud de los alelos se calcularon multiplicando la tasa evolutiva por la longitud más baja. El modelo evolutivo utilizado tanto para los intrones como para el citocromo *b* fue el HKY.

Se realizó una serie de análisis previos para determinar cuáles eran los *priors* más adecuados y el número de genealogías muestreadas necesarias para obtener una buena convergencia de los parámetros y asegurarse de que el muestreo se realizaba en todo el espacio paramétrico (tabla A12).

Cada análisis (tabla A11) se corrió tres veces de forma paralela y en cada una se muestrearon 100.000 genealogías. Finalmente, se combinaron las tres réplicas de cada análisis mediante la opción L-Mode implementada en el programa IMA2.

3.11 IDENTIFICACIÓN DEL PROCESO QUE CAUSA LA DISCREPANCIA CITO-NUCLEAR: INTROGRESIÓN O SEPARACIÓN INCOMPLETA DE LINAJES

Para poder identificar la causa de la discordancia cito-nuclear se analizaron las secuencias mitocondriales y nucleares mediante tres métodos diferentes.

En primer lugar, se utilizaron los modelos de aislamiento con migración, IM (IMa2) descritos en el *apartado 3.10.2*. Para este objetivo, uno de los parámetros que se tuvo en cuenta fue la tasa de migración (m), que es la tasa a la que una especie recibe un número determinado de individuos de otra especie y el otro fue la tasa de migración poblacional ($2Nm$), que es el producto de la tasa de migración por el número efectivo de copias de genes en la población receptora. Igualmente, las estimas del tiempo de especiación calculadas bajo un modelo con migración, IM (IMa2) y otro si migración, I (IMa2) junto con las calculadas con *BEAST sirvieron como información para discernir el origen de la discordancia cito-nuclear.

Además, se usó un método para distinguir procesos de introgresión por hibridación y separación incompleta de linajes basado en simulaciones paramétricas, siguiendo un proceso similar al propuesto por el método denominado JML (Joly *et al.* 2009). Este método está basado en la idea de que los procesos de hibridación deben dar lugar a distancias genéticas entre especies menores que los procesos de separación incompleta de linajes.

En primer lugar, se simularon secuencias mitocondriales a partir de un árbol de especies con los parámetros de tiempo de especiación y tamaños poblacionales obtenidos a partir de la reconstrucción de las secuencias de *M. lusitanicus* y *M. duodecimcostatus* que se consideró más adecuada. Esta reconstrucción fue mediante el uso de IMa2 (IM) y usando genes nucleares y todos los mitocondriales (análisis número 6, tabla A11). Usando el programa MCcoal versión 1.2 (Rannala & Yang 2003) y los parámetros del árbol de especies (θ *M. dudodecimcostatus* actual= 0,607936; θ *M. lusitanicus* actual= 2,856283; θ Ancestral= 0,213966; τ = 0,537633), se simularon 10.000 genealogías bajo un modelo nulo de coalescencia para dos especies, con 15 y 29 individuos

correspondientes a *M. duodecimcostatus* y *M. lusitanicus*, respectivamente. La tasa mutacional que se usó para el marcador mitocondrial fue de 0,019 (Soria-Carrasco 2011). Para cada una de las genealogías así obtenidas, se simularon secuencias de ADN usando el programa SeqGen version 1.3.2 (Rambaut & Grassly 1997). Las secuencias, de 1143 pares de bases de longitud, se simularon con un modelo GTR, con los valores de la gamma y las frecuencias nucleotídicas tomadas del alineamiento de citocromo *b* de 44 secuencias con el programa JModeltest (Posada 2008). También se usó una proporción de posiciones invariantes calculada a partir de este mismo alineamiento. A partir de estas secuencias se reconstruyeron árboles filogenéticos usando PhyML versión 2.4.4 (Guindon & Gascuel 2003), usando un modelo GTR + gamma. Por último, de cada uno de estos árboles de genes se estimó la distancia genética mínima entre dos pares de nodos terminales (individuos) que incluían las dos especies diferentes. Para este proceso se usó un programa Perl desarrollado para este propósito. Finalmente, de la distribución de 10.000 distancias genéticas mínimas se tomó la distancia genética correspondiente al 5% más bajo y se comparó con la mínima distancia genética del árbol de citocromo *b* real de los individuos de *M. lusitanicus* y *M. duodecimcostatus*. Si el valor obtenido a partir de la filogenia real estaba contenido dentro del 5% más bajo de la distribución nula se considera que esta distancia genética tan baja no puede explicarse por separación incompleta de linajes y, por lo tanto, podría deberse a algún evento de introgresión por hibridación. Las diferencias de este método con el del JML son que aquí se calcularon distancias genéticas y no número de posiciones diferentes, lo que hace el test más riguroso para secuencias de alta tasa evolutiva como son las secuencias mitocondriales.

Por último, se calculó un índice denominado GSI (*genealogical sorting index*) (Cummings *et al.* 2008). El GSI es un índice que cuantifica el grado de exclusividad ancestral o monofilia de un determinado clado en una filogenia enraizada (Cummings *et al.* 2008). Este índice varía de 1 a 0, donde 1 indica una máxima monofilia del clado seleccionado. La idea de usar este índice es que el grado de monofilia y, por tanto, el GSI será menor en un árbol con introgresión que en un árbol donde sólo esté presente la separación incompleta de linajes.

Este índice se calculó en el árbol de citocromo *b* de los 44 individuos (de los que se obtuvo también información nuclear), en el árbol de citocromo *b* del total de los 184

individuos y en las simulaciones paramétricas del anterior método. Para este último análisis se calculó un valor de GSI para cada una de las 10.000 simulaciones y, a partir de esta distribución de valores de GSI, se calculó la media y la desviación estándar, obteniendo un valor de GSI total para cada especie.

Tanto para el cálculo de la mínima distancia genética como para el cálculo del GSI se analizó la información mitocondrial de los 44 individuos por separado —a pesar de estar incluidos en el árbol mitocondrial de los 184 individuos— porque las 10.000 simulaciones generadas bajo un modelo nulo (sin migración) fueron reconstruidas con la información del árbol de especies de los 44, con lo cual la comparación entre resultados podría ser más directa.

IV. RESULTADOS Y DISCUSIÓN

1. DESARROLLO DE NUEVOS MARCADORES NUCLEARES PARA EL ORDEN *Rodentia*

1.1 RESULTADOS

1.1.1 Conjunto de intrones seleccionados

Se obtuvo un total de 12.513 genes ortólogos 1:1 para las cuatro especies de las que se usó información genómica: ratón, rata, conejillo de indias y humano (figura 17).

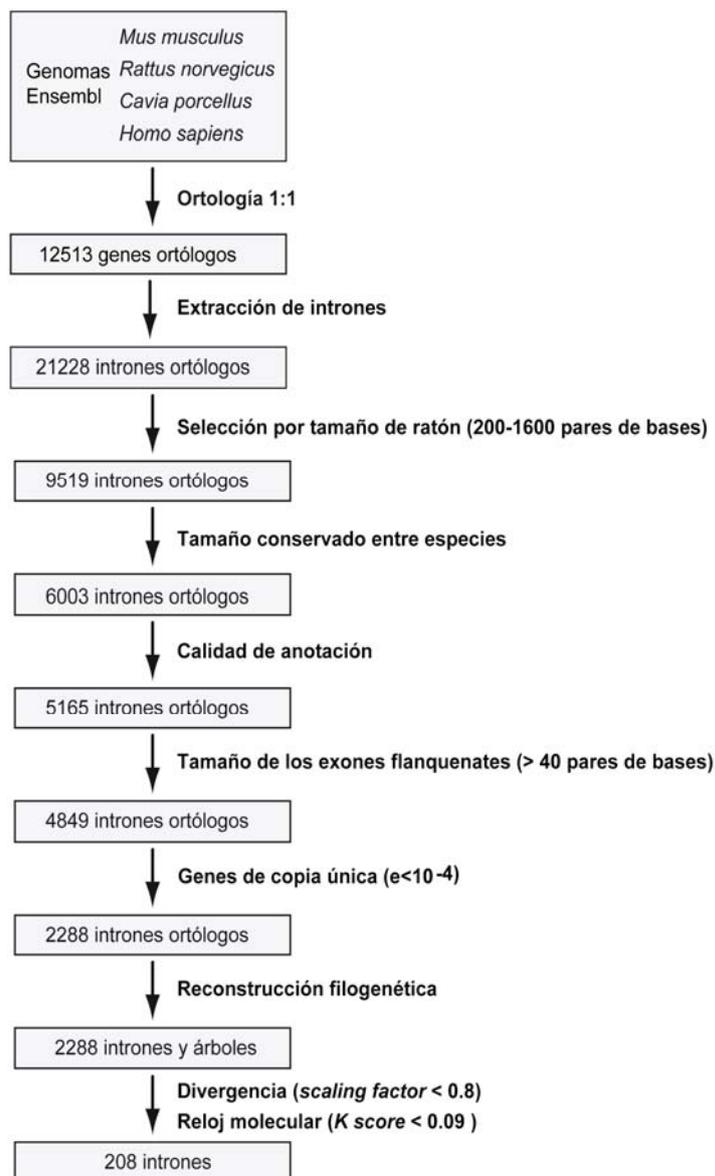


Figura 17. Conjunto de filtros bioinformáticos aplicados a tres genomas de roedores y al genoma de humano para la obtención de un conjunto final de 208 intrones.

De estos genes ortólogos se extrajeron 21.228 intrones con sus correspondientes exones. Después de aplicar el filtro que controlaba para la longitud en ratón (200-1.600 pares de bases) se seleccionaron 9.519 intrones. Tras la eliminación de aquellos intrones que eran muy diferentes en longitud entre las cuatro especies quedaron 6.003 intrones. De este conjunto, se eliminaron aquellos con mala calidad de secuencia o mala anotación, obteniendo 5.165 intrones. Después de descartar los intrones que estaban flanqueados por exones más pequeños de 40 nucleótidos, se obtuvo un conjunto de 4.849 intrones. Mediante búsquedas BLAST para eliminar intrones que no eran de copia única en el genoma se suprimió más de la mitad de ellos. Sobre los 2.288 intrones que quedaron se aplicó el resto de filtros de carácter filogenético.

Usando el árbol de referencia —el concatenado de todos los intrones— (figura A1) se obtuvieron los valores de *scaling factor* de los 2.288 intrones, que variaron entre 4,936 y 0,064 (figura A2). Para evitar seleccionar aquellos intrones que no fueran lo suficientemente variables, se seleccionaron los intrones con un *scaling factor* menor de 0,8. Los valores de *K score* variaron entre 0,467 y 0,012 (figura A2). Se eligieron aquellos intrones con un valor de *K score* menor de 0,09 para seleccionar aquellos intrones donde la forma del árbol y el reloj molecular se mantenía más parecida al árbol de referencia.

Después de aplicar todos estos filtros se obtuvo un conjunto de 208 intrones. Se elaboró un catálogo donde se recoge diversa información acerca de estos marcadores (Rodríguez-Prieto *et al.*, en preparación). Para cada intrón se incluye información sobre la función del gen al que pertenece, localización en el genoma, longitud del intrón y de los exones flanqueantes y parámetros que indican la divergencia del intrón (la longitud de la rama ratón-rata, el *scaling factor* y el *K score*). Asimismo, se recogen los alineamientos de los exones flanqueantes y del intrón para las cuatro especies así como el árbol filogenético del intrón. El examen de esta información permitirá seleccionar marcadores específicos para cada estudio y el diseño de cebadores con distintos grados de especificidad.

1.1.2 Validación experimental de los intrones para el orden Rodentia

Para probar si estos intrones funcionaban en distintas especies de roedores se escogieron diez de ellos y se diseñaron cebadores a partir de los exones. Estos intrones se aplicaron a 10 especies representativas de roedores (tabla A1). Mediante la técnica *touch-down* PCR se obtuvo, con el primer rango de temperaturas de hibridación (65-55° C), 76 de las 100 amplificaciones posibles (tabla 2). Para aquellas amplificaciones en las que las bandas eran muy débiles o inexistentes se probó una nueva amplificación a 65-50° C, pero no se obtuvo ningún resultado positivo. En amplificaciones inespecíficas (con dos o más bandas) se subió la temperatura a 65-60° C y se consiguieron dos amplificaciones correctas (de banda única) adicionales. Por tanto, se consiguió un 78% de éxito de amplificación. El único marcador que fue amplificado para todas las especies fue el Dhcr24-intrón-7 y el que menos éxito de amplificación tuvo fue el Rras-intrón-4. En el caso de las especies, para ninguna de ellas se consiguió una amplificación para todos los marcadores y el éxito varió entre el 30% y el 90% de las amplificaciones.

IV. Resultados y Discusión

Tabla 2. Conjunto de diez intrones amplificados en las diez especies del género *Rodentia* seleccionadas. Las amplificaciones positivas muestran la temperatura de hibridación a la que se realizó la amplificación mediante *touch-down* PCR. Las temperaturas subrayadas representan las amplificaciones positivas obtenidas posteriormente a la optimización de la temperatura de hibridación inicial. Nombre de la especie: **A**, *Octodontomys gliroides*; **B**, *Atherurus macrourus*; **C**, *Proechimys guairae*; **D**, *Myocastor coypus*; **E**, *Hydrochoerus hydrochaeris*; **F**, *Cynomys ludovicianus*; **G**, *Sciurus vulgaris* **H**, *Glis glis*; **I**, *Apodemus flavicollis*; **J**, *Microtus duodecimcostatus*.

	A	B	C	D	E	F	G	H	I	J
Abcb9 intrón 2	65-55	65-55	65-55	65-55	65-55	65-55	65-55	-	65-55	65-55
Agxt intrón 10	65-55	<u>65-60</u>	65-55	-	65-55	-	-	-	65-55	65-55
Catsper3 intrón 5	65-55	65-55	-	-	65-55	65-55	65-55	65-55	-	65-55
Dhcr24 intrón 7	65-55	65-55	65-55	65-55	65-55	<u>65-60</u>	65-55	65-55	65-55	65-55
Ivd intrón 8	65-55	65-55	65-55	-	65-55	65-55	65-55	65-55	65-55	65-55
Nadsyn1 intrón 4	65-55	65-55	65-55	65-55	-	65-55	65-55	65-55	65-55	65-55
Rras intrón 4	-	-	-	-	65-55	65-55	65-55	-	65-55	65-55
Smo intrón 9	65-55	65-55	-	-	65-55	65-55	65-55	65-55	65-55	65-55
Trpv4 intrón 8	65-55	65-55	-	-	65-55	65-55	65-55	-	65-55	-
Wls intrón 7	65-55	65-55	65-55	-	65-55	65-55	65-55	65-55	65-55	65-55

1.2.3 Análisis de la reconstrucción filogenética de los intrones amplificados

Los árboles filogenéticos de máxima verosimilitud obtenidos para cada uno de los 10 intrones son básicamente congruentes con las filogenias propuestas del orden *Rodentia* (Blanga-Kanfi *et al.* 2009; Honeycutt 2009; Churakov *et al.* 2010; Fabre *et al.* 2012) (figura A3). Este resultado indica que las nuevas secuencias amplificadas son ortólogas en todas las especies. Del mismo modo, los dos árboles filogenéticos, uno de máxima verosimilitud y otro de inferencia Bayesiana reconstruidos a partir de todas las secuencias concatenadas coinciden en términos generales con las filogenias de roedores (figura 18) (Blanga-Kanfi *et al.* 2009; Honeycutt 2009; Churakov *et al.* 2010; Fabre *et al.* 2012).

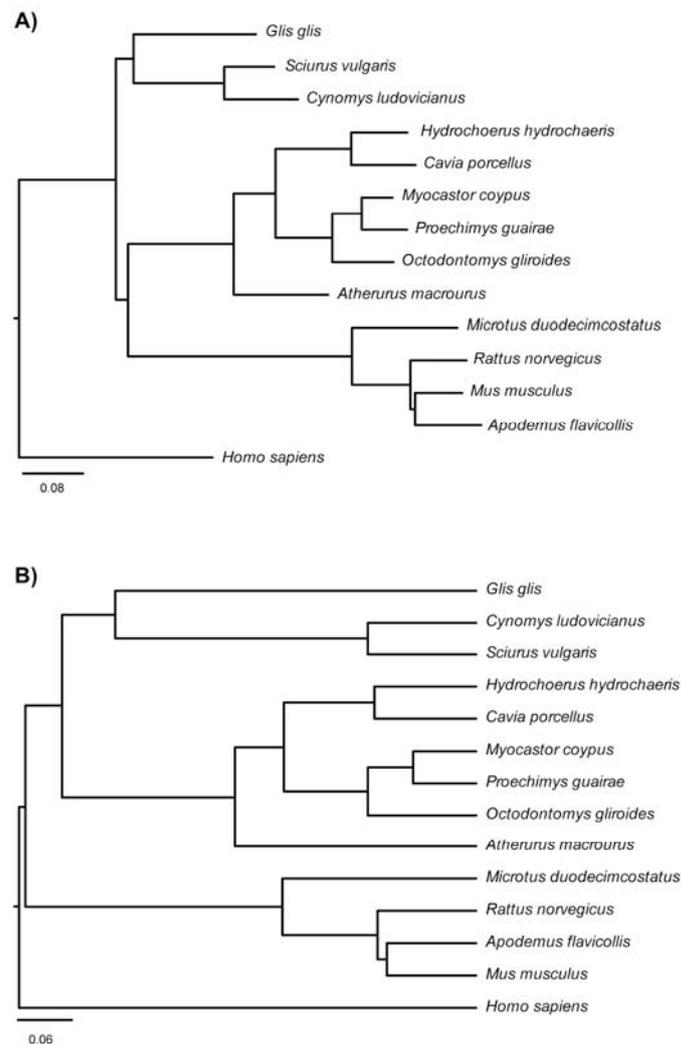


Figura 18. Análisis filogenético de los 10 intrones amplificados en las diez especies del género *Rodentia* más las secuencias correspondientes a ratón (*Mus musculus*), rata (*Rattus norvegicus*), conejillo de indias (*Cavia porcellus*) y humano (*Homo sapiens*). **(A)** Árbol filogenético de máxima verosimilitud reconstruido con todas las secuencias concatenadas. **(B)** Árbol de inferencia Bayesiana, en el que cada intrón fue introducido como una partición independiente. La escala indica número de sustituciones por posición.

1.2 DISCUSIÓN

1.2.1 Características generales del conjunto final de intrones

El dossier final de 208 intrones (Rodríguez-Prieto *et al.*, en preparación) obtenido mediante la aplicación de diferentes filtros bioinformáticos (de ortología, de calidad de la secuencia, de copia única, de carácter filogenético, etc.) es una herramienta muy útil que puede ser usada tanto para resolver cuestiones filogenéticas en especies cercanas como en taxa más alejados dentro del orden *Rodentia*. Para cada intrón se dispone de una serie de características útiles para decidir cuál de ellos seleccionar en función de la pregunta filogenética. De hecho, la amplificación experimental de 10 de estos 208 intrones resultó tener un gran éxito en varias especies de roedores. Se recomendaría para la correcta amplificación de estos intrones tener en cuenta dos factores antes del análisis experimental. En primer lugar, el éxito de la amplificación mejorará si, para el diseño de cebadores, se aumenta el número de secuencias de diferentes especies para el cálculo de bases degeneradas. En este estudio se incluyeron los exones disponibles en las bases de datos de diferentes especies de roedores, aparte de las cuatro de referencia, y de dos lagiformos. Gracias a estas secuencias adicionales se consigue que los cebadores reconozcan un mayor número de secuencias dentro de los roedores (y, aunque necesita una validación experimental, posiblemente también dentro de los lagomorfos). Otro importante factor para el éxito de la amplificación es la calidad del tejido a partir del cual se extrae el ADN genómico. Estudios iniciales realizados con material que no estaba en condiciones óptimas (como se demostraba por las bandas de PCR mitocondriales débiles) indicaron que eran muy pocos los intrones amplificables. Por tanto, parece claro que, como para muchas otras aplicaciones, la calidad del ADN condiciona en gran medida el éxito de la amplificación por PCR de los marcadores intrónicos.

Otro de los resultados que refleja las buenas características de esta selección intrónica son el árbol de máxima verosimilitud (obtenido con los 10 intrones concatenados) y el árbol de método Bayesiano (obtenido con los 10 intrones como particiones independientes) (figura 18). Estas dos reconstrucciones filogenéticas recogen la topología conocida de orden *Rodentia* (Blanga-Kanfi *et al.* 2009; Honeycutt 2009; Churakov *et al.* 2010; Fabre *et al.* 2012) excepto para el clado de los sciuridos y los

gliridos, que se situó o bien como basal (Blanga-Kanfi *et al.* 2009; Honeycutt 2009; Churakov *et al.* 2010; Fabre *et al.* 2012) o bien en una posición más interna (Meredith *et al.* 2011). Por tanto, se demuestra que estos intrones son válidos para reconstrucciones filogenéticas profundas aunque originariamente fueron desarrollados para resolver cuestiones filogenéticas entre especies cercanas del orden *Rodentia*.

1.2.2 Posibles aplicaciones de los intrones seleccionados

El interés de disponer de un conjunto de intrones caracterizado en función de su variabilidad, de su longitud, de los exones que los flanquean, etc. radica en la multitud de aplicaciones para las que pueden ser utilizados. En el contexto de su uso en la reconstrucción de árboles de especies, pueden usarse, junto a las calibraciones fósiles, para estimar tasas evolutivas de otros intrones para cada una de las especies de interés. Con esta información se pueden reconstruir árboles de especies en los que parámetros importantes como el tiempo de divergencia o parámetros demográficos se estimarán con más precisión (Brito & Edwards 2009; Degnan & Rosenberg 2009; Edwards 2009).

Por otro lado, se está empezando a combinar la potencia de la amplificación de marcadores óptimamente seleccionados con las técnicas de *Next Generation Sequencing* (NGS) para obtener un gran número de marcadores nucleares para la caracterización de individuos (McCormack & Faircloth 2013; O'Neill *et al.* 2013). Por tanto, disponer de un gran número de intrones ya caracterizados y optimizados que podrían ser amplificados de forma masiva generaría gran cantidad de información para estudiar procesos de especiación e historia evolutiva de un gran número de especies dentro del orden *Rodentia*.

2. VARIACIONES MORFOMÉTRICAS EXISTENTES ENTRE *M. duodecimcostatus* y *M. lusitanicus* RELACIONADAS CON SUS ÁREAS DE DISTRIBUCIÓN

2.1 RESULTADOS

2.1.1 Identificación y caracterización morfométrica de las dos especies de topillos en las zonas de distribución alopátrica

2.1.1.1 Asignación de especie y correlación con la distribución geográfica

Se observó que para el logCS existe una clara distribución bimodal para el carácter lateral (L) y para la mandíbula (J). Para el carácter ventral (V) se obtuvo un histograma de densidad con tres grupos. Por último, para el primer molar inferior (M1) y para el tercer molar superior (M3) se obtuvo una distribución completamente unimodal (figura 19A). Para el caso de la forma se observa una distribución de los individuos en dos grupos para L, J y M1. Sin embargo para el caso de V y M3 no se detectó una clara división en dos grupos, aunque para M3 un pequeño grupo de individuos parece separarse de la distribución principal (figura 19B). Los resultados del análisis discriminante indicaron que L, V y J son los caracteres que mejor discriminan entre las dos especies, ya que la correspondencia entre la especie *a priori* y la obtenida en el análisis es casi total (tabla 3). M1 y M3 no resultaron ser tan discriminantes a la hora de asignar la especie correspondiente a cada individuo.

Tabla 3. Asignación de especie a los individuos de las zonas de distribución alopátrica mediante un análisis discriminante (LDA) del tipo *leave-one-out* para los cinco caracteres. Se recoge la incongruencia entre la asignación de especie por su distribución geográfica y la obtenida con el LDA sobre la matriz de forma (Y) de cada carácter de manera independiente. D, *M. duodecimcostatus*; L; *M. lusitanicus*.

L			V			J			M1			M3		
	D	L		D	L		D	L		D	L		D	L
D	36	2	D	41	0	D	29	0	D	26	4	D	32	7
L	2	15	L	0	19	L	0	11	L	0	19	L	5	14

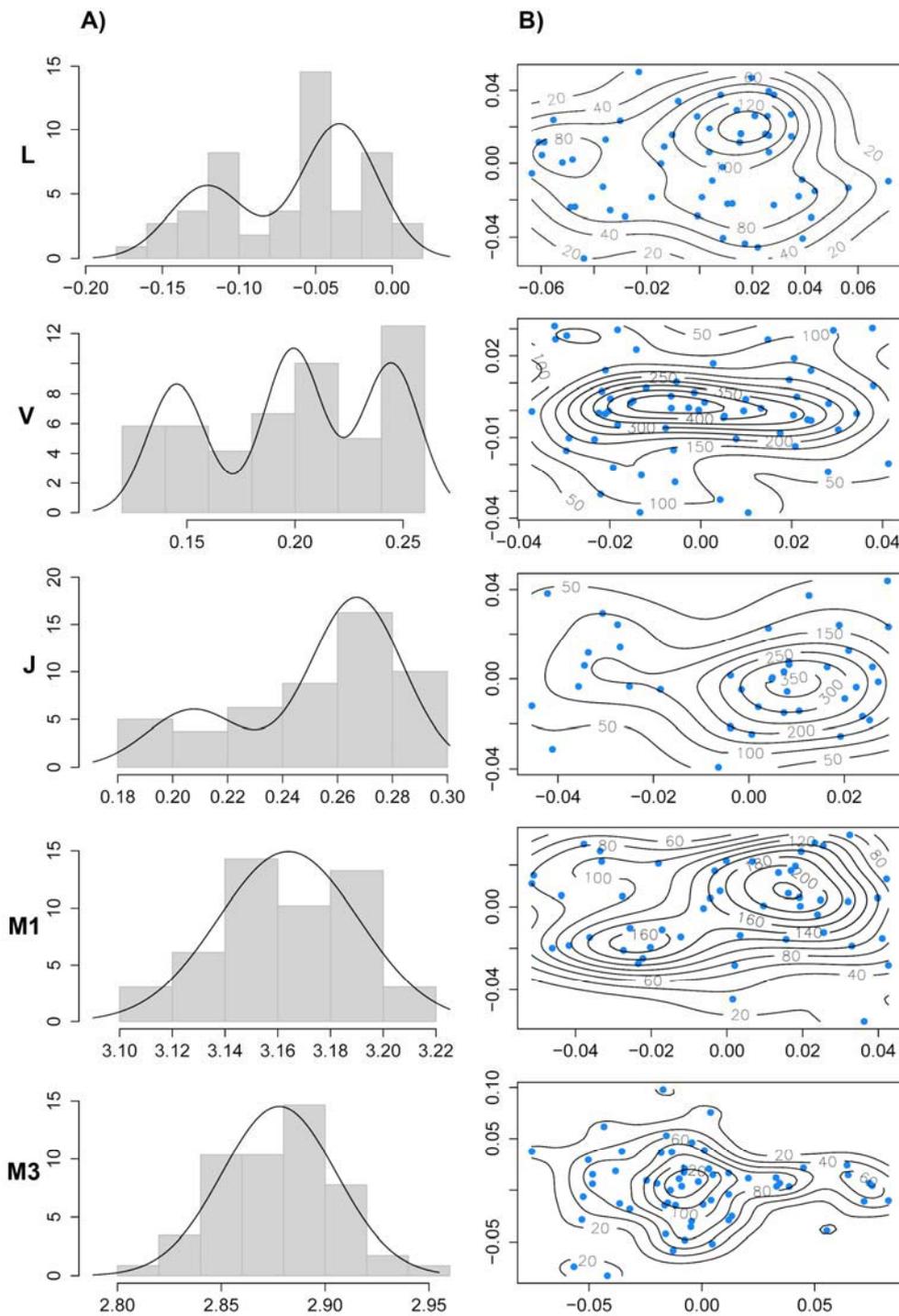


Figura 19. Clasificación de los individuos de las zonas de distribución alopatrica sin información *a priori* sobre la especie a la que pertenecen. **(A)** Densidad Kernel (KDE) en una dimensión con información del logCS para los cinco caracteres. **(B)** Densidad Kernel (KDE) en dos dimensiones con información de la matriz de forma (Y) para los cinco caracteres.

Tanto los resultados obtenidos mediante los métodos sin información *a priori* sobre la especie (KDE) como los que usan información *a priori* sobre la especie (análisis discriminante) detectaron, en general, la existencia de dos grupos. Por tanto, se concluyó que la asignación de especie por la distribución geográfica era correcta. En consecuencia, todos los análisis posteriores que se realizaron con los individuos de las zonas de distribución alopátrica usaron como referente de especie el correspondiente a cada individuo en función de su distribución geográfica.

2.1.1.2 Variación interespecífica e intraespecífica del tamaño

Tanto el resultado de la ANOVA como la representación de los diagramas de cajas mostraron que existe una variación interespecífica significativa, siendo *M. duodecimcostatus* más grande para todos los caracteres menos para el M3, en el que las dos especies son de igual tamaño. En la comparación intraespecífica, no se detectó una variación significativa entre las diferentes poblaciones de las que se disponían muestras de *M. duodecimcostatus* (figura 20, tabla 4).

Tabla 4. Análisis de la varianza (ANOVA). Variación interespecífica (sp) e intraespecífica (sp/pop) del logCS en los individuos de zonas de distribución alopátrica para los cinco caracteres. Se representa el valor del estadístico F de Fisher, con su *p-value* asociado y los grados de libertad del numerador y del denominador (g.l).

Modelo logCS~sp/pop	sp			sp/pop		
	F	<i>p-value</i>	g.l	F	<i>p-value</i>	g.l
L	70,01	<0,0001	1/51	0,12	0,87	2/51
V	83,44	<0,0001	1/56	1,50	0,23	2/56
J	100,43	<0,0001	1/36	1,23	0,30	2/36
M1	19,68	<0,0001	1/45	0,71	0,50	2/45
M3	1,14	0,2899	1/54	1,49	0,23	2/54

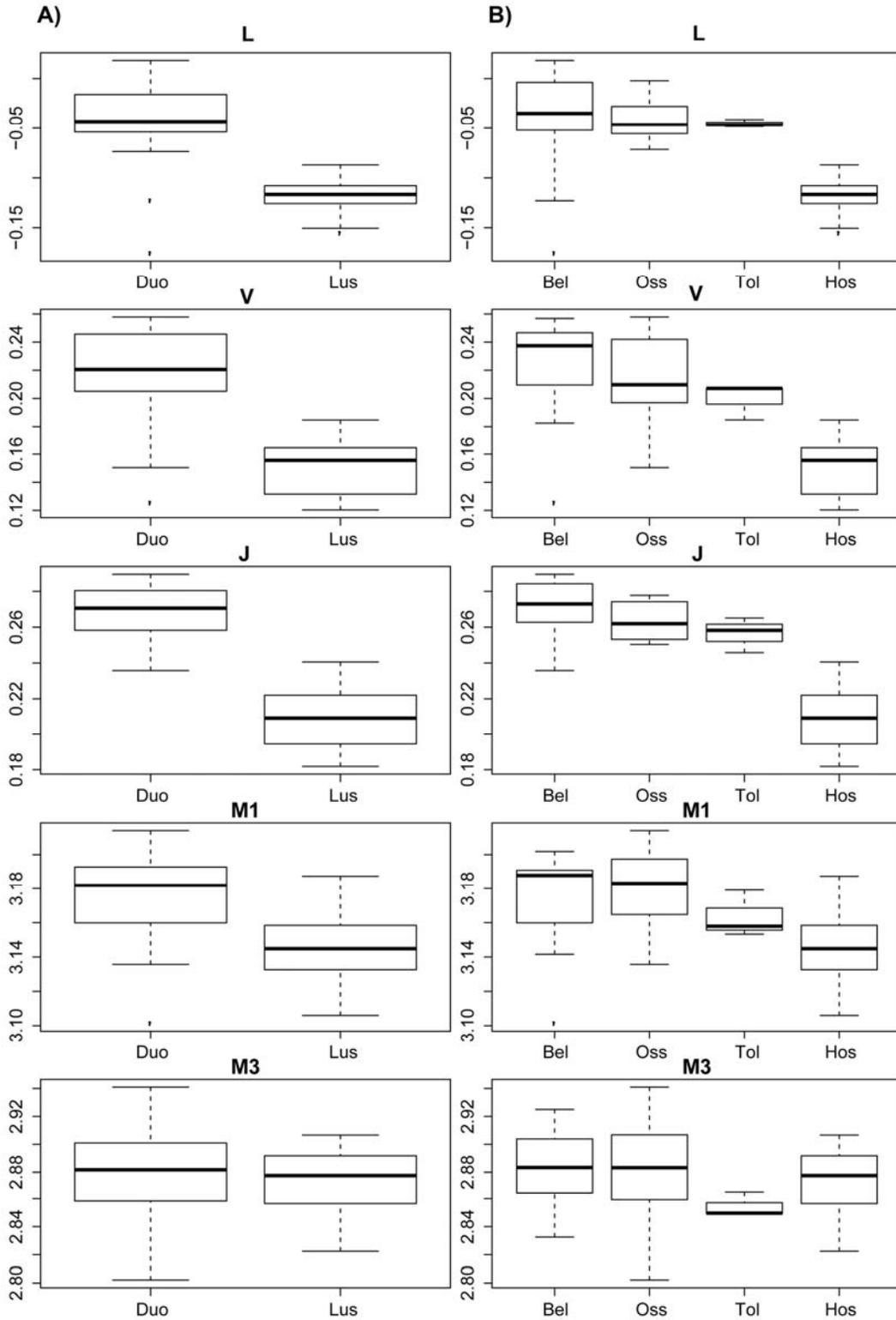


Figura 20. Valores de logCS del total de los individuos de las zonas distribución alopatrica para los cinco caracteres. (A) Comparación entre las dos especies de topillos (*M. duodecimcostatus*, Duo y *M. lusitanicus*, Lus). (B) Comparación entre las localidades de muestreo. Se representa la media (línea negra) y el primer y el tercer cuartil de la distribución de los valores.

2.1.1.3 Variación interespecífica e intraespecífica de la forma

Según lo resultados de la MANOVA (tabla 5) se observó que existe una variación interespecífica significativa de la forma para los todos los caracteres. Del mismo modo se comprobó que para el L, V y M3 las variaciones de forma pueden estar afectadas por procesos alométricos, ya que el logCS produce una variación significativa de la forma. Igualmente, la variación intraespecífica de la forma es significativa para los caracteres V, L y J entre las diferentes poblaciones de las que se disponían muestras de *M. duodecimcostatus*.

Tabla 5. Análisis multivariante de la varianza (MANOVA). Variación de la forma (Y) de los cinco caracteres de manera interespecífica (sp), intraespecífica (sp/pob) y en función del logCS para los individuos de zonas de distribución alopátrica. Se representa el valor aproximado del estadístico F de Fisher, con su *p-value* asociado y los grados de libertad del numerador y del denominador (g.l).

Modelo Y~logCS+sp/pob	sp			logCS			sp/pob		
	F	<i>p-value</i>	g.l	F	<i>p-value</i>	g.l	F	<i>p-value</i>	g.l
L	26,51	<0,0001	1/50	6,01	0,00011	1/50	4,45	<0,0001	2/50
V	15,86	<0,0001	1/55	3,56	0,00046	1/55	1,48	0,071	2/55
J	13,83	<0,0001	1/35	1,55	0,20	1/35	2,46	0,0055	2/35
M1	12,27	0,00022	1/44	1,58	0,24	1/44	1,41	0,20	2/44
M3	3,91	0,00027	1/53	1,94	0,044	1/53	1,14	0,31	2/53

Los resultados del análisis relativo de deformación (RW) indicaron una clara separación de las dos especies para 4 de los 5 caracteres analizados (todos menos el M3) (figura 21A). Entre el primer y el segundo componente del análisis se recogió entre el 66% (para L) y el 32% (para M3) de la variación de la forma.

Gracias a la representación gráfica de los *landmarks* se pudieron observar las variaciones en forma de cada especie para cada carácter (figura 21B, 21C). En esta representación gráfica se vio cómo quedaban representadas muchas de las características morfológicas tradicionales que se usan para la diferenciación entre las especies de topillos ibéricos.

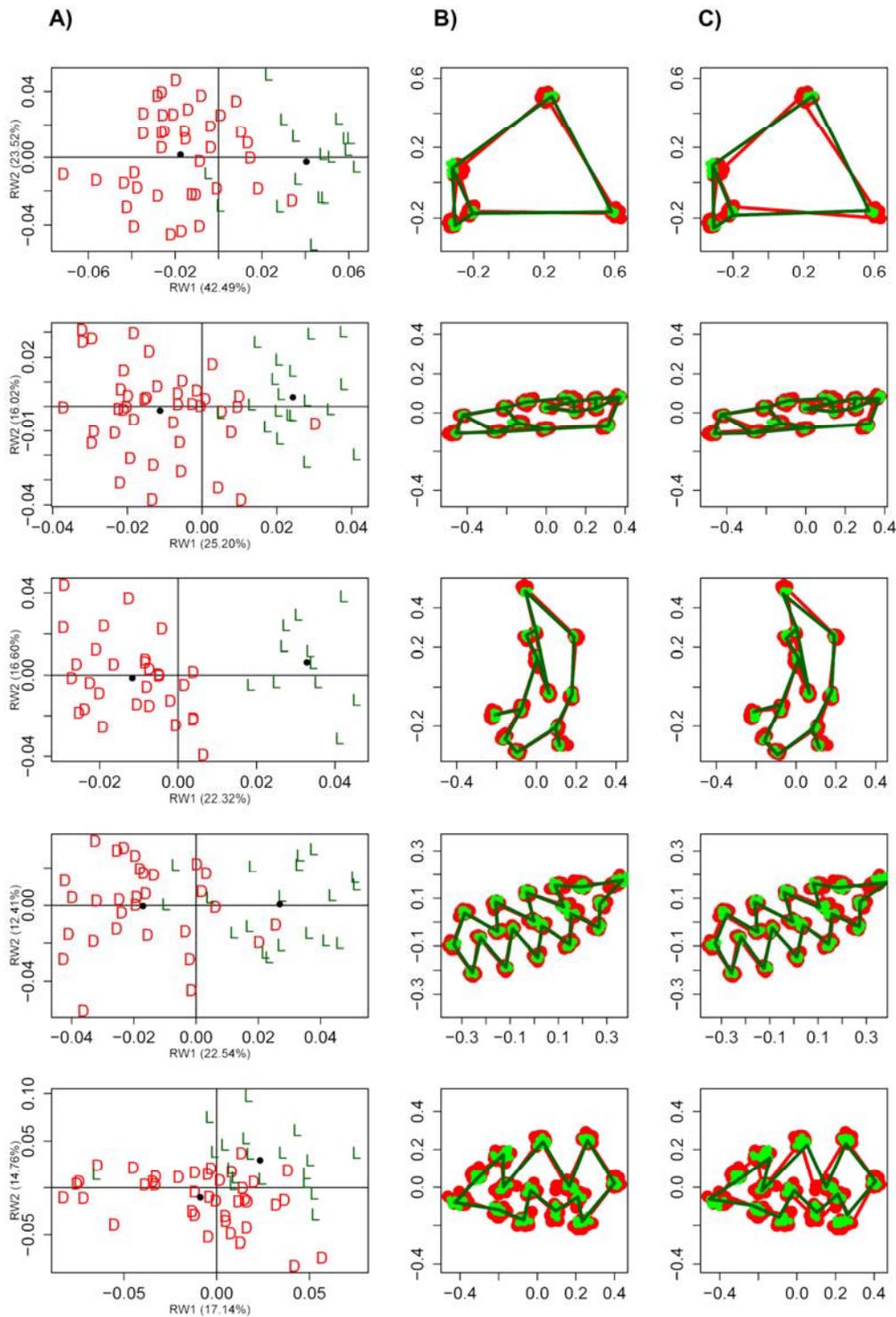


Figura 21. Análisis relativo de deformación para los cinco caracteres de los individuos de zonas de distribución alopatrica. (A) Representación del primer y del segundo componente de deformación (RW1 y RW2) para los cinco caracteres. En rojo, *M. duodecimcostatus* (D) y en verde *M. lusitanicus* (L). (B) Representación gráfica de la variación de la forma de los cinco caracteres. La línea que une los *landmarks* representa la forma media de cada especie. (C) Representación gráfica de la variación de la forma de los cinco caracteres. La línea que une los *landmarks* representa la forma del individuo más extremo, en forma, para cada especie.

IV. Resultados y Discusión

Para el L, se pudo observar como los dientes son de tipo más prognato en *M. duodecimcostatus* que en *M. lusitanicus*. Esta misma diferencia se recogió en el V. Para el J se observó la mayor robustez de la mandíbula de *M. duodecimcostatus* frente a la de *M. lusitanicus*. En el M1 se observó que el complejo anterocoronoide es más cerrado en *M. lusitanicus*, representado por una mayor cercanía entre los landmarks 2 y 20. En el M3 no se pudo reconocer una de las características utilizadas —mayor o menor prominencia del segundo triángulo bucal— en la identificación de especie para los representantes del género *Microtus* en la Península Ibérica. Se comprobó que efectivamente no existía una diferencia significativa interespecífica (valor de $F=0,0353$, $g.l=1/56$, $p\text{-value}=0,817$) entre el landmarks número 4 y la línea que unía los landmarks 2 y 6 (figura 22A). Por el contrario, sí que se observó que existía una diferencia en la forma global del M3 entre *M. duodecimcostatus* y *M. lusitanicus* (figura 22B).

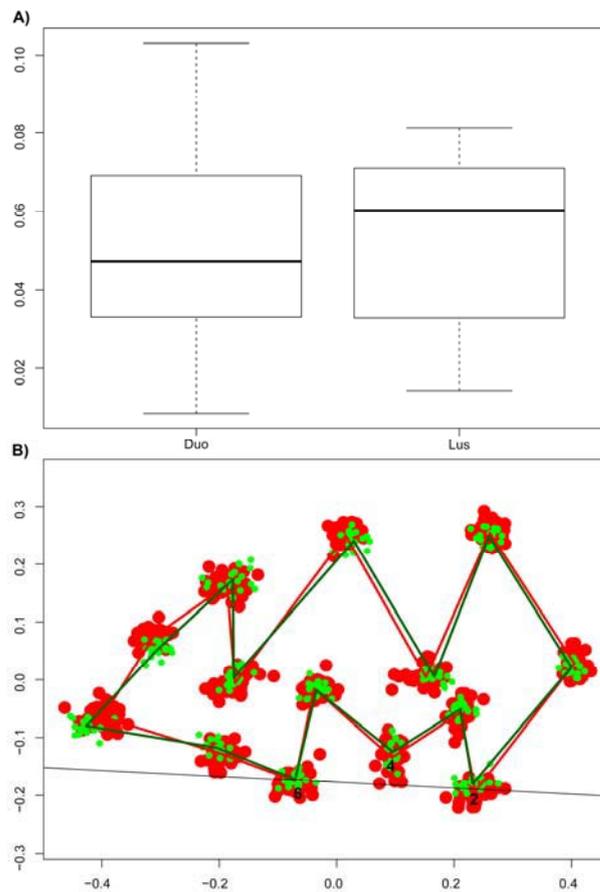


Figura 22. Distancia del landmark 4 a la línea entre el landmark 2 y el 6 para los individuos de las zonas de distribución alopatrica de las dos especies de topillos. **(A)** Valor de la distancia media (línea negra) y primer y tercer cuartil de la distribución de los valores para *M. duodecimcostatus*, Duo y para *M. lusitanicus*, Lus. **(B)** Representación gráfica de los landmarks de M3. Los landmarks rojos corresponden al topillo mediterráneo y los verdes al topillo lusitano.

2.1.2 Identificación y caracterización morfométrica de las dos especies de topillos en las zonas de distribución solapada

2.1.2.1 Asignación de especie

En el proceso de asignación de especie de los individuos de las zonas de distribución solapada se obtuvieron diferentes resultados. Se ha de tener en cuenta que la información relativa a tamaño y forma de estos individuos no se analizó de forma separada a la información de los individuos de distribución alopátrica sino que se realizó un análisis con toda la información de forma unificada.

En primer lugar, la agrupación de individuos sin información *a priori* sobre la especie basándose en el logCS de cada carácter indicó que sólo para V se podían diferenciar dos grupos de individuos. Para el resto de los caracteres se obtuvo una distribución unimodal de los datos (figura 23A). Para el caso de la forma se observó, para L, V y J, una clara disposición de los individuos en dos grupos, situación no reflejada ni para M1 ni para M3 (figura 23B).

Posteriormente se procedió a la asignación de especie a los individuos de las zonas de distribución solapada mediante un análisis discriminante. Los resultados obtenidos se pueden dividir en tres partes.

Primero, el análisis discriminante produjo una perfecta discriminación de los 32 individuos de referencia incluidos en el análisis (21 *M. duodecimcostatus* y 11 *M. lusitanicus*), coincidiendo completamente con la información *a priori* sobre la especie de cada individuo.

Segundo, con los individuos pertenecientes a zonas de distribución alopátrica a los que les faltaba información de alguno de los caracteres (28 individuos) se comprobó la validez del método de predicción de especie (tabla A6). De los 28 individuos, a 27 se les asignó la misma especie a la que pertenece por distribución geográfica. El único individuo, en el que la asignación fue diferente fue un individuo de Cataluña (C1210).

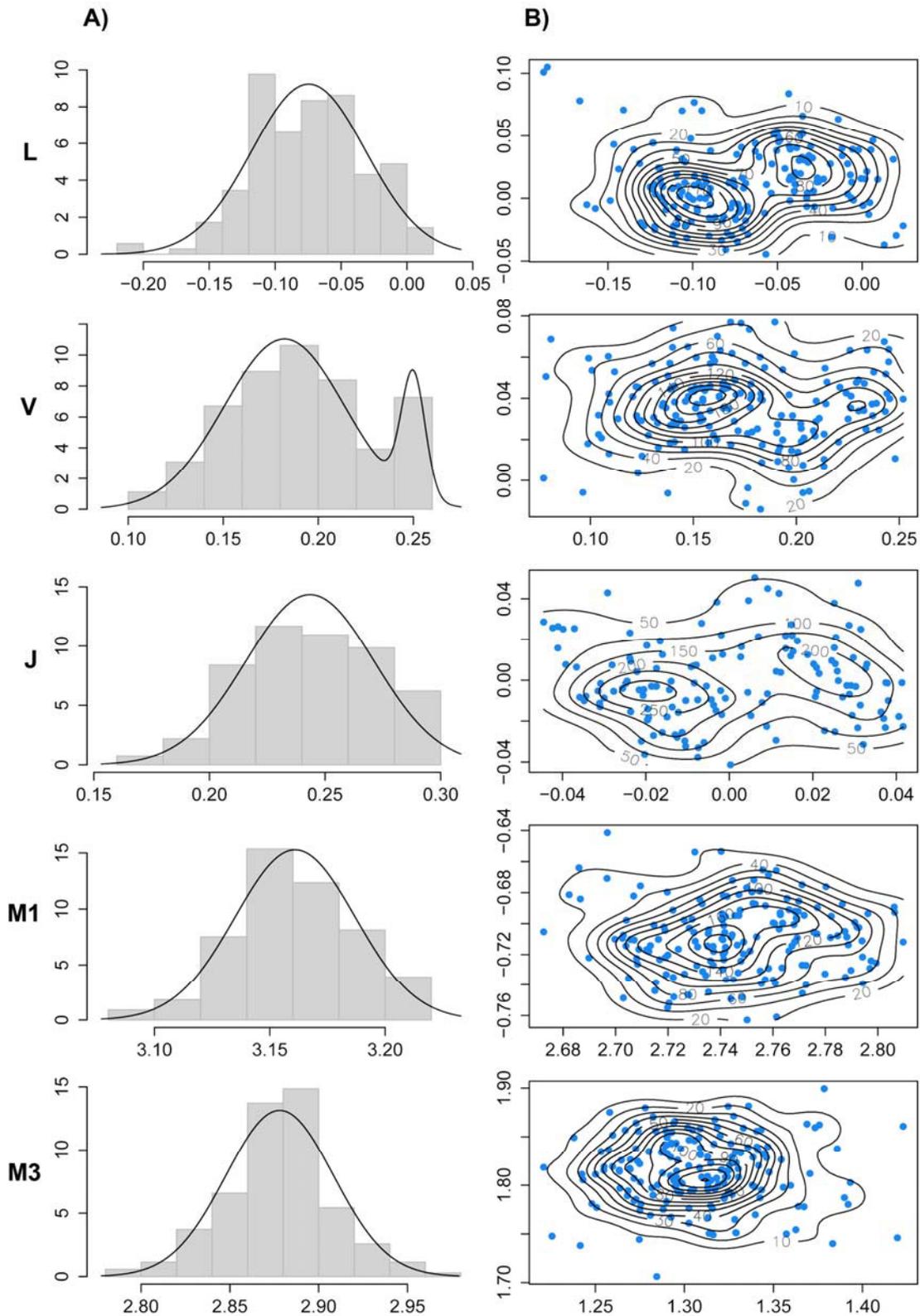


Figura 23. Clasificación de los individuos de las zonas de distribución alopátrica y de las zonas de distribución solapada sin información *a priori* sobre la especie a la que pertenecen. **(A)** Densidad Kernel (KDE) en una dimensión con información del logCS para los cinco caracteres. **(B)** Densidad Kernel en dos dimensiones con información de la matriz de forma (Y) de los cinco caracteres.

Posteriormente se asignó la especie a los 119 individuos de zonas de distribución solapada (tabla A6). Asociado a la especie que le correspondió a cada individuo se estimó el coeficiente de asignación posterior. Para la mayoría de los 119 individuos su coeficiente es máximo para una de las dos especies (99%). Solamente para tres individuos (C1948, C1250 y C1454) el coeficiente es menor del 99% para la especie que le correspondió (72%, 90% y 89%, respectivamente) (tabla A6).

Por último se obtuvieron los resultados del análisis de agrupación de individuos basados en algoritmos E-M. Se obtuvo que el modelo con un mayor BIC fue EEI con 2 grupos (BIC= 8496,669). Para cada carácter se obtuvo una representación gráfica de la distribución de los individuos en el espacio según su forma de acuerdo al modelo anteriormente calculado (figura 24A). Seguidamente se comprobó la coincidencia entre los resultados del LDA (en la asignación de especie) con los resultados de agrupación del análisis E-M (figura 24B, tabla A6). Para algunos individuos (C1210, C1252 y C1296) se produjo una discordancia en la asignación de especie mediante el LDA y la agrupación mediante el E-M. Esta discordancia era esperable ya que estos individuos se situaron en las zonas del espacio forma donde la probabilidad de una asignación errónea de especie es más alta (figura 24C). Por tanto, para el caso del individuo C1210, debido a que provenía de una zona de distribución alopátrica propia de *M. duodecimcostatus*, se le asignó la especie que le pertenecía por distribución geográfica. Los otros dos individuos no obtuvieron una asignación definitiva de especie, por lo que para los análisis posteriores a cada uno se le asignó una de las dos especies. Al resto de individuos se les asignó la especie según los resultados obtenidos con el LDA — confirmados por el algoritmo E-M— (tabla A6) y fueron considerados como tal en los sucesivos análisis.

Igualmente que para el LDA, se obtuvieron en el análisis E-M los coeficientes de agrupación posterior a cada uno de los grupos para cada individuo (tabla A6). En comparación con los coeficientes obtenidos para el LDA son menos discriminantes en cuanto a la proporción de pertenencia a un grupo (en este caso especie), con la existencia de muchos individuos con coeficientes intermedios para las dos especies.

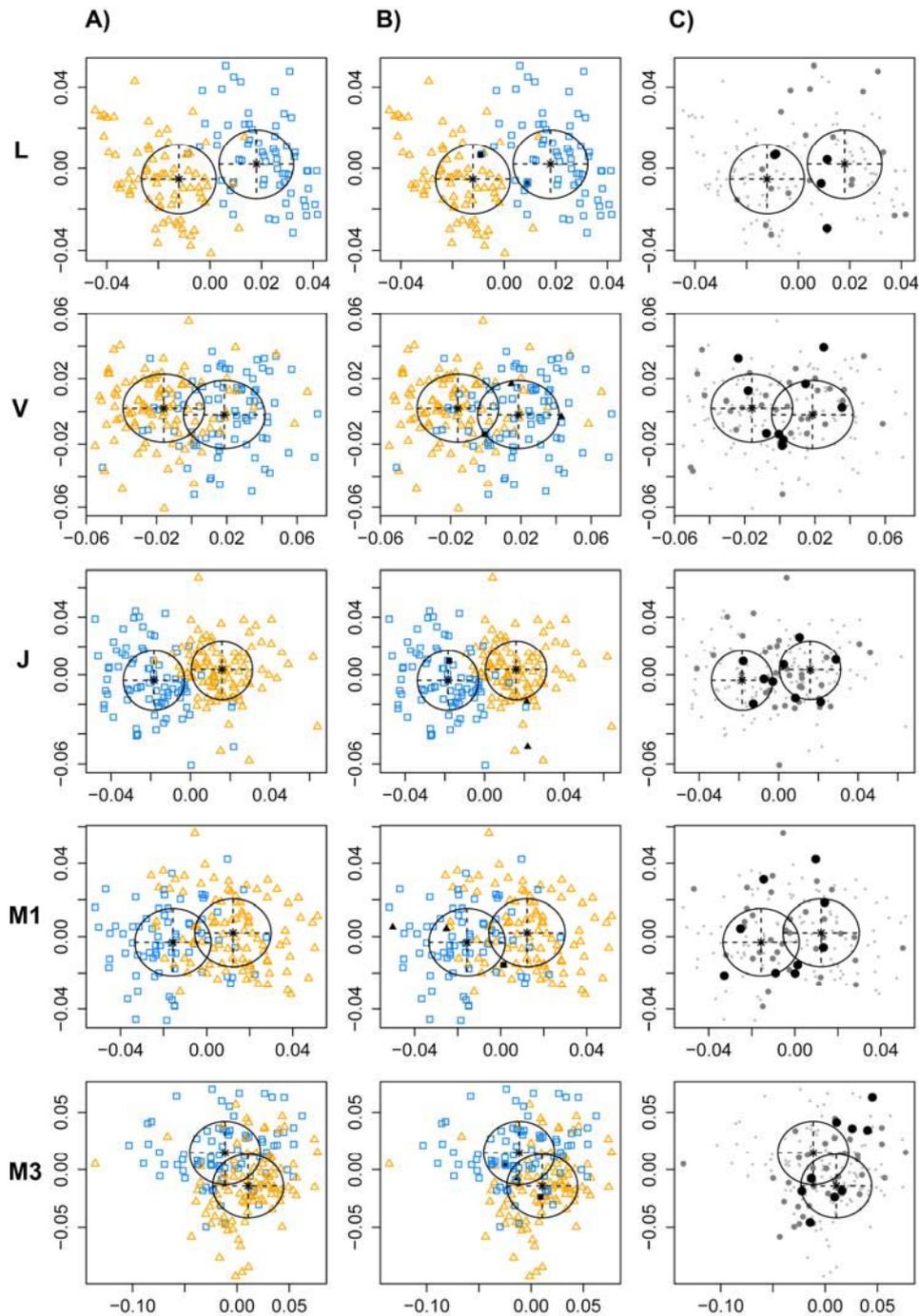


Figura 24. Clasificación de los individuos de las zonas de distribución alopátrica y de las zonas de distribución solapada de las dos especies, para los cinco caracteres, mediante algoritmos E-M (*expectation-maximization*). **(A)** Clasificación jerárquica aglomerativa de los individuos sin información *a priori* sobre la especie a la que pertenecen. Los colores se corresponden a los dos grupos que propone el modelo. **(B)** Comparación entre la asignación de especie obtenida por el LDA y por el algoritmo E-M. Los dos triángulos negros y el cuadrado negro se corresponden con individuos a los que se les asignó una especie diferente con los dos métodos (C1210, C1252 y C1296). **(C)** Representación en el espacio forma, de las zonas de incertidumbre en la asignación de especie. Del negro al gris claro, zonas de mayor probabilidad de una asignación de especie errónea a zonas de una mayor probabilidad de asignación de especie correcta.

2.1.2.2 Variación interespecífica e intraespecífica del tamaño

De forma general, y como ya se comentó para los individuos de las zonas distribución alopátrica, *M. duodecimcostatus* es más grande que *M. lusitanicus* para todos los caracteres, siendo la diferencia menor para el caso del M3 (figura 25A, tabla 6).

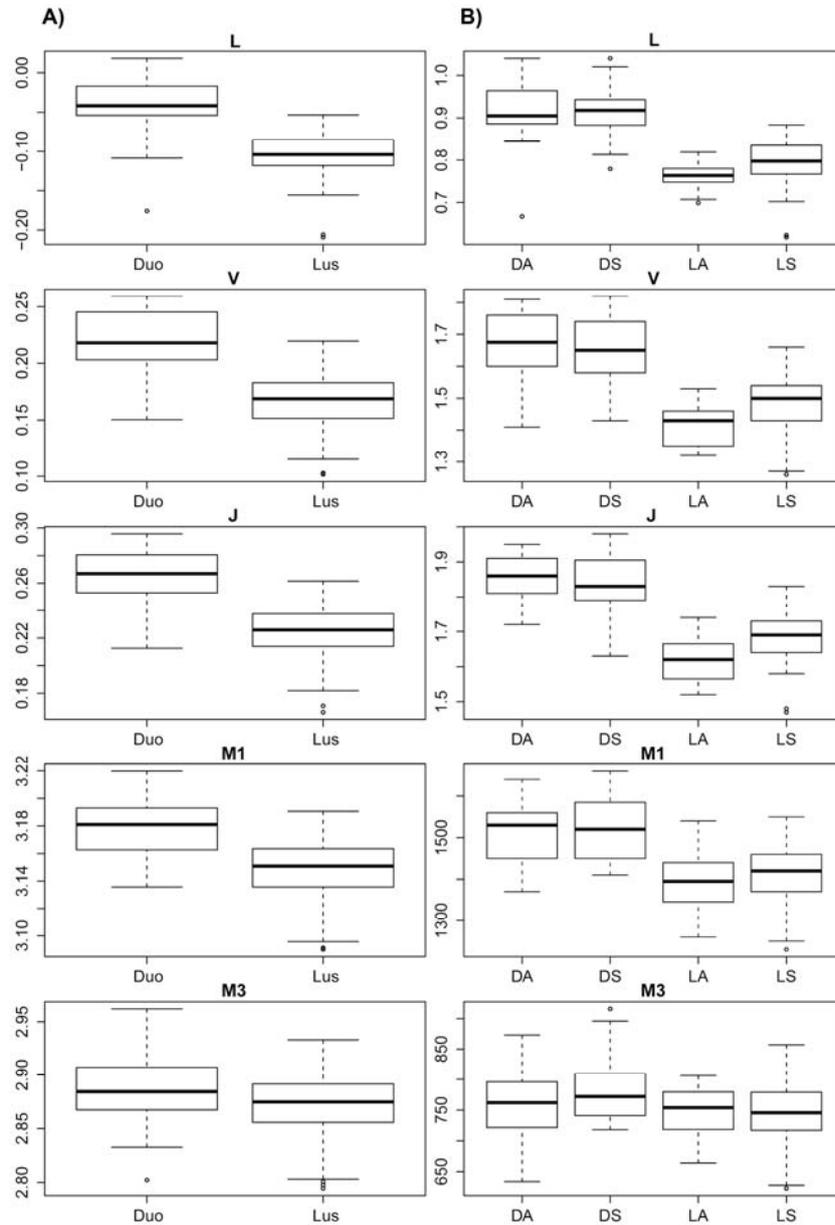


Figura 25. Valores de logCS del total de los individuos de las zonas distribución alopátrica y de las zonas de distribución solapada para los cinco caracteres. **(A)** Comparación entre las dos especies de topillos (*M. duodecimcostatus*, Duo y *M. lusitanicus*, Lus). **(B)** Comparación entre las dos especies en función de su distribución geográfica (*M. duodecimcostatus* de distribución alopátrica, DA; *M. duodecimcostatus* de distribución solapada, DS; *M. lusitanicus* de distribución alopátrica, LA y *M. lusitanicus* de distribución solapada, LS). Se representa la media (línea negra) y el primer y el tercer cuartil de la distribución de los valores.

Tabla 6. Análisis de la varianza (ANOVA). Variación interespecífica (sp) e intraespecífica según su distribución geográfica (sp*geo) del logCS en todos los individuos de las dos especies. Se representa el valor del estadístico F de Fisher, con su *p-value* asociado y los grados de libertad del numerador y del denominador (g.l).

Modelo logCS~sp*geo	sp			sp*geo		
	F	<i>p-value</i>	g.l	F	<i>p-value</i>	g.l
L	222,14	< 0,0001	1/170	0,13	0,034	1/170
V	198,67	< 0,0001	1/175	8,71	0,0036	1/175
J	180,06	< 0,0001	1/133	8,32	0,0046	1/133
M1	95,55	< 0,0001	1/162	0,11	0,74	1/162
M3	13,93	0,00026	1/171	2,39	0,12	1/171

En cuanto a la variación intraespecífica entre zonas alopátricas y solapadas, la información del análisis de la varianza (ANOVA) y de los diagramas de cajas indica que existe una diferencia en tamaño en los *M. lusitanicus* con una distribución en las zonas de distribución solapada, siendo más grandes que sus conoespecíficos de las zonas de distribución alopátrica para L, V y J, situación que no se observó para M1 y M3. Por el contrario, *M. duodecimcostatus* no sufre una variación en tamaño en función de su área de distribución (figura 25B, tabla 6).

2.1.2.3 Variación interespecífica e intraespecífica de la forma

Al analizar la variación de la forma de los cinco caracteres, mediante el análisis multivariante de la varianza (teniendo en cuenta la información de los individuos tanto de las zonas de distribución alopátrica como de las zonas de distribución solapada), se observó que existe una variación interespecífica significativa para los cinco caracteres. Del mismo modo, existe un efecto de alometría significativo para los cinco caracteres. Por último, se observó también para los cinco caracteres una variación significativa intraespecífica de la forma en función de la distribución geográfica de los individuos (zonas de distribución alopátrica o zonas de distribución solapada) (tabla 7).

Tabla 7. Análisis multivariante de la varianza (MANOVA). Variación de la forma (Y) de los cinco caracteres de manera interespecífica (sp), intraespecífica según su distribución geográfica (sp*geo) y en función del logCS en todos los individuos de las dos especies. Se representa el valor aproximado del estadístico F de Fisher, con su *p-value* asociado y los grados de libertad del numerador y del denominador (g.l).

Modelo Y~logCS+sp*geo	sp			logCS			sp*geo		
	F	<i>p-value</i>	g.l	F	<i>p-value</i>	g.l	F	<i>p-value</i>	g.l
L	19,48	< 0,0001	1/169	25,19	< 0,0001	1/169	6,33	< 0,0001	1/169
V	10,91	< 0,0001	1/174	27,85	< 0,0001	1/174	2,54	0,00024	1/174
J	14,75	< 0,0001	1/132	21,21	< 0,0001	1/132	2,99	< 0,0001	1/132
M1	7,96	< 0,0001	1/161	6,11	< 0,0001	1/161	3,06	< 0,0001	1/161
M3	11,81	< 0,0001	1/170	2,47	0,00051	1/170	2,39	0,00082	1/170

Los resultados del análisis relativo de deformación (RW) indican una menor diferenciación interespecífica que para el caso de la comparación de los individuos de zonas de distribución alopátrica, ya que el primer y segundo componente recogieron una menor variación de la forma, de entre el 60% (L) al 30% (M3). Esta situación queda igualmente reflejada a nivel gráfico, donde se observó una menor diferencia entre las nubes de puntos de cada una de las especies (figura 26A). La representación gráfica de los *landmarks* del mismo modo recogió la existencia de una mayor similitud interespecífica, a nivel de forma, entre los individuos de las zonas de distribución solapada para los cinco caracteres (figura 26C), que entre los individuos de las zonas de distribución alopátricas (figura 26B).

A pesar de esta mayor similitud en forma entre los individuos de las zonas de distribución solapada de las dos especies, los caracteres tradicionalmente utilizados para diferenciar las dos especies se siguieron identificando como ocurría para el caso de los individuos de las zonas de distribución alopátrica (*apartado 2.1.1.3*) (figura 26B). Del mismo modo que para los individuos de las zonas de distribución alopátrica, no se observó —para el total de los 179 individuos— ni una diferencia interespecífica (sp) significativa (valor de $F=0,7259$, $g.l=1/171$, $p-value=0,3954$) (figura 27A) ni intraespecífica (sp*geo) (valor de $F=0,4373$, $g.l=1/171$, $p-value=0,5093$) en la prominencia del segundo triángulo bucal del M3, aunque sí que se mantuvo la diferencia global en forma entre las dos especies (figura 27B).

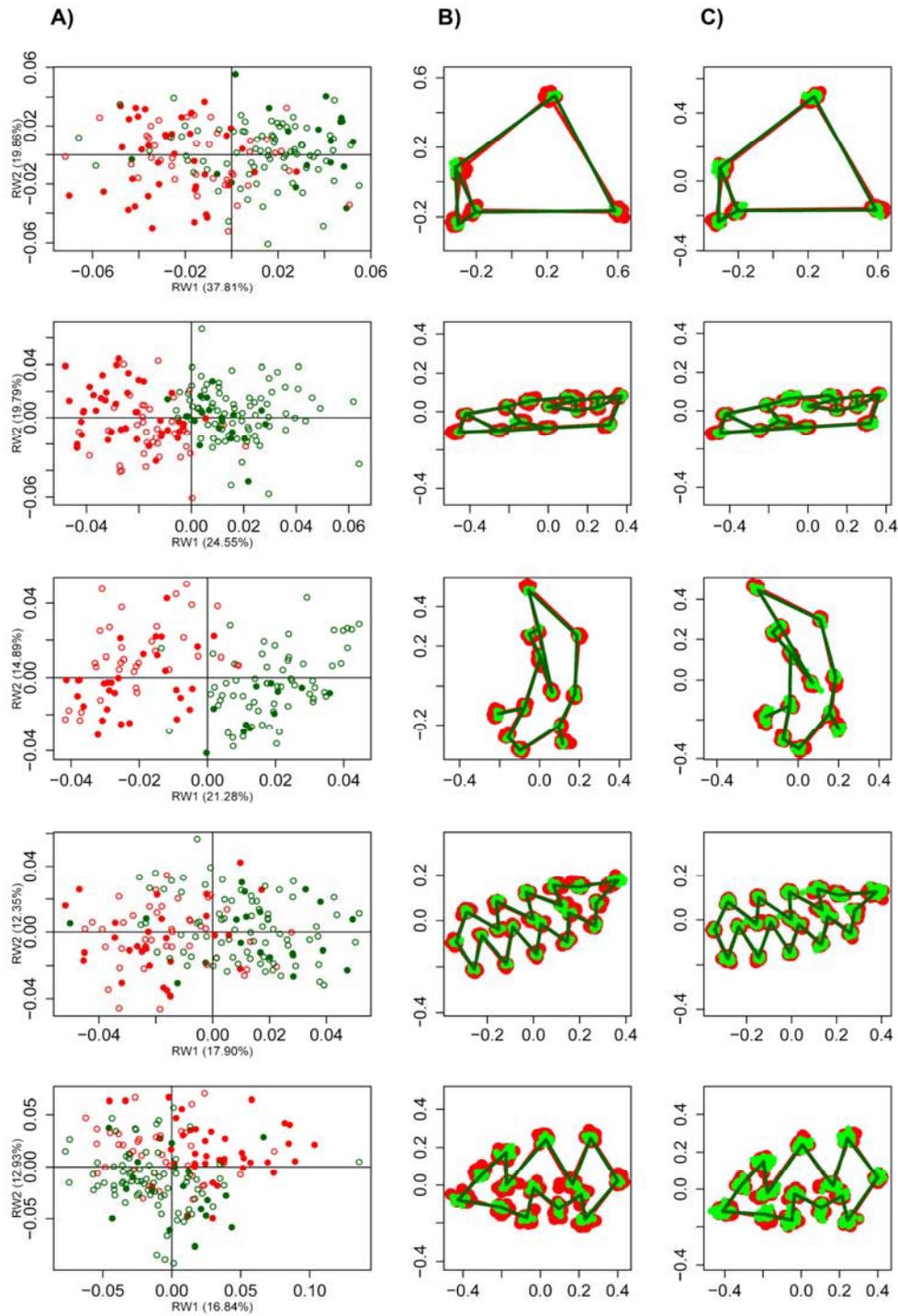


Figura 26. Análisis relativo de deformación para los cinco caracteres de los individuos de zonas de distribución alopatrica y de las zonas de distribución solapada de las dos especies de topillos. **(A)** Representación del primer y del segundo componente de deformación (RW1 y RW2) para los cinco caracteres. En rojo *M. duodecimcostatus*, en verde *M. lusitanicus*. Puntos, individuos de zonas de distribución alopatrica; círculos, individuos de zonas de distribución solapada. **(B)** Representación gráfica de la variación de la forma de los cinco caracteres. La línea que une los *landmarks* representa la forma media de los individuos de zonas de distribución alopatrica para cada especie. **(C)** Representación gráfica de la variación de la forma de los cinco caracteres. La línea que une los *landmarks* representa la forma media de los individuos de zonas de distribución solapada para cada especie

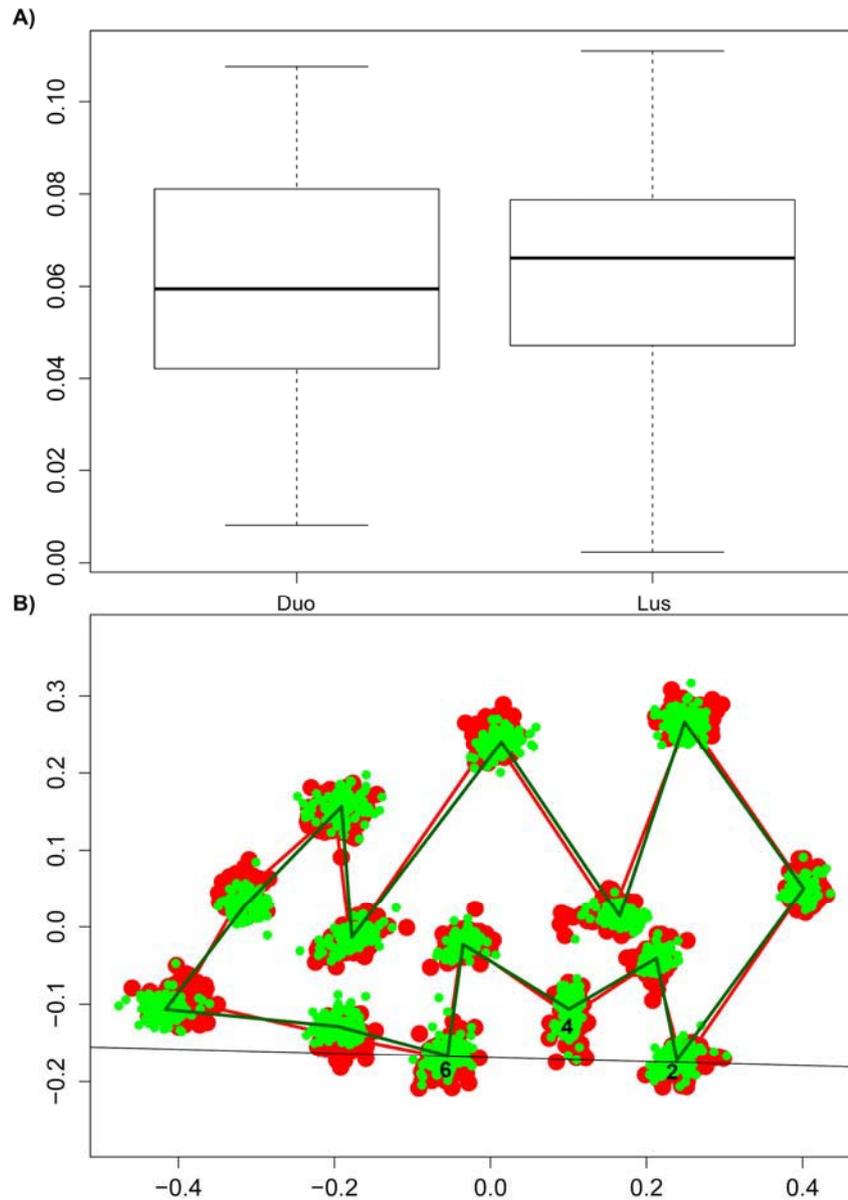


Figura 27. Distancia del *landmark* 4 a la línea entre el *landmark* 2 y el 6 para los individuos de las zonas de distribución alopátrica y solapada de las dos especies de topillos. **(A)** Valor de la distancia media (línea negra) y primer y tercer cuartil de la distribución de los valores para *M. duodecimcostatus*, Duo y para *M. lusitanicus*, Lus. **(B)** Representación gráfica de los *landmarks* de M3. Los *landmarks* rojos corresponden al topillo mediterráneo y los verdes al topillo lusitano.

De forma cuantitativa, los resultados obtenidos del análisis de la variación de los vectores de forma en el espacio reflejaron conclusiones similares a las anteriormente comentadas. Se observó una variación interespecífica significativa en el módulo del vector al comparar los individuos de las zonas de distribución alopátrica y las de distribución solapada. Es decir, para cuatro de los cinco caracteres (todos menos V) el módulo del vector era más pequeño al comparar *M. duodecimcostatus* con *M. lusitanicus* de las zonas de distribución solapada que de las zonas de distribución alopátrica (tabla 8A). El cambio más pronunciado se registró para M3, siendo este carácter donde se registra una mayor diferencia en el módulo del vector. Por otro lado, la variación intraespecífica sólo fue significativa para M3. Es decir, la cantidad de cambio que cada una de las especies sufre al comparar individuos de zonas alopátricas y solapadas es la misma menos para M3, donde *M. duodecimcostatus* presenta una mayor variación en el módulo del vector que *M. lusitanicus* (tabla 8B).

Tabla 8. Cuantificación de la variación del módulo de los vectores de forma en el espacio para los cinco caracteres. **(A)** Variación interespecífica. Se compara la variación del módulo entre *M. duodecimcostatus* y *M. lusitanicus* de las zonas de distribución alopátrica (DA-LA) y la variación del módulo entre *M. duodecimcostatus* y *M. lusitanicus* de las zonas de distribución solapada (DS-LS). Se representa el *p-value* asociado a cada cálculo. **(B)** Variación intraespecífica. Se compara la variación del módulo entre *M. duodecimcostatus* de las zonas de distribución alopátrica con los de las zonas de distribución solapada (DA-DS) y la variación del módulo entre *M. lusitanicus* de las zonas de distribución alopátrica con los de las zonas de distribución solapada (LA-LS). Se representa el *p-value* asociado a cada cálculo.

A)

Caracteres	Módulo DA-LA	Módulo DS-LS	<i>p-value</i>
L	0,05517	0,03558	0,008
V	0,03970	0,03727	0,564
J	0,04644	0,04264	0,001
M1	0,04392	0,02835	0,003
M3	0,05526	0,03752	0,02

B)

Caracteres	Módulo DA-DS	Módulo LA-LS	<i>p-value</i>
L	0,02472	0,01937	0,339
V	0,01953	0,01995	0,922
J	0,02369	0,01967	0,518
M1	0,02209	0,02140	0,842
M3	0,05023	0,02897	0,013

2.2 DISCUSIÓN

2.2.1 Individuos de las zonas de distribución alopátrica

Por lo que se refiere a los individuos de distribución alopátrica existe de manera general una diferencia de tamaño y forma entre las dos especies. Igualmente se podría concluir que los caracteres del cráneo (L y V) podrían ser firmes candidatos para la distinción entre especies, ya sea a nivel de tamaño como de forma. Del mismo modo J se presentaría como un buen candidato para esta diferenciación, pero el problema es que en muchas ocasiones es difícil recuperar las mandíbulas o se encuentran muy deterioradas. La capacidad discriminante de estos 3 caracteres (L, V y J) podría relacionarse con el hecho de que son estructuras que estas especies de topillos utilizan para excavar los diferentes terrenos. Debido a que estas dos especies tienden a ocupar diferentes hábitats en cuanto a la composición del substrato (Santos *et al.* 2009b) podrían presentarse modificaciones propias de especie en estos caracteres que les permitirían una mayor adaptación a su ambiente.

Los resultados del ANOVA para el logCS reflejaron de forma clara el mayor tamaño de *M. duodecimcostatus* para todos los caracteres respecto a *M. lusitanicus*. La distribución trimodal del logCS de V obtenida en el análisis KDE no refleja la existencia de dos únicas especies y propone la existencia de tres grupos diferentes. Pero al realizar una disposición de los individuos en esta distribución no se observó una correlación entre cada clase de tamaño y cada especie y sus grandes zonas de distribución (figura A4). Solamente se observa que los individuos de la población de Belianes (Cataluña) se distribuyen de manera preferencial en la clase de tamaño mayor. La existencia de estos individuos mucho más grandes hace que la distribución de densidad, en vez de ser bimodal (resultado esperado), se convierta en trimodal. El ANOVA no recoge esta diferencia intraespecífica porque la varianza dentro de cada población es tan grande que hace que la varianza total entre poblaciones no sea significativamente diferente de cero (tabla 4).

Se analizó la variación interespecífica de la forma mediante diferentes metodologías: MANOVA, análisis discriminante, análisis relativo de deformación y análisis KDE. Todas las metodologías sugieren que existe una diferencia significativa interespecífica a

nivel de forma. La diferencia de cada metodología radica en el hecho de que cada una indica cuáles serían los caracteres a tener en cuenta de manera preferencial para la distinción entre especies basándose en la forma.

Para el MANOVA los mejores caracteres serían J y M1 ya que el resto puede ver afectada su forma debido a la existencia de procesos alométricos (tabla 5). En el caso del LDA los caracteres que discriminan mejor entre especie son J y V (tabla 3). Por otro lado, el que recoge una mayor deformación para el RW1 en el análisis relativo de deformación es L (figura 21). Por último, la existencia de dos grupos de densidad de datos viene dada por L, J y M1 para el caso del análisis de KDE (figura 19).

2.2.2 Individuos de las zonas de distribución solapada

La predicción de especie obtenida fue especialmente satisfactoria con el LDA, cuyo resultado fue confirmado con el obtenido en el análisis de E-M (figura 24B). Es decir, la información contenida en la matriz de forma, con la información de los cinco caracteres a la vez, era lo suficientemente informativa para asignar de manera correcta la especie a la que pertenecía cada uno de los 119 individuos de las zonas de distribución solapada (tabla A6).

El análisis E-M permitió conocer cuáles eran las zonas de la distribución de la forma en el espacio, para cada carácter, donde la predicción de la especie es más difícil de realizar. Se obtuvo alrededor de cada nube de puntos que formaba un grupo (en este caso una especie) unos valores de incertidumbre para la clasificación (figura 24C). Los individuos en los que se produjo una discordancia en la asignación de especie entre el LDA y el análisis E-M se encontraban en estas áreas de incertidumbre. Por tanto, es importante tener en cuenta, a la hora de clasificar individuos, estas zonas de incertidumbre ya que los resultados de asignación de especie podrían verse afectados.

Para los coeficientes de clasificación *a posteriori* obtenidos del análisis E-M se observó que algunos individuos presentaban coeficientes intermedios entre *M. duodecimcostatus* y *M. lusitanicus* (tabla A6). A pesar de la existencia de individuos con proporciones intermedias, el análisis de E-M concluyó que la agrupación más plausible de los individuos eran dos grupos correspondientes a las dos especies (como se ha comentado

anteriormente) sin la existencia de un tercer grupo en el que se agruparan individuos de formas intermedias.

Se comprobó que existe una diferencia significativa en el tamaño de los individuos en función a la especie a la que pertenecen, pero la situación no resulta tan clara como para la comparación entre los individuos de las zonas de distribución alopátricas. En ese caso, los *M. lusitanicus* de las zonas de distribución solapada presentan un tamaño superior que sus conespecíficos de las zonas de distribución alopátrica. La distribución bimodal de las clases de tamaños observada para el caso de los individuos de las zonas de distribución alopátrica se transformó en una distribución unimodal, donde las clases de tamaño centrales quedarían ocupadas por los *M. lusitanicus* de las zonas de distribución solapada (figura 23A). Se observa este cambio en la distribución de los individuos para cuatro de las cinco caracteres (todos menos el ventral). En V se observa una distribución bimodal ya que se mantiene el grupo de individuos de la población de Belianes que presentan un tamaño netamente superior al resto de los individuos. Por tanto, el tamaño de estas estructuras no debería ser utilizado como única herramienta para la discriminación interespecífica en las zonas de distribución solapada ya que podría originar resultados erróneos en la distinción entre especies.

La variación interespecífica en forma se mantiene entre las dos especies al analizar los individuos de las zonas de distribución solapada. Es decir, se confirma que existe una diferenciación interespecífica constante a nivel de forma (para estos cinco caracteres) ya sea al comparar individuos de las zonas de distribución alopátrica como los de zonas de distribución solapada (figura 26A). Pero aún existiendo claras variaciones interespecíficas se observó que los individuos de las dos especies, en las zonas de distribución solapada, tienen una forma más parecida entre sí. Este resultado quedó reflejado en el análisis relativo de deformación, donde las nubes de individuos quedaban menos definidas para las dos especies (figura 26A) y los porcentajes de los dos primeros componentes son menores. Igualmente, la representación gráfica de los *landmarks* recogió este acercamiento en forma entre especies si se compara con los individuos de zonas de distribución alopátrica (figura 26B y 26C).

De manera cuantitativa, se confirmó esta situación mediante el cálculo del módulo del vector que expresa cantidad de cambio de forma. Se verificaba que existía una mayor

similitud en forma entre *M. duodecimcostatus* y *M. lusitanicus* de las zonas de distribución solapada que entre las dos especies en zonas de distribución alopátrica (tabla 8A). Es decir, se detecta un proceso de convergencia de la forma.

Se podría plantear la hipótesis de que esta situación se debe a la existencia de una forma más adecuada para los hábitats donde se distribuyen de forma solapada las dos especies de topillos y por tanto las dos especies la adoptan. Esta adaptación es perfectamente posible debido a la enorme plasticidad fenotípica que presentan estos animales. Además, al no encontrarse en una situación de competición por el nicho (Santos *et al.* 2010) podrían perfectamente adoptar la misma forma sin provocar una disminución en su eficacia biológica, sino más bien aumentándola (Spaeth 2009). Esta situación se reprodujo en cuatro de los cinco caracteres. Para el único donde no se cumplió fue en V, donde no se produce una variación significativa en forma entre las dos especies en una zona y otra de distribución. Este resultado se podría explicar debido a la existencia de fuertes presiones selectivas sobre la modificación de la estructura craneal de manera global.

Si se produce una convergencia en forma entre dos especies puede deberse a que una especie adopta la forma de la otra o que ambas adopten una forma intermedia. Esa última situación es la que puede haber sucedido entre estas dos especies de topillos. Al analizar la variación intraespecífica entre los individuos de zonas de distribución alopátrica y zonas de distribución solapada se observó que para ambas especies el módulo del vector de cambio de forma era el mismo para cuatro de los cinco caracteres (todos menos M3) (tabla 8B). Es decir, las dos especies han cambiado en la misma proporción su forma para adoptar una conformación fenotípica que es más similar entre ellas en las zonas de distribución solapada que la que presentan cada especie en las zonas de distribución alopátrica. Para el caso del M3, *M. duodecimcostatus* sufre un cambio de forma más pronunciado para conseguir la convergencia en forma del que sufre *M. lusitanicus*. Por tanto, M3 sería un carácter de gran interés para el análisis de variación intraespecífica de *M. duodecimcostatus*.

A pesar de la existencia de diferencias en tamaño y en forma entre estas dos especies de topillos, en muchas ocasiones son tremendamente sutiles sobre todo cuando se comparan individuos pertenecientes a las zonas de distribución solapadas (figura 26A y

26C). Por tanto, extraer información —del tamaño y de la forma— de diferentes estructuras permite maximizar la identificación de las diferencias existentes entre las especies y mejora considerablemente la asignación de especie a los diferentes individuos. Del mismo modo, el uso de la morfometría geométrica para cuantificar estas diferencias permite analizar el cambio de tamaño y de forma de las estructuras seleccionadas de forma integrativa permitiendo elevar la precisión de las clasificaciones de los individuos en una de las dos especies.

3. CARACTERIZACIÓN DE LAS RELACIONES FILOGENÉTICAS ENTRE *M. duodecimcostatus* y *M. lusitanicus* MEDIANTE EL USO DE SECUENCIAS MITOCONDRIALES Y NUCLEARES

3.1 RESULTADOS

3.1.1 Variabilidad interespecífica del citocromo b y de los intrones

La variabilidad interespecífica se analizó utilizando la información de las 184 secuencias mitocondriales y de los 431 alelos de las secuencias nucleares amplificadas (tabla 9). Para el intrón Agxt-5 se amplificaron 86 alelos pertenecientes a 43 individuos diferentes; para el intrón Catsper3-10, 54 alelos (32 individuos); para el intrón Dhcr24-7, 88 alelos (44 individuos); para el intrón Nadsyn1-4, 60 alelos (30 individuos); para el intrón Smo-9, 84 alelos (42 individuos) y para el intrón Wls-7, 59 alelos (30 individuos).

Se observó que tanto para el marcador mitocondrial como para los seis intrones amplificados *M. lusitanicus* presentaba una mayor diversidad nucleotídica y haplotípica respecto a *M. duodecimcostatus* (tabla 9). Del mismo modo, se obtuvo un mayor número de secuencias polimórficas para el caso del topillo lusitano (tabla 9). Por el contrario, no se detectaron diferencias significativas entre el número de secuencias heterocigotas amplificadas (medido como número de polimorfismos) a partir de las muestras de tejido fresco y a partir de las muestras de tejido óseo para todos los intrones para ninguna de las dos especies. (*M. duodecimcostatus*, Chi-cuadrado = 0,9308, g.l = 4, *p-value* = 0,9201; *M. lusitanicus*, Chi-cuadrado = 6,0701, g.l = 7, *p-value* = 0,5316).

En lo que se refiere a los test de neutralidad, se detectó una señal de posibles expansiones poblacionales para el caso de *M. lusitanicus* para 5 genes con el test de F_s de F_u , 2 de ellos corroborados con R2. El test de F_s de F_u para el citocromo *b* de *M. duodecimcostatus* también detectó una señal de una posible expansión poblacional (tabla 9).

Los resultados que se obtuvieron para los tests de recombinación fueron muy diferentes entre los dos métodos empleados. Mientras que el test de recombinación de los cuatro gametos (Rm) detectó eventos de recombinación para alguno de los loci nucleares, el estadístico ϕ_w no indicó la existencia de ningún evento de recombinación (tabla 9). Se deduce que el estadístico ϕ_w presenta una mayor capacidad de diferenciar entre eventos de mutaciones recurrentes a lo largo de las secuencias y procesos de recombinación (Bruen 2006).

Se reflejó la diversidad haplotípica de las secuencias mediante las genealogías de haplotipos construidas (figura 28). Los individuos del grupo DCNM (discordantes citomorfonuclear), se representaron en color gris. En la genealogía de haplotipos del marcador mitocondrial se observó que los individuos del grupo discordante presentaban haplotipos propios del clado formado por *M. duodecimcostatus* (figura 28G). Por el contrario, para las tres genealogías haplotípicas donde los intrones presentaban haplotipos propios de especie (Agxt-5, Catsper3-10 y Dhcr24-7), los individuos del grupo DCMN presentaban haplotipos del clado de *M. lusitanicus* (figura 28A, 28B, 28C). Para los otros tres intrones (Nadsyn1-4, Smo-9 y Wls-7) no se observó una diferenciación haplotípica por especie (figura 28D, 28E, 28F). Es destacable la diversidad haplotípica que se observó para el caso del intrón Catsper3-10, donde *M. lusitanicus* presentaba un elevado número de haplotipos únicos (16) frente a los dos haplotipos de *M. duodecimcostatus*. Igualmente este marcador nuclear es el que presentó un mayor número de diferencias (mutaciones) entre las dos especies (figura 28B).

Por último, se obtuvo que la distancia genética media entre *M. duodecimcostatus* y *M. lusitanicus* para el marcador mitocondrial —teniendo en cuenta la información de los 184 individuos— era del 3,2%.

IV. Resultados y Discusión

Tabla 9. Diversidad haplotípica y nucleotídica (π , Θ) nuclear y mitocondrial para las dos especies de topillos. PB: pares de bases totales; se indican las diferencias en longitud de los diferentes alelos. Tests de neutralidad para los marcadores nucleares y para el citocromo *b* (D Tajima, Fs de Fu y R2); se realizaron 1.000 simulaciones para obtener el *p-value*. Test de recombinación de los cuatro gametos (Rm) y estadístico ϕ_w para los marcadores nucleares. Para este último estadístico se muestran los *p-value* de la probabilidad de existencia de un proceso de recombinación en las secuencias reales y en 1.000 permutaciones. En negrita valores de *p-value* significativos. Duo, *M. duodecimcostatus*; Lus, *M. lusitanicus*.

Marcador	N	PB	H	S	π	Θ sitio	D Tajima	Fs de Fu	<i>p-value</i>	R2	<i>p-value</i>	Rm	ϕ_w
Agxt													1 1
Duo	30	442 441	2	1	0,001	0,001	1,578	1,672	0,880	0,254	0,986	0	
Lus	56	442 441	9	9	0,002	0,004	-0,177	-5,858	0,001	0,040	0,0349	0	
Catsper3													0,916 0,952
Duo	25	311 307	2	1	0,000	0,001	-1,157	-1,061	0,224	0,196	0,465	0	
Lus	29	311 291	16	18	0,016	0,016	-0,050	-5,258	0,017	0,114	0,434	3	
Dhcr24													0,462 0,714
Duo	30	340	7	5	0,003	0,004	-0,785	-3,136	0,008	0,092	0,150	1	
Lus	58	340	11	10	0,003	0,006	-1,499	-6,417	0,001	0,053	0,055	1	
Nadsyn1													0,9 1
Duo	22	500 491	3	2	0,001	0,001	-0,871	-0,870	0,271	0,121	0,193	0	
Lus	38	500 491	9	12	0,004	0,006	-0,821	-1,478	0,242	0,086	0,197	1	
Smo													0,079 0,17
Duo	30	398 384	5	4	0,001	0,003	-1,112	-2,120	0,045	0,072	0,053	0	
Lus	54	398	12	10	0,005	0,006	-0,250	-3,502	0,073	0,100	0,474	1	
Wls													0,599 0,851
Duo	23	367	5	4	0,002	0,003	-0,750	-1,550	0,101	0,102	0,122	0	
Lus	36	367	9	10	0,003	0,007	-1,885	-5,319	0,001	0,047	0,000	0	
Cit. b													
Duo	80	1143	37	73	0,007	0,013	-1,404	-1,177	0,008	0,056	0,066		
Lus	104	1143	69	146	0,025	0,031	-0,617	-1,556	0,01	0,080	0,365		

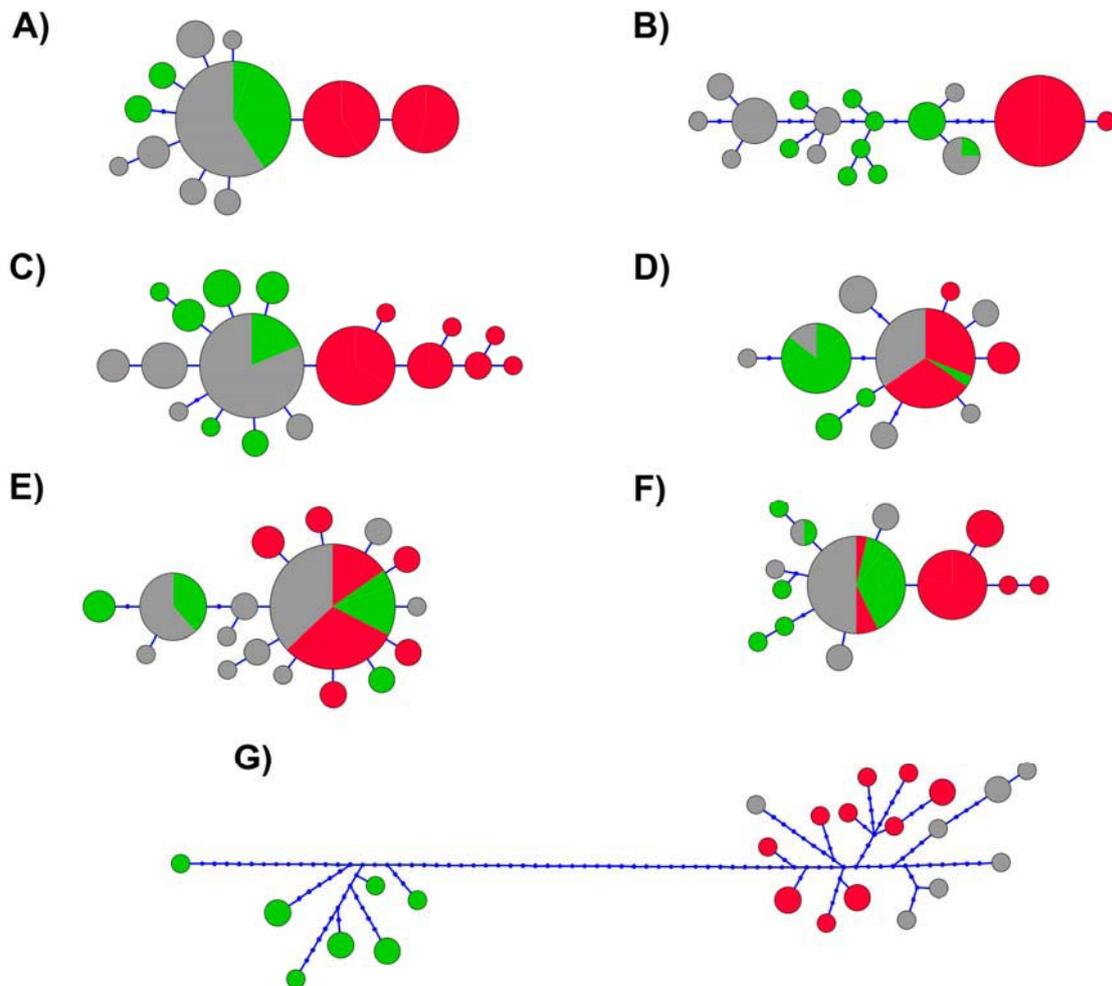


Figura 28. Genealogías haplotípicas de los seis intrones y del citocromo *b* amplificados en 44 individuos. El tamaño de los círculos es proporcional al número de individuos para cada haplotipo. Rojo, *M. duodecimcostatus*; verde, *M. lusitanicus*; gris, DCMN. (A) Agxt-intrón-10. (B) Catsper3-intrón-5. (C) Dhcr24-intrón-7. (D) Nadsyn1-intrón-4. (E) Smo-intrón-9. (F) Wls-intrón-7. (G) Citocromo *b*.

3.1.2 Análisis de la historia evolutiva intraespecífica e interespecífica: *Beast e IMA2

3.1.2.1 Uso de calibraciones fósiles en la filogenia del orden *Rodentia*

Se comprobó que para los cinco puntos de calibración utilizados las distribuciones efectivas conjuntas de los *priors* se encontraban dentro de las distribuciones de sus respectivos *priors* y de las distribuciones posteriores obtenidas en el análisis inicial (figura A7). Se concluyó por tanto que no existe ninguna interacción entre los *priors* de calibración utilizados y por tanto los parámetros obtenidos en el análisis, como las tasas de mutación, fueron calculados adecuadamente.

3.1.2.2 Cálculo de las tasas evolutivas de los intrones para el género *Microtus*

A partir de un árbol Bayesiano de secuencias de intrones de *Rodentia* y calibraciones fósiles tomadas de la literatura (figura 16) se obtuvo la media y la distribución estándar de la tasa evolutiva de cada intrón para el género *Microtus* (tabla 10).

Tabla 10. Tasas evolutivas, en sustituciones por posición por millón de años, de los seis intrones amplificados calculadas para el género *Microtus*.

Nombre intrón	Media	Desviación estándar
Agxt-intrón-10	0,00449	0,00131
Catsper3-intrón-5	0,00783	0,00172
Dhcr24-intrón-7	0,00445	0,00121
Nadsyn1-intrón-4	0,00547	0,00129
Smo-intrón-9	0,00594	0,00121
Wls-intrón-7	0,00541	0,00115

El intrón con la tasa de mutación (en unidades de sustituciones por posición por millón de años) más elevada fue Catsper3-5, y Dhcr24-7 el que presentó una menor tasa de mutación. Para el citocromo *b* no se estimó la tasa de mutación con varias especies del orden *Rodentia* debido a la saturación que presentaría este gen a estos niveles filogenéticos tan altos. Se utilizó, como tasa de mutación mitocondrial el valor de 0,019 sustituciones/posición·Ma (± 0.0045), calculado con calibraciones más recientes (Soria-Carrasco 2011). Por tanto, la tasa de mutación mitocondrial resultó ser un orden de magnitud más acelerada que cualquier de tasa de mutación de los intrones (tabla 10).

Estas tasas de mutaciones fueron las empleadas como *priors* en los análisis de *Beast e IMA2. De estos análisis, se seleccionó uno de *Beast y otro de IMA2, en el que se incluyeron todas las secuencias mitocondriales y nucleares del grupo concordante, CCMN, (análisis 1, tabla A11) y se obtuvieron las distribuciones posteriores de estas tasas de mutación calculadas en los análisis y se compararon con las estimadas e introducidas como *priors* en los análisis (figura 29).

Para todos los intrones la coincidencia entre las tasas estimadas y las distribuciones posteriores de cada tasa, para cada análisis, fue muy elevada (figura 29A). Para el caso

de dos intrones, *Catsper3-5* y *Dhcr24-7*, la estimada (mediante el uso de calibraciones fósiles) fue algo más baja que las obtenidas *a posteriori*. Para el resto de los cuatro intrones las tasas de mutación estimadas fueron más altas que las obtenidas *a posteriori* (figura 29A). Para el caso del citocromo *b* se observó que la tasa de mutación empleada era mucho más baja que las obtenidas *a posteriori* en los dos análisis (figura 29B). La obtenida con *Beast era casi dos veces superior y la obtenida con IMa2 casi tres veces superior.

3.1.2.3 Estima del tiempo de divergencia entre las dos especies de topillos

Se estimó con *Beast que *M. duodecimcostatus* y *M. lusitanicus* divergieron hace 0,6487 (0,3601-0,9791) millones de años, utilizando toda la información nuclear disponible (CCMN y DCMN) y la información mitocondrial únicamente del grupo CCMN (concordante cito-morfonuclear) (análisis 1, tabla A11). Se eligió esta metodología (*Beast) y este análisis para esta estima por diversos motivos. Primero, se empleó para su cálculo tanto los marcadores nucleares como el marcador mitocondrial recogiendo así la mayor información posible de los eventos de coalescencia (Sánchez-Gracia & Castresana 2012). Segundo, no incluir las secuencias mitocondriales del grupo DCMN en el análisis se consideró fundamental ya que la discordancia cito-nuclear podía provocar estimas erróneas del tiempo de especiación. En cuanto a la metodología, *Beast es más preciso en la estima de los parámetros ya que dispone de la posibilidad de implementar los modelos evolutivos y tipos de relojes moleculares más adecuados para los marcadores utilizados. Por último, se pudo implementar un modelo de coalescencia lineal y raíz constante.

A pesar de que el método que se eligió como más preciso fue *Beast, los resultados que se obtuvieron con IMa2, bajo las mismas condiciones (toda la información nuclear disponible y de la mitocondrial sólo la del grupo CCMN) (análisis 1, tabla A11) fueron bastantes similares. El tiempo de especiación calculado con un modelo de aislamiento con migración fue de 0,4595 (0,2946-0,7617) millones de años.

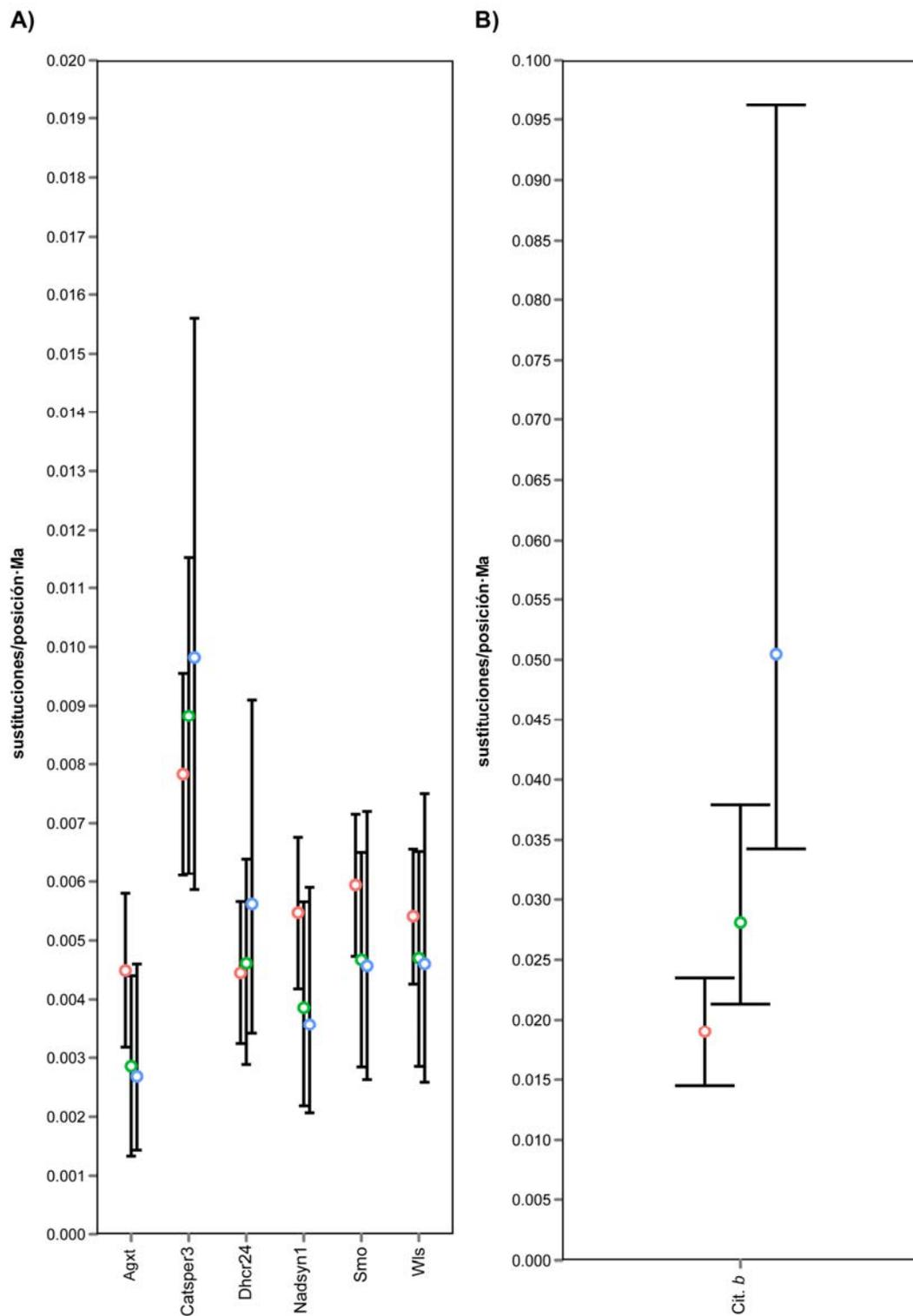


Figura 29. (A) Comparación de las tasas evolutivas de los seis intrones amplificados obtenidas mediante el uso de calibraciones fósiles (rojo), con las distribuciones posteriores de las tasas evolutivas obtenidas con *Beast (verde) y IMA2 (azul). (B) Comparación de la tasa evolutiva de citocromo *b* (Soria-Carrasco 2011) (rojo) con las distribuciones posteriores de las tasas evolutivas obtenidas con *Beast (verde) y IMA2 (azul).

Del análisis de *Beast descrito anteriormente se recogieron los valores de los TMRCAs (*time to the most recent ancestor*), de cada uno de los seis intrones y del citocromo *b* (tabla 11). Como se esperaba, para todos los marcadores su TMRCA fue más alto (entendido como más atrás en el tiempo respecto al tiempo actual) que el tiempo de especiación de *M. dudodecimcostatus* y *M. lusitanicus*.

Tabla 11. Tiempo al ancestro común más cercano (TMRCA) en millones de años de los seis intrones y el citocromo *b* calculados con *Beast.

Marcador	Mediana	95% HPD inferior	95% HPD superior
Agxt-intrón-10	0,961	0,525	1,636
Catsper3-intrón-5	1,273	0,704	2,133
Dhcr24-intrón-7	0,942	0,522	1,554
Nadsyn1-intrón-4	0,934	0,533	1,552
Smo-intrón-9	0,941	0,537	1,544
Wls-intrón-7	0,909	0,511	1,471
Citocromo <i>b</i>	0,802	0,509	1,145

3.1.2.4 Estima de los parámetros demográficos ancestrales y actuales de los dos topillos ibéricos

Los parámetros demográficos de las dos especies se estimaron tanto con *Beast como con IMA2 (análisis 1, tabla A11). En *Beast se utilizó un modelo de coalescencia linear y raíz constante para poder calcular tanto el tamaño poblacional ancestral como el actual de las dos especies. Debido a que en *Beast no se pudo implementar un tiempo de generación distinto de 1 los resultados obtenidos se dividieron por 0.5, que es el tiempo generacional aproximado de las dos especies de topillos ya que pueden presentar actividad sexual durante todo el año (Mira & Mathias 2007; Cotilla & Palomo 2007; Ventura *et al.* 2010). Los resultados obtenidos del tamaño poblacional efectivo, ancestral y actual, para *M. duodecimcostatus* y *M. lusitanicus* calculados con *Beast fueron muy similares a los obtenidos con IMA2 (tabla 12).

Se observó que ambas especies se encontrarían bajo procesos de expansión poblacional, ya que su tamaño efectivo poblacional actual es mucho más elevado que el de la población ancestral, situación más pronunciada para el caso de *M. lusitanicus* (tabla 12).

Tabla 12. Tamaños efectivos poblacionales (N_e) actuales y ancestrales de las dos especies de topillos, obtenidos con *Beast e IMA2, en millones de individuos.

Especie	*Beast	IMa2
<i>M. duodecimcostatus</i>	0,6738 (0,2398 - 1,604)	0,6622 (0,4245 - 1,0019)
<i>M. lusitanicus</i>	1,8194 (0,793 - 4,2228)	2,6185(1,8373 - 3,7664)
Ancestral	0,3393 (0,0014 - 2,4362)	0,2343 (0 - 0,7981)

3.1.2.5 Análisis de la incongruencia cito-nuclear bajo un escenario de introgresión o de separación incompleta de linajes

Para todos los análisis realizados bajo el modelo de aislamiento con migración (análisis 4, 5 y 6, tabla A11) no se obtuvieron valores significativamente distintos de cero tanto para la tasa de migración (m) como para la tasa de migración poblacional ($2Nm$). Este resultado indicaría que el número de individuos migrantes de una especie a otra es muy bajo y no fue detectado por el modelo IM de IMA2.

A pesar que el modelo de aislamiento con migración no detectó una migración significativa de forma global, se analizaron las secuencias mitocondriales mediante diferentes aproximaciones, ya que, el proceso de discordancia cito-nuclear parecía estar provocado únicamente por este marcador. Así, se pretendía demostrar si esta incongruencia era debida a una introgresión o podía explicarse por un modelo neutro de coalescencia y separación incompleta linajes.

En primer lugar, los diferentes análisis realizados para la estima del tiempo de especiación (análisis 2-6, tabla A11), indicaron el posible efecto de la hibridación de individuos sobre la estima del tiempo de divergencia entre las dos especies (figura 30). Además, esta migración de individuos de una especie a otra quedó registrada únicamente en el marcador mitocondrial y no en los intrones (figura 30). Al utilizar la información mitocondrial únicamente del grupo CCMN (figura 30A) (análisis 2, tabla A11), se obtuvo un tiempo de especiación, con *Beast e I, más elevado del obtenido por el mejor análisis (*apartado 3.1.2.3*), pero no incoherente ya que se tiene que considerar que posiblemente la tasa de mutación del marcador mitocondrial utilizada para los

análisis está subestimada (figura 29B). Seguidamente, se utilizó la información mitocondrial de los dos grupos, CCMN y DCMN, con un modelo I y con *Beast (figura 30B) (análisis 3, tabla A11). Para ambos modelos el tiempo de especiación bajó drásticamente para poder explicar la información contenida en las secuencias mitocondriales, bajo un escenario de separación incompleta de linajes. Este resultado indicó que se había violado una de las asunciones del modelo *Beast, ya que éste no permite la introducción de secuencias que hayan sufrido un proceso de intercambio génico posterior al evento de especiación.

Posteriormente se incluyeron los marcadores nucleares en los análisis para observar cómo podía variar la estima del tiempo de especiación según las diferentes condiciones de los análisis al incluir la información nuclear. Se calculó el tiempo de divergencia con la información de los marcadores nucleares (incluyendo el grupo CCMN y DCMN) (análisis 4, tabla A11) y se obtuvo prácticamente la misma estima con *Beast, con IM y con I (figura 30C). Al añadir la información mitocondrial del grupo CCMN (pero no del grupo discordante), a la información nuclear, (análisis 5, tabla A11) las estimas de tiempo fueron similares a las del análisis anterior para los tres métodos (figura 30D). Con *Beast se obtuvo un tiempo de especiación un poco superior, seguramente por lo argumentado en el *apartado 3.1.2.3*, es decir, el mejor uso que realiza *Beast de la información *a priori* sobre marcadores y tasas evolutivas.

El punto clave de estos análisis se encontró en el experimento en el que se utilizó la información nuclear total (CCMN y DCMN) y la información mitocondrial total (CCMN y DCMN) (figura 30E) (análisis 6, tabla A11). Al comparar los resultados obtenidos para cada uno de los métodos entre el análisis 5 (figura 30D) y el análisis 6 (figura 30E) se comprobó, que para el caso de *Beast, el tiempo de especiación obtenido era más reciente con las secuencias mitocondriales del grupo discordante. Este mismo resultado se obtuvo cuando se aplicó el modelo de I de IMA2. Sin embargo, para el caso del modelo IM, el tiempo de especiación obtenido fue similar entre los dos análisis. Este resultado indicó que para poder explicar la información contenida en las secuencias mitocondriales del grupo DCMN, bajo un modelo sin migración (I) y con la metodología *Beast (que tampoco considera migración), el tiempo de especiación tiene que sufrir un descenso considerable (figura 30E), indicando de nuevo la violación de la premisa anteriormente comentada. Sin embargo, si se permite la existencia de migración

(IM), aunque esta no sea significativamente distinta de cero, permite que la información contenida en las secuencias mitocondriales del grupo DCMN sea identificada como una huella de procesos de migración, posteriores a la especiación; y por tanto el tiempo de especiación calculado es similar al obtenido en los análisis que no incluyeron la información mitocondrial del grupo DCMN (figura 30D).

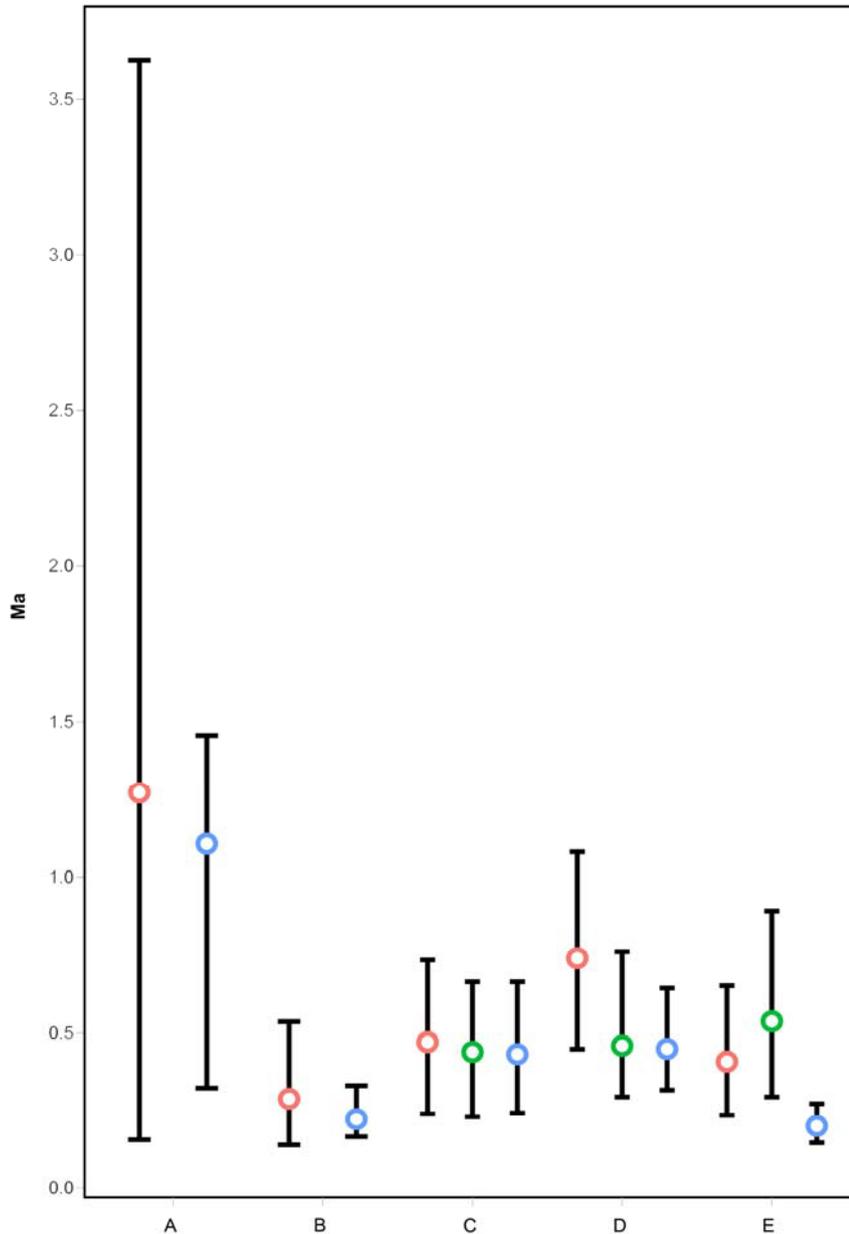


Figura 30. Comparación de las estimas del tiempo de especiación, en millones de años, de *M. duodecimcostatus* y *M. lusitanicus* calculados con *Beast (rojo), I (azul) e IMA2 (verde). **(A)** Citocromo *b* del grupo CCMN. **(B)** Citocromo *b* del grupo CCMN y DCMN. **(C)** Intrones del grupo CCMN. **(D)** Intrones y citocromo *b* del grupo CCMN. **(E)** Intrones y citocromo *b* del grupo CCMN y DCMN.

Analizando todo esta información en global, y teniendo en cuenta que el tiempo de especiación calculado con un modelo con migración (IM), con y sin las secuencias mitocondriales de los individuos discordantes, fue similar, mientras que el calculado bajo un modelo sin migración (I) y con *Beast variaba al incluir o no la información mitocondrial del grupo DCMN, se podría concluir que la información filogenética registrada en las secuencias mitocondriales del grupo DCMN se corresponde más con un escenario de introgresión por migración e hibridación que con un escenario de separación incompleta de linajes.

Por otro lado, se analizó la mínima distancia genética (Joly *et al.* 2009) para el citocromo *b* entre un individuo de *M. dudodecimcostatus* y otro de *M. lusitanicus* para ver si esta distancia sería compatible con un modelo nulo de coalescencia y separación incompleta de linajes o no. Esta distancia fue de 0,0039 sustituciones/posición para el árbol con información mitocondrial de los 184 individuos y de 0,0083 sustituciones/posición para el árbol mitocondrial de los 44 individuos. En ambos casos, la distancia genética resultó ser menor que el 5% más bajo obtenido a partir de las simulaciones nulas (0,0097). Un valor de distancia genética más bajo en el árbol real de citocromo *b* reforzaría la hipótesis de que un escenario de hibridación sería la explicación más plausible para la discordancia cito-nuclear observada.

Además, se calculó el índice GSI (*genealogical sorting index*) (tabla 13). El análisis se efectuó sobre la información mitocondrial, tanto de los 44 individuos como del total de los 184 individuos.

Tabla 13. Valores del índice GSI obtenidos para las 10.000 genealogías simuladas a partir el árbol de especies para el cálculo de la mínima distancia genética (media y desviación estándar); para el árbol de citocromo *b* de los 44 individuos de los que se obtuvo información nuclear y mitocondrial; y para el árbol de citocromo *b* del total de 184 individuos.

Especie	GSI simulaciones	GSI (44 individuos)	GSI (184 individuos)
<i>M. duodecimcostatus</i>	0,8971 (±.0,0157)	0,7383	0,9152
<i>M. lusitanicus</i>	0,8746 (±.0,0032)	0,5656	0,7976

Se observó, en primer lugar, que el valor de GSI de las simulaciones para las dos especies era bastante similar. Es decir, ambas especies presentarían un grado de monofilia recíproca similar bajo un escenario nulo (sin la existencia de procesos de migración) (tabla 13). Sin embargo, se observó que el índice de GSI para *M. lusitanicus* resultó ser más bajo que el de *M. duodecimcostatus* cuando se calculó sobre el árbol real de citocromo *b*. Esto ocurrió así tanto sobre el árbol de 44 individuos como sobre el de 184 individuos (tabla 13). Por tanto, para ambas reconstrucciones filogenéticas, *M. lusitanicus* presenta un menor grado de monofilia recíproca respecto a *M. duodecimcostatus*, apoyando un escenario de introgresión frente a uno de separación incompleta de linajes.

Por último, se representó el tiempo de especiación descrito en el apartado 3.1.2.3 en un árbol de inferencia Bayesiana reconstruido con Beast para el total de las 184 secuencias de citocromo *b* (figura 31). Bajo un escenario de separación incompleta de linajes, se esperaría que los nodos coalescencia entre las secuencias alélicas de los individuos del grupo DCMN y *M. duodecimcostatus* se situaran más atrás en el tiempo que el tiempo de divergencia de las dos especies (figura 7) (Joly *et al.* 2009). Bajo un escenario de hibridación se podrían esperar las dos situaciones, o que se situaran más atrás en el tiempo o más recientes que el tiempo de divergencia (figura 7) (Joly *et al.* 2009). Se observó que estas coalescencias son más recientes que el tiempo de especiación (figura 31). Por tanto, se podría rechazar la existencia de un escenario de separación incompleta de linajes y reforzar la hipótesis, —junto con los resultados comentados anteriormente— de que la incongruencia cito-nuclear que se observa en estas dos especies de topillos se podría deber a un proceso de introgresión por hibridación.

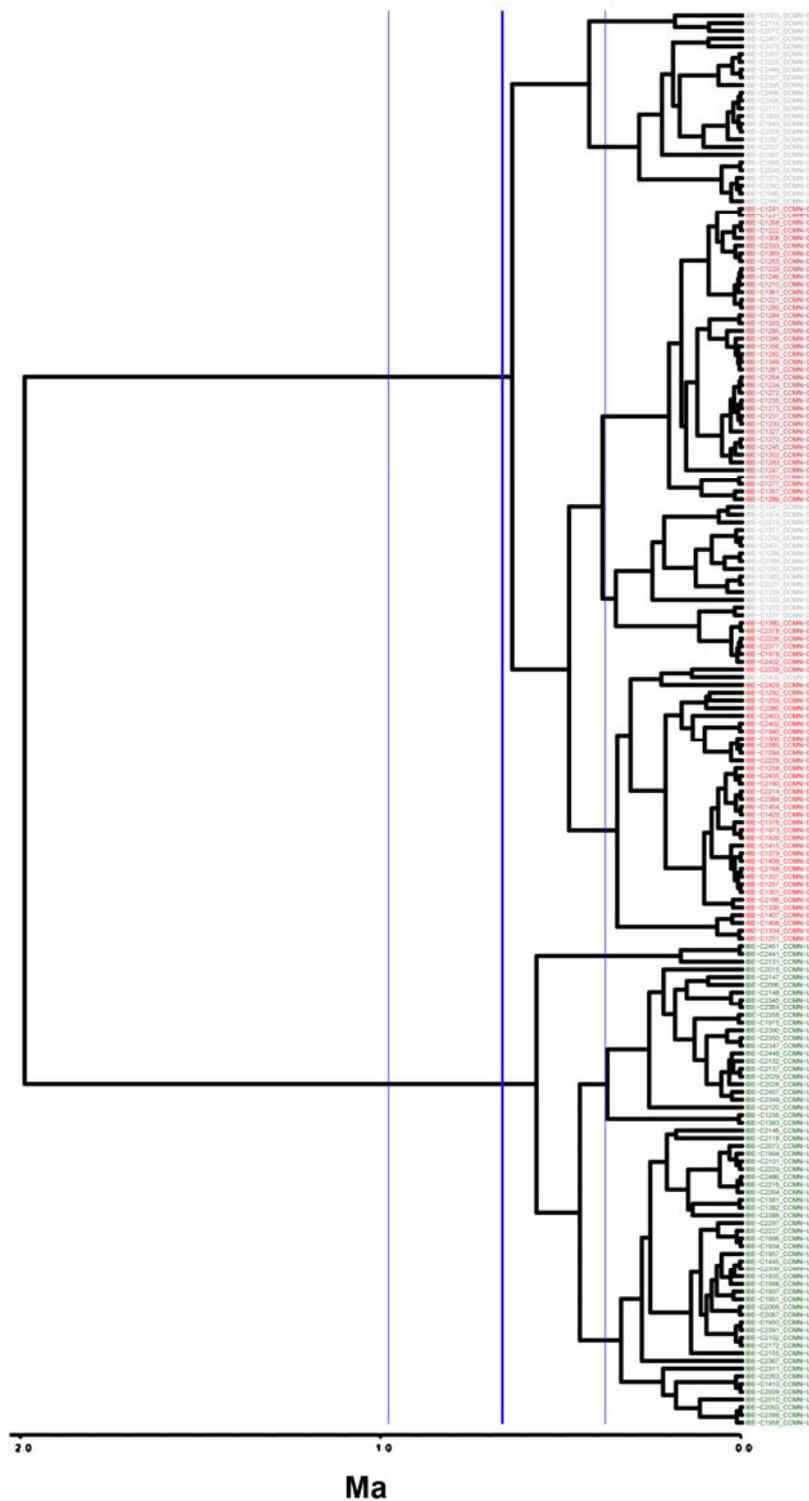


Figura 31. Árbol de inferencia Bayesiana del citocromo *b* del total de los 184 individuos. Se representa el tiempo de especiación, 0,6487 (0,3601-0,9791) en millones de años. En rojo se representa el clado formado por *M. duodecimcostatus*, en verde el clado de *M. lusitanicus* y en gris el clado de DCMN (individuos que presentan un discordancia entre la asignación de especie mediante el citocromo *b* [*M. duodecimcostatus*] y mediante los intrones y la morfología [*M. lusitanicus*]).

3.2 DISCUSIÓN

3.2.1 Diferentes escenarios para la historia evolutiva de *M. duodecimcostatus* y *M. lusitanicus*

Como el proceso de especiación es un continuo de un largo proceso (De Queiroz 2007) intentar describirlo no es una tarea fácil. Describir un evento de especiación implicaría de forma general, determinar cuáles ha sido las especies que se han originado, cuál es el tiempo de divergencia entre las especies, cuál fue el origen geográfico de la especiación y describir si se han desarrollado todos los procesos (genéticos, ecológicos, morfológicos, etc.) que permiten la existencia de las especies como entidades separadas. Para el caso concreto de la reconstrucción filogenética de un evento de especiación, el uso de las metodologías que sean capaces de integrar de un modo más preciso la información contenida en los marcadores nucleares utilizados promoverá la obtención de información más precisa y aproximada a la realidad de este evento de especiación.

M. duodecimcostatus y *M. lusitanicus* son dos especies hermanas de topillos subterráneos que se distribuyen a lo largo de la Península Ibérica. El intento de describir su proceso de especiación mediante reconstrucciones filogenéticas de árboles de especie dio como resultado diversas opciones para explicar su historia evolutiva.

Uno de los puntos de esta historia evolutiva que se intentó caracterizar con más precisión fue el tiempo de especiación. De acuerdo a los resultados, estas dos especies divergieron hace 0,6487 (0,3601-0,9791) Ma. Este periodo coincidiría con el Pleistoceno medio (Kurten 2007), que en el continente Europeo se ha caracterizado por ser un periodo donde se produjeron eventos de especiación a nivel global y en todos los taxa (Randi 2006). Por tanto, el origen de estas dos especies de topillos podría ser un ejemplo más de estos procesos especiativos que se desarrollaron en el continente Europeo.

Otra de las características importantes que se pueden describir en un evento de especiación es el origen geográfico del mismo. Determinar mediante reconstrucciones filogenéticas cuál fue la distribución geográfica de las poblaciones ancestrales que dieron lugar a las especies actuales es un área de estudio en el que existe mucha

controversia y en la que no hay un acuerdo firme sobre los métodos más adecuados a utilizar (Barraclough & Nee 2001; Losos & Glor 2003). Para el caso concreto de estas dos especies de topillos, con la información de las secuencias que se obtuvo y con la información actual de sus rangos de distribución se podrían plantear dos escenarios geográficos de especiación.

Actualmente, estas dos especies presentan de forma predominante una distribución parapatrica con fragmentos de zonas distribución solapada (Santos *et al.* 2009b; Santos *et al.* 2010). Existen dos explicaciones plausibles acerca del origen geográfico de la especiación de dos especies que presentan un rango actual de distribución parapatrica. Por un lado, el evento de especiación pudo ser alopatrico y que la distribución actual sea el resultado de una expansión posterior de sus rangos de distribución hasta distribuirse en zonas adyacentes (Bull 1991). Por otro lado, el origen geográfico de la especiación puede haber sido igualmente parapatrico como la distribución geográfica actual (Bull 1991).

La Península Ibérica se ha considerado tradicionalmente como un refugio glacial durante el Pleistoceno para muchas especies. De forma paralela se ha empezado a considerar que no ha sido un refugio continuo si no que en realidad se pueden distinguir sub-refugios dentro del Península Ibérica (Gómez & Lunt 2006). Dos de los refugios-dentro-refugio que se han descrito, dentro de la Península Ibérica, coinciden con parte de las actuales áreas de distribución alopatricas de las dos especies de topillos (noroeste de la Península y sur de los Pirineos). Por tanto, uno de los posibles escenarios evolutivos que se podría plantear sería una primera especiación alopatrica, en la que las dos especies evolucionaron a partir de las poblaciones ancestrales de forma aislada en estos sub-refugios durante un período glacial. Posteriormente, se producirían expansiones de los rangos de distribución geográficas de las dos especies hasta alcanzar la distribución actual. El otro escenario evolutivo plantearía una especiación parapatrica. Es decir, a partir de una población ancestral, estas dos especies de topillos se formaron como especies independientes en una situación de distribución geográfica adyacente. El proceso de especiación se hubiera visto reforzado por procesos de selección adaptativa (diferenciación ecológica o diferenciación sexual) (Smadja 2011) (figura 3).

Estos escenarios de especiación pueden ser igualmente plausibles y es difícil determinar cuál de ellos es el que se ajusta mejor a la verdadera historia evolutiva de las especies, o si existen otras alternativas posibles. Esta situación ocurre debido a que inferir el origen geográfico de la especiación a partir de los rangos de distribución actuales es un proceso complejo y arriesgado (Losos & Glor 2003).

Pero, a pesar de esta incertidumbre, el análisis meticuloso de la información contenida en las secuencias mitocondriales y nucleares de estas dos especies puede aportar información acerca de la relación filogenética que existe entre ellas y del punto en el que se encuentra el proceso de especiación.

M. lusitanicus es la especie que presentó una mayor diversidad haplotípica y nucleotídica. Esta mayor diversidad genética en la especie *M. lusitanicus* se vio igualmente reflejada en los resultados obtenidos sobre los tamaños efectivos poblacionales de las dos especies. La información combinada de las secuencias mitocondriales y nucleares analizadas bajo un escenario de coalescencia indicó que las dos especies sufrieron procesos de expansión poblacional, ya que sus tamaños efectivos poblacionales actuales son mayores que los registrados en las poblaciones ancestrales (tabla 12). Dentro de este escenario es la especie *M. lusitanicus* la que presenta un mayor grado de expansión poblacional frente a *M. duodecimcostatus*. Estos resultados podrían indicar la existencia de un proceso de especiación peripátrica, donde la especie *M. duodecimcostatus* se habría originado a partir de una poblacional marginal de la especie *M. lusitanicus*. Sin embargo, los resultados de las reconstrucciones filogenéticas, tanto de las secuencias mitocondriales como de las nucleares, indican que este escenario no es demasiado plausible. La señal filogenética de una especiación peripátrica debería contener más clados parafiléticos, al menos en las secuencias nucleares, de los observados (Losos & Glor 2003). Por otro lado, la mayor variabilidad genética podría ser un indicio de que el refugio glacial donde se situó *M. lusitanicus* durante las épocas de máximas glaciaciones era más grande que el que ocupó la especie *M. duodecimcostatus*. Debido a esto, la especie *M. lusitanicus* pudo mantener un acervo genético más elevado y mantener así una mayor diversidad nucleotídica y haplotípica durante su historia evolutiva respecto a *M. duodecimcostatus*. Este escenario se correspondería con una especiación alopátrica o parapátrica.

Existe un gran número de evidencias morfológicas, genéticas y filogenéticas que apuntan a la clara separación de estas dos especies, constituyéndose como entidades independientes. A pesar de esta clara separación se detectó la presencia de una discordancia cito-nuclear que podría reflejar la existencia de procesos de flujo génico posteriores al evento de especiación.

Aunque en trabajos anteriores se detectó una discordancia mitocondrial en estas dos especies (Bastos-Silveira *et al.* 2012), no se realizó ningún test para demostrar si eran debidos a una introgresión por hibridación o a una separación incompleta de linajes. Como una discordancia cito-nuclear puede ser la consecuencia de procesos tan diferentes como la separación incompleta de linajes o la introgresión por hibridación (Joly 2009) (figura 7), en este trabajo se aplicaron metodologías de reconstrucción de árboles de especies (*Beast), modelos de aislamiento con migración (IMa2) y tests apropiados que permitían cuantificar las diferencias genéticas existentes entre las dos especies para discernir cuál de estos dos procesos —separación incompleta de linajes o introgresión por hibridación— había sido el causante de la discordancia cito-nuclear observada.

Los resultados que se obtuvieron en el presente trabajo apuntaron que el origen de esta discordancia cito-nuclear se debió a la existencia de procesos introgresión por hibridación de la especie *M. duodecimcostatus* hacia *M. lusitanicus*, originado secuencias mitocondriales introgresadas en individuos de la especie *M. lusitanicus*. El hecho de detectar una introgresión asimétrica (individuos de la especie *M. lusitanicus* con ADN mitocondrial de *M. duodecimcostatus*) puede ser debido a diferentes situaciones. Una explicación plausible sería la existencia de procesos de expansión en los rangos de distribución de una de las especies y que al entrar en contacto de manera secundaria con la especie residente se produjera un intercambio de flujo génico debido a la inexistencia de barreras de aislamiento reproductivo bien definidas (Currat & Excoffier 2004; Petit & Excoffier 2009). Para que se verifique un introgresión diferencial o asimétrica, es decir, para que exista un proceso de introgresión sobre el ADN mitocondrial y no sobre el nuclear, debe existir, por un lado, una mayor dispersión de los machos y, por otro, una baja variabilidad intraespecífica de la especie invasora (Petit & Excoffier 2009). En el género *Microtus* se ha descrito la existencia de una dispersión mayor de los machos (Hahne *et al.* 2011, Le Galliard *et al.* 2012). Bajo estas

premisas, *M. lusitanicus* debería haber invadido el rango de distribución de *M. duodecimcostatus* y, además, los individuos dispersantes debían presentar una baja variabilidad genética permitiendo, de este modo, una introgresión masiva de ADN mitocondrial de *M. duodecimcostatus* a las poblaciones de *M. lusitanicus*.

La discordancia cito-nuclear es igualmente asimétrica o diferencial a nivel geográfico, ya que sólo se observó en individuos pertenecientes a las zonas de distribución solapada. Esta situación reforzaría la hipótesis de que una introgresión por hibridación haya sido el proceso más plausible que ha provocado la discordancia cito-nuclear, ya que bajo un escenario de separación incompleta de linajes podría existir una discordancia cito-nuclear en individuos de todo el área de distribución geográfica de las especies (Nosil 2012).

En todo caso, el análisis morfológico (capítulo 2) detecta tan solo dos grupos de individuos correspondientes a las dos especies y no un tercer grupo que podría corresponder a individuos híbridos (figura 24). Esto, junto al hecho de que no se detectó una señal de introgresión en el genoma nuclear, indicaría que las hibridaciones son antiguas y el tiempo ha diluido la señal de hibridación en el genoma (Marchant 1988).

Se intentó, por tanto, describir el proceso de especiación y la historia evolutiva de estas dos especies de topillos de la Península Ibérica analizándolo desde el mayor número de puntos de vista y utilizando las metodologías más adecuadas (métodos de coalescencia) (figura 32) para poder desgranar e identificar todos los procesos que se han sucedido para dar lugar a esta pareja de especies cercanas.

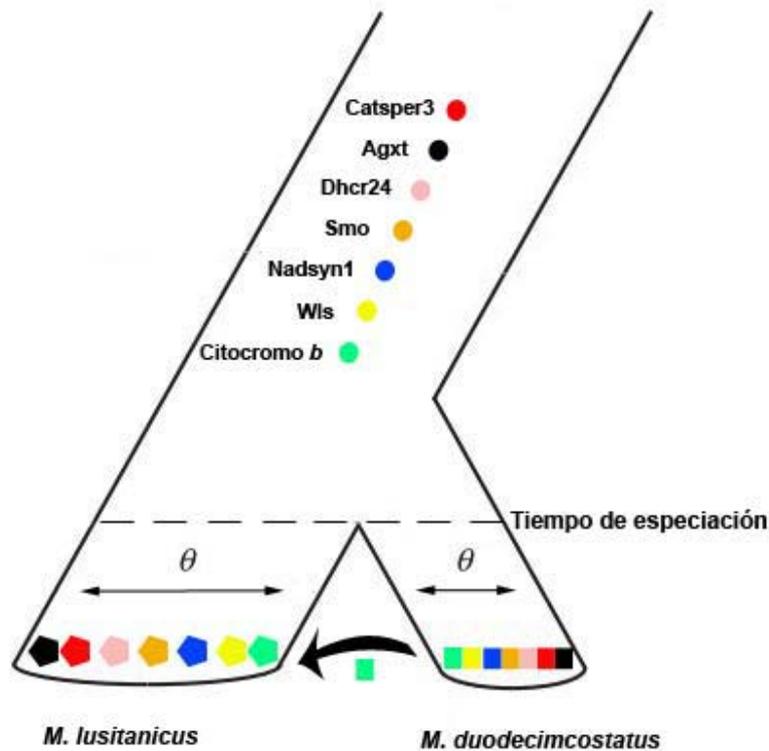


Figura 32. Modelización del árbol de especies de *M. duodecimcostatus* y *M. lusitanicus*. Se representan los TMRCA de los marcadores empleados y los tamaños poblacionales (θ) actuales de las dos especies.

Por tanto, estas dos especies se presentan como entidades independientes, en el que el proceso de especiación se ha verificado y que actualmente presentan una distribución parapátrica a lo largo de la Península Ibérica con áreas de distribución solapada. Esta distribución parapátrica se podría entender como un global de una serie de poblaciones con distribución alopátrica en las zonas más lejanas de la distribución de las dos especies y, a partir de las cuáles, las áreas de distribución de las dos especies se extienden hacia el centro peninsular llegando a entrar en contacto, originando las áreas de distribución solapada. Pero, a pesar de estas zonas solapadas, estas dos especies presentarían una distribución más parapátrica que simpátrica porque, a pesar de que comparten áreas de distribución, las características a pequeña escala del territorio ocupado (como las características del suelo, [Santos *et al.* 2009b]) son las que condicionan verdaderamente su distribución, manteniendo las poblaciones de cada especie en hábitats probablemente más similares que en las zonas alopátricas, pero presentando ciertas diferencias entre ellos. Esta hipótesis se vería reforzada por los

resultados obtenidos en el análisis de la variación de forma entre las dos especies, donde la diferencia es mayor entre los individuos de las zonas de distribución alopátrica. Esta diferencia disminuye en las zonas de distribución solapada en las que los individuos de ambas especies sufren una convergencia en forma, adoptando una forma más idónea al hábitat de las zonas de distribución solapada. Ya que los caracteres morfológicos analizados están muy relacionados con la actividad excavadora de estas dos especies de topillos subterráneos, la no total convergencia reflejaría la existencia de pequeñas diferencias entre los hábitats —principalmente en la composición del substrato— donde se distribuye cada especie.

V. CONCLUSIONES

1. Se analizaron todos los genomas de roedores bien anotados disponibles en las bases de datos para seleccionar, mediante una serie de filtros bioinformáticos, un conjunto final de 208 intrones útiles para resolver cuestiones filogenéticas entre especies cercanas dentro del orden *Rodentia* y para poder realizar estudios de especiación dentro de este grupo.
2. Los intrones seleccionados se caracterizaban por encontrarse en copia única en el genoma, presentar una longitud adecuada que permitía su amplificación, estar flanqueados por exones conservados y con una longitud suficiente para poder diseñar cebadores de PCR sobre ellos, y ser lo suficientemente divergentes como para ser informativos para resolver cuestiones filogenéticas entre especies cercanas.
3. Se llevó a cabo la amplificación de diez de estos 208 intrones en diez especies del orden *Rodentia* consiguiendo un éxito de amplificación del 78% con mínimas modificaciones de la temperatura inicial de hibridación, indicando la validez de los filtros aplicados para la selección de los intrones.
4. Las reconstrucciones filogenéticas de máxima verosimilitud y de inferencia Bayesiana realizadas con las secuencias amplificadas reflejaron las topologías propuestas en la literatura para el orden *Rodentia*, demostrando que los intrones seleccionados son útiles tanto para las reconstrucciones filogenéticas de especies cercanas como para las reconstrucciones de nodos más profundos, pudiendo ser útiles estos intrones, por tanto, para el cálculo de tasas evolutivas con el uso combinado de dataciones fósiles.
5. Para el estudio del proceso de especiación entre las especies hermanas de topillos ibéricos, *Microtus duodecimcostatus* y *Microtus lusitanicus*, y debido a su gran similitud morfológica, se analizó mediante la técnica de morfometría geométrica la información sobre la variación de la forma de cinco caracteres óseos seleccionados. Ello permitió una correcta asignación de especie a los individuos de las zonas de distribución alopátrica, ya que coincidía con su distribución geográfica, indicando la gran capacidad discriminante de estos caracteres. Del mismo modo, esta información permitió una clasificación robusta de los individuos de las zonas de distribución solapada.

6. *M. duodecimcostatus* presenta un tamaño mayor para cuatro de los cinco caracteres analizados respecto a *M. lusitanicus* pero no presenta una variación intraespecífica en el tamaño de las estructuras. Sin embargo, los individuos de *M. lusitanicus* de las zonas de distribución solapada presentan un tamaño superior respecto a sus conespecíficos de las zonas de distribución alopátrica para cuatro de los cinco caracteres, alcanzando una dimensión similar a la que presenta *M. duodecimcostatus*.

7. Existe una clara diferenciación en forma para los cinco caracteres analizados entre las dos especies de topillos aunque, en las zonas de distribución solapada existe una convergencia en forma para cuatro de los cinco caracteres. La cantidad de cambio que sufren estas estructuras es equivalente en las dos especies, demostrando una adquisición de la forma más óptima por parte de las dos especies para el ambiente donde se distribuyen en las zonas de distribución solapada.

8. El uso de la morfometría geométrica y el análisis combinado de cinco caracteres para la diferenciación morfológica entre estas dos especies de topillos permite recoger de un modo más integrativo las diferencias en forma existentes entre ellas.

9. Mediante el uso de secuencias mitocondriales y de seis intrones en el contexto de la teoría de la coalescencia y de la reconstrucción de árboles de especies, se estimó que *M. duodecimcostatus* y *M. lusitanicus* divergieron hace aproximadamente 0,65 millones de años, durante el Pleistoceno Medio. Para realizar esta estimación, se calcularon previamente las tasas evolutivas de cada uno de los intrones en un árbol de varias especies de roedores.

10. Por la distribución actual de las especies —parapátrica con zonas de distribución solapada— se puede proponer que sufrieron una especiación de tipo alopátrica y un contacto secundario posterior por expansión de las áreas de distribución o una especiación de tipo parapátrica. *M. lusitanicus* presenta una diversidad haplotípica y nucleotídica más alta que *M. duodecimcostatus* y, aunque las dos especies experimentaron procesos de expansión poblacional, el que sufre *M. lusitanicus* es mayor.

11. Los árboles de citocromo *b* indicaron que las secuencias de *M. lusitanicus* no formaban un grupo monofilético mientras que la información nuclear reflejaba la clasificación morfológica de las especies. El uso de distintos tests indicó que la discordancia cito-nuclear detectada se podría explicar por la existencia de procesos de introgresión por hibridación de la especie *M. duodecimcostatus* hacia *M. lusitanicus*, originando así secuencias mitocondriales introgresadas en individuos de la especie *M. lusitanicus*, siendo los machos del topillo lusitano los que invadieron las áreas de distribución del topillo mediterráneo.

12. Ya que la asignación morfológica detectó únicamente dos grupos morfológicos — correspondientes con las dos especies— y que la introgresión detectada es asimétrica — sólo se detectó en el genoma mitocondrial— se considera que los procesos de hibridación son antiguos y que el tiempo ha diluido la señal de hibridación, por lo que estas dos especies se pueden considerar como dos entidades independientes, verificándose así la finalización del proceso de especiación entre ellas.

VI. APÉNDICE

Tabla A1. Especies del orden *Rodentia* utilizadas para la extracción de ADN genómico, a partir de tejido fresco, para su posterior amplificación. AMCC, *Ambrose Monell Cryo Collection* del Museo de Historia Natural de Nueva York; BTAC, Banco de tejidos de Cataluña de la Universidad Autónoma de Barcelona.

Código individuo	Familia	Especie	Localidad	Donador
AMCC103825	<i>Octodontidae</i>	<i>Octodontomys gliroides</i>	Sama, Tarija, Bolivia	AMCC
AMNH275683 AMCC125125	<i>Hystriidae</i>	<i>Atherurus macrourus</i>	Huu Lien Nature Reserve, Lang Son, Vietnam	AMCC
AMCC175991	<i>Echymidae</i>	<i>Proechimys guairae</i>	Cabure, Serrania de San Luis, Venezuela	AMCC
2009/57	<i>Sciuridae</i>	<i>Cynomys ludovicamus</i>		BTAC
131/02	<i>Myocastoridae</i>	<i>Myocastor coypus</i>		BTAC
Z1107MS	<i>Cavidae</i>	<i>Hydrochoerus hydrochaeris</i>		BTAC
IBE-C2054	<i>Muridae</i>	<i>Apodemus flavicollis</i>	Gavorrano, Grosseto, Italia	
IBE-C2063	<i>Sciuridae</i>	<i>Sciurus vulgaris</i>	Vilazzano, Trento,Italia	
IBE-C2393	<i>Cricetidae</i>	<i>Microtus duodecimcostatus</i>	Belianes,Lleida, España	
IBE-C2161	<i>Gliridae</i>	<i>Glis glis</i>	Villa Lagarina, Trento,Italia	

Tabla A2. Modelos de sustitución nucleotídica y relojes moleculares calculados para cada uno de los diez intrones con la información de las secuencias de las diez especies de roedores.

Nombre intrón	Modelo de sustitución nucleotídica	Reloj molecular
Abcb9-intrón-2	HKY+I	Relajado
Agxt-intrón-10	HKY+G	Relajado
Catsper3-intrón-5	HKY+G	Relajado
Dhcr24-intrón-7	GTR+I	Relajado
Ivd-intrón-8	HKY+G	Estricto
Nadsyn1-intrón-4	HKY+G	Relajado
Ras-intrón-4	HKY+G	Estricto
Smo-intrón-9	HKY+G	Relajado
Trpv4-intrón-8	GTR+G	Estricto
Wls-intrón-7	GTR+G	Relajado

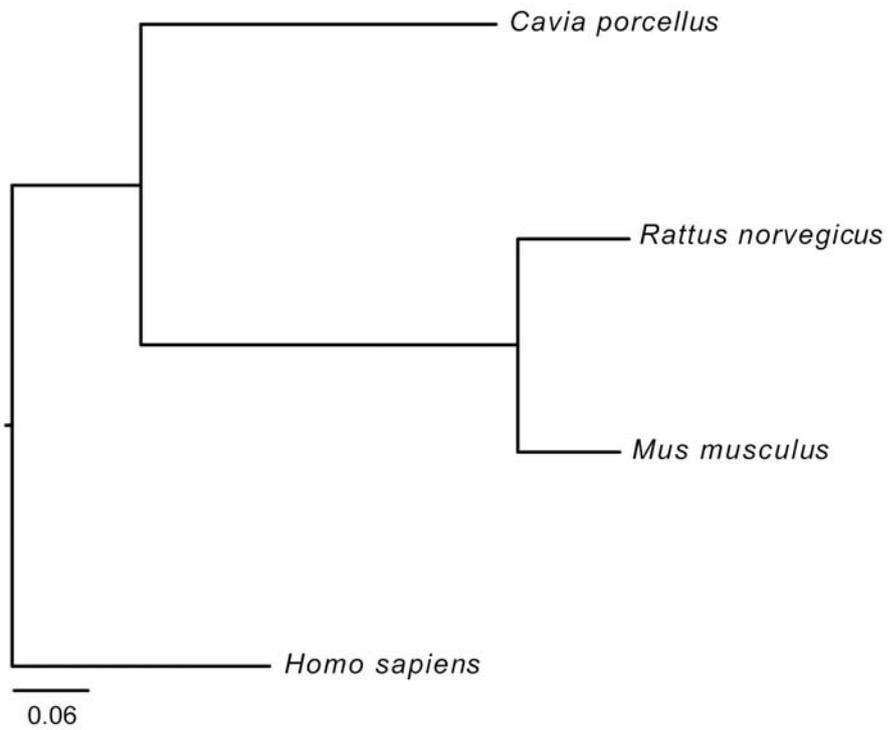


Figura A1 Árbol de máxima verosimilitud reconstruido con el concatenado de 2288 intrones obtenidos a partir de los filtros bioinformáticos. Árbol de referencia utilizado para calcular el *scaling factor* y el *K score* de cada intrón.

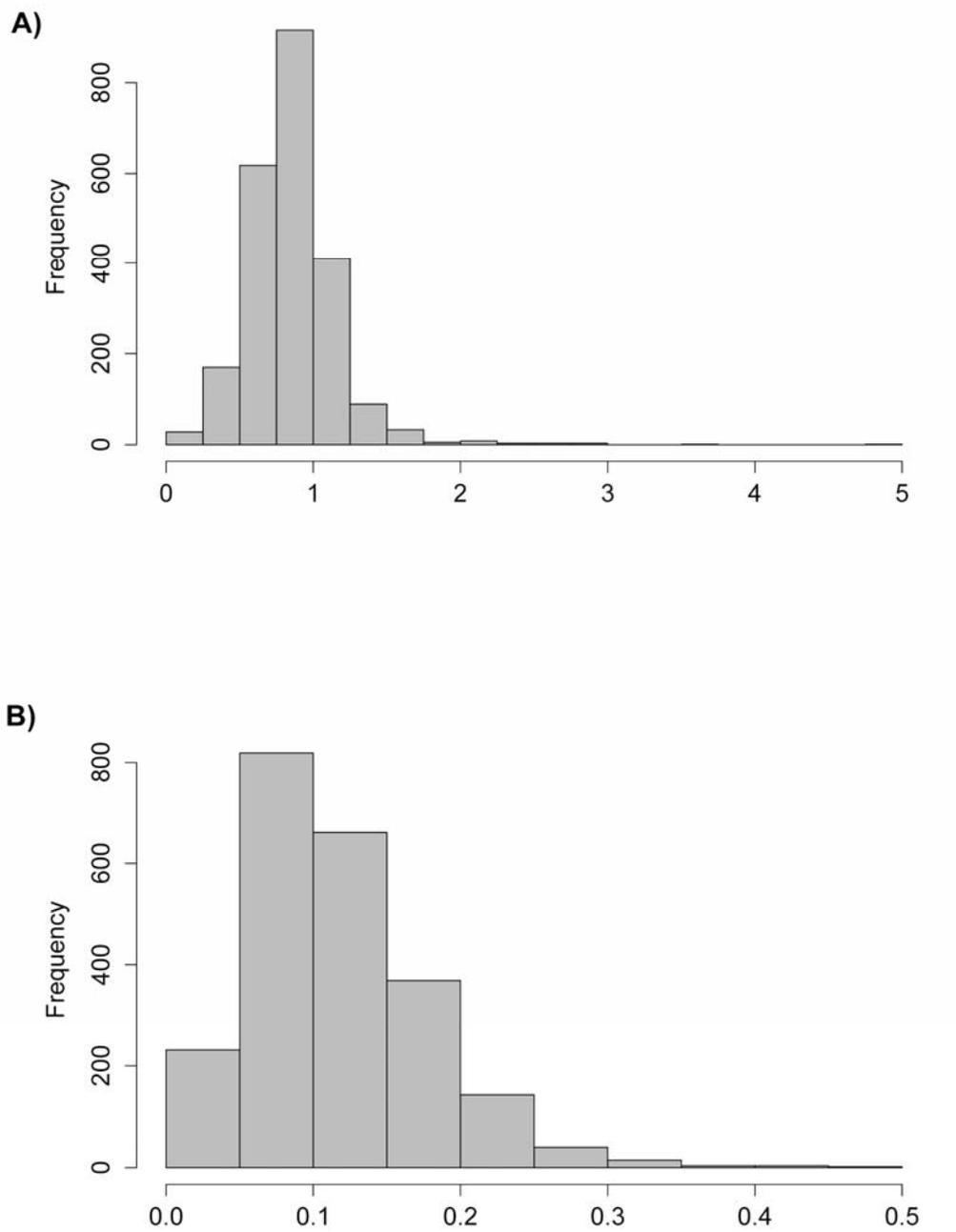


Figura A2. Distribución de frecuencias de los valores de *scaling factor* (A) y de *K score* (B) de los 2288 intrones.

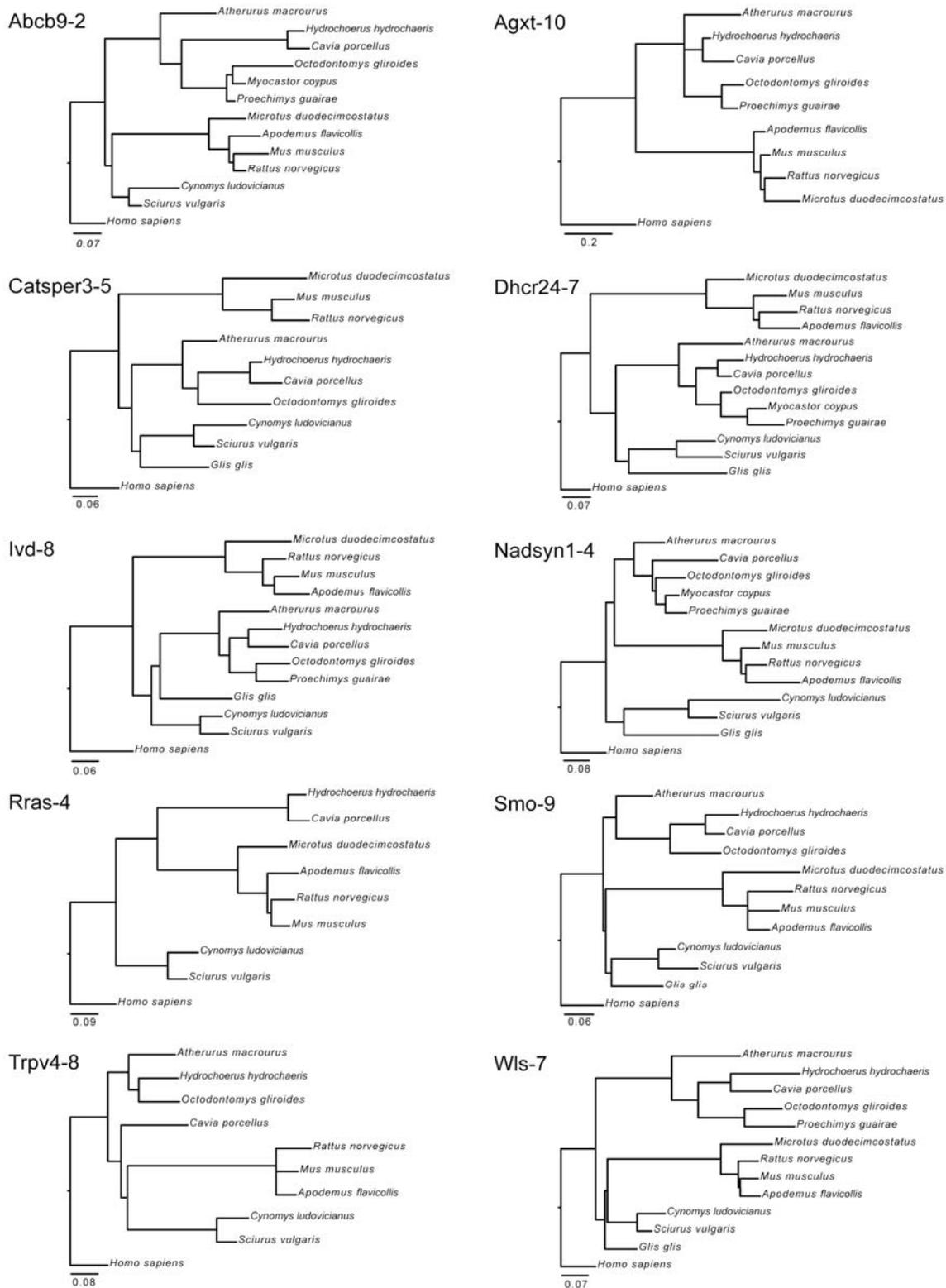


Figura A3. Árboles de máxima verosimilitud de los diez intrones, reconstruidos con las secuencias amplificadas en diez especies de roedores. La escala indica número de sustituciones por posición.

Tabla A3. Individuos recolectados de las dos especies de topillos. La línea horizontal divide los individuos de las zonas de distribución alopátrica (parte superior de la línea) y los de las zonas de distribución solapada (parte inferior de la línea). Se recoge el número de caracteres de los cuáles se pudo obtener información de cada individuo (X, presencia de información). En cursiva, los individuos de las zonas de distribución alopátrica de los que se obtuvo información para los cinco caracteres, y que forman el grupo de referencia. Localidades de muestreo: **Bel:** Belianes; **Ceb:** San Cebrían de Mazote; **Esc:** Escorial de la Sierra; **Hos:** Hospital de Incio; **Lum:** Lumbrales; **Men:** Menaza; **Mun:** Muñíos; **Oss:** Ossó de Sió; **Pal:** Palacios de Sil; **Pob:** Poblaciones de Arreba; **Por:** Porqueriza; **Sal:** Salvatierra de Tormes; **Toa:** Torrejón de Alba; **Tol:** Toledo y **Vil:** Villacantid.

Código individuo	Localidad	Lateral (L)	Ventral (V)	Mandíbula (J)	Primer molar inferior (M1)	Tercer molar superior(M3)
C1210	Bel	X	X	--	X	--
C1221	<i>Bel</i>	<i>X</i>	<i>X</i>	<i>X</i>	<i>X</i>	<i>X</i>
C1222	<i>Bel</i>	<i>X</i>	<i>X</i>	<i>X</i>	<i>X</i>	<i>X</i>
C1223	Bel	X	X	X	--	X
C1230	Bel	X	X	X	--	X
C1231	<i>Bel</i>	<i>X</i>	<i>X</i>	<i>X</i>	<i>X</i>	<i>X</i>
C1234	<i>Bel</i>	<i>X</i>	<i>X</i>	<i>X</i>	<i>X</i>	<i>X</i>
C1235	<i>Bel</i>	<i>X</i>	<i>X</i>	<i>X</i>	<i>X</i>	<i>X</i>
C1237	<i>Bel</i>	<i>X</i>	<i>X</i>	<i>X</i>	<i>X</i>	<i>X</i>
C1239	<i>Bel</i>	<i>X</i>	<i>X</i>	<i>X</i>	<i>X</i>	<i>X</i>
C1241	Bel	X	X	X	--	X
C1245	<i>Bel</i>	<i>X</i>	<i>X</i>	<i>X</i>	<i>X</i>	<i>X</i>
C1246	<i>Bel</i>	<i>X</i>	<i>X</i>	<i>X</i>	<i>X</i>	<i>X</i>
C1247	<i>Bel</i>	<i>X</i>	<i>X</i>	<i>X</i>	<i>X</i>	<i>X</i>
C1264	<i>Bel</i>	<i>X</i>	<i>X</i>	<i>X</i>	<i>X</i>	<i>X</i>
C1265	Bel	--	X	X	X	X
C1266	<i>Bel</i>	<i>X</i>	<i>X</i>	<i>X</i>	<i>X</i>	<i>X</i>
C1268	Bel	X	X	X	--	X
C1270	Bel	--	X	X	X	--
C1272	<i>Bel</i>	<i>X</i>	<i>X</i>	<i>X</i>	<i>X</i>	<i>X</i>
C1273	<i>Bel</i>	<i>X</i>	<i>X</i>	<i>X</i>	<i>X</i>	<i>X</i>
C1277	Bel	X	X	X	--	X
C2028	Hos	X	X	--	X	X
C2029	<i>Hos</i>	<i>X</i>	<i>X</i>	<i>X</i>	<i>X</i>	<i>X</i>
C2096	Hos	X	X	--	X	X
C2131	<i>Hos</i>	<i>X</i>	<i>X</i>	<i>X</i>	<i>X</i>	<i>X</i>
C2132	<i>Hos</i>	<i>X</i>	<i>X</i>	<i>X</i>	<i>X</i>	<i>X</i>
C2137	<i>Hos</i>	<i>X</i>	<i>X</i>	<i>X</i>	<i>X</i>	<i>X</i>
C2146	<i>Hos</i>	<i>X</i>	<i>X</i>	<i>X</i>	<i>X</i>	<i>X</i>
C2147	<i>Hos</i>	<i>X</i>	<i>X</i>	<i>X</i>	<i>X</i>	<i>X</i>
C2148	Hos	X	X	--	X	X
C2345	<i>Hos</i>	<i>X</i>	<i>X</i>	<i>X</i>	<i>X</i>	<i>X</i>
C2347	Hos	X	X	--	X	X
C2349	Hos	X	X	--	X	X
C2350	<i>Hos</i>	<i>X</i>	<i>X</i>	<i>X</i>	<i>X</i>	<i>X</i>
C2364	<i>Hos</i>	<i>X</i>	<i>X</i>	<i>X</i>	<i>X</i>	<i>X</i>
C2390	Hos	--	X	--	X	X
C2441	<i>Hos</i>	<i>X</i>	<i>X</i>	<i>X</i>	<i>X</i>	<i>X</i>
C2448	Hos	--	X	--	X	X
C2451	<i>Hos</i>	<i>X</i>	<i>X</i>	<i>X</i>	<i>X</i>	<i>X</i>
C2457	Hos	X	X	--	X	X
C1281	Oss	X	X	--	--	X
C1282	Oss	X	X	--	--	X
C1283	<i>Oss</i>	<i>X</i>	<i>X</i>	<i>X</i>	<i>X</i>	<i>X</i>
C1284	Oss	X	X	--	X	X
C1285	Oss	X	X	--	--	X
C1288	Oss	X	X	--	X	X
C1308	Oss	X	X	--	--	X

C1327	Oss	--	X	--	X	X
C1346	Oss	X	X	X	X	X
C1349	Oss	X	X	--	--	X
C1352	Oss	X	X	X	--	X
C1356	Oss	X	X	X	X	X
C1357	Oss	X	X	--	X	X
C1359	Oss	X	X	--	X	X
C1361	Oss	X	X	--	X	X
C1369	Oss	X	X	X	X	X
C2238	Tol	X	X	X	X	X
C2403	Tol	X	X	X	X	X
C2429	Tol	X	X	X	X	X
C1407	Ceb	X	X	--	X	X
C1408	Ceb	X	X	X	X	X
C1250	Esc	X	X	--	X	X
C1983	Esc	X	X	X	X	X
C2213	Esc	X	X	--	X	X
C2237	Esc	X	X	--	--	X
C1375	Lum	X	X	X	X	X
C1945	Lum	X	X	X	X	X
C1946	Lum	X	X	X	X	X
C1948	Lum	X	X	X	X	X
C1959	Lum	X	X	X	X	X
C1987	Lum	X	X	X	X	X
C2007	Lum	X	X	X	X	X
C2049	Lum	X	X	X	X	X
C2057	Lum	X	X	X	X	X
C2070	Lum	X	X	X	X	X
C2072	Lum	X	X	X	X	X
C2111	Lum	X	X	X	X	X
C2160	Lum	X	X	X	X	X
C2225	Lum	X	X	X	X	X
C2343	Lum	X	X	X	X	X
C2359	Lum	X	X	X	X	X
C2395	Lum	X	X	X	X	X
C2401	Lum	X	X	X	X	X
C2430	Lum	X	X	X	X	X
C2436	Lum	X	X	X	X	X
C2437	Lum	X	X	X	X	X
C2489	Lum	X	X	X	X	X
C2496	Lum	X	X	X	X	X
C1937	Men	X	X	X	X	X
C1950	Men	X	X	X	X	X
C1951	Men	X	X	X	X	X
C1957	Men	X	X	X	X	X
C1958	Men	X	X	X	X	X
C1988	Men	X	X	X	X	X
C2009	Men	X	X	X	X	X
C2010	Men	X	X	X	X	X
C2053	Men	X	X	X	X	X
C2067	Men	X	X	X	X	X
C2068	Men	X	X	--	X	X
C2073	Men	X	X	X	X	X
C2101	Men	X	X	X	X	X
C2102	Men	X	X	X	X	X
C2118	Men	X	X	X	X	X
C2155	Men	X	X	X	X	X
C2172	Men	X	X	X	X	X
C2263	Men	X	X	X	X	X
C2311	Men	X	X	X	X	X

VI. Apéndice

C2391	Men	X	X	X	X	X
C2399	Men	X	X	X	X	X
C2003	Mun	X	X	--	X	X
C2015	Mun	X	X	X	X	X
C2114	Mun	X	X	--	X	X
C2120	Mun	X	X	X	X	X
C1975	Pal	X	X	X	X	X
C2358	Pal	X	X	X	X	X
C1445	Pob	X	X	--	X	X
C1934	Pob	X	X	--	X	X
C1935	Pob	X	X	X	X	X
C1994	Pob	X	X	--	X	X
C1996	Pob	X	X	X	X	X
C2216	Pob	X	X	--	X	X
C2227	Pob	X	X	--	X	X
C2229	Pob	X	X	X	X	X
C2264	Pob	X	X	X	X	X
C2297	Pob	X	X	X	X	X
C2338	Pob	X	X	X	X	X
C2367	Pob	X	X	--	X	X
C2388	Pob	X	X	X	X	X
C2486	Pob	X	X	--	X	X
C1251	Por	X	X	--	--	X
C1258	Por	X	X	--	X	--
C1259	Por	X	X	X	X	X
C1292	Por	X	X	X	X	X
C1294	Por	X	X	X	X	X
C1300	Por	X	X	X	X	X
C1336	Por	X	X	X	X	X
C1380	Por	X	X	X	X	X
C1978	Por	X	X	--	X	X
C2166	Por	X	X	X	X	X
C2180	Por	X	X	X	X	X
C2236	Por	X	X	X	X	X
C2377	Por	X	X	X	X	X
C2378	Por	X	X	X	X	X
C2402	Por	X	X	X	X	X
C2432	Por	X	X	X	X	X
C2435	Por	X	X	X	X	X
C1257	Sal	X	X	X	X	X
C1293	Sal	X	X	--	X	X
C1296	Sal	X	X	X	X	X
C1301	Sal	X	X	X	X	X
C1335	Sal	X	X	--	X	X
C1337	Sal	X	X	X	X	X
C1339	Sal	X	X	X	X	X
C1341	Sal	X	X	X	X	X
C1374	Sal	X	X	X	X	X
C1377	Sal	X	X	--	X	X
C1379	Sal	X	X	X	X	X
C1415	Sal	X	X	X	X	X
C1429	Sal	X	X	X	X	X
C1438	Sal	X	X	X	X	X
C1454	Sal	X	X	X	X	X
C1939	Sal	X	X	--	X	X
C1940	Sal	X	X	X	X	X
C1973	Sal	X	X	X	X	X
C2168	Sal	X	X	X	X	X
C2169	Sal	X	X	X	X	X
C2214	Sal	X	X	X	X	X

C2228	Sal	X	X	X	X	X
C2380	Sal	X	X	X	X	X
C2384	Sal	X	X	X	X	X
C2385	Sal	X	X	X	X	--
C2431	Sal	X	X	X	X	X
C1252	Toa	X	X	--	X	X
C1291	Toa	X	X	X	X	X
C1297	Toa	X	X	X	X	X
C1334	Toa	X	X	X	X	X
C1376	Toa	X	X	X	X	X
C1410	Vil	X	X	X	X	X

Tabla A4. Descripción de los *landmarks* seleccionados para los cinco caracteres.

Lateral (L)	
1	Anterior alveolar margin of the left first upper molar
2	Exposed lesser curvature of the left upper incisor
3	Most anterior point on nasal bone
4	Lateral most point of the upper incisor tooth
5	Intersection between the upper part of the zygomatic arc and post-orbital constriction
Ventral (V)	
1	Rostralmost point of the upper incisor tooth next to the midline
2	Rostral end of the rostral palatine fissure
3	Upper end of the premaxilar and maxilar suture
4	Caudal end of the palatine fissure
5	Caudalmost point of the suture between palatine bones and the rostral border of mesopterygoid fossa
6	Intersection of the rostral curvature of the nasal process of the incisive and nasal bones in a ventral projection
7	Lower end of the premaxilar and maxilar suture
8	Rostral end of the zygomatic plate
9	Caudal end of the zygomatic plate
10	Jugal end of M1 molar
11	Jugal end of M2 molar
12	Caudalmost point of the molar row
13	Lingual end of M2 molar
14	Lingual end of M1 molar
15	Rostralmost point of the molar row
Mandíbula (J)	
1	Mandibular symphysis
2	Uppermost and most anterior point of the molar row
3	Most posterior point of the molar row
4	Most anterior point of the molar row
5	Diagastic process
6	Lower point of the articular process
7	Deepest point of mandibular notch
8	Most posterior point of angular process
9	Most posteroinferior point of condylar process
10	Most posterosuperior point of condylar process
11	Deepest point of sigmoid notch
12	Tip of coronoid process
13	Contact point between the lower part of the coronoid process and the molar row
Primer molar inferior (M1)	
1	Most anterior part
2	Lingual reentrant angle 1. LRA1
3	Triangle 1. T1
4	Lingual reentrant angle 2. LRA2
5	Triangle 3. T3
6	Lingual reentrant angle 3. LRA3

7	Triangle 5. T5
8	Lingual reentrant angle 4. LRA4
9	Triangle 7. T7
10	Lingual reentrant angle 5. LRA5
11	Talonid lingual triangle
12	Most posterior part
13	Talonid bucal triangle
14	Bucal reentrant angle 4. BRA4
15	Triangle 6. T6
16	Bucal reentrant angle 3. BRA3
17	Triangle 4. T4
18	Bucal reentrant angle 2. BRA2
19	Triangle 2. T2
20	Bucal reentrant angle 1. BRA1
Tercer molar superior (M3)	
1	Most anterior part
2	Triangle 1. T1
3	Bucal reentrant angle 1. BRA1
4	Triangle 3. T3
5	Bucal reentrant angle 2. BRA2
6	Triangle 5. T5
7	Bucal reentrant angle 3. BRA3
8	Most posterior part
9	Lingual reentrant angle 3. LRA3
10	Triangle 6. T6
11	Lingual reentrant angle 2. LRA2
12	Triangle 4. T4
13	Lingual reentrant angle 3. LRA1
14	Triangle 2. T2

VI. Apéndice

Tabla A5. Valores del primer y del segundo componente de deformación (RW1 y RW2) para cada uno de los cinco caracteres para todos los individuos. En negrita, valores medios utilizados para completar la información que les faltaba a algunos de los individuos (a partir de la información de los individuos de referencia). ZL1, ZL2, coeficientes del lateral; ZV1, ZV2, coeficientes del ventral; ZJ1, ZJ2, coeficientes de la mandíbula; ZM1, ZM2, coeficientes del primer molar inferior y ZM31, ZM32 coeficientes del tercer molar superior.

Código	ZL1	ZL2	ZV1	ZV2	ZJ1	ZJ2	ZM11	ZM12	ZM31	ZM32
C1210	0.0430	-0.0032	0.0217	-0.0483	0.0089	-0.0075	-0.0502	0.0052	-0.0252	0.0045
C1221	0.0403	0.0265	-0.0410	0.0052	0.0292	-0.0026	-0.0343	0.0055	-0.0330	-0.0183
C1222	0.0479	0.0338	-0.0328	0.0132	0.0044	-0.0060	-0.0022	0.0096	-0.0297	-0.0496
C1223	0.0216	0.0260	-0.0373	0.0171	0.0079	-0.0110	-0.0058	-0.0039	-0.0579	0.0640
C1230	0.0003	0.0184	-0.0480	0.0388	-0.0019	0.0222	-0.0058	-0.0039	-0.0381	0.0077
C1231	-0.0099	-0.0120	-0.0288	0.0105	0.0188	0.0125	-0.0148	-0.0175	-0.0739	-0.0055
C1234	0.0115	-0.0137	-0.0074	0.0188	0.0228	-0.0247	-0.0249	-0.0116	-0.0695	0.0044
C1235	-0.0280	-0.0117	-0.0305	-0.0107	0.0294	-0.0172	-0.0195	0.0228	-0.0242	0.0251
C1237	0.0442	0.0247	-0.0324	0.0192	0.0414	-0.0012	-0.0150	-0.0381	-0.0167	0.0004
C1239	0.0173	0.0292	-0.0465	0.0128	0.0272	-0.0083	-0.0161	-0.0350	-0.0166	0.0034
C1241	0.0160	-0.0459	-0.0436	-0.0228	0.0147	0.0119	-0.0058	-0.0039	-0.0169	0.0026
C1245	0.0700	-0.0281	-0.0387	0.0233	0.0261	-0.0002	-0.0454	-0.0169	-0.0509	0.0136
C1246	0.0247	0.0006	-0.0318	-0.0170	0.0282	-0.0026	-0.0518	0.0161	-0.0819	0.0419
C1247	0.0161	-0.0043	-0.0045	-0.0105	0.0091	-0.0067	-0.0149	-0.0386	-0.0169	0.0128
C1250	-0.0076	-0.0134	0.0084	-0.0157	0.0089	-0.0075	0.0139	0.0186	0.0110	0.0412
C1251	-0.0187	0.0011	-0.0428	-0.0108	0.0089	-0.0075	-0.0058	-0.0039	0.0446	0.0327
C1252	0.0147	0.0168	0.0212	-0.0184	0.0089	-0.0075	-0.0252	0.0044	-0.0132	-0.0075
C1257	-0.0508	-0.0344	0.0003	-0.0607	0.0309	0.0477	-0.0042	0.0036	-0.0117	0.0013
C1258	0.0569	0.0257	-0.0004	-0.0271	0.0089	-0.0075	-0.0161	0.0345	-0.0252	0.0045
C1259	0.0433	-0.0066	-0.0159	0.0065	0.0209	0.0062	-0.0007	0.0224	-0.0379	0.0042
C1264	0.0117	0.0367	-0.0418	-0.0131	0.0052	-0.0173	0.0097	-0.0219	-0.0571	0.0042
C1265	0.0043	0.0032	-0.0240	0.0336	0.0119	0.0427	-0.0226	0.0001	-0.0367	0.0049
C1266	0.0022	0.0133	-0.0208	-0.0165	0.0308	-0.0233	-0.0265	-0.0086	-0.0522	0.0086
C1268	0.0179	0.0145	-0.0118	-0.0021	0.0288	-0.0069	-0.0058	-0.0039	-0.0516	-0.0011
C1270	0.0043	0.0032	-0.0339	0.0321	0.0363	-0.0156	-0.0163	-0.0037	-0.0252	0.0045
C1272	0.0110	-0.0409	-0.0201	0.0285	0.0151	0.0220	-0.0320	-0.0304	-0.0894	0.0054
C1273	0.0258	0.0144	-0.0364	0.0054	0.0089	-0.0226	0.0057	-0.0020	-0.0135	0.0352
C1277	0.0413	0.0325	-0.0090	-0.0096	0.0321	-0.0314	-0.0058	-0.0039	0.0336	0.0632
C1281	0.0360	-0.0284	-0.0135	-0.0327	0.0089	-0.0075	-0.0058	-0.0039	-0.0346	0.0069
C1282	0.0358	0.0064	-0.0047	-0.0031	0.0089	-0.0075	-0.0058	-0.0039	-0.0251	0.0104
C1283	0.0285	0.0139	-0.0181	0.0275	0.0029	0.0165	0.0172	0.0256	-0.0286	0.0338
C1284	0.0417	-0.0355	-0.0344	0.0148	0.0089	-0.0075	-0.0173	-0.0333	-0.0298	-0.0153
C1285	0.0313	-0.0203	-0.0268	-0.0091	0.0089	-0.0075	-0.0058	-0.0039	0.0143	-0.0193
C1288	0.0013	-0.0209	-0.0029	-0.0047	0.0089	-0.0075	0.0097	0.0423	0.0452	0.0630
C1291	0.0201	0.0202	0.0040	0.0668	-0.0140	-0.0005	-0.0195	-0.0120	0.0771	0.0150
C1292	-0.0078	0.0020	-0.0146	-0.0186	0.0370	-0.0187	-0.0142	0.0001	0.0385	0.0319
C1293	-0.0214	0.0025	-0.0258	0.0403	0.0089	-0.0075	-0.0163	-0.0085	0.0011	0.0009
C1294	0.0283	0.0130	-0.0273	-0.0365	0.0315	0.0251	-0.0015	0.0210	0.0277	-0.0031
C1296	-0.0006	-0.0139	-0.0179	0.0093	-0.0092	0.0066	0.0013	-0.0155	0.0091	-0.0238
C1297	-0.0281	0.0003	0.0081	-0.0227	-0.0237	-0.0155	-0.0103	0.0331	-0.1358	0.0049
C1300	0.0366	0.0132	-0.0019	-0.0220	0.0046	0.0391	-0.0024	0.0021	-0.0498	-0.0074
C1301	0.0080	-0.0164	-0.0017	-0.0116	0.0070	0.0214	-0.0388	0.0017	0.0265	0.0663
C1308	0.0445	-0.0378	-0.0032	-0.0069	0.0089	-0.0075	-0.0058	-0.0039	0.0022	0.0665
C1327	0.0043	0.0032	0.0054	-0.0004	0.0089	-0.0075	-0.0469	0.0259	0.0090	0.0276
C1334	0.0154	-0.0135	-0.0064	-0.0121	0.0206	-0.0081	-0.0081	-0.0085	0.0237	0.0267
C1335	-0.0103	0.0041	-0.0045	0.0194	0.0089	-0.0075	0.0502	-0.0062	-0.0107	-0.0148

C1336	0.0269	0.0128	-0.0173	-0.0372	0.0252	0.0160	-0.0179	0.0155	0.0538	-0.0064
C1337	0.0015	-0.0520	0.0208	-0.0219	0.0330	0.0079	-0.0320	-0.0100	-0.0149	-0.0089
C1339	-0.0425	0.0135	0.0107	0.0186	-0.0084	-0.0014	0.0515	0.0113	0.0334	-0.0169
C1341	-0.0273	0.0198	0.0050	0.0209	-0.0016	0.0127	0.0240	-0.0262	0.0108	-0.0844
C1346	0.0427	-0.0166	-0.0115	-0.0142	0.0303	-0.0031	-0.0234	-0.0101	-0.1035	0.0213
C1349	0.0317	-0.0054	-0.0162	-0.0215	0.0089	-0.0075	-0.0058	-0.0039	0.0005	0.0170
C1352	0.0373	0.0316	-0.0212	0.0042	0.0253	-0.0253	-0.0058	-0.0039	-0.0243	0.0034
C1356	0.0154	-0.0218	-0.0251	-0.0080	0.0381	-0.0100	-0.0180	0.0079	-0.0386	0.0271
C1357	0.0552	-0.0254	-0.0420	-0.0031	0.0089	-0.0075	0.0007	-0.0020	-0.0031	0.0468
C1359	-0.0046	0.0041	-0.0285	0.0396	0.0089	-0.0075	-0.0236	-0.0088	-0.0848	0.0224
C1361	0.0342	-0.0500	-0.0243	0.0189	0.0089	-0.0075	-0.0022	0.0230	-0.0855	0.0403
C1369	0.0387	0.0037	-0.0277	0.0443	0.0123	0.0038	-0.0222	-0.0199	-0.0081	0.0304
C1374	-0.0110	0.0106	0.0195	0.0035	-0.0411	0.0264	0.0427	0.0211	0.0536	-0.0184
C1375	0.0164	-0.0122	0.0193	-0.0338	-0.0137	-0.0132	0.0365	0.0033	0.0265	0.0369
C1376	0.0332	0.0063	0.0123	-0.0055	0.0314	-0.0123	-0.0449	0.0018	0.0082	0.0301
C1377	-0.0324	0.0187	0.0129	-0.0181	0.0089	-0.0075	0.0175	0.0033	-0.0014	-0.0923
C1379	-0.0072	-0.0212	-0.0250	-0.0083	0.0149	0.0273	-0.0215	-0.0049	0.0246	0.0244
C1380	0.0149	-0.0284	-0.0161	-0.0213	0.0202	0.0079	-0.0241	-0.0219	-0.0068	0.0343
C1407	0.0192	0.0262	-0.0388	-0.0287	0.0089	-0.0075	-0.0330	0.0216	0.0111	0.0286
C1408	0.0313	0.0240	-0.0316	-0.0097	0.0027	0.0114	-0.0237	-0.0046	0.0354	0.0211
C1410	-0.0124	0.0026	-0.0012	0.0185	-0.0392	0.0077	0.0437	0.0103	0.0419	-0.0183
C1415	0.0000	-0.0126	-0.0169	-0.0189	0.0153	0.0141	-0.0082	0.0075	-0.0146	0.0600
C1429	0.0013	-0.0193	-0.0186	0.0006	0.0251	0.0379	0.0120	-0.0223	-0.0154	-0.0316
C1438	0.0162	0.0276	-0.0130	-0.0085	0.0375	-0.0012	-0.0367	-0.0344	-0.0743	0.0150
C1445	-0.0158	0.0029	0.0116	-0.0086	0.0089	-0.0075	0.0213	0.0017	0.0092	-0.0098
C1454	-0.0080	0.0096	-0.0127	0.0427	0.0417	-0.0227	0.0112	-0.0260	-0.0115	-0.0467
C1934	0.0659	0.0118	0.0618	0.0122	0.0089	-0.0075	0.0284	0.0309	0.0084	0.0275
C1935	0.0196	-0.0008	0.0241	0.0117	-0.0067	0.0280	0.0137	0.0097	0.0310	0.0087
C1937	0.0249	0.0393	0.0025	0.0070	0.0112	-0.0294	0.0001	-0.0202	-0.0145	-0.0458
C1939	0.0356	0.0119	-0.0265	-0.0402	0.0089	-0.0075	-0.0122	0.0204	0.0042	0.0049
C1940	0.0062	0.0151	-0.0181	-0.0275	0.0297	0.0111	-0.0021	-0.0194	0.0302	0.0122
C1945	-0.0233	-0.0003	0.0360	0.0033	-0.0348	-0.0101	0.0184	0.0278	-0.0133	0.0388
C1946	-0.0322	-0.0091	0.0174	-0.0163	-0.0196	-0.0141	0.0067	0.0190	0.0126	-0.0057
C1948	-0.0237	0.0325	-0.0128	-0.0197	-0.0087	0.0070	-0.0328	-0.0214	-0.0223	-0.0187
C1950	-0.0207	0.0198	0.0100	0.0055	-0.0074	-0.0330	-0.0201	0.0359	0.0305	-0.0066
C1951	-0.0270	0.0018	0.0152	-0.0510	-0.0311	-0.0113	-0.0008	0.0393	-0.0180	-0.0124
C1957	-0.0374	-0.0102	0.0295	-0.0575	-0.0122	-0.0158	0.0164	0.0137	0.0490	-0.0155
C1958	0.0057	-0.0251	0.0192	-0.0023	-0.0161	-0.0256	0.0031	-0.0065	0.0172	-0.0227
C1959	0.0189	-0.0041	0.0436	-0.0024	-0.0221	0.0042	0.0078	0.0013	-0.0222	-0.0259
C1973	0.0255	-0.0024	-0.0295	0.0077	0.0326	0.0058	0.0225	-0.0207	0.0282	0.0152
C1975	-0.0265	0.0261	0.0037	0.0055	-0.0300	-0.0087	0.0235	0.0257	0.0330	0.0105
C1978	0.0152	0.0077	-0.0272	-0.0409	0.0089	-0.0075	-0.0378	-0.0090	0.0482	-0.0003
C1983	0.0481	0.0350	0.0244	0.0021	-0.0445	0.0287	0.0144	-0.0056	0.0433	0.0466
C1987	-0.0199	-0.0260	0.0105	0.0232	-0.0211	-0.0299	-0.0257	0.0333	0.0744	-0.0153
C1988	-0.0187	0.0105	0.0127	0.0033	-0.0344	-0.0079	0.0058	-0.0045	0.0078	-0.0291
C1994	-0.0180	0.0126	0.0105	0.0264	0.0089	-0.0075	-0.0089	-0.0200	0.0406	0.0344
C1996	-0.0041	0.0275	0.0358	0.0232	-0.0010	-0.0089	0.0031	0.0074	-0.0030	0.0077
C2003	0.0105	-0.0102	0.0492	0.0072	0.0089	-0.0075	0.0214	-0.0188	-0.0348	0.0030
C2007	-0.0550	-0.0212	0.0176	0.0019	-0.0236	0.0093	0.0211	-0.0227	-0.0125	-0.0413
C2009	-0.0439	0.0110	0.0043	0.0153	-0.0021	-0.0051	0.0155	0.0187	0.0163	0.0015
C2010	-0.0174	0.0136	0.0161	0.0102	-0.0130	-0.0267	0.0237	-0.0012	0.0246	-0.0141
C2015	-0.0228	-0.0088	0.0360	-0.0093	-0.0176	-0.0030	0.0179	-0.0147	-0.0299	-0.0385
C2028	-0.0495	-0.0366	0.0041	0.0039	0.0089	-0.0075	-0.0124	-0.0305	-0.0212	-0.0408
C2029	-0.0162	0.0324	0.0117	-0.0011	-0.0187	-0.0030	0.0167	0.0121	-0.0022	-0.0002

VI. Apéndice

C2049	-0.0353	-0.0015	0.0090	-0.0046	-0.0198	0.0171	0.0262	-0.0047	0.0132	-0.0156
C2053	-0.0131	0.0132	0.0000	-0.0014	-0.0149	-0.0180	0.0341	-0.0047	0.0239	-0.0259
C2057	0.0013	-0.0453	0.0640	-0.0351	-0.0281	0.0075	-0.0075	-0.0146	0.0352	-0.0147
C2067	-0.0108	-0.0200	0.0307	0.0194	-0.0366	0.0064	-0.0055	0.0054	0.0034	-0.0663
C2068	-0.0156	-0.0607	0.0242	-0.0139	0.0089	-0.0075	0.0024	0.0109	-0.0088	-0.0528
C2070	-0.0339	-0.0006	0.0328	0.0343	-0.0292	0.0430	0.0185	-0.0236	0.0417	-0.0220
C2072	-0.0301	0.0067	0.0265	0.0148	-0.0398	0.0251	0.0373	-0.0193	0.0122	-0.0240
C2073	-0.0135	-0.0356	0.0063	-0.0053	-0.0075	-0.0377	0.0186	-0.0065	-0.0048	-0.0449
C2096	0.0063	-0.0133	0.0161	0.0022	0.0089	-0.0075	0.0284	-0.0265	-0.0393	-0.0278
C2101	-0.0263	-0.0196	0.0155	0.0356	-0.0175	0.0058	0.0090	-0.0190	0.0075	-0.0455
C2102	-0.0306	-0.0262	0.0081	0.0109	-0.0248	-0.0016	-0.0021	-0.0027	0.0360	-0.0162
C2111	-0.0255	0.0021	0.0219	-0.0089	-0.0411	0.0158	0.0013	-0.0152	0.0221	-0.0142
C2114	0.0122	-0.0045	0.0407	0.0218	0.0089	-0.0075	-0.0110	0.0242	-0.0350	-0.0114
C2118	-0.0009	0.0163	0.0157	-0.0040	-0.0175	-0.0030	-0.0172	0.0260	0.0296	-0.0047
C2120	0.0020	-0.0062	0.0256	-0.0060	-0.0182	-0.0271	0.0041	0.0082	0.0057	-0.0070
C2131	-0.0374	0.0175	0.0036	0.0165	0.0003	-0.0414	0.0388	-0.0110	-0.0665	0.0287
C2132	0.0000	-0.0191	0.0116	0.0002	-0.0125	-0.0280	-0.0234	0.0082	-0.0259	-0.0323
C2137	-0.0016	0.0556	0.0012	-0.0023	-0.0107	-0.0297	0.0474	-0.0228	-0.0084	-0.0186
C2146	-0.0477	0.0232	-0.0019	0.0011	-0.0196	-0.0195	0.0492	0.0077	0.0144	-0.0233
C2147	-0.0277	-0.0224	0.0151	-0.0158	-0.0018	-0.0169	0.0084	0.0075	-0.0122	-0.0041
C2148	-0.0244	0.0163	0.0156	-0.0111	0.0089	-0.0075	0.0168	-0.0123	0.0015	-0.0608
C2155	-0.0157	-0.0006	0.0113	-0.0055	-0.0194	-0.0183	0.0292	-0.0202	0.0567	-0.0225
C2160	-0.0069	-0.0308	0.0395	-0.0001	-0.0043	-0.0095	-0.0047	0.0062	0.0539	0.0043
C2166	0.0716	-0.0014	-0.0206	-0.0054	0.0254	0.0081	-0.0424	0.0059	0.0384	0.0216
C2168	0.0348	0.0200	-0.0028	-0.0095	0.0092	0.0450	-0.0280	0.0099	0.0624	0.0469
C2169	-0.0058	0.0230	0.0297	-0.0064	-0.0079	0.0074	0.0370	-0.0026	-0.0188	-0.0219
C2172	-0.0437	0.0029	-0.0057	0.0057	-0.0256	-0.0040	0.0100	0.0254	-0.0013	0.0567
C2180	0.0090	0.0262	-0.0068	0.0156	0.0209	-0.0055	0.0051	0.0136	-0.0235	-0.0179
C2213	-0.0305	0.0064	0.0102	-0.0185	0.0089	-0.0075	-0.0236	0.0054	0.0353	-0.0047
C2214	0.0258	-0.0077	-0.0184	-0.0032	0.0374	-0.0232	0.0277	-0.0003	0.0127	-0.0061
C2216	-0.0237	0.0194	0.0296	0.0297	0.0089	-0.0075	0.0051	-0.0235	0.0054	-0.0275
C2225	-0.0038	-0.0234	0.0248	-0.0177	-0.0216	-0.0048	0.0199	-0.0162	0.0100	-0.0372
C2227	-0.0022	0.0012	0.0215	-0.0070	0.0089	-0.0075	0.0134	-0.0002	0.0028	-0.0209
C2228	0.0097	0.0071	-0.0206	0.0042	-0.0040	0.0108	-0.0329	-0.0457	0.0219	0.0112
C2229	-0.0085	-0.0054	0.0132	0.0147	-0.0309	-0.0017	-0.0026	0.0288	0.0740	0.0265
C2236	0.0245	-0.0170	-0.0041	-0.0300	0.0245	0.0097	-0.0314	0.0056	-0.0238	0.0700
C2237	-0.0443	0.0378	0.0027	0.0076	0.0089	-0.0075	-0.0058	-0.0039	0.0570	-0.0072
C2238	0.0105	-0.0427	-0.0435	-0.0203	0.0135	0.0218	-0.0293	-0.0109	-0.0019	0.0032
C2263	-0.0442	-0.0180	0.0037	-0.0159	-0.0205	-0.0046	0.0195	0.0074	0.0290	-0.0139
C2264	-0.0172	0.0235	0.0001	0.0243	-0.0038	-0.0122	-0.0111	-0.0018	-0.0114	-0.0106
C2297	0.0585	-0.0075	0.0461	0.0152	0.0135	0.0156	-0.0054	0.0566	0.0310	0.0129
C2311	-0.0141	-0.0086	-0.0052	0.0000	-0.0423	0.0257	0.0135	0.0219	0.0412	-0.0274
C2338	-0.0272	0.0038	0.0038	0.0314	-0.0239	0.0108	0.0034	0.0319	0.0183	-0.0120
C2343	0.0215	0.0127	0.0261	-0.0215	-0.0105	-0.0324	-0.0102	-0.0119	0.0093	0.0446
C2345	-0.0484	-0.0082	0.0080	0.0276	-0.0359	-0.0083	0.0113	0.0245	-0.0168	-0.0763
C2347	-0.0473	0.0256	-0.0091	0.0141	0.0089	-0.0075	0.0283	-0.0137	-0.0112	-0.0482
C2349	-0.0469	0.0229	0.0032	-0.0210	0.0089	-0.0075	0.0099	-0.0066	-0.0435	-0.0582
C2350	-0.0324	-0.0072	0.0268	0.0056	-0.0143	-0.0047	0.0201	-0.0150	0.0432	-0.0497
C2358	-0.0165	-0.0029	0.0059	0.0253	-0.0238	0.0233	0.0377	-0.0297	-0.0039	-0.0515
C2359	-0.0219	0.0217	0.0003	0.0381	-0.0076	-0.0097	0.0132	-0.0158	-0.0031	-0.0065
C2364	-0.0524	0.0090	0.0060	0.0154	-0.0277	-0.0055	0.0327	0.0259	0.0456	0.0374
C2367	0.0260	0.0122	0.0408	0.0384	0.0089	-0.0075	0.0168	-0.0189	0.0358	0.0017
C2377	-0.0249	-0.0111	-0.0412	0.0042	-0.0030	0.0384	-0.0196	0.0117	0.0543	0.0223
C2378	0.0020	-0.0173	-0.0074	-0.0028	0.0112	0.0043	0.0133	-0.0059	0.0161	-0.0185

C2380	-0.0152	0.0105	-0.0186	0.0096	0.0215	0.0033	-0.0143	-0.0090	-0.0100	0.0330
C2384	-0.0025	-0.0350	-0.0122	-0.0299	0.0287	0.0270	-0.0199	0.0256	0.0070	0.0118
C2385	0.0403	-0.0160	-0.0172	-0.0108	0.0163	0.0193	-0.0079	0.0067	-0.0252	0.0045
C2388	0.0054	0.0175	0.0117	-0.0022	-0.0104	-0.0100	-0.0222	0.0267	0.0186	0.0143
C2390	0.0043	0.0032	0.0203	-0.0160	0.0089	-0.0075	0.0281	-0.0130	0.0000	0.0229
C2391	-0.0183	0.0084	0.0078	0.0083	-0.0245	-0.0005	0.0195	0.0108	0.0233	-0.0026
C2395	-0.0225	-0.0020	0.0004	0.0087	-0.0032	-0.0149	0.0282	0.0212	0.0420	0.0004
C2399	-0.0096	-0.0083	0.0215	-0.0152	-0.0061	-0.0305	-0.0060	-0.0028	0.0201	-0.0166
C2401	-0.0395	-0.0486	0.0222	0.0151	-0.0203	-0.0363	0.0042	0.0197	0.0156	-0.0274
C2402	0.0306	-0.0039	-0.0234	0.0171	0.0406	-0.0178	-0.0144	0.0016	0.0209	-0.0305
C2403	0.0158	-0.0239	-0.0113	-0.0119	0.0255	0.0211	-0.0344	-0.0034	-0.0073	0.0553
C2429	0.0472	0.0052	-0.0052	-0.0155	0.0379	-0.0014	-0.0451	-0.0122	-0.0174	0.0363
C2430	-0.0058	0.0160	0.0294	0.0006	-0.0166	0.0192	0.0265	-0.0081	-0.0379	-0.0180
C2431	-0.0197	0.0220	-0.0006	0.0133	-0.0160	0.0126	0.0392	-0.0317	0.0099	0.0138
C2432	0.0346	-0.0022	-0.0134	-0.0154	0.0259	0.0044	-0.0185	-0.0463	-0.0139	0.0204
C2435	0.0028	-0.0209	-0.0008	0.0205	0.0061	0.0504	-0.0222	0.0154	0.0542	0.0365
C2436	-0.0513	0.0065	0.0164	0.0071	-0.0371	0.0254	0.0071	-0.0150	-0.0167	-0.0310
C2437	-0.0371	0.0012	0.0224	0.0091	-0.0055	-0.0029	-0.0151	-0.0156	0.0337	-0.0346
C2441	-0.0407	0.0405	0.0283	-0.0049	-0.0113	-0.0256	0.0379	0.0183	0.0246	0.0048
C2448	0.0043	0.0032	0.0106	-0.0079	0.0089	-0.0075	0.0102	0.0302	-0.0248	-0.0471
C2451	-0.0427	-0.0053	0.0068	0.0059	-0.0207	-0.0075	0.0118	0.0092	0.0282	-0.0102
C2457	-0.0428	0.0096	-0.0014	0.0012	0.0089	-0.0075	-0.0025	-0.0182	0.0330	-0.0214
C2486	0.0354	0.0027	0.0288	0.0104	0.0089	-0.0075	-0.0145	0.0312	0.0260	0.0358
C2489	-0.0048	0.0164	0.0142	-0.0137	-0.0315	-0.0053	0.0194	0.0083	0.0251	-0.0276
C2496	-0.0408	-0.0043	0.0269	0.0381	-0.0310	-0.0131	0.0121	-0.0120	0.0049	-0.0887

Tabla A6. Asignación de especie a los individuos de las zonas de distribución alopatrica y de las zonas de distribución solapada mediante un análisis discriminante (LDA) y un algoritmo E-M (*expectation-maximization*). La línea horizontal divide a los individuos de las zonas de distribución alopatrica (parte superior) de los individuos de las zonas de distribución solapada (parte inferior). Se representan los coeficientes de asignación posterior de los dos métodos para todos los individuos. En negrita se representan los individuos en los que la asignación de especie fue opuesta entre los dos métodos. Localidades de muestreo: **Bel:** Belianes; **Ceb:** San Cebrián de Mazote; **Esc:** Escorial de la Sierra; **Hos:** Hospital de Incio; **Lum:** Lumbrales; **Men:** Menaza; **Mun:** Muñíos; **Oss:** Ossó de Sió; **Pal:** Palacios de Sil; **Pob:** Poblaciones de Arreba; **Por:** Porqueriza; **Sal:** Salvatierra de Tormes; **Toa:** Torrejón de Alba; **Tol:** Toledo y **Vil:** Villacantid.

Código individuo	Localidad	Especie LDA	Coef. Duo LDA	Coef. Lus LDA	Especie E-M	Coef. Duo E-M	Coef. Lus E-M
C1210	Bel	Lus	0,814582	0,185418	Duo	0,996950	0,003050
C1221	Bel	Duo	1,000000	0,000000	Duo	0,999862	0,000138
C1222	Bel	Duo	0,999620	0,000380	Duo	0,958173	0,041827
C1223	Bel	Duo	0,999997	0,000003	Duo	0,999964	0,000036
C1230	Bel	Duo	0,999999	0,000001	Duo	0,993855	0,006145
C1231	Bel	Duo	0,999996	0,000004	Duo	0,998011	0,001989
C1234	Bel	Duo	0,995155	0,004845	Duo	0,998433	0,001567
C1235	Bel	Duo	0,999979	0,000021	Duo	0,992229	0,007771
C1237	Bel	Duo	1,000000	0,000000	Duo	0,999872	0,000128
C1239	Bel	Duo	1,000000	0,000000	Duo	0,999381	0,000619
C1241	Bel	Duo	1,000000	0,000000	Duo	0,998991	0,001009
C1245	Bel	Duo	1,000000	0,000000	Duo	1,000000	0,000000
C1246	Bel	Duo	1,000000	0,000000	Duo	0,999999	0,000001
C1247	Bel	Duo	0,991589	0,008411	Duo	0,984845	0,015155
C1264	Bel	Duo	0,999936	0,000064	Duo	0,996530	0,003470
C1265	Bel	Duo	0,999998	0,000002	Duo	0,996400	0,003600
C1266	Bel	Duo	0,999934	0,000066	Duo	0,999139	0,000861
C1268	Bel	Duo	0,999509	0,000491	Duo	0,996359	0,003641
C1270	Bel	Duo	0,999997	0,000003	Duo	0,997684	0,002316
C1272	Bel	Duo	1,000000	0,000000	Duo	0,999922	0,000078
C1273	Bel	Duo	0,999960	0,000040	Duo	0,997961	0,002039
C1277	Bel	Duo	0,999303	0,000697	Duo	0,998741	0,001259
C2028	Hos	Lus	0,016923	0,983077	Lus	0,014898	0,985102
C2029	Hos	Lus	0,000009	0,999991	Lus	0,004867	0,995133
C2096	Hos	Lus	0,002738	0,997262	Lus	0,104069	0,895931
C2131	Hos	Lus	0,000018	0,999982	Lus	0,117057	0,882943
C2132	Hos	Lus	0,000207	0,999793	Lus	0,091821	0,908179
C2137	Hos	Lus	0,000006	0,999994	Lus	0,003059	0,996941
C2146	Hos	Lus	0,000000	1,000000	Lus	0,000024	0,999976
C2147	Hos	Lus	0,000146	0,999854	Lus	0,016797	0,983203
C2148	Hos	Lus	0,000030	0,999970	Lus	0,000433	0,999567
C2345	Hos	Lus	0,000000	1,000000	Lus	0,000005	0,999995
C2347	Hos	Lus	0,001592	0,998408	Lus	0,000783	0,999217
C2349	Hos	Lus	0,000484	0,999516	Lus	0,003107	0,996893
C2350	Hos	Lus	0,000000	1,000000	Lus	0,000007	0,999993
C2364	Hos	Lus	0,000001	0,999999	Lus	0,000126	0,999874
C2390	Hos	Lus	0,002189	0,997811	Lus	0,188696	0,811304
C2441	Hos	Lus	0,000000	1,000000	Lus	0,000029	0,999971
C2448	Hos	Lus	0,002656	0,997344	Lus	0,032251	0,967749
C2451	Hos	Lus	0,000006	0,999994	Lus	0,000230	0,999770
C2457	Hos	Lus	0,008478	0,991522	Lus	0,003171	0,996829
C1281	Oss	Duo	0,999896	0,000104	Duo	0,998126	0,001874

C1282	Oss	Duo	0,986017	0,013983	Duo	0,990255	0,009745
C1283	Oss	Duo	0,999041	0,000959	Duo	0,990408	0,009592
C1284	Oss	Duo	1,000000	0,000000	Duo	0,999238	0,000762
C1285	Oss	Duo	0,999940	0,000060	Duo	0,941947	0,058053
C1288	Oss	Duo	0,862781	0,137219	Duo	0,711858	0,288142
C1308	Oss	Duo	0,999783	0,000217	Duo	0,999555	0,000445
C1327	Oss	Duo	0,770092	0,229908	Duo	0,954201	0,045799
C1346	Oss	Duo	1,000000	0,000000	Duo	0,999997	0,000003
C1349	Oss	Duo	0,999718	0,000282	Duo	0,991925	0,008075
C1352	Oss	Duo	0,999640	0,000360	Duo	0,997275	0,002725
C1356	Oss	Duo	1,000000	0,000000	Duo	0,999885	0,000115
C1357	Oss	Duo	1,000000	0,000000	Duo	0,999952	0,000048
C1359	Oss	Duo	0,999880	0,000120	Duo	0,999496	0,000504
C1361	Oss	Duo	0,999997	0,000003	Duo	0,999967	0,000033
C1369	Oss	Duo	0,999998	0,000002	Duo	0,999644	0,000356
C2238	Tol	Duo	1,000000	0,000000	Duo	0,999312	0,000688
C2403	Tol	Duo	1,000000	0,000000	Duo	0,999895	0,000105
C2429	Tol	Duo	0,999999	0,000001	Duo	0,999984	0,000016
C1407	Ceb	Duo	0,999999	0,000001	Duo	0,998947	0,001053
C1408	Ceb	Duo	0,999997	0,000003	Duo	0,994013	0,005987
C1250	Esc	Lus	0,102799	0,897201	Lus	0,449140	0,550860
C1983	Esc	Lus	0,000082	0,999918	Lus	0,130082	0,869918
C2213	Esc	Lus	0,007445	0,992555	Lus	0,023733	0,976267
C2237	Esc	Lus	0,000674	0,999326	Lus	0,001137	0,998863
C1375	Lum	Lus	0,000200	0,999800	Lus	0,062611	0,937389
C1945	Lum	Lus	0,000000	1,000000	Lus	0,002458	0,997542
C1946	Lum	Lus	0,000001	0,999999	Lus	0,000882	0,999118
C1948	Lum	Duo	0,723397	0,276603	Duo	0,511867	0,488133
C1959	Lum	Lus	0,000000	1,000000	Lus	0,006484	0,993516
C1987	Lum	Lus	0,000003	0,999997	Lus	0,000294	0,999706
C2007	Lum	Lus	0,000001	0,999999	Lus	0,000052	0,999948
C2049	Lum	Lus	0,000063	0,999937	Lus	0,000446	0,999554
C2057	Lum	Lus	0,000000	1,000000	Lus	0,000275	0,999725
C2070	Lum	Lus	0,000000	1,000000	Lus	0,000015	0,999985
C2072	Lum	Lus	0,000000	1,000000	Lus	0,000016	0,999984
C2111	Lum	Lus	0,000001	0,999999	Lus	0,000278	0,999722
C2160	Lum	Lus	0,000001	0,999999	Lus	0,001618	0,998382
C2225	Lum	Lus	0,000002	0,999998	Lus	0,000551	0,999449
C2343	Lum	Lus	0,000068	0,999932	Lus	0,644702	0,355298
C2359	Lum	Lus	0,000304	0,999696	Lus	0,011372	0,988628
C2395	Lum	Lus	0,000336	0,999664	Lus	0,001319	0,998681
C2401	Lum	Lus	0,000000	1,000000	Lus	0,000064	0,999936
C2430	Lum	Lus	0,000003	0,999997	Lus	0,006508	0,993492
C2436	Lum	Lus	0,000001	0,999999	Lus	0,000129	0,999871
C2437	Lum	Lus	0,000004	0,999996	Lus	0,000281	0,999719
C2489	Lum	Lus	0,000001	0,999999	Lus	0,000266	0,999734
C2496	Lum	Lus	0,000000	1,000000	Lus	0,000001	0,999999
C1937	Men	Lus	0,007927	0,992073	Lus	0,202852	0,797148
C1950	Men	Lus	0,000008	0,999992	Lus	0,004485	0,995515
C1951	Men	Lus	0,000002	0,999998	Lus	0,003405	0,996595
C1957	Men	Lus	0,000000	1,000000	Lus	0,000043	0,999957
C1958	Men	Lus	0,000009	0,999991	Lus	0,005182	0,994818
C1988	Men	Lus	0,000000	1,000000	Lus	0,000389	0,999611
C2009	Men	Lus	0,000122	0,999878	Lus	0,001626	0,998374

VI. Apéndice

C2010	Men	Lus	0,000000	1,000000	Lus	0,000321	0,999679
C2053	Men	Lus	0,000030	0,999970	Lus	0,000542	0,999458
C2067	Men	Lus	0,000000	1,000000	Lus	0,000045	0,999955
C2068	Men	Lus	0,000417	0,999583	Lus	0,004052	0,995948
C2073	Men	Lus	0,000031	0,999969	Lus	0,001813	0,998187
C2101	Men	Lus	0,000006	0,999994	Lus	0,000231	0,999769
C2102	Men	Lus	0,000043	0,999957	Lus	0,000583	0,999417
C2118	Men	Lus	0,000077	0,999923	Lus	0,013451	0,986549
C2155	Men	Lus	0,000001	0,999999	Lus	0,000075	0,999925
C2172	Men	Lus	0,001431	0,998569	Lus	0,070640	0,929360
C2263	Men	Lus	0,000034	0,999966	Lus	0,000219	0,999781
C2311	Men	Lus	0,000729	0,999271	Lus	0,000331	0,999669
C2391	Men	Lus	0,000014	0,999986	Lus	0,001094	0,998906
C2399	Men	Lus	0,000006	0,999994	Lus	0,005166	0,994834
C2003	Mun	Lus	0,000001	0,999999	Lus	0,070929	0,929071
C2015	Mun	Lus	0,000000	1,000000	Lus	0,000399	0,999601
C2114	Mun	Lus	0,000012	0,999988	Lus	0,164241	0,835759
C2120	Mun	Lus	0,000001	0,999999	Lus	0,006198	0,993802
C1975	Pal	Lus	0,000002	0,999998	Lus	0,000423	0,999577
C2358	Pal	Lus	0,000066	0,999934	Lus	0,000227	0,999773
C1445	Pob	Lus	0,001175	0,998825	Lus	0,011501	0,988499
C1934	Pob	Lus	0,000000	1,000000	Lus	0,185980	0,814020
C1935	Pob	Lus	0,002932	0,997068	Lus	0,042937	0,957063
C1994	Pob	Lus	0,010010	0,989990	Lus	0,126447	0,873553
C1996	Pob	Lus	0,000000	1,000000	Lus	0,013827	0,986173
C2216	Pob	Lus	0,000001	0,999999	Lus	0,001183	0,998817
C2227	Pob	Lus	0,000245	0,999755	Lus	0,015072	0,984928
C2229	Pob	Lus	0,000014	0,999986	Lus	0,001534	0,998466
C2264	Pob	Lus	0,002628	0,997372	Lus	0,078877	0,921123
C2297	Pob	Lus	0,001339	0,998661	Lus	0,434461	0,565539
C2338	Pob	Lus	0,000067	0,999933	Lus	0,001065	0,998935
C2367	Pob	Lus	0,000003	0,999997	Lus	0,010657	0,989343
C2388	Pob	Lus	0,001511	0,998489	Lus	0,203883	0,796117
C2486	Pob	Lus	0,005607	0,994393	Lus	0,753161	0,246839
C1251	Por	Duo	0,999986	0,000014	Duo	0,887384	0,112616
C1258	Por	Duo	0,984487	0,015513	Duo	0,996434	0,003566
C1259	Por	Duo	0,999979	0,000021	Duo	0,998664	0,001336
C1292	Por	Duo	0,999688	0,000312	Duo	0,951038	0,048962
C1294	Por	Duo	1,000000	0,000000	Duo	0,988517	0,011483
C1300	Por	Duo	0,999517	0,000483	Duo	0,993229	0,006771
C1336	Por	Duo	0,999992	0,000008	Duo	0,932571	0,067429
C1380	Por	Duo	0,999999	0,000001	Duo	0,999419	0,000581
C1978	Por	Duo	0,999976	0,000024	Duo	0,962602	0,037398
C2166	Por	Duo	1,000000	0,000000	Duo	0,999920	0,000080
C2180	Por	Duo	0,798096	0,201904	Duo	0,663798	0,336202
C2236	Por	Duo	0,999998	0,000002	Duo	0,999972	0,000028
C2377	Por	Duo	0,999999	0,000001	Duo	0,734595	0,265405
C2378	Por	Duo	0,920809	0,079191	Duo	0,172911	0,827089
C2402	Por	Duo	0,999983	0,000017	Duo	0,962134	0,037866
C2432	Por	Duo	0,999998	0,000002	Duo	0,999681	0,000319
C2435	Por	Duo	0,999575	0,000425	Duo	0,810740	0,189260
C1257	Sal	Duo	0,999864	0,000136	Duo	0,601583	0,398417
C1293	Sal	Duo	0,990539	0,009461	Duo	0,648930	0,351070
C1296	Sal	Duo	0,915932	0,084068	Lus	0,167926	0,832074

C1301	Sal	Duo	0,999866	0,000134	Duo	0,997375	0,002625
C1335	Sal	Lus	0,013792	0,986208	Lus	0,016006	0,983994
C1337	Sal	Duo	0,983642	0,016358	Duo	0,943691	0,056309
C1339	Sal	Lus	0,000001	0,999999	Lus	0,000017	0,999983
C1341	Sal	Lus	0,000076	0,999924	Lus	0,000047	0,999953
C1374	Sal	Lus	0,000000	1,000000	Lus	0,000010	0,999990
C1377	Sal	Lus	0,000008	0,999992	Lus	0,000045	0,999955
C1379	Sal	Duo	0,999999	0,000001	Duo	0,981768	0,018232
C1415	Sal	Duo	0,999990	0,000010	Duo	0,998810	0,001190
C1429	Sal	Duo	0,999992	0,000008	Duo	0,833715	0,166285
C1438	Sal	Duo	0,999997	0,000003	Duo	0,999966	0,000034
C1454	Sal	Duo	0,892732	0,107268	Duo	0,205679	0,794321
C1939	Sal	Duo	0,999973	0,000027	Duo	0,993991	0,006009
C1940	Sal	Duo	0,999982	0,000018	Duo	0,953518	0,046482
C1973	Sal	Duo	0,999999	0,000001	Duo	0,981673	0,018327
C2168	Sal	Duo	0,999972	0,000028	Duo	0,989561	0,010439
C2169	Sal	Lus	0,000001	0,999999	Lus	0,001585	0,998415
C2214	Sal	Duo	0,999474	0,000526	Duo	0,884528	0,115472
C2228	Sal	Duo	0,999734	0,000266	Duo	0,966005	0,033995
C2380	Sal	Duo	0,999739	0,000261	Duo	0,984052	0,015948
C2384	Sal	Duo	0,999998	0,000002	Duo	0,985454	0,014546
C2385	Sal	Duo	0,999998	0,000002	Duo	0,998891	0,001109
C2431	Sal	Lus	0,001704	0,998296	Lus	0,009064	0,990936
C1252	Toa	Lus	0,012300	0,987700	Duo	0,650738	0,349262
C1291	Toa	Lus	0,005457	0,994543	Lus	0,045700	0,954300
C1297	Toa	Lus	0,000218	0,999782	Lus	0,863785	0,136215
C1334	Toa	Duo	0,998344	0,001656	Duo	0,964981	0,035019
C1376	Toa	Duo	0,993234	0,006766	Duo	0,998096	0,001904
C1410	Vil	Lus	0,000006	0,999994	Lus	0,000061	0,999939

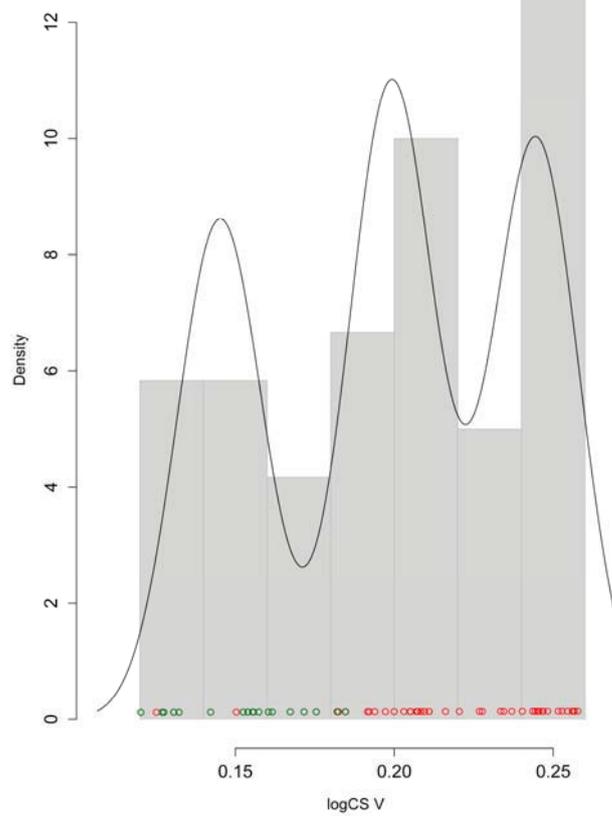


Figura A4. Distribución de frecuencias de los valores de logCS de V de los individuos de las dos especies de topillos, *M. duodecimcostatus* (rojo) y *M. lusitanicus* (verde) de las zonas de distribución alopátrica.

Tabla A7. Individuos de las dos especies de topillos de los que se extrajo ADN genómico a partir de tejido fresco.

Código individuo	Tejido	Localidad
C1256	Músculo	Matarrosa del Sil
C1381	Músculo	Matarrosa del Sil
C1382	Músculo	Matarrosa del Sil
C1383	Músculo	Matarrosa del Sil
C2393	Cola	Belianes

Tabla A8. Conjunto de cebadores diseñados para la amplificación de las secuencias mitocondriales y de las secuencias nucleares de las dos especies de topillos. Las tres parejas de cebadores para la amplificación del citocromo *b* fueron diseñados específicamente para el género *Microtus*. Los cebadores de los intrones son los mismos que se desarrollaron en el primer capítulo para todo el orden *Rodentia*. Se probaron diferentes temperaturas de hibridación con tal de optimizar el proceso de amplificación.

Marcador	Secuencias cebadores (forward y reverse)	Temperaturas de hibridación
Primer fragmento citocromo <i>b</i> (1-340)	ACCAATGACATGAAAAATCATCGTT TACTATTYGCYGTMATRGC	50° C
Segundo fragmento citocromo <i>b</i> (323-769)	ACATARTMGAAACATGAAAYATA ACYCCTGCAAATCCACTCAMEYA	50° C
Tercer fragmento citocromo <i>b</i> (679-1143)	AAAGAYTYCTMGGRGTCCTT GTCTTGTAACCAGTAATGAAAA	50° C
Agxt-intrón-10	GGCTACAACCTGGAGGGACATC TGCAGGGCCTCCYTCAGGGCCT	57, 59, 61, 65-50° C
Catsper3-intrón-5	TGCTKGCMTCTTCATCTT AGRATYAYTGCTTCTYCTCC	65-50° C
Dhcr24-intrón-7	CAGGACATGCTGGTGCCCATGAA CCTGGCTGGCTGGGCAGGATGAA	57, 59, 61, 62, 65-50° C
Nadsyn1-intrón-4	GTYCGYTACAAYTGCAGAGT TCCTKSHCCAAGGGGTRAACCA	55, 57, 59, 65-50° C
Smo-intrón-9	GCCACCCTGCTCATCTGGAGGCG TTGGCRATCATCTTGCTYTTCTTGA	57, 59, 65-50° C
Wls-intrón-7	AAACACATYGCMGGSTAYTGGAA TCYGTKCCAACRTCYGTRGTCCA	57, 59, 65-50° C

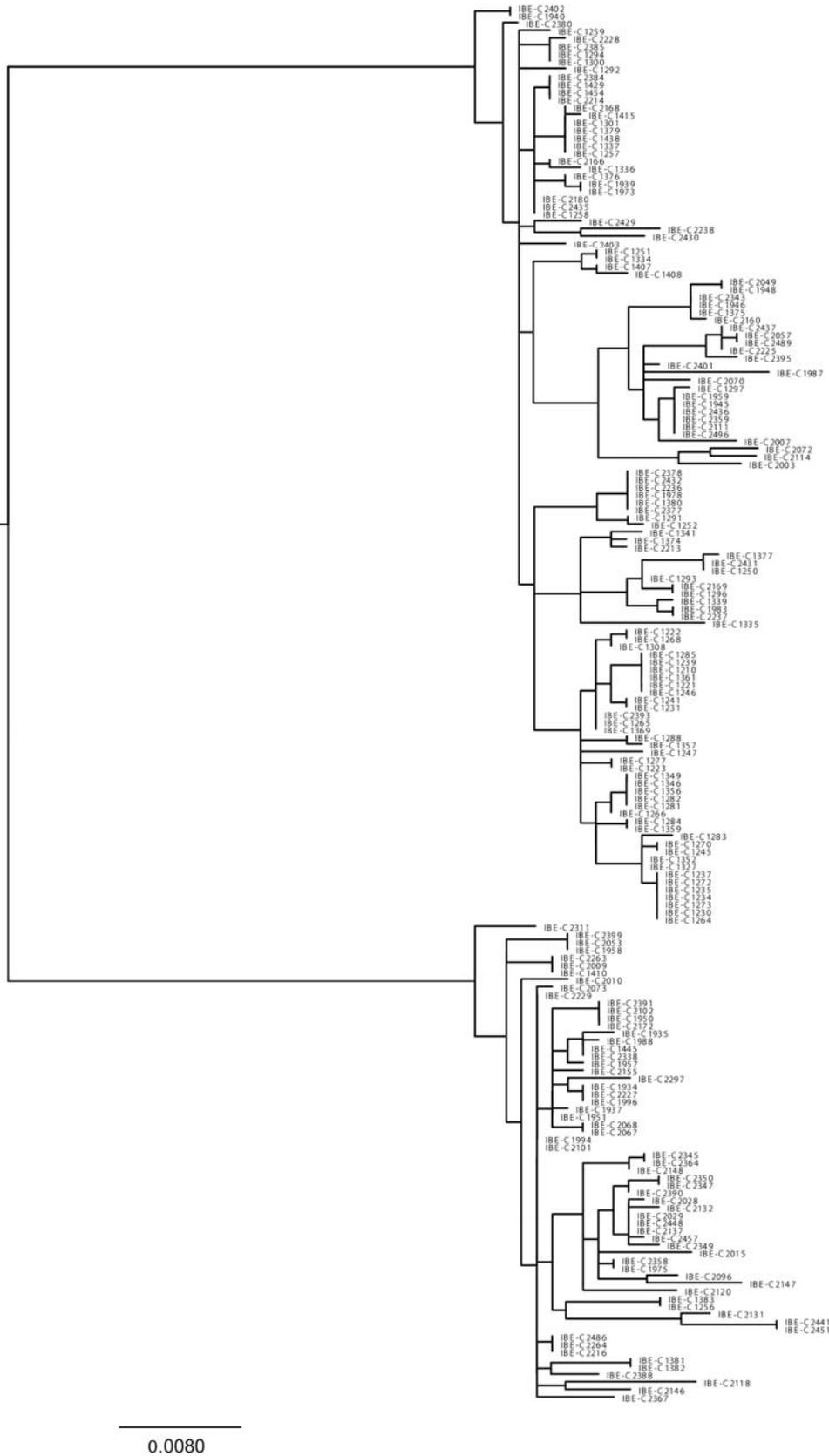


Figura A5. Árbol de máxima verosimilitud reconstruido con las 184 secuencias mitocondriales de las dos especies de topillos. La escala indica tasas de sustitución por posición.

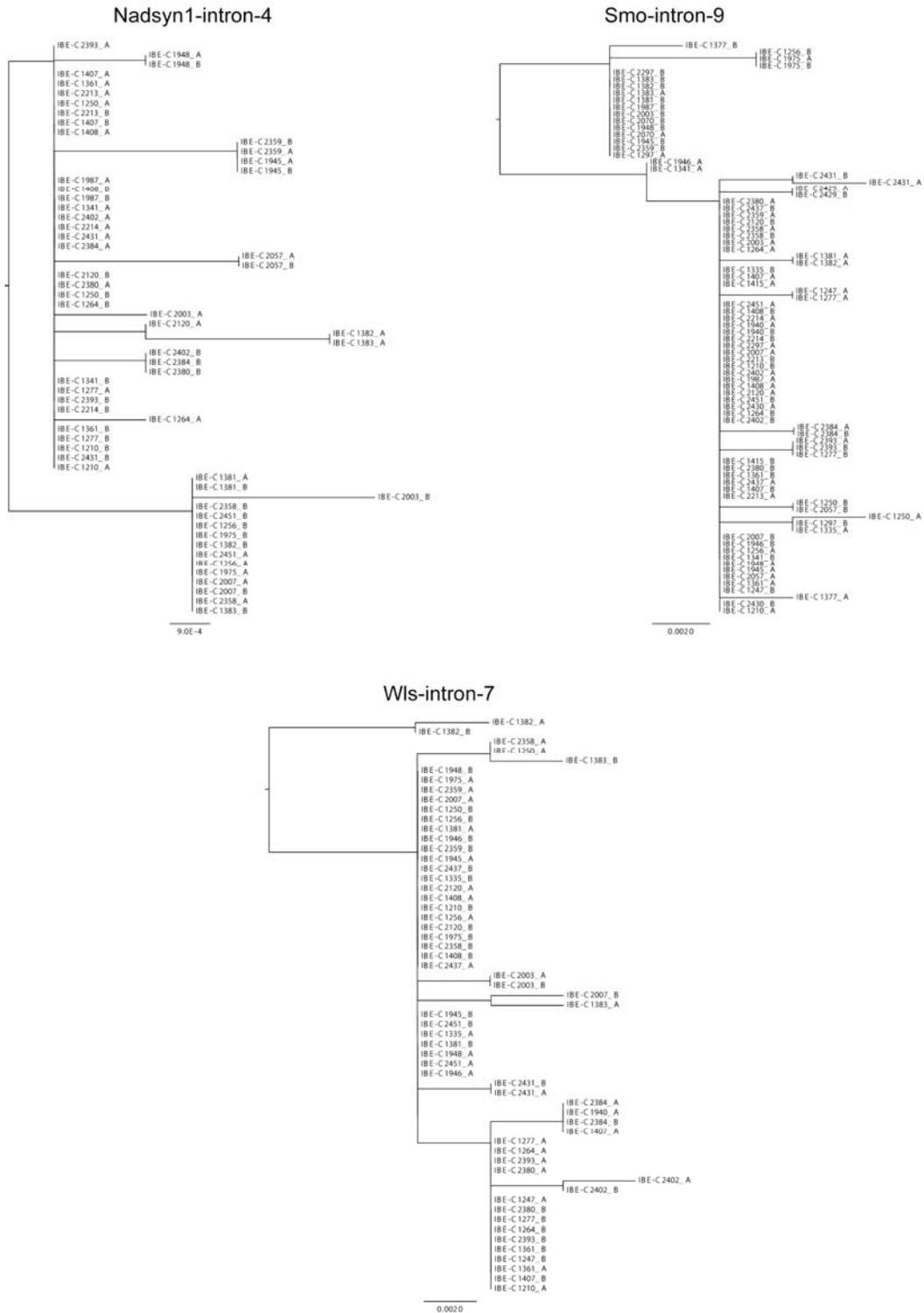


Figura A6. Árboles de máxima verosimilitud reconstruidos con las secuencias de los intrones amplificados en *M. duodecimcostatus* y *M. lusitanicus*. Las letras A y B indican las dos variantes alélicas de una misma secuencia para cada individuo. La escala indica tasas de sustitución por posición.

Tabla A9. Asignación de especie genética (mediante el citocromo *b*) y morfológica, de los individuos de las zonas de distribución alopátrica y de las zonas de distribución solapada. **(A)** Total de los 184 individuos recolectados. La línea horizontal divide a los individuos de las zonas de distribución alopátrica (parte superior) de los individuos de las zonas de distribución solapada (parte inferior). En cursiva, los individuos que presentan una incongruencia entre la especie genética y la morfológica. **(B)** Total de los 44 individuos para los que se realizó además una asignación de especie con la información contenida en las secuencias nucleares (intrones). En la parte superior de la primera línea horizontal se presentan los individuos de las zonas de distribución alopátrica, en la parte media de la tabla, los individuos de las zonas de distribución solapada y en la parte baja de la tabla los individuos de los que se extrajo ADN genómico a partir de tejido fresco. En cursiva los individuos pertenecientes al grupo DCMN (discordantes cito-morfonuclear). Localidades de muestreo: **Bel:** Belianes; **Ceb:** San Cebrián de Mazote; **Esc:** Escorial de la Sierra; **Hos:** Hospital de Incio; **Lum:** Lumbrales; **Men:** Menaza; **Mun:** Muñíos; **Oss:** Ossó de Sió; **Pal:** Palacios de Sil; **Pob:** Poblaciones de Arriba; **Por:** Porqueriza; **Sal:** Salvatierra de Tormes; **Toa:** Torrejón de Alba; **Tol:** Toledo y **Vil:** Villacantid.

A)

Código individuo	Localidad	Especie citocromo <i>b</i>	Especie morfología
C1210	Bel	Duo	Duo
C1221	Bel	Duo	Duo
C1222	Bel	Duo	Duo
C1223	Bel	Duo	Duo
C1230	Bel	Duo	Duo
C1231	Bel	Duo	Duo
C1234	Bel	Duo	Duo
C1235	Bel	Duo	Duo
C1237	Bel	Duo	Duo
C1239	Bel	Duo	Duo
C1241	Bel	Duo	Duo
C1245	Bel	Duo	Duo
C1246	Bel	Duo	Duo
C1247	Bel	Duo	Duo
C1264	Bel	Duo	Duo
C1265	Bel	Duo	Duo
C1266	Bel	Duo	Duo
C1268	Bel	Duo	Duo
C1270	Bel	Duo	Duo
C1272	Bel	Duo	Duo
C1273	Bel	Duo	Duo
C1277	Bel	Duo	Duo
C2028	Hos	Lus	Lus
C2029	Hos	Lus	Lus
C2096	Hos	Lus	Lus
C2131	Hos	Lus	Lus
C2132	Hos	Lus	Lus
C2137	Hos	Lus	Lus
C2146	Hos	Lus	Lus
C2147	Hos	Lus	Lus
C2148	Hos	Lus	Lus
C2345	Hos	Lus	Lus
C2347	Hos	Lus	Lus
C2349	Hos	Lus	Lus
C2350	Hos	Lus	Lus
C2364	Hos	Lus	Lus
C2390	Hos	Lus	Lus

C2441	Hos	Lus	Lus
C2448	Hos	Lus	Lus
C2451	Hos	Lus	Lus
C2457	Hos	Lus	Lus
C1281	Oss	Duo	Duo
C1282	Oss	Duo	Duo
C1283	Oss	Duo	Duo
C1284	Oss	Duo	Duo
C1285	Oss	Duo	Duo
C1288	Oss	Duo	Duo
C1308	Oss	Duo	Duo
C1327	Oss	Duo	Duo
C1346	Oss	Duo	Duo
C1349	Oss	Duo	Duo
C1352	Oss	Duo	Duo
C1356	Oss	Duo	Duo
C1357	Oss	Duo	Duo
C1359	Oss	Duo	Duo
C1361	Oss	Duo	Duo
C1369	Oss	Duo	Duo
C2238	Tol	Duo	Duo
C2403	Tol	Duo	Duo
C2429	Tol	Duo	Duo
C1407	Ceb	Duo	Duo
C1408	Ceb	Duo	Duo
<i>C1250</i>	<i>Esc</i>	<i>Duo</i>	<i>Lus</i>
<i>C1983</i>	<i>Esc</i>	<i>Duo</i>	<i>Lus</i>
<i>C2213</i>	<i>Esc</i>	<i>Duo</i>	<i>Lus</i>
<i>C2237</i>	<i>Esc</i>	<i>Duo</i>	<i>Lus</i>
<i>C1375</i>	<i>Lum</i>	<i>Duo</i>	<i>Lus</i>
<i>C1945</i>	<i>Lum</i>	<i>Duo</i>	<i>Lus</i>
<i>C1946</i>	<i>Lum</i>	<i>Duo</i>	<i>Lus</i>
C1948	Lum	Duo	Duo
<i>C1959</i>	<i>Lum</i>	<i>Duo</i>	<i>Lus</i>
<i>C1987</i>	<i>Lum</i>	<i>Duo</i>	<i>Lus</i>
<i>C2007</i>	<i>Lum</i>	<i>Duo</i>	<i>Lus</i>
<i>C2049</i>	<i>Lum</i>	<i>Duo</i>	<i>Lus</i>
<i>C2057</i>	<i>Lum</i>	<i>Duo</i>	<i>Lus</i>
<i>C2070</i>	<i>Lum</i>	<i>Duo</i>	<i>Lus</i>
<i>C2072</i>	<i>Lum</i>	<i>Duo</i>	<i>Lus</i>
<i>C2111</i>	<i>Lum</i>	<i>Duo</i>	<i>Lus</i>
<i>C2160</i>	<i>Lum</i>	<i>Duo</i>	<i>Lus</i>
<i>C2225</i>	<i>Lum</i>	<i>Duo</i>	<i>Lus</i>
<i>C2343</i>	<i>Lum</i>	<i>Duo</i>	<i>Lus</i>
<i>C2359</i>	<i>Lum</i>	<i>Duo</i>	<i>Lus</i>
<i>C2395</i>	<i>Lum</i>	<i>Duo</i>	<i>Lus</i>
<i>C2401</i>	<i>Lum</i>	<i>Duo</i>	<i>Lus</i>
<i>C2430</i>	<i>Lum</i>	<i>Duo</i>	<i>Lus</i>
<i>C2436</i>	<i>Lum</i>	<i>Duo</i>	<i>Lus</i>
<i>C2437</i>	<i>Lum</i>	<i>Duo</i>	<i>Lus</i>
<i>C2489</i>	<i>Lum</i>	<i>Duo</i>	<i>Lus</i>
<i>C2496</i>	<i>Lum</i>	<i>Duo</i>	<i>Lus</i>
C1937	Men	Lus	Lus
C1950	Men	Lus	Lus

C1951	Men	Lus	Lus
C1957	Men	Lus	Lus
C1958	Men	Lus	Lus
C1988	Men	Lus	Lus
C2009	Men	Lus	Lus
C2010	Men	Lus	Lus
C2053	Men	Lus	Lus
C2067	Men	Lus	Lus
C2068	Men	Lus	Lus
C2073	Men	Lus	Lus
C2101	Men	Lus	Lus
C2102	Men	Lus	Lus
C2118	Men	Lus	Lus
C2155	Men	Lus	Lus
C2172	Men	Lus	Lus
C2263	Men	Lus	Lus
C2311	Men	Lus	Lus
C2391	Men	Lus	Lus
C2399	Men	Lus	Lus
C2003	<i>Mun</i>	<i>Duo</i>	<i>Lus</i>
C2015	Mun	Lus	Lus
C2114	<i>Mun</i>	<i>Duo</i>	<i>Lus</i>
C2120	Mun	Lus	Lus
C1975	Pal	Lus	Lus
C2358	Pal	Lus	Lus
C1445	Pob	Lus	Lus
C1934	Pob	Lus	Lus
C1935	Pob	Lus	Lus
C1994	Pob	Lus	Lus
C1996	Pob	Lus	Lus
C2216	Pob	Lus	Lus
C2227	Pob	Lus	Lus
C2229	Pob	Lus	Lus
C2264	Pob	Lus	Lus
C2297	Pob	Lus	Lus
C2338	Pob	Lus	Lus
C2367	Pob	Lus	Lus
C2388	Pob	Lus	Lus
C2486	Pob	Lus	Lus
C1251	Por	Duo	Duo
C1258	Por	Duo	Duo
C1259	Por	Duo	Duo
C1292	Por	Duo	Duo
C1294	Por	Duo	Duo
C1300	Por	Duo	Duo
C1336	Por	Duo	Duo
C1380	Por	Duo	Duo
C1978	Por	Duo	Duo
C2166	Por	Duo	Duo
C2180	Por	Duo	Duo
C2236	Por	Duo	Duo
C2377	Por	Duo	Duo
C2378	Por	Duo	Duo
C2402	Por	Duo	Duo

C2432	Por	Duo	Duo
C2435	Por	Duo	Duo
C1257	Sal	Duo	Duo
C1293	Sal	Duo	Duo
C1296	<i>Sal</i>	<i>Duo</i>	<i>Duo/Lus</i>
C1301	Sal	Duo	Duo
C1335	<i>Sal</i>	<i>Duo</i>	<i>Lus</i>
C1337	Sal	Duo	Duo
C1339	<i>Sal</i>	<i>Duo</i>	<i>Lus</i>
C1341	<i>Sal</i>	<i>Duo</i>	<i>Lus</i>
C1374	<i>Sal</i>	<i>Duo</i>	<i>Lus</i>
C1377	<i>Sal</i>	<i>Duo</i>	<i>Lus</i>
C1379	Sal	Duo	Duo
C1415	Sal	Duo	Duo
C1429	Sal	Duo	Duo
C1438	Sal	Duo	Duo
C1454	Sal	Duo	Duo
C1939	Sal	Duo	Duo
C1940	Sal	Duo	Duo
C1973	Sal	Duo	Duo
C2168	Sal	Duo	Duo
C2169	<i>Sal</i>	<i>Duo</i>	<i>Lus</i>
C2214	Sal	Duo	Duo
C2228	Sal	Duo	Duo
C2380	Sal	Duo	Duo
C2384	Sal	Duo	Duo
C2385	Sal	Duo	Duo
C2431	<i>Sal</i>	<i>Duo</i>	<i>Lus</i>
C1252	<i>Toa</i>	<i>Duo</i>	<i>Duo/Lus</i>
C1291	<i>Toa</i>	<i>Duo</i>	<i>Lus</i>
C1297	<i>Toa</i>	<i>Duo</i>	<i>Lus</i>
C1334	Toa	Duo	Duo
C1376	Toa	Duo	Duo
C1410	Vil	Lus	Lus

B)

Código individuo	Localidad	Especie citocromo <i>b</i>	Especie morfología	Especie intrones	Grupo
C1210	Bel	Duo	Duo	Duo	CCMN
C1247	Bel	Duo	Duo	Duo	CCMN
C1264	Bel	Duo	Duo	Duo	CCMN
C1277	Bel	Duo	Duo	Duo	CCMN
C1361	Oss	Duo	Duo	Duo	CCMN
C2429	Tol	Duo	Duo	Duo	CCMN
C2451	Hos	Lus	Lus	Lus	CCMN
C1407	Ceb	Duo	Duo	Duo	CCMN
C1408	Ceb	Duo	Duo	Duo	CCMN
C2402	Por	Duo	Duo	Duo	CCMN
C1415	Sal	Duo	Duo	Duo	CCMN
C1940	Sal	Duo	Duo	Duo	CCMN
C2214	Sal	Duo	Duo	Duo	CCMN
C2380	Sal	Duo	Duo	Duo	CCMN
C2384	Sal	Duo	Duo	Duo	CCMN
C2120	Mun	Lus	Lus	Lus	CCMN
C1975	Pal	Lus	Lus	Lus	CCMN
C2358	Pal	Lus	Lus	Lus	CCMN
C2297	Pob	Lus	Lus	Lus	CCMN
C2486	Pob	Lus	Lus	Lus	CCMN
C1250	<i>Esc</i>	<i>Duo</i>	<i>Lus</i>	<i>Lus</i>	<i>DCMN</i>
C2213	<i>Esc</i>	<i>Duo</i>	<i>Lus</i>	<i>Lus</i>	<i>DCMN</i>
C1945	<i>Lum</i>	<i>Duo</i>	<i>Lus</i>	<i>Lus</i>	<i>DCMN</i>
C1946	<i>Lum</i>	<i>Duo</i>	<i>Lus</i>	<i>Lus</i>	<i>DCMN</i>
C1987	<i>Lum</i>	<i>Duo</i>	<i>Lus</i>	<i>Lus</i>	<i>DCMN</i>
C2007	<i>Lum</i>	<i>Duo</i>	<i>Lus</i>	<i>Lus</i>	<i>DCMN</i>
C2057	<i>Lum</i>	<i>Duo</i>	<i>Lus</i>	<i>Lus</i>	<i>DCMN</i>
C2070	<i>Lum</i>	<i>Duo</i>	<i>Lus</i>	<i>Lus</i>	<i>DCMN</i>
C2359	<i>Lum</i>	<i>Duo</i>	<i>Lus</i>	<i>Lus</i>	<i>DCMN</i>
C2430	<i>Lum</i>	<i>Duo</i>	<i>Lus</i>	<i>Lus</i>	<i>DCMN</i>
C2437	<i>Lum</i>	<i>Duo</i>	<i>Lus</i>	<i>Lus</i>	<i>DCMN</i>
C2003	<i>Mun</i>	<i>Duo</i>	<i>Lus</i>	<i>Lus</i>	<i>DCMN</i>
C1335	<i>Sal</i>	<i>Duo</i>	<i>Lus</i>	<i>Lus</i>	<i>DCMN</i>
C1341	<i>Sal</i>	<i>Duo</i>	<i>Lus</i>	<i>Lus</i>	<i>DCMN</i>
C1377	<i>Sal</i>	<i>Duo</i>	<i>Lus</i>	<i>Lus</i>	<i>DCMN</i>
C2431	<i>Sal</i>	<i>Duo</i>	<i>Lus</i>	<i>Lus</i>	<i>DCMN</i>
C1297	<i>Toa</i>	<i>Duo</i>	<i>Lus</i>	<i>Lus</i>	<i>DCMN</i>
C1293	<i>Sal</i>	<i>Duo</i>	<i>Duo</i>	<i>Lus</i>	<i>DCMN</i>
C1948	<i>Lum</i>	<i>Duo</i>	<i>Duo</i>	<i>Lus</i>	<i>DCMN</i>
C2393	Bel	Duo	Tejido	Duo	CCMN
C1256	Mat	Lus	Tejido	Lus	CCMN
C1381	Mat	Lus	Tejido	Lus	CCMN
C1382	Mat	Lus	Tejido	Lus	CCMN
C1383	Mat	Lus	Tejido	Lus	CCMN

Tabla A10. Modelos de sustitución nucleotídica y relojes moleculares calculados para cada uno de los seis intrones y para el citocromo *b* con la información de las secuencias amplificadas para las dos especies de topillos.

Nombre intrón	Modelo de sustitución nucleotídica	Reloj molecular
Agxt-intrón-10	HKY+G	Estricto
Catsper3-intrón-5	HKY+I	Estricto
Dhcr24-intrón-7	HKY+I	Estricto
Nadsyn1-intrón-4	HKY+G	Estricto
Smo-intrón-9	HKY+G	Estricto
Wls-intrón-7	HKY+I	Estricto
Citocromo <i>b</i>	GTR+I+G	Relajado

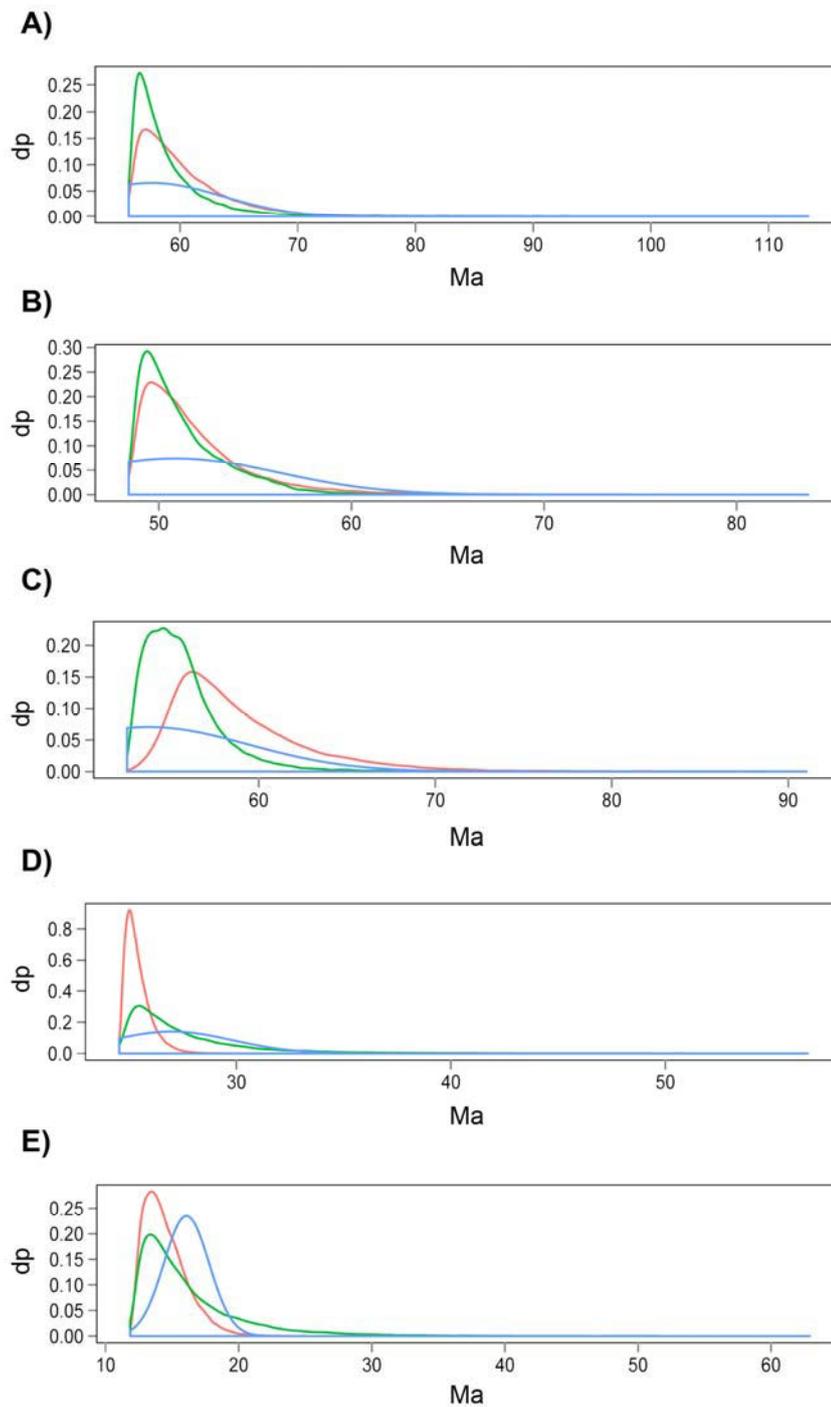


Figura A7. Distribuciones de las edades calculadas para cada nodo calibrado. En azul, distribuciones de los *priors* introducidos en el análisis; en rojo las distribuciones posteriores del análisis y en verde las distribuciones efectivas conjuntas de los *priors* del análisis. **(A)** Nodo clado *Rodentia*. **(B)** Nodo clado *Sciuridae*. **(C)** Nodo clado *Hystricognathi*. **(D)** Nodo clado *Octodontoidea*. **(E)** Nodo clado *Hydrocaeridae-Caviidae*. Escala temporal en millones de años.

VI. Apéndice

Tabla A11. Combinación de individuos y marcadores incluidos para cada uno de los análisis realizados con *Beast e IMA2 (IM e I) para el cálculo de diferentes parámetros. (CCMN, individuos concordantes cito-morfonuclear; DCMN, individuos discordantes cito-morfonuclear). (a) Análisis en los que se implementó un modelo de coalescencia lineal y raíz constante en *Beast.

Número análisis	*Beast		IM (IMa2)		I (IMa2)		Parámetro estimado
	Intrones	Cit. <i>b</i>	Intrones	Cit. <i>b</i>	Intrones	Cit. <i>b</i>	
1	CCMN DCMN	CCMN	CCMN DCMN	CCMN	--	--	Tasas de mutación ^(a) Tamaño efectivo poblacional ^(a) Tiempo de especiación ^(a) TMRCA ^(a)
2	--	CCMN	--	--	--	CCMN	Tiempo de especiación
3	--	CCMN DCMN	--	--	--	CCMN DCMN	Tiempo de especiación
4	CCMN DCMN	--	CCMN DCMN	--	CCMN DCMN	--	Tiempo de especiación
5	CCMN DCMN	CCMN	CCMN DCMN	CCMN	CCMN DCMN	CCMN	Tiempo de especiación
6	CCMN DCMN	CCMN DCMN	CCMN DCMN	CCMN DCMN	CCMN DCMN	CCMN DCMN	Tiempo de especiación

Tabla A12. *Priors* utilizados para cada uno de los análisis realizados bajo un modelo de aislamiento con migración (IM) o sin migración (I). (CCMN, individuos concordantes cito-morfonuclear; DCMN, individuos discordantes cito-morfonuclear).

Modelo	Intrones	Cit. <i>b</i>	<i>Priors</i>
IM (IMa2)	CCMN DCMN	--	-q10 -t5 -m2 -hfg -hn15 -ha0,96 -hb0,9 -p23567 -y1,0 -b500000 -l100000
	CCMN DCMN	CCMN	-q20 -t5 -m2 -hfg -hn15 -ha0,96 -hb0,9 -p23567 -y1,0 -b500000 -l100000
	CCMN DCMN	CCMN DCMN	-q20 -t5 -m2 -hfg -hn15 -ha0,96 -hb0,9 -p23567 -y1,0 -b500000 -l100000
I (IMa2)	--	CCMN	-q900 -t30 -m0 -hfg -hn15 -ha0,96 -hb0,9 -p23567 -y1,0 -b200000 -l50000
	--	CCMN DCMN	-q400 -t50 -m0 -hfg -hn15 -ha0,96 -hb0,9 -p23567 -y1,0 -b200000 -l50000
	CCMN DCMN	--	-q10 -t5 -m0 -hfg -hn15 -ha0,96 -hb0,9 -p23567 -y1,0 -b500000 -l100000
	CCMN DCMN	CCMN	-q30 -t5 -m0 -hfg -hn15 -ha0,96 -hb0,9 -p23567 -y1,0 -b500000 -l100000
	CCMN DCMN	CCMN DCMN	-q20 -t5 -m0 -hfg -hn15 -ha0,96 -hb0,9 -p23567 -y1,0 -b500000 -l100000

VII. BIBLIOGRAFÍA

- Ácz R, Ubányi G, Ozár V (2005) Morphometric differences among root vole (*Muridae: Microtus oeconomus*) populations in Hungary. *Acta Zoologica Academiae Scientiarum Hungaricae*, **51**, 135–149.
- Adams DC (2004) Character displacement via aggressive interference in appalachian salamanders. *Ecology*, **85**, 2664–2670.
- Adams DC, Funk DJ (1997) Morphometric inferences on sibling species and sexual dimorphism in *neochlamisus bebbianae* leaf beetles: multivariate applications of the thin-plate spline. *Systematic Biology*, **46**, 180–194.
- Adams DC, Nistri A (2010) Ontogenetic convergence and evolution of foot morphology in European cave salamanders (Family: *Plethodontidae*). *BMC Evolutionary Biology*, **10**, 216.
- Adams DC, Otarola-Castillo E (2013) geomorph: an R package for the collection and analysis of geometric morphometric shape data. *Methods in Ecology and Evolution*, **4**, In Press.
- Adams DC, Rohlf FJ, Slice DE (2004) Geometric morphometrics: Ten years of progress following the “revolution”. *Italian Journal of Zoology*, **71**, 5–16.
- Albertson RC, Markert JA, Danley PD, Kocher TD (1999) Phylogeny of a rapidly evolving clade: the cichlid fishes of Lake Malawi, East Africa. *Proc. Natl. Acad. Sci. U S A.*, **96**, 5107–5110.
- Almança C (1973) La distinction entre *Pitymys lusitanicus* (Gerbe) et *P. duodecimcostatus* (de Sélys-Longchamps) d'après quelques indices craniométriques. *Rev. Biol.*, **9**, 207-216.
- Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research*, **25**, 3389-3402.
- Andersson L (1990) The driving force: Species concepts and ecology. *Taxon*, **39**, 375–382.
- Auguie B (2012) gridExtra: functions in Grid graphics. R Packaged version 0.9.
- Aulagnier S, Amori G, Hutterer R, Kryštufek B, Yigit N, Mitsain G, Palomo LJ (2008) *Microtus lusitanicus*. In: *IUCN 2012. IUCN Red List of Threatened Species. Version 2012.2.*
- Aulagnier S, Palomo LJ (2008) *Microtus duodecimcostatus*. In: *IUCN 2012. IUCN Red List of Threatened Species. Version 2012.2.*

- Avice JC (2000) *Phylogeography: The history and formation of species*. Harvard Univ. Press, Cambridge, MA.
- Avice JC, Ball RM Jr (1990) Principles of genealogical concordance in species concepts and biological taxonomy. *Oxf. Surv. Evol. Biol.*, **7**, 45–67.
- Banfield J, Raftery AE (1993) Model-based Gaussian and non-Gaussian clustering. *Biometrics*, **49**, 803-821.
- Barracough TG, Nee S (2001) Phylogenetics and speciation. *Trends in Ecology & Evolution*, **16**, 391-399.
- Bastos-Silveira C, Santos SM, Monarca R, Mathias MDL, Heckel G (2012) Deep mitochondrial introgression and hybridization among ecologically divergent vole species. *Molecular Ecology*, **21**, 5309–23.
- Baum DA, Shaw KL (1995) Genealogical perspectives on the species problem. In: *Experimental and molecular approaches to plant biosystematics* (eds. Hoch PC, Stephenson AG), pp. 289-303. Missouri Botanical Garden, St. Louis.
- Benton MJ, Donoghue PCJ, Asher RJ (2009) Calibrating and constraining molecular clocks. In: *The Timetree of Life* (eds. Hedges SB, Kumar S), pp. 35-86. Oxford University Press, Oxford, New York.
- Berlocher SH, Feder JL (2002) Sympatric speciation in phytophagous insects: moving beyond controversy? *Annu. Rev. Entomol.*, **47**, 773–815.
- Blanco JC (1998). *Mamíferos de España*. Planeta. España.
- Blanga-Kanfi S, Miranda H, Penn O, Pupko T, DeBry RW, Huchon D (2009) Rodent phylogeny revised: analysis of six nuclear genes from all major rodent clades. *BMC Evolutionary Biology*, **9**, 71.
- Bookstein FL (1986) Size and shape spaces for landmark data in two dimensions. *Statistical Science*, **1**, 181–242.
- Bookstein FL (1991) *Morphometric tools for landmark data: geometry and biology*. Cambridge University Press, Cambridge.
- Bookstein FL (1996). A standard formula for the uniform shape component in landmark data. In: *Advances in Morphometrics* (eds. Marcus LF, Corti M, Loy A et al.), pp. 153–168. Plenum Press, New York.
- Bookstein FL (1998) A hundred years of morphometrics. *Acta Zool. Acad. Sci. Hung.*, **44**, 7-59.

- Borghi CE, Giannoni SM, Martínez-Rica JP (1994) Habitat segregation of three sympatric fossorial rodents in the Spanish Pyrenees. *Z Saugetierkunde*, **59**, 52–57.
- Bradley RD, Baker RJ (2001) A test of the genetic species concept: cytochrome *b* sequences and mammals. *J. Mammal.*, **82**, 960–973.
- Brito PH, Edwards SV (2009) Multilocus phylogeography and phylogenetics using sequencebased markers. *Genetica*, **135**, 439–455.
- Brown WM, George JrM, Wilson AC (1979) Rapid evolution of animal mitochondrial DNA. *Proc. Natl. Acad. Sci. USA*, **76**, 1967–1971.
- Bruen TC, Philippe H, Bryant D (2006) A simple and robust statistical test for detecting the presence of recombination. *Genetics*, **172**, 2665–81.
- Brunet-Lecomte P, Brochet G, Chaline J, Delibes M (1987) Morphologie dentaire comparée de *Pitymys lusitanicus* et *Pitymys duodecimcostatus* (*Arvicolinae*, *Rodentia*) dans le Nord-Ouest de l'Espagne. *Mammalia*, **51**, 145–158.
- Brunet-Lecomte P, Chaline J (1990) Relations phylogénétiques et évolution des campagnols souterrains d'Europe (*Terricola*, *Arvicolinae*, *Rodentia*). *C. R. Acad. Sci. Paris*, **311(II)**, 745–750.
- Brunet-Lecomte P, Chaline J (1991) Morphological evolution and phylogenetic relationships of the European ground voles (*Arvicolinae*, *Rodentia*). *Lethaia*, **24**, 45–53.
- Brunet-Lecomte P, Chaline, J (1992) Morphological convergences versus biochemical divergences in the Holarctic ground voles: *Terricola* and *Pitymys* (*Arvicolinae*, *Rodentia*). *N. Jb. Geol. Paläont.*, **12**, 721–734.
- Bryant D, Bouckaert R, Felsenstein J, Rosenberg N A, RoyChoudhury A (2012) Inferring species trees directly from biallelic genetic markers: bypassing gene trees in a full coalescent analysis. *Molecular Biology and Evolution*, **29**, 1917–32.
- Buckley TR, Cordeiro M, Marshall DC, Simon C (2006) Differentiating between hypotheses of lineage sorting and introgression in New Zealand alpine cicadas (*Maoricicada Dugdale*). *Systematic Biology*, **55**, 411–425.
- Bull M (1991) Ecology of parapatric distributons. *Ann. Rev. Ecol. Syst.*, **22**, 19–36.
- Bush GL (1969) Sympatric host race formation and speciation in frugivorous flies of the genus *Rhagoletis* (*Diptera*, *Tephriidae*). *Evolution*, **23**, 237–251.

- Bush GL (1994) Sympatric speciation in animals: new wine in old bottles. *Trends in Ecology & Evolution*, **9**, 285–8.
- Casas-Marce M, Revilla E, Godoy J (2009) Searching for DNA in museum specimens: a comparison of sources in a mammal species. *Molecular Ecology Resources*, **10**, 502–507.
- Castresana J (2000) Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Molecular Biology and Evolution*, **17**, 540–552.
- Caumul R, Polly PD (2005) Phylogenetic and environmental components of morphological variation: skull, mandible, and molar shape in marmots (*Marmota*, *Rodentia*). *Evolution*, **59**, 2460–72.
- Chaline J (1966) Un exemple d'évolution chez les Arvicolidés (*Rodentia*): Les lignées *Allophaiomys*, *Pitymys* et *Microtus*. *C. R. Acad. Sci., Paris D*, **263**, 1022–1024.
- Chaline J (1972) Les rongeurs du Pléistocène moyen et supérieur de France. Cahiers de Paléontologie du CNRS, Paris.
- Chaline J (1974) Esquisse de l'évolution morphologique, biométrique et chromosomique du genre *Microtus* (*Arvicolinae*, *Rodentia*) dans le Pléistocène de l'hémisphère nord. *Bull. Soc. Géol. France*, **16**, 440–450.
- Chaline J (1985) Evolutionary data on steppe Lemmings (*Arvicolinae*, *Rodentia*). In: *Multidisciplinary analysis of evolutionary relationships among rodents* (eds. Lockett P, Harten-berger JL), pp. 631–341. Plenum Pub. Corp., New York.
- Chaline J, Brunet-lecomte P, Montuire S, Viriot L, Courant F (1999) Anatomy of the arvicoline radiation (*Rodentia*): palaeogeographical, palaeoecological history and evolutionary data. *Ann. Zool. Fennici*, **36**, 239–267.
- Chaline J, Graf J (1988) Phylogeny of the *Arvicolidae* (*Rodentia*): Biochemical and Paleontological. *American Society of Mammalogists*, **69**, 22–33.
- Chesser RT, Zink RM (1994) Modes of speciation in birds – a test of Lynch's method. *Evolution*, **48**, 490–497.
- Churakov G, Sadasivuni MK, Rosenbloom KR, Huchon D, Brosius J, Schmitz J (2010) Rodent evolution: back to the root. *Molecular Biology and Evolution*, **27**, 1315–1326.
- Collyer ML, Adams DC (2007) Analysis of two-state multivariate phenotypic change in ecological studies. *Ecology*, **88**, 683–92.

- Cotilla I, Palomo LJ (2007) *Microtus duodecimcostatus* (de Selys-Longchamps, 1839). In: *Atlas y libro rojo de los mamíferos terrestres de España* (eds. Palomo LJ, Gisbert J, Blanco JC), pp. 422–425. Madrid: Dirección General para la Biodiversidad - SECEM – SECEMU.
- Coyne JA, Orr HA (2004) *Speciation*. Sinauer Associates, Sunderland, MA.
- Creer S, Malhotra A, Thorpe R, Pook C (2005) Targeting optimal introns for phylogenetic analyses in non-model taxa: experimental results in Asian pitvipers. *Cladistics*, **21**, 390–395.
- Cummings MP, Neel MC, Shaw KL (2008) A genealogical approach to quantifying lineage divergence. *Evolution; International Journal of Organic Evolution*, **62**, 2411–22.
- Curat M, Excoffier L (2004) Modern humans did not admix with Neanderthals during their range expansion into Europe. *PLoS Biol*, **2**, 2264–2274.
- Danley PD, Kocher TD (2001) Speciation in rapidly diverging systems: lessons from Lake Malawi. *Mol. Ecol.*, **10**, 1075–1086.
- Dayrat B (2005) Towards integrative taxonomy. *Biol. J. Linn. Soc.*, **85**, 407–415.
- Degnan JH, Rosenberg NA (2009) Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends in Ecology and Evolution*, **24**, 332–340.
- De Queiroz K (1999) The general lineage concept of species and the defining properties of the species category. In: *Species: New interdisciplinary essays* (eds. Wilson RA), pp.49–89. MIT Press, Cambridge.
- De Queiroz K (2007) Species concepts and species delimitation. *Systematic biology*, **56**, 879–86.
- Dobzhansky T (1950) Mendelian populations and their evolution. *Am. Nat*, **84**, 401–418.
- Donoghue MJ (1985) A critique of the biological species concept and recommendations for a phylogenetic alternative. *Bryologist*, **88**, 172–181.
- Drummond AJ AB, Cheung M, Heled J, Kearse M, Moir R, Stones- Havas S, Thierer T, Wilson A (2009) Geneious v4.
- Drummond AJ, Ho SYW, Phillips MJ, Rambaut A (2006) Relaxed phylogenetics and dating with confidence. *PLoS Biology*, **4**, e88.
- Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, **7**, 214.

- Dryden IL, Mardia KV (1998) Statistical shape analysis. John Wiley & Sons, New York.
- Edwards SV (2009) Is a new and general theory of molecular systematics emerging? *Evolution*, **63**, 1-19.
- Ellerman JR, Morrison-Scott TS (1951) Checklist of Palearctic and Indian Mammals. British Museum, London, 810 pp.
- Fabre P-H, Hautier L, Dimitrov D, Douzery EJ (2012) A glimpse on the pattern of rodent diversification: a phylogenetic approach. *BMC Evolutionary Biology*, **12**, 88.
- Feder JL (1998) The apple maggot fly, *Rhagoletis pomonella* flies in the face of conventional wisdom about speciation? In: *Endless Forms: Species and Speciation* (eds. Howard DJ, Berlocher SH), pp. 130–144. Oxford University Press, New York.
- Felsenstein J (1981) Evolutionary trees from DNA sequences: a maximum likelihood approach. *Journal of Molecular Evolution*, **17**, 368-376.
- Fink S, Fischer MC, Excoffier L, Heckel G (2010) Genomic scans support repetitive continental colonization events during the rapid radiation of voles (*Rodentia: Microtus*): the utility of AFLPs versus mitochondrial and nuclear sequence markers. *Systematic Biology*, **59**, 548–72.
- Fisher RA (1935) The logic of inductive inference. *J. R. Stat. Soc.*, **98**, 39-82.
- Fitzpatrick BM, Fordyce JA, Gavrilets S (2008) What, if anything, is sympatric speciation? *Journal of Evolutionary Biology*, **21**, 1452–9.
- Flicek P, Amode MR, Barrell D, *et al.* (2012) Ensembl 2012. *Nucleic Acids Research*, **40**, D84-90.
- Flot J-F, Tillier A, Samadi S, Tillier S (2006) Phase determination from direct sequencing of length-variable DNA regions. *Molecular Ecology Notes*, **6**, 627–630.
- Fraley C, Raftery AE (2002) Model-based Clustering, Discriminant Analysis and Density Estimation. *Journal of the American Statistical Association*, **97**, 611-631.
- Fraley C, Raftery AC, Murphy TB, Scrucca L (2012) mclust Version 4 for R: Normal Mixture Modeling for Model-Based Clustering, Classification, and Density Estimation Technical Report No. 597, Department of Statistics, University of Washington.

- Fu YX, Li WH (1993) Statistical tests of neutrality of mutations. *Genetics*, **133**, 693–709.
- Fujita MK, Leaché AD, Burbrink FT, McGuire JA, Moritz C (2012) Coalescent-based species delimitation in an integrative taxonomy. *Trends in Ecology & Evolution*, **27**, 480–8.
- Futuyma DJ, Mayer GC (1980) Non-allopatric speciation in animals. *Syst. Zool.*, **29**, 254–271.
- Gaubert P, Taylor PJ, Fernandes CA, Bruford MW, Veron G (2005) Patterns of cryptic hybridization revealed using an integrative approach: a case study on genets (*Carnivora, Viverridae, Genetta spp.*) from the southern African subregion. *Biological Journal of the Linnean Society*, **86**, 11–33.
- Gavrilets S (2003) Models of speciation: what have we learned in 40 years? *Evolution*, **57**, 2197–2215.
- Gavrilets S (2004) *Fitness Landscapes and the Origin of Species*. Princeton University Press, Princeton, NJ.
- Geraldes A *et al.* (2008) Reduced introgression of the Y chromosome between subspecies of the European rabbit (*Oryctolagus cuniculus*) in the Iberian Peninsula. *Mol. Ecol.*, **17**, 4489–4499.
- Gómez A, Lunt DH (2006) Refugia within refugia: patterns of phylogeographic concordance in the Iberian Peninsula. In: *Phylogeography of Southern European Refugia* (eds. Weiss S, Ferrand, N), Chapter 3, pp. 155–188. Kluwer Academic Publishers, Amsterdam.
- Good JM *et al.* (2008) Ancient hybridization and mitochondrial capture between two species of chipmunks. *Mol. Ecol.*, **17**, 1313–1327.
- Guédon G, Paradis E, Croset H (1992) Capture-recapture study of a population of the Mediterranean pine vole (*Microtus duodecimcostatus*) in Southern France. *Zeitschrift für Säugetierkunde*, **57**, 364–72.
- Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* **52**, 696–704.
- Hahne J, Jenkins T, Halle S, Heckel G (2011) Establishment success and resulting fitness consequences for vole dispersers. *Oikos*, **120**, 95–105.
- Hailer F, Kutschera VE, Hallström BM, Klassert D, Fain SR, Leonard JA, Arnason U, Janke A (2012) Nuclear genomic sequences reveal that polar bears are an old and distinct bear lineage. *Science*, **336**, 344–347.

- Hanski I, Gaggiotti OE (2004) Metapopulation biology: Past, present, and future. In: *Ecology, genetics, and evolution of metapopulations* (eds. Hanski I, Gaggiotti OE), pp. 3–22. Elsevier, Amsterdam.
- Hasegawa M, Kishino H, Yano T (1987). Man's place in *Hominoidea* as inferred from molecular clocks of DNA. *J. Mol. Evol.*, **26**, 132–147.
- Heled J, Drummond AJ (2010) Bayesian inference of species trees from multilocus data. *Molecular Biology and Evolution* **27**, 570-580.
- Hennig W (1966) Phylogenetic systematics. University of Illinois Press, Urbana.
- Herrera CM (1973) Biometría y Sistemática de *Pitymys duodecimcostatus* de Sélys-Longchamps, 1839 (*Mammalia, Rodentia*) en el sur de España. *Arq Mus Bocage, Série II*, 261-276.
- Hey J (2010) Isolation with migration models for more than two populations. *Molecular Biology and Evolution* **27**, 905-920.
- Holder MT, Anderson JA, Holloway AK (2001) Difficulties in detecting hybridization. *Systematic Biology*, **50**, 978–982.
- Holland BR, Benthin S, Lockhart PJ, Moulton V, Huber KT (2008) Using supernetworks to distinguish hybridization from incomplete lineage sorting. *BMC Evolutionary Biology*, **8**, 202.
- Honeycutt RL (2009) Rodents (*Rodentia*). In: *The Timetree of Life* (eds. Hedges SB, Kumar S), pp. 490-494. Oxford University Press, Oxford, New York.
- Hotelling H, (1933) Analysis of a complex of statistical variables into principal components. *J. Educ. Psychol.*, **24**, 417-441, 498-520.
- Huson DH, Klöpper T, Lockhart PJ, Steel MA (2005) Reconstruction of reticulate networks from gene trees. In: *Proceedings of the Ninth International Conference on Research in Computational Molecular Biology*, pp. 233–249. Springer, Heidelberg.
- Hudson RR, Coyne JA (2002) Mathematical consequences of the genealogical species concept. *Evolution*, **56**, 1557–1565.
- Hudson RR, Turelli M (2003) Stochasticity overrules the three-times rule: genetic drift, genetic draft, and coalescence times for nuclear loci versus mitochondrial DNA. *Evolution*, **57**, 182–190.
- Hull DL (1980) Individuality and selection. *Annu. Rev. Ecol. Syst.*, **11**, 311–332.
- Igea J, Juste J, Castresana J (2010) Novel intron markers to study the phylogeny of closely related mammalian species. *BMC Evolutionary Biology*, **10**, 369.

- Igea J, Aymerich P, Fernández-González A *et al.* (2013) Phylogeography and postglacial expansion of the endangered semi-aquatic mammal *Galemys pyrenaicus*. *BMC Evolutionary Biology*, **13**, 115.
- International Human Genome Sequencing Consortium (2001) Initial sequencing and analysis of the human genome. *Nature*, **409**, 860-921.
- Ishida Y, Oleksyk TK, Georgiadis NJ *et al.* (2011) Reconciling apparent conflicts between mitochondrial and nuclear phylogenies in African elephants. *PloS One*, **6**, e20642.
- Jaarola M, Martinková N, Gündüz I *et al.* (2004) Molecular phylogeny of the speciose vole genus *Microtus* (*Arvicolinae*, *Rodentia*) inferred from mitochondrial DNA sequences. *Molecular Phylogenetics and Evolution*, **33**, 647–63.
- Jaarola M, Searle JB (2004) A highly divergent mitochondrial DNA lineage of *Microtus agrestis* in southern Europe. *Heredity*, **92**, 228–34.
- Jacobsen F, Omland KE (2011) Species tree inference in a recent radiation of orioles (Genus *Icterus*): multiple markers and methods reveal cytonuclear discordance in the northern oriole group. *Molecular Phylogenetics and Evolution*, **61**, 460–9.
- Jacobsen F, Omland KE (2012) Extensive introgressive hybridization within the northern oriole group (Genus *Icterus*) revealed by three-species isolation with migration analysis. *Ecology and Evolution*, **2**, 2413-2429.
- Joly S, McLenachan PA, Lockhart PJ (2009) A statistical approach for distinguishing hybridization and incomplete lineage sorting. *The American Naturalist*, **174**, E54-70.
- Kaliontzopoulou A, Adams DC, Meijden A, Perera A, Carretero MA. (2011) Relationships between head morphology, bite performance and ecology in two species of *Podarcis* wall lizards. *Evolutionary Ecology*, **26**, 825–845.
- Kass RE, Raftery AE (1995) Bayes factor. *Journal of the American Statistical Association*, **90**, 773-795.
- Katoh K, Toh H (2008) Recent developments in the MAFFT multiple sequence alignment program. *Briefings in Bioinformatics* **9**, 286-298.
- Kingman JFC (1982) The coalescent. *Stochastic Processes and their Applications*, **13**, 235–248.
- Knowles LL, Kubatko LS (2010) Estimating species trees. Practical and theoretical aspects. Wiley-Blackwell, Hoboken, N.J.

- Korbie DJ, Mattick JS (2008) Touchdown PCR for increased specificity and sensitivity in PCR amplification. *Nature Protocols*, **3**, 1452-1456.
- Kretzoi M (1969) Skizze einer Arvicoliden-Phylogenie. *Vertebrata Hungarica*, **11**, 155–193.
- Kuhner MK, Yamato J, Felsenstein J. (1998) Maximum likelihood estimation of population growth rates based on the coalescent. *Genetics*, **149**, 429–434.
- Kumar S (2005) Molecular clocks: four decades of evolution. *Nature reviews. Genetics*, **6**, 654–62.
- Kurten B (2007) Pleistocene mammals of Europe. Weidenfeld & Nicolson. London.
- Ladoukakis ED, Zouros E (2001) Recombination in animal mitochondrial DNA: evidence from published sequences. *Mol Biol Evol*, **18**, 2127–2131.
- Leaché AD et al. (2009) Quantifying ecological, morphological, and genetic variation to delimit species in the coast horned lizard species complex (*Phrynosoma*). *Proc. Natl. Acad. Sci. U.S.A.*, **106**, 12418–12423.
- Le Galliard J-F, Remy A, Ims RA, Lambin X (2012) Patterns and processes of dispersal behaviour in arvicoline rodents. *Molecular Ecology*, **21**, 505–523.
- Lemon J (2006) Plotrix: a package in the red light district of R. *R-News*, **6(4)**, 8-12.
- Levins R (1970) Extinction. In: *Some mathematical problems in biology* (eds. Desternhaber M), pp. 77-107. American Mathematical Society, Providence, Rhode Island.
- Li C, Riethoven JJM, Ma L (2010) Exon-primed intron-crossing (EPIC) markers for nonmodel teleost fishes. *BMC Evolutionary Biology*, **10**, 90.
- Librado P, Rozas J (2009) DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, **25(11)**, 1451-1452.
- Linnen CR, Farrell BD (2007) Mitonuclear discordance is caused by rampant mitochondrial introgression in *Neodiprion* (*Hymenoptera: Diprionidae*) sawflies. *Evolution*, **61**, 1417–38.
- Losos JB, Glor RE (2003) Phylogenetic comparative methods and the geography of speciation. *Trends in Ecology & Evolution*, **18**, 220–227.
- Lucy D, Aykroyd R (2012) GenKern: Functions for generating and manipulating binned kernel density estimates. R package version 1.2-10.
- Machado CA, Kliman RM, Markert JA, Hey J (2002) Inferring the history of speciation from multilocus DNA sequence data: the case of *Drosophila pseudoobscura* and close relatives. *Molecular Biology and Evolution*, **19**, 472–488.

- Maddison WP (1997) Gene Trees in Species Trees. *Systematic Biology*, **46**, 523.
- Madureira ML (1981) Discriminant analysis in Portuguese pine voles: *Pitymys lusitanicus* Gerbe and *Pitymys duodecimcostatus* de Selys-Longchamps (*Mammalia: Rodentia*). *Arq Mus Bocage, Série A I*, 111–122.
- Madureira ML (1982) New data on the taxonomy of Portuguese *Pitymys*. *Arq Mus Bocage, Série A I*, 341–368.
- Mallet J (1995) A species definition for the modern synthesis. *Trends Ecol. Evol.*, **10**, 294–299.
- Mallet J (2008) Hybridization, ecological races and the nature of species: empirical evidence for the ease of speciation. *Philosophical transactions of the Royal Society of London. Series B, Biological Sciences*, **363**, 2971–86.
- Mallet J, Meyer a, Nosil P, Feder JL (2009) Space, sympatry and speciation. *Journal of Evolutionary Biology*, **22**, 2332–41.
- Marchant AD (1988) Apparent introgression of mitochondrial DNA across a narrow hybrid zone in the *Caledia captiva* species-complex. *Heredity*, **60**, 39-46.
- Markova E, Malygin V, Montuire S et al. (2009) Dental Variation in Sibling Species *Microtus arvalis* and *M. rossiaemeridionalis* (*Arvicolinae, Rodentia*): Between-Species Comparisons and Geography of Morphotype Dental Patterns. *Journal of Mammalian Evolution*, **17**, 121–139.
- Martínez JJ, Di Cola V (2011) Geographic distribution and phenetic skull variation in two close species of *Graomys* (*Rodentia, Cricetidae, Sigmodontinae*). *Zoologischer Anzeiger - A Journal of Comparative Zoology*, **250**, 175–194.
- Maruki T, Kumar S, Kim Y (2012) Purifying selection modulates the estimates of population differentiation and confounds genome-wide comparisons across single-nucleotide polymorphisms. *Molecular Biology and Evolution*, **29**, 3617-3623.
- Mayden RL (1997) A hierarchy of species concepts: The denouement in the saga of the species problem. In: *Species: The units of biodiversity* (eds. M. F. Claridge, H. A. Dawah, and M. R. Wilson), pp. 381–424. Chapman and Hall, London.
- Mayr E (1942) Systematics and the origin of species. Columbia University Press, New York.
- McCormack JE, Faircloth BC (2013) Next-generation phylogenetics takes root. *Molecular Ecology*, **22**, 19-20.

- McDevitt AD, Mariani S, Hebblewhite M, Decesare JJ, Morgantini L, Seip D, Weckworth BV, Musiani M (2009) Survival in the Rockies of an endangered hybrid swarm from diverged caribou (*Rangifer tarandus*) lineages. *Molecular Ecology*, **18**, 665–679.
- Meier R, Willmann R (2000) The Hennigian species concept. In: *Species concepts and phylogenetic theory* (eds. Wheeler QD, Meier R), pp. 30–43. Columbia University Press, New York.
- Melo-Ferreira J, Boursot P, Carneiro M et al. (2012) Recurrent introgression of mitochondrial DNA among hares (*Lepus spp.*) revealed by species-tree inference and coalescent simulations. *Systematic Biology*, **61**, 367–81.
- Meredith RW, Janecka JE, Gatesy J, et al. (2011) Impacts of the Cretaceous Terrestrial Revolution and KPg extinction on mammal diversification. *Science*, **334**, 521–524.
- Michener CD (1970) Diverse approaches to systematics. *Evol. Biol.*, **4**, 1–38.
- Mikkola H (1983) Owls of Europe. T & AD Poyser, London.
- Mira A, Mathias ML (2007) *Microtus lusitanicus* (Gerbe, 1879). In: *Atlas y libro rojo de los mamíferos terrestres de España* (eds. Palomo LJ, Gisbert J, Blanco JC), pp. 418–421. Madrid: Dirección General para la Biodiversidad - SECEM – SECEMU.
- Mishler, BD (1985) The morphological, developmental, and phylogenetic basis of species concepts in bryophytes. *Bryologist*, **88**, 207–214.
- Mitchell-Jones AJ, Amori G, Bogdanowicz W, Krystufek B, Reijnders PJH, Spitzenberger F, Stubbe M, Thissen JBM, Vohralik V, Zima J (1999) *The atlas of European mammals*. London: Poyser T, Poyser AD.
- Mouse Genome Sequencing Consortium (2002) Initial sequencing and comparative analysis of the mouse genome. *Nature*, **420**, 520–562.
- Musser GG, Carleton MD (1993) Family *Muridae*. In: *Mammal species of the world: a taxonomic and geographic reference* (eds. Wilson DE, Reeder DM), pp. 501–706. WashingtonDC: Smithsonian Institution Press.
- Nadachowski A, Zagorodnyuk I (1996) Recent *Allophaiomys*-like species in the Palaeartic: Pleistocene relicts or a return to an initial type? *Acta Zool. Cracov.*, **39**, 387–394.

- Nappi A, Brunet-Lecomte P, Montuire S (2005) Dental morphology of *Microtus (Terricola)* voles from Calabria (Southern Italy) and relationships with *M. (T.) savii* (Rodentia, Arvicolidae). *Hystrix It. J. Mamm*, **16**, 75–85.
- Nichols R (2001) Gene trees and species trees are not the same. *Trends Ecol. Evol.*, **16**, 358–364.
- Niethammer J (1970) Über Kleinsäuger aus Portugal. *Bonn. Zool. Beitr.*, **21**, 89-118.
- Niethammer J, Krapp F (1982) Handbuch der Säugetiere Europas. *Wiesbaden (Germany): Akademische Verlagsgesellschaft*.
- Nosil P (2012) Ecological Speciation. Oxford University Press, Oxford, New York.
- Nowak RM (1999) Walker's Mammals of the World, sixth ed. The Johns Hopkins University Press, Baltimore, ML.
- O'Neill EM, Schwartz R, Bullock CT, Williams JS, Shaffer HB, Aguilar-Miguel X, Parra-Olea G, Weisrock DW (2013) Parallel tagged amplicon sequencing reveals major lineages and phylogenetic structure in the North American tiger salamander (*Ambystoma tigrinum*) species complex. *Molecular Ecology*, **22**, 111-129.
- Padial JM et al. (2010) The integrative future of taxonomy. *Front. Zool.*, **7**, 16.
- Palumbi SR, Baker CS (1994) Contrasting population structure from nuclear intron sequences and mtDNA of humpback whales. *Molecular Biology and Evolution*, **11**, 426-435.
- Pamilo P, Nei M (1988) Relationships between gene trees and species trees. *Molecular Biology and Evolution*, **5**, 568–83.
- Pearson K, (1895) Note on regression and inheritance in the case of two parents. *Proc. R. Soc. Lond.*, **58**, 240-242.
- Pearson K, (1901) On lines and planes of closest fit to systems of points in space. *Phil. Mag. Ser.*, **6**, 2, 559-572.
- Peters JL, Zhuravlev Y, Fefelov I, Logie A, Omland KE (2007) Nuclear loci and coalescent methods support ancient hybridization as cause of mitochondrial paraphyly between gadwall and falcated duck (*Anas spp.*) *Evolution*, **61**, 1992–2006.
- Peters JL, Zhuravlev YN, Fefelov I, Humphries EM, Omland KE (2008) Multilocus phylogeography of a holarctic duck: colonization of north america from eurasia by gadwall (*Anas strepera*). *Evolution*, **62**, 1469-1483.

- Petit RJ, Excoffier L (2009) Gene flow and species delimitation. *Trends in Ecology & Evolution*, **24**, 386–93.
- Piras P, Marcolini F, Raia P, Curcio M, Kotsakis T (2010) Ecophenotypic variation and phylogenetic inheritance in first lower molar shape of extant Italian populations of *Microtus (Terricola) savii* (Rodentia). *Biological Journal of the Linnean Society*, **99**, 632–647.
- Pompanon F, Bonin A, Bellemain E, Taberlet P (2005) Genotyping errors: causes, consequences and solutions. *Nature Reviews. Genetics*, **6**, 847–59.
- Porter CA, Goodman M, Stanhope MJ (1996) Evidence on mammalian phylogeny from sequences of exon 28 of the von Willebrand factor gene. *Mol. Phylogenet. Evol.*, **5**, 89–101.
- Posada D (2008) jModelTest: phylogenetic model averaging. *Molecular Biology and Evolution*, **25**, 1253-1256.
- Poulton EB (1903) What is a species? *Trans. Entomol. Soc. Lond.*, **1903**, 77–116.
- Quinn G, Keough MJ (2002) *Experimental Design and Data Analysis for Biologists*. Cambridge University Press, Cambridge, UK.
- Rambaut A, Grassly NC (1997) Seq-Gen: An application for the Monte Carlo simulation of DNA sequence evolution along phylogenetic trees. *Comput. Appl. Biosci.*, **13**, 235-238.
- Ramos-Onsins SE, Rozas J (2002) Statistical properties of new neutrality tests against population growth. *Molecular Biology and Evolution*, **19**, 2092–100.
- Randi E (2006) Phylogeography of South European Mammals. In: *Phylogeography of Southern European Refugia* (eds. Weiss S, Ferrand, N), Chapter 3, pp. 101-126. Kluwer Academic Publishers, Amsterdam.
- Rannala B, Yang Z (2003) Bayes estimation of species divergence times and ancestral population sizes using DNA sequences from multiple loci. *Genetics*, **164**, 1645–1656.
- Rat Genome Sequencing Project Consortium (2004) Genome sequence of the Brown Norway rat yields insights into mammalian evolution. *Nature*, **428**, 493-521.
- Ridley M (1989) The cladistic solution to the species problem. *Biol. Philos*, **4**, 1–16.
- Rodríguez C, Peris SJ (2007) Habitat associations of small mammals in farmed landscapes: implications for agrienvironmental schemes. *Anim. Biol.*, **57**, 301–314.

- Rodríguez-Prieto A, Igea J, Castresana J Development of rapidly evolving intron markers to study the speciation and phylogeography of rodents. (En preparación).
- Rohlf FJ (2005a). tpsDig2. Department of Ecology and Evolution, State University of New York at Stony Brook.
- Rohlf FJ (2005b). tpsRelw, relative warps analysis. Department of Ecology and Evolution, State University of New York at Stony Brook.
- Rohlf FJ, Bookstein FL (2003) Computing the uniform component of shape variation. *Syst. Biol.*, **52**, 66-69.
- Rohlf FJ, Marcus LF (1993) A Revolution in Morphometrics. *Tree*, **8**, 129–132.
- Rohlf FJ, Slice D (1990) Extensions of the Procrustes methods for the optimal superimposition of landmarks. *Syst. Zool.*, **39**, 40-59.
- Rosen DE (1979) Fishes from the uplands and intermontane basins of Guatemala: Revisionary studies and comparative geography. *Bull. Am. Mus. Nat. Hist.*, **162**, 267–376.
- Rosenberg NA, Nordborg M (2002) Genealogical trees, coalescent theory and the analysis of genetic polymorphisms. *Nat. Rev. Genet.*, **3**, 380–390.
- Rychlik L, Ramalhinho G, Polly PD (2006) Response to environmental factors and competition: skull, mandible and tooth shapes in Polish water shrews (*Neomys*, Soricidae, *Mammalia*). *Journal of Zoological Systematics and Evolutionary Research*, **44**, 339–351.
- Sacchi R, Pupin F, Rosa DP, Fasola M (2007) Bergmann's rule and the Italian Hermann's tortoises (*Testudo hermanni*): latitudinal variations of size and shape. *Amphibia-Reptilia*, **28**, 43–50.
- Salicini I, Ibáñez C, Juste J (2011) Multilocus phylogeny and species delimitation within the Natterer's bat species complex in the Western Palearctic. *Molecular Phylogenetics and Evolution*, **61**, 888-898.
- Salzburger W, Ewing GB, Von Haeseler A (2011) The performance of phylogenetic algorithms in estimating haplotype genealogies with migration. *Molecular Ecology* **20**, 1952-1963.
- Sánchez-Gracia A, Castresana J (2012) Impact of deep coalescence on the reliability of species tree inference from different types of DNA markers in mammals. *PLoS One*, **7**, e30239.
- Sanders KL, Lee MSY (2007) Evaluating molecular clock calibrations using Bayesian analyses with soft and hard bounds. *Biology Letters*, **3**, 275–9.

- Sang T, Zhong Y (2000) Testing hybridization hypotheses based on incongruent gene trees. *Systematic Biology*, **49**, 422–434.
- Santos SM (2009a) Factors influencing the distribution of the lusitanian and the mediterranean pine voles (*Microtus lusitanicus* and *M. duodecimcostatus*) in Portugal: A multiscale approach. Tesis doctoral. Universidade de Lisboa. Portugal.
- Santos SM, Mira AP, Mathias ML (2009b) Factors influencing large-scale distribution of two sister species of pine voles (*Microtus lusitanicus* and *Microtus duodecimcostatus*): the importance of spatial autocorrelation. *Canadian Journal of Zoology*, **87**, 1227–1240.
- Santos SM, Mathias ML, Mira AP (2010) Local coexistence and niche differences between the Lusitanian and Mediterranean pine voles (*Microtus lusitanicus* and *M. duodecimcostatus*). *Ecological Research*, **25**, 1019–1031.
- Sato A, Tichy H, O’Higin C, Grant PR, Grant BR, Klein J (2001) On the origin of Darwin’s finches. *Mol. Biol. Evol.*, **18**, 299–311.
- Savolainen V, Anstett MC, Lexer C, Hutton I, Clarkson JJ, Norup MV, Powerll MP, Springate D, Salamin N, Baker WJ (2006) Sympatric speciation in palms on an oceanic island. *Nature*, **441**, 210–213.
- Schlick-Steiner BC et al. (2010) Integrative taxonomy: a multisource approach to exploring biodiversity. *Annu. Rev. Entomol.*, **55**, 421–438.
- Schliwen UK et al. (1994) Sympatric speciation suggested by monophyly of crater lake cichlids. *Nature*, **368**, 629–632.
- Shawyer C (1998) The Barn owl. Arlequin Press, Essex.
- Seehausen O (2004) Hybridization and adaptive radiation. *Trends in Ecology & Evolution*, **19**, 198–207.
- Sequeira F, Sodré D, Ferrand N et al. (2011) Hybridization and massive mtDNA unidirectional introgression between the closely related Neotropical toads *Rhinella marina* and *R. schneideri* inferred from mtDNA and nuclear markers. *BMC Evolutionary Biology*, **11**, 264.
- Serb JM, Alejandrino A, Otárola-Castillo E, Adams DC (2011) Morphological convergence of shell shape in distantly related scallop species (*Mollusca: Pectinidae*). *Zoological Journal of the Linnean Society*, **163**, 571–584.
- Shurtliff QR (2011) Mammalian hybrid zones: a review. *Mammal Review*, **43**, 1–21.

- Sidlauskas B (2008) Continuous and Arrested Morphological Diversification in Sister Clades of Characiform Fishes: a Phylomorphospace Approach. *Evolution*, **62**, 3135–3156.
- Simpson GG (1961) Principles of animal taxonomy. Columbia University Press, New York.
- Sites JW, Marshall JC (2004) Operational criteria for delimiting species. *Annu. Rev. Ecol. Evol. Syst.*, **35**, 199–227.
- Slice DE, Bookstein FL, Marcus LF, Rohlf F J (1996) Appendix I: A glossary for geometric morphometrics. In: *Advances in Morphometrics* (eds. Marcus LF, Corti M, Loy A et al.), pp. 531–551. Plenum Press.
- Smadja CM, Butlin RK (2011) A framework for comparing processes of speciation in the presence of gene flow. *Molecular ecology*, **20**, 5123–40.
- Smedley D, Haider S, Ballester B, Holland R, London D, Thorisson G, Kasprzyk A (2009) BioMart - biological queries made easy. *BMC Genomics*, **10**, 22.
- Smith HM (1955) The perspective of species. *Turtlex News*, **33**, 74–77.
- Sneath PHA, Sokal RR (1973) Numerical taxonomy: The principles and practice of numerical classification. Freeman WH, San Francisco.
- Sokal RR, Crovello TJ (1970) The biological species concept: A critical evaluation. *Am. Nat.*, **104**, 127–153.
- Soria-Carrasco V (2011) Global analyses of mammalian diversification from molecular phylogenies. Tesis doctoral. Universitat Barcelona, Spain.
- Soria-Carrasco V, Talavera G, Igea J, Castresana J (2007) The K tree score: quantification of differences in the relative branch length and topology of phylogenetic trees. *Bioinformatics*, **23**, 2954–2956.
- Spaeth PA (2009) Morphological convergence and coexistence in three sympatric North American species of *Microtus* (*Rodentia: Arvicolinae*). *Journal of Biogeography*, **36**, 350–361.
- Stamatakis A (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic Analyses with thousands of taxa and mixed models. *Bioinformatics*, **22**, 2688–2690.
- Stanhope MJ, Czelusniak J, Si JS, Nickerson J, Goodman M (1992). A molecular perspective on mammalian evolution from the gene encoding Interphotoreceptor Retinoid Binding Protein, with convincing evidence for bat monophyly. *Mol. Phylogenet. Evol.*, **1**, 148–160.

- Stephens M, Smith N, Donnelly P (2001) A new statistical method for haplotype reconstruction from population data. *American Journal of Human Genetics*, **68**, 978-989.
- Swofford DL (2003) *PAUP*: Phylogenetic Analysis Using Parsimony (*and Other Methods)*. Version 4. Sinauer Associates, Sunderland, Massachusetts.
- Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**, 585-595.
- Talavera G, Castresana J (2007) Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Systematic Biology*, **56**, 564-577.
- Tamura K, Peterson D, Peterson N *et al.* (2011) MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Molecular Biology and Evolution*, **28**, 2731–2739.
- Taylor I (1994) *Barn owls: predator–prey relationships and conservation*. Cambridge University Press, Cambridge, UK.
- Van der Meulen AJ (1973) Middle Pleistocene smaller mammals from the Monte Peglia (Orvieto, Italy) with special reference to the phylogeny of *Microtus* (*Arvicolinae*, *Rodentia*). *Quaternaria*, **17**, 1–144.
- Van der Meulen AJ (1978) *Microtus* and *Pitymys* (*Arvicolinae*) from Cumberland cave, Maryland, with a comparison of some New and Old World species. *Annals of Carnegie Museum*, **47**, 101–145.
- VanValen L (1976) Ecological species, multispecies, and oaks. *Taxon*, **25**, 233–239.
- Veiga JP (1986) Interannual fluctuations of three microtine populations in Mediterranean environments: the effect of rainfall. *Mammalia*, **50**, 114–116.
- Venter JC, Adams MD, Myers EW, *et al.* (2001) The sequence of the human genome. *Science*, **291**, 1304-1351.
- Ventura J, Gisbert J, Jiménez L (2010) Breeding characteristics of the Lusitanian pine vole, *Microtus lusitanicus*. *Animal Biology*, **60**, 1–14.
- Wallace SC (2006) Differentiating *Microtus Xanthognathus* and *Microtus Pennsylvanicus* Lower First Molars Using Discriminant Analysis of Landmark Data. *Journal of Mammalogy*, **87**, 1261–1269.
- Wand MP, Jones MC (1995). *Kernel smoothing*. Chapman & Hall, London, United Kingdom.

- Wang RL, Wakeley J, Hey J (1997) Gene flow and natural selection in the origin of *Drosophila pseudoobscura* and close relatives. *Genetics*, **147**, 1091–1106.
- Warnes GR *et al.* (2012) gtools: Various R programming tools. R package version 2.7.0.
- Warnes GR *et al.* (2013). gdata: Various R programming tools for data manipulation. R version 2.12.0.2.
- Webb TJ, Gaston KJ (2000) Geographic range size and evolutionary age in birds. *Proc. R. Soc. Lond. Ser. B*, **267**, 1843–1850.
- Wiens JJ, Penkrot TA (2002) Delimiting species using DNA and morphological variation and discordant species limits in spiny lizards (*Sceloporus*). *Syst. Biol.*, **51**, 69–91.
- Wright S (1940) The statistical consequences of Mendelian heredity in relation to speciation. In: *The new systematics* (eds. Huxley J), pp. 161–183. Oxford University Press, London.
- Wyner YM, Johnson SE, Stumpf RM, Desalle R (2002) Genetic assessment of a white-collared x red-fronted lemur hybrid zone at Andringitra, Madagascar. *American Journal of Primatology*, **67**, 51–66.
- Zagorodnyuk IV (1989) Taksonomiya, rasprostranenie i morfologicheskaya izmenchivost' polevok roda *Terricola* vostochnoi Evropy. [Taxonomy, distribution and morphological variation of the *Terricola* voles in East Europe]. *Vestnik Zool.*, **5**, 3–14.
- Zakrzewski RJ (1985) The fossil record. In: *Biology of New World Microtus* (eds Tamarin R.H), pp. 1–51, No. 8. American Society of Mammalogists. Special Publication.
- Zheng SH, Zhang ZQ (2000) Late Miocene–Early Pleistocene micromammals from Wenwanggou of Lingtai, Gansu, China. *Vertebrata Pal Asiatica*, **38**, 58–71.
- Zuckerlandl E, Pauling L (1965) In: *Evolving Genes and Proteins* (eds. Bryson V, Vogel H J), pp. 97–166. Academic Press, New York.