UNIVERSIDAD AUTÓNOMA DE MADRID
FACULTAD DE CIENCIAS, DEPARTAMENTO DE BIOLOGÍA

ESTUDIO FILOGEOGRÁFICO DE UNA ESPECIE LITORAL, DISYUNTA, CORSO-SARDA / IBÉRICA: *Armeria pungens* (PLUMBAGINACEAE)

TESIS DOCTORAL
ROSALÍA PIÑEIRO PORTELA

DIRECTORES:
GONZALO NIETO FELINER
JAVIER FUERTES AGUILAR

TUTOR:
HELIOS SAINZ OLLERO

Madrid, 2007

INDICE

ANEXOS

INTRODUCCIÓN GENERAL

El estudio de los procesos evolutivos en el nivel intraespecífico

La genética de poblaciones clásica explica la estructura genética de las especies por la interacción entre: 1) el patrón de flujo génico actual entre poblaciones y 2) la adaptación a las condiciones ecológicas actuales (Schaal *et al*., 1998). La genética de poblaciones se centra, pues, en la estima de parámetros de la diversidad genética asumiendo condiciones de equilibrio histórico de las poblaciones (equilibirio Hardy-Weinberg).

La aparición de la filogeografía en 1987 revoluciona los estudios poblacionales añadiendo una perspectiva temporal (Schaal *et al*., 1998; Schaal y Olsen, 2000; Avise, 2000; Widmer y Lexer, 2001). Esta disciplina estudia las frecuencias génicas y relaciones filogenéticas entre poblaciones (filo) en un contexto geográfico explícito (geografía), lo que permite considerar la influencia de los eventos históricos (fluctuación de tamaños poblacionales, colonizaciones, expansiones del área de distribución, etc.), además del flujo génico y la selección natural, en la evolución intraespecífica (Avise 1989; Schaal y Olsen, 2000).

La filogeografía se basa en la teoría de la coalescencia, según la cual en una población de tamaño constante, reaparecen nuevos alelos por mutación o se pierden alelos por deriva génica, de manera que todos los alelos actuales derivan de un alelo ancestral común, o lo que es lo mismo en un sentido retrospectivo, coalescen en ese único alelo (Schaal *et al*., 1998; Hare, 2001).

El papel del flujo génico en la evolución intraespecífica

Bajo la visión Neo-Darwinista, se reformuló el concepto biológico de especie, (Dobzhansky, 1937; Mayr, 1942; Stebbins, 1950), que considera las especies como entidades reproductivas aisladas. Se creía, pues, que era el flujo génico extensivo entre las poblaciones de una misma especie el responsable del mantenimiento de las especies.

Sin embargo, en los años 70 se llevan a cabo un buen número de experimentos de estima directa de la dispersión, que muestran que ésta raramente tiene lugar a lo largo

de todo el rango de distribución de una especie [Ehrlich y Raven (1969), Sokal (1973), Raven (1973) y Levin (1979), citados por Grant (1981)]. Esto implicaba que el flujo génico no es extensivo y, con ello, insuficiente para compensar la deriva génica y la adaptación a las condiciones locales, enfatizando la acción de la evolución a nivel poblacional (Ehrlich y Raven, 1969).

Posteriormente, a partir de estudios de paternidad (ver Ellstrand, 2003) y, sobre todo, aplicando filogenias moleculares, surgen estimas indirectas de flujo génico muy superiores a lo que se creía en base a las medidas directas. No obstante, se siguen documentando casos en los que el flujo génico es insuficiente para mantener la integridad a nivel de especie [Slatkin (1985), Alltrand y Ellam (1993), Levin (2000), citados por Morjan y Rieseberg (2004)]. En este contexto, la visión de la hibridación ha ido cambiando hasta su consideración actual como fenómeno frecuente y bien documentado (Rieseberg y Wendel, 1993; Mallet, 2005) que puede afectar a la estructura poblacional mucho después de que la divergencia "irreversible" entre especies haya comenzado. Las consecuencias de la hibridación pueden ser no adaptativas, como en el caso del reforzamiento de barreras reproductivas (Dobzhansky, 1937) o la extinción de especies de distribución restringida (Rhymer y Simberloff, 1996; Levin *et al*., 1996; Huxel, 1999). Pero en trabajos modernos se ha puesto de relieve su posible papel adaptativo a través de diversos mecanismos, como el incremento de la variabilidad genética de los linajes, la fijación de la heterocigosidad, la purga de alelos deletéreos y, sobre todo, por la creación de nuevas combinaciones genéticas (Anderson y Stebbins, 1954; Lewotin y Birch, 1966; Arnold, 1997, 2004; Ellstrand, 2003). Este último mecanismo se hace evidente por la existencia de híbridos estabilizados de segunda generación (incluso de generaciones superiores) o, en última instancia, la detección de especies híbridas, fenómeno muy frecuente en plantas en el nivel poliploide, aunque algo menos en el homoploide (Arnold, 1997; Rieseberg, 1997). En definitiva, la hibridación parece jugar un papel en la evolución de especies con barreras reproductivas débiles.

Evolución intraespecífica de *Armeria pungens y A. maderensis*

Esta memoria doctoral es un estudio de filogeografía de la especie *Armeria pungens* y del endemismo insular *Armeria maderensis*. Para ello se han buscado marcadores moleculares variables a nivel intraespecífico y se ha tratado de extraer su señal filogenética para reconstruir la historia evolutiva de estas dos especies. En el caso de *A. pungens*, se han utilizado tres marcadores nucleares de características diversas y tres regiones del ADN cloroplástico, que se han contrastado con un detallado estudio morfométrico.

A partir de las distancias genéticas y relaciones genealógicas entre poblaciones reveladas por los marcadores moleculares, así como patrones de la diversidad genética, se han identificado los distintos linajes genéticos y se ha intentado reconstruir (i) los patrones de flujo génico entre ellos, (ii) la relación entre la estructura genética y las condiciones ecológicas, y, finalmente, (iii) los eventos históricos acontecidos.

Desde la revisión taxonómica del género *Armeria* en la Península Ibérica en 1990, el grupo de investigación en el que se encuadra esta tesis doctoral ha seguido una línea de investigación en el Real Jardín Botánico de Madrid sobre la evolución en este género. A partir de las pautas de variación morfológica y eco-geográfica de las especies de *Armeria*, se propuso como punto de partida un escenario evolutivo en el que la hibridación jugaba un papel fundamental (Nieto Feliner, 1987, 1988). Seguidamente, se procedió a un programa de hibridación experimental del que se concluyó que las barreras reproductivas entre las distintas especies son sumamente débiles y la recuperación de la viabilidad polínica en los retrocruces es rápida (Nieto Feliner *et al*., 1996; Nieto Feliner, 1997). A continuación, se inició el estudio con marcadores ITS utilizando un enfoque filogenético, del que se concluyó que este marcador sigue un patrón geográfico independiente de la taxonomía en *Armeria.* La interpretación más probable es que este patrón resulte de la hibridación entre especies seguida de un proceso de evolución concertada por el que se fija una de las copias de uno de los parentales en una zona geográfica (Fuertes Aguilar *et al*., 1999). Finalmente, se pasó al uso de marcadores moleculares utilizando un enfoque filogeográfico. Se estudiaron tres especies distribuidas a distintas alturas del macizo Sierra Nevada y se comprobó que la variación genética, también a esta escala más fina, sigue un patrón geográfico, sugiriendo eventos de reticulación en la evolución de estas tres especies.

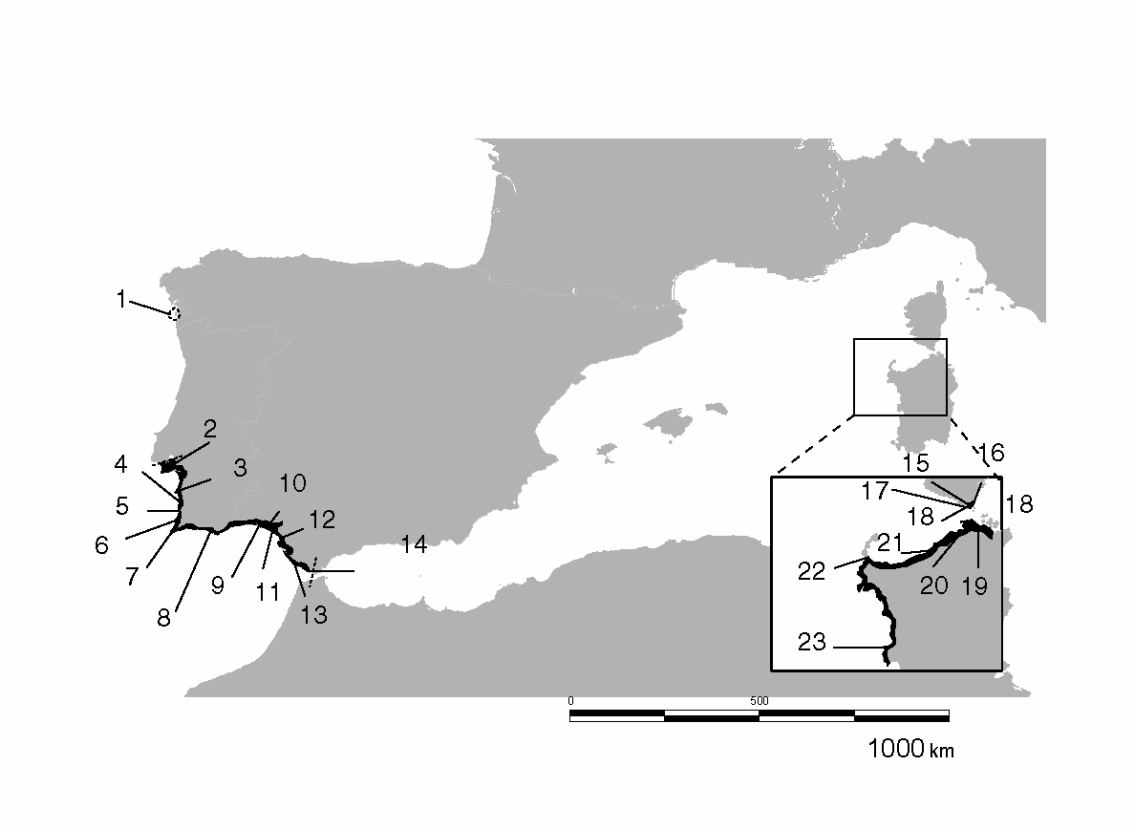
En este trabajo, se ha escogido un sistema lineal simple dentro del género, con lo que se espera un papel más limitado de la evolución reticulada que haga más fácil su detección. En *Armeria* siempre ha de tenerse en cuenta la influencia potencial mediante reticulación del los congéneres simpátricos. En nuestro caso, esto último se ha abordado utilizando marcadores ITS y cloroplásticos.

Rango geográfico actual de *Armeria pungens* y de sus congéneres costeros en el W del Mediterráneo

Armeria pungens es un arbusto de arenales costeros cuya área principal se localiza en el SW de la Península Ibérica, entre el río Tajo y el estrecho de Gibraltar. Presenta, asimismo, dos áreas disyuntas en islas continentales: el archipiélago de las islas Cíes, en el Atlántico, y las islas de Córcega y Cerdeña, en el Mediterráneo (Fig. 1). Ambos sistemas de islas tienen historias geológicas muy distintas. Las islas Cíes, frente a la ría de Vigo, se separaron de las costas de Galicia muy recientemente, tras las últimas glaciaciones, y actualmente están a tan sólo 2,5 Km de la costa. Hay constancia de que las islas han estado sometidas a la acción antrópica desde épocas prehistóricas y en la actualidad la vegetación natural está notablemente degradada. Debido a esto y a su escaso aislamiento no poseen ningún elemento florístico estrictamente endémico (Fernández de la Cigoña, 1988; Guitián y Guitián, 1990; Luaces y Toscano, 1998).

Córcega y Cerdeña están mucho más aisladas. Concretamente, Córcega se sitúa a 160 Km de Provenza y 82 Km de la Toscana. Entre ambas islas hay una distancia de 12 Km. La historia geológica de la placa corso-sarda es compleja; en el Oligoceno superior la placa estaba unida a las costas de Provenza, y hace 16 m.a. comenzó la rotación en el sentido contrario de las agujas del reloj hasta su posición actual (Krijgsman, 2002). Desde entonces, las conexiones con el continente han tenido lugar durante la desecación del Mesiniense y, aunque no todos los autores están de acuerdo (e.g. Gamisans, 1991), podrían haberse producido también durante las glaciaciones. La flora actual de Córcega y Cerdeña presenta afinidad con las floras de las Baleares, Provenza, los Alpes, los Pirineos y la Península Ibérica, lo que a menudo se considera consecuencia de esta posición de la placa sardo-corsa en el Terciario (Briquet, 1910; Cardona y Contrandiopoulos, 1979). En la actualidad Gamisans (1991) recoge 75 táxones endémicos.

Figura 1 Distribución actual simplificada de *Armeria pungens***.** Se indica la localización de cada una de las 23 poblaciones muestreadas en este trabajo.



En la Península Ibérica, las especies litorales y sublitorales de *Armeria* se restringen a las costas cantábrica y atlántica (Fig. 2; Bernis, 1955; Devesa, 1987; Nieto Feliner, 1990). En los sectores central y SW se concentra una elevada diversidad de especies, todas endémicas, pertenecientes a la sección *Macrocentron* (sección a la que pertenece también *A. pungens*). Al norte del río Tajo se localizan las especies costeras *A. welwitschii*, especie de duna distribuida desde la desembocadura del Tajo hasta Cabo Montego y *A. berlengensis*, emparentada con la anterior y restringida a los acantilados graníticos costeros de las islas Berlengas. En esta zona se sitúa también la especie, no perteneciente a la sección *Macrocentron*, *A. pseudoarmeria*, endémica del Cabo da Roca y alrededores. En la costa SW, por debajo de la desembocadura del Tajo, se hallan seis especies sublitorales de la sección *Macrocentron*: *A. pinifolia* y *A. rouyana* , al N del cabo S. Vicente, y *A. hirta*, *A. gaditana* *A. macrophylla* y *A. velutina* al S de dicho cabo.

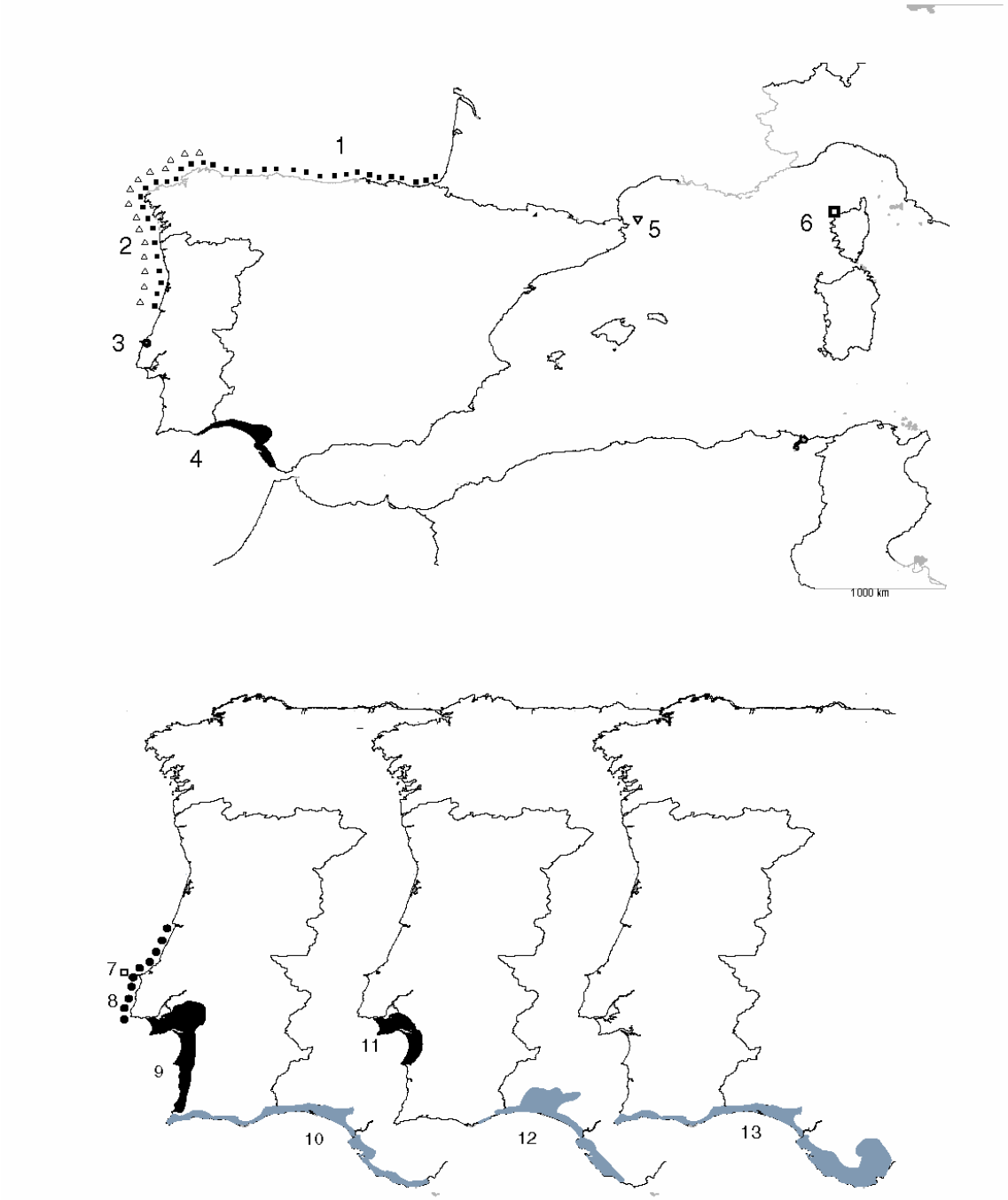
Los sectores N y NW de la Península Ibérica están dominados por el agregado *maritima*, representado por *A. maritima,* especie de amplia distribución que en la Península Ibérica se localiza en zonas de acantilado o de marisma, y por el endemismo del NW Ibérico, *A. pubigera*, típico de acantilados costeros.

Las islas Cíes son la única localidad donde el agregado *maritima* y la sec. *Macrocentron* se ponen en contacto, ya que en este archipiélago conviven las especies *A. maritima*, *A. pubigera* y *A. pungens*.

En la vertiente mediterránea el género *Armeria* se ausenta de la costa, de modo que falta en las zonas litorales del E de la Península Ibérica, S de Francia y de la Penínula Itálica. La única excepción es *A. ruscinonensis* ssp. *ruscinonensis*, endémica de los acantilados costeros del NE de la Península Ibérica y SE de Francia (Bianchini, 1982; Pinto da Silva, 1972). En el N de Marruecos aparece la especie costera *A. simplex*, también perteneciente a la sección *Macrocentron* y distribuida en Argelia y Marruecos. (Nieto Feliner, 2002).

En Córcega y Cerdeña hay descritas seis especies endémicas de *Armeria*. *A. leucocephala, A. multiceps, A. sardoa, A. soleirolii*, *A. sulcitana* y *A. morisii.* A morisii parece estar morfológica y ecológicamente aislada de las demás, mientras que las otras cinco especies muestran gran similitud morfológica (Bernis, 1954). De acuerdo a una filogenia de ITS (Fuertes Aguilar *et al.*, 1999) este grupo de armerias endémicas está emparentado con especies del E de la Península Ibérica (*A. fontquerii* y *A. trachyphylla*). *A. soleirolii* es la única especie costera, concretamente crece en los acantilados de los alrededores de Calvi, al N de Córcega (Briquet y Litardière, 1955; Arrigoni, 1970; Bianchini, 1982; Gamisans y Jeanmonod, 1993; Paradis y Culioli, 2003).

Figura 2 Distribución simplificada de las especie litorales y sublitorales de *Armeria* en la Península Ibérica e islas de Córcega y Cerdeña. **1**. *A. maritima*; **2**. *A. pubigera*; **3**. *A. pseudoarmeria*; **4**. *A. gaditana*; **5**. *A. ruscinonensis* ssp. *ruscinonensis*; **6**. *A. soleirolii*; **7**. *A. berlengensis*; **8**. *A. welwitschii*; **9**. *A. pinifolia*; **10**. *A. macrophylla*; **11**. *A. rouyana*; **12**. *A. velutina*; **13**. *A. hirta* (según Bernis, 1955; Arrigoni, 1970; Devesa, 1987; Nieto Feliner, 1990).



Estructura de la memoria doctoral

Esta memoria doctoral se ha organizado en seis capítulos. Los cinco primeros son formalmente independientes y poseen formato de manuscrito científico, de modo que pueden leerse por separado o, incluso, alterar el orden de su lectura. Los capítulos 2-5 están en distintas fases de elaboración para ser publicados en revistas científicas. Los capítulos 1-4 tratan distintos aspectos relacionados con la filogeografía de *Armeria pungens*, cada uno con distintos marcadores obtenidos a partir de un muestreo de 23 poblaciones que cubre todo el rango de distribución de la especie. El capítulo 5 es un trabajo de conservación de *Armeria maderensis.* En el capítulo 6 se integra y relaciona la información de todos los capítulos anteriores.

En el capítulo 1 se examina variabilidad morfológica dentro de *A. pungens* con métodos de análisis multivariante a fin de evaluar su coherencia con la geografía y el patrón de variación genética. Con la misma metodología se lleva a cabo la comparación con otras especies del género, utilizando medidas de 24 especies obtenidas previamente en el grupo de investigación. Se ha puesto especial interés en la comparación con especies afines de la sección *Macrocentron*.

El capítulo 2 es una aproximación a la filogeografía de *Armeria pungens* con marcadores AFLP. Se pone énfasis en explicar la disyunción Atlántico-Mediterránea de la especie y se contrastan las hipótesis de vicarianza *vs*. dispersión con base en las distancias genéticas entre áreas y estimas de diversidad genética. Tras obtener resultados que apoyan la hipótesis de dispersión a larga distancia, se intenta sopesar la importancia relativa de la probabilidad de dispersión *vs*. limitación ecológica durante el establecimiento. Para ello se ha investigado la posible correlación de los linajes genéticos con las rutas migratorias de aves entre el Atlántico y el Mediterráneo o con las condiciones ecológicas de las distintas áreas. Para esto último se ha construido un modelo bioclimático basado en SIG que permite comparar el hábitat potencial de los distintos linajes identificados dentro de de *A. pungens*. Este capítulo está publicado desde mayo en *Molecular Ecology.*

El capítulo 3 es un trabajo más detallado de filogeografía de *A. pungens* basado en ITS y tres secuencias no codificantes del cloroplasto. El objetivo es obtener una genealogía de ADN cloroplástico para examinar si los fenómenos de colonización han

dejado huella genética en el ADN de herencia materna. Se pretende confirmar los resultados del capítulo anterior respecto a la introducción en Córcega y Cerdeña, así como reconstruir más detalladamente la historia evolutiva en la Península ibérica. Además, la comparación entre marcadores cloroplásticos y nucleares permite detectar eventos de hibridación con otras especies, por lo que se han incluido en el estudio 10 táxones simpátricos. Este capítulo está en una fase muy avanzada de redacción para ser enviado a *New Phytologist*.

En el capítulo 4 se añade a la información anterior el marcador nuclear de bajo número de copias *GapC*. Se pretende confirmar la coherencia de este marcador con los marcadores nucleares estudiados en los capítulos 2 y 3: AFLP e ITS, respectivamente. Gran parte del esfuerzo de esta investigación se centra en resolver las limitaciones metodológicas y de análisis de este tipo de marcadores para extraer la señal filogeográfica que puedan contener.

El capítulo 5 es un análisis basado en AFLP del endemismo de isla oceánica *A. maderensis*. Se han examinado la estructura geográfica de la variabilidad genética así como sus niveles de diversidad para evaluar su status de conservación. Como referencia, se han tomado parámetros de la diversidad genética de A. pungens recalculados a partir de los datos AFLP del capítulo 3. Los marcadores AFLP de ambas especies se han obtenidos usando exactamente los mismos protocolos y combinaciones de primers. Este capítulo se ha enviado a *Biological Conservation*.

El capítulo 6 es una discusión general de todos los resultados de la memoria doctoral. Se integra la información filogeográfica aportada por todos los marcadores estudiados en *A. pungens*, discutiendo las limitaciones y ventajas de cada uno de ellos. También justifica por qué la comparación entre *A. pungens* y *A. maderensis*, a priori interesante por su distribución respectiva en islas continentales y en una isla oceánica, resultó menos útil de lo que se esperaba, debido a la historia evolutiva tan dispar de ambas especies.

REFERENCIAS

Anderson E, Stebbins JR (1954) Hybridization as an Evolutionary Stimulus. *Evolution* **8**, 378-288.

Arnold ML (1997) *Natural Hybridization and Evolution.* Oxford University Press, Oxford.

Arnold ML (2004) Transfer and origin of adaptations through natural hybridization: were Anderson and Stebbins right? *Plant Cell* **16**, 562–570.

Arrigoni PV (1970) Contributto alla conoscenza delle Armerie sardo-corse. *Webbia* **25**, 137–182.

Avise J (1989) Gene trees and organismal histories: a phylogenetic approach to population biology. *Evolution* **43**, 1192.

Avise J (2000) *Phylogeography: the History and Formation of Species.* Harvard University Press, Cambridge, MA.

Bernis F (1954) Revisión del género *Armeria* Willd. con especial referencia a los grupos ibéricos. *Anales del Instituto Botánico A. J. Cavanilles* **11**, 5–288.

Bernis F (1955) Revisión del género *Armeria* Willd. con especial referencia a los grupos ibéricos. Parte segunda (descriptiva de los grupos ibéricos). *Anales del Instituto Botánico Cavanilles* **12**, 77–252.

Bianchini F (1982) *Armeria*. In: *Flora de Italia* (ed. S. P), pp. 294-301. Edagricole, Bologna.

Briquet J (1910) *Prodrome de la Flore Corse.* Georg, Genève, Bâle, Lyon.

Briquet J, De Litardiére R (1955) *Armeria*. In: *Prodrome de la Flore Corse*. Tome III. (eds. Briquet J, De Litardiére R), pp. 1-12. Éditions Paul Lechevalier. Paris.

Cardona MA, Contandriopoulos J (1979) Endemism and evolution in the islands of the Western Mediterranean. In: *Plants and Islands* (ed. Bramwell D). Academic Press, London.

Devesa JA (1987) *Armeria*. In: *Flora Vascular de Andalucía Occidental* Vol. I. (eds. Valdés B, Talavera S, Fernández-Galiano E), pp. 297-302. Ketres, Barcelona.

Dobzhansky T (1937) *Genetics and the Origin of Species.* Columbia University Press, New York.

Ehrlich P, Raven P (1969) Differentiation of populations. *Science* **165**, 1228-1232.

Ellstrand NC (2003) Current knowledge of gene flow in plants: implications for transgene flow. *Philosophical Transactions of the Royal Society B: Biological Sciences* **358**, 1163-1170.

Fernández de la Cigoña Núñez E (1988) *Islas Cíes, Parque Natural de Galicia (historia, etnografía, geografía, flora y fauna)* Fernández de la Cigoña Núñez, Vigo.

Fuertes Aguilar J, Rossello JA, Nieto Feliner, G (1999) Molecular evidence for the compilospecies model of reticulate evolution in *Armeria* (Plumbaginaceae). *Systematic Biology* **48**, 735-754.

Gamisans J (1991) *La végétation de la Corse in Compléments au Prodrome de la Flore Corse, annexe 2.* Conservatoire et Jardin botaniques de la ville de Genève, Genève.

Gamisans J, Jeanmonod D (1993) *Catalogue des plantes vasculaires de la Corse in Compléments au Prodrome de la Flore Corse, annexe 3.* Conservatoire et Jardin botaniques de la ville de Genève.

Grant V (1981) *Plant Speciation.* Columbia University Press, New York.

Guitián J, Guitián P (1990) *A Paisaxe Vexetal das Illas Cíes.* Consellería de Agricultura, Ganadería e Montes. Xunta de Galicia.

Hare M (2001) Prospects for nuclear gene phylogeography. *Trends in Ecology & Evolution* **16**, 700-706.

Huxel G (1999) Rapid displacement of native species by invasive species: effects of hybridization. *Biological Conservation* **89**, 143-152.

Krijgsman W (2002) The Mediterranean: mare nostrum of Earth sciences. *Earth and Planetary Science Letters* **205**, 1-12.

Levin D, Francisco-Ortega J, Jansen RK (1996) Hybridization and the extinction of rare plant species. *Conservation Biology* **10**, 10-16.

Lewontin R, Birch L (1966) Hybridization as a source of variation for adaptation to new environments. *Evolution* **20**, 315-336.

Luaces Anca J, Toscano Novella MC (1998) *Islas Cíes.* Ediciones Nigra Trea, Vigo.

Mallet J (2005) Hybridization as an invasion of the genome. *Trends in Ecology & Evolution* **20**, 229-237.

Mayr E (1942) *Systematics and the origin of species.* Columbia University Press, New York.

Morjan CL, Rieseberg LH (2004) How species evolve collectively: implications of gene flow and selection for the spread of advantageous alleles. *Molecular Ecology* **13**, 1341-1356.

Nieto Feliner G (1987) El género *Armeria* (Plumbaginaceae) en la Península Ibérica: aclaraciones y novedades para una síntesis. *Lagascalia* **15**, 233-236.

Nieto Feliner G (1988) Flujo génico en *Armeria* (Plumbaginaceae) en la Península Ibérica. *Lagascalia* **15 (suppl.)**, 233-236.

Nieto Feliner G (1990) *Armeria*. In: *Flora Iberica* (eds. Castroviejo S, Laínz M, López G*, et al.*), p. 642–721. Consejo Superior de Investigaciones Científicas, Madrid.

Nieto Feliner G (1997) Natural and experimental hybridization in *Armeria* (Plumbaginaceae): *A. salmantica*. *International Journal of Plant Sciences* **158**, 585–592.

Nieto Feliner G (2002) *Armeria* Willd. In: *Catalogue des plantes vasculaires du Nord du Maroc, incluant des clés d'identification* (eds. Valdés B, Rejdali M, Kadmiri AAE, Jury JL, Montserrat JM). Consejo Superior de Investigaciones Científicas, Madrid.

Nieto Feliner G, Izuzquiza A, Lansac AR (1996) Natural and experimental hybridization in *Armeria* (Plumbaginaceae): A. *villosa* subsp. *carratracensis*. *Plant Systematics and Evolution* **201**, 163–177.

Paradis G, Culioli JM (2003) *Armeria pungens* en Corse: état de ses stations et effectifs en 2002. *Le Monde des Plantes* **478**, 1-8.

Rhymer JM, Simberloff D (1996) extinction by hybridization and introgression. *Annual Review of Ecology and Systematics* **27**, 83-109.

Pinto da Silva R (1972) *Armeria* Willd. In: *Flora Europaea.* (eds. Tutin T, Heywood V, Burges N *et al*.). Cambridge University Press, Cambridge, New York.

Rieseberg LH (1997) Hybrid origins of plant species. *Annual Review of Ecology & Systematics* **28**, 359.

Rieseberg LH, Wendel JF (1993) Introgression and its consequences in plants. In: *Hybrid zones and the evolutionary process* (ed. Harrison R), pp. 70-109. Oxford University Press, New York.

Schaal BA, Hayworth DA, Olsen KM, Rauscher JT, Smith WA (1998) Phylogeographic studies in plants: problems and prospects. *Molecular Ecology* **7**, 465-474.

Schaal BA, Olsen KM (2000) Gene genealogies and population variation in plants. *Proceedings of the National Academy of Sciences USA* **97**, 7024-7029.

Sokal RR (1973) The species problem reconsidered. *Systematic zoology* **22**, 360-374.

Stebbins GL (1950) *Variations and Evolution in Plants.* Columbia University Press, New York.

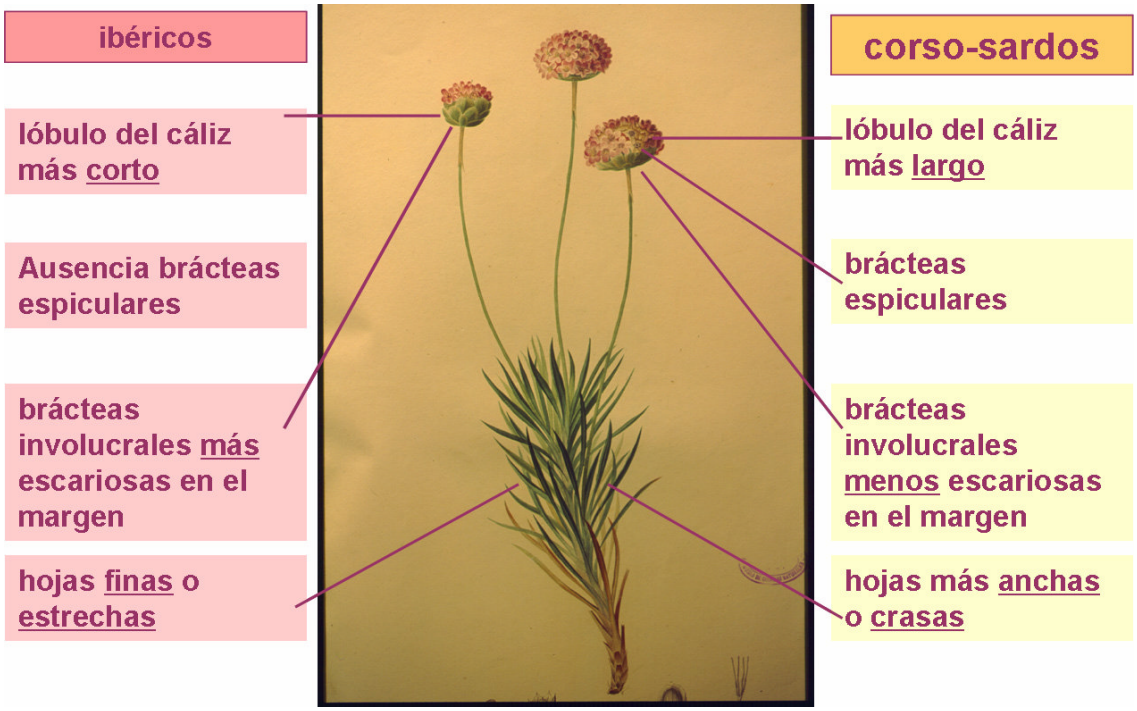
Widmer A, Lexer C (2001) Glacial refugia: sanctuaries for allelic richness, but not for gene diversity. *Trends in Ecology & Evolution* **16**, 267-269.

1. ANÁLISIS MULTIVARIANTE DE LOS CARACTERES MORFOLÓGICOS EN *Armeria pungens*

1.1 INTRODUCCIÓN

Armeria pungens fue descrita por primera vez por Link a partir de ejemplares ibéricos en el año 1800 bajo el nombre de *Statice pungens*. Casi simultáneamente, en 1801, Ventenant describió *Statice fasciculata* a partir de pliegos corsos (Bernis, 1955). La distinción taxonómica de ambas especies se mantuvo a lo largo de todo el siglo XIX, principalmente en base a la presencia de brácteas espiculares (brácteas que axilan las espículas, cimas escorpioides muy reducidas) en *S. fasciculata* y su ausencia en *S. pungens*. Sin embargo, algunos autores de la época, como Cosson en 1850, Willkomm en 1870 y Daveau en 1889 (ver Bernis, 1955), ya habían apuntado el escaso poder discriminante de dicho carácter. Otros caracteres diagnóstico discutidos en la literatura botánica de la época eran la presencia de hojas más anchas, lóbulos del cáliz más largos o margen escarioso de las brácteas involucrales más estrecho en *S. fasciculata* (ver Bernis, 1955; Briquet y Litardière, 1955; Arrigoni, 1970; Figs. 1.1 y 1.2).

Figura 1.1 Caracteres diferenciadores de *Statice pungens* (SW Iberia) y *S. fasciculata* (Córcega y Cerdeña) discutidos en la literatura botánica del s.XIX.



A partir del siglo XX, ya bajo el inequívoco nombre de *Armeria*, todos los autores (Fiore, Coste, Sampaio, Coutinho) coinciden en desechar el poder diferenciador de las brácteas espiculares y fundir ambos táxones en una especie única (ver Bernis, 1955). Mientras que en algunas obras todavía se mantienen táxones intraespecíficos separando las poblaciones ibéricas de las corso-sardas, otras las asimilan completamente.

En 1955 Francisco Bernis lleva a cabo la monografía de los grupos ibéricos. Como el propio autor reconoce, su tratamiento es muy sintético, pues reúne en tan sólo 7 especies toda la diversidad de *Armeria* en la Península Ibérica, a pesar de que éste es uno de los centros de diversidad del género. Subordinadas a estas 7 especies distingue un gran número de subespecies, variedades y formas. Sin embargo, a *A. pungens* le asigna *status* de especie y ni siquiera reconoce categorías subespecíficas. En palabras de Bernis, "no hay en el género *Armeria* otro grupo, de entre los muchos que habitan en la región Mediterránea, que varíe tan poco dentro de un área tan dilatada (y desmenuzada) como es la que posee en la actualidad la sp. *pungens*". Tan sólo menciona la pauta de variación de la longitud del lóbulo calicino, a menudo importante carácter diagnóstico en el género *Armeria*, que disminuye desde los cálices largos de Córcega y Cerdeña, pasando por los de longitud media del SW de la Península Ibérica hasta los lóbulos más cortos en las Islas Cíes.

En su "Prodrome de la Flore Corse", Briquet y Litardière (1955) destacan nuevamente este patrón de variación del lóbulo calicino como único carácter variable en *A. pungens*. Sin embargo, en la revisión de las armerias sardo-corsas de Arrigoni (1970) únicamente se describe el mayor tamaño de las hojas y escapos de las plantas sardas frente a las corsas o ibéricas. Aunque también se apunta la gran variabilidad de la longitud del lóbulo calicino, no se menciona ningún patrón geográfico.

En tratamientos más modernos, se localizan individuos excepcionalmente pubescentes en el SW de Portugal, correspondientes a la var. *velutina* de Coutinho (Pinto da Silva *et al.*, 1972). Esta variedad se recoge en *Flora Europaea* como el resultado de la hibridación entre *A. pungens* y *A. pinifolia,* mientras que tanto la monografía Bernis (1955) como *Flora Iberica* (Nieto Feliner, 1990) la atribuyen a la hibridación entre *A. pungens* y *A. rouyana*. Finalmente, en un trabajo sobre vegetación

costera de Portugal, Asensi *et al.* (1993) la mencionan como uno de los rasgos diferenciadores de las dunas de Bordeira, junto a *Thymus camphoratus* y *Calendula suffruticosa*.

En este trabajo se pretende evaluar la variabilidad morfológica de *Armeria pungens* utilizando un muestreo detallado a nivel poblacional y métodos de análisis multivariante. Los objetivos específicos son: a) analizar los patrones de variación morfológica dentro de la especie; b) evaluar la congruencia de estos caracteres con la geografía y la variación genética; c) evaluar y en su caso confirmar la cohesión morfológica de *A. pungens* y su distinción del resto de especies del género.

1.2 MATERIAL Y MÉTODOS

1.2.1 Recolección del material vegetal

Se recolectaron 241 especímenes de 21 poblaciones naturales a lo largo de toda su área de distribución tratando de respetar una distancia de 10 m entre individuos (de diez a dieciocho individuos por población). Los pliegos se depositaron en el herbario del Real Jardín Botánico de Madrid (MA). Los datos corológicos de las recolecciones se adjuntan en la Tabla 1.1.

1.2.2 Selección de caracteres para el estudio morfométrico

Se examinaron un total de 19 variables morfológicas (15 cuantitativas continuas, Tabla 1.2, y 4 cualitativas, Tabla 1.3). Las medidas de las brácteas involucrales y del cáliz se representan detalladamente en la Fig. 1.2. Las variables se seleccionaron tomando como referencia trabajos de morfometría en el género *Armeria* (Moore y Yates, 1974; Papanicolaou y Kokkini, 1982; Lefèbvre y Vekemans, 1995; Nieto Feliner, 1996, 1997) o bien los tratamientos taxonómicos ya mencionados que discuten diferencias intraespecíficas en *A. pungens* (Bernis 1954, 1955; Briquet y Litardière, 1955; Arrigoni, 1970; Pinto da Silva *et al.*, 1972; Asensi *et al.*, 1993). Las medidas fueron efectuadas con un calibre digital Brown y Sharpe Plus (modelo 599-571-3) y una lupa binocular.

El registro del carácter indumento del cáliz resultó difícil debido a que los límites entre las tres estados establecidos para este carácter no eran claros. La confusión se produjo principalmente entre los estados subholótrico (indumento en los nervios y surcos intercostales) y holopleurótrico (sólo en los nervios medios), dado que en ocasiones el cáliz presenta hileras muy nutridas, aunque únicamente en las costillas. Los cálices hemipleurótricos fueron mucho más fáciles de identificar, y se registraron siempre que se apreció un menor revestimiento en los nervios secundarios.

1.2.3 Análisis morfológico de *A. pungens*

Como método exploratorio de ordenación de la variabilidad, se llevó a cabo un análisis de componentes principales (PCA) basado en una matriz de correlación de los caracteres cuantitativos usando el paquete SPSS (versión 14.0, SPSS Inc). Cuatro cocientes calculados a partir de las variables originales se incorporaron también al análisis. Para evitar problemas de dependencia de caracteres que puedan afectar a los resultados del PCA (Wiley, 1981), una de las medidas originales de cada conciente fue eliminada: longitud de la hoja, longitud de la bráctea involucral, ancho del margen de la bráctea involucral interna y longitud del cáliz. Se proyectaron las puntuaciones de los 241 individuos frente a los dos primeros componentes principales. La variabilidad de los caracteres cualitativos se examinó mediante el cálculo del porcentaje de individuos para cada estado de carácter dentro de cada área geográfica.

1.2.4 Análisis morfológico a nivel de género

La variabilidad morfométrica de *A. pungens* se reanalizó en el contexto de la morfología del género. Para ello se generó una nueva matriz incluyendo los 241 individuos de *A. pungens* más 55 especímenes de 24 especies de *Armeria* obtenidos de Fuertes Aguilar *et al.* (1999). Con la matriz ampliada se efectuó un nuevo PCA.

Figura 1.2 Detalle de los caracteres morfométricos medidos en *Armeria pungens* relativos a las brácteas involucrales y al cáliz. La numeración sigue la Tabla 1.2.

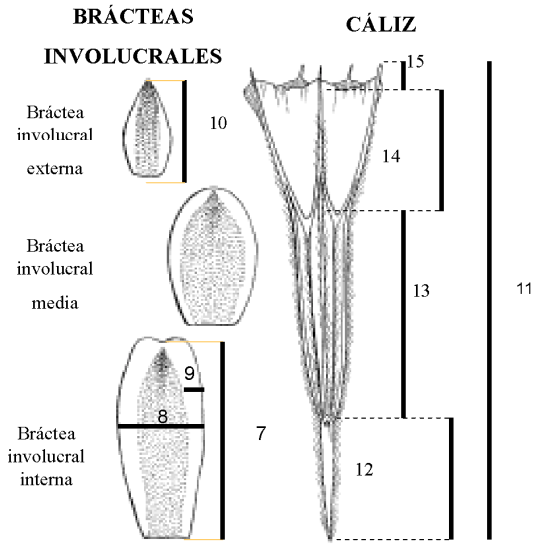


Tabla 1.1 Recolección de pliegos de *Armeria pungens* utilizados en el análisis multivariante de los caracteres morfológicos.

Código población	Localidad, origen y datos de recolección	número de colector (nº de individuos)
1- Cíes	PONTEVEDRA, Vigo, Illas Cíes, Illa de Monteagudo, dunas estabilizadas, 29TNG0875, 10 m,. 20-V-2000, *G. Nieto, J. Fuertes & X.R. García*	GN4271-4277, GN4294 y GN 4296-4298 (11)
1- Cíes	PONTEVEDRA, Vigo, Illas Cíes, Illa do Faro, junto al cementerio, dunas cubiertas con *Pteridium*, 29TNG0875, 10 m., 20-V-2000, *G. Nieto, J. Fuertes & X.R. García*	GN 4313-4319 (7)
2- Albufeira	ESTREMADURA, Setúbal, Lagoa de Albufeira, orilla sur de la laguna, arenales, 29SMC8462 5, 17-V-2002, *G.Nieto & J.Fuertes*	GN4496-4505 (10)
3- Sines	BAIXO ALENTEJO, Sines, cra. de Cercal a Sines, dunas, 29SNB1896, 17-V-2002, *G.Nieto & J.Fuertes*	GN4485-4495 (11)
4- Furnas	BAIXO ALENTEJO, Vilanova de Milfontes, praia Furnas, orilla sur del río Mira, dunas, 5 m., 29SBN1875, 17-V-2002, *G.Nieto & J.Fuertes*	GN4475-4484 (10)
5- Zambujeira	BAIXO ALENTEJO, Zambujeira do Mar, dunas fósiles en bordes de acantilado, 20 m., 29SBN1954, 17-V-2002, *G.Nieto & J.Fuertes*	GN4465-4474 (10)
6- Bordeira	ALGARVE, Carrapateira, praia de Bordeira, dunas estabilizadas, 29SBN0817, 10m., 17-V-2002, *G.Nieto & J.Fuertes*	GN 4455-4464 (10)
7- S. Vicente	ALGARVE, Vila do Bispo, Cabo Sao Vicente, acantilados y litosuelos calizos muy venteados, 29SNA0197, aprox. 40 m., 17-V-2002, *G.Nieto & J.Fuertes*	GN4445-4454 (10)
8- Garrrao	ALGARVE, Almansil, praia do Garrao (al W. de la de Ançao), dunas, 29SBN8303, 16-V-2002, *G.Nieto & J.Fuertes*	GN4433-4444 (10)
9- P. Umbría	HUELVA, Punta Umbría, Playa de Punta Umbría, dunas, 29SPB1878, 5 m., 16-V-2002, *G.Nieto & J.Fuertes*	GN4421-4432 (12)
10- C. Maneli	HUELVA, Almonte, cra. Matalascañas-Mazagón "cuesta Maneli", dunas, 29SQB0505, 5 m., 16-V-2002, *G.Nieto & J.Fuertes*	GN4406-4419 (14)

Tabla 1.1 (Continuación) Recolección de pliegos de *Armeria pungens* utilizados en el análisis multivariante de caracteres morfológicos.

11- Matalascañas	HUELVA, Almonte, Torre de la Higuera, dunas estabilizadas, 29SQA1699, dunas, 5 m., 16-V-2002, *G.Nieto & J.Fuertes*	GN4394-4405 (12)
12- Doñana	HUELVA, Almonte, P.N. de Doñana, pr. El Inglesillo, dunas, 29SQA3179, 12-VI-2002, *Sébastien Gatelier*	GN4590-4599 (10)
13- Trafalgar	CÁDIZ, Tarifa, Barbate, Cabo Trafalgar, dunas estabilizadas, 29SQA6608, 10 m., 15-V-2002, *G. Nieto & J. Fuertes*	GN4379-4390 (12)
14- Camarinal	CÁDIZ, Tarifa, Punta del Camarinal, dunas estabilizadas, 30STE4896, 10 m., 15-V-2002, *G. Nieto & J. Fuertes*	GN4362-4376 (15)
15- Piantarella	CORSE, Bonifacio, Playa Piantarella, barra de arena, con abundantes restos de Posidonia, 6-VI-2002, *G. Nieto & J. Fuertes*	GN4508-4518 (11)
16- P. Sperone	CORSE, Bonifacio, Petit Spérone, sistema dunar encajado en vaguada, 6-VI-2002, G. *Nieto & J. Fuertes*	GN4519-4528 (10)
17- I. Piana	CORSE, Bonifacio, Grand Spérone, 6-VI-2002, G. *Nieto & J. Fuertes*	
18- G. Sperone	CORSE, Bonifacio, Ile de Piana, 6-VI-2002, G. *Nieto, J. Fuertes*	
19- P. Liscia	SARDEGNA, Porto Pozzo, spiaggia di Porto Liscia, arenal estabilizado en lentiscar-sabinar, 7-VI-2002, *G.Nieto & J.Fuertes*	GN4530-4542 (13)
20- R. Maiore	SARDEGNA, S. de Sta. Teresa Gallura, spiaggia di Rena Maiore, dunas estabilizadas, 10 m., 7-VI-2002, *G.Nieto & J.Fuertes*	GN4544-4554 (11)
21- B. Mare	SARDEGNA, Badesi di Mare, dunas estabilizadas, junto a viñedo, 10 m., 7-VI-2002, *G.Nieto & J.Fuertes*	GN4556-4565 (10)
22- Stintino	SARDEGNA, Pr. Porto Torres, Stintino, al S del Stagno di Pilo playa muy alteradadunas estabilizadas, 10 m., 7-VI-2002, *G.Nieto & J.Fuertes*	GN4566-4575 (19)
23- C. Mannu	SARDEGNA, NW de Oristano, Capo Mannu entre cabo y Su Pallesu, dunas estabilizadas, 10 m., 7-VI-2002, *G.Nieto & J.Fuertes*	GN4578-4587 (10)

Tabla 1.2 Estadísticos descriptivos de los 15 caracteres morfométricos y cuatro ratios medidos en *Armeria pungens*. Se indican la media y la desviación típica (entre paréntesis) para el total de la especie y para las distintas áreas geográficas por separado, N = tamaño muestral,
(1) promedio de dos medidas por individuo.

	Variables morfométrícas	Total N=241		Islas Cíes N=28		Portugal N=51		Golfo de Cádiz N=72		Camarinal N=15		Córcega-Cerdeña N=75	
1	Longitud hoja (mm) (1)	46,11	(14,22)	51,14	(9,69)	42,35	(14,32)	40,98	(12,89)	65,69	(18,97)	48,97	(10,56)
2	Ancho hoja (mm) (1)	2,08	(0,61)	1,93	(0,51)	1,70	(0,59)	2,23	(0,54)	2,08	(0,63)	2,28	(0,57)
3	Longitud escapo (mm)	252,65	(74,78)	303,06	(79,63)	230,46	(83,17)	241,18	(68,12)	240,00	(44,52)	272,13	(68,36)
4	Diámetro base del escapo (mm)	17,9	(4,30)	17,00	(2,20)	15,70	(3,40)	18,60	(4,60)	19,50	(4,60)	19,00	(4,40)
5	Diámetro del involucro (mm)	29,33	(3,71)	29,01	(3,50)	28,06	(3,77)	29,78	(3,43)	25,73	(3,39)	30,73	(3,27)
6	Longitud vaina involucral (mm) (1)	26,30	(5,09)	26,55	(3,79)	24,82	(5,04)	28,22	(5,37)	22,16	(3,72)	26,43	(4,58)
7	Longitud bráctea involucral interna (mm) (1)	12,27	(1,52)	11,87	(1,39)	11,96	(1,55)	12,25	(1,49)	11,06	(1,40)	12,89	(1,34)
8	Ancho bráctea involucral interna (mm) (1)	6,83	(1,16)	8,21	(1,54)	6,14	(0,92)	6,61	(0,94)	6,53	(0,94)	7,33	(0,95)
9	Margen bráctea involucral interna (mm) (1)	1,57	(0,44)	1,93	(0,40)	1,34	(0,22)	1,52	(0,52)	1,85	(0,29)	1,66	(0,40)
10	Longitud bráctea involucral externa (mm) (1)	4,72	(1,00)	4,60	(1,05)	4,29	(0,95)	4,71	(0,90)	3,85	(0,63)	5,29	(0,90)
11	Longitud cáliz (mm) (1)	10,46	(1,03)	10,16	(0,59)	9,44	(1,02)	11,01	(0,78)	10,65	(1,24)	10,80	(0,61)
12	Longitud espolón del cáliz (mm) (1)	2,63	(0,47)	2,51	(0,34)	2,37	(0,55)	2,82	(0,39)	2,75	(0,46)	2,67	(0,40)
13	Longitud tubo del cáliz (mm) (1)	3,64	(0,45)	3,45	(0,27)	3,35	(0,31)	4,04	(0,35)	3,28	(0,38)	3,61	(0,41)
14	Longitud limbo del cáliz (mm) (1)	3,37	(0,52)	3,78	(0,31)	3,03	(0,41)	3,40	(0,47)	3,76	(0,57)	3,43	(0,51)
15	Longitud lóbulo cáliz (incluyendo arista) (mm) (1)	1,06	(0,33)	0,74	(0,18)	0,86	(0,23)	1,00	(0,22)	1,00	(0,26)	1,38	(0,28)
16	ratio ancho / longitud hoja	23,75	(9,33)	27,65	(6,94)	27,87	(12,59)	18,96	(6,15)	32,51	(8,74)	22,30	(5,86)
17	ratio diámetro involucro / longitud bráctea involucral interna	1,15	(0,23)	1,11	(0,15)	1,17	(0,24)	1,08	(0,20)	1,19	(0,21)	1,20	(0,24)
18	ratio doble margen / ancho bráctea involucral interna	0,39	(0,07)	0,39	(0,07)	0,36	(0,07)	0,39	(0,07)	0,35	(0,07)	0,41	(0,07)
19	ratio longitud bráctea involucral externa / bráctea involucral interna	0,46	(0,11)	0,48	(0,10)	0,44	(0,08)	0,46	(0,14)	0,57	(0,07)	0,45	(0,10)
20	ratio lóbulo cáliz / longitud cáliz	0,10	(0,03)	0,07	(0,02)	0,09	(0,02)	0,09	(0,02)	0,09	(0,02)	0,13	(0,03)

Tabla 1.3 Caracteres cualitativos examinados en *Armeria pungens*. El número y porcentaje (entre paréntesis) de individuos que presentan cada uno de los estados de carácter se indica para las distintas áreas geográficas. N- tamaño muestral.

	Islas Cíes N=18	Portugal N=61	Golfo de Cádiz N=72	Camarinal N=15	Córcega-Cerdeña N=75
Indumento hojas					
glabro			**63 (87.5%)**	1(6.7%)	6 (8%)
ciliado	18 (**100%**)	**51 (80.4%)**	9 (12.5%)	3 (20%)	**69 (92%)**
pubérulo				**11(73.3%)**	
pubescente		10 (19.6%)			
Indumento escapo					
glabro	18 (100%)	50 (82%)	72 (100%)	5 (33,3%)	75 (100%)
pubérulo		1 (1,6%)		10 (66,7%)	
pubescente		10 (16,4%)			
Brácteas y/o bractéolas espiculares					
presentes	18 (100%)	51 (80.4%)	55 (76.4%)	9 (60%)	46 (61.3%)
ausentes		10 (19.6%)	17 (23.6%)	6 (40%)	29 (38.7%)
Indumento cáliz					
subholotrico	7 (60.7%)	33 (45.1%)	63 (87.5%)	6 (40%)	50 (66.7%)
holopleurotrico	11 (39.3%)	21 (41.2%)	9 (12.5%)	5 (33.3%)	17 (22.7%)
hemopleurotrico		7 (13.7%)		4 (26.7%)	8 (10.7%)

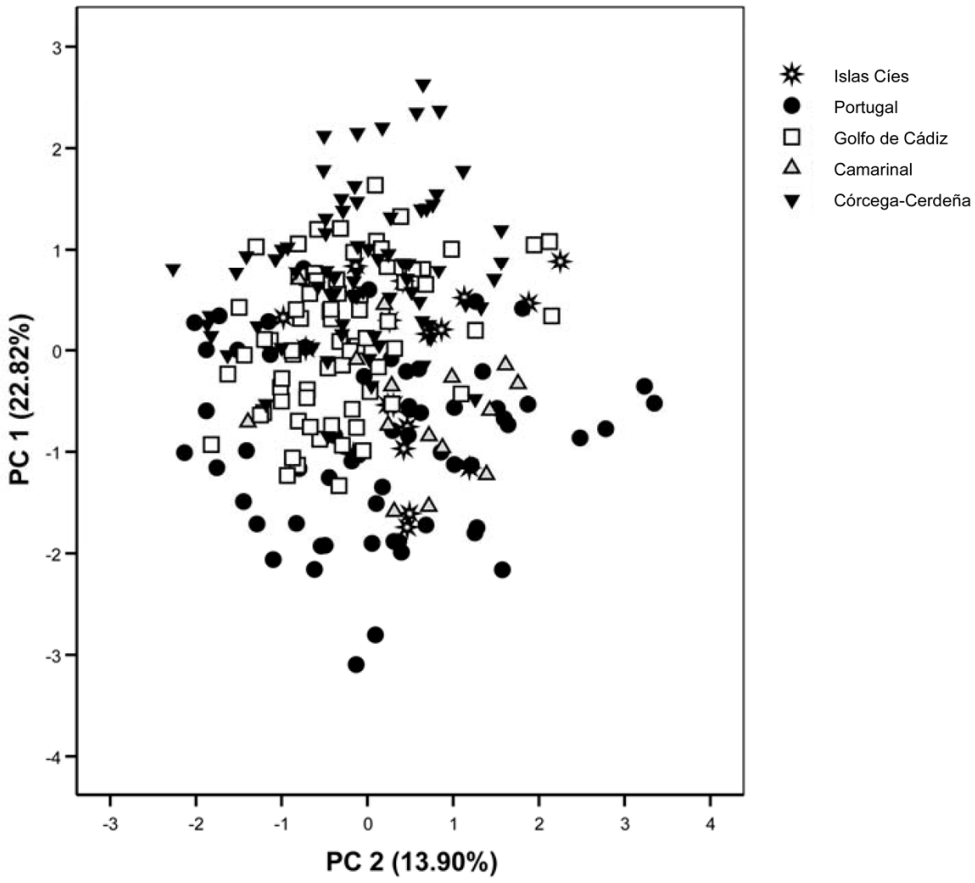
1.3 RESULTADOS

Los tres primeros componentes principales extraídos en el PCA para *A. pungens* explican un 49,44% de la varianza (Tabla 1.4). Buena parte de las variables originales muestran correlación positiva con el primer componente, PC 1, que puede, pues, interpretarse como indicador de tamaño general. El diagrama de dispersión de los 241 individuos frente a los dos primeros componentes principales (Fig. 1.3) no presenta ninguna discontinuidad aparente, lo que revela la ausencia de patrones geográficos en la variación morfológica de *A. pungens*. No obstante, las plantas de Córcega y Cerdeña tienden a presentar mayores tamaños, como indican sus mayores puntuaciones para PC 1, i.e., el Mediterráneo alberga los individuos de mayor tamaño de la especie.

Tabla 1.4 Análisis de componentes principales (PCA) de *Armeria pungens* en base a 15 caracteres morfométricos. Los factores de carga indican la correlación de cada variable con cada uno de los tres primeros componentes principales. El autovalor y el porcentaje de varianza indican la contribución de cada componente a la varianza total de los datos.

Variables morfométricas	PC1	PC2	PC3
Ancho hoja	0,59	-0,41	0,20
Longitud escapo	0,47	0,75	0,18
Diámetro base del escapo	0,59	0,41	0,20
Diámetro del involucro	0,62	-0,06	0,26
Ancho bráctea involucral interna	0,46	0,01	0,37
Longitud bráctea involucral externa	0,71	-0,03	-0,27
Longitud espolón del cáliz	0,42	0,12	0,42
Longitud tubo del cáliz	0,33	-0,27	0,28
Longitud limbo del cáliz	0,37	-0,08	0,52
Longitud lóbulo cáliz (incluyendo arista)	0,67	-0,02	-0,54
ratio ancho / longitud hoja	-0,22	0,85	-0,10
ratio diámetro involucro/ longitud bráctea involucral interna	-0,03	-0,59	0,07
ratio doble margen / ancho bráctea involucral interna	0,45	-0,07	-0,44
ratio longitud bráctea involucral externa/ bráctea involucral interna	-0,12	0,12	0,15
ratio lóbulo cáliz / longitud cáliz	0,52	-0,01	-0,70
autovalor	3,42	2,09	1,91
% varianza	22,82	13,90	12,72
% varianza acumulada	22,82	36,72	49,44

Figura 1.3 Análisis de componentes principales (PCA) de *Armeria pungens* en base a 15 caracteres morfométricos. Diagrama de dispersión de 241 individuos frente a los dos primeros componentes principales, que explican el 22.82% y 13.90% de la varianza, respectivamente.



Por el contrario, los caracteres cualitativos relativos al indumento de los órganos muestran un claro patrón geográfico (Tabla 1.4; Fig. 1.4). Dos poblaciones se desmarcan claramente de las demás: Bordeira (pop. 6), con escapos y hojas siempre pubescentes, y la población del límite meridional de la especie (Camarinal, pop. 14), con mayor incidencia de hojas y escapos pubérulos. Además, las poblaciones del Golfo de Cádiz (pops. 8 a 13) poseen hojas de glabras a subglabras, en contraste con las hojas ciliadas en el nervio medio dominantes en el resto del rango geográfico (Islas Cíes, Portugal y Córcega-Cerdeña: pops. 1, 2-7 y 15-23, respectivamente). La presencia de brácteas y bractéolas espiculares, siempre atrofiadas, no muestra ningún patrón geográfico.

Figura 1.4 Distribución del carácter indumento de las hojas en las distintas áreas de distribución de *Armeria pungens*.

En el PCA para el género *Armeria* (55 accesiones de *Armeria* sp. + 241 de *A. pungens*), los dos primeros componentes principales explican un 59,82% de la varianza (Tabla 1.5). El diagrama de dispersión frente a PC1 y PC2 muestra que *A. pungens*, a pesar incluir un número extremadamente superior de ejemplares, no se solapa en el espacio morfométrico con el resto de especies (Fig. 1.5).

Figura 1.5 Análisis de componentes principales (PCA) del género *Armeria* (55 especímenes de 24 especies; Fuertes Aguilar *et al.*, 1999) y *Armeria pungens* (241 individuos de este estudio) en base a 14 caracteres morfométricos. Diagrama de dispersión de 241 individuos frente a los dos primeros componentes principales, que explican el 41.99% y 17.82% de la varianza, respectivamente.

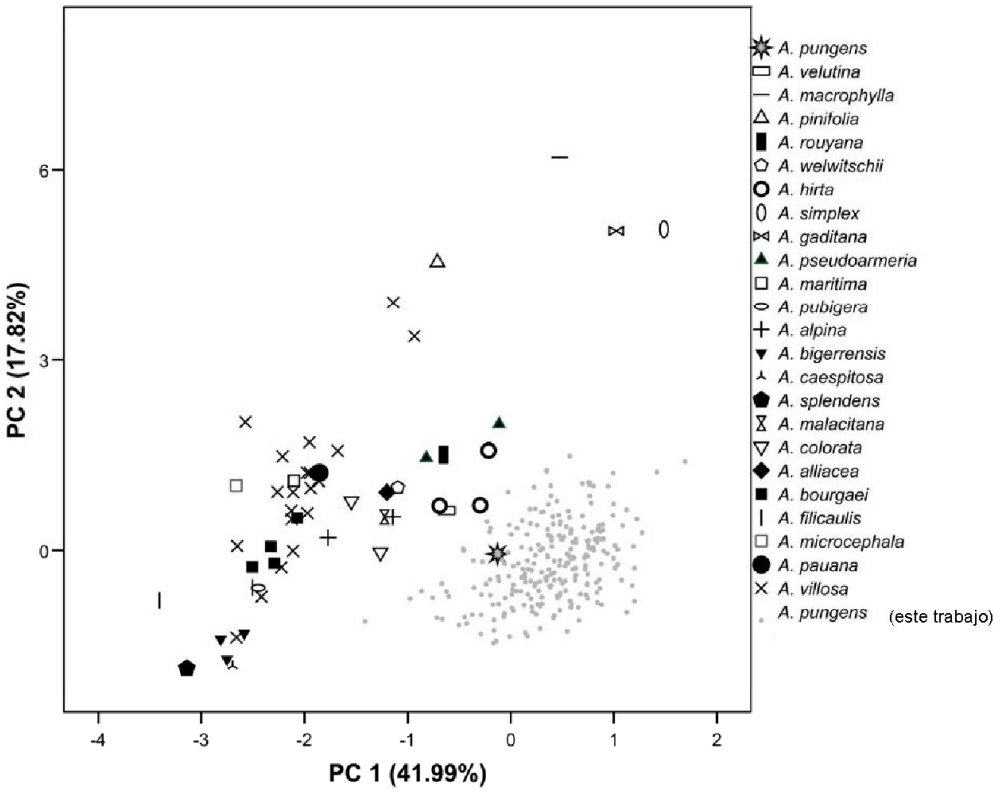


Tabla 1.5 Análisis de componentes principales (PCA) del género *Armeria* (55 especímenes de 24 especies; Fuertes Aguilar *et al.*, 1999) y *Armeria pungens* (241 individuos de este estudio) en base a 14 caracteres morfométricos. Los factores de carga indican la contribución de cada variable a cada uno de los tres primeros componentes principales.

Morphometric variables	PC1	PC2	PC3
Longitud hoja	0,08	0,89	-0,23
Longitud escapo	0,11	0,87	0,04
Diámetro base del escapo	0,57	0,54	0,19
Diámetro del involucro	0,87	0,13	0,20
Longitud bráctea involucral interna	0,78	0,27	0,19
Ancho bráctea involucral interna	0,74	-0,04	0,03
Longitud cáliz	0,84	0,14	0,10
Longitud espolón del cáliz	0,88	-0,11	-0,17
Longitud tubo del cáliz	0,80	-0,22	0,06
Longitud limbo del cáliz	0,83	-0,31	-0,06
ratio ancho / longitud hoja	0,01	0,42	-0,80
ratio diámetro involucro/ longitud bráctea involucral interna	0,33	0,12	0,32
ratio longitud bráctea involucral externa/ bráctea involucral interna	-0,51	0,27	0,43
ratio lóbulo cáliz / longitud cáliz	-0,68	0,37	0,36
autovalor	5,88	2,5	1,27
% varianza	41,99	17,82	9,08
% varianza acumulada	41,99	59,82	68,89

1.4 DISCUSIÓN

1.4.1 Variabilidad morfológica de *A. pungens* en el contexto de los tratamientos taxonómicos previos

Los resultados del estudio morfométrico a nivel de género, confirman la clara entidad morfológica de *Armeria pungens* respecto de las otras especies, incluidas las especies litorales y subitorales de la sección *Macrocentron* (*A. gadiatana, A. hirta, A. macrophylla, A. pinifolia, A. rouyana, A. velutina* y *A. welwitschii,*). Además, el análisis morfométrico para *A. pungens* corrobora la ausencia de diferencias significativas entre las distintas áreas de distribución, tal como han venido apuntando las revisiones taxonómicas de los últimos años. El detallado muestreo de este trabajo descarta, pues, definitivamente el poder diagnóstico de las brácteas o bractéolas espiculares y del ancho del margen escarioso de las brácteas involucrales internas.

Con respecto a los caracteres señalados en la literatura botánica más moderna como variables en *A. pungens*. (i) el patrón de hojas y escapos mayores en Cerdeña descrito por Arrigoni (1970) podría deberse simplemente al mayor porte general de los individuos corso-sardos puesto de manifiesto en el análisis multivariante; (ii) similar sesgo podría explicar también el patrón geográfico de disminución del lóbulo del cáliz desde Cerdeña hasta Galicia comentado por Bernis (1955) y Briquet y Litardière (1955).

Con respecto al tipo de indumento de las hojas, los individuos pubescentes de Bordeira (pop. 6) corresponden sin ninguna duda a la variedad *velutina* de Coutinho. (Pinto da Silva *et al.*, 1972). Sorprendentemente, la diferenciación de un linaje de hojas glabras en el Golfo de Cádiz *vs.* un linaje de hojas ciliadas en el resto del área de distribución (Islas Cíes, Portugal, Córcega y Cerdeña) no se discute en ninguno de los trabajos taxonómicos realizados hasta ahora. Tampoco se habían observado las plantas pubérulas de la población de Camarinal (pop. 14).

1.4.2 Valor del indumento de las hojas como marcador evolutivo

La morfología aporta escasa información acerca de la evolución de *A. pungens* dadas las escasas diferencias halladas entre los individuos de distintas áreas geográficas. Tan sólo el indumento de las hojas sigue una pauta de variación congruente con la geografía, lo que sugiere que es un reflejo de los procesos evolutivos que han tenido lugar en esta especie, es decir, que tiene un valor taxonómico a este nivel.

Los caracteres relativos a la pubescencia de distintos órganos parecen estar fijados genéticamente en *Armeria*, probablemente bajo control mono u oligogénico Así lo sugieren tanto experimentos de transplantes que examinan el indumento de cáliz y hojas (Lefèbvre, 1971), como experimentos con híbridos artificiales que investigan el indumento de escapos (Lefèbvre, 1976; Phillipp, 1974) y de hojas (Nieto Feliner, 1996). Por ello, no es raro que en varios trabajos anteriores de morfometría en *Armeria* se haya encontrado un patrón geográfico del indumento de los órganos a nivel intraspecífico, como es el caso de (i) los escapos de *A. maritima* en Dinamarca (Lefèbvre, 1969) o Estados Unidos (Lefèvbre y Vekemans, 1995), (ii) la pubescencia de las hojas asociada a la salinidad de los suelos en *A. maritima* (Lefèbvre, 1974) o (iii) la pubescencia de *A. welwitschii* en la parte central de su área de distribución (Cristina Tauleigne, comunicación personal).

En nuestro caso, la prueba del significado evolutivo del indumento de las hojas en *A. pungens*, es la perfecta correlación de este carácter con los linajes genéticos identificados en *A. pungens* con base en marcadores moleculares (ver capítulos 2, 3 y 4).

Según los datos genéticos, *A. pungens* presenta un área ancestral en el SW de Portugal desde la que se han colonizado las islas mediterráneas de Córcega y Cerdeña por dispersión a larga distancia. Todas estas poblaciones constituyen el linaje portugués-corso-sardo (pops. 2-7 y 15-23) y se caracterizan por tener hojas ciliadas en el nervio medio, exceptuando Bordeira (pop. 6; ver más abajo).

También a partir de las poblaciones portuguesas se produjo la colonización lineal hacia el sur del área contigua del Golfo de Cádiz (pops. 8-13), asociada a un fuerte efecto fundador y divergencia por deriva génica. La diferenciación de este linaje en el Golfo de Cádiz implicó un aumento en la frecuencia del carácter hojas glabras, aunque también se ha detectado la presencia minoritaria de individuos con hojas ciliadas.

La población pubérula en el límite meridional de *A. pungens* (Camarinal pop. 14) deriva, según sugieren los datos genéticos, de la hibridación reciente con *A. macrophylla.* Esta última es la única especie con hojas y escapos pubérulos de las que habitan la franja litoral y sublitoral del Golfo de Cádiz (*A. gadiatana*, *A. hirta*, *A. macrophylla* y *A. velutina*), lo que confirma su implicación en el evento de reticulación.

La población más septentrional, Islas Cíes (pop. 1), presenta hojas con la misma apariencia, indumento ciliado, que las del linaje portugués-corso-sardo de *A. pungens*. Esto está en consonancia con los marcadores de ADN nuclear (ITS, AFLP y *GapC*). Los marcadores cloroplásticos sugieren, por el contrario, que se ha producido hibridación con *A. welwitschii* y/o *A. berlengensis*. La congruencia de la morfología con los marcadores nucleares frente a la señal del ADN cloroplástico es típica de los fenómentos de captura cloroplástica por hibridación ancestral entre dos especies, seguida de introgresión hacia una de ellas (ver recopilación de 39 ejemplos en plantas en Rieseberg y Wendell, 1993; Arnold, 1997, Arnold *et al.*,1999).

Finalmente, en la población de Bordeira se ha fijado la pubescencia. Este carácter no parece haberse adquirido por hibridación, ya que ninguna de las especies del litoral Ibérico para las que se han descrito poblaciones pubescentes (*A. pubigera*, *A. berlengensis*, *A. velutina* y *A. welwitschii*) se encuentra en sus proximidades. No obstante, a la hora de detectar fenómenos de hibridación considerando los caracteres morfológicos, ha de tenerse en cuenta que en los tratamientos taxonómicos del género *Armeria* la variación intraspecífica del indumento de los órganos no se ha recogido exhaustivamente (Pinto da Silva, 1972; Guinochet y Vilmorin, 1973; Bianchini, 1982; Devesa, 1987; Nieto Feliner, 1990). Tampoco los marcadores moleculares (ITS, AFLP, cpDNA. *GapC*) sugieren que se haya producido hibridación, y ni siquiera revelan la diferenciación genética de esta población con respecto a las poblaciones vecinas.

1.5 REFERENCIAS

Arrigoni PV (1970) Contributto alla conoscenza delle Armerie sardo-corse. *Webbia* **25**, 137–182.

Arnold ML (1997) *Natural Hybridization and Evolution.* Oxford University Press, Oxford.

Arnold ML, Bulger MR, Burke JM, Hempel AL, Williams JH (1999) Natural hybridization: how low can you go and still be important? *Ecology* **80**, 371-381.

Asensi Marfil A, Díez Garretas B, Van der Maarel E (1993) Dry coastal ecosystems of Portugal. In: *Dry Coastal Ecosystems* (ed. Van der Maarel E). Elsevier, Amsterdam, London, New York, Tokyo.

Bernis F (1954) Revisión del género *Armeria* Willd. con especial referencia a los grupos ibéricos. *Anales del Instituto Botánico A. J. Cavanilles* **11**, 5–288.

Bernis F (1955) Revisión del género *Armeria* Willd. con especial referencia a los grupos ibéricos. Parte segunda (descriptiva de los grupos ibéricos). *Anales Instituto Botánico A. J. Cavanilles* **12**, 77–252.

Bianchini, F (1982) *Armeria*. In: *Flora de Italia*. (eds. Pignatti S.), pp. 294-301. Edagricole, Bologna.

Briquet J, De Litardiére R (1955) *Armeria*. In: *Prodrome de la Flore Corse*. Tome III. (eds. Briquet J, De Litardiére R), pp. 1-12. Éditions Paul Lechevalier. Paris.

Devesa, J A (1987) *Armeria*. In: *Flora Vascular de Andalucía Occidental* Vol. I. (eds. Valdés B, Talavera S, Fernández-Galiano E), pp. 297-302. Ketres, Barcelona.

Fuertes Aguilar J, Rossello JA, Nieto Feliner G (1999) Molecular evidence for the compilospecies model of reticulate evolution in *Armeria* (Plumbaginaceae). *Systematic Biology* **48**, 735-754.

Guinochet M, Vilmorin R (1973) *Flore de France*. Centre Nacional de la Recherche Scientifique, Paris.

Lefèbvre C (1969) Populations d'*Armeria* Willd. le long de la mer baltique et de la mer du nord. *Bulletin de la Société Royale de Botanique de Belgique* **102**, 293-303.

Lefèbvre C (1971) Premiers résultats sur les relations génétiques existant entre des populations d' *Ameria* d'Europe Occidentale et Centrale. *Bulletin de la Société royale de Botanique Belgique* **104**, 263.

Lefèbvre C (1974) Relations génétiques entre des populations d'*Armeria maritima* (Mill.) Willd. *Comptes Rendus Hebdomadaires des Séances de l'Académie des Sciences. Série D* **278**: 1573-1576.

Lefèbvre C (1976) Breeding System and Population Structure of *Armeria maritima* (Mill.) Willd. on a Zinc-Lead Mine. *New Phytologist* **77**, 187-192.

Lefèbvre C, Vekemans X (1995) A numerical taxonomic study of *Armeria maritima* (Plumbaginaceae) in North America and Greenland. *Canadian Journal of Botany* **73**, 1583–1595.

Moore DM, Yates B (1974) *Armeria* L. in South America. *Botaniska Noticer* **127**, 83-192.

Nieto Feliner G (1990) *Armeria*. In: *Flora Iberica: Plantas de la Península e Islas Baleares*. Vol. II. (eds. Castroviejo S, Laínz M, López González G, *et al.*), pp. 642-721. Real Jardín Botánico. CSIC, Madrid.

Nieto Feliner G (1996) Natural and experimental hybidization in *Armeria* (Plumbaginaceae): *A. villosa* subsp. *carratracensis*. *Plant Systematics and Evolution* **201**, 163-177.

Nieto Feliner G (1997) Natural and experimental hybridization in *Armeria* (Plumbaginaceae)*: Armeria salmantica. International Journal of Plant Sciences* **158** (5): 585-592.

Papanicolaou K, Kokkini S (1982) *Armeria* (Plumbaginaceae) in Greek mountains. *Willdenowia* **12**, 213-219.

Philipp M (1974) Morphological and genetical studies in the *Armeria maritima* aggregate. *Botanisk Tidsskrift* **69**, 40–51.

Pinto da Silva R (1972) *Armeria* Willd. In: *Flora Europaea*. Vol. 3. (eds. Tutin TG, Heywood VH, Burges NA, *et al.*). Cambridge University Press, Cambridge, New York.

Rieseberg LH, Wendel JF (1993) Introgression and its consequences in plants. In: *Hybrid zones and the evolutionary process* (ed. Harrison R), pp. 70-109. Oxford University Press, New York.

Wiley EO (1981) *Phylogenetics: the theory and practice of phylogenetic systematics*. John Wiley and Sons, New York.

2. ECOLOGY MATTERS: ATLANTIC-MEDITERRANEAN DISJUNCTION IN THE SAND-DUNE SHRUB *Armeria pungens*

2.1 INTRODUCTION

The Mediterranean flora has high levels of plant diversity and endemism (80% of European plant endemics are Mediterranean) combined with a striking variety of species distribution patterns including frequent disjunct distributions (Blondel and Aronson, 1999; Médail and Quézel, 1999; Thompson, 2005). The historical interpretation of such diverse patterns is often difficult, since they reflect the overlapping effects of several processes that operated at different spatial and temporal scales. From a spatial perspective, the Mediterranean landscapes provide high habitat variability (Blondel and Aronson, 1999). At a temporal scale, different palaeoclimatic and geological processes have played a role in the evolution of Mediterranean plant communities. Most often invoked are the isolation of microplates resulting from Tertiary tectonic movements, the Messinian salinity crises at the Miocene-Pliocene boundary, 5.96-5.33 million years (Myr) ago, (Hsü *et al.*, 1977; Krijgsman *et al.*, 1999) the establishment of a Mediterranean climate type at the Plio-Pleistocene boundary, 3.2-2.8 Myr ago (Suc, 1984), and sea level changes associated with Pleistocene glaciations (Hewitt, 2000). The negative impacts of humans on certain plant communities during the last several thousand years may have been a further relevant factor (Martínez and Montero, 2004).

Systems that simplify the spatial component, such as coastal species with linear distribution ranges, allow identifying influential changes in a more clear way. Phylogeographic studies of some of these systems stress the importance of ecological requirements of each species modulated by Pleistocene climatic oscillations to explain differences in their genetic structures (Clausing *et al.*, 2000; Kadereit *et al.*, 2005). Other linear systems are best interpreted in terms of the geological history of their habitats (e.g., *Hypochaeris salzmanianna*, in Tremetsberger *et al.*, 2004). Many more comparative phylogeographic studies of plants in the Mediterranean are needed to clarify the relative contribution of historical vs. ecological factors in shaping present-day species distributions.

The present phylogeographic study examines the coastal plant species *Armeria pungens*, which has a linear geographic distribution with two significant disjunctions (Fig. 2.1.B). *Armeria*

pungens is a perennial shrub that grows on sandy maritime dunes (exceptionally on sandy soils over limestone maritime cliffs) with a main geographical distribution along a 500 km coastal stripe in SW Iberia, from the mouth of the Tagus river to the Gibraltar Strait. It also occurs on two continental archipelagos: in the Atlantic, on the Cíes Islands (offshore Galician coast, Northern Spain), and in the Mediterranean, on S Corsica and N Sardinia).

Taxonomic boundaries within *Armeria* are sometimes difficult to establish due to frequent hybridisation, even between distantly related species (Bernis, 1954; Nieto Feliner, 1990, Nieto Feliner *et al.*, 2001; Tauleigne-Gomes and Lefèbvre, 2005). However, this is not the case for *A. pungens*. All taxonomic treatments of this genus have recognized *A. pungens* as a distinct species with a disjunct distribution pattern (Bernis, 1955; Arrigoni, 1970; Pinto da Silva, 1972; Nieto Feliner, 1990). Thus, the fact that the Iberian and Corsican populations were initially described as separate species was only due to the circumstance that they were discovered and described almost simultaneously in the beginning of the XIXth century (Bernis, 1955; Arrigoni, 1970).

An interesting feature of our study system is the different geologic histories of the two disjunct areas, which opens several a priori possibilities for the occurrence of *A. pungens*. While Cíes archipelago was recently isolated from the mainland coasts after the last glaciations (Luaces and Toscano, 1998) and is only 2.5 km apart, the isolation of Corsica and Sardinia from Eastern Spain and Southern France dates back to the Tertiary, (c.16 Myr ago, Krijgsman, 2002). Land bridges with N Italy have been proposed during the Messinian salinity crisis and during the glaciations, although the latter are considered very unlikely by some authors (Gamisans, 1991). The strong affinities of Corso-Sardinian floras with other territories, mainly with the Balearic islands but also with Provence, Alps, Pyrenees and Iberian Peninsula, have traditionally been explained by the position of the Corso-Sardinian plate before the Miocene (Briquet, 1910; Cardona and Contandriopoulos, 1979; Gamisans, 1991). In contrast, the relative contribution of long-distance dispersal to the flora of Corsica and Sardinia has been insufficiently addressed in classic Mediterranean biogeography. Recent molecular phylogenies suggest that dispersal has been important in shaping contemporary distribution patterns in many plant groups (Winkworth *et al.*, 1999; Brochmann *et al.*, 2003; Schönswetter *et al.*, 2004; de Queiroz, 2005). The presence of early-divergent lineages in Corsica and Sardinia has been confirmed with molecular data for some plant

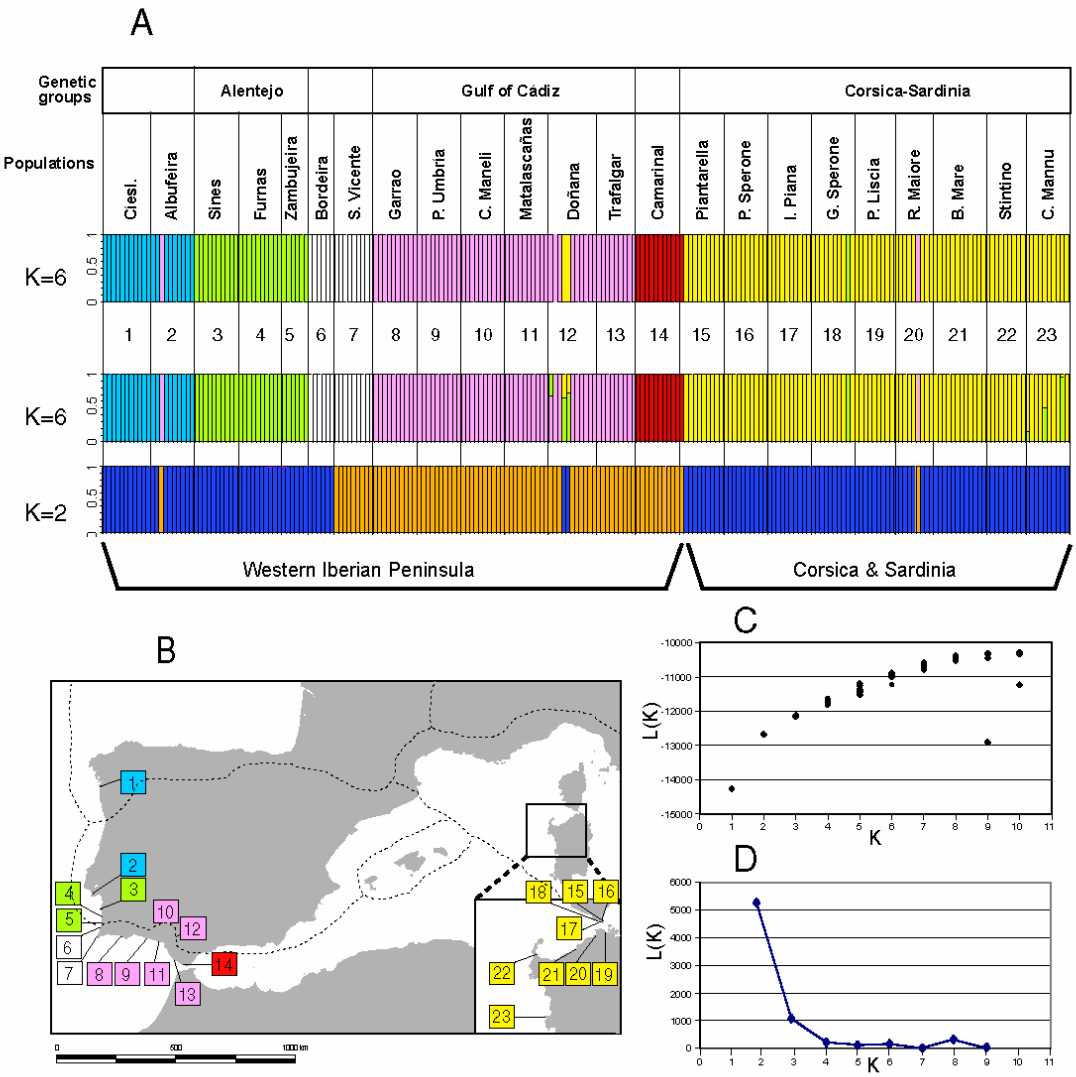
groups [e.g. *Doronicum corsicum* (Álvarez Fernández *et al.*, 2001)] but not for others [e.g. *Erodium maritimum* and *E. corsicum* (Fiz *et al.*, 2006)].

The well-defined taxonomic status of *A. pungens*, its linear geographical distribution, together with the well known geological history of the islands where it occurs provides a simple model to test scenarios of vicariance (old presence favoured by land connections) versus long-distance dispersal (recent arrival). For this purpose, we assessed the current genetic variation across *A. pungens* populations and examined them in the context of the extant distribution patterns and geologic history.

Predictive species–climate modelling has become very popular in recent years as a powerful tool to assess biodiversity patterns, species distribution, and increasingly the potential impacts of climate change (Huntley *et al.*, 1995; Guisan and Thuiller, 2005). Bioclimatic envelope models (BEM), describing the relationship between a species' observed distribution and climate, are supported on the niche theory of Hutchinson (Luoto *et al.*, 2005). Climate envelopes based on empirical correlation between species' distributions and selected climate variables allow characterising eco-physiological conditions in particular habitats without experimental analyses, unpractical in natural populations (Bakkenes *et al.*, 2002). In this work we use this technique as a way of comparing genetic and ecological data from *A. pungens* in the West-Mediterranean. This combined approach has been applied from a phylogenetic perspective to assess the role of ecology on speciation within several groups (Peterson, 1999; Graham, 2004; Yesson and Culham, 2006).

Specifically, the following issues were addressed: (i) Are disjunct populations on continental islands in the Atlantic and the Mediterranean the result of past fragmentations involving tectonic landmass movements or are they the result of long-distance dispersal? (ii) If due to long-distance dispersal, which is the main source area? To address these questions we needed to determine the genetic structure of *A. pungens*, i.e., to recognize genetic groups of populations and clarify the similarities among them, considering the distribution of the genetic variation within and among populations. (iii) Finally, we assessed the contribution of bioclimatic factors to the current distribution of *A. pungens.* For this purpose, using BEM tools, we modelised the climatic conditions of the *A. pungens* present range, as well as those of their genetically characterized groups of populations, to explore their potential distribution habitats.

Figure 2.1 Genetic structure of *A. pungens* inferred by bayesian clustering of AFLP data. **A**: Assignment of 221 individuals into K genetically distinguishable groups. Each individual is represented by a vertical bar coloured according to the assigned group(s). The 23 populations are identified by name and number following Table 2.1. The upper band represents the highest probability partition yielded by BAPS. The middle and lower bands show the most stable and likely assignments estimated by STRUCTURE (at K=2 and K=6) (see text). **B**: Geographic location of sampled populations of *A. pungens*, numbered as in Table 2.1 and colour-coded following bayesian clustering results for 6 groups in Fig. 2.1.A. Doted lines mark approximate boundaries of floristic provinces in Takhtajan (1986). **C**: Log probability of data $L(K)$ as a function of K for 10 STRUCTURE runs at K=1-6 or 5 runs at K= 7-10. **D**: Rate of change in the probability between successive runs, ΔK, as a function of K (see Evanno *et al.*, 2005).



2.2 MATERIALS AND METHODS

2.2.1 The species

Armeria pungens belongs to section *Macrocentron* Boiss., a small but morphologically well differentiated group with long-spurred calyces represented by ten species in the Iberian Peninsula (Nieto Feliner, 1990), two in North Africa (Nieto Feliner, 2002) and possibly one in the Eastern Mediterranean (Pinto da Silva, 1972). In Corsica and Sardinia *A. pungens* is the only representative of the section (Arrigoni, 1970). Chromosome counts on specimens from all main geographical areas have always revealed a diploid number, 2n=18, (Arrigoni, 1970; Lago and Castroviejo, 1993 and references therein), as reported in most counts for the genus.

Although reproductive biology of *A. pungens* has not been studied in detail, it is well documented that, with just some exceptions mainly in the Artic and North and South America (Baker, 1966; Moore and Yates, 1974), almost all species of *Armeria* are obligate outcrossers, as determined by an incompatibility system that involves a pollen-stigma dimorphism (Baker, 1966). Pollination by unspecific insects has been reported for the close relative *A. velutina* (Herrera, 1988) as well as for the widespread *A. maritima* (Eisikowitch and Woodell, 1975; Woodell and Dale, 1993). Wind pollination has also been suggested (Woodell and Dale, 1993) but it seems unlikely to be effective due to the low number of pollen grains per anther (Tauleigne-Gomes and Lefèbvre, 2005; G. Nieto Feliner, unpublished data). Parachute-like calyces containing the fruit suggest wind dispersal, although recorded distances are low in *A. maritima* (Philipp *et al.*, 1992), and animal dispersal may also occur facilitated by stiff hairs of fruiting calyces. *Armeria* fruits lack any special mechanism for water transport such as corky tissues.

2.2.2 Sampling, DNA extraction

Leaves and ripe fruits were collected in twenty-three populations of *A. pungens* spanning the whole species distribution (Table 2.1). Where possible, a distance of 10 metres between individuals was kept. Voucher specimens of all sampled individuals were collected and deposited in the herbarium of the Royal Botanical Garden of Madrid (MA). Leaves were dried and stored at –80ºC, whereas fruits were germinated with giberelic acid after a cold treatment of one month. Seedlings were cultivated in a greenhouse in the Botanical Garden of Madrid from 2002 to 2004. Most AFLP fingerprints were obtained from young fresh leaves grown from seeds belonging to separate

individuals in the field. When germination failed, frozen dried leaves were used instead. An average of ten individuals per population were sampled (Table 2.1). Genomic DNA was extracted from dried or fresh leaves using DNeasy Plant Minikit (Quiagen, Valencia, California, USA) according to the manufacturer´s instructions. The approximate quantity and quality of the isolated DNA was determined by 1.5% TAE (Tris-Acetate-EDTA)-agarose electrophoresis and ethidium bromide staining. Negative controls were performed to monitor contamination.

Table 2.1 Collection data of 23 populations of *Armeria pungens* included in the AFLP study. N=sampling size (approximately ten individuals per population were included) . The geographical location of each population is represented in Figure 2.1.B. Coordinates of each population were recorded as presence data for the climatic model.

Population code	Site location, collector (year)	Coordinates latitude/longitude	N
1- Cíes	Spain, Pontevedra, Illas Cíes, RPP, IMA, IMF, LMR, MSA and PGM (2003)	42.2245 / -8.89666	11
2- Albufeira	Portugal, Estremadura, Lagoa de Albufeira, GNF and JFA (2002) and RPP, AC and PEG (2003)	38.5111 / -9.17712	10
3- Sines	Portugal, Baixo Alentejo, Sines-Cercal, RPP, AC and PEG (2003)	37.9160 / -8.78885	10
4- Furnas	Portugal, Baixo Alentejo, Vilanova de Milfontes, Praia das Furnas, RPP, AC and PEG (2003)	37.7266 / -8.78937	10
5- Zambujeira	Portugal, Baixo Alentejo, Zambujeira do Mar, GNF and JFA (2002) and RPP, AC and PEG (2003)	37.5372 / -8.77858	6
6- Bordeira	Portugal, Algarve, Praia do Bordeira, GNF and JFA (2002)	37.2038 / -8.90348	6
7- S. Vicente	Portugal, Algarve, Cabo de São Vicente, GNF and JFA (2002) and RPP, AC and PEG (2003)	37.0235 / -8.98239	9
8- Garrrao	Portugal, Algarve, Praia de Garrao, GNF and JFA (2002) and RPP, AC and PEG (2003)	37.0739 /-8.07122	10
9- P. Umbría	Spain, Huelva, Punta Umbría, Playa de Punta Umbría, GNF and JFA (2002)	37.1956 / -6.98823	10
10- C. Maneli	Spain, Huelva, Mazagón-Matalascañas, "cuesta Maneli", GNF and JFA (2002)	37.0729 / -6.68781	10
11- Matalascañas	Spain, Huelva, Torre de la Higuera, Playa de Torre la Higuera, GNF and JFA (2002) and RPP, AC and PEG (2003)	37.0164 / -6.56591	10
12- Doñana	Spain, Huelva, Doñana National Park, pr. El Inglesillo, SG (2002)	36.9678 / -6.39901	11
13- Trafalgar	Spain, Cadiz, Cabo de Trafalgar, GNF and JFA (2002) and RPP, AC and PEG (2003)	36.1839 / -6.03606	9
14- Camarinal	Spain, Cadiz, Punta Camarinal, GNF and JFA (2002)	36.0798 / -5.79252	10
15- Piantarella	France, Corsica, Bonifacio, Piantarella, GNF and JFA (2002)	41.3761 / 9.22254	10
16- P. Sperone	France, Corsica, Bonifacio, Petit Sperone, CBNMP(2002)	41.3691 / 9.22269	10
17- I. Piana	France, Corsica, Bonifacio, île Piana, CBNMP (2002)	41.3735 / 9.23012	10
18- G. Sperone	France, Corsica, Bonifacio, Grand Sperone, CBNMP (2002)	41.3682 / 9.21417	10
19- P. Liscia	Italy, Sardinia, Porto Pozzo, Porto Liscia, GNF and JFA (2002)	41.2000 / 9.31667	9
20- R. Maiore	Italy, Sardinia, S of Santa. Teresa di Gallura, Spiaggia di Rena Maiore, GNF and JFA (2002)	41.1441 / 9.13725	9
21- B. Mare	Italy, Sardinia, Badesi Mare, GNF and JFA (2002)	40.9667 / 8.88333	12
22- Stintino	Italy, Sardinia, Pr. Porto Torres, Stintino, S Stagno di Pilo, GNF and JFA (2002)	40.9372 / 8.22750	9
23- C. Mannu	Italy, Sardinia, NW of Oristano, Capo Mannu, between the cape and Su pallosu, GNF and JFA (2002)	40.0750 / 8.37129	10

Abbreviations of collectors: AC, A. Costa; CBNMP, Conservatoire Botanique National Mediterranéen de Porquerolles; GNF, G. Nieto Feliner; IMA, I. Martínez Arcos; IMF, I. Martínez Fernández; JFA, J. Fuertes Aguilar; LMR, L. Muriel Ríos; MSA; M. Souto Alonso; PEG, P. Escobar García; PGM, P. García Meijide; RPP, R. Piñeiro Portela; SG, S. Gatelier.

Table 2.2 Presence data of *A. pungens*, based on herbarium specimens, used to build the climatic model in addition to the populations sampled for AFLP (Table 2.1). All voucher specimens are from the Royal Botanical Garden of Madrid (MA).

	Site location, collector, date (voucher)	Coordinates latitude/longitude
a	Portugal: Baixo Alentejo, Peninsula de Troia, pr. Malha da costa, Malato-Beliz *et al.*, 27.6.1971 (MA 306078)	38.4300 / -8.83323
b	Portugal: Baixo Alentejo, Praia de Melides, M. Henna and L. Loidi, 19.5.1990 (MA 486568)	38.1324 / -8.78824
c	Portugal: Baixo Alentejo, Sines, pr. lighthouse, E. Monasterio, F. Muñoz Garmendia and J. Pedrol, 10.4.1988 (MA 448985)	37.9612 / -8.87979
d	Portugal: Baixo Alentejo, Cabo Sardao, Malato Beliz and J.A. Guerra, 17.6.1978 (MA 238063)	37.6004 / -8.82370
e	Portugal: Baixo Alentejo, Praia do Carvalhal, Malato Beliz and J.A. Guerra, 16.5.1984 (MA 421137)	37.5012 / -8.78999
f	Portugal: Algarve, Odeceixe, F. Bernis, 24.4.1949 (MA 394446)	37.4381 / -8.80147
g	Portugal: Algarve, Praia de Esteveira, A. Herrero, 27.3.1991 (MA 649267)	37.4291 / -8.81279
h	Portugal: Algarve, Praia de Monte Clérigo, F. Bellot, 19.4.1968 (MA 498827)	37.3930 / -8.82418
i	Portugal: Algarve, Ponta de Sagres, V.J. Arán and M.J. Tohá, 28.3.1999 (MA 643449)	37.0235 / -8.99363
j	Portugal: Algarve, Praia de Armançao de Pêra, Praia grande, L. Medina, S. Nisa and M. Pardo, 7.6.2001 (MA 691371)	37.0937 / -8.32977
k	Portugal: Algarve, Praia do Ançao, Faro-Ferreiras, A. Moura, 28.5.1986 (MA 394453)	37.0285 / -8.03805
l	Spain: Huelva, La Rábida, Pinar del Palo, 29.4.1949, F. Bernis (MA 394442)	37.2127 / -6.93144
m	Spain: Huelva, Mazagón, Fernández Casas, 15.8.1969 (MA 413843)	37.0918 / -6.73224
n	Spain: Huelva, La Barra, C. Vicioso, 21.4.1943 (MA 145356)	37.1745 / -7.37170

2.2.3 AFLP protocols: *Eco*RI/*Mse*I and *Kpn*I/*Mse*I

In order to cover as many genomic regions as possible, two different restriction enzyme pairs were combined: *Kpn*I/*Mse*I and *Eco*RI/*Mse*I. The restriction with *Kpn*I, insensitive to methylation, in addition to the extensively used *Eco*RI, only affected by overlapping methylation (5-methylcytosine in CpG formed with the adjacent sequence) reduces polymorphism due to different levels of methylation (Roberts *et al.*, 2005). By this strategy, we minimised effects associated to gene expression (Reyna-López *et al.*, 1997; Xiong *et al.*, 1999; Cervera *et al.*, 2002).

An initial screening of 37 combinations of selective primers on four individuals from different geographic areas was performed. Three combinations were selected that yielded clear and evenly distributed bands: (6-FAM)*Eco*RI+acc/*Mse*I+cacc, (6-FAM)*Eco*RI+acg-*Mse*I+ctac, (6-FAM)*Kpn*I+atc/*Mse*I+cag (MWG-Biotech AG). To prevent mismatch amplifications with the *Mse*I primer with four selective nucleotides (Vos *et al.*, 1995), preselective primers with two selective bases were used.

Restriction and ligation of genomic DNA with *Eco*RI and *Mse*I were performed according to Gaudeul *et al.* (2000) with few modifications. Double digestion with *Kpn*I and *Mse*I was performed independently from the ligation. The restriction reaction was incubated at 37ºC for 2.5 hours and consisted of 0.2 µl NE1 buffer , 0.1 µl 1mg/ml BSA, 1U *Mse*I, 5U *Kpn*I (New England BioLabs) and 10 µl template DNA (final volume 20 µl). Ligation, also incubated for 2.5 hours at 37ºC, comprised 2.0 µl 10x T4 DNA ligase buffer, 1.0 µl BSA (1mg/ml), 2.0 µl 10 µM *Mse*I adapter pair, 2.0 µl *Kpn*I adapter pair, 2U T4 DNA ligase, and 15 µl restriction reaction (final volume 25 µl). *Kpn*I adapter sequences 5´-CTCGTAGACTGCGTACAGTAC-3´ and 5´-TGTACGCAGTCTAC-3´, as well as the adapter-matching primers 5´-GACTGCGTACAGTACCA and 5´(6-FAM)-GACTGCGTACAGTACCATC-3´were designed following the rules for good primer design (Dieffenbach *et al.*, 1993). Restriction-ligation products were diluted 10 times in purified H_2O. Pre-amplification and selective amplification were performed as described in Gaudeul *et al.* (2000) with small modifications. Restriction-ligation and amplification reactions were incubated in a thermocycler GeneAmp PCR System 9700 (PE Biosystems, Foster City, California, USA). The fragments were separated using an ABI 3700 sequencer (Applied Biosystems) using 1 µl of PCR product and GeneScan-500 ROX size standard.

A reproducibility test was performed by reextracting DNA from one individual per population, in 22 of the 23 populations, and repeating the whole AFLP procedure. The error rate was calculated for every primer combination as the number of phenotypic differences related to the total number of phenotypic comparisons, and subsequently averaged over the three combinations. This rate was further used to evaluate the quality of the fingerprints.

2.2.4 AFLP analysis

Amplified bands were aligned with the internal size standard using the ABI Prism GeneScan Analysis Software v.3.1 (Applied Biosystems). Subsequently, fragments of each primer combination were scored automatically with Genographer v. 1.6.0 (Montanta State University, http://hordeum.oscs.montana.edu/genographer/) either as present (1) or absent (0), and manually corrected. Peaks were recorded in a range from 50 to 500 bp.

2.2.4.1 *Genetic distance- degree of similarity between individuals and populations*

To distinguish genetically similar groups of individuals in our AFLP data set and clarify the similarity among them, first a descriptive comparison of phenotypes was made by constructing a pair-wise similarity matrix between all individuals using Dice's coefficient and subjecting it to a principal coordinates analysis (PCoA). Secondly, a pair-wise distance matrix was computed between populations using net nucleotide differences between populations (Nei and Li, 1979, equation 25), as implemented in ARLEQUIN 3.0 ('Nei's net average number of differences between populations'; Excoffier *et al.*, 2005), and visualised through a neighbour-joining (NJ) tree. The PCoA and the NJ analyses were performed with NTSYS pc 2.0 (Rohlf, 1998).

2.2.4.2 *Genetic differentiation- the distribution of genetic variation among populations and regions*

We also used a MCMC (Markov Chain Monte Carlo) Bayesian clustering method to identify genetically similar groups, i.e., having distinctive allele frequencies, using STRUCTURE version 2.0 (Pritchard *et al.*, 2000). This approach assumes Hardy-Weinberg and linkage equilibrium within groups. Following user's manual recommendations for dominant markers a missing value was added to each marker and the ancestry model of no admixture was chosen. For the allele frequency model, we set

the default option of correlated allele frequencies (Falush *et al.*, 2003) as advised by the manual when they are expected to be similar in the different groups. Since clustering of individuals was different across independent runs and even very long chains of 10^6 did not stabilize the results, we run 10 simulations for each number of groups, K, from K=1 to 6 and 5 replicates for K=7 to 10, using a burnin period of 10^5 and runs lengths of 10^6. We applied two criteria to choose the best value of K in our data set: the estimated posterior log probability of the data, L(K), and the stability of assignment patterns across runs. Since L(K) continued to grow slightly with increasing values of K, the criterion of selecting the K that maximises the probability of the data was difficult to apply. We therefore calculated another *ad hoc* quantity based on the rate of change in the probability between successive K, ΔK, as proposed by Evanno *et al.* (2005). The results were compared with those provided by another bayesian clustering method, BAPS 3.2 (Corander *et al.*, 2006), which uses stochastic optimisation instead of MCMC to find the optimal partition (with highest estimated probability). Given the faster performance of this algorithm, the simulation was started from K=2 to K=23 as the maximum number of diverged groups, with three replicates for each K.

In order to quantify the genetic differentiation with an alternative method not assuming Hardy-Weinberg equilibrium or independence of markers, the proportion of molecular variance within and among populations, AMOVA, was analysed with ARLEQUIN 3.0. AMOVA based on pair-wise distances between individuals were carried out for the whole data set considering three levels, among regions, among populations within regions and within populations. For subsets of the data, two levels, within and among populations, were considered.

2.2.4.3 *Genetic diversity within populations*

The amount of genetic variation was measured for each population: (i) in terms of allele richness, by the direct observation of percentage of polymorphic loci (P), i.e., variable across the data set; (ii) in terms of similarity, with Shannon´s index (H_{sh}) (Shannon, 1948); and (iii) in terms of allele frequencies, assuming Hardy-Weinberg equilibrium for each locus, using Nei´s unbiased gene diversity H_s (Nei, 1978). Since every AFLP band represents a different location in the genome, both Shannon´s and Nei´s indices were calculated for every locus and then averaged over all loci as implemented in POPGENE version 1.32 (Yeh and Boyle, 1997) and TFPGA version 1.3

(Miller, 1997), respectively. Shannon index was standardised to the lowest sampling (6 individuals).

2.2.5 Bioclimatic envelope model

The potential distribution of *A. pungens* according to climatic factors was determined through a GIS modelling analysis undertaken with Idrisi Kilimanjaro software (Clark Labs, Worcester, USA) and DIVA-GIS (Hijmans *et al.*, 2001). The BIOCLIM algorithm was used to build the distribution model (Nix, 1986; Busby, 1991; McMahon *et al.* 1995). This algorithm first characterises the environmental conditions of the actual distribution of the species and then identifies additional sites that fall within the already defined environmental hyperspace (Barry and Elith, 2006). We considered as presence data the geographic location of our own sampling sites (Table 2.1), which represent most of the known populations of the species. These geographic data were supplemented with those obtained from herbarium specimens (Table 2.2). Altogether, it is considered that presence data used are representative for characterising the environmental conditions of a species with a linear distribution range and confined to a restricted ecological habitat.

Only climatic variables were used to construct the distribution model because we were not only interested in characterising the environmental conditions of *A. pungens* with respect to other species but also to explore if geographically cohesive ecological subsets could be identified corresponding to genetic subgroups. For this purpose, environmental variables such as substrate (sand), elevation above sea level and the sea influence through salt accumulation are largely invariant across the distribution range of the species. Among the possible climatologic variables, we chose three with a more likely direct physiological significance on seed germination as well as plant growth and survival: monthly minimum temperature (Tmin), monthly maximum (Tmax) temperature y and monthly precipitation (Médail and Quézel, 2003). Monthly minimum temperature may represent a threshold for over winter survival and for seed germination. Monthly maximum temperature is critical during Mediterranean summers when it is associated to drought. Monthly precipitation was chosen as an estimate of moisture availability for both seed germination and plant growth. These three environmental predictors were obtained from WorldClim data set (www.worldclim.org), which contains records from 1960 to 1990, with 30 seconds

resolution (Hijmans *et al.*, 2005). Since each of these variables was considered for 12 months, the models were based on 36 data layers. For each of the 36 variables, we considered the ranges between (1) maximum and minimum values; (2) mean ± 1 standard deviation; (3) 95% of the confidence interval; (4) mean ± 1.96 standard deviation, and combined them either in an excluding way (AND) or in an inclusive way (OR). We chose to use the range between maximum and minimum values and the 'AND' combination, which of the ones recovering the actual distribution area was the most selective of the four criteria. Thus, localities were selected as matching the climatic conditions of the presence data only when their values for each of the 36 variables fell within the ranges determined by the maximum and minimum values on the basis of the presence data.

In accordance with the two purposes pursued here with the species distribution modelling, two approaches were followed. To explore whether the current distribution of *A. pungens* covers most of its environmental envelope or, alternatively, suitable sites are found elsewhere, all presence data were used to construct the model. To explore if geographically cohesive climatic profiles could be identified within the actual range of *A. pungens* matching the genetic subgroups, different subsets of the whole data in Tables 2.1 and 2.2 were used to construct the models. These subsets were: 1) Gulf of Cadiz, that is, data from the Southern coasts of Iberia from S. Vicente Cape to Gibraltar; 2) Atlantic populations north of S. Vicente Cape; and 3) Corso-Sardinian populations.

2.3 RESULTS

2.3.1 AFLP profiles

Highly reproducible AFLP patterns were obtained for all 22 replicates undertaken in the reproducibility test. An average error rate of 4.7% was estimated across all three primer pairs, in agreement with previous reports below 5% (Blears *et al.*, 1998; Bonin *et al.*, 2004, Skrede *et al.*, 2006). Seven unreliable fragments, non reproducible in at least 5 out of 22 allelic comparisons, were removed leading to a total data set of 223 bands. Only one of the bands was monomorphic across the 221 individuals assayed and no identical multilocus phenotypes were found between individuals or populations. The average number of fragments per individual was 56.82,

and 80.33 per primer combination. Number of private and rare (< 10% of the individuals) fragments per population as well as the percentage of fragments shared between populations are provided in appendix 2.1.

In order to further assess the quality of the data, we recalculated the error rate after discarding potentially unreliable bands following four different criteria: i) slight size differences among putative homologous bands across individuals, ii) low intensity bands, iii) changing intensity of one band across samples and iv) bands of high (upper 10%) or small (lower 10%) molecular weight (Bagley *et al.*, 2001; Bonin *et al.*, 2004). Bands of changing intensity across samples resulted to be the least reproducible, but discarding them implied the loss of too much information (an average of 20.5 bands per primer combination) while only achieving an improvement in reproducibility of less than 1%. We therefore decided to keep the entire data set of 223 bands, since we considered it preferable to maximise the phylogeographical signal by sampling more loci even if maintaining a small amount of noise.

2.3.2 Genetic groups *vs*. geographical disjunctions

Using the increase in the probability of the data and the stability of the assignments as criteria, STRUCTURE revealed that two groups are most appropriate to interpret our AFLP data set, which can be further subdivided into a total of six groups (Fig. 2.1.A). Most individuals were assigned with high probability to one of the clusters and individuals from the same population were always assigned to the same cluster, with the exception of a few outliers (assigned to a different cluster than the remaining individuals from the same population).

The largest increase in the posterior probability of the data occurred at K=2. In estimations of more than two groups it continued to increase slightly. This can be observed directly through the graphic representation of L(K) over 10 or 5 runs for each K value (Fig. 2.1.C) and is even more clear when representing ΔK, which exhibited a maximum value at K=2 (Fig. 2.1.D).

The assignment of individuals across replicate runs only provided stable results for K=2 and K=6 (results not shown). At K=2 all ten runs inferred exactly the same two clusters not matching the main geographical disjunction (Fig. 2.1.A): (i) Gulf of Cadiz

(pops. 7 to 14) and (ii) Cíes islands, Portugal and Corsica-Sardinia (pops. 1, 2 to 6 and 15 to 23). Five outliers were obtained. At K=6 the six subgroups identified are compatible with the two groups revealed under K=2, with the exception of S. Vicente Cape population which represents the geographic boundary between the two K=2 groups (Fig. 2.1.B): (i) Cíes islands with Albufeira (pops.1 and 2), (ii) Alentejo (pops. 3 to 5), (iii) S. Vicente Cape with Bordeira (pops. 7 and 6), (iv) Gulf of Cadiz (pops. 8 to 13), (v) Camarinal (pop. 14) and (vi) Corsica-Sardinia (pops. 15 to 23). These six clusters were consistent in eight out of ten simulations; the two remaining runs produced different solutions but showed much lower log likelihood. Seven individual outliers were obtained, but they seem not meaningful since the assignment of five of them varied across runs. For K higher than six, clustering was very unstable and basically split Corso-Sardinian individuals into several extra groups lacking any pattern. BAPS optimal partition estimate (Fig. 2.1.A) coincided with STRUCTURE by showing the same 6 groups.

Also, the results of AMOVA (Table 2.3) were consistent with the groupings inferred by the STRUCTURE and BAPS packages. In fact, when considering the two main STRUCTURE groups (Corsica-Sardinia and Cíes islands together with Portugal in one group vs. Gulf of Cadiz in the other), the proportion of variance explained by differences between regions was 17.16%. This percentage is larger than that obtained analysing the two main geographical groups: Atlantic vs. Mediterranean (13.86%) and even larger than considering the three main disjunct areas: Cíes islands vs. SW Iberian Peninsula vs. Corsica-Sardinia (16.94%), which indicates that the link between Corsica-Sardinia and Portugal is strong. The grouping of K=6 is also highly supported. This population clustering maximises the among regions variance component (23.18%), relative to the among-populations/within-regions component (21.14%). In at least 10 other designs we tried, variation was always higher among populations/within regions than among regions (results not shown), as was in the two clusterings at K=6 exhibiting low likelihood.

2.3.3 Genetic distance among groups

The same six genetic groups were revealed both after the PCoA of individuals (Fig. 2.2) and the NJ tree of populations (Fig. 2.3). In the Iberian Peninsula, the southernmost populations, in the Gulf of Cadiz, formed a highly compact and distinct

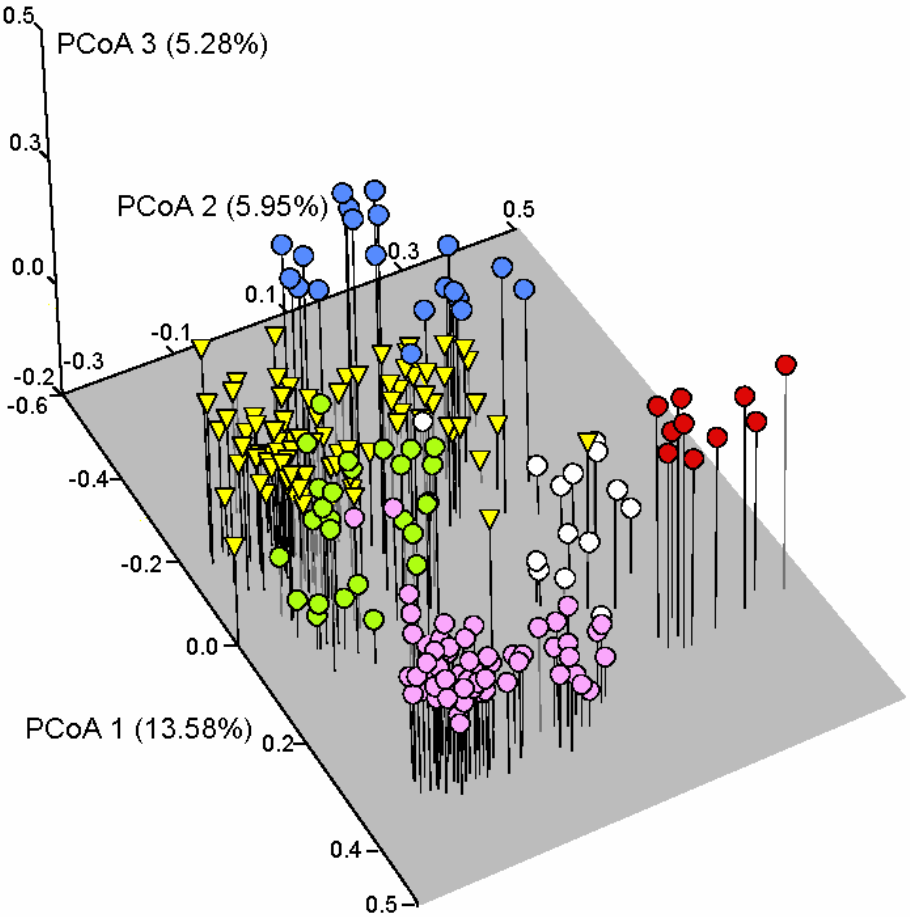
group, whereas the remaining populations showed a decreasing degree of similarity towards the North: S. Vicente-Bordeira group, followed by Alentejo group and, finally, Albufeira-Cíes islands group. Therefore, within the Iberian range, the genetic structure of *A. pungens* was shown to be strongly congruent with the linear distribution of the species along the coast, i.e., increasing the geographical distance among populations leads to a decrease in the genetic distances. In contrast, the disjunct Corso-Sardinian group did not follow this trend of genetic distance correlated with geographic distance since it fell between Alentejo group and Albufeira-Cíes islands group in the PCoA and close to Albufeira-Cíes islands in the NJ tree. Finally, although the Camarinal population was most closely related to the Gulf of Cadiz group it stood well apart from all other populations (see discussion below).

2.3.4 Genetic diversity and differentiation

The estimates for within-population diversity of *A. pungens* were 28.86% (P), 0.11 (H_s) and 0.12 (H_{sh}). The value for AMOVA-derived among-population variation, $\Phi_{ST,}$ was 0.41.

The partitioning of genetic variation within and among populations revealed different patterns in the different areas. Among-population variation resulted to be higher in the Atlantic than in the Mediterranean. Separate AMOVA analyses for every disjunct area or genetic subcluster (Table 2.3) showed that Corso-Sardinian populations are the least differentiated (23.76%), as compared to the Iberian ones. This was also true when the distinct population of Camarinal was excluded from the analysis (results not shown). Shorter branches between Corso-Sardinian populations in the NJ (Fig. 2.3) support this conclusion. Finally, the percent of variance explained by differences among the Cíes islands and Albufeira is high (42.25%). These results suggest that even if the genetic proximity of these two populations is indicated by both bayesian and genetic distance analyses, they are still significantly differentiated.

Figure 2.2 Genetic distance among individuals of *A. pungens* based on AFLP. Principal co-ordinate analysis based on a pair-wise similarity matrix between individual phenotypes using Dice. Scatterplot of 221 individuals against the first three principal axes indicating the percentage of the variance explained by each axis. Colour codes of each individual according to its assignment to any of the 6 genetic groups depicted in Fig. 2.1.



With respect to within-population diversity (Table 2.4), the lowest average estimates corresponded to the Gulf of Cadiz area (P=21.45 %, H_s=H_{sh}=0.08) as compared to the remaining areas. The populations from Corsica and Sardinia on the other hand appeared to be as much or even more variable (P=33.08 %, H_s=0.12, H_{sh}=0.14) than the Iberian. No apparent influence of population size was observed. For very small populations like Albufeira (pop. 2), and Piantarella (pop. 15, the smallest in Corsica, with only 34 individuals cf. Paradis and Culioli, 2003), remarkable levels of

genetic diversity were obtained. Cíes islands populations showed to be less diverse (P=26.46 %, H_s =H_{sh}=0.09) than the closest populations from Alentejo and S. Vicente-Bordeira ranges, but still more diverse than those of the Gulf of Cadiz. The divergent population of Camarinal was the most diverse of the Iberian range (P=34.53 %, H_s=H_{sh}=0.14; see discussion). As expected, Shannon´s index estimates were slightly higher than Nei´s unbiased gene diversity for most populations. The general trend was the same for all the three indices, with only minor differences found among them.

Figure 2.3 Genetic distance among populations of *A. pungens* based on AFLP. Neighbour-joining tree built from a pair-wise distance matrix between 23 populations based on Nei and Li distance (1979).

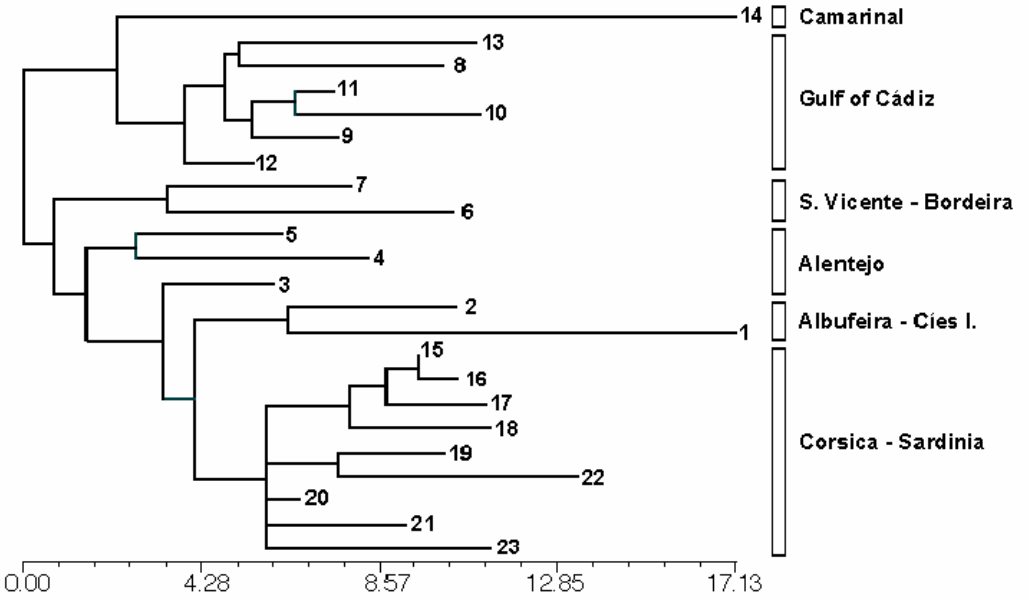


Table 2.3 Analysis of molecular variance (AMOVA) of *A. pungens* based on AFLP considering three hierarchical levels for the whole AFLP data set and two levels for various subsets of the data. The significance of variance components and Φ−statistics was $p<0.0001$ for all tests. Genetic groups are defined based on bayesian and distance analyses. Geographic groups are defined on the basis of disjunctions.

TESTED GROUPS	PERCENTAGE OF MOLECULAR VARIANCE		
	Among groups	*Among populations* within groups	Among individuals within populations
GENETIC GROUPS			
2 groups: Gulf of Cadiz vs. Cíes islands+ Portugal + Corsica-Sardinia	17.16	29.25	53.59
6 groups: Cíes-Albufeira / Alentejo / S. Vicente-Bordeira / Gulf of Cadiz / Camarinal / Corsica-Sardinia	23.18	21.14	55.68
GEOGRAPHICAL GROUPS			
2 groups: Iberian Penisula vs. Corsica-Sardinia	13.86	31.35	54.78
3 groups: Cíes islands / SW Iberian Peninsula / Corsica-Sardinia	16.59	28.86	54.54
GENETIC GROUPS			
2 groups			
Gulf of Cadiz (8 populations)	-	39.54	60.46
Cies islands + Portugal + Corsica-Sardinia (15 populations)	-	33.30	66.70
6 groups			
Cies islands + Albufeira (2 populations)	-	42.25	57.75
Alentejo (3 populations)	-	27.82	72.18
S. Vicente + Bordeira (2 populations)	-	28.72	71.28
Gulf of Cadiz (6 populations)	-	32.10	67.90
Corsica-Sardinia (9 populations)	-	23.76	76.24
Camarinal (1 population)	-	-	-
GEOGRAPHICAL GROUPS			
2 groups			
Iberian Peninsula (14 populations)	-	44.37	55.63
Corsica-Sardinia (9 populations)	-	23.76	76.24
3 groups			
Cies islands (1 population)	-	-	-
SW Iberian Peninsula (13 populations)	-	44.81	58.19
Corsica-Sardinia (9 populations)	-	23.76	76.24
Total range (23 populations)	-	41.17	58.83

2.3.5 Bioclimatic envelope model

2.3.5.1 *Total data approach*

The analysis of the whole data (Fig. 2.4.A) shows that the current distribution of *A. pungens* covers most of its environmental envelope. Localized spots of climatically potential sites are also found in Minorca, Majorca and Algeria.

2.3.5.2 *Subsets of the total data*

When only the data from the Southern coasts of Iberia from S. Vicente Cape to Gibraltar (Gulf of Cadiz) are used to build the model (Figure 2.4.B), the analysis failed to find any other similar area within the Western Mediterranean. This result indicates that the Gulf of Cadiz has different environmental conditions than the rest of the current geographic range of *A. pungens*. When only the Atlantic populations north of S. Vicente are considered (Fig. 2.4.C), populations from the Gulf of Cadiz are not selected by the envelope whereas those from Corsica and Sardinia are. This result is reciprocal since when only Corso-Sardinian populations are used as input (Fig. 2.4.D), Atlantic populations north of S. Vicente are selected but those from the Gulf of Cadiz are excluded.

An inspection of monthly mean values for the variables helps us to identify the parameters responsible for differences and similarities among the climatic profile of *A. pungens* across populations of its range. When looking at the monthly variation of the analysed climatic variables we observe that while Tmin do not exhibit any clear covariation between Corso-Sardinian and Atlantic populations, Tmax and Precipitation show differences between Atlantic- Corso-Sardinian populations on one side and Gulf of Cadiz populations on the other. For precipitation, the almost total lack of rainfall in the Gulf of Cadiz during July (0-1 mm, mean 0.7) and August (1-3 mm, mean 1.9) differs from range values in Atlantic populations (1-24 mm, mean 4.5 in July; 1-6 mm, mean 3.6 in August) and Corso-Sardinian populations (3-33 mm, mean 7.4 in July; 3-16 mm, mean 9.6 in August) (Fig 2.5.A). A similar pattern is observed in the same months for the maximum temperature where the Gulf of Cadiz populations reach 24.7-30.5 ºC (July, mean 28.3) and 24.9-30.5 ºC (August, mean 28.6) while the range values in Atlantic populations is 23.1-28.9 ºC (July, mean 27.2) and 23.2.-29.2 ºC (August, mean 27.5). Similarly, Corso-Sardinian values for maximum temperature are 25.8-28.6 (July, mean27.2) and 26.4-28.9 (August, mean 27.6) (Fig 2.5.B). The patterns shown by both variables underlines the fact that summer drought is the most important factor segregating the Gulf of Cadiz climatic conditions from those of the remaining populations of *A. pungens*.

Table 2.4 Within-population genetic diversity of *A. pungens* based on AFLP. Populations arranged according to main genetic groups based on Bayesian and distance analyses.

Population	*P*	H_s	H_{sh}
I. Cíes-Albufeira			
1- Cíes I.	26.46	0.09	0.09
2- Albufeira	28.70	0.10	0.13
mean	**27.58**	**0.10**	**0.11**
II.Alentejo			
3- Sines	36.77	0.13	0.12
4- Furnas	25.56	0.11	0.10
5- Zambujeira	26.46	0.12	0.13
mean	**29.60**	**0.12**	**0.12**
III. Bordeira-S. Vicente			
6- Bordeira	26.46	0.11	0.14
7- S. Vicente	32.29	0.12	0.15
mean	**29.37**	**0.12**	**0.15**
IV. Gulf of Cadiz			
8- Garrao	19.28	0.07	0.09
9- P. Umbría	29.60	0.11	0.12
10- C. Maneli	19.28	0.06	0.06
11- Matalascañas	17.49	0.07	0.07
12- Doñana	24.66	0.10	0.09
13- Trafalgar	18.39	0.06	0.09
mean	**21.45**	**0.08**	**0.09**
V. Camarinal			
14- Camarinal	**34.53**	**0.14**	**0.14**
VI. Corsica-Sardinia			
15- Piantarella	43.95	0.17	0.14
16- P. Sperone	25.11	0.10	0.11
17- I. Piana	36.77	0.14	0.16
18- G. Sperone	34.08	0.12	0.16
19- R. Maiore	34.98	0.14	0.14
20- P.Liscia	34.53	0.13	0.17
21- B. Mare	31.84	0.11	0.14
22- Stintino	28.25	0.10	0.14
23- C.Mannu	28.25	0.11	0.13
mean	**33.08**	**0.12**	**0.14**

Notes: P = percentage of polymorphic loci, H_s = Nei´s unbiased gene diversity; H_{sh} = Shannon´s index standarised to 6 individuals.

Figure 2.4 Potential distribution modelling of *A. pungens* according to climatic factors based on: A) all presence data; B) Gulf of Cadiz data; C) Atlantic locations north of S. Vicente Cape; and D) Corso-Sardinian data. The locations used to build each model are indicated by numbers (those from which individuals have been sampled in this study, see Table 2.1) and letters (those taken from the literature or herbaria, see Table 2.2). Areas selected by BIOCLIM algorithm are coloured in black; those with an arrow in cases A, C and D were not included as presence data.

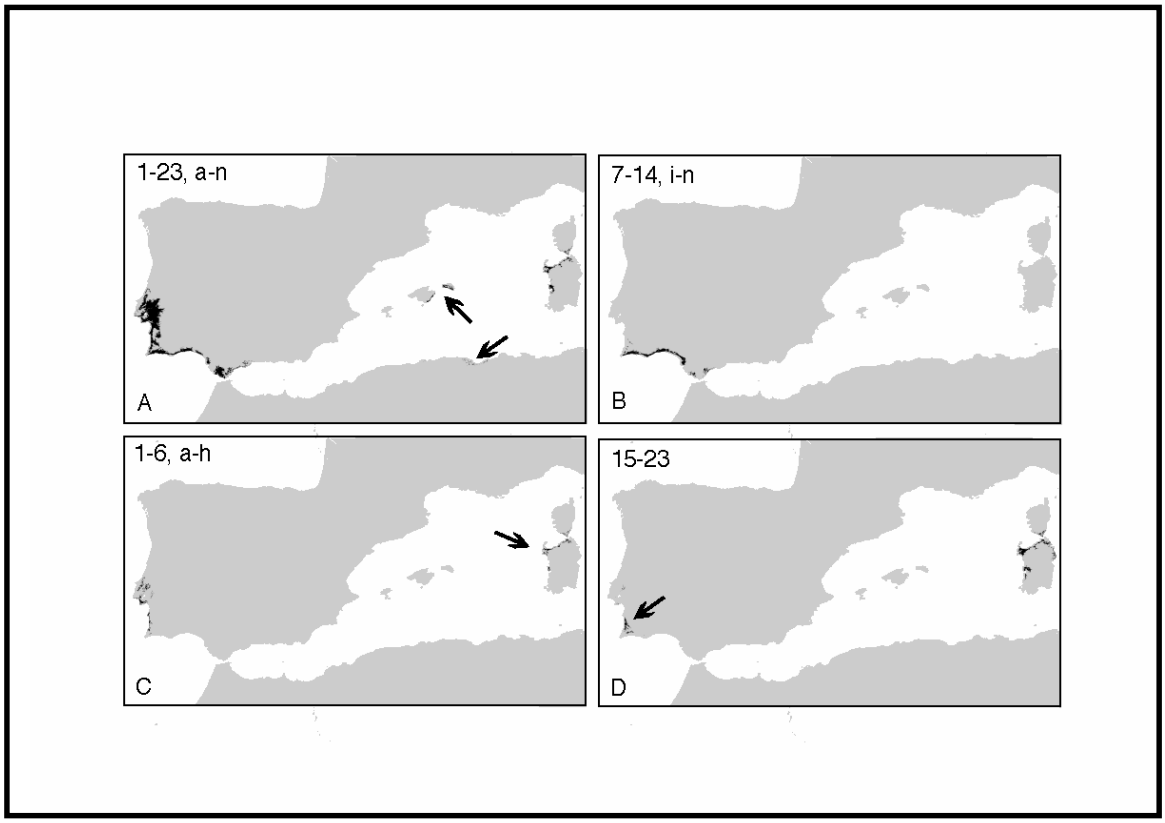
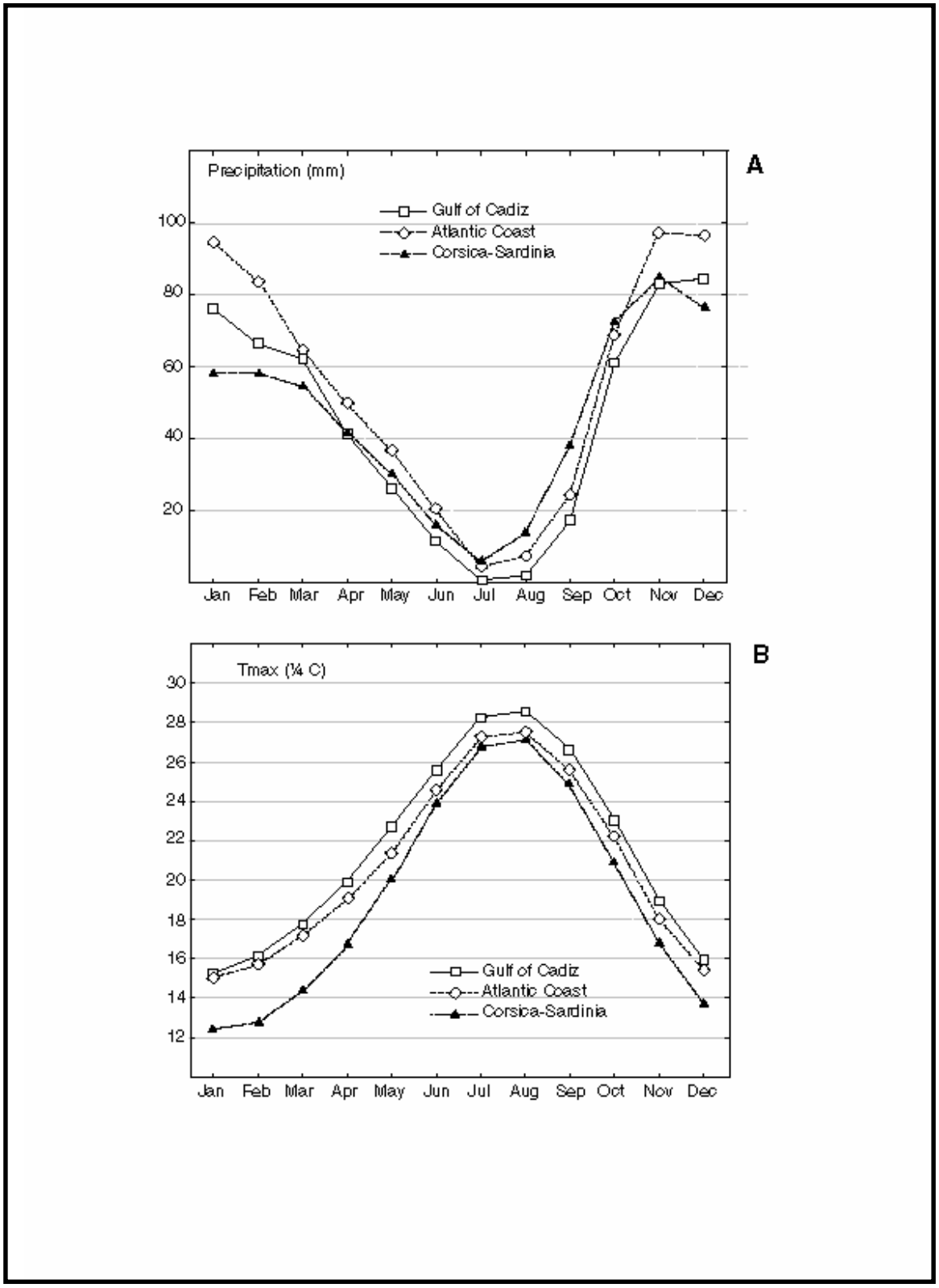


Figure 2.5 Monthly variation of A) mean precipitation and B) mean maximum temperature of pooled populations of *Armeria pungens* from the Atlantic coasts of Iberia, Corsica-Sardinia and the Gulf of Cadiz.



2.4 DISCUSSION

2.4.1 Long-distance dispersal of *A. pungens* into continental islands

2.4.1.1 *Corsica and Sardinia*

It is very likely that populations of *A. pungens* from Corsica and Sardinia are the descendants of individuals dispersed from Portugal. Such a link is suggested by the fact that Portugal together with Corsica, Sardinia and Cíes islands harbour one of the two main genetic lineages of *A. pungens*, the other one being currently located in the Gulf of Cadiz. This is inferred from the most likely Bayesian partition of the AFLP data at K=2 (Fig. 2.1) and supported by the AMOVA showing a higher molecular variance between the two genetic lineages than among the main disjunct geographical areas (Table 2.3). Interestingly, the indument of the leaves correlates this genetic pattern: glabrous or subglabrous leaves in the Gulf of Cadiz (except for the puberulent plants from Camarinal) as compared to leaves with ciliate middle veins in the remaining areas (see chapter 1). Within Portugal, populations from Alentejo region are the closest genetically to the Corso-Sardinian ones, as indicated by the distance analysis (Figs. 2.2 and 2.3).

Genealogical relationships among alleles and thus the direction of dispersal cannot be directly inferred with unordered AFLP (Schaal and Olsen, 2000). However, the distribution of the genetic variation among populations suggests that the dispersal took place from Alentejo, with higher AFLP differentiation among populations, into Corsica or Sardinia, with a lower level of population differentiation. Previously obtained nuclear ribosomal ITS data is consistent with this hypothesis. ITS sequence variation follows a clear geographic structure in *Armeria* that is independent of taxonomy (Fuertes Aguilar *et al.*, 1999; Fuertes Aguilar and Nieto Feliner, 2003). This consists of (i) the occurrence of the same or very similar ITS sequences in different species within the same geographic area and (ii) the presence of different ITS sequences, depending on the geographic origin, in species with wider geographic distributions. This pattern has been interpreted as due to extensive gene flow among congeners, favoured by weak internal reproductive barriers, and biased concerted evolution (Nieto Feliner *et al.*, 2001). *Armeria pungens* from Corsica and Sardinia presents the same ITS sequence as the Iberian populations (see chapter 3) instead of the ITS found in four of the Corso-Sardinian endemics of *Armeria* (Fuertes Aguilar and

Nieto Feliner, 2003). This exception to the geographic structure of ITS variation suggests that contact between *A. pungens* and the other species of *Armeria* has not occurred, probably due to a relatively recent arrival of *A. pungens* into the archipelago.

The distribution of the relatives of *A. pungens* is also consistent with this direction of dispersal into Corsica or Sardinia. The SW of the Iberian Peninsula constitutes the main geographical range of *A. pungens* and its congeners of section *Macrocentron*. They often occur in sympatry or parapatry, suggesting that the divergence from a common ancestor took place in this area. In contrast, the remaining species of *Armeria* from Corsica and Sardinia are not related to section *Macrocentron* (Bernis, 1955). All of them, except for the coastal species *A. soleirolii*, have very similar morphology and occur on ecologically isolated sites, never in sympatry (Arrigoni, 1970).

Taking together the dating of the counter-clockwise rotation of the Corsica-Sardinia microplate from the continent (ended c. 16 Myr ago; Krijgsman, 2002) and the low genetic distinctness of Corso-Sardinian populations of *A. pungens* with respect to their Iberian conspecifics, the conclusion is that migration from Portugal was probably by long-distance dispersal. Land bridges with N Italy proposed during the Messinian salinity crises (c. 5 Myr ago) and during the Pleistocene glaciations, do not seem to have played any role in the evolutionary history of *A. pungens* since neither this species nor any representative of Section *Macrocentron* occur in the Italian Peninsula.

The remarkable within-population variation in Corsica-Sardinia does not fit population genetics predictions. Colonised islands are expected to be less diverse than their continental sources since they harbour only a subset of the initial variation carried by immigrants (Wright, 1931). However, no evidence for founder events was detected in our case. This is an unexpected result, which has several possible causes. Island colonization by several founders is likely to have contributed to a comparatively high genetic diversity within the Corso-Sardinian archipelago (Brown and Marshall, 1981). In fact, this is a prerequisite for strictly allogamous plants like *Armeria* to colonize islands (Baker's rule; Baker, 1955). But the question concerning whether subsequent colonization or gene flow has taken place following the initial establishment cannot be answered with our data. A number of molecular studies support reduced diversity in

colonised islands (Glover and Barret, 1987; Schwaegerle and Schaal, 1979; Inoue and Kawahara, 1990; Richardson *et al.*, 2003). But there are also examples of introduced plants having equal or even more genetic diversity than source areas, for which different causes have been proposed. Several independent introductions is one of them, as suggested for *Rubus alceifolius* in Madagascar (Amsellem *et al.*, 2000). Others include a bottleneck in the source area, as in *Pinus luchuensis* (Chiang *et al.*, 2006), a scenario that does not seem to hold for *A. pungens*.

2.4.1.2 *Cies islands*

The genetic link of the isolated population on offshore Cíes islands with the closest mainland population of Albufeira revealed by the bayesian and distance analyses, suggests that the former originated from the latter (currently 500 km apart) by long-distance dispersal. The low genetic diversity of this population might indicate a founder effect.

2.4.1.3 *Dispersal agents*

Once the distribution of genetic diversity excludes a vicariance scenario, identifying the dispersal agents that caused the disjunction may be relevant. Given the characteristics of the fruits of *A. pungens*, dispersal may have taken place by seabird movements along the Atlantic and the Mediterranean. This scenario is consistent with the establishment of important colonies of seabirds on offshore islands close to the areas of introduction, e.g., Cíes islands themselves (Viada, 1998), Piana island (Lavezzi archipelago) in Corsica and Asinara island (Maddalena archipelago) in Sardinia (Monbailliu and Sultana, 1993). Alternatively, humans may have played the role of dispersers. The first known connections between Corsica-Sardinia and Iberia are the trade routes established by Phoenician colonies since the end of the Bronze Age (Morgenroth, 1999; González Ruibal, 2004). In Cíes islands there are also remains of stable settlements since the end of the Bronze Age (Luaces and Toscano, 1998) and contacts between this area and the Mediterranean increased from the end of the second millennium onwards (González Ruibal, 2004).

2.4.2 Isolation by distance in SW Iberian Peninsula

An important feature of our results is the occurrence of the main phylogeographic break of *A. pungens* in the SW of the Iberian Peninsula. The two main

genetic lineages meet at S. Vicente Cape: the Gulf of Cadiz lineage, on the one hand, and populations north of this cape, on the other. Within both lineages further genetic groupings can be distinguished. All of them strongly conform to the linear distribution of *A. pungens* along the SW coast of Iberia, indicating that there is isolation by distance (Wright, 1943). Climatic conditions correlate the genetic and morphological (glabrous or subglabrous leaves) distinctness of the populations from the Gulf of Cadiz (see below; Fig. 2.4, 2.5).

Excluding the Camarinal population, which is discussed below, the Gulf of Cadiz lineage has experienced genetic drift, as shown by the considerable differentiation between populations apparent in the NJ tree of populations (Fig. 2.3) and the low amounts of within-population diversity shown by genetic diversity parameters, which yielded even lower values than in the isolated population on the Cíes Islands (Table 2.4). Genetic drift might be a consequence of reduced population sizes and/or founder events in the Gulf of Cadiz associated with the glaciations and postglacial period or destructive episodic tsunamis. Tsunamis have been reported in this area every 1500-2000 years according to sedimentary and historical records, and were likely associated with a subduction zone beneath the Gulf of Cadiz and Gibraltar Strait (Luque *et al.*, 2002 and references therein; Gutscher *et al.*, 2005).

2.4.3 Introgression in the Camarinal population

We propose that the distinctness of the population from Punta Camarinal, detected on the basis of the AFLP data is due to introgression from another species of section *Macrocentron,* namely *Armeria macrophylla.* This hypothesis explains the exceptionally high level of genetic diversity harboured by this population and is consistent with data from plastid as well as nuclear ribosomal sequences (see chapter 1). A morphological character that is absent in the remaining populations of *A. pungens* (puberulent leaves and scapes) occurs both in the Camarinal population of *A. pungens* and in *A. macrophylla*. While both species occurs on sandy soils, *A. pungens* is mostly of sand dunes beaches while *A. macrophylla* usually occurs in pine forest understory or shrubland. The introgressive explanation is also spatially feasible since the Camarinal population occurs on a fossil dune only a few hundred meters apart from a population of *A. macrophylla*.

2.4.4 Interpreting genetic structure with bioclimatic modelling

AFLP analysis strongly supports recent long-distance dispersal between Portugal and Corsica-Sardinia but does not provide clues to understand either the mechanisms involved or the explanation for the current absence of *A. pungens* in intermediate sites (Mediterranean Iberian coast, Balearic Islands). Another question which AFLPs do not help to solve is why populations from the Gulf of Cadiz were not the source for the colonization of the Thyrrenian islands despite being the closest geographically. The answers to these questions are provided by our modelling studies. These firstly conclude that bioclimatic conditions of Corso-Sardinian populations match those of the Portuguese populations from the Iberian Peninsula and deviate from the Gulf of Cadiz (Fig. 2.4, 2.5), in concordance with the genetic structure of *A. pungens* based on AFLP data (Fig. 2.1-2.3). Summer drought seems to be a key factor for this similarity pattern (Fig. 2.5). The bioclimatic model also shows that climate conditions where populations thrive are not found nowadays along the West Mediterranean coast except for a few spots in S Minorca, SE Majorca and Gulf of Bejaia (Algeria) (Fig 2.4). Causes for the current absence of *A. pungens* in these sites include the lack of specific habitats (sand dunes) which are not included in the model and/or the failure of diaspores to arrive there but also suggest the possibility of a past occurrence. It has been pointed out that BEM has limitations in explaining potential species distributions when plant-animal interactions and human impact are involved (Hampe, 2004; Pearson *et al.*, 2006). However, the strong correspondence between genetic lineages and bioclimatic groups supports the confidence on the bioclimate modelling as an accurate way of describing the habitat variation within *A. pungens*.

Dispersal agents sometimes provide convincing explanations for disjunct species distribution patterns but this is usually so inasmuch as they depart from stochasticity, e.g., constant winds (Muñoz *et al.*, 2004). We have examined the “habits” of potential dispersal agents to try to elucidate why dispersal of *A. pungens* involved the farthest populations. However, we are not aware of any specific link for potential dispersers (seabirds, winds, sea currents and routes of commerce) between Portugal and Corsica-Sardinia that excludes the Gulf of Cadiz. Therefore, the occurrence of populations in Corsica-Sardinia and their absence in the intermediate sites are more likely due to selection of successful genotypes in similar habitats than to inability of dispersers to bring diaspores from the Gulf of Cadiz or to disseminate them in such intermediate

sites. This is consistent with previous classical and recent studies stressing the role of the Gibraltar strait as a biogeographic barrier not just between Iberia and North Africa but also between Atlantic and Mediterranean lineages (Takhtajan, 1986; Rivas-Martínez *et al.*, 2002; Kadereit *et al.*, 2005). Altogether, our data suggest that the climate is one of the main factors shaping the current genetic structure in *A. pungens* and that bioclimate envelope analysis may be a useful tool to explore the present role of environmental factors even in places with long history of human influence.

2.5 REFERENCES

Álvarez Fernández I, Fuertes Aguilar J, Panero JL, Nieto Feliner G (2001) A phylogenetic analysis on *Doronicum* (Asteraceae, Senecioneae) based on morphological, nuclear ribosomal (ITS) and chloroplast (*trn*L-F) evidence. *Molecular Phylogenetics and Evolution* **20**, 41–64.

Amsellem L, Noyer JL, Le Bourgeois T, Hossaert-McKey M (2000) Comparison of genetic diversity of the invasive weed *Rubus alceifolius* Poir. (Rosaceae) in its native range and in areas of introduction, using amplified fragment length polymorphism (AFLP) markers. *Molecular Ecology* **9**, 443–455.

Arrigoni PV (1970) Contributto alla conoscenza delle Armerie sardo-corse. *Webbia* **25**, 137–182.

Bagley MJ, Anderson SL, May B (2001) Choice of methodology for assessing genetic impacts of environmental stressors: polymorphism and reproducibility of RAPD and AFLP fingerprints. *Ecotoxicology* **10**, 239–244.

Baker HG (1955) Self-incompatibility and establishment after "long-distance" dispersal. *Evolution* **9**, 347–349.

Baker HG (1966) The evolution, functioning and breakdown of heteromorphic incompatibility systems, I. The Plumbaginaceae. *Evolution* **20**, 349–368.

Bakkenes M, Alkemade J, Ihle F, *et al.* (2002) Assessing the effects of forecasted climate change on the diversity and distribution of European higher plants for 2050. *Global Change Biology* **8**, 390–407.

Barry S, Elith J (2006) Error and uncertainty in habitat models. *Journal of Applied Ecology* **43**, 413–423.

Bernis F (1954) Revisión del género *Armeria* Willd. con especial referencia a los grupos ibéricos. *Anales del Instituto Botánico AJ Cavanilles* **11**, 5–288.

Bernis F (1955) Revisión del género *Armeria* Willd. con especial referencia a los grupos ibéricos. Parte segunda (descriptiva de los grupos ibéricos). *Anales del Instituto Botánico AJ Cavanilles* **12**, 77–252.

Blears MJ, De Grandis SA, Lee H, Trevors JT (1998) Amplified fragment length polymorphism (AFLP): a review of the procedure and its applications. *Journal of Industrial Microbiology and Biotechnology* **21**, 99–114.

Blondel J, Aronson J (1999) *Biology and Wildlife of the Mediterranean Region.* Oxford University Press, New York.

Bonin A, Bellemain E, Bronken Eidesen P*, et al.* (2004) How to track and assess genotyping errors in population genetics studies. *Molecular Ecology* **13**, 3261–3273.

Briquet J (1910) *Prodrome de la Flore Corse.* Georg, Genève, Bâle, Lyon.

Brochmann C, Gabrielsen TM, Nordal I, *et al.* (2003) Glacial survival or tabula rasa? The history of North Atlantic biota revisited. *Taxon* **52**, 417–450.

Brown AHD, Marshall DR (1981) Evolutionary changes accompanying colonisation in plants. In: *Evolution Today, Proceedings of the Second International Congress of Systematic and Evolutionary Biology* (eds. Scudder GEC, Reveal JL), pp. 351–363. Hunt Institute for Botanical Documentation, Carnegie-Mellon University, Pittsburgh.

Busby JR (1991) BIOCLIM - A bioclimatic analysis and prediction system. In: *Nature Conservation: Cost Effective Biological Surveys and Data Analysis* (eds. Margules CR, Austin MP), pp. 64–68. CSIRO, Canberra.

Cardona MA, Contandriopoulos J (1979) Endemism and evolution in the islands of the western Mediterranean. In: *Plants and Islands* (ed. Bramwell D), Academic Press, London.

Cervera MT, Ruiz-García L, Martínez-Zapater JM (2002) Analysis of DNA methylation-sensitive AFLP markers. *Molecular Genetic and Genomics* **268**, 543–552

Chiang YC, Hung KH, Schaal BA*, et al.* (2006) Contrasting phylogeographical patterns between mainland and island taxa of the *Pinus luchuensis* complex. *Molecular Ecology* **15**, 765–779.

Clausing G, Vickers K, Kadereit JW (2000) Historical biogeography in a linear system: genetic variation of Sea Rocket (*Cakile maritima*) and Sea Holly (*Eryngium maritimum*) along European coasts. *Molecular Ecology* **9**, 1823–1833.

Corander J, Marttinen P, Mäntyniemi S (2006) Bayesian identification of stock mixtures from molecular marker data. *Fishery Bulletin* **104**, 550–558.

de Queiroz A (2005) The resurrection of oceanic dispersal in historical biogeography. *Trends in Ecology and Evolution* **20**, 68–73.

Dieffenbach CW, Lowe TM, Dveksler GS (1993) General Concepts for PCR Primer Design. *PCR Methods and Applications* **3**, S30–S37.

Ehrich D (2006) Aflpdat: a collection of R functions for convenient handling of AFLP data. *Molecular Ecology Notes* **6**, 603–604.

Eisikowitch D, Woodell SRJ (1975) Some aspects of pollination ecology of *Armeria maritima* (Mill.) Willd. in Britain. *New Phytologist* **74**, 307–322.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software Structure: a simulation study. *Molecular Ecology* **14**, 2611–2620.

Excoffier, Laval LG, Schneider S (2005) Arlequin v. 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* **1**, 47–50.

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

Fiz O, Vargas P, Alarcón ML, Aldasoro JJ (2006) Phylogenetic relationships and evolution in *Erodium* (Geraniaceae) based on *trn*L–*trn*F sequences. *Systematic Botany* **31,** 739–763.

Fuertes Aguilar J, Nieto Feliner G (2003) Additive polymorphisms and reticulation in an ITS phylogeny of thrifts (*Armeria*, Plumbaginaceae). *Molecular Phylogenetics and Evolution* **28**, 430–447

Fuertes Aguilar J, Rossello JA, Nieto Feliner G (1999) Molecular evidence for the compilospecies model of reticulate evolution in *Armeria* (Plumbaginaceae). *Systematic Biology* **48**, 735–754.

Gamisans J (1991) *La végétation de la Corse in Compléments au Prodrome de la Flore Corse,* annexe 2. Conservatoire et Jardin botaniques, Genève.

Gaudeul M, Taberlet P, Till-Bottraud I (2000) Genetic diversity in an endangered alpine plant, *Eryngium alpinum* L. (Apiaceae), inferred from amplified fragment length polymorphism markers. *Molecular Ecology* **9**, 1625–1637.

Glover D, Barret SCH (1987) Genetic variation in continental and island populations of *Eichhornia paniculata* (Pontederiaceae). *Heredity* **59**, 7–17.

González Ruibal A (2004) Facing two seas: Mediterranean and Atlantic contacts in the North-West of Iberia in the first millennium BC. *Oxford Journal of Archaeology*, **23**, 287–317.

Graham CH, Ron SR, Santos JC, *et al.* (2004) Integrating phylogenetics and environmental niche models to explore speciation mechanisms in dendrobatid frogs. *Evolution* **58**, 1781-1793.

Guisan A, Thuiller W (2005) Predicting species distribution: offering more than simple habitat models. *Ecology Letters* **8**, 993–1009.

Gutscher MA (2005) Destruction of Atlantis by a great earthquake and tsunami? A geological analysis of the Spartel Bank hypothesis. *Geology* **33**, 685–688.

Hampe A (2004) Bioclimate envelope models: what they detect and what they hide. *Global Ecology and Biogeography* **13**, 469–471.

Herrera J (1988) Pollination relationships in southern Spanish Mediterranean shrublands. *Journal of Ecology* **76**, 274–287.

Hewitt G (2000) The genetic legacy of the Quaternary ice ages. *Nature* **405**, 907–913.

Hijmans RJ, Cameron SE, Parra JL, *et al.* (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* **25**, 1965–1978

Hijmans RJ, Guarino L, Cruz M, Rojas E (2001) Computer tools for spatial analysis of plant genetic resources data: 1. DIVA-GIS. *Plant Genetic Resources. Newsletter* **127**, 15–19.

Huntley B, Cramer W, Morgan AV, *et al.* (1997). Predicting the response of terrestrial biota future environmental changes. In: *Past and Future Rapid Environmental Changes: the Cpatial and Evolutionary Responses of Terrestrial Biota* (eds. Huntley B, Cramer W, Morgan AV, *et al.*), pp. 487–504. Springer-Verlag, Berlin.

Hsü KJ, Montadert L, Bernoulli D, *et al.* (1977) History of the Mediterranean salinity crisis. *Nature* **267**, 399–403.

Inoue K, Kawahara T (1990) Allozyme differentiation and genetic structure in island and mainland Japanese populations of *Campanula punctata* (Campanulaceae). *American Journal of Botany* **77**, 1440–1448.

Kadereit JW, Arafeh R, Somogyi G, Westberg E (2005) Terrestrial growth and marine dispersal? Comparative phylogeography of five coastal plant species at a European scale. *Taxon* **54**, 861–876.

Krijgsman W (2002) The Mediterranean: Mare Nostrum of Earth sciences. *Earth and Planetary Science Letters* **205**, 1–12.

Krijgsman W, Hilgen FJ, Raffi I, Sierro FJ, Wilson DS (1999) Chronology, causes and progression of the Messinian salinity crisis. *Nature*, **400**, 652–655.

Lago E, Castroviejo S (1993) Estudio citotaxonómico de la flora de las costas gallegas. *Cadernos da Área de Ciencias Biolóxicas*, **1**.

Luaces Anca J, Toscano Novella MC (1998) *Islas Cíes.* Ediciones Nigra Trea, Vigo.

Luque L, Lario J, Civis J, *et al.* (2002) Sedimentary record of a tsunami during Roman times, Bay of Cadiz, Spain. *Journal of Quaternary Science* **17**, 623–631.

Martínez F, Montero G (2004) The *Pinus pinea* L. woodlands along the coast of South-western Spain: data for a new geobotanical interpretation. *Plant Ecology* **175**, 1–18.

McMahon JP, Hutchinson MF, Nix HA, Ord KD (1995) *ANUCLIM, User's guide*. Centre for Resource and Environmental Studies, The Australian National University, Canberra.

Médail F, Quézel P (1999) Biodiversity hotspots in the Mediterranean basin: setting global conservation priorities. *Conservation Biology* **13**, 1510–1513.

Médail F, Quézel P (2003) *Écologie et biogéographie des forêts du bassin méditerranéen*. Éditions scientifiques et médicales Elsevier SAS, Paris.

Miller MP (1997) *Tools for Population Genetic Analyses (TFPGA) 1.3: A Windows Program for the Analysis of the Allozymes and Molecular Genetic Data*. Computer software distributed by the author.

Monbailliu X, Sultana J (1993) Sites required to ensure satisfactory populations levels of seabirds in the Mediterranean. In: *Status and Conservation of Seabirds. Proceedings of the 2nd Mediterranean Seabird Symposium. Calviá, 21-26 March 1989* (eds. Aguilar JS, Monbailliu X, Paterson AM), pp. 303–314. Seo, Madrid.

Moore DM, Yates B (1974) *Armeria* L. in South America. *Botaniska Notiser* **127**, 83–192.

Morgenroth U (1999) Southern Iberia and the Mediterranean trade-routes. *Oxford Journal of Archaeology* **18**, 395–401.

Muñoz J, Felicísimo ÁM, Cabezas F, *et al.* (2004) Wind as a long-distance dispersal vehicle in the Southern Hemisphere. *Science* **304**, 1144-1147.

Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* **89,** 583–590.

Nei M, Li WH (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences USA* **76**, 5269–5273.

Nieto Feliner G (1990) *Armeria*. In: *Flora Iberica* (eds. Castroviejo S, Laínz M, López G, *et al.*), Vol 2, pp. 642–721. Consejo Superior de Investigaciones Científicas, Madrid.

Nieto Feliner G, Fuertes Aguilar J, Rosselló JA (2001) Can extensive reticulation and concerted evolution result in a cladistically structured molecular data set? *Cladistics* **17**, 301–312.

Nieto Feliner G (2002) *Armeria* Willd. In: *Catalogue des Plantes Vasculaires du Nord du Maroc, incluant des clés d'identification* (eds. Valdés B, Rejdali M, Kadmiri AAE, *et al.*). pp. 183-185. Consejo Superior de Investigaciones Científicas, Madrid.

Nix HA (1986) A biogeographic analysis of Australian Elapid Snakes. In: *Atlas of Elapid Snakes of Australia* (ed. Longmore R), pp. 4–15. Australian Government Publishing Service, Canberra.

Paradis G, Culioli JM (2003) *Armeria pungens* en Corse: état de ses stations et effectifs en 2002. *Le Monde des Plantes* **478**, 1–8.

Pearson RG, Thuiller W, Araújo MB, *et al.* (2006) Model-based uncertainty in species range prediction. *Journal of Biogeography* **33**, 1704–1711.

Peterson AT, Soberón J, Sánchez-Cordero V (1999) Conservatism of ecological niches in evolutionary time. *Science* **285**, 1265-1267.

Philipp M, Madsen HES, Siegismund HR (1992) Gene flow and population structure in *Armeria maritima*. *Heredity* **69**, 32–42.

Pinto da Silva AR (1972) *Armeria*. In: *Flora Europaea* (eds. Tutin TG, Heywood VH, Burges NA*, et al.*), Vol 3, pp. 30–38. Cambridge University Press, Cambridge.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* **155**, 945–959.

Reyna-López GE, Simpson J, Ruiz-Herrera J (1997) Differences in DNA methylation patterns are detectable during the dimorphic transition of fungi by amplification of restriction polymorphisms. *Molecular and General Genetics* **253**, 703–710.

Richardson JE, Fay MF, Cronk QCB, Chase MW (2003) Species delimitation and the origin of populations in island representatives of *Phylica* (Rhamnaceae). *Evolution* **57**, 816–827.

Rivas-Martínez S, Días TE, Fernández-González F, Izco J, Loidi M, Lousa M, Penas A (2002) Vascular plant communities of Spain and Portugal. *Itinera Geobotanica* **15,** 5–432

Roberts RJ, Vincze T, Posfai J, Macelis D (2005) REBASE Restriction enzymes and DNA methyltransferases. *Nucleic Acids Research* **33**, D230–D232.

Rohlf FJ (1998) *NTSYS-PC Numerical Taxonomy and Multivariate Analysis System.* Exeter Publications, Setauket, New York.

Schaal BA, Olsen M (2000) Gene genealogies and population variation in plants. *Proceedings of the National Academy of Sciences USA* **97**, 7024–7029.

Schönswetter P, Tribsch A, Niklfeld H (2004) Amplified fragment length polymorphism (AFLP) suggests old and recent immigration into the Alps by the arctic-alpine annual *Comastoma tenellum* (Gentianaceae). *Journal of Biogeography* **31**, 1673–1681.

Schwaegerle KE, Schaal BA (1979) Genetic variability and founder effect in the Pitcher Plant *Sarracenia purpurea*. *Evolution* **33**, 1210–1218.

Shannon CE (1948) A mathematical theory of communication. *Bell Systems Technical Journal* **27**, 379-423, 623-656.

Skrede I, Bronken Eidesen P, Piñeiro Portela R, Brochmann C (2006) Refugia, differentiation and postglacial migration in arctic-alpine Eurasia, exemplified by the mountain avens (*Dryas octopetala* L.). *Molecular Ecology* **15**, 1827–1840.

Suc JP (1984) Origin and evolution of the Mediterranean vegetation and climate in Europe. *Nature* **307**, 429–432.

Takhtajan AL (1986) *The Floristic Regions of the World.* University California Press, Berkeley.

Tauleigne-Gomes C, Lefèbvre C (2005) Natural hybridisation between two coastal endemic species of *Armeria* (Plumbaginaceae) from Portugal. 1. Populational in situ investigations. *Plant Systematics and Evolution* **250**, 215–230.

Thompson JD (2005) *Plant evolution in the Mediterranean*. University of Oxford Press, Oxford.

Tremetsberger K, Talavera S, Stuessy TF*, et al.* (2004) Relationship of *Hypochaeris salzmanniana* (Asteraceae, Lactuceae), an endangered species of the Iberian Peninsula, to *H. radicata* and *H. glabra* and biogeographical implications. *Botanical Journal of the Linnean Society* **146**, 79–95.

Viada C (1998) *Áreas Importantes para las Aves en España*. Sociedad Española de Ornitología SEO, Madrid.

Vos P, Hogers R, Bleeker M, *et al.* (1995) AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research* **23**, 4407–4414.

Winkworth RC, Robertson AW, Ehrendorfer F, Lockhart PJ (1999) The importance of dispersal and recent speciation in the flora of New Zealand. *Journal of Biogeography* **26**, 1323–1325.

Wolfe AD, Liston A (1998) Contributions of the polymerase chain reaction to plant systematics. In: *Molecular Systematics of Plants II. DNA Sequencing.* (eds. Soltis DE, Soltis PS, Doyle JJ), pp. 43–86. Kluwer Academic Publishers, New York.

Woodell SRJ, Dale A (1993) *Armeria maritima* (Mill.) Willd. (*Statice armeria* L.; *S. maritima* Mill.). *Journal of Ecology* **81**, 573–588.

Wright S (1931) Evolution in Mendelian populations. *Genetics* **16**, 97–159.

Wright S (1943) Isolation by distance. *Genetics* **28**, 114.

Xiong LZ, Xu CG, Saghai Maroof MA, Qifa Zhang (1999) Patterns of cytosine methylation in an elite rice hybrid and its parental lines, detected by a methylation-sensitive amplification polymorphism technique. *Molecular and General Genetics* **261**: 439–446

Yeh FC, Boyle TJB (1997) Population genetic analysis of co-dominant and dominant markers and quantitative traits. *Belgian Journal of Botany* **129**, 157.

Yesson C, Culham A (2006) A phyloclimatic study of Cyclamen. *BMC Evolutionary Biology* **6** (art. no.72).

1

3. COLONIZATION FOUNDER EVENTS AND HYBRIDIZATION IN MARGINAL AREAS DRIVE EVOLUTIONARY HISTORY OF RECENT EXPANSION IN THE LINEAR DISTRIBUTED SHRUB *Armeria pungens*

3.1 INTRODUCTION

Understanding of expanding processes has benefited from new views of gene flow acquired in the last decades. First, molecular data have shown that long-distance dispersal is substantially more significant (Winkworth *et al.*, 1999, Ellstrand, 2003) than formerly inferred with direct measurements (Ehrlich and Raven, 1969; Grant, 1981). Second, the perception of interspecific gene flow has parallelly changed and is nowadays viewed as a well documented phenomenon (Rieseberg and Wendel, 1993; Mallet, 2005) that may shape genetic structure of populations long after irreversible species divergence begins. Modern research on hybridization has revealed its ability to improve fitness through creation of genetic novelty, increase of diversity, fixation of heterozygosity or purging of alleles (Anderson and Stebbins, 1954; Arnold, 1997, 2004; Ellstrand, 2003). Among hybridization evolutionary outcomes, expansion of species triggered by creation of fit hybrid genotypes is well supported by the frequent observation of introgressed genotypes in marginal or recently colonised areas (Abbott, 1992; Fuertes Aguilar *et al.*, 1999; Ellstrand and Schierenbeck, 2000; Moody and Les, 2002; Abbott *et al.*, 2003; Choler *et al.*, 2004; Petit *et al.*, 2004; Albert *et al.*, 2006; Lavergne and Molofsky, 2007; Tiébré *et al.*, 2007)

Dispersal and establishment into new areas are usually associated to founder events. Two derived genetic consequences are recognised driven by genetic drift: increase of genetic divergence and loss of diversity. The idea that founder events may promote divergence was first proposed by Mayr (1942) in his genetic revolution model and subsequently complemented by other founder events models of speciation developed by Carson (1968, in Templeton, 1996) and Templeton (1980). This theoretical framework, although having received much attention, has been challenged by experimental data, leading to an extreme controversy over the evolutionary effect of founder events. In addition, genetic assessments of natural systems having experienced documented founder events provide valuable information (Hawley *et al.*, 2006). Postglacial colonizations and introductions on islands are among the best studied of such systems. Glaciations provide a temporal frame for emergence of infraspecific

genetic structure, either by selection and drift in isolated refugia or, alternatively, by drift during leading edge postglacial expansion (e.g. Tregenza, 2002). Islands provide examples of spatially simplified founder events (Baker and Moeed, 1987; Clegg *et al.*, 2002; Abbott and Double, 2003).

The effect of founder events on genetic diversity and divergence are dissimilar and strongly dependent on the age and intensity of the founder event, as well as on patterns of gene flow following the initial establishment. Older colonizations usually lead to more evident genetic footprints than recent ones due to long-term low population sizes (Clegg *et al.*, 2002). The genetic effects of the founder event are notably stronger, for instance, in sequential founder events than in single step introductions (Barton and Charlesworth, 1984; Friar *et al.*, 2000; Clegg *et al.*, 2002; Pruett and Winker, 2005). Likewise, novel selective conditions in the colonised area might intensify founder events. Finally, subsequent patterns of gene flow among founded populations or with the source area are also relevant e.g. several waves of colonization usually minimize genetic impoverishment or divergence (Francisco-Ortega *et al.*, 2000; DeYoung *et al.*, 2003; Arnold, 2004)

In this study we assess the evolutionary history of the sand-dune shrub *Armeria pungens*, which has a main distribution along the SW Iberian coast and two disjunct areas on continental archipelagos: Cíes islands in the Atlantic and S Corsica-N Sardinia in the Mediterranean. Using AFLP and bioclimatic modelling (chapter 2) we discovered that the Atlantic-Mediterranean disjunction of *A. pungen*s was originated by long-distance dispersal from Portugal into Corsica-Sardinia, probably favoured by similar climatic conditions in both areas. In contrast, a remarkably distinct impoverished genetic lineage was found in the Gulf of Cadiz area, despite the absence of physical barriers with contiguous Portuguese populations. This area was found to be climatically different in the GIS modelling. However, the AFLP did not resolve whether the genetic structure in the Gulf of Cadiz lineage could be explained by recent colonization process or, alternatively, by long-term isolation associated to low population sizes. Finally, AFLP yielded evidence of introgression in the southernmost population of the species, given its distinctness and high levels of diversity.

We intend to obtain a cpDNA genealogy in order to examine colonization events in the history of *A. pungens.* First, we aim at assessing the genetic footprints left in the maternal DNA by long-distance dispersal and establishment in Corsica and Sardinia. Second, we intend to distinguish between as scenario of recent linear colonization along the coast *vs*. long-term evolution in the Gulf of Cadiz.

Chloroplast DNA provides molecular markers particularly suitable to investigate founder events due to its low effective population size and maternal inheritance (McCauley, 1995). Additionally, the possibility to generate cpDNA genealogies may resolve ancestor-descendant relationships among areas (Schaal *et al.*, 1998) and allows for comparison of haplotype and nucleotide diversities (Pons and Petit, 1996)

Finally, we aim at testing for introgression in the southernmost population of *A. pungens* (pop. 14- Punta de Camarinal), as suggested by high AFLP diversity levels and remarkable divergence from the remaining populations. In the northernmost population (pop1- Cies islands), morphologically intermediate plants between *A. pungens* and *A. pubigera* were identified. We intend to sequence the other species of *Armeria* arround northernmost and southernmost populations in order to identify putative hybridization events in marginal areas.

3.2 MATERIALS AND METHODS

3.2.1 Sampling strategy

Dried leaves and fruits from 23 populations spanning the entire range of *A. pungens* were used in the genetic study (Table 3.1). Average sampling size per population was 2 for ITS (59 in total) and 5 individuals for *trn*L-F, *mat*K and *trn*S-*trn*fM markers (112 in total). Exactly the same individuals were sampled for *mat*K and *trn*S-fM surveys, whereas some different individuals were included in the case of *trn*L-F. During collection a distance of 10 m. among plants was kept whenever possible. Two hundred and forty one specimens from twenty-one of the populations (ten to eighteen per population) were collected for the morphometric study and are deposited in the herbarium of the Royal Botanic Garden of Madrid (MA). Sample sizes and collection details are reported in Table 3.1.

Table 3.1 A: Collection data of 23 populations of *Armeria pungens* included in the cpDNA and ITS study. N=sampling size.

Population code	Site location, collector,(year)	N *trn*L-F	N *trn*S-fM + *mat*K	N ITS
1- Cíes	Spain, Pontevedra, Illas Cíes, RPP, IMA, IMF, LMR, MSA & PGM (2003)	5	5	4
2- Albufeira	Portugal, Estremadura, Lagoa de Albufeira, GNF & JFA (2002) and RPP, AC & PEG (2003)	5	5	2
3- Sines	Portugal, Baixo Alentejo, Sines-Cercal, RPP, AC & PEG (2003)	5	5	2
4- Furnas	Portugal, Baixo Alentejo, Vilanova de Milfontes, Praia das Furnas, RPP, AC & PEG (2003)	3	3	2
5- Zambujeira	Portugal, Baixo Alentejo, Zambujeira do Mar, GNF & JFA (2002) and RPP, AC & PEG (2003)	5	5	2
6- Bordeira	Portugal, Algarve, Praia do Bordeira, GNF & JFA (2002)	5	5	2
7- S. Vicente	Portugal, Algarve, Cabo de São Vicente, GNF & JFA (2002) and RPP, AC & PEG (2003)	5	5	2
8- Garrrao	Portugal, Algarve, Praia de Garrao, GNF & JFA (2002) and RPP, AC & PEG (2003)	5	5	2
9- P. Umbría	Spain, Huelva, Punta Umbría, Playa de Punta Umbría, GNF & JFA (2002)	5	5	4
10- C. Maneli	Spain, Huelva, Mazagón-Matalascañas, "cuesta Maneli", GNF & JFA (2002)	5	5	2
11- Matalascañas	Spain, Huelva, Torre de la Higuera, Playa de Torre la Higuera, GNF & JFA (2002) and RPP, AC & PEG (2003)	4	4	3
12- Doñana	Spain, Huelva, Doñana National Park, pr. El Inglesillo, SG (2002)	5	5	4
13- Trafalgar	Spain, Cadiz, Cabo de Trafalgar, GNF & JFA (2002) and RPP, AC & PEG (2003)	5	5	3
14- Camarinal	Spain, Cadiz, Punta Camarinal, GNF & JFA (2002)	5	5	3
15- Piantarella	France, Corsica, Bonifacio, Piantarella, GNF & JFA (2002)	5	5	3
16- P. Sperone	France, Corsica, Bonifacio, Petit Sperone, CBNMP(2002)	5	5	4
17- I. Piana	France, Corsica, Bonifacio, île Piana, CBNMP (2002)	5	5	2
18- G. Sperone	France, Corsica, Bonifacio, Grand Sperone, CBNMP (2002)	5	5	2
19- P. Liscia	Italy, Sardinia, Porto Pozzo, Porto Liscia, GNF & JFA (2002)	5	5	2
20- R. Maiore	Italy, Sardinia, S of Santa. Teresa di Gallura, Spiaggia di Rena Maiore, GNF & JFA (2002)	5	5	2
21- B. Mare	Italy, Sardinia, Badesi Mare, GNF & JFA (2002)	5	5	2
22- Stintino	Italy, Sardinia, Pr. Porto Torres, Stintino, S Stagno di Pilo, GNF & JFA (2002)	5	5	2
23- C. Mannu	Italy, Sardinia, NW of Oristano, Capo Mannu, between the cape and Su pallosu, GNF & JFA (2002)	5	5	2

Table 3.1 (continued) B: Collection data of the 20 populations of 10 coastal and subcoastal *Armeria* included in the cpDNA study. N=sampling size.

Population	Species	Site location, collector (year)	N *trn* L-F, *trn*S-fM, *mat*k	N ITS	Voucher
PB1	*A. pubigera*	Pontevedra, Illas Cíes, Monteagudo RPP (2003) and (2005)	2	2	RP17/05, RP226-1/03
PB2	*A. pubigera*	Pontevedra, Illas Cíes, Faro, RPP(2005)	1	1	RP63/05
MT	*A. maritima*	Pontevedra, Illas Cíes, Monteagudo, Lago dos Nenos, RPP(2005)	3	1	RP11/05, RP18/05, RP26/06
PUXPB	*A. pungens* X *A. pubigera*	Pontevedra, Illas Cíes, Monteagudo, S playa de Figueiras , RPP(2006)	4	4	RP10/06, RP14/06, GN4293, 95
	A. pungens X *A. pubigera*?	Pontevedra, Illas Cíes, Monteagudo, S playa de Figueiras, GNF, JFA, XRG (2000)	1		GN4292
W1	*A. welwitschii*	Portugal, Baixo Alentejo, foz del río Lis (orilla N), dunas, GNF & JFA (1998)	3		GN3950-52
W2	*A. welwitschii*	Portugal, Estremadura, pr. Pedro de Muel, 1Km. N the lighthouse, GNF & JFA (1998)	1		GN3954
W3	*A. welwitschii*	Portugal, Nazaré, Faro de Nazaré, GNF & JFA (1998)	2		GN3957-58
W4	*A. welwitschii*	Portugal, Cabo Carvoeiro, pr. Peniche, GNF & JFA (1998)	2		GN3962-64
B1	*A. berlengensis*	Portugal, Isla Berlenga, 80 m, Fisterra-Cora do sono, GNF & JFA (1998)	1		GN3967
B2	*A. berlengensis*	Portugal, Isla Berlenga, Casceiro? Do Mosteiro, GNF & JFA (1998)	1		GN3971
R1	*A. rouyana*	Portugal, Ribatejo, way to Salvaterra de Mayos, GNF & JFA (1998)	2		GN3959, 61
R2	*A. rouyana*	Portugal, Alentejo, Santo André, pr. Lake, GNF & JFA (1998)	1		GN3946
V1	*A. velutina*	Huelva,Doñana,carretera de circunvalación de Matalascañas, GNF & JFA (1997)	2		GN3882, 84
H1	*A. hirta*	Cádiz, Arcos-Medina Sidonia, GNF & JFA (1997)	2		GN3864, 66
H2	*A. hirta*	Cádiz, Conil de la Frontera, Pinar de Roche, pr. El colorado?, GNF & JFA (1997)	1		GN3873
H3	*A. hirta*	Málaga, Manilva, Punta de la Chullera, GNF & JFA (1997)	1		GN3857
M1	*A. macrophylla*	Cádiz, Conil de la Frontera, pr. Chiclana, GNF & JFA (1997)	2		GN3876, 78
M2	*A. macrophylla*	Cádiz, carretera de Vejer a Trafalgar, pr. Caños de Meca, GNF & JFA (1997)	1		GN3870
M3	*A. macrophylla*	Cádiz, Tarifa, Punta Camarinal, GNF(2005)	3	3	GN4669-71
G1	*A. gaditana*	Conil de la Frontera, GNF & JFA (1997)	2		GN3880, 81

Abbreviations of collectors: AC, A. Costa; CBNMP, Conservatoire Botanique National Mediterranéen de Porquerolles; GNF, G. Nieto Feliner; IMA, I. Martínez Arcos; IMF, I. Martínez Fernández; JFA, J. Fuertes guilar; LMR, L. Muriel Ríos; MSA; M. Souto Alonso; PEG, P. Escobar García; PGM, P. García Meijide; RPP, R. Piñeiro Portela; SG, S. Gatelier.

Populations of all coastal *Armeria*s from Northern and Southern Atlantic Iberia were included in the study in order to test for potential interspecific gene flow (putative maternal donors) of the northernmost and southernmost Iberian populations of *A. pungens*. To examine the northernmost population, sympatric species from the Cíes islands were sampled (*A. maritima*, *A. pubigera,* as well as plants with intermediate appearance between *A. pungens* and *A. pubigera*). Populations of three species (*A. welwitschii*, *A. berlengensis* and *A. rouyana)* from the central portion of the Iberian Atlantic coast were also studied. To assess the southernmost population, eight populations belonging to all four sublittoral species in this area (*A. hirta*, *A. velutina*, *A. macrophylla* and *A. gaditana*) were included. A total of 38 individuals from 17 populations belonging to 10 different taxa were included in the genetic study at the interspecific level. The local endemics *A. pseudoarmeria* and *A. euskadensis* were the only coastal Iberian species not included in the study. Both are allopatric with respect to *A. pungens*, and thus do not seem a priori relevant to our purposes.

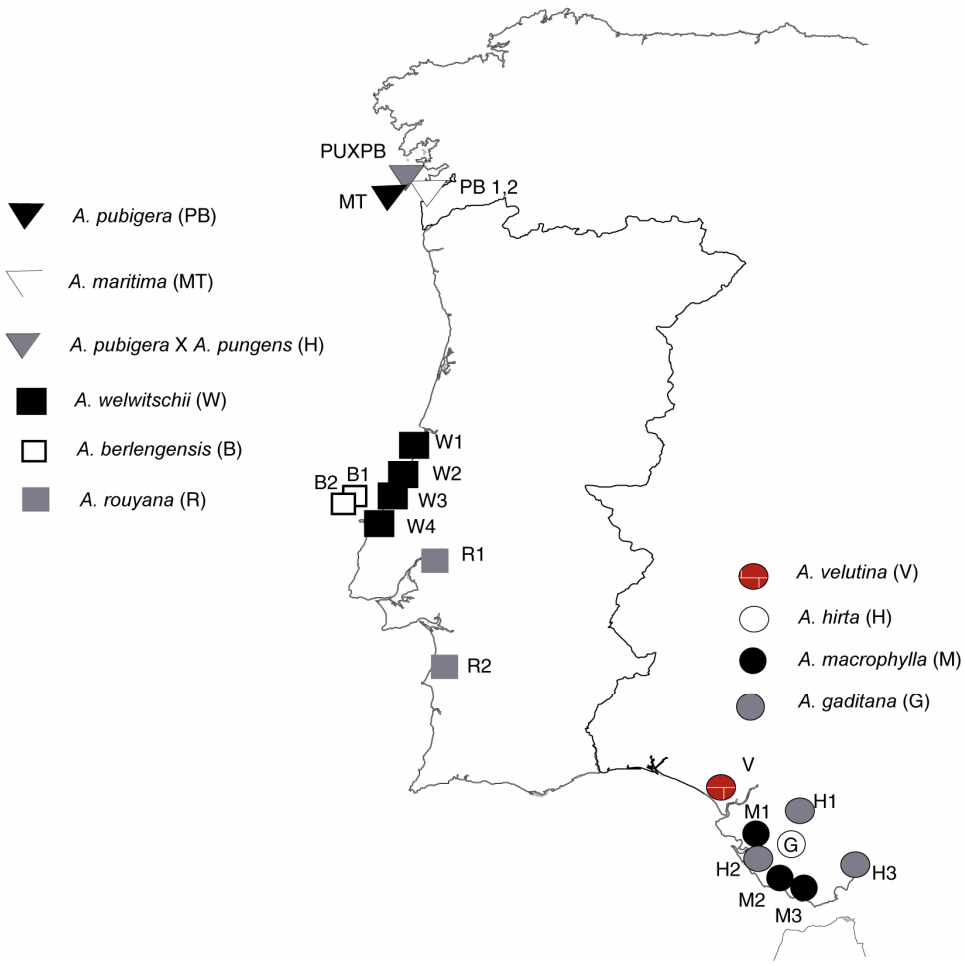
3.2.2 DNA isolation, PCR amplification

DNA isolation from dried or fresh leaves was performed with DNeasy Plant Mini Kit (Qiagen). The quantity (10-100 ng/µl).and quality of the isolated DNA was either determined by agarose electrophoresis or with a spectrophotometer Gene Quant II UV (Amersham Biosciences). Contamination was monitored with negative controls.

ITS and three chloroplast regions were amplified with universal primers: ITS (primers P1A and P4: Francisco-Ortega *et al.*, 1999), *trn*L-F intergenic spacer (primers e and f from Taberlet *et al.*, 1991), *trn*S-fM and *trn*K-K regions (Demesure *et al.*, 1995). PCR reactions were performed in a volume of 50 µl containing 2 µl of template DNA, 5 µl buffer 10x or 10µl buffer 5x, 3 µl $MgCl_2$ (25 mM), 2.5 µl 2 mM dNTPs, 2 µl of each primer 10 µM and 0.4 µl Taq (Promega). Amplifications were also performed with Ready-To-Go beads (Amersham Biosciences) in a final volume of 25 µl, with 1-5 µl template DNA, 0.5-2 µl 10 µM of each primer, 1µl Dimethyl Sulfoxide (ITS and *trn*L-F) or 1µl BSA (*trn*SfM and *trn*K-K). PCR conditions for ITS followed Lihovà *et al.* (2004) with 52-54ºC annealing temperature. *trn*L-F profile followed Gutiérrez Larena *et al.* (2002), with 50ºC annealing temperature. Only primer e was used for sequencing. Cycling conditions for *trn*S-fM and *trn*K-K consisted of 5 min at 94ºC, 25 or 30 cycles

of 30s 94ªC, 30s-2 min 55Cº-62ºC and 1.5-3 min 72ºC, with a final elongation cycle of 10 min 72ºC (Table).

Figure 3.1 Location of the 20 populations of coastal and subcoastal species of *Armeria* sampled for chloroplast DNA as potentially involved in interspecific gene flow with populations of *A. pungens* in the northernmost (pop. 1- Cíes) and southernmost (pop. 14- Camarinal) part of the species distribution. Population names and sampling data are reported in Table 1.1.



*trn*S-fM was amplified with universal primers. Low quality DNA extractions had to be amplified in two fragments by combining universal and internal specific primers: trnS-intfM (pr.800 bp) and intS-*trn*fM (pr.400 bp). Sequencing was performed with *trn*S and intS. The latter aimed at avoiding long mononucleotide strings at the end of the sequence. A preliminary *trn*K-K matrix, 2274 bp long, containing 46 sequences of *A. pungens* (two individuals per population) was yielded using internal (intK2-intK3) and external (extK2-extK3) specific primers. All fourteen variable sites, except for one substitution in a single individual at position 532, were located within *mat*K, as deduced from comparison with *Limonium latifolium* (genebank AY514861). Therefore, only *mat*K encoding gene was amplified and sequenced in the remaining samples using intk2-intk3. To deal with low quality DNA isolations, PCR products obtained with extK2-extK3 were used as template in the amplification with intK2-intK3. PCR conditions are shown in Table 3.2. Primer design was performed with Primer3 software (Table 3.3).

PCR products were checked by agarose electrophoresis and ethidium bromide or Syber green staining. Thermocyclers used were Gene Amp 9700 (Applied Biosystems), PTC-200 (MJ research) or Mastercycler (Eppendorf). Purification of PCR products was performed with Ultraclean PCR clean-up kit (MoBio, USA), QIAquick PCR Purification Kit (QiaGen) or Pertfectprep Gel Cleanup (Eppendorf). Sequencers used were ABI Prism 377 or ABI Prism 3730 (Applied Biosystems).

Table 3.2 Template quantities and PCR cycle conditions for each primer combination of specific *trn*S-fM and *mat*K primers for *Armeria.*

Primer combination	Template quantities	PCR conditions: nr cycles , annealing, elongation
*trn*S-fM *A. pungens*	2 μl DNA (1:0, 1:10)	25 cycles, 30s 62ºC, 1.5 min at 72ºC
*trn*S-fM other spp.	3-5 μl DNA (1:0)	30 cycles, 2 min 55ºC, 3 min at 72ºC.
*mat*K *A. pungens*	1μl DNA (1:10)	25 cycles, 30s 58-60ºC, 1.5 min. at 72ºC
*mat*K other spp.	5 μl DNA (1:0)	30 cycles, 2 min 55-58ºC, 3 min at 72ºC

Tabla 3.3 Sequences of the six specific primers for amplification of *trn*S-fM and *mat*K regions in *Armeria.*

Primer	Sequence
intS:	5'- GGG GGT CAA ACT TAA ACT TCT TG-3'
intfM:	5'-CCG CAT TTC TTC CTT CTC AA-3'
intK2:	5'- GGC AAG ACG ATT TCT TAT ATC CA-3'
intK3:	5'-TGA TTC AAA CCC TAC CAC ATG A-3'
extK2:	5'-AAG CAA GGA TTC GTC CAC AC-3'
extK3:	5'- TGC ACA CAG CTT TCC CTA TG-3'

3.2.3 Analysis of cpDNA within *Armeria* pungens

112 consensus sequences of *A. pungens* were obtained for each marker and manually aligned with Bioedit 5.0.9 (Hall, 1999) taking into account the guidelines by Kelchner (2000). Genealogical relationships among haplotypes were estimated in a statistical parsimony network with TCS 1.21 (Clement *et al.*, 2000). Phylogeographic signal contained in indels was incorporated into the analysis. For gaps associated to mononucleotide repeats of different lengths, every contiguous gap position was coded as a single substitution. The only indel found was counted as a single mutation. Given the low taxonomic focus of the analysis (Provan *et al.*, 2001) the risk of homoplasy was considered unlikely.

Population genetics estimates were calculated using the programs dnaSP 4.10.7 (Rozas *et al.*, 2003) and Haplonst (Pons and Petit, 1996).Within-population diversity was estimated (i) in terms of allele frequencies with Nei´s haplotypic diversity, Hd (ii) in terms of weighted sequence divergence, with nucleotide diversity, Pi. Population subdivision was estimated for different geographical regions based on haplotype frequencies only, Gst and both on haplotype frequencies and genetic distance between haplotypes, Nst (Pons and Petit, 1996). Difference between the two estimates was assessed with a haplotypes identity permutation test implemented in PERMUT (www.pierroton.inra.fr). Allelic richness, A, corrected for sample size N=28 with rarefaction method of El Mousadik and Petit (1996) was calculated with FSTAT 2.9.3.2 (Goudet, 2001).

3.2.4 Analysis of cpDNA variation across species

The chloroplast DNA matrix containing 78 sequences was constructed by adding 38 *Armeria* sp. sequences to a reduced matrix of 40 *A. pungens* sequences (selected out of 112 by deleting identical sequences of each population). Indels were treated as missing values.

Phylogenetic analyses were performed using Maximum Parsimony and Bayesian approaches. Maximum-Parsimony analyses were carried out with PAUP v. 4.0b10 (Swofford, 2002) under Fitch parsimony, heuristic searches with 100 random taxon addition replicates, holding 100 trees at each step, and TBR branch swapping on all trees (options steepest descent off and mulltrees on). 18540 trees were saved and the best score was 128. Results were compared with more restrictive searches setting mulltrees off option (118 saved trees), and by saving 100 trees per replicate (3600 trees retained). Strict consensus trees from different search strategies showed exactly the same topology. Branch support was calculated by fast boostraping with 10,000 resamples. Bootstrap supports were compared with those produced by coding gaps as presence/absence characters (see section 3.2.3).

Previous to the Bayesian analyses the cpDNA matrix was imported into Modeltest 1.1b (Nylander, 2003) to select the model of evolution that best fits the data under the criteria hierarchical likelihood ration test (hLRT) and Akaike Information Criterion (AIC). Subsequently the Bayesian analysis was run with Mr. Bayes 3.1 (Ronquist and Huelsenbeck, 2003) under default settings for 10^6 generation. Convergence of the two independent runs was diangnosed using standard deviation of split frequencies and trees obtained before discarded. The remaining trees were compiled in a majority rule consensus tree, using posterior probability as a measure of clade support.

3.2.5 Analysis of ITS sequences within *A. pungens*

59 ITS consensus sequences of *A. pungens* were generated from forward and reverse sequences and aligned. Intraindividual polymorphisms, detected as double peaks, were carefully checked on both reverse and forward electropherograms and coded using IUPAC ambiguity codes according to Fuertes Aguilar *et al.* (1999).

3.2.6 Analysis of ITS sequences across species

The ITS sequences of *A. pungens* were reanalysed in the frame of the phylogenetic study of the genus *Armeria*, comprising 133 sequences representing 70% of the species (Fuertes Aguilar and Nieto Feliner, 2003). In that study, all coastal Iberian species of sect *Macrocentron* shared a unique ribotype (R1), also widespread across many other inland western Iberian species, with only a few substitutions among species. This large Iberian clade, combining moderate genetic variation and high morphological diversification, suggests a mixture of recent radiation of the group and extensive hybridization and biased concerted evolution toward ribotype R1. Coastal species from sect. *Maritima* formed an independent clade together with the morphologically related alpina group (R7). Finally, Corso-Sardinian species *A. leucocephala*, *A. multiceps*, *A. soleirolii* and *A. sardoa* fell within a different ribotype (R5).

Taking into account the evolution of ITS (biparental inheritance, concerted evolution; Dover, 1982), the geographic structure of ITS or the detection of additive polymorphic sites (APS) involoving parsimony informative sites might effectively identify hybrid events (Nieto Feliner *et al.*, 2001).

Therefore, in populations where *A. pungens* is known to be sympatric with species from other ITS clades, individuals were checked for putative additivities revealing hybridization. Specifically, additive polymorphis sites revealing coexistence R1 (characteristic of A. pungens) and R7 (characteristic of A. maritime/ clade) were checked in the Cies Islands while in Corsica and Sardinia additivities between R1 and R5 were inspected. For this purpose, 10 additional sequences from congeners close to the Cies Islands (pop. 1) and Camarinal (pop. 14) were obtained (Table 1.1 B)

3.2.7 cpSSR test

Amplification protocols for 8 universal primer combinations of chloroplast simple sequence repeats (ccmp, Weising and Gardner, 1999; ccSSR Chung and Staub, 2003) were optimised in a volume of 10 µl containing, 1.5µl 15-250 ng/µl (ccmp1, SSR13,20) or 5-25 ng/µl (ccmp 2,3,5,6; SSR7,15) template DNA, 1 µl buffer 10x, 0.6

μl MgCl2 (25 mM), 1 μl 2 mM dNTPs, 0.5 μl of each primer 10 μM and 0.1 μl Taq (Promega). Primers were 5'-labelled with either HEX or FAM. PCR products were run on an ABI Sequencer and allele sizes scored with Genemapper 3.7 (Applied Biosystems). Individuals representing the main genetic groups on the basis of previous ALFP data (chapter 2) were genotyped in order to evaluate the level of polymorphism.

3.3 RESULTS

3.3.1 Genetic variation within *A. pungens*

3.3.1.1 *Pilot study of choroplast microsatellites*

Universal chloroplast microsatellites primers yielded amplified fragments of comparable length to *Nicotiana tabacum* and other species. The 8 loci analysed (Table 3.4) exhibited low variation within *A. pungens*, which we considered is not enough to draw any phylogeographic conclusions. Double peaks were consistently obtained for ccmp5 and ccSSR15, but protocols were no further optimised since this technique was discarded to address our questions.

Table 3.4 Chloroplast microsatellite analysis of *Armeria pungens*. Pilot study using 8 loci.

Primer	Allele Sizes	Annealing temp. (C°)	Polymorphism test	N individuals (areas[1] included in the analysis)
Ccmp2	201 202	50	No geographical pattern for both size variants	17 (I,II,III,IV,V,VI)
Ccmp3	112 113	52	Cies (pop. 1)exclusive size variant	24 (I,II,III,IV,V,VI)
Ccmp5	93-94 94-95 96-97	50	Cies (pop.1) and Doñana (pop12) exclusive size variants	16 (I,II,III,IV,V,VI)
Ccmp6	102	52	Monomorphic	22 (I,II,III,IV,V,VI)
SSR15	268-269	52	Monomorphic	22 (I,II,III,IV,V,VI)
SSR20	313	50	Monomorphic	12 (I,II,III,IV,VI)
SSR13	261	50	Monomorphic	6 (II,III,IV,VI)
SSR7	311	50	Monomorphic	5 (I,II,III,V)

Ccmp conserved chloroplast microsatellite primers by Weising and Gardner (1999)
ccSSR consensus chloroplast simple sequence repeats primes by Chung and Staub (2003)
[1]areas are considered according to six AFLP genetic groups (chapter 2): I (pos 1-2), II (pops 3-5), III (pops 5-7), IV (pops 8-13), V (pop14),VI (pops. 15-23).
ccmp5 and ccSSR15 yielded double peaks, the size of both peaks is separated by -.

3.3.1.2 *cpDNA analysis within A. pungens*

The final cpDNA matrix for *A. pungens* consisted of concatenated sequences from three regions (*trn*L-F, *trn*S-fM and *mat*K) from 112 individuals resulting in a total of 3037 bp. Twenty nucleotide substitutions, three mononucleotide repeats and one indel defined 16 haplotypes within *A. pungens* (Table 3.5). As expected from its fast evolution (Hilu and Liang, 1997; Kelchnner, 2000; Sang, 2002, Hausner *et al.*, 2006), *mat*K showed higher variation.

Phylogenetic relationships between the 16 cpDNA haplotypes revealed three different cpDNA lineages within *A. pungens* following a clear geographic structure (Fig. 3.2). Two extremely divergent haplotypes were detected in the northernmost (pop.1-haplotype P) and southernmost (pop. 14-haplotype Q) populations of the Iberian range of the species: "Cies islands" and "Camarinal" lineages. Both lineages differ by ten and eight mutational steps, respectively from the third group of related haplotypes that can be considered, "Core *A. pungens*", which comprises the remaining Iberian populations as well as the Corso-Sardinian ones. Two genetic lineages can be further distinguished within this core, which are congruent with indumentum of the leaves: i) "Gulf of Cadiz" lineage and ii) "Portuguese-Corso-Sardinian" lineage. The "Gulf of Cadiz" lineage consisted of one haplotype across all six populations (pops. 8 to 13-haplotype N), except for a single ancestral haplotype in one individual (pop. 8-haplotype E). Within the "Portuguese-Corso-Sardinian" lineage, Portuguese populations harboured a high number of haplotypes (pops. 2 to 7- haplotypes A, B, C, F, G, H, I, J, K, L, M, D), three of them shared with the almost genetically uniform Corso-Sardinian islands: haplotype A in pop 4. and highly dominant in Corsica-Sardinia (pops 15 to 22); haplotype C in pops. 5 and 19; haplotype G in pops. 2, 6, 7 and 23. In addition, two local exclusive haplotypes were found in Corsica and Sardinia (pop. 15-haplotype M and pop. 23-haplotype D).

The high level of population subdivision within *A. pungens* (Gst=0.73), points out the remarkable genetic structure of the species, which shows sharp differences across regions and lineages regarding population diversity and differentiation. For maternal markers in angiosperms an average Gst=0.637 has been reported (Petit *et al.*,

2005). “Portuguese” populations exhibited high intrapopulation (average hs=0.48) and regional (A_{28}=10; ht=0.90) diversities, resulting therefore in a moderate level of regional differentiation (Gst=0.47) (Table 3.6). For most of the populations haplotypic diversity was very high (Hd=0.67-0.80) although two of them were fixed for one haplotype (pops. 2 and 6) yielding, thus, and Hd value of 0 (Table 3.7). Given the dominance of haplotype A, diversity in “Corsica-Sardinia” was significantly lower both at population (hs=0.08; Hd=0) and regional scales (A_{28}=4; ht=0.26), resulting thus also in a moderate regional differentiation (Gst=0.66). Finally, genetic variation was almost zero along the “Gulf of Cadiz” range (Hd=0.07; A_{28}=2; ht=0; Gst=0), whereas “Cies islands” and “Camarinal” populations were completely monomorphic across all three chloroplast markers.

Table 3.5 List of 16 chloroplast haplotypes recognised within *Armeria pungens* with indication of the variable nucleotides in each of the three chloroplast regions sequenced.

haplotype	*trn* L-F				*trn* S-fM					*mat*k														
	194	215	223	250	216	660-61	832	1089-90	1098-99	863	917	962	973	1037	1061	1189	1378	1397	1460	1676	1836	2106	2113	2124-29
A	T	G	T	G	**G**	..	**C**	**C.**	..	G	G	C	T	T	C	C	C	C	A	G	A	G	A	.
B	T	C	T	G	**G**	..	**C**	**C.**	..	G	G	C	T	T	C	C	C	C	A	G	A	G	A	.
C	T	G	T	G	**G**	..	**C**	**C.**	..	G	T	C	T	T	C	C	C	C	A	G	C	G	A	.
D	T	G	T	G	**G**	..	**C**	**C.**	..	G	T	C	T	T	C	C	C	C	A	G	A	G	A	.
E	T	C	T	G	**G**	..	**C**	**C.**	..	G	T	C	T	T	C	C	c	C	A	G	A	G	A	.
F	T	C	T	G	**G**	..	**C**	**C.**	..	G	T	C	T	T	C	C	C	C	A	G	C	A	A	.
G	T	C	T	G	**G**	..	**C**	**C.**	..	G	T	C	T	T	C	C	C	C	A	G	C	G	A	.
H	T	G	T	G	**G**	..	**C**	**CC**	..	G	t	C	T	g	C	C	C	C	A	G	A	G	A	.
I	T	C	T	G	**G**	..	**C**	**CC**	..	G	T	C	T	T	C	C	C	C	A	G	A	G	A	.
J	T	C	T	G	**G**	..	**C**	**C.**	..	G	T	C	T	T	C	C	C	C	A	G	C	A	C	.
K	T	G	T	G	**G**	..	**C**	**C.**	..	G	T	C	T	T	C	C	C	C	A	G	C	A	A	.
L	T	G	T	G	**G**	..	**C**	**C.**	..	G	T	C	T	T	C	C	C	C	A	G	C	A	C	.
M	T	G	T	G	**G**	..	**G**	**C.**	..	G	T	C	T	T	C	C	C	C	T	G	C	G	A	ATAAT
N	T	C	T	G	**G**	**AA**	**C**	**C.**	..	G	T	C	G	T	C	C	A	C	A	G	A	G	A	
O	G	G	G	G	**G**	**A.**	**G**	..	**A.**	T	T	C	T	T	C	A	C	C	T	T	C	G	A	AATAAT
P	G	G	T	T	**A**	..	**G**	**CC**	**AA**	G	T	A	T	T	T	C	C	A	T	G	C	G	A	.

Genetic differentiation in *A. pungens* taking into account allele similarity resulted significantly higher (Nst=0.85) than based on allele frequencies (confirmed by permutation of haplotypes identities test, U=Gst/Nst=2.15; Pons & Petit 1996). In the frame of the island model, this significant difference supports historical (non equilibrium) population genetic structure (Petit *et al.*, 2005). In Portugal, there is no difference between Gst (0.47) and Nst (0.48) indicating mutation-drift equilibrium. In Corsica-Sardinia Nst (0.86) is also higher than Gst (0.66), indicating that despite general low variation in the region, closely related haplotypes are more likely to be found in the same population rather than in different populations.

Figure 3.2 A: Geographic distribution of the 16 *A. pungens* chloroplast haplotypes based on three non coding regions sequenced in 112 individuals (*trn*L-*trn*F, *trn*S-*trn*fM, *mat*K) from 23 populations (numbered as in Table 3.1). **B:** Statistical parsimony network showing the genealogical relationships between haplotypes.

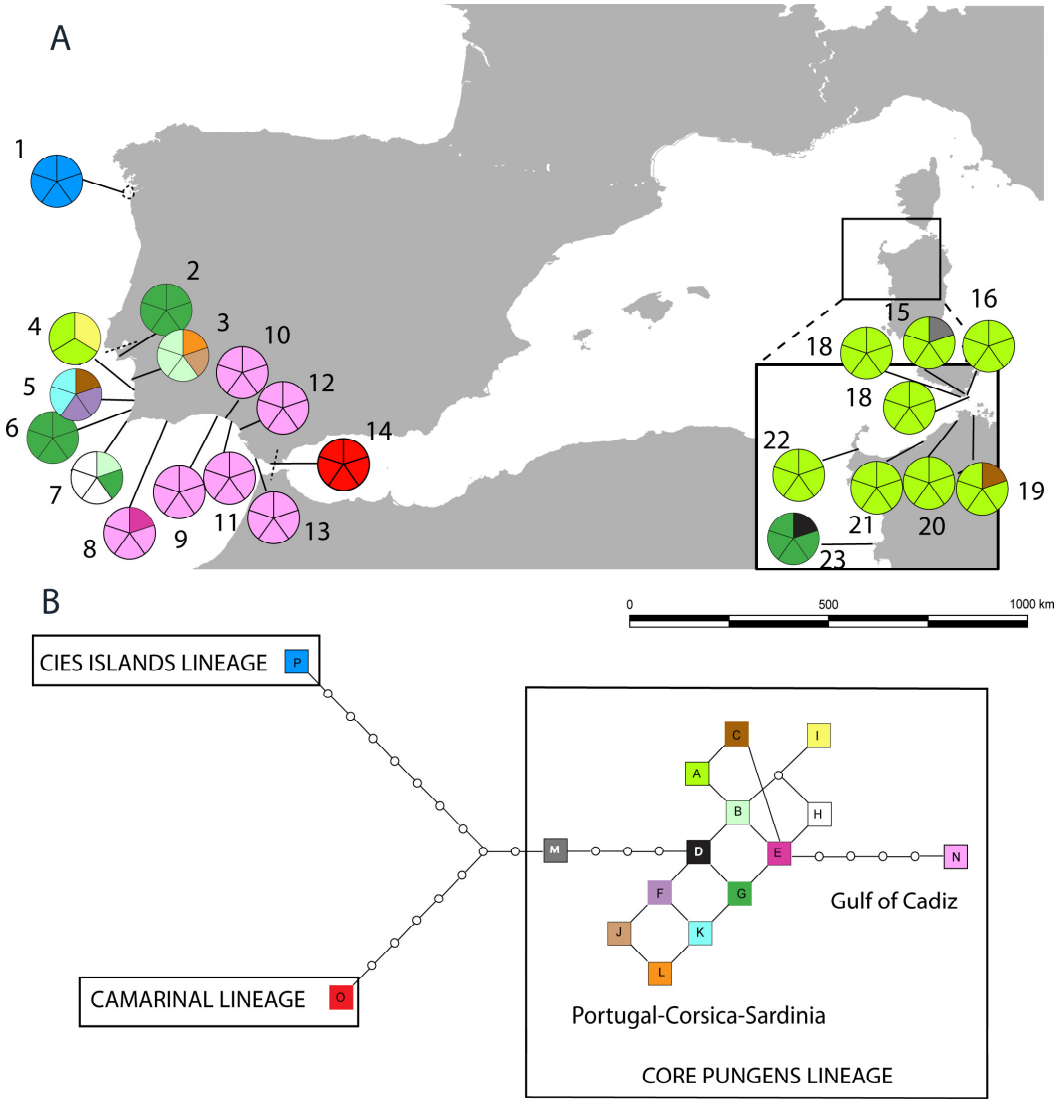


Table 3.6 Regional genetic diversity and differentiation in *Armeria pungens*.

Region	N indivs (pops)	A (N=28)	hs	ht	Gst	vs	vt	Nst
Portugal	28 (6)	10	0.48 (0.1522)	0.90 (0.1131)	0.47 (0.1066)	0.47 (0.1561)	0.90 (0.1668)	0.48 (0.0701)
Gulf of Cadiz	29 (6)	2	0	0	0	0	0	0
Corsica-Sardinia	45 (9)	4	0.09 (0.0588)	0.26 (0.1751)	0.66 (0.0648)	0.04 (0.0238)	0.27 (0.1947)	0.86 (0.0452)

N= sampling size
A = allelic richness corrected for sample size (N = 28) with rarefaction method of El Mousadik and Petit (1996)
hs = average within population diversity based on haplotype frequencies only.
ht = total genetic diversity based on haplotype frequencies only.
Gst = genetic differentiation based on haplotype frequencies only.
vs = average within population diversity based on haplotype frequencies and genetic distance between haplotypes.
vt = total genetic diversity based on haplotype frequencies and genetic distance between haplotypes.
Nst = genetic differentiation based on haplotype frequencies and genetic distance between haplotypes.

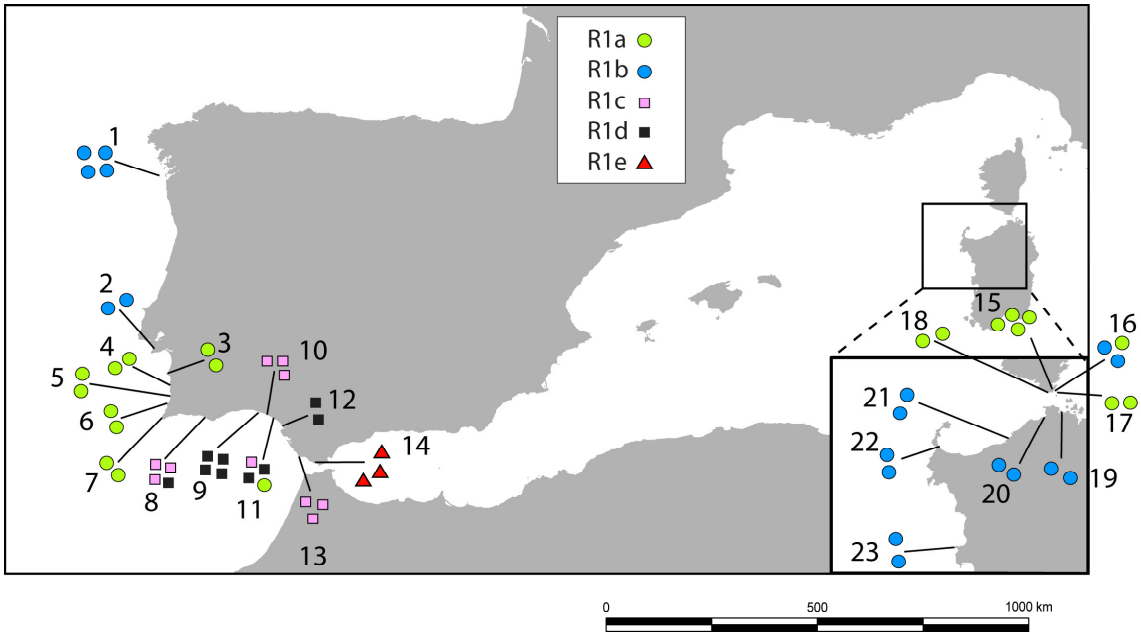
Table 3.7 Within-population cpDNA diversity in *Armeria pungens*. N = sample size; Hd = haplotype diversity; pi = nucleotide diversity.

Population	N	Hd	pi*1000
Cies Islands			
pop. 1	5	0.00	0.00
Portugal			
pop. 2	5	0.00	0.00
pop. 3	5	0.70	1.01
pop. 4	3	0.67	0.48
pop. 5	5	0.80	0.51
pop. 6	5	0.00	0.00
pop. 7	5	0.70	0.29
Gulf of Cadiz			
pop. 8	5	0.40	0.29
pop. 9	5	0.00	0.00
pop. 10	5	0.00	0.00
pop. 11	4	0.00	0.00
pop. 12	5	0.00	0.00
pop. 13	5	0.00	0.00
Camarinal			
pop. 14	5	0.00	0.00
Corsica-Sardinia			
pop. 15	5	0.40	
pop. 16	5	0.00	0.58
pop. 17	5	0.00	0.00
pop. 18	5	0.00	0.00
pop. 19	5	0.40	0.00
pop. 20	5	0.00	0.14
pop. 21	5	0.00	0.00
pop. 22	5	0.00	0.00
pop. 23	5	0.40	0.00
Total range	112	0.19	0.14

3.3.1.3 *ITS analysis within A. pungens*

Fifty nine nrITS sequences of *A. pungens* aligned according to Fuertes Aguilar *et al.* (1999) resulted in a matrix of 629 bp. No ITS variation fixed within genomes was found within *A. pungens*. However, five polymorphic sites, interpreted as intragenomic occurrence of more than one ITS sequence (ribotype) (Fuertes Aguilar *et al.*, 1999; Nieto Feliner and Rosselló, 2007), were detected (R1a, R1b, R1c, R1d, R1e; Table 3.7). These polymorphic positions showed a geographic structure (Fig.3.3). The southernmost population (pop. 14-Camarinal) presented two exclusive polymorphic sites in positions 380 and 448 (R1e), a divergent pattern that is supported by both morphology and cp variation (see below). The northernmost population (pop.1-Cies) shared a polymorphic site in position 99 (R1b) with the northernmost population from the mainland (pop.2-albufeira) as well as Sardinia (pops. 19-23) and two Corsican individuals (pop. 16). in agreement with similar morphology, but not supporting cp DNA divergence. Finally, in congruence with morphology and cpDNA, two polymorphic sites were restricted to the Gulf of Cadiz area (pops. 8-13): positions 62 and 392 (R1c and R1d), polymorphic in 16 and 9 individuals, respectively, out of 17 individuals sampled.

Figure 3.3 Geographical distribution of the nuclear ribosomal ITS sequence types within *Armeria pungens*, which differ only by polymorphic sites. Polymorphic sites defining sequence types R1a R1b R1c R1d and R1e are shown in Table 3.7.



3.3.2 Genetic variation above the species level

3.3.2.1 *cpDNA analysis across species*

The cpDNA matrix with 78 sequences across species had 3057 bp, with 94 variable sites, 23 of them parsimony informative. The strict consensus tree of the 18540 trees retained in the parsimony analysis is shown in Fig.3.4.

The strict consensus tree represents four major clades (labelled 1 to 4) which provide a picture of how the different genetic lineages within *A. pungens* (Fig. 3.2) relate to the ten congeners included in the study. Relationships between clades are not resolved since all of them lie in a polytomy. Additionally, 9 single terminals can be observed. Bootstrap supports of these four clades significantly decrease when indels are incorportated into the analysis, which suggests that above the species level indels show homoplasy. Clade 2 (70% bootstrap) encompasses the divergent *A. pungens* from the Cies islands ("Cies islands Lineage"), as well as two putative sympatric hybrids *A. pungens* X *A. pubigera*, and the two populations of *A berlengensis* (B1, B2) and three of *A. welwistchii* (W1,W2,W3, except for two individuals from population W1). The other congeneres from Cies (*A. pubigera* and *A. maritima*), together with the remaining putative hybrids *A. pungens* X *A. pubigera*, fell apart from *A. pungens* samples, as single terminals in polytomy with the four major clades. Clade 3 (95% bootstrap) includes the remaining population of *A. welwitschii*, W4. Clade 4 (69% bootstrap) is restricted to southern *Armeria*s, including divergent *A. pungens* from Camarinal ("Camarinal Lineage") as well as sequences from all species of *Armeria* in this area (*A. macrophylla, A. hirta, A. gaditana*). A close relationship between "Camarinal lineage" and its sympatric population of *A. macrophylla* is highly supported. Clade 1 corresponds to the "Core *A. pungens* lineage". Subclade 1a (93% bootstrap) comprises all six populations from the "Gulf of Cadiz" lineage. The remaining sequences of clade 1 form a polytomy to the subclade containing the "Portuguese-Corso-Sardinian" lineage of *A. pungens* together with populations R1 and R3 (*A. rouyana*), V1 (*A velutina*) and M1 (*A. macrophylla*), as well as single individuals from G1 (*A. gaditana*) and W1 (*A. welwitchsii*). One individual from Corsica is basal to clade 1.

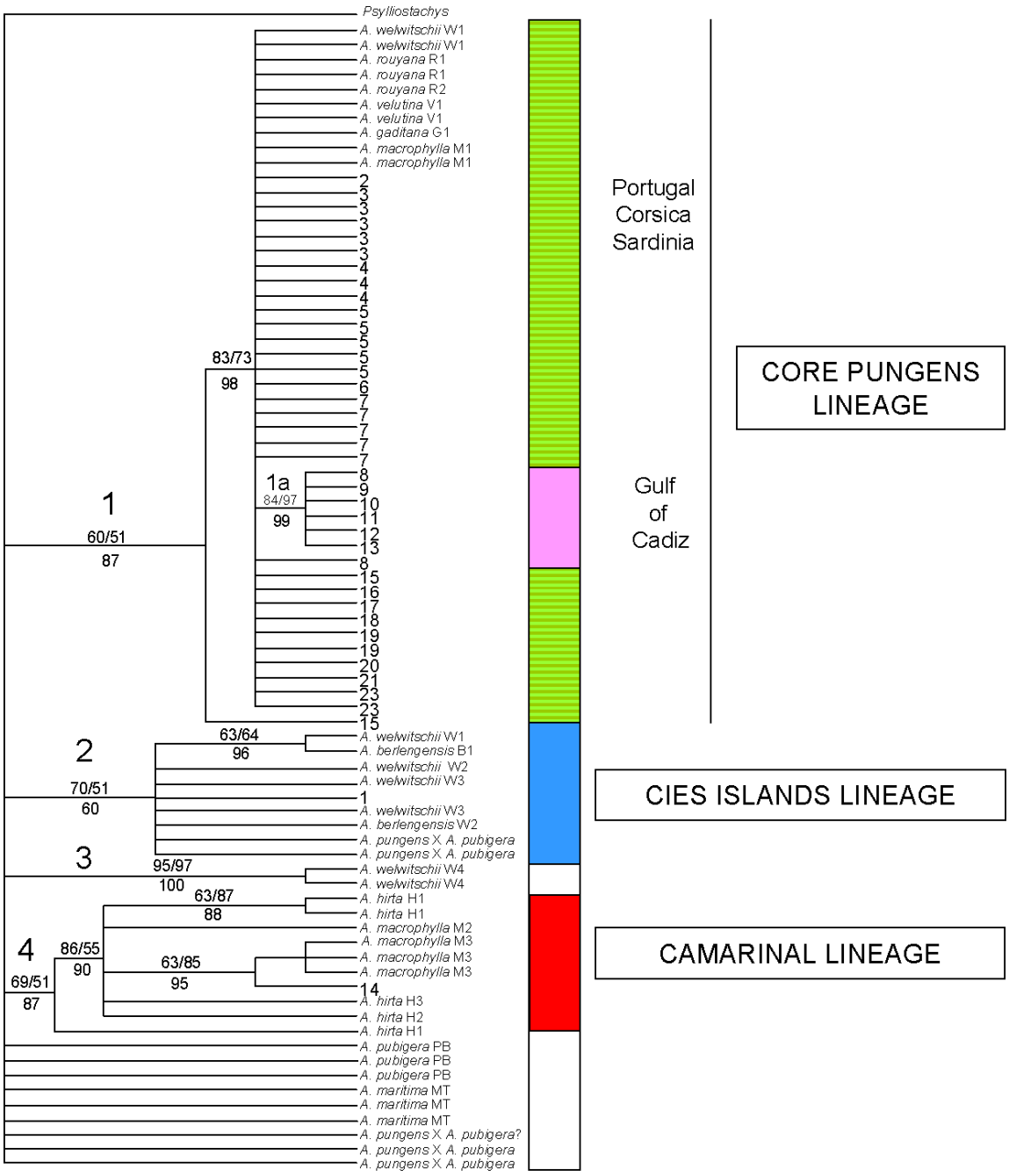
3.3.2.2 *ITS analysis across species*

The single ribotype found across the whole range of *A. pungens* corresponded to the widespread Western Iberian ribotype (R1) in *Armeria* (Fuertes Aguilar and Nieto Feliner, 2003; Table 3.7). In the Cies islands *A. pungens* showed a typical R1 variant, with a particular polymorphism in position 99, whereas putative hybrids *A. pungens* X *A. pubigera* showed an additive polymorphic site (APS) for all five parsimony informative positions between clades R1 and R5 plus a polymorphic site at position 99. *Armeria maritima* and *A. pubigera* also presented APS for the mentioned five informative positions. *Armeria pungens* from Camarinal showed a R1 with two particular polymorphic positions shared with parapatric *A. hirta* and *A. macrophylla*. Finally, *A. pungens* from Corsica and Sardinia showed a typical R1 and no sign of additivities with R5 were detected in the four parsimony informative sites distinguishing both clades.

Table 3.7 Nucleotide variation in nuclear ribosomal ITS region in *Armeria*. Polymorphic sites are coded according to the IUPAC code. Geographical distribution of polymorphic sites within *A. pungens* is represented in Fig. 3.3.

	62	99	236	303	380	381	390	391	392	398	403	415	416	417	448	520
ARMERIA SP. (Fuertes Aguilar and Nieto Feliner, 2003)																
clade R1 Western Iberian	T	C	C	C	T	C	C	T	T	T	G	C	G	G	C	A
clade R5a Corsica-Sardinia	.	.	.	.	.	.	.	C	C	.	.	.	T	T	.	.
clade R7 *maritima-alpina*	.	.	.	.	.	T	T	C	C	.	.	.	.	.	.	G
ARMERIA PUNGENS (N=59)																
R1a	T	C	C	C	T	C	C	T	T	T	G	C	G	G	C	A
R1b	.	Y	.	.	.	.	.	.	.	.	.	.	.	.	.	.
R1c	Y	.	.	.	.	.	.	.	Y	.	.	.	.	.	.	.
R1d	Y	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
R1e	.	.	.	.	Y	.	.	.	.	.	.	.	.	.	Y	.
SYMPATRICS pr. pops. 1 and 14 (N=10)																
A. pungens X *A. pubigera* Cies (pr. pop.1)	.	Y	.	.	.	Y	Y	Y	Y	.	.	.	.	.	.	R
A.maritima, A. pubigera Cies (pr. pop.1)	.	.	.	.	.	Y	Y	Y	Y	.	.	.	.	.	.	R
A. macrophylla (pr. pop 14, Camarinal)	.	.	.	.	Y	.	.	.	.	.	.	.	.	.	Y	.

Figure 3.4 Strict consensus of 18540 MP trees depicting the phylogenetic relationships between the chloroplast haplotyes of 23 populations of *A. pungens* and 20 populations of 10 coastal and subcoastal congeners. Bootstrap supports are indicated above branches based on two treatments of indels (missing data/presence-absence). Posterior probabilities of the Bayesian analysis are indicated below branches. The main *A. pungens* lineages are indicated according to Fig. 4.3 B. *A. pungens* individuals are labelled with population numbers as in Table 1.1A and Fig. 3.2. The name and population of origin of the other species follow Table 1.1B and Fig. 3.1.



3.4 DISCUSSION

3.4.1 Waves of expansion from Portugal into Corsica-Sardinia and the Gulf of Cadiz

The Portuguese region harbours the highest number of plastid DNA variants of the species, ten out of sixteen, but they are confined within the same lineage (Fig. 3.2; Gst=Nst, Table 3.6), which indicates long-term stability of populations. No reasons to suspect introgression from sympatric species into these populations are apparent from this genetic pattern. This seems to be the case even for Bordeira (pop. 6), which does show typical Portuguese plastid DNA (Fig. 3.2) and ITS variants (Fig. 3.3) despite exhibiting pubescent leaves and scapes, unlike the other conspecific populations (chapter 1).

In contrast to the stability of Portuguese populations, recent colonization associated to a founder event can be inferred from the current genetic structure of Corso-sardinian populations. Only five chloroplast haplotypes were found in the archipelago, three of them shared with Portugal: the dominant haplotype A, fixed for most of the populations, and the less frequent haplotypes C and G. This pattern evidences long-distance dispersal of the Portuguese lineage (chapter 2; Fig. 3.2), also consistent with the presence in Corsica-Sardinia of ciliate leaves (chapter 1) and typical Western Iberian ITS ribotype (R1). The latter suggests lack of nuclear gene flow from the remaining Corso-Sardinian congeners, which present a different ribotype (R5; Fuertes Aguilar and Nieto Feliner, 2003; Fig. 3.3; Table 3.7).

The plastid DNA sequences were especially informative in interpreting the history of the Gulf of Cadiz. Genealogical data reveal an ancestral rare haplotype (E) related to Southern Portugal and a frequent and exclusive haplotype (N) also related to such lineage but separated by several mutational steps (Fig. 3.3). This genealogy indicates the eastward linear colonization of the Gulf of Cadiz from southern populations of the Portuguese area (near Pop 7- Vicente). During process of colonization a severe founder effect associated to genetic drift probably took place leading to the loss of the four haplotypes missing between haplotypes N and E. Two ITS polymorphic sites (Fig. 3.3; Table 3.7) and glabrous to subglabrous leaves consistently differentiate this lineage with respect to the Portuguese one (chapter 1).

Possible alternative scenarios do not fit well the available evidence. In particular, since no differentiation among populations within the Gulf of Cadiz was detected in plastid DNA, ITS or morphology a scenario of long-term low population sizes is unlikely. Isolation of populations near the Gibraltar Strait and westward recolonization seems also unlikely because no sign of a contact zone can be observed in the border with the Portuguese lineage (pops. 6, 7, 8): neither intermediate morphology (chapter 1), nor mixture of plastid DNA haplotypes (Fig. 3.2) nor increased levels of nuclear (chapter 2) or plastid diversity (Table 3.7).

3.4.2 Hybridization at the margins of the distribution

3.4.2.1 *Putative Ancient hybrids in the Cíes islands*

The placement of plastid DNA sequences of *A. pungens* from the northernmost population in the Cies islands (pop. 1) apart from all other populations of the species in a clade together with *A. welwitschii* and *A. berlengesis* (Fig 3.4) was unexpected based on morphological similarity of *A. pungens* from Cíes with the conspecifics from the mainland, either considering quantitative characters or the presence ciliate leaves (chapter 1). In addition, this population shares a polymorphic site in ITS with the closest population of *A. pungens* from mainland Iberia (pop.2-albufeira), as well as with several Corso-Sardinian populations (Fig. 3.3). This confirms previous findings from AFLP data (chapter 2).

There is thus inconsistency between morphology and biparentally (AFLP and ITS) inherited markers on one side, which indicate similarity between *A. pungens* from Cíes and *A. pungens* from the mainland and maternally inherited plastid evidence, on the other, which relates *A. pungens* from Cíes with *A. welwitschii/berlengensis*. A likely explanation for this pattern is that *A. pungens* from the Cíes islands may have captured the chloroplast of the foreign *A. welwitschii/ berlengensis* lineage, a phenomenon not uncommon in plants (Rieseberg and Wendel, 1993; Arnold, 1997).

However, distinguishing stabilised introgression from incomplete sorting of ancestral polymorphisms is not an easy task (Wendel and Doyle, 1998). In ambiguous cases, independent evidence should be explored (Schaal *et al.*, 1998; Bittkay and Comes, 2005). In the genus *Armeria*, interspecific hybridization in natural populations is well-

documented, in most cases involving species that are currently sympatric (e.g. Bernis, 1954; Pinto da Silva, 1972; Nieto Feliner, 1990; Fuertes Aguilar *et al.*, 1999; Tauleigne-Gomes and Lefèbvre, 2005). However, at least in one case, past hybridization between currently allopatric species as a consequence of range shifts forced by Pleistocene climatic oscillations has been consistently documented in *Armeria* (Gutiérrez Larena *et al.*, 2002).

Based on the habitat similarities of *A. pungens* and *A welwitschi* (sandy coastal sites) and the dynamic history of Quaternary vegetations in Europe, it is conceivable that *A. welwitschii* showed a different northern geographic boundary in the past which favoured contacts of both species. In fact, a number of studies on different plant groups report historical introgression between currently allopatric species (Rieseberg and Wendel, 1993; Arnold *et al.*, 1999; Cronn and Wendel, 2004; Álvarez and Wendel 2006). However, the possibility of common ancestry with *A. welwitschii* cannot be ruled out.

Interpretation of the plastid phylogeny in terms of hybridization might be complicated by some degree of sorting of ancestral polymorphisms. Plastid haplotypes from two individuals of *A. welwitschii* are closely related to the core *A. pungens.* Surprisingly, they belong to the population pop W1, the most distant geographically from *A. pungens*. (Fig. 3.4). In addition, haplotypes of the southern populations of *A. velutina* (pop. V1), *A. macrophylla* (pop. M1) and *A. gaditana* (pop. G1) are also placed within the core A. pungens. Haplotypes shared across species in geographically apart locations are likely to represent ancestral polymorphisms. The populations of *A. rouyana* also exhibit haplotypes related to the core *A. pungens*. *A. rouyana* and *A. pungens* are currently sympatric and morphologically intermediate individuals between both species have been described (Bernis, 1954). However, testing for introgression *vs*. lineage sorting in this case would require a population-level sampling of chloroplast and nuclear markers of *A. rouyana*.

3.4.2.2 *Recent hybrids in Cies and Camarinal*

The hybrid origin of morphologically intermediate plants in the contact zone between *A. pungens* and *A. pubigera* in the Cies islands is strongly confirmed by genetic data, namely, a mixture of plastid haplotypes from both parents types (clades III and IV,

Fig. 3.4) in the hybrid zone and additive polymorphisms in ITS sequences (Table 3.7) that can be straightforwardly attributed to hybridization (western Iberian ribotype 1 and *martima-alpina* ribotype 7; Fuertes Aguilar and Nieto Feliner, 2003). The intermediate morphology and genetic composition suggests that there is ongoing hybridization in the Cies islands, where no signs of introgression of *A. pubigera* into *A. pungens* were found. Based on the spatial location of the hybrids in the ecotone between the habitats of the two progenitor species (sand dunes for *A. pungens*; granite cliffs for *A. pubigera*), this process can be well described as a hybrid zone.

Therefore, it is proposed that *A. pungens* from Cies islands has been involved in at least two different events of hybridization that have occurred at markedly different time periods and spatial scales: first with *A. welwitschii/berlengensis* lineage and recently with *A. pubigera*.

In the case of the Camarinal population available evidence is conclusive in revealing introgression from its congener *A. macrophylla*. The occurrence of a fixed exclusive plastid DNA haplotype (within *A. pungens*) shared with the sympatric *A. macrophylla* (Fig. 3.4) indicates that the latter has been the ovule donor (Nieto Feliner *et al.*, 2002). Furthermore, Camarinal population of *A. pungens* exhibits puberulent leaves and scapes (chapter 1) that are absent in any other conspecific population but are found in *A. macrophylla*. ITS sequence data are compatible with the scenario proposed in that the Camarinal population of *A. pungens* shares two polymorphic sites with *A. marcrophylla*.

3.4.3 Contrasting expanding mechanisms long-distance dispersal, linear colonization and hybridization in marginal areas

According to the current genetic structure, *A. pungens* seems to have an ancestral Atlantic lineage, represented by highly diverse populations in SW Portugal, that has undergone several waves of expansion. The substantial and rather clear phylogenetic signal in the plastid DNA data is puzzling in the context of the complex history of Mediterranean plant groups, in particular those with weak internal barriers (e.g. Albadalejo *et al.*, 2005). The linear –one dimensional- distribution of our study system has likely simplified the evolutionary processes involved.

Still, range expansions of *A. pungen*s were not uniform. On the one hand, plastid DNA and ITS data confirm that a subset of Portuguese genotypes reached Corsica-Sardinia through long-distance dispersal and successfully established there. On the other hand, in SW Iberia, linear colonization of the Gulf of Cadiz (from S. Vicent Cape to Gibraltar) has taken place from the former margin of the species pr. Pop. 7 (St. Vicente). Colonization apparently lead to fixation of new allele frequencies and loss of diversity by drift. Additionally, incongruence between plastid and nuclear DNA (together with morphology) suggests that *A. pungens* colonising the Cies islands has captured the plastid genome of a *A. welwitschii*/*berlenguensis* lineage. Finally, recent hybridization events have been detected in the northernmost and southernmost Iberian populations of *A. pungens*, involving the sympatric congeners *A. pubigera* (in the Cies islands) and *A. macrophylla* (in Camarinal).

3.4.3.1 *Role of founder events in population differentiation*

The depicted scenario of recent expansions makes unlikely long-term isolation as the main force driving the remarkable divergence of the Gulf of Cadiz lineage. Rather, differentiation of ITS and cpDNA and loss of chloroplast diversity seem to have been shaped by the severity of the founder event in this area. In contrast, the founder event associated to the establishment in Corsica-Sardinia was apparently less severe and did not promote genetic differentiation, only leading to loss of cpDNA diversity. AFLP revealing moderatee loss of nuclear diversity in Corsica-Sardinia (chapter 2) is not incongruent with this scenario, given the lower sensitivity to bottlenecks of the nuclear genome as compared to the chloroplast. Since a single chloroplast haplotype (haplotype A; Fig 3.2), shared with Portugal, is fixed for most Corso-Sardinian populations, there is no evidence of subsequent introductions in the archipelagothat could alternatively exlplain the lack of differentiation.

In a previous study using climatic envelope modelling, Portuguese and Corso-Sardinian locations of *A. pungens* were shown to have similar climatic conditions, with less severe summer droughts than the Gulf of Cadiz. We propose that the different environmental conditions in the Gulf of Cadiz may have intensified the colonising founder effect by selecting fit genotypes from all the diversity of source Portuguese populations. Heavy metal tolerant genotypes of *A. maritima* are known to have rapidly diverged under selection pressures from non-tolerant neighbouring populations, despite

substantial gene flow between them (Vekemas and Lefèbvre, 1997). In contrast, Portuguese individuals may have already been adapted to similar conditions in Corsica-Sardinia when they reached the archipelago.

Therefore, as a result of the preadaptation of dispersed genotypes, no divergence was associated to the colonization of Corsica-Sardinia, whereas in the Gulf of Cadiz, range expansion meant evolutionary change, probably promoted by a severe founder event in novel environmental conditions.

3.4.3.2 *Expansion through hybridization in marginal areas*

The past and ongoing hybridization processes reported in this study at both (northern and southern) ends of the Iberian range of *A. pungens* have generated evolutionary change probably allowing for range expansion. This income of genetic variation, resulting from hybridization, may be fit for novel or intermediate niches. In *Armeria*, this natural process is facilitated by low internal reproductive barriers demonstrated in artificial crossing experiments (Baker, 1966; Lefèbvre, 1974; Nieto Feliner *et al.*, 1996; Nieto Feliner, 1997; Vekemans and Lefèbvre, 1997).

Data from *A. pungens* do not imply that all introgression events are adaptive in *Armeria*. However, the whole scenario is better explained in adaptive terms. This seems to be also the case for the Southern Spanish endemic *A villosa,* which has five geographically restricted subspecies, as well as one subspecies spanning the whole distributional range. Three of the local subspecies have been suggested to result from introgression of the widespread one with other sympatric congeners either in novel habitats (ssp. *carratracensis* in serpetines, ssp. *bernisii* at higher altitudes) or in marginal areas (ssp. *alboi*) (Fuertes Aguilar *et al.* 1999; Nieto Feliner *et al.* 2001; Gutiérrez Larena *et al.* 2002).

3.5 REFERENCES

Abbott RJ (1992) Plant invasions, interspecific hybridization and the evolution of new plant taxa. *Trends in Ecology & Evolution* **7**, 401-405.

Abbott CL, Double MC (2003) Genetic structure, conservation genetics and evidence of speciation by range expansion in shy and white-capped albatrosses. *Molecular Ecology* **12**, 2953-2962.

Abbott RJ, James JK, Milne RI, Gillies ACM (2003) Plant introductions, hybridization and gene flow. *Philosophical Transactions of the Royal Society B: Biological Sciences* **358**, 1123-1132.

Albaladejo RG, Fuertes Aguilar J, Aparicio A, Nieto Feliner G (2005) Contrasting nuclear-plastidial phylogenetic patterns in the recently diverged Iberian *Phlomis crinita* and *P. lychnitis* lineages (Lamiaceae). *Taxon* **54**, 987-998.

Albert V, Jonsson B, Bernatchez L (2006) Natural hybrids in Atlantic eels (*Anguilla anguilla*, *A. rostrata*): evidence for successful reproduction and fluctuating abundance in space and time. *Molecular Ecology* **15**, 1903-1916.

Álvarez I, Wendel J (2006) Cryptic interspecific introgression and genetic differentiation within *Gossypium* aridum (Malvaceae) and its relatives. *Evolution* **60**, 505-517.

Anderson E, Stebbins JR (1954) Hybridization as an evolutionary stimulus. *Evolution* **8**, 378-288.

Arnold ML (1997) *Natural Hybridization and Evolution.* Oxford University Press, Oxford.

Arnold M (2004) Transfer and origin of adaptations through natural hybridization: were Anderson and Stebbins right? *The Plant cell* **16**, 562-570.

Arnold ML, Bulger MR, Burke JM, Hempel AL, Williams JH (1999) Natural Hybridization: how low can you go and still be important? *Ecology* **80**, 371-381.

Baker HG (1966) The evolution, functioning and breakdown of heteromorphic incompatibility systems, I. The Plumbaginaceae. *Evolution* **20**, 349–368.

Baker A, Moeed A (1987) Rapid genetic differentiation and founder effect in colonizing populations of common mynas *Acridotheres tristis*. *Evolution* **41**, 525–538.

Barton N, Charlesworth B (1984) Genetic revolutions, founder effects, andsSpeciation. *Annual Review of Ecology and Systematics* **15**, 133.

Bernis F (1954) Revisión del género *Armeria* Willd. con especial referencia a los grupos ibéricos. *Analales del Instituto Botánico Cavanilles* **11**, 5–288.

Bittkau C, Comes HP (2005) Evolutionary processes in a continental island system: molecular phylogeography of the Aegean *Nigella arvensis* alliance (Ranunculaceae) inferred from chloroplast DNA. *Molecular Ecology* **14**, 4065-4083.

Choler P, Erschbamer B, Tribsch A, Gielly L, Taberlet P (2004) Genetic introgression as a potential to widen a species' niche: Insights from alpine *Carex curvula*. *Proceedings of the National Academy of Sciences* **101**, 171-176.

Chung S, Staub JE (2003) The development and evaluation of consensus chloroplast primer pairs that possess highly variable sequence regions in a diverse array of plant taxa. *Theoretical and Applied Genetics* **107**, 757-767.

Clegg SM, Degnan SM, Kikkawa J*, et al.* (2002) Genetic consequences of sequential founder events by an island-colonizing bird. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 8127-8132.

Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology* **9**, 1657-1659.

Cronn R, Wendel JF (2004) Cryptic trysts, genomic mergers, and plant speciation. *New Phytologist* **161**, 133-142.

Demesure B, Sodzi N, Petit R (1995) A set of universal primers for amplification of polymorphic non-coding regions of mitochondrial and chloroplast DNA in plants. *Molecular Ecology* **4**, 129-131.

DeYoung R, Demarais S, Honeycutt R*, et al.* (2003) Genetic consequences of white-tailed deer (*Odocoileus virginianus*) restoration in Mississippi. *Molecular Ecology* **12**, 3237–3252.

Dover G (1982) Molecular drive: a cohesive mode of species evolution. *Nature* **299**, 111-117.

Ehrlich P, Raven P (1969) Differentiation of populations. *Science* **165**, 1228-1232.

Ellstrand NC (2003) Current knowledge of gene flow in plants: implications for transgene flow. *Philosophical Transactions of the Royal Society B: Biological Sciences* **358**, 1163-1170.

Ellstrand NC, Schierenbeck KA (2000) Hybridization as a stimulus for the evolution of invasiveness in plants? *Proceedings of the National Academy of Sciences of the United States of America* **97**, 7043-7050.

El Mousadik A, Petit R (1996) High level of genetic differentiation for allelic richness among populations of the argan tree [*Argania spinosa* (L.) Skeels] endemic to Morocco. *Theoretical and Applied Genetics* **92**, 832–839.

Francisco-Ortega J, Fuertes-Aguilar J, Gomez-Campo C, Santos-Guerra A, Jansen RK (1999) Internal transcribed spacer sequence phylogeny of *Crambe* L. (Brassicaceae): molecular data reveal two Old World disjunctions. *Molecular Phylogenetics and Evolution* **11**, 361-380.

Francisco-Ortega J, Santos-Guerra A, Kim S-C, Crawford DJ (2000) Plant genetic diversity in the Canary Islands: a conservation perspective. *American Journal of Botany* **87**, 909-919.

Friar E, Landoux T, Roalson E, Robichaux R (2000) Microsatellite analysis of a population crash and bottleneck in the Mauna Kea silversword, *Argyroxiphium sandwicense* ssp. *sandwicense* (Asteraceae), and its implications for reintroduction. *Molecular Ecology* **9**, 2027–2034.

Fuertes Aguilar J, Nieto Feliner G (2003) Additive polymorphisms and reticulation in an ITS phylogeny of thrifts (*Armeria*, Plumbaginaceae). *Molecular Phylogenetics and Evolution* **28**, 430-447.

Fuertes Aguilar J, Rossello JA, Nieto Feliner G (1999) Molecular evidence for the compilospecies model of reticulate evolution in *Armeria* (Plumbaginaceae). *Systematic Biology* **48**, 735-754.

Goudet J (2001) *FSTAT, a Program to Estimate and Test Gene Diversities and Fixation Indices.* Available from http://www.unil.ch/izea/softwares/fstat.html. Updated from Goudet (1995).

Grant V (1981) *Plant Speciation.* Columbia University Press, New York.

Gutiérrez Larena B, Fuertes Aguilar J, Nieto Feliner G (2002) Glacial-induced altitudinal migrations in *Armeria* (Plumbaginaceae) inferred from patterns of chloroplast DNA haplotype sharing. *Molecular Ecology* **11**, 1965-1974.

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis. *Nucleic Acids Symposium Series* **41**, 95-98.

Hausner G, Olson R, Simon D*, et al.* (2006) Origin and evolution of the chloroplast *trn*K (*mat*K) intron: a model for evolution of group II intron RNA structures. *Molecular Biology and Evolution* **23**, 380-391.

Hawley DM, Hanley D, Dhondt AA, Lovette IJ (2006) Molecular evidence for a founder effect in invasive house finch (*Carpodacus mexicanus*) populations experiencing an emergent disease epidemic. *Molecular Ecology* **15**, 263-275.

Hilu K, Liang H (1997) The *mat*K gene: sequence variation and application in plant systematics. *American Journal of Botany* **84**, 830-839.

Kelchner SA (2000) The evolution of non-coding chloroplast DNA and its application in plant systematics. *Annals of the Missouri Botanical Garden* **87**, 482-498.

Lavergne S, Molofsky J (2007) From the Cover: Increased genetic variation and evolutionary potential drive the success of an invasive grass. *Proceedings of the National Academy of Sciences* **104**, 3883-3888.

Lefèbvre C (1974) Relations génétiques entre des populations d'*Armeria maritima* (Mill.) Willd. *Comptes Rendus Académie Sciences Paris Série D* **278**, 1573-1576.

Lihovà J, Fuertes Aguilar J, Marhold K, Nieto Feliner G (2004) Origin of the disjunct tetraploid *Cardamine amporitana* (Brassicaceae) assessed with nuclear and chloroplast DNA sequence data. *American Journal of Botany* **91**, 1231-1242.

Mallet J (2005) Hybridization as an invasion of the genome. *Trends in Ecology & Evolution* **20**, 229-237.

Mayr E (1942) *Systematics and the Origin of Species.* Columbia University Press, New York.

McCauley D (1995) The use of chloroplast DNA polymorphism in studies of gene flow in plants. *Trends in Ecology & Evolution* **10**, 198-202.

Moody ML, Les DH (2002) Evidence of hybridity in invasive watermilfoil (*Myriophyllum*) populations. *Proceedings of the National Academy of Sciences* **99**, 14867-14871.

Nieto Feliner G (1990) *Armeria*. In: *Flora Iberica* (eds. Castroviejo S, Laínz M, López G*, et al.*), p. 642–721. Consejo Superior de Investigaciones Científicas, Madrid.

Nieto Feliner G (1997) Natural and experimental hybridization in *Armeria* (Plumbaginaceae): A. salmantica. *International Journal of Plant Sciences* **158**, 585–592.

Nieto Feliner G, Fuertes Aguilar J, Rosselló JA (2001) Can extensive reticulation and concerted evolution result in a cladistically structured molecular data set? *Cladistics* **17**, 301–312.

Nieto Feliner G, Fuertes Aguilar J, Rossello JA (2002) Reticulation or divergence: the origin of a rare serpentine endemic assessed with chloroplast, nuclear and RAPD markers. *Plant Systematics and Evolution* **231**, 19-38.

Nieto Feliner G, Izuzquiza A, Lansac AR (1996) Natural and experimental hybridization in *Armeria* (Plumbaginaceae): A. *villosa* subsp. *carratracensis*. *Plant Systematics and Evolution* **201**, 163–177.

Nieto Feliner G, Rossello JA (2007) Better the devil you know? Guidelines for insightful utilization of nrDNA ITS in species-level evolutionary studies in plants. *Molecular Phylogenetics and Evolution* **en prensa**.

Nylander JAA (2003) MrModeltest 1.0b. A Simplified Version of David Posada's "Modeltest 3.06". Uppsala, Sweden: Department of Systematic Zoology, http://www.ebc.uu.se/systzoo/staff/nylander.html.

Petit RJ, Bodenes C, Ducousso A, Roussel G, Kremer A (2004) Hybridization as a mechanism of invasion in oaks. *New Phytologist* **161**, 151-164.

Petit RJ, Duminil J, Fineschi S*, et al.* (2005) Comparative organization of chloroplast, mitochondrial and nuclear diversity in plant populations. *Molecular Ecology* **14**, 689-701.

Pinto da Silva A (1972) *Armeria*. In: *Flora Europaea* (eds. Tutin TG, Heywood VH, Burges NA*, et al.*), p. 30–38. Cambridge University Press, Cambridge.

Pons O, Petit RJ (1996) Measuring and testing genetic differentiation with ordered versus unordered alleles. *Genetics* **144**, 1237-1245.

Provan J, Powell W, Hollingsworth PM (2001) Chloroplast microsatellites: new tools for studies in plant ecology and evolution. *Trends in Ecology & Evolution* **16**, 142-147.

Pruett CL, Winker K (2005) Northwestern song sparrow populations show genetic effects of sequential colonization. *Molecular Ecology* **14**, 1421-1434.

Rieseberg LH, Wendel JF (1993) Introgression and its consequences in plants. In: *Hybrid Zones and the Evolutionary Process* (ed. Harrison R), pp. 70-109. Oxford University Press, New York.

Ronquist F, Huelsenbeck JP (2003) MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**, 1572-1574.

Rozas J, Sánchez-DelBarrio JC, Messeger S, Rozas R (2003) DNA polymorphisms analysis by the coalescent and other methods. *Bioinformatics* **19**, 2496-2497.

Sang T (2002) Utility of low-copy nuclear gene sequences in plant phylogenetics. *Critical Reviews in Biochemistry & Molecular Biology* **37**, 121-147.

Schaal BA, Hayworth DA, Olsen KM, Rauscher JT, Smith WA (1998) Phylogeographic studies in plants: problems and prospects. *Molecular Ecology* **7**, 465-474.

Swofford D (2002) *Phylogenetic Analysis Using Parsimony (and other Methods), Version 4*. Sinauer Associates, Sunderland MA.

Taberlet P, Gielly L, Pautou G, Bouvet J (1991) Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Molecular Biology* **17**, 1105-1109.

Tauleigne-Gomes C, Lefèbvre C (2005) Natural hybridization between two coastal endemic species of Armeria (Plumbaginaceae) from Portugal. 1. Populational in situ investigations. *Plant Systematics and Evolution* **250**, 215-230.

Templeton AR (1980) The theory of speciation via the founder principle. *Genetics* **94**, 1011-1038.

Templeton AR (1996) Experimental evidence for the genetic-transilience model of speciation. *Evolution* **50**, 909-915.Thompson JD (2005) *Plant evolution in the Mediterranean* University of Oxford Press, Oxford.

Tiébré M-S, Vanderhoeven S, Saad L, Mahy G (2007) Hybridization and sexual reproduction in the invasive alien *Fallopia* (Polygonaceae) complex in Belgium. *Annals of Botany* **99**, 193-203.

Tregenza T (2002) Divergence and reproductive isolation in the early stages of speciation. *Genetica* **116**, 291-300.

Vekemans X, Lefèbvre C (1997) On the evolution of heavy-metal tolerant populations in *Armeria maritima*: evidence from allozyme variation and reproductive barriers. *Journal of Evolutionary Biology* **10**, 175-191.

Weising K, Gardner RC (1999) A set of conserved PCR primers for the analysis of simple sequence repeat polymorphisms in chloroplast genomes of dicotyledonous angiosperms. *Genome* **42**, 9-19.

Wendel JF, Doyle JJ (1998) Phylogenetic incongruence: window into genomes history and molecular evolution. In: *Molecular Systematics of Plants II. DNA Sequencing* (eds. D S, Soltis P, Doyle JJ). Kluwer Academy Publications, Boston.

Winkworth RC, Robertson AW, Ehrendorfer F, Lockhart PJ (1999) The importance of dispersal and recent speciation in the flora of New Zealand. *Journal of Biogeography* **26**, 1323-1325.

4. USE OF LOW-COPY NUCLEAR GENE *GapC* IN PHYLOGEOGRAPHY OF *Armeria pungens*

4.1 INTRODUCTION

During the last decade, low-copy nuclear genes have been regarded as promising markers for phylogeographic studies (Schaal *et al.*, 1998, Schaal and Olsen, 2000; Hare, 2001; Sang, 2002). Their main reported advantages as compared to ITS and cpDNA markers are the potential high levels of variation, biparental inheritance with little concerted evolution and the possibility to compare unlinked genes of the nuclear genome (Sang, 2002; Small *et al.*, 2004). However, as examples using nuclear markers at the intraspecific level have accumulated, lack of coalescence of nuclear alleles from a single species has been observed to be a relatively common phenomenon (Olsen and Schaal, 1999; Comes and Abbott, 2001; Isoda and Shiraishi, 2001; Caicedo and Schaal, 2004; Small *et al.*, 2004, McKinnon *et al.*, 2005; Bartish *et al.*, 2006; Friar *et al.*, 2006; Syring *et al.*, 2007), which might seriously limit the utility of low-copy nuclear markers for phylogeographic studies (Maddison and Knowles, 2006).

Processes causing polyphyletic patterns of sequences from the same species may have different origins such as imperfect taxonomy, introgressive hybridization, paralogy or incomplete lineage sorting of ancestral polymorphisms (see Funk and Omland, 2003 for a review).

Occasional interspecific hybridization between otherwise distinct species and subsequent gene flow towards one of them, may lead to introgression of alleles from one species into the genome of another. The heterogeneous origin of the alleles will be reflected in a polyphyletic gene tree (Doyle *et al.*, 1990; Small *et al.*, 2004; Vriesendorp and Bakker, 2005). This process can be clearly detected when the phylogenetic signal is congruent across independent nuclear markers or sympatric populations of different species. But it becomes more difficult to detect if it is an old introgression event (Rieseberg and Ellstrand, 1993).

Random fixation of ancestral polymorphic alleles through multiple speciation events causing incomplete lineage sorting is especially challenging at low taxonomic ranks (Comes and Abbott, 2001; Sang, 2002). Within a recently diverged species each allele has often a particular history resulting in polyphyletic allele lineages (Sang, 2002). This phenomenon is random, so it will affect differently to unlinked markers. It is also far more frequent in nuclear genes as compared to the chloroplast due to the higher effective population size of the diploid and biparentally inherited nuclear genome (Sang 2002; Funk and Omland, 2003). Over time, new alleles will be created by mutation whereas coalescence will be eventually reached through extinction of certain allele lineages by drift (Funk and Omland, 2003). In theory, time to monophyly in a nuclear gene in a diploid organism is expected to be four-fold as compared to a chloroplast marker (Palumbi *et al.*, 2001; Hare, 2001).

Additionally, given that nuclear genes tend to evolve in multigene families, unnoticed amplification of paralogous copies (Wendel and Doyle, 1998) may be retrieved when duplication is followed by random deletion in some of the descendent lineages. The resulting gene tree may depict polyphyletic relationships among paralogs. But within the species level a complete duplication/deletion cycle is extremely rare to occur and the number of paralog copies of terminals of the same species is expected to be constant. Therefore, under the species level, mixing of paralogs is less of a concern but it may still arise as a result of failed sampling (while picking clones or through PCR failure; Sang, 2002) and random amplification of different paralogs in different terminals. To avoid mixing of paralog copies it is important to characterise first the gene family in a particular species.

A handicap when assessing the copy number of a nuclear gene is that it cannot be extrapolated from previous knowledge on other taxa but it has to be assessed empirically. The number of paralogs may be extremely different in close lineages and thus has to be determined for every single species by cloning and phylogenetic evaluation (Small and Wendel, 2000; Hughes *et al.*, 2006).

For not closely related organisms the best test of orthology is congruence with other markers in phylogenetic patterns within each of the putative orthologs (Patterson, 1988). But working at intraspecific levels this is difficult to observe, and one has to look at less definitive approaches such as chromosomal position through gene mapping (available for model organisms), degree of divergence among sequence types, differences in size or gene coding patterns among types, maximum number of sequence types retrieved per individual in relation to the ploidy level, congruence with expected heterozygosity levels or southern blotting with locus specific hybridization probes (Hare, 2001; Small *et al.*, 2004).

When undetected, the above described polyphyletic gene patterns may produce well supported yet incorrect phylogenetic relationships. The gene tree is thus assumed to reflect the species tree, compromising phylogeographic inference (Wendel and Doyle, 1998). However, if introgression, incomplete lineage sorting or paralogy are taken into account the phylogeographic signal within a genetic nuclear marker may be retrieved by separating the variation attributed to these processes, from the variation reflecting true phylogeographic events such as founder effects, fluctuation in population sizes, colonisations of new areas, etc (Sang, 2002). This can be achieved by separating ancestral from more recent variation in the case of random sorting of ancestral polymorphisms or by treating each copy type as an independent marker in the case of having amplified several paralogs (Hare, 2001; Caicedo and Schaal, 2004; Maddison and Knowles, 2006).

The objetive of the present study is to infer the phylogeographic history of the sand-dune shrub *Armeria pungens* using low copy gene, *GapC*. We discuss the constraints for phylogeographic inference based on this nuclear marker and compare it with other molecular markers and morphology.

Table 4.1 Collection data of 22 populations of *Armeria pungens* included in the phylogeographic study based on *GapC*. N=sampling size

Population code	Site location, collector (year)	N
1- Cíes	Spain, Pontevedra, Illas Cíes, RPP, IMA, IMF, LMR, MSA & PGM (2003)	6
2- Albufeira	Portugal, Estremadura, Lagoa de Albufeira, GNF & JFA (2002) and RPP, AC & PEG (2003)	3
3- Sines	Portugal, Baixo Alentejo, Sines-Cercal, RPP, AC & PEG (2003)	3
4- Furnas	Portugal, Baixo Alentejo, Vilanova de Milfontes, Praia das Furnas, RPP, AC & PEG (2003)	3
5- Zambujeira	Portugal, Baixo Alentejo, Zambujeira do Mar, GNF & JFA (2002) and RPP, AC & PEG (2003)	5
6- Bordeira	Portugal, Algarve, Praia do Bordeira, GNF & JFA (2002)	3
7- S. Vicente	Portugal, Algarve, Cabo de São Vicente, GNF & JFA (2002) and RPP, AC & PEG (2003)	3
8- Garrrao	Portugal, Algarve, Praia de Garrao, GNF & JFA (2002) and RPP, AC & PEG (2003)	4
9- P. Umbría	Spain, Huelva, Punta Umbría, Playa de Punta Umbría, GNF & JFA (2002)	3
10- C. Maneli	Spain, Huelva, Mazagón-Matalascañas, "cuesta Maneli", GNF & JFA (2002)	3
11- Matalascañas	Spain, Huelva, Torre de la Higuera, Playa de Torre la Higuera, GNF & JFA (2002) and RPP, AC & PEG (2003)	3
12- Doñana	Spain, Huelva, Doñana National Park, pr. El Inglesillo, SG (2002)	2
13- Trafalgar	Spain, Cadiz, Cabo de Trafalgar, GNF & JFA (2002) and RPP, AC & PEG (2003)	4
14- Camarinal	Spain, Cadiz, Punta Camarinal, GNF & JFA (2002)	4
15- Piantarella	France, Corsica, Bonifacio, Piantarella, GNF & JFA (2002)	3
16- P. Sperone	France, Corsica, Bonifacio, Petit Sperone, CBNMP(2002)	3
17- I. Piana	France, Corsica, Bonifacio, île Piana, CBNMP (2002)	-
18- G. Sperone	France, Corsica, Bonifacio, Grand Sperone, CBNMP (2002)	1
19- P. Liscia	Italy, Sardinia, Porto Pozzo, Porto Liscia, GNF & JFA (2002)	3
20- R. Maiore	Italy, Sardinia, S of Santa. Teresa di Gallura, Spiaggia di Rena Maiore, GNF & JFA (2002)	3
21- B. Mare	Italy, Sardinia, Badesi Mare, GNF & JFA (2002)	3
22- Stintino	Italy, Sardinia, Pr. Porto Torres, Stintino, S Stagno di Pilo, GNF & JFA (2002)	3
23- C. Mannu	Italy, Sardinia, NW of Oristano, Capo Mannu, between the cape and Su pallosu, GNF & JFA (2002)	3

Abbreviations of collectors: AC, A. Costa; CBNMP, Conservatoire Botanique National Mediterranéen de Porquerolles; GNF, G. Nieto Feliner; IMA, I. Martínez Arcos; IMF, I. Martínez Fernández; JFA, J. Fuertes Aguilar; LMR, L. Muriel Ríos; MSA; M. Souto Alonso; PEG, P. Escobar García; PGM, P. García Meijide; RPP, R. Piñeiro Portela; SG, S. Gatelier.

4.2 MATERIALS AND METHODS

4.2.1 Sampling

Leaves and ripe fruits from 73 individuals were collected in 22 populations spanning the whole geographical range of *A. pungens* (Table 4.1). Leaves were dried with silica gel and seeds were grown in the Botanical Garden of Madrid in order to produce fresh leaves. DNA was isolated from dry or fresh leaves using DNeasy®Plant Mini Kit (Qiagen), following manufacturer´s instructions.

4.2.2 Design and selection of locus-specific primers

For amplification of *GapC* (cytosolic glyceraldehyde 3-phosphate dehydrogenase) region, preliminary amplifications were performed with primer GPDX9R from Strand *et al.* (1997) and consensus primers (GP5F, GP2F, GP4R) designed from GenBank sequences comparing *Antirrhinum majus*, X59517; *Arabidopsis thaliana*, M64119; *Atriplex nummularia*, U02886; *Mesembryanthemum crystallinum*, J05223; *Ranunculus acris*, X60345). Once obtained the first sequences from *A. pungens*, specific primers (A6F, A1F, A7R, A4R, A8R) were designed using Primer3 Software, (Whitehead Institute for Biomedical Research). Position and description of sequences and annealing temperatures of the nine primers used are indicated in Fig. 4.1 and Table 4.2.

Table 4.2 Description of the nine *GapC* primers tested in *A. pungens*

Primer forward	Sequence (5'-3')	TmºC
GP5F	5'-TGC TCC CAT GTT TGT TGT CG-3'	57.3
GP2F	5'-GCA TTG TTG AGG GTC TTA TG-3'	55.3
A6F	5'-CAA TTG CCT TGC ACC TTT G-3'	54.5
A1F	5'-GCT GCC TCC TTC AAC ATC AT-3'	57.3
Primer reverse		
GPDX9R	5'-AAG CAA TTC CAG CCT TGG-3'	53.7
GP4R	5'-ATG TGG CGG ATC AAG TCA AC-3'	57.3
A7R	5'-GCA ACA CTT TGC CAA CAG C-3'	56.7
A4R	5'-TTC CCT TCA TCT TGC CAT TC-3'	55.3
A8R	5'-AAC ACG GGT GCT GCA ATC-3'	56.0

Table 4.3. Primer combinations tested for amplification of *GapC* in *A. pungens*. The copy paralagos amplified by each pair is indicated.

Primer combination	*GapC* **Copy**
GP5F- GPDX9R	1
A6F- A7R	1
GP2F- GP4R	2
GP2F- A4R	2
A1F-A8R	2
A1F- A4R	2
GP2F- GPDX9R	3

Seven combinations of the above primers were tested on three samples of *A. pungens* from divergent lineages (pops. 5, 9 and 19). Ninety three PCR products were cloned and sequenced. All seven primers amplified a partial DNA sequence of the *GapC* gene between exons 6 and 9. Three different *GapC* copies were recognisable: 1, 2 and 3. Their structure is depicted in Fig. 4.1, and primer pairs amplifying each copy are indicated in Table 4.3.

The three *GapC* copies were significantly different both in length (Table 4.4) and above all in nucleotide composition. The introns could not be aligned, whereas exons were easily aligned but coded for different proteins (results not shown). Given the high levels of divergence, copies 1, 2 and 3 seem to be different paralogs.

Figure 4.1 Scheme of the three *GapC* paralog copies 1, 2 and 3 found in *A. pungens.* Exons are represented by grey boxes and introns by lines. Nucleotide positions and exon numbers are referred to the Gene Bank sequence U02886.2 of *Atriplex nummularia*. The position of the primers used in this study are indicated: A) the universal primer GPDX9R (Strand *et al.*, 1997) and the primers designed from gene bank (GP5F, GP2F, GP4R) are placed on the *GapC* scheme for *Atriplex nummularia*; B) the specific primers for *Armeria* (A1F, A6F, A4R, A7R and A8R) are placed on the *GapC* schemes of the paralog copies of *A. pungens*.

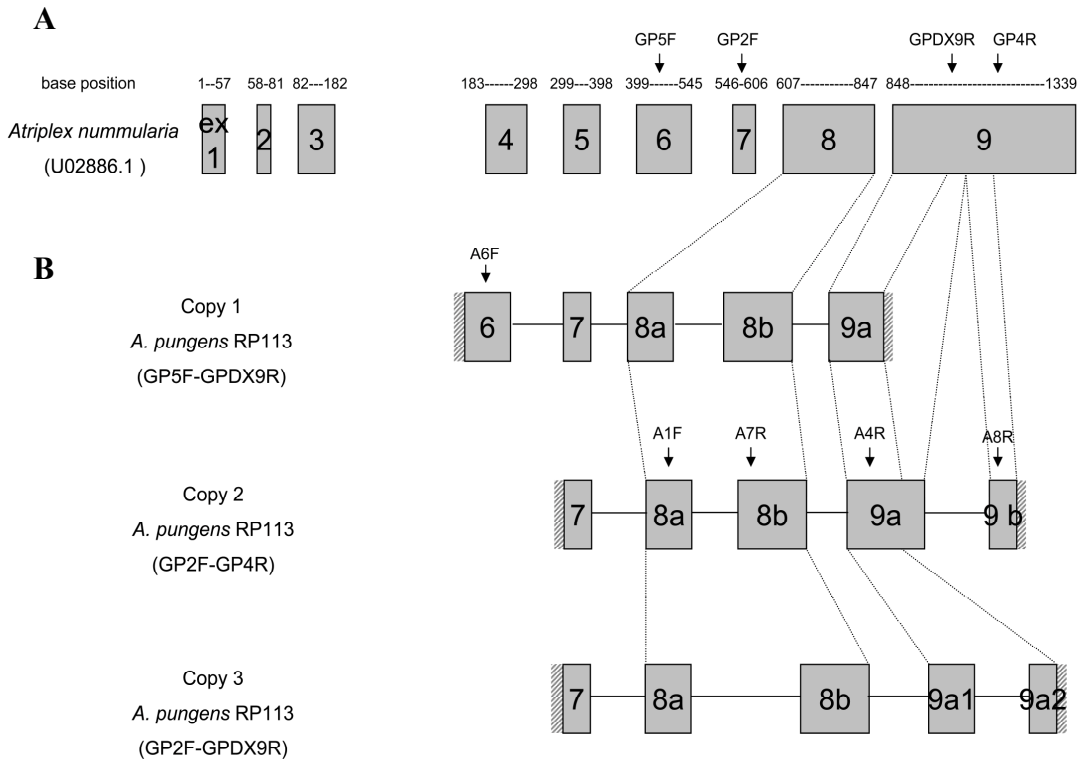
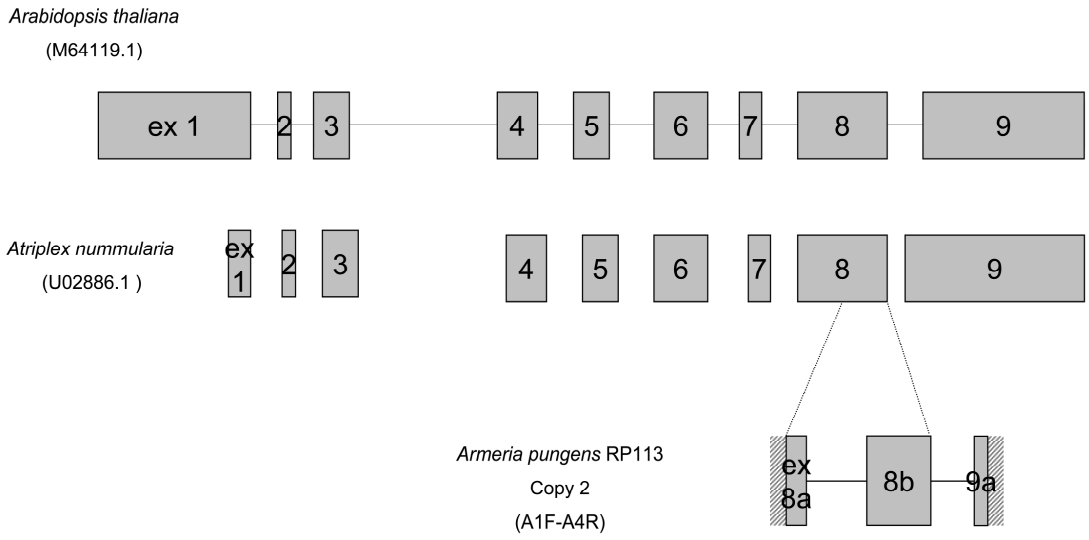


Table 4.4 Intron and exon length differences of the three *GapC* paralog copies 1, 2 and 3 found in *A. pungens* taking as a reference the Gene Bank sequence M64119 of *Arabidopsis thaliana*.

(All data are bps)	**ex 6**	**intr 6-7**	**ex 7**	**intr 7-8**	**ex 8**			**intr 8-9**	**ex 9**				
Arabidopsis thaliana (M64119)	147	95	61	94	241			89	444				
	ex 6	**intr 6-7**	**ex 7**	**intr 7-8**	**ex 8a**	**intr 8a-8b**	**ex 8b**	**intr 8b-9a**	**ex 9a**			**intr 9a-9b**	**ex 9b**
A.pungens copy 1	<106	112-115	61	77	98	98	143	78	119>				
A.pungens copy 2			<42	91	98	106-108	143	80-89	174			131-133	30>
									ex 9a1	**intr 9a1-9a2**	**ex 9a2**		
A.pungens copy 3			<42	116-121	98	133-237-245	143	134-137	85	111-116	35>		

For the present phylogeographic study, the combination A1F-A4R was chosen. It specifically amplifies a partial region of the Copy 2 in *Armeria* (Fig. 4.2).

Figure 4.2 Scheme of the partial region of *GapC* gene used for filogeografic inference of *A. pungens*, as compared to *Arabidopsis thaliana* and *Atriplex nummularia*. This region corresponds to the paralog copy 2 found in *A. pungens*, and is specifically amplified in *Armeria* with primers A1F-A4R.



4.2.3 PCR Amplification with A1F-A4R

For PCR amplification, one or more reaction of 25 µl per sample had been prepared, using puRe Taq® Ready-To-Go® PCR Beads, Amersham Biosciences Corp., with the following components (in final concentrations): 1.5 mM $MgCl_2$; 0.2 mM of each dNTP; 0.2 µM of each primer; 2.5 units of puReTaq®DNA Polymerase. The amount of total genomic DNA was variable, depending on the sample. The proper quantity had been determined empirically and was around 200-300 ng per reaction.

Cycling conditions were an initial denaturation at 95ºC (5 min); then 38 cycles of 95ºC (30 sec), 53ºC (30 sec), 72ºC (1 min); and a final extension at 72ºC (5 min). Samples for direct sequencing had been purified with UltraClean PCR Clean-up Kit (Mo-Bio Laboratories). When more than one band were present and for cloning, samples had been purified from 1.8 % agarose gel, using Perfectprep Gel Cleanup (Eppendorf).

4.2.4 Direct sequencing

Primers A1F-A4R were used for direct sequencing of *GapC* PCR products from 71 individuals of *A. pungens*. Forty one of the sequences were unambiguous whereas double electropherogram profiles were detected in the remaining 30. Sequences in double electropherograms were separated by haplotype substraction, using unambiguous sequences and single site polymorphic sequences as a reference (Clark, 1990; Caicedo and Schaal, 2004). A total of 101 sequences were obtained.

4.2.5 Genealogical relationships and allele frequencies

A final matrix containing the 101 sequences was aligned with Bioedit 5.0.9 (Hall, 1999). Alignment was straightforward although indels could be aligned in two slightly different ways. Indels were coded as presence/absence data following Simmons and Ochoterena (2000). Genealogical relationships among alleles were represented through a statistical parsimony network built with TCS 1.21 (Clement *et al.,* 2000) treating indels as missing values. Frequency and geographic distribution of alleles were examined.

4.2.6 Genetic diversity

From the final matrix of 101 sequences haplotype diversity (equations 8.4 and 8.12 but replacing 2n by n; Nei, 1987,) was calculated with dnaSP 4.10.7 (Rozas *et al.*, 2003). Sequences from the same genetic lineage were pooled in order to reach adequate sample sizes and regional genetic diversity estimates were calculated.

4.2.7 Cloning

In order to assess the different sequence types amplified by primers A1F-A4R, 11 out of 30 individuals yielding double peaks in electropherograms of direct sequences

were also cloned. Cloned accessions had diverse geographic origins (Table 4.5). Two additional individuals, not amplified by direct sequencing, were also cloned.

Cloning procedure was performed on purified samples with pGEM-T Easy Vector System II kit (Promega). Colonies screening was carried out by direct PCR of the colonies with universal vector primers T7 and SP6, preparing 20 μl reaction per colony with the following components (in final concentrations): 2.5 mM $MgCl_2$; 0.2 mM of each dNTP; 0.2 μM of each primer; 0.8 U of HotMaster Taq DNA Polymerase (Eppendorf). 1 μl of a 10 μl ultra pure water solution of the picked colony had been used as template. Cycling conditions were 94ºC (2 min); then 35 cycles of 94ºC (30 sec), 50ºC (30 sec), 65ºC (2 min); and a final extension at 65ºC (20 min). Direct sequenced colony PCR products had been purified with UltraClean PCR Clean-up Kit (Mo-Bio Laboratories).

Sixteen average colonies (from 7 to 24; Table 4.5) were picked per sample to produce a total of 216 cloned sequences from 13 individuals. All cloned variants seemed to be functional except for two clones from individuals from populations 6 and 9, harbouring a stop codon position and a mutation in the splicing sequence, respectively. These putative pseudogenes were discarded.

The final matrix of 214 clones per 415 bp was imported in PAUP v. 4.0b10 for Macintosh (Swofford, 2002) a parsimony heuristic search was performed with 100 stepwise random addition replicates, and TBR branch swapping (selecting options “multrees” and “steepest descent”). A strict consensus tree of the 2479 saved trees was generated with PAUP and fast boostrap calculated based on 10000 permutations. In addition, a NJ distance tree was performed

Nucleotide substitutions exclusive to one sample were ignored as putative Taq polymerase errors by creating majority rule consensus sequences for each copy type of every single accession. Seventeen pairs of consensus clones and direct sequences from all the 11 accessions were compared for: i) number and type of copies retrieved ii) the reproducibility in terms of nucleotide composition of each copy.

4.2.8 Recombinants

Putative recombinants were identified from long terminal branches and unexpected position in the parsimony network (see Popp *et al.*, 2005) as well as from the visualisation of the sequence. Additionally, the methods recommended in Posada and Crandall (2001) and Posada (2002) to detect recombination in low divergent data sets were applied: i) Maximum chi square test, based on substitutions distribution, was applied to alignments of three sequences at a time (2 putative parental + 1 putative recombinant) as implemented in START version 2 (Jolley *et al.*, 2001). Significance was established with permutation test (1000 replicates); ii) The Recombinant Detection Program (RDP) method (Martin *et al.*, 2005), based on phylogenetic position, was applied to the entire alignment using only internal sequences as a reference; iii) The Homoplasy test could not be applied since at least 10 parsimony informative synonymous sites per locus are needed (Jolley *et al.*, 2001).

4.3 RESULTS

4.3.1 Structure of the *GapC* region in *A. pungens*

The partial *GapC* region amplified by primers A1F-A4R included exon 8b and the two flanking introns. The structure of this region as compared to *Arabidopsis thaliana* is represented in Fig. 4.2

4.3.2 Identification of *Gap C* sequence types

Evaluation of Maximum Parsimony (MP) tree and NJ distances of the 214 clones, allowed identifying three main clades revealing three types of sequences amplified by primers A1F-A4R: I, II and III (Fig. 4.3). Phylogenetic position and direct visualization of the nucleotide sequence, indicated the existence of 12 recombinant clones between Types I and II (Table 4.5). Interestingly, in the individual harbouring only type I sequences, no recombinant clones were retrieved. Ten of the recombinants were confirmed with the Maximum Chi square method and randomization test. In contrast, the RDP method applied to the entire data set did not recognise any recombinant. The low number of parsimony informative sites in our data set may explain this failure.

The three sequence types had approximately the same size taking exons and introns into account (I 377 bp, II 368 and III 373) and coded exactly for the same protein, i.e., substitutions in the exon are synonymous changes (results not shown). Distances were low and approximately equivalent among types (Fig. 4.3 B). Type III was closer to Type II if the recombinants are included, and to Type I if they are removed (resuts not shown). In contrast, differences in the frequency of appearance were evident among types. Type I appeared in 159 individuals, Type II in 37 and Type III was extremely rare, appearing only in 6 cases (Table 4.5). The MP analysis revealed high bootstrap supports for the three clades (89%, 91% and 86%, respectively; Fig. 4.3 A).

Table 4.5 Comparison of direct sequencing and cloning procedures for *GapC* gene in *A. pungens*

SAMPLE	DIRECT SEQUENCING	CLONING				
Pop- individual	Sequence types	type I	type II	type III	recombinants I,II	total clones
2-1	I,II	15	1		2	18
2-2	I,II	20	3		1	24
5	I,II	10	0		1	11
6	I,II	13	0	1		14
9	I,II	9	0	4	1	14
8-1	I,II	20	3			23
8-2	I,II	15	4		1	20
11	not available		5		1	5
12-1	not available	16	2		1	19
12-2	I,II	16	2	1	1	20
15-1	I,II	15	2		3	20
15-2	II,II		15			15
17	I,II	10	0			10

Figure 4.3 Relationships among *GapC* sequence Types in *A. pungens*, based on 214 clones from 11 individuals. Three sequence types can be distinguished: I, II and III. Terminals are labelled according to the population of origin (Table 1) followed by the individual number. A: MP tree with bootstrap supports above 50 indicated over the branches.

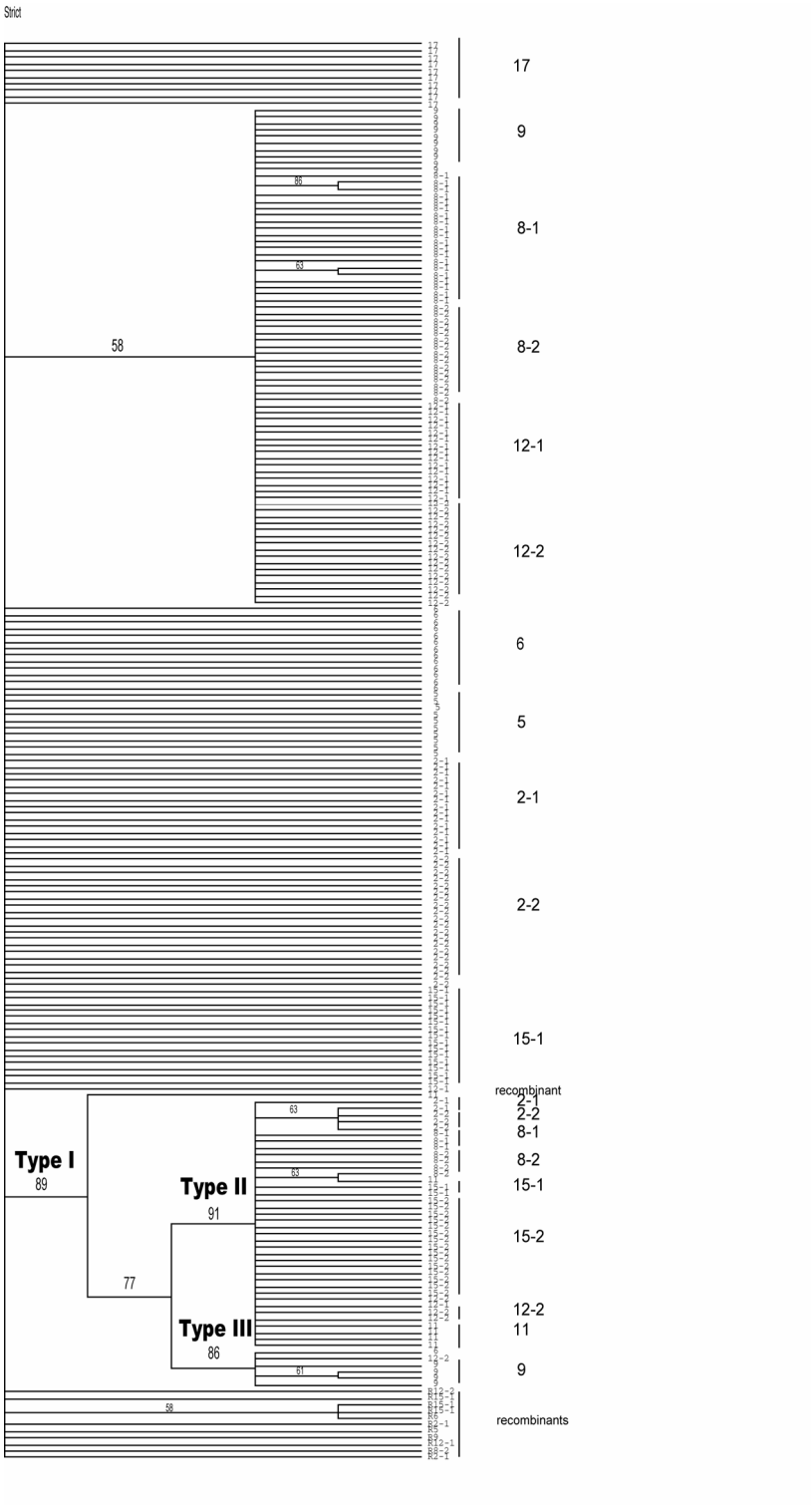


Figure 4.3 (continued): NJ distance tree with middle point rooting.

4.3.3 Direct sequencing *vs.* cloning

4.3.3.1 *Sequence types retrieved*

Comparison between the results of the 11 individuals analysed by direct sequencing and cloning procedures are summarised in Table 4.5. In six out of eleven individuals, exactly the same two sequence types were retrieved with both methods. In the remaining five, the one of the sequences of the less frequent Type II, was not retrieved in the clones, despite being found in the direct sequence. Remarkably, for these five individuals less colonies were picked than for the other eight (15 or less), which suggests a sampling failure of the Type II. However, the rare copy III complicates the results. It was detected in six out of 214 clones from three individuals (1 out of 15, 1 out of 20 and 4 out of 11 clones, respectively). In two of them, direct sequencing revealed types I and II whereas cloning revealed types II and III. In the third one, clones from all three Types I, II and III were retrieved from the same individual.

4.3.3.2 *Nucleotide composition*

Nucleotide composition of the 17 pairs of equivalent direct sequences and consensus clones was identical in 16 cases. Only one sequence presented a difference in a single nucleotide in one non parsimony informative site in (individual 2-1, sequence Type II). Since this mutation is exclusive to that clone in the whole data set it seems to be due by a Taq error. Only one clone of sequence Type II was retrieved from this accession, therefore no consensus could be created in order to correct for Taq artefacts (Smith *et al.*, 1997; Eyre-Walker *et al.*, 1998).

In summary, when the same copy types were retrieved both through direct sequencing and cloning, exactly the same nucleotide composition was found. However, both methods failed the amplification of one of the three copy types that was retrieved with the alternative method in three and four cases respectively.

4.3.4 Variation of *GapC* in *A. pungens*: genealogical relationships, haplotype frequencies and genetic diversity

The final alignment of the 101 *GapC* direct sequences of *A. pungens* was 380 bp long. Thirty two variable sites, 25 nucleotide substitutions and seven indels, were found that determine 13 allelic variants. The mutations among the 13 alleles are reported in Table 4.6. Clones were not included since they added only some extra sequences in clade Type III, which lacks enough sampling (see below).

The allele genealogy estimated with TCS, is represented in Fig 4.4. TCS algorithm did not support within the limits of 95% statistical parsimony the connections among the three main sequence types I, II and III and distinguished three independent clades corresponding to them. Variable sites distinguishing alleles congruent with geography were contained in exon 8 and in the first intron. The distinction of three clades (=sequence types) was mainly determined by the intron 2, although it was consistent with the sequence of intron 1 and the exon.

A TCS was also tried not taking indels into account (as presence/absence data) yielding the same alleles and patterns were depicted (results not shown). Only minor differences were found: i) the three sequence types I, II, and III are connected within the 95% parsimony limits, ii) less steps separate alleles E, F, G and L. Additionally, no difference was observed in the two slightly different alignments (results not shown).

Among the direct sequences, putative recombinants between Types I and Type II were identified from intermediate sequence and/or unexpected position in the parsimony network. They corresponded to alleles D (two sequences), allele E (three sequences) and allele M (one sequence). These sequences were tested for recombination with MaxChi square test. Additionally, four divergent sequences corresponding to alleles F (one sequence), K (one sequence) and G (two sequences) were additionally tested. Out of the six tested alleles, haplotypes M and E, were confirmed as recombinants by the MaxChi squared test.

Table 4.6 Single substitutions and indels among the 13 *GapC* alleles represented in the 101 sequences of *A. pungens* obtained. Alleles are named according to the TCS network in Figure 4.4. Exons positions are no coloured. Substitutions in the intron are coloured in grey

Allele	35	81	93	94	104	115	151	172	202	226	232	244	271	285	289	291	292	293	294	299	300	301	306	312	313	324	325	334	335	339	349	357
A	A	T	A	A	TT	.	A	C	T	A	A	G	C	.	G	A	C	A	A	A	G	AAAAA	T	G	AGATGTTTG	G	C	T	A	T	GT	T
B	A	T	A	A	TT	.	A	T	T	G	A	G	C	.	G	A	C	A	A	A	G	AAAAA	T	G	AGATGTTTG	G	C	T	A	T	GT	T
C	G	T	A	A	TT	.	A	C	T	A	A	G	C	.	G	A	C	A	A	A	G	AAAAA	T	G	AGATGTTTG	G	C	T	A	T	GT	T
D	A	T	A	A	TT	.	A	C	T	G	G	G	C	.	G	A	C	A	A	A	G	AAAAA	T	G	AGATGTTTG	G	C	T	A	T	GT	T
E	A	T	.	T	TT	GT	A	C	T	G	G	G	C	.	G	A	C	A	A	T	G	AAAAA	T	G	AGATGTTTG	G	C	T	A	T	GT	T
F	G	T	A	T	TT	GT	A	C	T	G	G	A	C	.	C	A	A	C	A	T	A	AAAAA	G	A	.	G	A	T	C	A	.	T
G	A	A	A	T	.	GT	A	C	T	G	G	G	C	.	G	A	C	A	A	T	G	AGAAA	T	G	AGATGTCTG	G	C	A	A	T	.	T
H	C	T	A	T	TT	GT	A	C	C	G	G	A	T	.	G	A	A	A	G	T	G	AAAAA	T	A	.	G	A	T	A	T	GT	T
I	A	T	A	T	TT	GT	A	C	T	G	G	A	C	.	C	A	A	C	A	T	A	AAAAA	G	A	.	G	A	T	C	A	.	T
J	A	A	A	T	TT	GT	A	C	T	G	G	A	C	.	C	A	A	C	A	T	A	AAAAA	G	A	.	G	A	T	C	A	.	T
K	A	A	A	T	TT	GT	A	C	T	G	G	A	C	.	C	A	A	C	A	T	A	AAAAA	G	A	.	G	A	T	C	A	.	C
L	A	A	A	T	TT	GT	A	C	T	G	G	A	C	.	C	C	C	A	A	T	A	AAAAA	G	A	.	G	A	T	C	A	.	T
M	A	A	A	T	TT	GT	C	C	T	G	G	A	C	T	C	C	C	T	A	T	T	.	T	G	AGATGTTTG	T	C	T	A	T	.	T

4.3.4.1 *Phylogeographic patterns of Type I and Type II*

GapC variation in *A. pungens* exhibited a strong geographic pattern in the 12 alleles corresponding to sequence Types I and II (Figs. 4.4 and 4.5). An ancestral allele, A, was found in 17 out of 22 populations from all geographic regions. Allele C, closely related to the ancestral allele A, was fixed in the Cies islands (pop. 1), and shared with the closest population in the mainland (pop. 2- Albufeira) as well as in one Sardinian population (pop. 19- Porto Liscia). Alleles B and I characterise and relate Portuguese populations with the Corso-Sardinian ones, together with the local allele D restricted to two neighbouring populations from Portugal (pops. 5 and 6). Outside Portugal and Corsica-Sardinia, allele B is exceptionally found in one sequence from Camarinal (pop. 14). Allele J is present in all six population from the Gulf of Cadiz and is exclusive to this region (pops. 8 to 13), together with the local related alleles K and L (pop. 13- Trafalgar). Finally, distinct minor alleles E and G appear in Camarinal (pop. 14). G is quite divergent and private to this population, whereas E is shared with a Portuguese population (pop. 5-Zambujeira) and intermediate between clades I and II (confirmed to be recombinant by the Maxchi test). Two outlier sequences were obtained: allele F. in pop. 13, and allele M in population 10 (identified as a recombinant in the MaxChi squared test).

4.3.4.2 *Type III*

A single allele (H) could be distinguished within the sequence Type III. No phylogeographic signal can be retrieved from this clade since it was found only in a few individuals.

Figure 4.4 *GapC* haplotype network for *A. pungens* constructed with TCS, based on 101 direct sequences form 71 individuals. The relative size of the circles represent the frecuency of each allele. The position of nucleotide substitutions (bars) and indels (crosses) separating haplotypes are indicated according to Table 4.6. The connections among the three main clades corresponding to the three sequence Types I, II and III are represented with dotted lines, since they are no supported within the limits of 95% statistical parsimony.

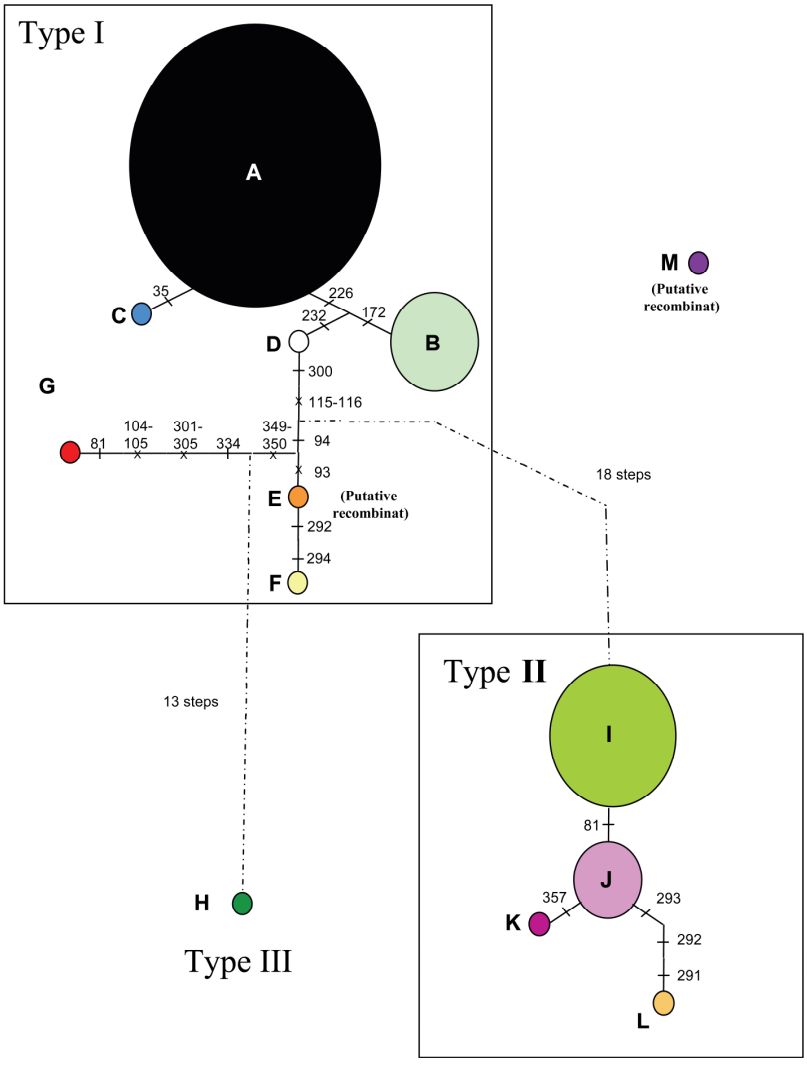
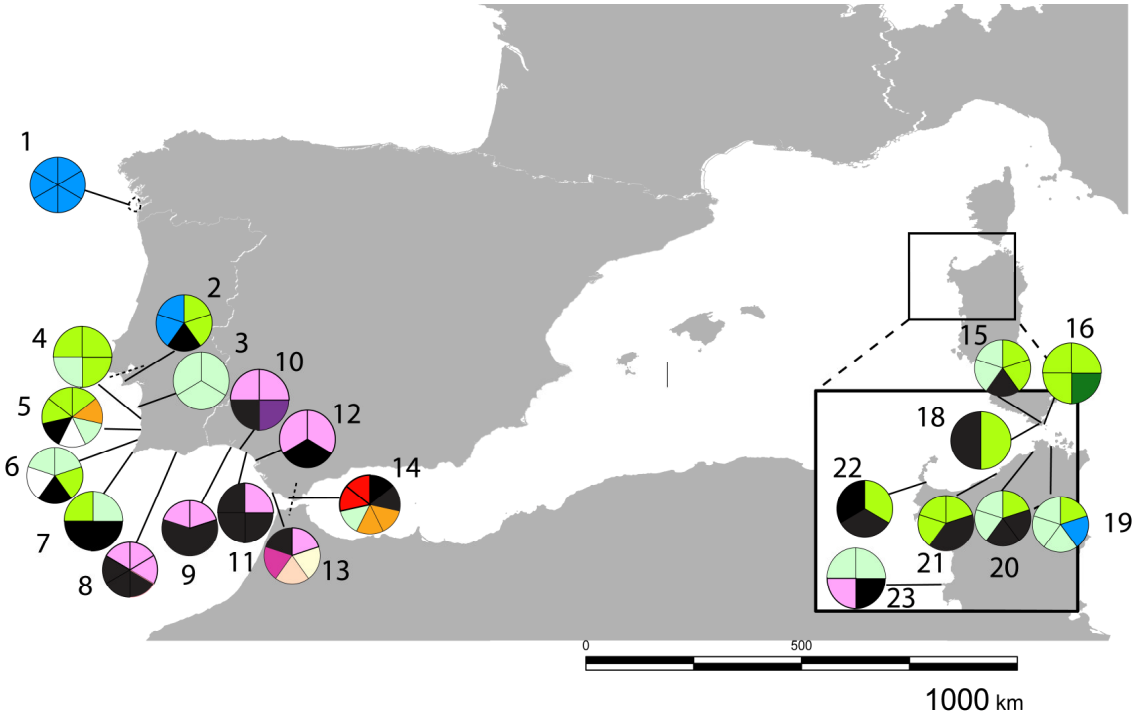


Figure 4.5 Allele frequencies and geographic distribution of *GapC* in *A. pungens*. Allele colours follow the TCS network in Figure 4.4. Populations are indicated as in Table 4.1.



4.3.4.3 *Genetic diversity*

Haplotype diversity for *GapC* in *A. pungens* was higher in Portuguese populations, slightly lower in Corsica-Sardinia, remarkably lower in the Gulf of Cadiz and zero in the Cies Islands. Camarinal population exhibited the highest diversity levels (Table 4.7).

Table 4.7 Haplotype diversity of *GapC* in different geographical areas of *A. pungens*. Populations contained in each area are indicated.

	Nr. individuals (sequences)	Hd
Cies Islands (pop.1)	6 (6)	0
Portugal (pops. 2-7)	20 (28)	0.775 (0.046)
Gulf of Cadiz (pops. 8-13)	19 (27)	0.655 (0.059)
Camarinal (pop. 14)	4 (7)	0.857 (0.102)
Corsica-Sardinia (pops. 15-23)	22 (33)	0.739 (0.037)

4.4 DISCUSSION

4.4.1 Lack of coalescence of the *GapC* at the species level

Unlike chloroplast markers, nuclear markers provide the opportunity to compare multiple unlinked genes. The degree of congruence among them may inform about the kind of processes acting on the genome. The polyphyletic pattern of *GapC* tree in *A. pungens* is not in agreement with the phylogeographic signal contained in other nuclear markers used in this species (AFLP and ITS; chapters 2 and 3). In addition, no congruence of the three main clades (sequence Types I, II and III) with the geography is observed. Therefore, the processes shaping the variation of *GapC* in *A. pungens* seem to be specific to this nuclear region, whereas hybridisation or phylogeographic events can be discarded, since they act at the whole nuclear genome level and usually result in geographic patterns of variation. Instead, mechanisms with a local effect in the genome are more likely to be involved, such as incomplete sorting of ancestral polymorphisms or presence of several paralogs. Distinguishing both possibilities is not straightforward.

The three sequence Types are only slightly different in terms of nucleotide composition or size and identical in their gene coding patterns, this strongly suggests that they are different alleles from the same orthologous locus that has undergone random sorting of ancestral alleles. However, given the diploid condition of *A. pungens*, a maximum of two alleles of each locus is expected per individual, while in the three cloned individuals where the Type III sequence was retrieved, Types I and II were also detected, either through direct sequencing or cloning. The detection of the three sequence Types in the same diploid individual together with the low incidence of Type III (7 out of 315 sequences and 4 out of 72 individuals) point at this type as a putative recent paralogous locus that amplifies only occasionally, whereas the status of Types I and II fits better that of uncoalesced alleles of the same locus.

Furthermore, taking into account the expected level of heterozygosity of an outcrossing plant like *A. pungens,* if Types I and II were considered paralogs the frequency of homozygous individuals would be strikingly high, but analysed as alleles would mean 41 (57.7%) homozytes *vs.* 30 (42.2 %) heterozygotes, which makes more sense.

4.4.2 Interpreting population history from polyphyletic gene patterns

Although the three distinct clades (=sequence Types) do not follow a geographic pattern, allele variants within clades I and II clearly do so. The number of terminals in clade III is too small for any pattern to be detected. This means that sequence Types I and II contains a phylogeographic signal that, interestingly, is in perfect agreement with the genetic groups revealed by AFLP and ITS (chapters 2 and 3). Particular allelic variants characterise Portuguese-Corso-Sardinian lineage (alleles B, D and I) *vs.* the Gulf of Cadiz (alleles J, K, L). In addition, the allele C relates the isolated population in the Cies Islands (pop. 1) with the northernmost population of the mainland (Albufeira; pop. 2), exactly like in ITS and AFLP, as well as with the Sardinian region (pop. 15), in agreement with ITS. Finally, the divergent alleles found in Camarinal (G and E; pop 14) fit the distinctness of this populations based on AFLP, ITS, cpDNA and morphologic data, demonstrated to be due to recent or on-going introgression from the sublitoral species *A. macrophylla* (see chapters 1, 2 y 3).

Allelic variants within each non coalesced lineage seem to have originated through recent mutations that mark recent phylogeohraphic events, whereas older random sorting of ancestral polymorphisms is responsible for the differentiation among allelic lineages. Therefore, two genetic patterns generated at different temporal scales are probably overlapping. By examining both genealogical relationships and allele frequencies we intend to separate the ancestral variation from phylogeographic history. If Types I and II were paralogs instead of non coalesced allelic lineages, they could be analysed as different markers providing independent phylogeographic information. The analysis procedure and main phylogeographic conclusions would not change.

4.4.3 Inference of phylogeographic events limited by low levels of phylogenetically useful variation

Compared to ITS and cpDNA (chapter 3), *GapC* sequences have higher apparent levels of nucleotide substitutions in *A. pungens*. However, the phylogeographic signal is partly blurred by the non coalescence of allelic lineages and, eventually, levels of phylogeographic useful variation provided by *GapC* are not enough to depict the detailed evolutionary history of the species.

However, the phylogeographic information is very consistent, since *GapC* allelic variants in *A. pungens* accurately correspond to the evolutionary lineages of *A. pungens* revealed by nuclear markers and morphology.

In sum, the analysis of the *GapC* region within *A. pungens* strengthens the phylogeographic inferences drawn from AFLP, ITS, cpDNA and morphology. However, The probable involvement of incomplete lineage sorting and/or presence of paralogous loci, recommend caution in considering the results for this marker alone.

4.5 REFERENCES

Bartish IV, Kadereit JW, Comes HP (2006) Late Quaternary history of *Hippophaë rhamnoides* L. (Elaeagnaceae) inferred from chalcone synthase intron (*Chsi*) sequences and chloroplast DNA variation. *Molecular Ecology* **15**, 4065-4083.

Caicedo AL, Schaal BA (2004) Population structure and phylogeography of *Solanum pimpinellifolium* inferred from a nuclear gene. *Molecular Ecology* **13**, 1871-1882.

Clark A (1990) Inference of haplotypes from PCR-amplified samples of diploid populations. *Molecular Biology and Evolution* **7**, 111-122.

Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology* **9**, 1657-1659.

Comes HP, Abbott RJ (2001) Molecular phylogeography, reticulation and lineage sorting in Mediterranean *Senecio* sect. *Senecio* (Asteraceae). *Evolution* **55**, 1943-1962.

Doyle JJ, Doyle JL, Brown AHD (1990) Chloroplast DNA Polymorphism and Phylogeny in the B Genome of *Glycine* Subgenus *Glycine* (Leguminosae). *American Journal of Botany* **77**, 772-782.

Eyre-Walker A, Gaut R, Hilton H, Feldman D, Gaut B (1998) Investigation of the bottleneck leading to domestication of maize. *Proceedings of the National Academy of Sciences of the USA* **95**, 4441-4446.

Friar E, Elizabeth A, Prince LM, *et al.* (2006) Ecological speciation in the East Maui-endemic *Dubautia* (Asteraceae) species. *Evolution* **60**, 1777-1792.

Funk DJ, Omland KE (2003) Species-level paraphyly and polyphyly: frequency, causes, and consequences, with insights from animal mitochondrial DNA. *Annual Review of Ecology, Evolution and Systematics* **34**, 397-423.

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis. *Nucleic Acids Symposium Series* **41**, 95-98.

Hare M (2001) Prospects for nuclear gene phylogeography. *Trends in Ecology & Evolution* **16**, 700-706.

Hughes CE, Eastwood RJ, Donovan Bailey C (2006) From famine to feast? Selecting nuclear DNA sequence loci for plant species-level phylogeny reconstruction. *Philosophical Transactions of the Royal Society B: Biological Sciences* **361**, 211-225.

Isoda K, Shiraishi S (2001) Allelic sequence polymorphisms in the intron region of the nuclear-encoded *GapC* gene preceded the speciation of three closely related *Abies* species (Pinaceae). *Theoretical and Applied Genetics* **102**, 244-250.

Jolley K, Feil E, Chan M, Maiden M (2001) Sequence Type Analysis and Recombinational Tests Version 2. *Bioinformatics* **17**, 1230-1231.

Maddison W, Knowles L (2006) Inferring phylogeny despite incomplete lineage sorting. *Systematic Biology* **55**, 21-30.

McKinnon G, Potts BM, Steane DA, Vaillancourt RE (2005) Population and phylogenetic analysis of the cinnamoyl coA reductase gene in *Eucalyptus globulus* (Myrtaceae). *Australian Journal of Botany* **53**, 827.

Martin DP, Williamson C, Posada D (2005) RDP2: recombination detection and analysis from sequence alignments. *Bioinformatics* **21**, 260–262.

Nei M (1987) *Molecular Evolutionary Genetics* Columbia University Press, New York.

Olsen KM, Schaal BA (1999) Evidence on the origin of cassava: Phylogeography of *Manihot esculenta*. *Proceedings of the National Academy of Sciences USA* **96**, 5586-5591.

Palumbi SR, Cipriano F, Hare M (2001) Predicting nuclear gene coalescence from mitochondrial data: the three-times rule. *Evolution.* **55**, 859–868.

Patterson C (1988) Homology in classical and molecular biology. *Molecular Biology and Evolution* 5, 603-625.

Popp M, Erixon P, Eggens F, Oxelman B (2005) Origin and evolution of a circumpolar polyploid species complex in *Silene* (Caryophyllaceae) inferred from low copy

nuclear RNA polymerase introns, rDNA and chloroplast DNA. *Systematic Botany* **30**, 302-313.

Posada D, Crandall K (2001) Evaluation of methods for detecting recombination from DNA sequences: computer simulations. *Proceedings of the National Academy of Sciences USA* **98**, 13757-13762.

Posada D (2002) Evaluation of methods for detecting recombination from DNA sequences: real data. *Molecular Biology and Evolution* **19**, 708-717.

Rieseberg LH, Ellstrand NC (1993) What can molecular and morphological markers tell us about plant hybridization. *Critical Reviews in Plant Sciences* **12**, 213–241.

Rozas J, Sánchez-DelBarrio JC, Messeger S, Rozas R (2003) DNA polymorphisms analysis by the coalescent and other methods. *Bioinformatics* **19**, 2496-2497.

Sang T (2002) Utility of low-copy nuclear gene sequences in plant phylogenetics. *Critical Reviews in Biochemistry & Molecular Biology* **37**, 121-147.

Schaal BA, Hayworth DA, Olsen KM, Rauscher JT, Smith WA (1998) Phylogeographic studies in plants: problems and prospects. *Molecular Ecology* **7**, 465-474.

Schaal BA, Olsen KM (2000) Gene genealogies and population variation in plants. *Proceedings of the National Academy of Sciences USA* **97**, 7024-7029.

Small R, Cronn RC, Wendel JF (2004) Use of nuclear genes for phylogeny reconstruction in plants. *Australian Systematic Botany* **17**, 145-170.

Small R, Wendel JF (2000) Phylogeny, duplication, and intraspecific variation of *Adh* sequences in new world diploid cottons (*Gossypium* L., Malvaceae). *Molecular Phylogenetics and Evolution* **16**, 73-84.

Smith D, McAllister J, Casino C, Simmonds P (1997) Virus `quasispecies': making a mountain out of a molehill? *Journal of General Virology* **78**, 1511-1519.

Simmons MP, Ochoterena H (2000) Gaps as characters in sequence-based phylogenetic analyses. *Systematic Biology* **49**, 369-381.

Strand AE, Leebens-Mack J, Milligan BG (1997) Nuclear DNA-based markers for plant evolutionary biology. *Molecular Ecology* **6**, 113-118.

Swofford D (2002) *Phylogenetic Analysis Using Parsimony (and Other Methods), Version 4*. Sinauer Associates, Sunderland MA.

Syring J, Farrell K, Businský R, Cronn R, Liston A (2007) Widespread genealogical nonmonophyly in species of *Pinus* subgenus *strobus*. *Systematic Biology* **56**, 163-181.

Vriesendorp B, Bakker FT (2005) Reconstructing patterns of reticulate evolution in angiosperms: what can we do? *Taxon* **54**, 493-604.

Wendel JF, Doyle JJ (1998) Phylogenetic incongruence: window into genome history and molecular evolution. In: *Molecular Systematic of plants II: DNA sequencing* (eds. Soltis PS, Soltis DE, Doyle JJ), pp. 265- 296. , Kluwer, Dordrecht.

5. GRAZING-INDUCED GENETIC BOTTLENECK THREATENS THE MACARONESIAN ENDEMIC *Armeria maderensis*

5. 1 INTRODUCTION

Propensity to extinction of island plants is well known. For instance, in the reviews by Primack (1998) and WCMC (1992), most species gone extinct between 1600 and the 1990's occured on islands (Frankham *et al.*, 2002). Understanding the causes for this susceptibility is an interesting challenge for evolutionary biologists. The most straightforward reasons are the small geographic ranges and low numbers of populations, which prevent a possible buffering effect of a large range whenever it faces a catastrophic event. Also, isolated evolution in the absence of predators and competitors makes island species less resistant to the arrival of foreign organisms both predators and competitors (Frankham *et al.*, 2002). Many island taxa have originated via long-distance dispersal of mainland ancestors (Winkworth *et al.*, 1999; Charbonnel *et al.*, 2002; Cowie and Holland, 2006) and have therefore experienced a reduction in the genetic diversity through founder events. Besides, island colonisers are prone to undergo adaptive radiations, producing several ecologically differentiated species that can easily hybridise (Sato *et al.*, 2001; Baldwin and Sanderson, 1998; Rees *et al.*, 2001; Kambysellis and Craddock, 1997). This has been traditionally regarded as a threat for the long-term survival of species, although nowadays it is also viewed as a mechanism able to trigger invasiveness (Ellstrand and Schierenbeck, 2000). In volcanic islands, the risk of volcanism must also be considered. However, in addition to the mentioned natural processes, the primary cause of plant extinction on islands is the negative impact of humans through habitat destruction and grazing of introduced animals as well as foreign competitors (Coblentz, 1978; Fowler de Neira and Johnson, 1985; Rieseberg and Swensen, 1996; Cronk, 1997; Stuessy *et al.*, 1998, Donlan *et al.*, 2002; Wagner *et al.*, 2005).

Madeira, a volcanic island originated within the African plate in the Tertiary (approx 5.3 My; Geldmacher *et al.*, 2000), comprises 1226 described vascular plant species, of which 234 are Macaronesian endemics and 157 exclusive to the island (Vieira, 1992; Press and Short, 1994; Jardim and Francisco, 2000). For altitudes higher

than 1300 m, Jardim and Francisco (2000) reported 54 Madeiran endemics. One of those alpine endemics is the plant *Armeria maderensis* Lowe, which has an extremely restricted distribution in rocky areas exposed to humid winds in the central sector of the main mountain range of Madeira (Vieira, 1992; Press and Short, 1994; Jardim and Francisco, 2000). Its distribution spans a small area between the highest peaks, Ruivo (1862 m) and Areeiro (1818 m), which is part of the Natural Park of Madeira, a geological and high altitude vegetation reserve, also part of a Special Protected Area (SPA) included in the Natura 2000 Net as a Communitary Important Site (CIS). *Armeria maderensis* occurs only above 1600 meters, i.e.,-strictly in the supratemperate belt (Mesquita *et al.*, 2004). Climax vegetation in this belt includes several rupicolous communities which may develop where tree cover and grazing are absent. One of them is the *Armerio maderensis-Parafestucetum albidae***,** dominated by *Armeria maderensis* Lowe, *Deschampsia maderensis* (Hack and Bornm.) Buschm, *Parafestuca albida* (Lowe) Alexeev, *Anthoxanthum maderensis* Teppner and *Anthyllis lemmaniana* Lowe. Moderate grazing of this community seems to result in the replacement by the association *Viola rivianae-agrostietum castellanae*, dominated by non-endemics (Costa *et al.*, 2004; Silva *et al.*, unpublised). In locations exposed to heavy grazing and soil erosion, even poorer annual communities are established (*Leontodo longirrostris-Ornithopetum perpusilli*). Remarkably, *A. maderensis* has not been reported to occur on these poor alpine pastures.

Very early in the history of Madeira, in 1419, goats and sheep were introduced and spread across the mountains (Sousa, 2003). The introduction of goats on islands became a frequent way of providing fresh meat throughout the world, mostly in 17th and 18th centuries. This livestock grazing affects directly the composition of plant communities because herbivores harvest selectively, eliminating some plants and altering the distribution of others. The introduction of herbivores has also indirect effects on the vegetation because it may pave the way for exotic annual plants to invade the habitat (Holmgren, 2002; Silva *et al.*, unpublised). Qualitative changes on the vegetation induced by goats have been repeatedly described, e.g., in the Robinson Crusoe Island (Cuevas and Le Quesne, 2006), Sta Catalina islands (California), Galápagos, New Zealand, Hawaii and Sta Helena (Coblentz, 1978).

The impact of long-term grazing on the Madeiran alpine vegetation is evident from the fact that never grazed areas ("mangas") have the highest floristic diversity in the island (Silva *et al.*, 2005). This effect was observed decades ago by authors who claimed the need of introducing controls on grazing activities in Madeira (Silva and Menezes, 1946; Andrada, 1990).

In 1994, the regional government initiated the removal of goats in Madeiran Mountains in the frame of a conservation program according to EU Habitats Directive 92/43/EEC. The process was completed in 2003 (Sousa, 2003). A general increase of floristic diversity could be already observed between 2001 and 2006 (Silva *et al.*, unpublised). In particular, *Armeria maderensis* has been observed (Silva *et al.*, unpublished) to descend from inaccessible rocky crevices, which were the exclusive habitat before, to floristically poor alpine pastures (*Viola rivianae-agrostietum castellanae* and *Leontodo longirrostris-Ornithopetum perpusilli*).

This study uses AFLP data to assess the genetic architecture of *Armeria maderensis* to help interpret its evolutionary history and evaluate its conservation perspectives. Our sampling was performed in 2003, just before goats were definitively removed and after almost 600 years of grazing. Specifically, the following questions are investigated:

a) Is genetic variation structured in *A. maderensis*?
b) Is it genetically depauperate?, and in that case, how much in comparison with other widespread *Armeria*?
c) How can we interpret its genetic structure or lack thereof taking into account life history traits and phylogenetic position?
d) What may be the influence of long-term grazing on the genetic structure? What prospects can be envisaged for its conservation?

5.2 MATERIALS AND METHODS

5.2.1 Study system

Armeria maderensis is the only endemic representative of the genus in Macaronesia. Morphologically, the most unique feature is the lack of imbrication of involucre bracts, which are few, narrow and seemingly arranged in a single row so that it can be hardly said that they constitute an involucre. The insertion of the flower pedicel on the calyx is frequently more truncate and thus less spurred than in other congeners. Another apparent feature is the patent arrangement of the inner flower in each spikelet within the glomerule. These morphological characters are useful to identify *A. maderensis* but to the extent that they are not shared with other species, they do not help in finding its closest relatives.

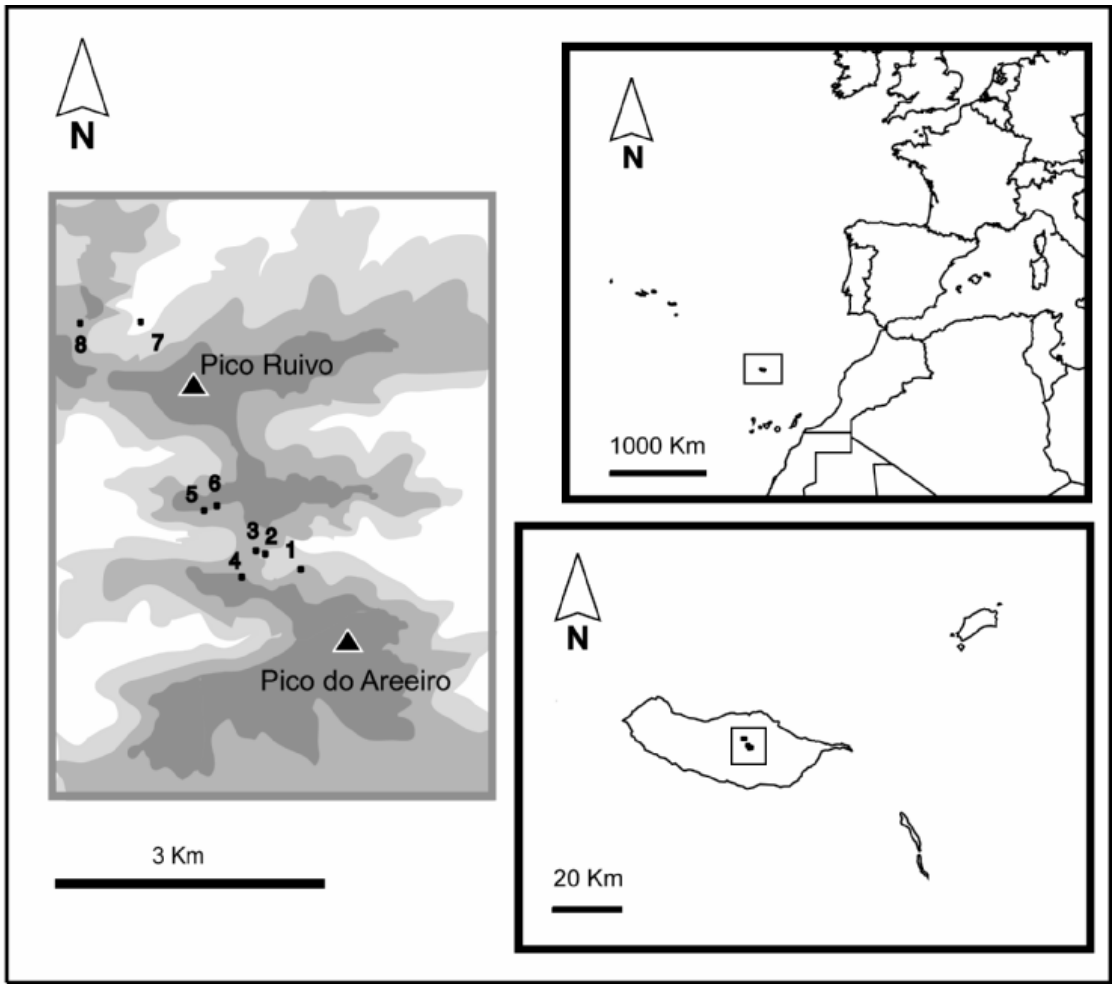
Based on the number of cases in which closest relatives to other Macaronesian endemics have been investigated with phylogenetic methods (reviewed in Juan *et al.*, 2000; Emerson, 2002; Carine *et al.*, 2004) and on the high taxonomic diversity for *Armeria* existing in the Western Iberian Peninsula, it is reasonable to hypothesize that *A. maderensis* originated from a Western Mediterranean ancestor. The position of *A. maderensis* in a phylogeny based on nuclear ribosomal ITS is consistent with this hypothesis. However, the identification of the specific Western Mediterranean lineage of *Armeria* having originated *A. maderensis* is precluded due to the irresolution in deeper nodes of the tree. The strong ITS divergence of *A. maderensis* from its Western Mediterranean congeners (99% bootstrap) clearly indicates that the arrival of the putative ancestor did not take place in recent times (Fuertes Aguilar and Nieto Feliner, 2003; Nieto Feliner *et al.*, 2001).

5.2.2 Sampling strategy, DNA isolation, AFLP protocol

As mentioned above, the sampling was performed in the small distribution area between peaks Ruivo and Areeiro before goats were removed from the island (Table 5.1; Fig. 5.1). To our knowledge, no previous expeditions had attempted to localise or

quantify the populations of *A. maderensis.* Nine sites were found, the most distant ones being less than 3 km apart. We succeeded at collecting fruits from eight of these sites (Table 5.1). An additional site, harbouring a few individuals on a vertical north-facing wall (pr. Site 8, on the path to Torrinhas peak), was inaccessible even for professional climbers of the National Park.

Figure 5.1 Location of sampled sites of *A. maderensis* for the AFLP study numbered as in Table 5.1 *Armeria maderensis* is restricted to exposed North windy rocky areas at highest altitudes in the central sector of the dorsal chain of Madeira. Maximal distance between sites is less that 3 km.



Sites 1, 3, 7 and 8 consisted of very few scattered individuals (Table 5.1, Fig. 5.1), whereas sites 2, 4, 5 and 6 corresponded to true subpopulations. They were highly inaccessible, which limited the number of individuals collected at each site (e.g.

population 4 had to be sampled by equipped climbers). In total, we were able to sample ripe fruits from 44 separate mother plants.

Table 5.1 Details of collecting work of *A. maderensis* for the AFLP study. Location of sampled sites is represented in Figure 5.1.

SITE	Site, habitat, *collector (date)	Coordinates Altitude (m)	N Collected/ genotyped
1	Pico Areeiro, stony pastures in the summit, GNF, AC, JFA, RL (July 2003)	32º 44' 16.6''N 16º 55' 42.5'' W 1779	2/2
2	Pico Areeiro, rock crevices, GNF, AC, JFA, RL (July 2003)	32º 44' 21.5'' N 16º 55' 55.3'' W 1668	12/6
3	From Pico Areeiro to Pico Ruivo, rock crevices; GNF, AC, JFA, RL (July 2003)	32º 44' 23.7'' N 16º 55' 58.8'' W 1753	1/1
4	From Pico Areeiro to Pico Ruivo, Manga Grande, rock crevices; GNF, AC, JFA, RL (July 2003)	32º 44' 12.7'' N 16º 56' 03.3'' W	5/4
5	From Pico Areeiro to Pico Ruivo, pr. Pico do Gato, rock crevices; GNF, AC, JFA, RL (July 2003)	32º 44' 35.5'' N 16º 56' 17.5'' W 1779	12/8
6	Pr. Pico Areeiro, western slope of Pico do Gato, rock crevices; GNF, AC, JFA, RL (July 2003)	32º 44' 39.3'' N 16º 56' 13.4'' W 1599	10/8
7	From Pico Ruivo to Pico das Torrinhas, stony open shrubs and pastures; GNF, AC, JFA, RL (July 2003)	32º 45' 42.2'' N 16º 56' 41.2'' W 1739	1/1
8	From Pico Ruivo to Pico das Torrinhas, rocky crevices in walls; GNF, AC, JFA, RL (July 2003)	32º 45' 42.2'' N 16º 57' 02.9'' W 1741	1/1

*Abbreviations of collectors: AC, A. Costa; GNF, G. Nieto Feliner; JFA, J. Fuertes Aguilar; RL, R. Lansac.

Seeds were germinated after a cold treatment of one month. Seedlings were cultivated in the greenhouse at the Botanical Garden of Madrid between October 2003 and June 2005 in order to provide fresh leaves for DNA isolation. DNA was extracted with Plant DNeasy Minikit.

The AFLP protocol was performed according to Gaudeul *et al.* (2000) for *Eco*RI/*Mse*I enzyme combination and chapter 2 for *Kpn*I/*Mse*I combination. The three primer combinations used were: (6-FAM)*Eco*RI+ acc/*Mse*I + cacc, (6-FAM)*Eco*RI + acg-*Mse*I + ctac, (6-FAM)*Kpn*I + atc/*Mse*I + cag. Protocols and selective primers were exactly the same as those used in the phylogeographic study of the congener *A. pungens* (chapter 3). In the end, 31 individuals from different mother plants were successfully genotyped with the three primer combinations. This sample size was due (1) to failure in amplifying all the three primer combinations for some samples and (2) avoidance to sample individuals from the same mother plant. Given the restricted distribution of *A. pungens*, the sampling strategy followed probably represents appropriately the genetic variation of the species.

A reproducibility test for each primer pair was performed by reextracting DNA from 7, 6 and 5 individuals, respectively, and repeating the whole procedure. The error rate was calculated as the total number of loci differences relative to the total number of loci comparisons and subsequently averaged over the three combinations. To further assess and eventually refine the quality of the AFLP data, potentially not homologous bands were checked following four different criteria: (i) slight size differences among putative homologous bands across individuals, (ii) low intensity bands, (iii) changing intensity of one band across samples, and (iv) bands of high (upper 10%) or small (lower 10%) molecular weight (Bagley *et al.*, 2001; Bonin *et al.*, 2004). Once identified the bands falling into any of those categories, the error rate improvement after removing bands from each category was calculated (see chapter 2).

5.2.3 Genetic structure of *A. maderensis*

Dice similarity among individuals was calculated and visualised with a PCoA (NTSYSpc 2.1; Rohlf, 1998). In addition, both Jaccard and simple matching coefficients were calculated. These indices gave almost identical results than Dice

coefficient (r=0.99 with Jaccard and r=0.98 with simple matching). A minimum spanning tree (MST) based on Dice was imposed on the PCoA to detect local distortions. Nei and Li distance (1979) was also calculated with PAUP 4.0b10 (Swofford, 2002) and then clustering performed with NJ algorithm. Both Dice and Nei and Li indices, algebraically equivalent, are based on the assumption that shared absences are more likely to be non homologous (Wolfe and Liston, 1998).

Bootstrap support was calculated based on 10000 permutations. Correlation between Nei and Li distances and geographic distances was calculated by performing a Mantel test with NTSYSpc 2.1. Significance was tested by randomization (1000 permutations). Bayesian clustering analyses were performed with STRUCTURE 2.1 and 2.2 (Pritchard *et al.*, 2000; Falush *et al.* 2003; Falush *et al.* 2007) and BAPS 3.2 (Corander *et al.*, 2006). With STRUCTURE we used run lengths of 10^6 following a burnin of 10^5. Long burnin periods were necessary to reach convergence of the alpha statistic before the end of the burnin phase. Each phenotype was coded by a single allele and a missing datum according to the indications in the manual (1-missing for dominant markers and 2-missing for recessive). We selected the admixture ancestry model and the model of correlated allele frequencies. The latter is appropriate in cases where weak genetic structure is expected. Simulations from K=1 to K=8 were run, i.e. the maximum number of genetic groups tested equal to the number of geographical sites. 10 repetitions where run at each K. The number of genetic groups (K) in our data set was inferred taking into account the estimated posterior log probability of the data, *L(K)*, as well as the stability of the assignment patterns of individuals into K groups across repetitions. BAPS simulations were started from K=1 to K=8 as the maximum number of groups, with four replicates at each K.

We tried alternative possibilities with STRUCTURE: (i) the configuration of non-admixture and correlated allele frequencies with the same data set; (ii) coding of alleles proposed in Evanno *et al.* (2005), adding missing data only to dominant alleles but not to the recessive homozygous (1-missing for dominant markers and 0-0 for recessive) run in STRUCTURE 2.2 (Falush *et al.*, 2003). Five replicates at each K were performed in these alternative analyses.

5.2.4 Genetic diversity of *A. maderensis*

The lack of a consistent population subdivision in the distribution of *A. maderensis* makes difficult the calculation of diversity parameters on a per-population basis. Only four of the sites found could be considered true subpopulations. Three of them were accessible enough for appropriate sampling (pops. 2, 5 and 6). Therefore, within-population diversity estimates are reported for these three sites. In addition, all genotyped individuals were pooled in order to calculate the total species diversity.

Four genetic diversity estimates were computed at both population and species level: a) allelic richness, from the percentage of polymorphic loci, *P* (POPGENE 3.2; Yeh and Boyle, 1997) and band richness, A (AFLPDiv 1.0; Coart *et al.*, 2005); b) Gene diversity of Nei (1973), H Nei (POPGENE 3.2) c) allele similarity using Shannon index (1948), Sh (POPGENE 3.2); and d) Bayesian genetic diversity, H bayes (HICKORY 1.0.4; Holsinger *et al.*, 2002; Holsinger and Lewis, 2005). Standard deviations of diversity estimates are reported. Both monomorphic and polymorphic loci were included in the calculations, except for the Bayesian diversity estimates calculated with HICKORY, which takes only polymorphic loci into account.

Finally, Nei´s unbiased gene diversity was calculated only at the population level with TFPGA (Miller, 1997). This measure corrects for different sampling sizes by multiplying the index per 2n/2n-1, where n is the sample size (Nei, 1987: equation 8.4). The bias is effectively reduced for sampling sizes < 50 and large number of loci available, which fits our sampling strategy.

Nei's and Shannon´s measures were calculated for each locus and averaged over loci. Since the concept of heterozygosity cannot be applied to dominant markers, average Nei´s gene diversity is simply a measure of genetic variation. We also calculated Nei´s diversity estimate using Lynch and Milligan's method (1994) implemented with TFPGA (Miller, 1997), which attempts to correct the bias generated by dominant markers by pruning frequent loci for the estimation of allele frequencies. For the Bayesian approach, the full model, the f=0 model (no inbreeding), the theta=0 (no differentiation among populations) and the free model were compared using the Deviance Information Criterion (DIC). Models showing smaller DIC values represent a

better compromise between model fit and model complexity. The data set at the genetic lineage level fitted better the full model, and the data set at the population level produced a lower DIC value for the model (see below how data sets were constructed). Therefore, we estimated diversity parameters under the full and the f=0 models, respectively. In order to check for stability of results, several runs for each analysis were tried (Holsinger and Lewis, 2005). Identical results were obtained in separate runs.

5.2.5 Comparisons of the genetic diversity levels

5.2.5.1 *Comparison with the widespread congener A. pungens*

Genetic diversity estimates of the single genetic lineage of *A. maderensis* were compared with each of the genetic lineages detected within the widespread congener *A. pungens* (chapter 2). Determination of the genetic structure of *A. pungens* was based on the same Bayesian and genetic distance analyses used in the current study (chapter 2). One of the lineages (lineage V, Camarinal population) was not included, given its low sample size and hybrid origin (chapters 2 and 3). For comparability purposes, independent absence/presence matrices were edited for each lineage from the complete AFLP matrices for *A. pungens* and *A. maderensis.* For each genetic lineage, two matrices were edited, one at the population and another at the total genetic lineage level, and diversity parameters were recalculated from them. According to the lowest sample sizes, the number of individuals was standardised to N=6 at the population level and to N=15 at the total genetic lineage level (=species level in *A. maderensis*). This was achieved by random exclusion of individuals, except in the estimation of band richness, where the rarefaction method of El Mousadik and Petit (1996) was used. Loci absent in all individuals in one lineage were removed to avoid underestimation of the genetic variability.

5.2.5.2 *Comparison with reference values of AFLP diversity in plants*

Within-population Nei´s unbiased gene diversity was also calculated for *A. maderensis.* This estimate was the one chosen by Nybom (2004) for comparison of population parameters across studies based on dominant markers. Following Nybom's approach, unbiased Nei's diversity is here reported only for polymorphic bands.

5.3 RESULTS

5.3.1 AFLP profiles

By recalculating the error rate after removing each of the categories of potentially not reproducible bands, those bands changing intensity across samples were observed to be the least reliable (Table 5.2). Discarding them meant a decrease of the error rate below 5% (3.1%). Accordingly, 67 unreliable bands were discarded. An additional unreliable band, non reproducible in most comparisons, was eliminated. Subsequently, individuals that were not amplified for one of the primer combinations (5, 6, and 5, respectively) were removed, which meant the loss of 32 bands. A final data set of 31 individuals per 90 markers was retained for analysis. Markers spanned from 56 bp to 447 bp and only 64.44 % of them were polymorphic. No identical phenotypes among individuals were detected.

5.3.2 Genetic structure

PCoA and NJ tree representing among-individual genetic distances showed no evident geographic structure of the genetic variation within *A. maderensis.* In the PCoA based on Dice similarities among 31 individuals no discontinuities were observed (Fig. 5.2.A). The NJ based on Nei and Li distance among individuals showed neither a geographic pattern nor bootstrap support (Fig. 5.2.B). The Mantel test (r = -0.01608; p random Z < observed Z = 0.6134) confirmed the lack of linear correlation between genetic and geographic distances. Both BAPS and STRUCTURE inferred a single Bayesian cluster within *A. maderensis*, corroborating the absence of genetic structure in the AFLP data (Fig 5.2). BAPS found the optimal partition at K=1 (results not shown). In the STRUCTURE assignments, individuals were evenly assigned into the K groups in simulations from K=2 to K=8, which clearly indicates lack of population structure (Fig. 5.3.A). Consistently, *L(K)* did not show a maximum value at any specific K. Instead, runs at K=1, K=2, K=6, K=7 and K=8 yielded very similar posterior probabilities between *L(K)*=-589.3 and *L(K)*=-586.7 (Fig 5.3.B). A sharp decrease of the posterior probability is observed at K=3 followed by a gradual recovery from at K=3 to K=6 and subsequent stabilisation. This might suggest the existence of 6 genetic groups, but the random assignment of individuals at K=6 discards this possibility. Indentical lack of genetic structure was inferred using the configuration of non-admixture and

correlated allele frequencies as well as the coding of alleles adding missing data only to dominant alleles.

Table 5.2 Error rate in the AFLP data set of *Armeria maderensis*, calculated as the total number of loci differences relative to the total number of loci comparisons and subsequently averaged over the three combinations.
Error rate after removing potentially not homologous bands according to four different criteria: (i) slight size differences among putative homologous bands across individuals, (ii) low intensity bands, (iii) changing intensity of one band across samples, and (iv) bands of high (upper 10%) or small (lower 10%) molecular weight.

	Nr. bands retained	% error rate
All bands included	190	7.1
Type of bands removed:		
Different size among samples	147	6.4
Low intensity	144	6.6
Changing intensity among samples	123	**3.1**
Upper 10% molecular weight	172	6.9
Lower 10% molecular weight	172	6.7
Upper 10% molecular weight + Lower 10% molecular weight	154	6.4
Upper 20% molecular weight	154	7.1
Lower 20% molecular weight	154	6.1
Upper 20% molecular weight + Lower 20% molecular weight	132	5.6

Figure 5.2 Genetic distance analyses of AFLP markers of *A. maderensis*. **A:** Principal Coordinates Analysis (PCoA) representing Dice similarities between AFLP phenotypes of *A. maderensis*. Individuals are plotted against the first and second principal axes. The percentage of variance accounted for each axis is indicated. A maximum Spanning Tree (MST) has been superimposed.

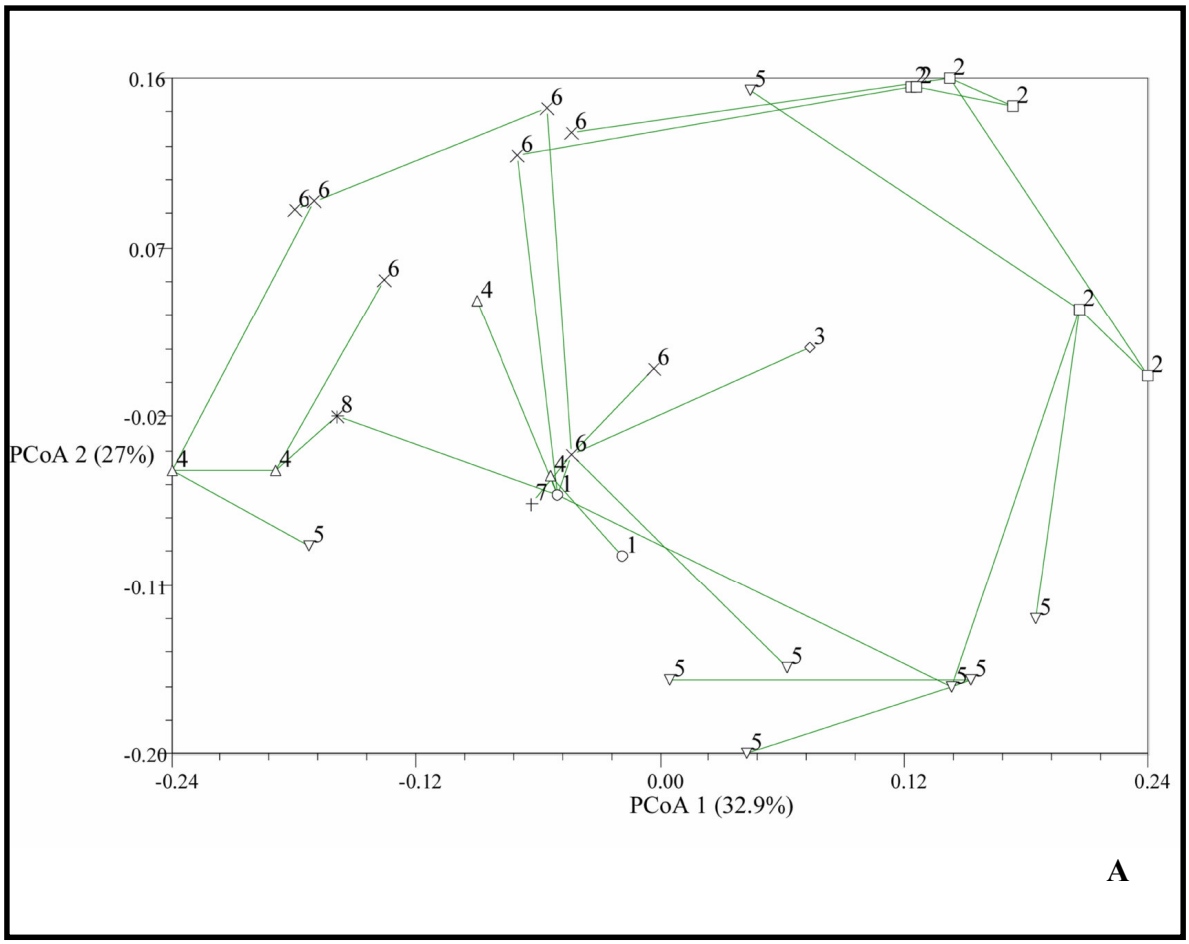


Figure 5.2 (Continued) Genetic distance analyses of AFLP markers of *A. maderensis*. **B**: Unrooted Neighbour-Joining (NJ) tree showing Nei and Li (1979) distances between AFLP phenotypes of *A. maderensis*. Individuals are identified by their geographical location in Table 5.1 and Figure 5.1. Bootstrap is represented only in clusters with more than 50 % support.

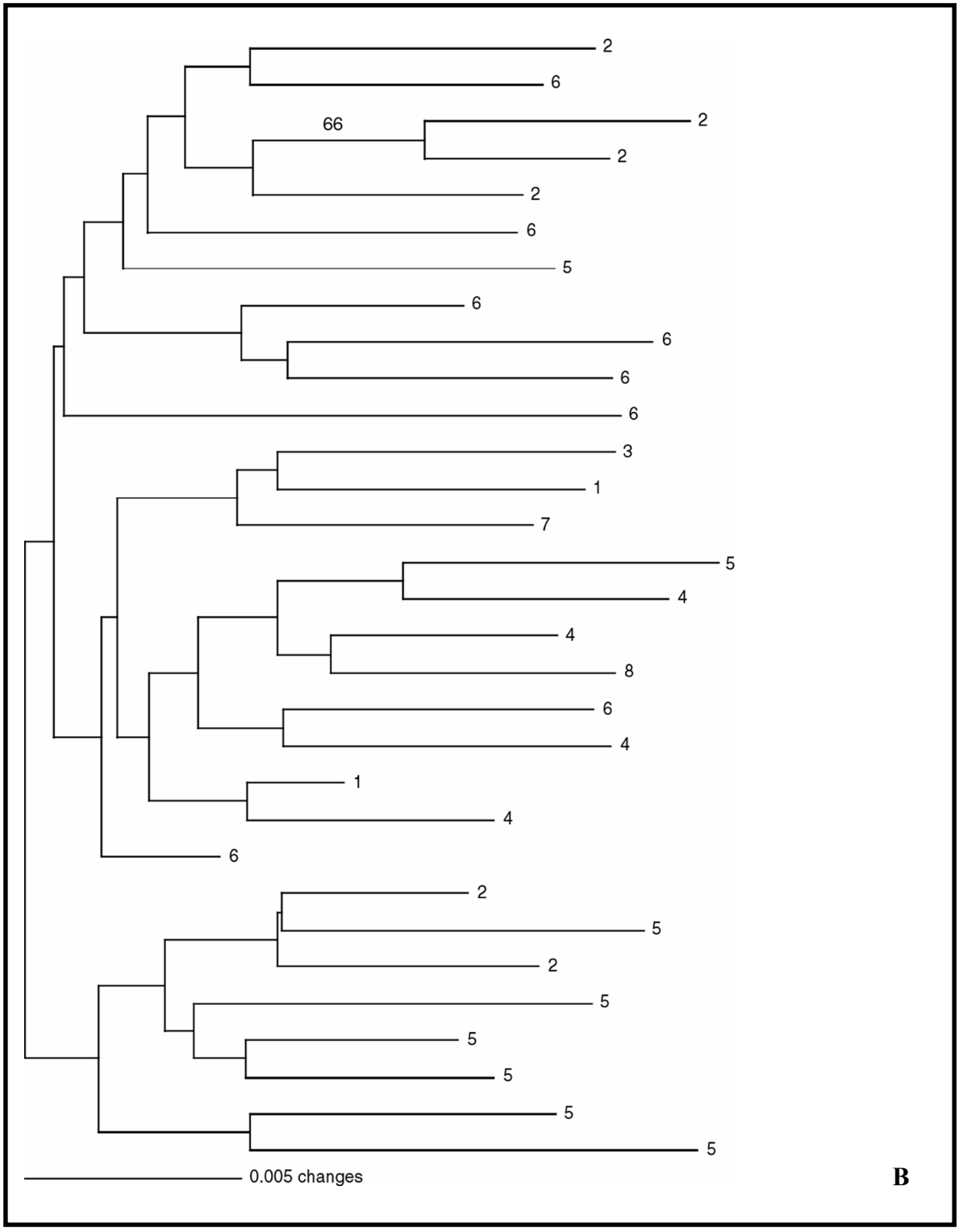
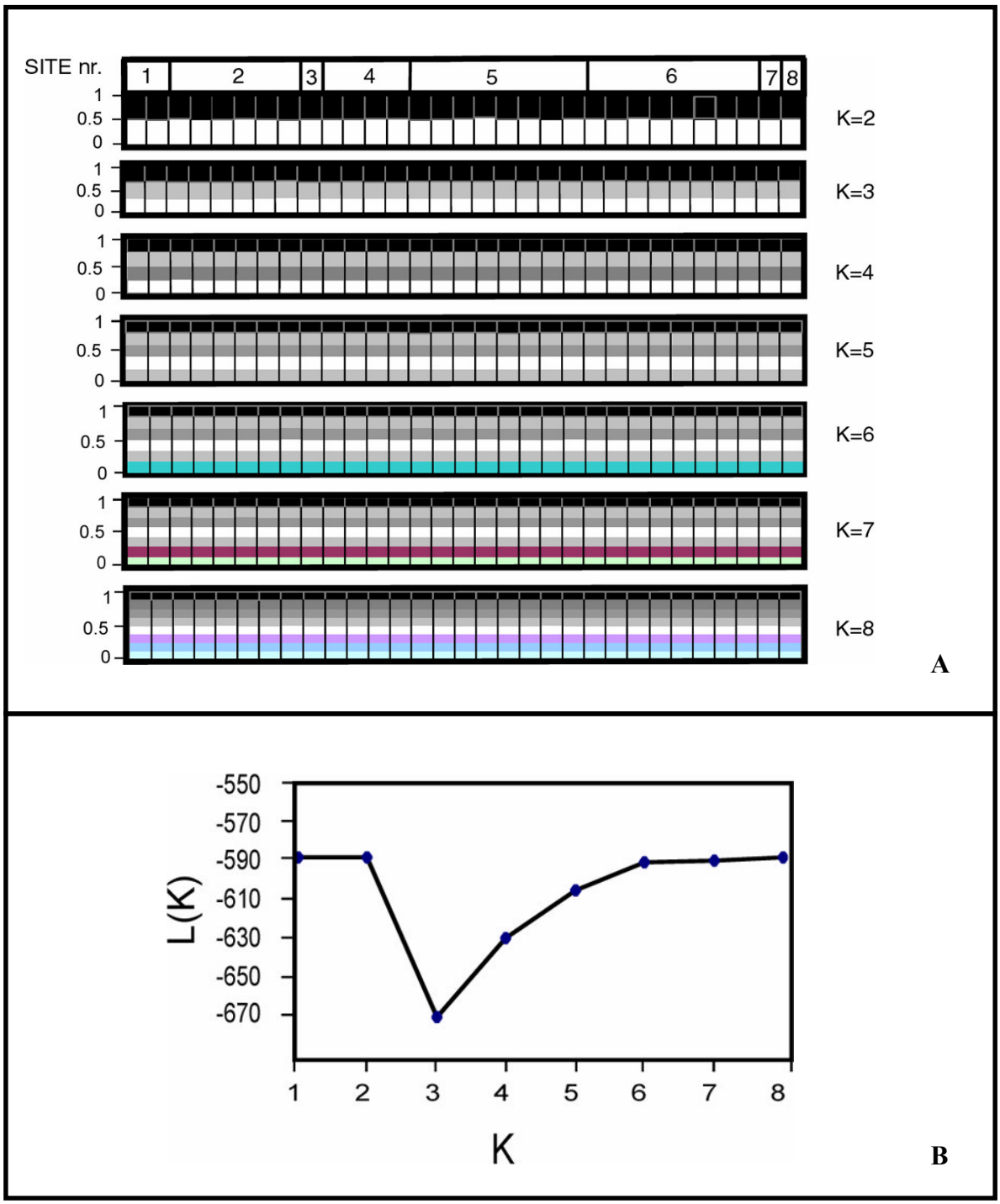


Figure 5.3 AFLP Structure of *A. maderensis* estimated with Bayesian Clustering with STRUCTURE. **A**: Assignment of 31 individuals into K groups. Every individual is represented with a vertical bar divided in different shaded intensities corresponding to the estimated assignment probabilities to each group. Sampled sites are numbered as in Table 5.1. Ten replicates at each K produced nearly identical assignment patterns. The ones represented here are the highest probability runs at each K. **B**: Log probability of the data as a function of K averaged over 10 STRUCTURE runs from K=1 to K=8.



5.3.3 Genetic diversity

When the total diversity estimates of *A. maderensis* are compared with each of the genetic lineages of the widespread congener *A. pungens,* diversity of *A. maderensis* resulted to be slightly lower than that in the lineage with the lowest diversity of *A. pungens* (lineage IV; Table 5.3). This pattern holds for percentage of polymorphisms, band richness, Nei´s gene diversity and Shannon index but not for the Bayesian diversity estimates. The remaining lineages of *A. pungens* showed significantly higher diversity levels for the mentioned parameters, especially lineages I, II and III followed by intermediate levels of diversity in lineage VI. In every case, correction of Nei's index for dominant markers using estimation of allele frequencies by Lynch and Milligan's (1994) method gave almost identical results than based on the squared root of the frequency of the recessive allele (results not shown). The conclusion of a lower diversity of *A. maderensis* as compared to lineages within *A. pungens,* drawn from the total diversity measures, hold for all the five diversity estimates used on a per-population basis, as shown in Table 5.4. However, average within-population diversity of *A. maderensis* must be considered cautiously, since the extremely restricted distribution only allowed for appropriate population sampling in three sites. For comparison, we estimated the diversity levels of the different lineages of *A. pungens* including also those bands that were absent in all individuals in one lineage. In this case, the diversity levels of the different lineages of *A. pungens* slightly decreased but were still higher that in *A. maderensis* (results not shown).

The unbiased within-population gene diversity of Nei (1978) for *A. maderensis* averaged over populations 2, 5 and 6 was 0.19.

Table 5.3 Comparison of the total AFLP diversity *Armeria maderensis* and the widespread congener *Armeria pungens*. The unique Bayesian genetic lineage of *A. maderesis* and five Bayesian lineages of *A. pungens* (recalculated from chapter 2) were analysed. Sample size was standardised to 15 individuals according to the size of the smallest lineage. Standard deviation is reported in brackets.

N=15	*[1]P	*[2]A	*[3]H_T Nei	*[4]H_T Bayes	*[5]Sh
A. maderensis	**47.77**	**1.46**	**0.08** (0.14)	**0.44** (0.06)	**0.14** (0.20)
A. pungens lineage I (mainland+ island)	73.91	1.75	0.21 (0.18)	0.45 (0.005)	0.33 (0.26)
A. pungens lineage II (mainland)	81.81	1.78	0.21 (0.17)	0.43 (0.005)	0.33 (0.23)
A. pungens lineage III (mainland)	83.9	1.84	0.24 (0.18)	0.45 (0.005)	0.37 (0.24)
A. pungens lineage IV (mainland)	**51.85**	**1.58**	**0.14** (0.18)	**0.43** (0.006)	**0.22** (0.26)
A. pungens lineage VI (Island)	77.44	1.63	0.19 (0.18)	0.43 (0.006)	0.30 (0.25)

*[1]P= percentage of polymorphic loci

*[2]A = Bayesian band richness (Rarefaction method of El Mousadik and Petit. 1996)

*[3]H_T Nei = total gene diversity of Nei (1973)

*[4]Sh = allele similarity Shannon index (1948)

*[5]H_T Bayes = Bayesian total diversity

Table 5.4 Within-population AFLP diversity of the only Bayesian lineage of *A. maderensis* and five Bayesian lineages of *A. pungens*. Standard deviation is reported in brackets. Population sample sizes was standardised to 6 individuals.

N=6	*[1]P	*[2]A	*[3]H_s Nei	*[4]H_s Bayes	*[5]Sh
A. maderensis					
Pop. 2	17.78	1.178	0.06 (0.14)	0.1509 (0.01)	0.09 (0.20)
Pop. 5	24.44	1.273	0.07 (0.14)	0.176 (0.01)	0.11 (0.20)
Pop. 6	27.78	1.256	0.06 (0.11)	0.1606 (0.01)	0.12 (0.19)
Mean	**22.33**	**1.24**	**0.06**	**0.16**	**0.11**
***A. pungens* lineage I**					
Pop. 1	43.47	1.377	0.14 (0.19)	0.42 (0.01)	0.21 (0.28)
Pop. 2	50.43	1.444	0.18 (0.26)	0.42 (0.01)	0.20 (0.29)
Mean	**46.95**	**1.41**	**0.16**	**0.42**	**0.20**
***A. pungens* lineage II**					
Pop. 3	51.51	1.497	0.18 (0.20)	0.42 (0.006)	0.27 (0.29)
Pop. 4	35.61	1.352	0.11 (0.17)	0.42 (0.006)	0.17 (0.25)
Pop. 5	43.18	1.447	0.13 (0.18)	0.42 (0.006)	0.21(0.26)
Mean	**43.43**	**1.43**	**0.14**	**0.42**	**0.22**
***A. pungens* lineage III**					
Pop. 6	50.00	1.500	0.18 (0.20)	0.42 (0.007)	0.26 (0.29)
Pop. 7	54.24	1.529	0.19 (0.20)	0.42(0.007)	0.28 (0.29)
Mean	**52.12**	**1.51**	**0.19**	**0.42**	**0.27**
***A. pungens* lineage IV**					
Pop. 8	26.67	1.264	0.09 (0.17)	0.42 (0.005)	0.14 (0.25)
Pop. 9	38.52	1.388	0.14 (0.19)	0.42(0.005)	0.20 (0.28)
Pop. 10	23.70	1.239	0.09 (0.17)	0.42 (0.005)	0.13 (0.24)
Pop. 11	20.00	1.227	0.07 (0.16)	0.42 (0.005)	0.11 (0.23)
Pop. 12	34.07	1.308	0.11 (0.18)	0.42(0.005)	0.17 (0.25)
Pop. 13	25.18	1.456	0.09 (0.18)	0.42 (0.005)	0.14 (0.25)
Mean	**28.02**	**1.31**	**0.10**	**0.42**	**0.15**

Table 5.4 (Continued) Within-population AFLP diversity of the only Bayesian lineage of *A. maderensis* and five Bayesian lineages of *A. pungens*. Standard deviation is reported in brackets. Population sample sizes was standardised to 6 individuals.

***A. pungens* lineage VI**					
Pop. 15	50.00	1.465	0.16 (0.19)	0.43 (0.005)	0.25 (0.27)
Pop. 16	29.27	1.291	0.11 (0.18)	0.43 (0.005)	0.16 (0.26)
Pop. 17	39.02	1.417	0.13 (0.18)	0.43 (0.005)	0.19 (0.26)
Pop. 18	28.66	1.359	0.10 (0.18)	0.43 (0.005)	0.15 (0.26)
Pop. 19	42.07	1.396	0.15 (0.19)	0.43 (0.005)	0.22 (0.28)
Pop. 20	42.68	1.427	0.15 (0.19)	0.43 (0.005)	0.22 (0.28)
Pop. 21	35.37	1.344	0.13 (0.19)	0.43 (0.005)	0.19 (0.27)
Pop. 22	35.98	1.331	0.13 (0.19)	0.43 (0.005)	0.20 (0.28)
Pop. 23	31.71	1.330	0.11 (0.17)	0.43 (0.005)	0.16 (0.25)
Mean	**37.20**	**1.37**	**0.13**	**0.43**	**0.19**

*[1]P= percentage of polymorphic loci

*[2]A = Bayesian band richness (Rarefaction method of El Mousadik and Petit. 1996)

*[3]H_S Nei = within-population gene diversity of Nei (1973)

*[4]Sh = allele similarity Shannon index (1948)

*[5]H_S Bayes = Bayesian winthi-population diversity

5.4 DISCUSSION

5.4.1 Is *Armeria maderensis* genetically depauperate?

5.4.1.1 *Genetic diversity comparisons*

To assess if *A. maderensis* is genetically depauperate we need to compare its genetic diversity with a reference value. In plants, two approaches have been followed for interspecific comparisons. The most frequent one is the direct comparison of genetic diversity parameters of extant species obtained with protein markers (Hamrick and Godt, 1989), RAPD, AFLP or SSR (Nybom, 2004). More recently, a second approach intending to control for phylogenetic constraints has been applied.

Direct comparisons of diversity across studies have allowed defining general limits for plant species diversity, which can be used as reference values in conservation projects. In their review of studies based on AFLP markers, Nybom (2004) reported

ranges of within-population unbiased Nei's gene diversity in plants between Hs=0.15 and Hs=0.31, which places *A. maderensis,* with Hs=0.19, close to the lowest limit. However, for comparability purposes Nybom´s calculations were based only on polymorphic markers, which probably overestimates the diversity of a plant like *A. maderensis*. This also explains the inconsistency of the results produced by HICKORY. The frequent monomorphic markers in *A. maderensis* determine low levels of genetic diversity and therefore removing them implies an overestimation of the variation.

These meta-analyses have found significant differences associated to life history traits. Therefore, the classic procedure has been to compare diversity values for a particular species with average diversity values estimated for plants with similar life history traits, (e.g., perennial outcrossing in the case of *A. maderensis*). However, if the phylogenetic position is not accounted for, direct comparisons across species can produce biased results. Empirical similarity of population genetic parameters within related taxa was found based on allozymes by Gitzendanner and Soltis (2000) and cpDNA by Petit *et al.* (2005). Therefore, the control for phylogenetic constraints is a suitable approach. For this purpose, two alternative procedures have been proposed. Comparison with the most closely related widespread congener, if a phylogeny is available, is suggested in Gitzendanner and Soltis (2000). Another possibility is to perform phylogenetically independent contrasts (PIC; Felsenstein, 1985). Aguinagalde *et al.* (2005) used this method to assess cpDNA Gst trends in European temperate trees and shrubs and found little association with life history traits when phylogenetic position was accounted for, thus challenging the conclusions of former meta-analyses (Petit *et al.*, 2005). In addition to the phylogenetic position, which can be regarded as the pre-speciation history of a species, the intraspecific phylogeographic history (after speciation) is also reflected in the genetic structure. Significant factors in shaping the genetic architecture of a species are for instance one or several introductions of ancestors into oceanic islands, founder events, historical low population sizes, differentiation of intraspecific lineages by spatial isolation or hybridisation events.

Given the unresolved phylogenetic position of *A. maderensis* (Fuertes Aguilar and Nieto Feliner, 2003), its origin can only be traced back to an unkown Western Mediterranean ancestor. Following suggestions by Gitzendanner and Soltis (2000) to

correct for phylogenetic bias, we have chosen one of these western Mediterranean congeners, *A. pungens,* for comparison. This allows for a more effective evaluation of the diversity levels of *A. maderensis* than the direct contrast with reference values derived from comparisons across species (within life history categories, within archipelagos, etc).

Comparison with *A. pungens* has more advantages. *A. pungens* has a mainland distribution as well as two disjunct areas on continental islands, which provides a reference of genetic diversity for *Armeria* on islands (Table 5.3). In addition, its phylogeographic history is well known based on morphological and phylogeographic data (chapters 2 and 3). *A. pungens* presents an Atlantic lineage that has evolved in SW Portugal for a long time, as indicated by its high genetic diversity (Lineages II and III, Table 5.3). The remaining distributional areas in the Atlantic (lineages I and IV) as well as the disjunct Mediterranean range (lineage VI) have originated through recent expansions from the original Portuguese populations. Colonization events, either along the coast or by long-distance dispersal, meant the loss of genetic diversity, especially in the case of the Gulf of Cadiz (lineage IV; Table 5.3), and to a lesser extent in Corsica-Sardinia islands (lineage VI, Table 5.3). The loss of genetic diversity is stronger in the Gulf of Cadiz probably because colonisation was linked to adaptation to different environmental conditions in this area. The genetic composition of the population of the Cies islands (pop.1, lineage I) is probably complicated by introgression events (chapter 3), which may explain its relatively high diversity levels (Table 5.3). The fact that genetic diversity in *A. maderensis* is slightly lower than in the extremely impoverished lineage IV of *A. pungens* (Table 5.3), strongly suggests a bottleneck in the history of *A. maderensis*.

Aside from the above mentioned historical factors, other aspects should be taken into account when comparing genetic diversity across species. These include loci choice, sampling strategies or the way of calculating population diversity parameters, among others. Concerning the latter, it matters if, e.g. all loci or only polymorphic loci are considered (Culley *et al.*, 2002), the technique used for calculation of Shannon index or how diversity is averaged over different loci. In our case we are using exactly the same AFLP protocols and even the same selective primer combinations in both *A.*

maderensis and *A. pungens* studies. Sampling sizes and population genetic parameters have also been standardised.

In sum, both comparing diversity estimates with other studies based on AFLP and with the widespread congener (using exactly the same markers, protocols, analyses) the conclusion that *A. maderensis* has a remarkably low diversity is beyond any reasonable doubt. The fact that AFLP markers have been obtained using two different enzyme combinations increased the genomic regions represented in the fingerprints.

5.4.1.2 *Genetic structure*

In addition to low diversity levels, no intraspecific genetic structure is currently observed within *A. maderensis*. Recently, the Madeiran endemic grass *Deschampsia maderensis*, one of the companions of *A. maderensis* in typical *Armerio maderensis-Parafestucetum albidae* communities, has also been investigated using AFLP markers (P. Catalán *et al.*, unpubl.). In contrast to our study, the wind pollinated *D. maderensis* showed strong intraspecific differentiation between two neighbouring lineages located on northern and southern slopes. This differential pattern may be due to the adaptation of grass species to grazing pressures. In addition, *D. maderensis* has a slightly larger distributional area than *A. pungens* as well as a wider altitudinal range (1300-1860 m). Whatever the causes for such differences may be, concurrence of species with different phylogeographic histories in the same specialized habitat has been documented (Schönswetter *et al.*, 2006). A similar lack of genetic structure using allozymes and ITS data was found in *Matthiola maderensis,* despite conspicuous morphological differences among populations from central, northern and southern slopes (Caujapé and Marrero, personal communication).

5.4.2 Threats for *Armeria maderensis* and future prospects

We considered the possibility that reduced genetic diversity in *A. maderensis* is actually reflecting a shift to autogamy during establishment after long-distance dispersal. The breakdown of self-incompatibility is observed in *Armeria* in circumpolar areas. It has been hypothesised that this mechanism favours establishment after long distance dispersal or in areas poor in pollinators (Baker, 1966). However, two different pollen-stigma morphs that are associated to the incompatibility system in *Armeria* have

been detected in *A maderensis* plants (G. Nieto Feliner, personal observation), suggesting that incompatibility has been maintained after colonisation. Mechanisms favouring outcrossing have been reported for other Macaronesian species (Francisco-Ortega *et al.*, 2000).

Therefore, the low genetic diversity observed in *A. maderensis* does not seem to result from its breeding system but from population declines. The fact that *A. maderensis* has started to occupy more accessible horizontal pastures since the removal of goats in 2003, suggests that the confinement to vertical habitats (mangas) in the last decades was induced by the long-term pressure of goats. This points at grazing as the most probable cause of low genetic diversity levels in *A. maderensis*, although no fragmentation of the range has taken place. As indicated by the strong distinctness of *A. maderensis* from its closest mainland relatives (Fuertes Aguilar and Nieto Feliner, 2003), the colonisation of Madeira is probably old and thus the associated founder event cannot account for the current low diversity revealed by our data.

The observed loss of genetic diversity may raise the risk of inbreeding depression, which could reduce survival and fecundity in the short term. However, to assess accurately the amount of inbreeding, co-dominant markers would be required. Long-term adaptation to environmental changes might be also compromised (Frankham *et al.*, 2002; Peterson and McCracken, 2005). To prevent these risks, we recommend monitoring the populations of *A. maderensis* in the following years, to survey its recovery. Future genetic assessments of *A. maderensis* may be performed and the present genetic study used as a reference of the diversity levels immediately before eradication of grazing.

If the evolution of the populations was observed to be non appropriate, the observed lack of genetic structure would allow for a reinforcement of subpopulations with mixed seeds without danger of outbreeding depression. Although no germination tests have been performed for *A. maderensis*, germination rates of seeds are usually high in *Armeria* (Woodell and Dale, 1974), and are maximized when seeds are placed in the refrigerator for one month (personal observation).

In conservation projects on islands involving the removal of herbivores, a general increase of the cover of vegetation or specific richness has been quantified (Mueller-Dombois and Spatz, 1975; Coblentz, 1977; Schofield, 1989; Lorvelec and Pascal, 2005). More specific responses have also been reported associated with particular features, such as the increase of palatable plants (Donlan *et al.*, 2002). Aside the practical application of this genetic study to the conservation of *A. maderensis*, this could be a good opportunity to address the theoretical question of how genetic variation parallels the recovery in number of individuals after a population bottleneck.

5.5 REFERENCES

Aguinagalde I, Hampe A, Mohanty A*, et al.* (2005) Effects of life-history traits and species distribution on genetic structure at maternally inherited markers in European trees and shrubs. *Journal of Biogeography* **32**, 329-339.

Andrada EC (1990) *Repovoamento florestal no Arquipélago da Madeira (1952-1975)*. Direcção-Geral das Florestas, Secretaria de Estado da Agricultura, Ministério da Agricultura Pescas e Alimentação, Lisboa.

Bagley MJ, Anderson SL, May B (2001) Choice of methodology for assessing genetic impacts of environmental stressors: Polymorphism and reproducibility of RAPD and AFLP fingerprints. *Ecotoxicology* **10**, 239-244.

Baker HG (1966) The evolution, functioning and breakdown of heteromorphic incompatibility systems, I. The Plumbaginaceae. *Evolution* **20**, 349–368.

Baldwin B, Sanderson M (1998) Age and rate of diversification of the Hawaiian silversword alliance (Compositae). *Proceedings of the National Academy of Sciences USA* **95**, 9402-9406.

Bonin A, Bellemain E, Bronken Eidesen P*, et al.* (2004) How to track and assess genotyping errors in population genetics studies. *Molecular Ecology* **13**, 3261-3273.

Carine MA, Russell SJ, Santos-Guerra A, Francisco-Ortega J (2004) Relationships of the Macaronesian and Mediterranean floras: molecular evidence for multiple colonizations into Macaronesia and back-colonization of the continent in *Convolvulus* (Convolvulaceae). *American Journal of Botany* **91**, 1070-1085.

Charbonnel N, Angers B, Rasatavonjizay R, *et al.* (2002) The influence of mating system, demography, parasites and colonization on the population structure of *Biomphalaria pfeifferi* in Madagascar. *Molecular Ecology* **11**, 2213-2228.

Coart E, Glabeke SV, Petit RJ, Bockstaele EV, Roldán-Ruiz I (2005) Range wide versus local patterns of genetic diversity in hornbeam (*Carpinus betulus* L.). *Conservation Genetics* **6**, 259-273.

Coblentz BE (1977) Some range relationships of feral goats on Santa Catalina Island, California. *Journal of Range Management* **30**, 415-419.

Coblentz BE (1978) The effects of feral goats (*Capra hircus*) on island ecosystems. *Biological Conservation* **13**, 279-286.

Corander J, Marttinen P, Mäntyniemi. S (2006) Bayesian identification of stock mixtures from molecular marker data. *Fishery Bulletin*, **in press**.

Costa JC, Capelo J, Jardim R, *et al.* (2004) Catálogo sintaxonómico e florístico das comunidades vegetais da Madeira e Porto Santo. In: A paisagem vegetal da Ilha da Madeira (ed. Capelo J). *Quercetea* **6**, 61-183.

Cowie RH, Holland BS (2006) Dispersal is fundamental to biogeography and the evolution of biodiversity on oceanic islands. *Journal of Biogeography* **33**, 193-198.

Cronk QCB (1997) Islands: stability, diversity, conservation. *Biodiversity and Conservation* **6**, 477-493.

Cuevas JG, Le Quesne C (2006) Low vegetation recovery after short-term cattle exclusion on Robinson Crusoe Island. *Plant Ecology* **183**, 105-124.

Culley TM, Wallace LE, Gengler-Nowak KM, Crawford DJ (2002) A comparison of two methods of calculating G_{ST}, a genetic measure of population differentiation. *American Journal of Botany* **89**, 460-465.

Donlan CJ, Tershy BR, Croll DA (2002) Islands and introduced herbivores: conservation action as ecosystem experimentation. *The Journal of Applied Ecology* **39**, 235-246.

El Mousadik A, Petit R (1996) High level of genetic differentiation for allelic richness among populations of the argan tree [*Argania spinosa* (L.) Skeels] endemic to Morocco. *Theoretical and applied genetics* **92**, 832–839.

Ellstrand NC, Schierenbeck KA (2000) Hybridization as a stimulus for the evolution of invasiveness in plants? *Proceedings of the National Academy of Sciences USA* **97**, 7043-7050.

Emerson BC (2002) Evolution on oceanic islands: molecular phylogenetic approaches to understanding pattern and process. *Molecular Ecology* **11**, 951-966.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software structure: a simulation study. *Molecular Ecology* **14**, 2611-2620.

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

Falush D, Stephens M, Pritchard JK (2007) Inference of population structure using multilocus genotype data: Dominant markers and null alleles**.** *Molecular Ecology Notes* **7,** 574-578.

Felsenstein J (1985) Phylogeny and the comparative method. *The American naturalist* **125**, 1-15.

Fowler de Neira LE, Johnson MK (1985) Diets of giant tortoises and feral burros on Volcan Alcedo, Galápagos. *The Journal of Wildlife Management* **49**, 165–169.

Francisco-Ortega J, Santos-Guerra A, Kim S-C, Crawford DJ (2000) Plant genetic diversity in the Canary Islands: a conservation perspective. *American Journal of Botany* **87**, 909-919.

Frankham J, Ballou JD, Briscoe DA (2002) *Introduction to Conservation Genetics.* Cambridge University Press, Cambridge.

Fuertes Aguilar J, Nieto Feliner G (2003) Additive polymorphisms and reticulation in an ITS phylogeny of thrifts (*Armeria*, Plumbaginaceae). *Molecular Phylogenetics and Evolution* **28**, 430-447.

Gaudeul M, Taberlet P, Till-Bottraud I (2000) Genetic diversity in an endangered alpine plant, *Eryngium alpinum* L. (Apiaceae), inferred from amplified fragment length polymorphism markers. *Molecular Ecology* **9**, 1625-1637.

Geldmacher J, vand den Bogaard P, Hoernle K, Schmincke H (2000) The $^{40}Ar/^{39}Ar$ age dating of the Madeira Archipelago and hotspot track (eastern North Atlantic). *Geochemistry, Geophysics, Geosystems* **1**, Paper number 1999GC000018.

Gitzendanner MA, Soltis PS (2000) Patterns of genetic variation in rare and widespread plant congeners. *American Journal of Botany* **87**, 783-792.

Hamrick J, Godt M (1989) Allozyme diversity in plant species. In: *Plant Population Genetics, Breeding and Genetic Resources* (eds. Brown A, Clegg M, Kahler A, Weir B), pp. 43–63. Sinauer, Sunderland MA.

Holmgren M (2002) Exotic herbivores as drivers of plant invasion and switch to ecosystem alternative states. *Biological Invasions* **4**, 25-33.

Holsinger KE, Lewis PO, Dey DK (2002) A Bayesian approach to inferring population structure from dominant markers. *Molecular Ecology* **11**, 1157-1164.

Holsinger KE, Lewis PO (2005) *Hickory: A Package for Analysis of Population Genetic Data v1.0.4*, University of Connecticut.

Jardim R, Francisco D (2000) *Flora endémica da Madeira.* Múchia Publicaçôes.

Juan C, Emerson BC, Oromí P, Hewitt GM (2000) Colonization and diversification: towards a phylogeographic synthesis for the Canary Islands. *Trends in Ecology & Evolution* **15**, 104-108.

Kambysellis M, Craddock E (1997) Ecological and reproductive shifts in the diversification of the endemic Hawaiian *Drosophila*. In: *Molecular Evolution and Adaptive Radiation* (eds. Givnish T, Sytsma K), pp. 475-509. Cambridge University Press, Cambridge, UK.

Lorvelec O, Pascal M (2005) French attempts to eradicate non-indigenous mammals and their consequences for native biota. *Biological Invasions* **7**, 135-140.

Lynch M, Milligan B (1994) Analysis of population genetic structure with RAPD markers. *Molecular Ecology* **3**, 91-99.

Mesquita S, Capelo J, Sousa J (2004) Bioclimatologia da Ilha da Madeira: abordagem numérica. In: *A Paisagem Vegetal da Ilha* (ed. Capelo J).

Miller MP (1997) *Tools for Population Genetic Analyses (TFPGA). A Windows program for the analysis of the allozymes and molecular genetic data*. Department of Biological Science. Northern Arizona University.

Mueller-Dombois D, Spatz G (1975) The influence of feral goats on lowland vegetation in Hawaii Volcanoes National Park. *Phytocoenologica* **3**, 1-29.

Nei M (1973) Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences USA* **70**, 3321-3323.

Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* **89**, 583-590.

Nei M (1987) *Molecular evolutionary genetics.* Columbia University Press, New York.

Nei M, Li WH (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences USA* **76**, 5269-5273.

Nieto Feliner G, Fuertes Aguilar J, Rosselló JA (2001) Can extensive reticulation and concerted evolution result in a cladistically structured molecular data set? *Cladistics* **17**, 301–312.

Nybom H (2004) Comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. *Molecular Ecology* **13**, 1143-1155.

Peterson PM, McCracken CL (2005) Genetic consequences of reduced diversity: heterozytosity loss, inbreeding depression and effective population size. In: *Plant Conservation. A Natural History Approach* (eds. Krupnick GA, Kress WJ). The University of Chicago Press.

Petit RJ, Duminil J, Fineschi S*, et al.* (2005) Comparative organization of chloroplast, mitochondrial and nuclear diversity in plant populations. *Molecular Ecology* **14**, 689-701.

Press JR, Short MJ (1994) *Flora of Madeira*. The Natural History Museum, London.

Primack RB (1998) *Essentials of Conservation Biology.* Sinauer Associates, Sunderland, MA.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* **155**, 945-959.

Rees D, Emerson B, Oromí P, Hewitt G (2001) The diversification of the genus *Nesotes* (Coleoptera: Tenebrionidae) in the Canary Islands: evidence from mtDNA. *Molecular Phylogenetics and Evolution* **21**, 321-326.

Rieseberg LH, Swensen S (1996) Conservation genetics of endangered island plants. In: *Conservation Genetics. Case Histories from Nature* (eds. Avise J, Hamrick JL), pp. 305-327. Chapman & Hall, New York.

Rohlf FJ (1998) *NTSYS-PC numerical taxonomy and multivariate analysis system*. Exeter Publications, Setauket, N.Y.

Sato A, Tichy H, O'hUigin C*, et al.* (2001) On the Origin of Darwin's Finches. *Molecular Biology and Evolution* **18**, 299-311.

Schofield EK (1989) Effects of introduced plants and animals on island vegetation: examples from Galápagos Archipelago. *Conservation Biology* **3**, 227-239.

Schönswetter P, Popp M, Brochmann C. (2006) Rare arctic-alpine plants of the European Alps have different immigration histories: the snow bed species *Minuartia biflora* and *Ranunculus pygmaeus. Molecular Ecology* **15**, 709–720.

Shannon CE (1948) A mathematical theory of communication. *Bell Systems Technical Journal* **27**, 379-423, 623-656.

Sousa P (2003) 50 anos a servir a floresta. Governo Regional da Madeira. Secretaria Regional do Ambiente e Recursos Naturais. Direcção Regional de Florestas, Funchal.

Silva FA, Menezes CA (1946) *Elucidário Madeirense*, Fac-símile 1998 edn., Funchal.

Silva M, Menezes D, King S, *et al.* (2007) Vascular plant communities in the mountains of Madeira (Portugal). Diversity and vegetation recovery after 4 years without grazing. **Unpublished**.

Stuessy TF, Swenson U, Crawford DJ, *et al.* (1998) Plant conservation in the Juan Fernández Archipelago, Chile. *Aliso* **16**, 89-102.

Swofford D (2002) *Phylogenetic Analysis Using Parsimony (and Other Methods), Version 4*. Sinauer Associates, Sunderland MA.

Vieira R (1992) *Flora da Madeira : o interesse das plantas endémicas macaronésicas*. Reservas e Conservaçâo da Natureza.

Wagner WL, Mix D, Price J (2005) Pacific oceanic islands. In: *Plant Conservation. A Natural History Approach* (eds. Krupnick GA, Kress WJ). The University of Chicago Press.

WCMC - World Conservation Monitoring Centre (1992) *Global biodiversity: status of the earth's living resources*. Chapman and Hall, London.

Winkworth RC, Robertson AW, Ehrendorfer F, Lockhart PJ (1999) The importance of dispersal and recent speciation in the flora of New Zealand. *Journal of Biogeography* **26**, 1323-1325.

Wolfe AD, Liston A (1998) Contributions of the polymerase chain reaction to plant systematics. In: *Molecular Systematics of Plants II. DNA Sequencing.* (eds. Soltis DE, Soltis PS, Doyle JJ), p. 43–86. Kluwer Academic Publishers, New York.

Woodell SRJ, Dale A (1993) Armeria maritima (Mill.) Willd. (*Statice armeria* L.; *S. maritima* Mill.). *Journal of Ecology* **81**, 573-588.

Yeh FC, Boyle TJB (1997) Population genetic analysis of codominant and dominant markers and quantitative traits. *Belgian Journal of Botany*, **129**, 157.

6. DISCUSIÓN GENERAL: FILOGEOGRAFÍA DE *Armeria pungens* y *Armeria maderensis*

6.1 UTILIZACIÓN E INTERPRETACIÓN DE LOS MARCADORES PARA INFERENCIA FILOGEOGRÁFICA EN *Armeria pungens*

Los estudios de filogeografía en plantas han sido menos exitosos que en animales debido a la dificultad de hallar marcadores moleculares neutros adecuados para el nivel intraespecífico (Schaal y Olsen, 2000). La diferencia fundamental es que el ADN mitocondrial (ADNmt) de animales contiene marcadores suficientemente variables para estudios por debajo del nivel de especie (Avise *et al.*, 1987; McCauley, 1995; Zhang y Hewitt, 2003), mientras que en plantas los genes mitocondriales presentan tasas de substitución extraodinariamente bajas (Demesure *et al*., 1995). Además, la organización del genoma mitocondrial en plantas es muy lábil en los distintos grupos, lo que supone una limitación para el diseño de cebadores (Demesure *et al*., 1995).

De este modo, los marcadores para plantas han de buscarse en los genomas nuclear y cloroplástico. Dado que suelen ser poco variables a nivel intraespecífico, lo más recomendable es utilizar varios marcadores independientes a fin de buscar cambios informativos en una muestra diversa de los genomas (Hare, 2001; Zhang y Hewitt, 2003). Además, el contraste de diversos loci no ligados ayuda a discernir patrones específicos de la evolución del marcador de aquellos que son un reflejo de la evolución del organismo. Ésta es la aproximación que hemos seguido para estudiar la historia evolutiva de *Armeria pungens*, utilizando marcadores AFLP (3 combinaciones de cebadores -223·loci- en 221 individuos), secuencias de ITS (59 individuos), el gen nuclear *GapC* (101 secuencias de 71 individuos), secuencias de tres regiones no codificantes de ADN cloroplástico, ADNcp (336 secuencias de 112 individuos y 120 secuencias de 38 de especies simpátricas) y, finalmente, microsatélites cloroplásticos, SSRcp, (8 loci en 5-24 individuos).

Cuando se manejan múltiples marcadores cada uno puede proporcionar información de distinta relevancia para las cuestiones que nos estemos planteando: diversidad genética y diferenciación, hibridación, filogeografía, poliploidía, filogenia,

genotipado de individuos, etc. Asimismo, cada tipo de marcador ha de ser analizado de acuerdo a sus características propias: nivel de polimorfismos, evolución (familia multigénica, evolución concertada, etc.), forma de transmisión (mono o biparental), herencia dominante o codominante o tendencia a retención de polimorfismos ancestrales.

6.1.1 Herencia materna *vs.* biparental

Los marcadores nucleares y cloroplásticos difieren en su modo de transmisión y tasas evolutivas, por lo que proporcionan información evolutiva complementaria.

La estructura del ADNcp está altamente conservada, tanto en el orden como en número de genes, lo que permite el diseño de cebadores universales que amplifican regiones no codificantes (Demesure *et al*., 1995; Taberlet *et al*., 1991; Grivet *et al*., 2001). Los genes cloroplásticos son esencialmente de copia simple (Havey *et al*., 1998) dentro del genoma, aunque en cada célula hay un elevado número de copias todas iguales. Por ello, no suelen causar problemas de mezcla de genes ortólogos (similitud debida a antepasado común) y parálogos (similitud debida a duplicación). El ADN nuclear, en cambio, tiene un gran tamaño y complejidad, incluyendo secuencias repetidas que constituyen familias multigénicas. La presencia de copias ortólogas y parálogas de la mayoría de estos genes les resta utilidad filogenética.

El genoma cloroplástico, al ser transmitido sólo as través de la semilla en la mayoría de las angiospermas y poseer tamaños poblacionales efectivos muy pequeños (la molécula de ADNcp tiene tan sólo de 100 a 200 Kb) es muy sensible a los procesos de deriva génica (Birky, 1995; McCauley, 1995; Petit *et al*., 2003; Small *et al*., 2004). Por ello, y a pesar de sus bajas tasas de sustitución, suele mostrar tiempos de coalescencia más cortos que el ADN nuclear, i.e., se fija más rápido dentro de una población (Clegg *et al*., 1994). Esto provoca que los parámetros de diferenciación interpoblacional derivados de marcadores nucleares y cloroplásticos no sean directamente comparables (el genoma cloroplástico tiende a presentar niveles de diferenciación netamente mayores; Maruyama y Fuerst, 1984; Petit *et al*., 1993; McCauley, 1995), aunque sí lo sean las diferencias relativas entre áreas de distribución (Ennos, 1994; Petit *et al*., 2005)

Los menores tiempos de coalescencia otorgan dos ventajas a los marcadores cloroplásticos sobre los genes nucleares, que alcanzan la coalescencia en períodos teóricamente cuatro veces más largos en organismos diploides (Schaal y Olsen, 2000); Hare, 2001). En primer lugar, por debajo del nivel de especie los marcadores nucleares a menudo presentan patrones polifiléticos que no reflejan la historia evolutiva de los organismos. En segundo lugar, los marcadores cloroplásticos, al ser más sensibles a los fenómenos de deriva génica, son mucho mejores indicadores de cuellos de botella y efectos fundadores (McCauley, 1995), cuyo conocimiento ayuda a interpretar la variabilidad genética y la propia historia evolutiva de los organismos.

Debido a esta simplicidad estructural, el ADN cloroplástico ha sido el más utilizado para estudios de filogeografía en plantas (Schaal *et al*., 1998; Soltis *et al*., 2006). No obstante, la herencia uniparental del ADN cloroplástico implica que no proporcione la historia completa del organismo y además, todos los genes cloroplásticos están ligados, mientras que el ADN nuclear proporciona la posibilidad de comparar múltiples genes con historias evolutivas independientes (Wendel y Doyle, 1998; Hare, 2001). Otro inconveniente es que, incluso en las regiones no codificantes o en los genes que evolucionan más rápido dentro del cloroplasto, como *matK* y *ndhF*, las tasas de mutación son bajas (Schaal *et al*., 1998; Sang, 2002; Small *et al*., 2004).

La distinta forma de herencia de los marcadores nucleares y cloroplásticos tiene una interesante aplicación en estudios de evolución reticulada. El ADNcp de los híbridos revela sólo la mitad del parentesco en plantas de origen híbrido o poliploide. Por ello, las identifica como cercanas a uno de los progenitores sin llegar a revelar la historia de hibridación. Si el evento de hibridación va seguido de introgresión y fijación de una molécula de ADNcp de otro linaje (captura cloroplástica) la señal recuperada a partir del ADN nuclear y cloroplástico estarán en conflicto (Arnold, 1997).

6.1.2 Marcadores de huella genética (*fingerprinting*) *vs.* secuencias

La técnica utilizada para su obtención influye tanto en la información que pueden proporcionar los marcadores como en los procedimientos más apropiados para su análisis.

Los métodos de huella genética se basan en el tamaño de fragmentos del genoma. Algunas de las técnicas más populares en filogeografía siguen este principio, como RFLP (*restriction fragment length polymophism*), RAPD (*random amplified polymorphic DNA*), AFLP (*amplified fragment length polymorphisms*). ISSR (*intersimple sequence repeats*) o SSR (*simple sequence repeats*, microsatélites). En todos ellos, al igual que en las isoenzimas, la información ancestral es ambigua (no puede polarizarse), y los patrones genealógicos no pueden deducirse fácilmente (Schaal y Olsen, 2000; Zhang y Hewitt, 2003). También presentan riesgo de falta de homología cuando alelos de tamaños idénticos corresponden a distintas regiones del genoma (Doyle *et al*., 1998; Zhang y Hewitt, 2003) pero éste es un problema de menor entidad en estudios intraespecíficos o entre especies relacionadas. Una ventaja de estos marcadores es que debido a su menor complejidad técnica y coste, permiten investigar un elevado número de loci y/o individuos que normalmente no podría abordarse con secuenciación.

Las secuencias son marcadores de gran calidad, que contienen información útil para reconstruir genealogías. La evolución del ADN intraespecífico, sin embargo, tiene lugar en un marco de reticulación (tokogenético, Henning, 1968), no de aislamiento y divergencia, de modo que los análisis filogenéticos tradicionalmente usados en sistemática molecular no son los más convenientes para filogeográfía (Zhang y Hewitt, 2003). Las redes de haplotipos son más apropiadas, puesto que permiten trabajar con los niveles reducidos de variabilidad genética y los elevados muestreos típicos de las investigaciones filogeográficas, así como representar relaciones reticuladas, e.g. por la presencia de polimorfismos ancestrales junto a los derivados (Schaal y Olsen, 2000; Posada y Crandall, 2001; Zhang y Hewitt, 2003; Vriesendorp y Bakker, 2005). Enraizar una filogenia a estos niveles taxonómicos es muy difícil porque se carece de la perspectiva necesaria para diferenciar la identidad de alelos por descendencia, por transferencia o por aparición independiente. En estos casos, es posible enraizar las redes de haplotipos considerando las predicciones de la teoría de la coalescencia (Posada y Crandall, 2001). Según ésta, los alelos ancestrales corresponden a los más frecuentes y con mayor número de conexiones con otros alelos (Hudson, 1990; Castelloe y Templeton, 1994). Sin embargo, este enfoque no funciona en escenarios donde haya retención de polimorfismos ancestrales o procesos de flujo génico (Schaal *et al*., 1998). Además, el método bootstrap para evaluar el apoyo de los clados pierde eficacia (Zhang

y Hewitt, 2003; Posada y Crandall, 2001). Con datos intraspecíficos la divergencia alélica, la frecuencia alélica y el número de alelos son parámetros clave para reconstruir las relaciones evolutivas (Zhang y Hewitt, 2003).

6.1.3 Marcadores para inferencia filogeográfica en *A. pungens*

6.1.3.1 *La morfología*

La información más valiosa aportada por la morfología es el encuadre de *A. pungens* dentro del género. Los caracteres morfométricos separan de manera muy clara *A. pungens* del resto de especies de *Armeria*, incluidas las especies litorales y sublitorales de la sección *Macrocentron* que ocupan el SW de la Península Ibérica.
Sin embargo, la diferenciación morfológica entre los distintos linajes de *A. pungens* es escasa, lo que limita su eficacia como marcador evolutivo. Tan sólo el indumento de las hojas parece reflejar los distintos linajes que han surgido a lo largo de la evolución de nuestra especie.

6.1.3.2 *AFLP*

Son muy útiles porque tienen una buena cobertura del genoma, fundamentalmente nuclear, proporcionando una gran cantidad de polimorfismos (una media de 238 loci en una revisión de 27 estudios de AFLP; Nybom, 2004), sobre todo cuando se utilizan varias combinaciones de enzimas de restricción. Presentan menos problemas de homología que los RAPD, pues la perfecta hibridación de los primers con los adaptadores permite utilizar elevadas temperaturas de anillamiento en las reacciones de amplificación. Además, requieren protocolos de optimización relativamente sencillos por lo que resultan ideales como primera aproximación filogeográfica a plantas no modelo.

Su gran desventaja es que, como se ha comentado al describir los métodos de huella genética, no permiten hacer inferencia filogenética (Zhang y Hewitt, 2003; Lowe *et al*., 2004), de forma que no informan acerca de las áreas ancestrales o rutas de colonización. Éstas, han de deducirse indirectamente de la comparación de las distancias genéticas entre poblaciones y áreas, así como de los patrones diversidad. Otro inconveniente es su carácter dominante, que impide distinguir los loci heterocigotos de los homocigotos dominantes, por lo que para estimar frecuencias alélicas y parámetros

de la diversidad genética es necesario asumir condiciones de equilibrio. Aunque algunas aproximaciones, como el índice de Shannon (1948) o el porcentaje de loci polimórficos, no requieren estas asunciones, la eficacia de los AFLP para estimas de diversidad ha sido muy debatida (Lowe *et al*., 2004).

En el caso de *A. pungens*, los AFLP proporcionan una representación satisfactoria de los principales linajes de ADN nuclear. Las distancias genéticas entre áreas así como los niveles relativos de diversidad y diferenciación permiten hacer las siguientes hipótesis biogeográficas:

- La proximidad genética de poblaciones **portuguesas** y **sardo-corsas** indica una dispersión reciente a larga distancia. La evidente mayor diversidad y estructura genética en el área portuguesa, apuntan a que éste es el punto de partida de la dispersión.

- El **Golfo de Cádiz** lo ocupa un linaje claramente diferenciado, a pesar de la ausencia de barreras espaciales con las poblaciones portuguesas. La ínfima diversidad genética y considerable diferenciación entre poblaciones de este linaje sugiere que ha sido afectado por deriva génica, asociada a un proceso de colonización reciente o a bajos tamaños poblacionales a más largo plazo. Como la información genética no se puede polarizar, no queda claro cuál de los dos procesos ha originado el depauperado linaje genético del Golfo de Cádiz.

- La población de **Camarinal** se presenta genéticamente muy diferenciada del resto de poblaciones, aunque más próxima a las del Golfo de Cádiz, y con niveles relativamente altos de diversidad con respecto a estas últimas. La hibridación con táxones simpátricos podría explicar este patrón, sin embargo esto sólo puede confirmarse utilizando marcadores de otro tipo.

- Finalmente, la población de las **Islas Cíes** presenta un patrón genético ambiguo. Es genéticamente próxima a las poblaciones más cercanas del continente, lo que podría apuntar a dispersión a larga distancia, pero el grado de divergencia con respecto a éstas es notable. La diversidad genética basada en los AFLP es moderada.

6.1.3.3 *ITS*

Los genes ribosómicos constituyen una familia multigénica. Presentan de cientos a miles de copias organizadas en tándem, potencialmente en uno o varios cromosomas. Cada unidad de repetición contiene los genes 18S, 5.8S y 26S, los espaciadores internos ITS1 e ITS2 y el espaciador intergénico IGS (Baldwin *et al*., 1995; Álvarez y Wendel, 2003).

Una de las peculiaridades del ADN ribosómico nuclear, es que las distintas copias tienden a evolucionar conjuntamente tanto a nivel intragenómico como intraespecífico (o mejor, dentro de grupos reproductivos), mediante mecanismos de homogeneización. Este modo de evolución se denomina "evolución concertada" (Arnheim, 1983; Elder y Turner, 1995). Los mecanismos moleculares responsables son el sobrecruzamiento desigual y la conversión génica, que son mecanismos de corrección entre genes (Dover, 1982; Feliciello *et al*., 2005). Debido a la relativamente baja presión selectiva a que están sometidos en comparación con las secuencias codificantes, los espaciadores ribosómicos presentan una elevada tasa de cambio, alcanzando una divergencia de hasta un 39% entre especies filogenéticamente muy emparentadas (Baldwin *et al*., 1995). No obstante, sufren cierta presión selectiva, ya que están involucrados en la maduración del ARN (Van Nues *et al*., 1994; Bena *et al*., 1998).

Las elevadas tasas de sustitución unidas a la evolución concertada de los ITS, resultan, al tiempo, en bajos niveles de divergencia intraspecífica y elevados niveles de divergencia interespecífica. Las frecuentes mutaciones de los espaciadores ribosómicos se fijarán o perderán por los mecanismos homogeneizadores, favoreciéndose la acumulación de diferencias interespecíficas. Por todo ello, estos marcadores se aplican fundamentalmente al estudio de relaciones filogenéticas entre especies emparentadas.

Sin embargo, en ocasiones la evolución concertada no llega a completarse (Álvarez y Wendel, 2003), lo que conduce a la coexistencia de varias copias de ITS en un mismo organismo. Esto puede ocurrir como resultado de fallos o bajas tasas en los mecanismos de homogeneización (Trick y Dover, 1984), hibridación intraspecífica, reciente divergencia o duplicación de loci funcionales. En el tiempo de homogeneización influyen, por un lado, el número de repeticiones y su disposición en el

genoma y, por otro, la estructura genética de las poblaciones y eventos evolutivos que las afectan.

La evolución concertada no completa ha de ser, pues, tenida en cuenta a la hora de detectar fenómenos como la evolución reticulada, especiación híbrida y origen de poliploides (Rieseberg, 1997; Fuertes Aguilar *et al*., 1999). Cuando varias copias de ITS surgen dentro de un genoma por hibridación, la rapidez y dirección de homogeneización no puede predecirse. Y no es siempre congruente entre linajes (Álvarez y Wendel, 2003).

A pesar de su ya esperada escasa variación a nivel intraespecífico, el análisis de ITS resulta informativo para el estudio de la disyunción de *A. pungens* porque es el marcador mejor conocido en el género *Armeria*. La región ITS presenta en *Armeria* moderada variabilidad entre especies y un patrón de variación coherente con la geografía, e independiente de especie, que se interpreta como debido a la radiación reciente de este género unida a fenómenos de hibridación interespecífica y evolución concertada. Con frecuencia, se observan polimorfismos intraindividuales atribuidos a la evolución concertada no completa tras un evento reciente de hibridación.

En teoría, al tratarse de secuencias usadas para inferencia de filogenias interespecíficas podrían usarse para enmarcar *A. pungens* dentro de relaciones filogenéticas en *Armeria*. Lamentablemente, el bajo grado de variabilidad de ITS tanto en *Armeria* como en *A. pungens* no permite discernir qué especies son las más emparentadas, ya que gran parte de las especies del W de la Península Ibérica, incluidas todas las litorales de la sección *Macrocentron*, presentan el mismo ribotipo, con tan sólo algunas sustituciones entre especies. Por el mismo motivo no es posible establecer firmemente cuáles son las poblaciones ancestrales de *A. pungens*. Aunque el hecho de que las poblaciones disyuntas del Mediterráneo también presenten este ribotipo de W Ibérico sugiere que las poblaciones ancestrales son las de la Península.

Las posiciones polimórficas intraindividuales detectadas en *A. pungens* revelan la existencia de varias copias de ITS dentro de un mismo individuo que difieren entre sí en una o dos bases. Su congruencia con la geografía y con los demás marcadores nucleares

sugiere que contienen señal de eventos evolutivos, ya sea de hibridación o divergencia, que todavía no han sido homogeneizados:

- Los dos polimorfismos exclusivos en el **Golfo de Cádiz** podrían deberse a procesos recientes de reticulación o diferenciación. La segunda hipótesis es apoyada por los demás marcadores moleculares.

- Los polimorfismos exclusivos de **Camarinal**, son compartidos por poblaciones simpátricas de *A. macrophylla*, lo que apunta a la hibridación de ambas especies como causa.

- La comparación con congéneres de las **Islas Cíes** y de **Córcega y Cerdeña,** que poseen ribotipos muy diferenciados de *A. pungens,* revela que no se ha producido introgresión del genoma de otras armerias en las poblaciones isleñas de *A. pungens*. En ambas islas *A. pungens* presenta un ribotipo cercano de las poblaciones portuguesas de *A. pungens* y del resto de especies del W de la Península Ibérica, lo que constituye una excepción al patrón geográfico general de ITS en *Armeria*.

6.1.3.4 *GapC*

Los genes nucleares de bajo número de copia son marcadores de herencia biparental. Presentan gran diversidad de tasas de sustitución, con lo que potencialmente pueden hallarse marcadores con resolución suficiente para estudios a nivel poblacional (Sang, 2002; Small *et al*., 2004). Una de sus mayores ventajas es que permiten comparar marcadores independientes en el genoma nuclear. Así, examinando la congruencia entre loci, pueden distinguirse los patrones que únicamente reflejan la historia de un gen concreto de aquellos que manifiestan la historia de los organismos. Sin embargo, estos marcadores presentan muchas dificultades prácticas, como la necesidad de diseñar cebadores y de evaluar la ortología de cada gen cada vez que se aborda un nuevo estudio (Zhang y Hewitt, 2003; Schlüter *et al*., 2004; Small *et al*., 2004; Hughes *et al*., 2006; Álvarez *et al*., no publicado).

Como se ha mencionado, los marcadores nucleares a menudo presentan retención de polimorfismos ancestrales entre especies cercanas y, como se organizan en familias génicas afectadas por fenómenos de duplicación, el riesgo de amplificar

múltiples parálogos dentro de un mismo organismo se incrementa. Tanto la compartición de polimorfismos ancestrales como la mezcla de secuencias ortólogas y parálogas pueden complicar la inferencia filogenética.

En el caso de *A. pungens*, la variación de *GapC* se muestra profundamente afectada por una mezcla de ambos procesos combinada con bajos niveles de variabilidad, lo que deriva en un árbol genealógico complejo que difumina la señal filogeográfica. La comparación con otros marcadores nucleares descarta que este patrón genético sea causado por hibridación. No obstante, dentro de cada clado se detecta claramente una pauta de variación congruente con la geografía y absolutamente congruente con la morfología y con los otros marcadores de ADN nuclear (AFLP e ITS) reforzando las inferencias filogeográficas:

- Tres alelos (B, D e I) relacionan las poblaciones de **Portugal** con las de **Córcega-Cerdeña**, mientras que otros tres (J, K y L) son exclusivos del **Golfo de Cádiz**. No obstante, el escaso número de posiciones filogenéticamente informativas impide determinar las relaciones filogenéticas entre estos dos linajes.

- La población de **Camarinal** presenta alelos en común con el resto de *A. pungens* (A, B y E) junto a otro divergente (G) que podría haberse incorporado su genoma por hibridación con un especie simpátrica, al tratarse de un gen de herencia biparental.

- Finalmente, la población de las **Islas Cíes** presenta único alelo (C) compartido con la más septentrional de las poblaciones portuguesas en el continente y con una población sarda, un vínculo también reflejado en la variación de ITS y AFLP, así como en la indumentación de las hojas.

6.1.3.5 *Secuencias de ADN cloroplástico*

Con tan sólo algunas excepciones, el ADNcp se divide en dos regiones de copia simple (LSC large single copy y SSC small single copy) separadas por una región invertida (Grivet *et al*., 2001). LSC es la más variable y, por ello, la más útil para estudios a niveles taxonómicos bajos. Su longitud en *Quercus robur* es de 89200 bp y en *Nicotiana tabacum* de 86686 bp. En LSC se localizan las tres regiones de ADN

cloroplástico que hemos secuenciado en *A. pungens*, hasta obtener un total de 3037 bp: *trn*L-F, *trn*S-fM y *mat*K.

Las tasas evolutivas del cloroplasto son generalmente bajas (Olmstead y Palmer, 1994; Sang, 2002); incluso las regiones no codificantes parecen estar sometidas a ciertas constricciones funcionales, de forma que hay zonas más conservadas y otras con mayor incidencia de delecciones y sustituciones. Comprender los mecanismos de mutación del ADNcp es crítico para asegurar la homología del alineamiento de secuencias (e.g. detectar fragmentos invertidos) y para asegurarnos de que estamos utilizando marcadores neutros que no contienen constricciones funcionales (Kelchner, 2000; Simmons y Ochotorena, 2000). Sin embargo, en alineamientos de secuencias de una misma especie o de especies muy relacionadas la homología suele estar asegurada.

El marcador *trnL-F* corresponde al espaciador entre los genes *trn*L(UAA) y *trn*F(GAA). Es muy variable en longitud y número de mutaciones puntuales en función del grupo estudiado, aunque se transcribe, y por ello está relativamente conservado para mantener la estructura secundaria (Quandt y Stech, 2004). La existencia de cebadores universales (Taberlet *et al*., 1991) ha propiciado su secuenciación extensiva en angiospermas, frecuentemente en el marco de estudios poblacionales (e.g. Gutiérrez Larena *et al*., 2002; Alsos *et al*., 2005; Koch *et al*., 2006). La región *trn*S-fM en *Nicotiana sylvestris* (Gene Bank) incluye el espaciador intergénico *trn*S-*psb*Z, el gen *psb*Z, el espaciador *psb*Z-*trn*G, el gen *trn*G y, finalmente, el espaciador intergénico *trn*G-*trn*fM. Posee 1700 bp en *Quercus robur* y 1254 bp en *Nicotiana tabacum*. Finalmente, el gen *matK* se localiza en el interior del intrón *trn*K (Sugita *et al*., 1985; Jonson y Soltis, 1994; Hausner *et al*., 2006). Codifica para una versión degenerada de una proteína retrotranscriptasa, que parece conferir al intrón capacidad para invadir nuevos sitios en el genoma. Probablemente también interviene en el splicing del intrón (Hausner *et al*., 2006). El intrón *trn*K es una de las regiones cloroplásticas de evolución más rápida (Hilu y Liang, 1997; Kelchner, 2000), por eso es común su uso en filogeografía (Sang, 2002). Se ha observado que evoluciona aproximadamente al doble de velocidad que *rbc*L en Polemoniaceae y al triple en Saxifragaceae (Johnson y Soltis, 1994).

En el caso de *A. pungens*, las secuencias de ADNcp constituyen un marcador evolutivo muy valioso. Las rápidas tasas evolutivas de *mat*K se presentan también en *A. pungens*, donde de las 24 posiciones variables en los tres marcadores cloroplásticos analizados, 15 se localizan en esta región.

- Debido a su sensibilidad a los procesos de deriva génica, el ADN cloroplástico permite identificar dentro del área de distribución de *A. pungens* dos zonas que han sido recientemente colonizadas: **Córcega-Cerdeña** y el **Golfo de Cádiz,** cada una con un haplotipo dominante presente en todas sus poblaciones.

- Al ser marcadores ordenados polarizables, es posible establecer una **genealogía** fiable de las secuencias cloroplásticas que apunta a **Portugal** como área ancestral de la planta. Se aprecia, además, que mientras que la colonización de Córcega-Cerdeña ha sido llevada a cabo por un subconjunto de los haplotipos del litoral portugués, la colonización lineal del Golfo de Cádiz fue acompañada de un proceso de divergencia a partir de los haplotipos portugueses más próximos.

- En las poblaciones situadas en los límites norte (**Cíes**) y sur (**Camarinal**) de la especie, la comparación de los patrones de ADN nuclear y cloropástico revela varios procesos de reticulación en distintas épocas. Estos eventos se confirman con la secuenciación de los mismos marcadores cloroplásticos en las otras especies de *Armeria* que habitan estas áreas: (1) introgresión de *A. macrophylla* en la población de Camarinal; (2) existencia de una zona híbrida en el área de contacto entre *A. pungens* y *A. pubigera* en las Islas Cíes; (3) posible hibridación ancestral entre *A. pungens* de Cíes y *A. welwistchii* o *A. berlenguensis*. En este último caso, y dado que todos los genes cloroplásticos están ligados, no es posible descartar completamente que el patrón polifilético de la población de Cíes se deba a la compartición de polimorfismos ancestrales entre dos especies emparentadas.

6.1.3.6 *Microsatélites cloroplásticos (cpSSR)*

Los microsatélites cloroplásticos son repeticiones mono o dinucleotídicas, principalmente (A)n o (T)n (Ishii *et al*., 2001), generalmente localizadas en regiones espaciadoras o intrones del cloroplasto. Como otros marcadores de huella genética,

presentan riesgo de que alelos de tamaño idéntico nos sean homólogos (*i.e.*, que sean idénticos no por divergencia sino por mutación independiente), aunque esto es poco frecuente al nivel taxonómico de nuestro estudio. El análisis de microsatélites ha de realizarse teniendo en cuenta las distancias entre las variantes de tamaño y su frecuencia, asumiendo un modelo de mutación paso a paso por adición o delección de unidades de repetición. De acuerdo a este modelo, las variantes con pequeñas diferencias de longitud están más estrechamente emparentadas que las que poseen mayor diferencia (Slatkin, 1995).

Las tasas de mutación de este tipo de marcadores parece ser muy heterogéneas en función del grupo de organismos o loci considerados. Pero en teoría los microsatélites podrían proporcionar marcadores más variables que las mutaciones puntuales y por ello ser útiles a nivel poblacional (Provan *et al*., 2001; Estoup *et al.*, 2002).

La mayoría de las investigaciones que utilizan microsatélites cloroplásticos se centran en especies de Pinaceae (Provan *et al*., 2001), ya que la secuenciación completa del cloroplasto de *Pinus thunbergii* ha facilitado la creación cebadores para esta familia (Powell *et al*., 1995; Vendramin *et al*., 1996). Posteriormente se diseñaron primers universales para angiospermas (Weising y Gardner, 1999; Chung y Staub, 2003). Todos fueron probados en *A. pungens*, y los protocolos de nueve de ellos consiguieron optimizarse; no obstante, ninguno proporcionó variabilidad suficiente.

6.2 FILOGEOGRAFÍA DE *A. pungens*

6.2.1 Linajes genéticos en *A. pungens*

La apreciación más inmediata de la evaluación de cinco tipos de marcadores moleculares en *A. pungens* es la baja variabilidad general, que sugiere que los procesos evolutivos que han afectado a esta planta son recientes.

Todos los marcadores moleculares estudiados permiten evaluar las distancias genéticas entre poblaciones y, sin excepción, señalan que las áreas de **Portugal y Córcega-Cerdeña** están ocupadas por el mismo linaje, sugiriendo dispersión a larga distancia entre ambas, mientras que el **Golfo de Cádiz** es ocupado por un segundo linaje claramente distinto, tanto genética como morfológicamente (Fig. 6.1). El

indumento de las hojas también distingue ambos linajes: hojas glabras y ciliadas en el nervio medio, respectivamente.

Por otra parte, las poblaciones situadas en los límites septentrional y meridional de la especie presentan características particulares:

- Al norte, la población de las **Islas Cíes** se caracteriza por un haplotipo cloroplástico notablemente divergente del resto de la especie, mientras que su morfología y su genoma nuclear son típicos de *A. pungens*.

- **Camarinal,** en el extremo sur, es una población diferente de las demás, tanto en su morfología (hojas pubérulas) como en sus genomas nuclear y cloroplástico.

6.2.2 Rutas de Colonización: relaciones genealógicas entre poblaciones y patrones de diversidad genética

6.2.2.1 *Córcega-Cerdeña*

Además de la evidente proximidad con Portugal, revelada por todos los marcadores genéticos y por la morfología, el reparto de la diversidad genética entre y dentro de las poblaciones, proporciona información complementaria acerca del proceso de colonización. Entre los marcadores estudiados, los AFLP y las secuencias de cloroplasto son los únicos suficientemente variables para estimas de diversidad intrapoblacional (Fig. 6.2) y diferenciación genética, e indican que la variación genética de Córcega y Cerdeña es menor y está menos estructurada que en Portugal, lo que sugiere un sentido de dispersión desde Portugal hacia Córcega-Cerdeña.

Por otra parte, el contexto de variación a nivel de género proporcionado por la región ITS nos informa de que *A. pungens* tiene en todo su rango de distribución un ribotipo típico de las armerias del W de la Península Ibérica, bien diferenciado del que presentan tanto sus congéneres en Córcega y Cerdeña (*A. leucocephala, A. morisii, A. multiceps, A. sardoa, A. soleirolii* y *A. sulcitana*) y como los de las Islas Cíes (*A. pubigera* y *A. maritima*). Esto confirma el sentido de dispersión inferido en base los patrones de diversidad de AFLP y ADNcp.

6.2.2.2 *Golfo de Cádiz*

Todos los marcadores revelan que esta región es genéticamente distinta del resto (Fig. 6.1). Sin embargo, tan sólo los marcadores de secuencia permiten reconstruir las relaciones genealógicas entre las distintas áreas, i.e., polarizar la información estableciendo cuáles son ancestrales y cuáles derivadas. En el caso de *A. pungens*, aplicando la teoría de la coalescencia, las secuencias de cloroplasto indican que las poblaciones del Golfo de Cádiz son derivadas del área portuguesa. Lamentablemente, las demás secuencias, ITS y *GapC*, no son lo suficientemente variables para precisar relaciones genealógicas entre alelos. Finalmente, las estimas de diversidad basadas en AFLP y ADNcp y *GapC* indican que este linaje presenta una diversidad genética muy baja (Fig 6.2).

En resumen, a partir las distancias genéticas descritas, junto a los patrones de diversidad genética y el indumento de las hojas, queda claro que *A. pungens* ha sufrido dos procesos recientes de expansión desde un área ancestral, actualmente localizada en las costas atlánticas portuguesas:

- por dispersión a larga distancia a Córcega y Cerdeña, cuyas poblaciones poseen un subconjunto de la diversidad portuguesa.

- al Golfo de Cádiz, mediante un proceso de colonización lineal acompañado de divergencia y pérdida extrema de diversidad por deriva génica asociada a efecto fundador.

6.2.3 Reticulación en los límites del área de distribución: comparación de marcadores de herencia materna y biparental

La comparación de marcadores nucleares y cloroplásticos permite detectar eventos de reticulación en las poblaciones situadas en los límites del área de distribución.

En el límite sur, Camarinal, la morfología y todos los marcadores estudiados sugieren una clara divergencia de esta población del resto de la especie (Fig. 6.1) y una elevada diversidad genética (Fig. 6.2). Además la comparación con otras especies de *Armeria* en base a ITS y, sobre todo, a la morfología (hojas pubérulas) y la

compartición de haplotipos de ADNcp, indican que se debe a una introgresión reciente de *A. macrophylla*.

En el límite norte, la población de las Islas Cíes presenta patrones contradictorios entre la morfología y los tres marcadores nucleares por un lado, que la relacionan con el linaje corso-sardo-portugués, y las secuencias de cloroplasto por otro, que la aproximan a *A. welwitschii* y *A. berlengensis* (Fig. 6.1). Todo apunta a una introgresión antigua con captura cloroplástica de *A. welwistchii* y *A. berlenguensis*, lo que parece perfectamente posible si el área de alguna de estas especies hubiera alcanzado latitudes más septentrionales en el pasado. Esta incongruencia de la morfología y ADN nuclear frente al ADN cloroplástico, es típica de eventos de captura cloroplástica (Rieseberg y Wendel, 1993; Arnold, 1997; Arnold *et al*., 1999).

Finalmente, también en las Islas Cíes, se halló una zona híbrida en el área de contacto entre *A. pungens* y *A. pubigera.* La morfología y ecología intermedia de estos individuos con respecto a ambas especies se confirma con los datos genéticos, ya que presentan una mezcla de haplotipos cloroplásticos, así como polimorfismos aditivos de ITS con respecto a los ribotipos parentales.

Figura 6.1 Distribución de los linajes genéticos revelados en *A. pungens* en base a tres marcadores de ADN nuclear (AFLP, ITS y *GapC*) y tres regiones del cloroplasto (*trn*L-F, *trn*SfM, *mat*K). **A:** AFLP (seis linajes, ver capítulo 2); **B:** ITS (cinco "ribotipos", ver capítulo 3).

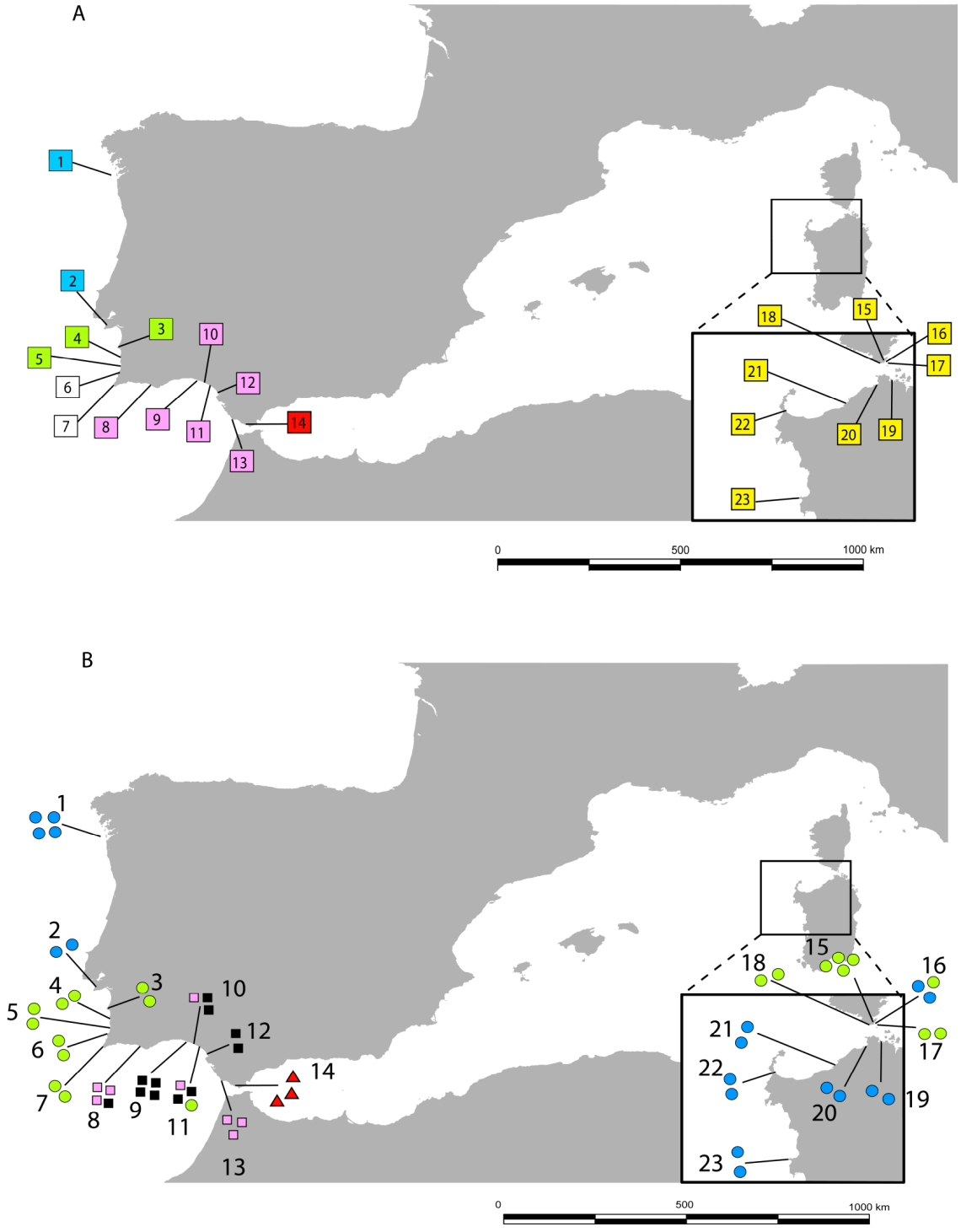


Figura 6.1 (Continuación) Distribución de los linajes genéticos revelados en *A. pungens*. **C:** cpDNA (16 haplotipos, ver capítulo 3); **D:** GapC (13 alelos, ver capítulo 4).

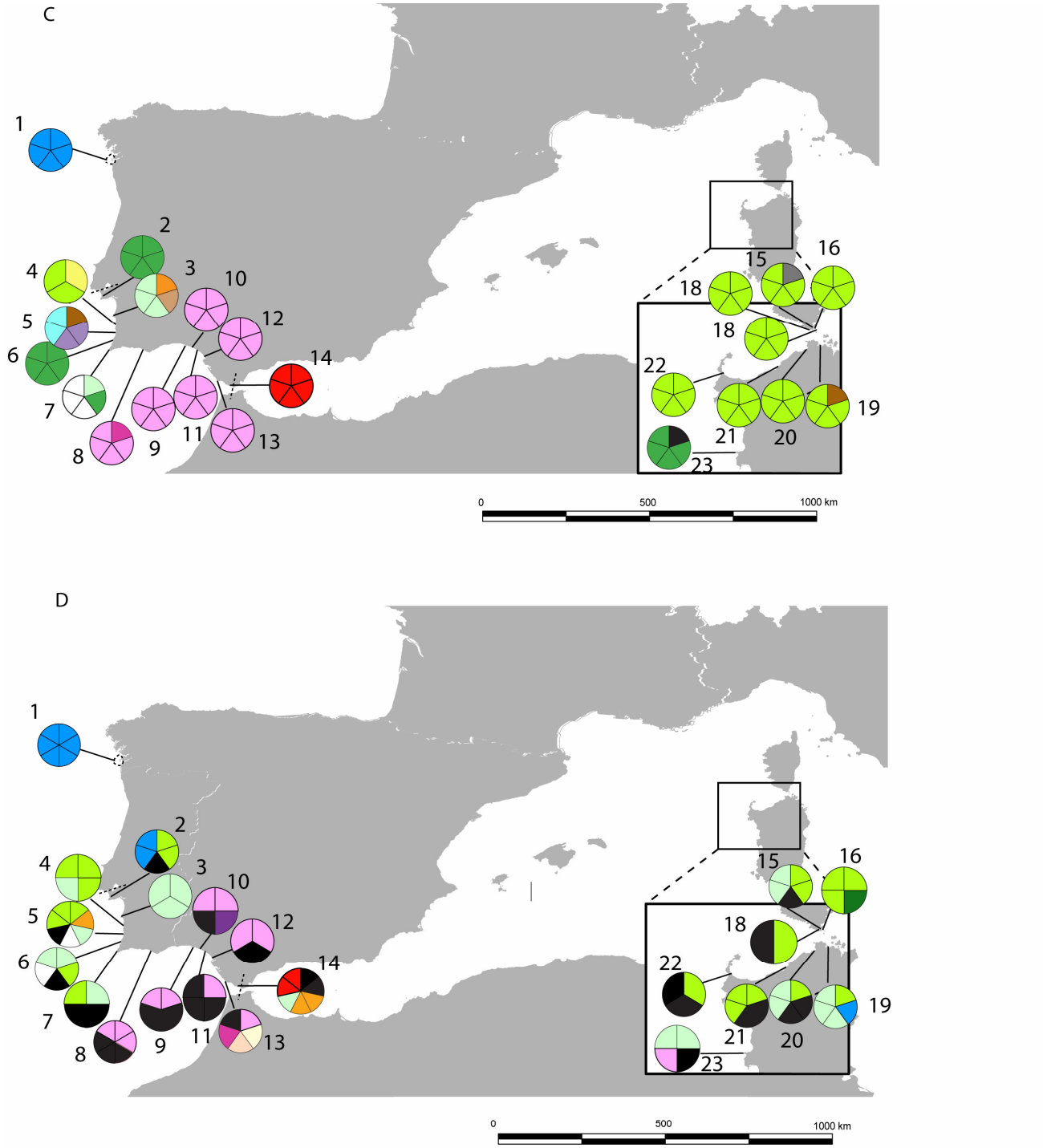
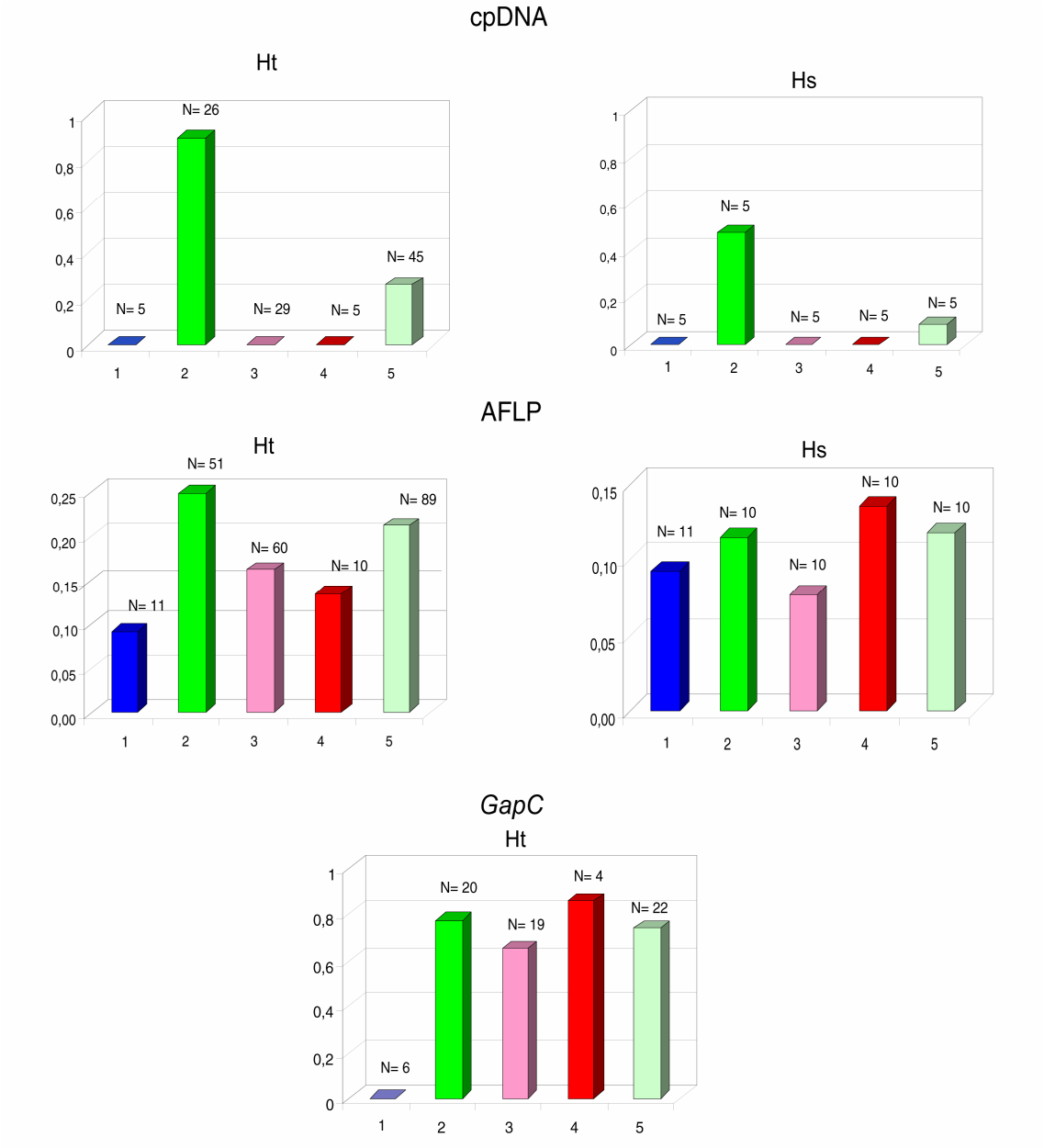


Figura 6.2 Diversidad total (Ht) de los distintos linajes genéticos de *A. pungens* en base a cpDNA, AFLP y GapC, indicando el tamaño muestral total para cada región. Diversidad intrapoblacional media (Hs) en base a cpDNA, y AFLP, indicando la media tamaño muestral poblacional para cada región. **1:** Islas Cíes; **2:** Portugal; **3:** Golfo de Cádiz; **4:** Camarinal; **5:** Córcega-Cerdeña.



6.3 COMPARACIÓN CON LA FILOGEOGRAFÍA DE *Armeria maderensis*

El proyecto (BOS2001–1839) donde se enmarca esta tesis doctoral se planteó como un estudio filogeográfico de dos especies de un mismo género en el que la hibridación interespecífica es frecuente y juega un papel evolutivo, de forma que puede hablarse de que *Armeria* tiene evolución reticulada (Fuertes Aguilar *et al*., 1999; Nieto Feliner *et al*., 2001; Fuertes Aguilar y Nieto Feliner, 2003). Pero las dos especies se escogieron por representar dos casos de estudio con características muy diversas.

Había tres facetas que, a priori, parecían más heterogéneas:

(1) La amplitud de las áreas de distribución era la primera de ellas (3 km de distancia máxima entre poblaciones en *A. maderensis* frente a 1300 km en *A. pungens* que no obstante incluyen una larga disyunción).

(2) La susceptibilidad de recibir flujo génico de otras especies en un género con unas barreras reproductivas internas débiles (Nieto Feliner *et al*., 1996) parecía también muy diferente. Se esperaba escasa en *A. maderensis* debido a su aislamiento (a 850 km del Cabo San Vicente y c. 700 Km de la costa africana) y alta en *A. pungens* que es simpátrica o casi con una decena de congéneres a lo largo de su área de distribución.

(3) Las características biogeográficas de las poblaciones isleñas de ambas especies difieren también por el hecho de ser *A. maderensis* endémica de una isla oceánica mientras que *A. pungens* se distribuye por dos archipiélagos continentales además de la Península Ibérica.

A la postre, las diferencias entre ambas especies han resultado ser mayores de lo previsto a priori y la comparación ha perdido algo de sentido pero hay resultados de un caso de estudio que tienen implicaciones en el otro, resultan complementarios o ayudan a interpretar los patrones del otro.

Del aislamiento de *A. maderensis* y la baja probabilidad de recibir flujo génico de otros congéneres pueden ser buen reflejo las diferencias en la secuencia única de ITS detectada en esta especie con respecto al resto del género, con cuatro mutaciones exclusivas. Sin embargo, esta diferenciación no se detecta en el gen nuclear *GapC* (datos no incluidos en esta memoria doctoral). Seguramente, *GapC* ha sufrido una diferenciación alélica que antecede claramente a la diferenciación de *A. maderensis* (y, por supuesto de *A. pungens*) y simplemente no contiene señal filogenética útil para

arrojar luz en las relaciones filogenéticas de esta especie con el resto del género. En cambio, la alta tasa de mutación en los ITS ha permitido la acumulación de marcadores en esta región que atestiguan el aislamiento de *A. maderensis*, aislamiento que –con todo– representa una historia evolutiva más reciente que aquella en la que se diferenciaron las especies del género.

Como contraste, en *A. pungens* nuestro estudio ha revelado una situación muy diferente en cuanto a la susceptibilidad de recibir flujo génico de otros congéneres ya que documentamos –sobre todo con base en las secuencias cloroplásticas— tres casos de hibridación en distintas épocas. Tal vez lo más interesante de esto sea que los casos documentados se sitúan en los extremos norte y sur del área de distribución ibérica de la especie (dos en Islas Cies y uno en Punta Camarinal). Por eso, es inevitable plantearse la hipótesis de que pueda ser ésta una estrategia de colonización de nichos ligeramente distintos en los márgenes del área de distribución. Esta hipótesis encajaría desde luego en un género con frecuentes casos de reticulación muchos de los cuales son interpretables adaptativamente (e.g., Nieto Feliner *et al*., 2002).

Hay dos piezas de información que resultan especialmente ilustrativas y complementarias entre ambos sistemas. La documentación de migración a larga distancia en *A. pungens* (de Portugal a Córcega/Cerdeña) de c. 1500 Km da credibilidad a la colonización de una la isla oceánica de Madeira que, como se ha dicho, está a c. 850 km de la Península Ibérica. En cambio, el tiempo de ambas colonizaciones parece ser muy distinto. Mientras que los datos moleculares y morfológicos sugieren poderosamente una colonización reciente del archipiélago corso-sardo por parte de *A. pungens,* los datos de ITS y los morfológicos indican que la colonización de Madeira por parte de *A. maderensis* ha debido ser muy anterior. En efecto, a la distancia mínima de cuatro mutaciones en los ITS frente a cualquier congénere hay que sumar que este endemismo maderense posee varios rasgos morfológicos exclusivos y que el conjunto de la morfología no permite identificar inequívocamente al pariente más próximo.

En lo que se refiere a la magnitud y distribución de la variabilidad genética, la utilización de las mismas combinaciones de cebadores y protocolos de AFLP para las dos especies nos ha permitido una estima comparativa muy fiable de cuáles son los niveles de diversidad. Sin embargo, los factores históricos que han influido en estos

niveles son muy dispares en ambas especies y por ello un sistema no ilustra mucho sobre las causas en el otro. En *A. maderensis*, la baja diversidad genética y ausencia de estructuración, debida principalmente a pastoreo de cabras unido a un nicho reducido, ha borrado señal filogeográfica, incluida la que podía ilustrar sobre la colonización de la isla. En *A. pungens*, en cambio, los AFLP proporcionan una información más que satisfactoria sobre la historia evolutiva de esta especie que indica migraciones a larga distancia, efectos fundadores, patrones de diversidad genética diferenciales en los distintos núcleos geográficos, etc (ver apartado 6.2.2).

6.4 REFERENCIAS

Alsos IG, Engelskjøn T, Gielly L, Taberlet P, Brochmann C (2005) Impact of ice ages on circumpolar molecular diversity: insights from an ecological key species. *Molecular Ecology* **14**, 2739-2753.

Álvarez I, Costa A, Nieto Feliner G (2007) Selecting single-copy nuclear genes for plant phylogenetics: a preliminary analysis for the Senecionae (Asteraceae). **No publicado**.

Alvarez I, Wendel J (2003) Ribosomal ITS sequences and plant phylogenetic inference. *Molecular Phylogenetics and Evolution* **29**, 417-734.

Arnheim, N., 1983. Concerted evolution of multigene families. In: *Evolution of Genes and Proteins.*, pp. 38–61. (eds. Nei M and Koehn R eds.) Sinauer, Sunderland, MA.

Arnold ML (1997) *Natural Hybridization and Evolution.* Oxford Univ. Press, Oxford.

Arnold ML, Bulger MR, Burke JM, Hempel AL, Williams JH (1999) Natural hybridization: how low can you go and still be important? *Ecology* **80**, 371-381.

Avise JC, Reeb CA, Saunders NC (1987) Geographic population structure and species differences in mitochondrial DNA of Mouthbrooding Marine Catfishes (Ariidae) and Demersal Spawning Toadfishes (Batrachoididae). *Evolution* **41**, 991-1002.

Baldwin B, Sanderson M, Porter JM*, et al.* (1995) The ITS region of nuclear Ribosomal DNA: a valuable source of evidence on angiosperm phylogeny. *Annals of the Missouri Botanical Garden* **82**, 247-277.

Bena G, Jubier MF, Olivieri I, Lejeune B (1998) Ribosomal external and internal transcribed spacers: combined use in the phylogenetic analysis of *Medicago* (Leguminosae). *Journal of Molecular Evolution* **46**, 299-306.

Birky CW (1995) Uniparental inheritance of mithocondrial and chloroplast genes - mechanisms and evolution. *Proceedings of the National Academy of Sciences of the USA* **92**, 11331-11338.

Castelloe J, Templeton A (1994) Root probabilities for intraspecific gene trees under neutral coalescent theory. *Molecular Phylogenetics and Evolution* **3**, 102-113.

Chung S, Staub JE (2003) The development and evaluation of consensus chloroplast primer pairs that possess highly variable sequence regions in a diverse array of plant taxa. *Theoretical and Applied Genetics* **107**, 757-767.

Clegg M, Gaut B, Learn G, Jr, Morton B (1994) Rates and patterns of chloroplast DNA evolution. *Proceedings of the National Academy of Sciences of the USA* **91**, 6795-6801.

Demesure B, Sodzi N, Petit R (1995) A set of universal primers for amplification of polymorphic non-coding regions of mitochondrial and chloroplast DNA in plants. *Molecular Ecology* **4**, 129-131.

Doyle J, Morgante M, Tingey S V, Powell W (1998) Size homoplasy in chloroplast microsatellites of wild perennial relatives of soybean (*Glycine* subgenus *Glycine)*. *Molecular Biology and Evolution* **15**, 215-218.

Dover G (1982) Molecular drive: a cohesive mode of species evolution. *Nature* **299**, 111-117.

Elder JF, Turner BJ (1995) Concerted evolution of repetitive DNA sequences in eukaryotes. *Quarterly Review of Biology* **70**, 297–320.

Ennos RA (1994) Estimating the relative rates of pollen and seed migration among plant populations. *Heredity* **72**, 250-259.

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

Feliciello I, Picariello O, Chinaly G (2005) The first characterisation of the overall variability of repetitive units in a species reveals unexpected features of satellite DNA. *Gene* **349**, 153-164.

Fuertes Aguilar J, Nieto Feliner G (2003) Additive polymorphisms and reticulation in an ITS phylogeny of thrifts (*Armeria*, Plumbaginaceae). *Molecular Phylogenetics and Evolution* **28**, 430-447.

Fuertes Aguilar J, Rossello JA, Nieto Feliner G (1999) Molecular evidence for the compilospecies model of reticulate evolution in *Armeria* (Plumbaginaceae). *Systematic Biology* **48**, 735-754.

Grivet D, Heinze B, Vendramin GG, Petit RJ (2001) Genome walking with consensus primers: application to the large single copy region of chloroplast DNA. *Molecular Ecology Notes* **1**, 345-349.

Gutiérrez Larena B, Fuertes Aguilar J, Nieto Feliner G (2002) Glacial-induced altitudinal migrations in *Armeria* (Plumbaginaceae) inferred from patterns of chloroplast DNA haplotype sharing. *Molecular Ecology* **11**, 1965-1974.

Hare M (2001) Prospects for nuclear gene phylogeography. *Trends in Ecology & Evolution* **16**, 700-706.

Hausner G, Olson R, Simon D*, et al.* (2006) Origin and evolution of the chloroplast *trn*K (*mat*K) intron: a model for evolution of group II intron RNA structures. *Molecular Biology and Evolution* **23**, 380-391.

Havey MJ, McCreight JD, Rhodes B, Taurick G (1998) Differential transmission of the *Cucumis* organellar genomes. *Theoretical and Applied Genetics* **97**, 122-128.

Henning W (1968) *Elementos de una Sistemática Filogenética* Editorial Universitaria de Buenos Aires.

Hilu K, Liang H (1997) The *mat*K gene: sequence variation and application in plant systematics. *American Journal of Botany* **84**, 830-839.

Hudson RR (1990) *Gene Genealogies and the Coalescent Process*. Oxford University Press, Oxford.

Hughes CE, Eastwood RJ, Donovan Bailey C (2006) Review. From famine to feast? Selecting nuclear DNA sequence loci for plant species-level phylogeny reconstruction. *Philosophical Transactions of the Royal Society B: Biological Sciences* **361**, 211-225.

Ishii T, Mori N, Ogihara Y (2001) Evaluation of allelic diversity at chloroplast microsatellite loci among common wheat and its ancestral species. *Theoretical and Applied Genetics* **103**, 896-904.

Johnson L, Soltis DE (1994) *mat*K DNA sequences and phylogenetic reconstruction in saxifragaceae s. str. *Systematic Botany* **19**, 143-156.

Kelchner SA (2000) The evolution of non-coding chloroplast DNA and its application in plant systematics. *Annals of the Missouri Botanical Garden* **87**, 482-498.

Koch MA, Kiefer C, Ehrich D*, et al.* (2006) Three times out of Asia Minor: the phylogeography of *Arabis alpina* L. (Brassicaceae). *Molecular Ecology* **15**, 825-839.

Lowe A, Harris S, Ashton P (2004) *Ecological Genetics: Design, Analysis, and Application.* Blackwell Science, Victoria, Australia.

Maruyama T, Fuerst PA (1984) Population bottlenecks and nonequilibrium models in population genetics. I. allele numbers when populations evolve from zero variability. *Genetics* **108**, 745-763.

McCauley D (1995) The use of chloroplast DNA polymorphism in studies of gene flow in plants. *Trends in Ecology & Evolution* **10**, 198-202.

Nieto Feliner G, Izuzquiza A, Lansac AR (1996) Natural and experimental hybridization in *Armeria* (Plumbaginaceae): *A. villosa* subsp. *carratracensis*. *Plant Systematics and Evolution* **201**, 163–177.

Nieto Feliner G, Fuertes Aguilar J, Rosselló JA (2001) Can extensive reticulation and concerted evolution result in a cladistically structured molecular data set? *Cladistics* **17**, 301–312.

Nieto Feliner G, Fuertes Aguilar J, Rossello JA (2002) Reticulation or divergence: the origin of a rare serpentine endemic assessed with chloroplast, nuclear and RAPD markers. *Plant Systematics and Evolution* **231**, 19-38.

Nybom H (2004) Comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. *Molecular Ecology* **13**, 1143-1155.

Olmstead R, Palmer J (1994) Chloroplast DNA systematics: a review of methods and data analysis. *American Journal of Botany* **81**, 1205-1224.

Petit RJ, Aguinagalde I, de Beaulieu J-L*, et al.* (2003) Glacial refugia: hotspots but not melting pots of genetic diversity. *Science* **300**, 1563-1565.

Petit RJ, DuminilL J, Fineschi S*, et al.* (2005) Comparative organization of chloroplast, mitochondrial and nuclear diversity in plant populations. *Molecular Ecology* **14**, 689-701.

Petit RJ, Kremer A, Wagner DB (1993) Geographic structure of chloroplast DNA polymorphisms in European oaks. *Theoretical and Applied Genetics* **87**, 122-128.

Posada D, Crandall K (2001) Intraspecific gene genealogies: trees grafting into networks. *Trends in Ecology & Evolution* **16**, 37-45.

Powell W, Morgante M, Andre C*, et al.* (1995) Hypervariable microsatellites provide a general source of polymorphic DNA markers for the chloroplast genome. *Current Biology* **5**, 1023-1029.

Provan J, Powell W, Hollingsworth PM (2001) Chloroplast microsatellites: new tools for studies in plant ecology and evolution. *Trends in Ecology & Evolution* **16**, 142-147.

Quandt D, Stech M (2004) Molecular evolution and phylogenetic utility of the chloroplast *trn*T-*trn*F region in bryophytes. *Plant Biology* **6**, 545–554.

Rieseberg LH (1997) Hybrid origins of plant species. *Annual Review of Ecology & Systematics* **28**, 359-389.

Rieseberg LH, Wendel JF (1993) Introgression and its consequences in plants. In: *Hybrid Zones and the Evolutionary Process* (ed. Harrison R), pp. 70-109. Oxford University Press, New York.

Sang T (2002) Utility of low-copy nuclear gene sequences in plant phylogenetics. *Critical Reviews in Biochemistry & Molecular Biology* **37**, 121-147.

Schaal BA, Hayworth DA, Olsen KM, Rauscher JT, Smith WA (1998) Phylogeographic studies in plants: problems and prospects. *Molecular Ecology* **7**, 465-474.

Schaal BA, Olsen KM (2000) Gene genealogies and population variation in plants. *Proceedings of the National Academy of Sciences USA* **97**, 7024-7029.

Schlüter PM, Stuessy TF, Paulus HF (2004) Making the first step: practical considerations for the isolation of low-copy nuclear sequence markers. *Taxon* **54**, 766-770.

Shannon CE (1948) A mathematical theory of communication. *Bell Systems Technical Journal* **27**, 379-423, 623-656.

Simmons MP, Ochoterena H (2000) Gaps as characters in sequence-based phylogenetic analyses. *Systematic Biology* **49**, 369-381.

Slatkin M (1995) A measure of population subdivision based on microsatellites allele frequencies. *Genetics* **139**, 457-462.

Small R, Cronn RC, Wendel JF (2004) Use of nuclear genes for phylogeny reconstruction in plants. *Australian Systematic Botany* **17**, 145-170.

Soltis D, Morris A, McLachlan J, Manos P, Soltis P (2006) Comparative phylogeography of unglaciated eastern North America. *Molecular Ecology* **15**, 4261-4293.

Sugita M, Shinozaki K, Sugiura M (1985) Tobacco Chloroplast tRNALys(UUU) gene contains a 2.5-Kilobasepair intron: an open reading frame and a conserved boundary sequence in the intron. *Proceedings of the National Academy of Sciences of the USA* **82**, 3557-3561.

Taberlet P, Gielly L, Pautou G, Bouvet J (1991) Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Molecular Biology* **17**, 1105-1109.

Trick M, Dover GA (1984) Unexpectedly slow homogenisation within a repetitive DNA family shared between two subspecies of tsetse fly. *Journal of Molecular Evolution* **20**, 322-329.

van Nues RW, Rientjes JM, van der Sande CA*, et al.* (1994) Separate structural elements within Internal Transcribed Spacer 1 of *Saccharomyces cerevisiae* ribosomal DNA. *Journal of Molecular Biology* **223**, 899–910.

Vendramin GG, Lelli LL, Rossi PP, Morgante MM (1996) A set of primers for the amplification of 20 chloroplast microsatellites in Pinaceae. *Molecular Ecology* **5**, 595-598.

Vriesendorp B, Bakker FT (2005) Reconstructing patterns of reticulate evolution in angiosperms: what can we do? *Taxon* **54**, 493-604.

Weising K, Gardner RC (1999) A set of conserved PCR primers for the analysis of simple sequence repeat polymorphisms in chloroplast genomes of dicotyledonous angiosperms. *Genome* **42**, 9-19.

Wendel JF, Doyle JJ (1998) Phylogenetic incongruence: window into genomes history and molecular evolution. In: *Molecular Systematics of Plants II. DNA Sequencing* (eds. D S, Soltis P, Doyle JJ). Kluwer Academy Publications, Boston.

Zhang DX, Hewitt GM (2003) Nuclear DNA analyses in genetic studies of populations: practice, problems and prospects. *Molecular Ecology* **12**, 563.

Conclusiones

Los diferentes tópicos que se abordaron y analizaron sobre los LVM y LTM permiten establecer un conjunto de ideas que resultan adecuadas para el tratamiento de datos ordinales en estos modelos. En primer lugar, se ha demostrado que no resulta indiferente la adopción de una u otra aproximación teórica (UV o IRT) para ajustar modelos con datos ordinales. Los resultados de las aplicaciones realizadas muestran un mejor comportamiento en la estimación y evaluación del ajuste de los LTM ajustados desde la IRT.

Por otra parte, se aclaró que el método de estimación empleado afecta significativamente la evaluación del ajuste, debiendo distinguirse entre métodos de información completa y limitada. Se determinó, para la escala empleada, que ajustando las funciones normal (NOR) y logística (POM) desde la IRT se encuentran mejores resultados en la evaluación del ajuste que empleando el GLS y WLS desde la aproximación UV. Bajo condiciones simuladas, se resalta la estabilidad de los métodos ajustados desde la IRT e inconsistencias bajo la aproximación UV. Tomando en cuenta que estos resultados se derivaron a partir de la evaluación del ajuste, son similares a los obtenidos por otros autores en

trabajos en los que también se consideran la estabilidad de los parámetros, (Moustaki et al., 2004; Maydeu-Olivares., 2005).

Al existir diversas alternativas, teóricas y metodológicas, para el tratamiento de los datos ordinales, los investigadores deberían evaluar sus datos y verificar sus distribuciones. La presencia de no normalidad puede ser abordada aproximando los datos con funciones logísticas desde la IRT, pues como se demostró en esta investigación, estos modelos ofrecen mejores resultados en comparación con los datos con métodos lineales.

Entre el conjunto de factores que afectan la bondad del ajuste, se ha señalado que el procedimiento de simulación empleado en los procesos de validación de los índices, tiene un peso significativo en la valoración de los mismos. Al evaluar el efecto conjunto del procedimiento de simulación y los métodos de estimación en LTM tratados desde la IRT, se demostró que el empleo de técnicas de remuestreo, específicamente la implementación no paramétrica de Bootstrap (tal y como está disponible en el software comercial), no resulta adecuada para la evaluación de los índices de ajuste en estos modelos. Lo contrario sucede con las implementaciones Monte Carlo, basadas en la generación de datos ordinales a partir de las matrices policóricas y de covarianzas asintóticas y de los valores de los umbrales para cada ítem, a través del cual se obtienen soluciones convergentes y óptimas que permiten evaluar la precisión de los índices de ajuste. Aunque estos resultados hacen referencia directamente a los efectos sobre la bondad del ajuste, están en línea con los obtenidos por Hartmann (2005); Langeheine et al., (1996) y Young (1994), sobre la estimación de parámetros.

En cuanto al método de estimación de información completa (adoptando las funciones NOR y POM), los pesos de los factores en el modelo evaluado no resultaron con diferencias significativas, lo que implica un comportamiento similar en los métodos de estimación aproximados desde la IRT. Estos resultados son congruentes con los obtenidos en el capítulo II en el que se compararon

métodos de información limitada y de información completa, y en donde a favor de los últimos, se concluía acerca de los óptimos ajustes obtenidos. Por otra parte se sugiere que, en el tratamiento de datos ordinales, se considere la aplicación de métodos de estimación que aproximen las funciones logística y normal desde la IRT, tal y como lo plantean otros autores (Jöreskog & Moustaki, 2001; 2006).

Existen muchas limitaciones en cuanto a la disponibilidad de herramientas de software que permitan obtener índices de ajuste adaptados al comportamiento de datos ordinales en LTM, tratados desde la IRT. Por lo que se opta por la creación de nuevos códigos de programas y la utilización simultánea de herramientas. El procedimiento alternativo propuesto para la generación de muestras por simulación, empleando tanto Bootstrap como Monte Carlo, permite obtener los índices de ajuste para un conjunto de muestras y evaluar la distribución de los mismos, utilizando herramientas de software alternativas al LISREL/PRELIS (Jörekog & Sörbom, 2006), que permiten superar las limitaciones en este campo. El procedimiento se caracteriza por su funcionabilidad y portabilidad, permitiendo, a través de la implementación de un conjunto detallado de instrucciones, obtener resultados que permitan evaluar el ajuste.

El tema de los índices de bondad del ajuste en LVM representa un aporte especial de este trabajo, pues a partir de la polémica planteada sobre el tema, que versa principalmente sobre lo inadecuadas que resultan las medidas generales de evaluación del ajuste χ^2o G^2 para evaluar escalas que presentan un número elevado de ítems categóricos y que cuestiona que estas medidas no dan información acerca de ítems anómalos, se analiza su utilidad y aplicabilidad para la evaluación de modelos específicos y posteriormente se proponen nuevas medidas que superan estas limitaciones.

Considerando estas limitaciones, en el presente trabajo se examinaron las distintas medidas alternativas de ajuste que permitían valorar los modelos, indagando acerca de sus elementos de cálculo y condiciones de aplicación. Este

análisis permitió la formulación de nuevas medidas de ajuste (φ_{LR} y φ_{GF}) cuya aplicación proporciona unas serie de beneficios: (1) Obtener un valor único que permite evaluar con precisión el ajuste global del modelo; (2) Los valores obtenidos son interpretables a partir de una regla práctica que permite establecer un límite entre buenos y pobres ajustes. Esta información se traduce en una herramienta valiosa para el investigador pues puede valorar la viabilidad de la solución obtenida; (3) Existe un rango de interpretación que sugiere la presencia de ítems anómalos. Esto permite que al investigar, se pueda evaluar el comportamiento de un ítem (o par de ellos) verificando la presencia de casos Heywood y, considerar sí con la eliminación del mismo se encontraría un mejor ajuste; (4) Al no ser necesaria la inspección individual del ajuste por ítem (pares o tríos), los nuevos coeficientes resultan apropiados cuando se necesita evaluar un número elevado de soluciones, cómo las obtenidas en un estudio de simulación o en la evaluación de modelos alternativos Por lo tanto, el análisis de sus distribuciones facilita las conclusiones acerca de la evaluación del ajuste del modelo; (5) Su cálculo es simple, lo que facilita su implementación mediante software.

En el presente trabajo se hace uso del programa PRELIS para obtener los valores de S_{LR}, S_{GF}, LR_{biv} y $GF_{biv,}$ que son necesarios para su cálculo, pero resulta viable el desarrollo de un programa independiente o la implementación de pequeñas rutinas en software comercial.

Los nuevos estadísticos aprovechan la información suministrada por otros índices como: χ^2_{LR}, χ^2_{GF}, S_{LR}, S_{GF}, LR_{biv} y GF_{biv} y la emplean como elementos de cálculo, esto permite establecer una coherencia teórica y metodológica que se sustenta en la polémica sobre el tema, estableciendo una conexión con las reglas de interpretación de estos estadísticos, propuestas por diversos autores (Drasgow et al., 1995; Chernyshenko et al., 2001; Bartholomew et al., 2002; Bartholomew & Knott, 1999; Jöreskog & Moustaki, 2001), lo que supone una ampliación y mejora en la pirámide de ajuste de LVM.

Aunque estas medidas han sido planteadas para el tratamiento de datos ordinales desde la IRT utilizando FIML como método de estimación, los coeficientes φ_{LR} y φ_{GF} son independientes del método de estimación y pueden adoptarse con otros métodos de estimación como el MML. Esto supone ampliar su área de aplicación y la consideración de los mismos como medidas de ajuste de LVM en general.

Esta tesis plantea un conjunto de líneas de investigación futuras a seguir caracterizadas principalmente por:

- El desarrollo de procedimientos de software que permitan obtener índices específicos para evaluar LVM desde la IRT, que consideren el fenómeno del sparseness, tanto en modelos unidimensionales como multidimensionales, que contengan implementadas rutinas de simulación Monte Carlo para que permitan la validación posterior de estos índices

- La verificación del comportamiento de los coeficientes de ajuste φ_{LR} y φ_{GF} con ítems binarios.

- La comparación entre la implementación de Bootstrap paramétrico y la simulación Monte Carlo, en procedimientos de validación de los índices de ajuste.

- La utilización de los coeficientes de ajuste φ_{LR} y φ_{GF} para el ajuste de modelos politómicos de la TRI como el de respuesta graduada de Samejima y el de crédito parcial

- La comparación de los coeficientes de ajuste φ_{LR} y φ_{GF}, empleando FIML, MML u otros métodos de estimación.

En este sentido, se resalta que el tratamiento de datos ordinales en LVM está en pleno desarrollo y expansión y que cada día nos encontramos con nuevas estrategias que permiten aproximarnos con mayor precisión y rigor a las

realidades que pretendemos medir con nuestros datos y que por lo tanto el conocimiento específico de los mismos por parte de los investigadores puede resultar clave en la consolidación de los conocimientos derivados de la investigación.