Phylogeography of Two Lusitanian Sea Stars

David John Darrock

A thesis submitted to Cardiff University for the higher degree of Doctor of Philosophy

September 2010

Cardiff School of Biosciences
Cardiff University
DECLARATION

This work has not previously been accepted in substance for any degree and is not concurrently submitted in candidature for any degree.

Signed ................................. (candidate) Date 06/04/11

STATEMENT 1

This thesis is being submitted in partial fulfillment of the requirements for the degree of PhD.

Signed ................................. (candidate) Date 06/04/11

STATEMENT 2

This thesis is the result of my own independent work/investigation, except where otherwise stated.
Other sources are acknowledged by explicit references.

Signed ................................. (candidate) Date 06/04/11

STATEMENT 3

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed ................................. (candidate) Date 06/04/11

STATEMENT 4: PREVIOUSLY APPROVED BAR ON ACCESS

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loans after expiry of a bar on access previously approved by the Graduate Development Committee.

Signed ................................. (candidate) Date 06/04/11
Dedicated to Lorna and Rhys
SUMMARY

The first comprehensive genetic study of North East Atlantic and Mediterranean sea stars, *Asterina gibbosa* and *Asterina phylactica*, is presented here, based on mitochondrial DNA (mtDNA) and Amplified Fragment Length Polymorphism (AFLP) data. MtDNA analysis revealed that the two putative species are distinct however there is incomplete lineage sorting, with the two most common haplotypes being shared across both species. MtDNA revealed low divergence between populations especially among most Atlantic populations, with no significant differentiation between the two basins. Although, both species possessed private haplotypes within both basins, the most common haplotype within both species is found throughout the entire geographical range of both species. Two mitochondrial haplogroups were identified, both of which showed evidence for a population expansion, occurring during the Pleistocene epoch. Haplogroup 1 was dominated by *A. gibbosa* (88%) whereas haplogroup 2 was dominated by *A. phylactica* (84%). The mtDNA results tentatively suggest that one *Asterina* population may descend from a population that survived the last glacial maximum (LGM) in one or more northern refugium. The AFLP data showed that *A. gibbosa* and *A. phylactica* are genetically distinct, with no apparent hybridization between species, with the possible exception of a single individual found at Rovinj, Croatia which was identified as being an *A. phylactica* individual at the time of sampling, but the allocation test assigned it to the Naples, Italy, *A. gibbosa* population. This could be the result of introgression or the individual could have been incorrectly classified as *A. phylactica* at the time of sampling. The AFLP data showed that there is gene flow occurring but it appears to be restricted, particularly within the Mediterranean basin, with no apparent gene flow occurring between the Atlantic and Mediterranean basins.

There was no evidence with either marker to conclude that the brooding behaviour of *A. phylactica* provides a different pattern of genetic diversity within populations or differentiation between populations to the crawl away behaviour of *A. gibbosa*. The analysis suggests that gene flow is slightly more restricted for *A. phylactica* than *A. gibbosa* and, for *A. gibbosa*, it is more restricted in the Mediterranean than the Atlantic. The study identifies both *A. phylactica* and *A. gibbosa* populations that would be suitable to receive conservation status, based upon their unique genetic characteristics.
A phylogeographic study of the sea star Asterina gibbosa using Mitochondrial DNA

2.1 Abstract
2.2 Introduction
2.3 Methods
  2.3.1 Sampling
  2.3.2 DNA extraction
  2.3.3 mtDNA amplification and sequencing
  2.3.4 Authentic sequence verification
  2.3.5 Analysis of Genetic diversity and differentiation
  2.3.6 Analysis of Population Demography
2.4 Results
  2.4.1 Sequences and Haplotype Analysis
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.4.2 Genetic Diversity and Population Structure</td>
<td>75</td>
</tr>
<tr>
<td>2.4.3 Haplogroup Analysis</td>
<td>78</td>
</tr>
<tr>
<td>2.4.4 Population Demography</td>
<td>80</td>
</tr>
<tr>
<td>2.5 Discussion</td>
<td>81</td>
</tr>
<tr>
<td>2.6 References</td>
<td>85</td>
</tr>
<tr>
<td>Population genetics of an enigmatic sea star <em>Asterina phylactica</em></td>
<td>96</td>
</tr>
<tr>
<td>3.1 Abstract</td>
<td>97</td>
</tr>
<tr>
<td>3.2 Introduction</td>
<td>97</td>
</tr>
<tr>
<td>3.3 Materials and methods</td>
<td>100</td>
</tr>
<tr>
<td>3.3.1 Sampling</td>
<td>100</td>
</tr>
<tr>
<td>3.3.2 mtDNA Laboratory Procedures</td>
<td>102</td>
</tr>
<tr>
<td>3.3.3 AFLP Laboratory Procedures</td>
<td>103</td>
</tr>
<tr>
<td>3.3.4 Data analysis</td>
<td>105</td>
</tr>
<tr>
<td>3.3.4.1 mtDNA</td>
<td>105</td>
</tr>
<tr>
<td>3.3.4.2 AFLP</td>
<td>106</td>
</tr>
<tr>
<td>3.4 Results</td>
<td>108</td>
</tr>
<tr>
<td>3.4.1 Mitochondrial DNA</td>
<td>108</td>
</tr>
<tr>
<td>3.4.2 AFLP</td>
<td>119</td>
</tr>
<tr>
<td>3.5 Discussion</td>
<td>123</td>
</tr>
<tr>
<td>3.6 References</td>
<td>128</td>
</tr>
<tr>
<td>A comparative genetic analysis of the cryptic sea stars <em>Asterina gibbosa</em> &amp; <em>Asterina phylactica</em> using mitochondrial and AFLP markers</td>
<td>137</td>
</tr>
<tr>
<td>4.1 Abstract</td>
<td>138</td>
</tr>
<tr>
<td>4.2 Introduction</td>
<td>139</td>
</tr>
<tr>
<td>4.3 Methods</td>
<td>143</td>
</tr>
<tr>
<td>4.3.1 Sampling</td>
<td>143</td>
</tr>
<tr>
<td>4.3.2 Laboratory Procedures</td>
<td>145</td>
</tr>
<tr>
<td>4.3.3 Data Analysis</td>
<td>145</td>
</tr>
<tr>
<td>4.3.3.1 mtDNA</td>
<td>146</td>
</tr>
<tr>
<td>4.3.3.2 AFLP</td>
<td>146</td>
</tr>
<tr>
<td>4.4 Results</td>
<td>148</td>
</tr>
<tr>
<td>4.4.1 Mitochondrial DNA</td>
<td>148</td>
</tr>
<tr>
<td>4.4.2 AFLP Data</td>
<td>162</td>
</tr>
<tr>
<td>4.5 Discussion</td>
<td>169</td>
</tr>
</tbody>
</table>
ACKNOWLEDGEMENTS

I am most grateful for all of the support, guidance, comments and patience that has been provided by Professor Michael W. Bruford. I am grateful to Erika Baus for all of her help and support, particularly with the AFLP techniques and sampling. I am grateful to Dr. Roland Emson for his help, guidance and for providing the inspiration to perform this work. I am grateful to Joanna Cable for her guidance and support. I would like to thank Carlos Fernandes, Nicola Anthony and Helen Wilcock for their training and assistance in the lab. I would like to thank all of my former lab mates, most notably Rhys, Katherine, Jiang Xiang, Simon, Tim, Farius, Trini, Yosha, Mireille, Ciara and Andy. I would like to thank Dr Maria Ina Arnone who kindly provided Asterina gibbosa specimens from Naples, Italy. Finally, I would like to say thank you to my parents for their assistance in collecting specimens from around the coast of the UK and to my wife, Lorna for her help and support throughout the project.
1.1. General Introduction

This thesis concerns the phylogeography of two congeneric species of Asterinid cushion star – *Asterina gibbosa* (Pennant, 1897) and *Asterina phylactica* (Emson and Crump, 1979), using genetic data derived from mitochondrial DNA sequences and Amplified Fragment Length Polymorphisms. Both species lay eggs that hatch into benthic brachiolariae larvae, with no method for long-range dispersal other than rafting as juveniles or adults. Both species inhabit the intertidal zone along the North East Atlantic Ocean and Mediterranean Sea, an area with a turbulent geological and oceanographic history, coupled with strong currents and modern day barriers to dispersal.

In general, species of marine invertebrates with no, or limited, pelagic larval dispersal are expected to consist of highly structured populations with little or no gene flow (Benzie, 2000). By comparison, marine invertebrates with high potential for dispersal should exhibit little phylogeographic structure (Benzie, 2000) as a high level of gene flow would result in limited genetic differentiation between populations (Avise et al., 1987; Bohonak, 1999). There are many examples which show a conflict between actual and expected patterns of gene flow (Solé-Cava et al. 1994; Grant and da Silva-Tatley 1997; Uthicke and Benzie 2000; Lazoski et al. 2001; Thornhill et al., 2008; Hunter, 2010). Some brooding marine invertebrates have been shown to have large ranges with little genetic differentiation over large distances, for example, the brooding sea star *Astrotoma agassizii* shows high levels of genetic continuity across geographical distances of greater than 500 km (Hunter and Halanych, 2008).

1.2. The Family Asterinidae

There are 21 genera and 116 species worldwide within the Family Asterinidae (Gray, 1840) (O’Loughlin and Waters, 2004). Most species within the Family can be found in the Pacific, *A. gibbosa* and *A. phylactica* being the only species found in the North Eastern Atlantic Ocean and Mediterranean Sea.
Traditionally, morphology has been used as a basis for taxonomic classification, however, some morphological characters subject to strong selection are phenotypically plastic (O’Loughlin and Waters, 2004) or phylogenetically informative for some groups but uninformative and unreliable for others (Mah, 2000). Due to the unreliability of traditional taxonomy, researchers have turned to molecular data to reassess systematic relationships (Hart et al., 1997; Dartnall et al., 2003; O’Loughlin and Waters, 2004; Waters et al., 2004). There have been a number of taxonomic revisions of this Family (reviewed in O’Loughlin and Waters, 2004). The most recent phylogeny was constructed using sequences of the mitochondrial DNA (mtDNA) cytochrome oxidase I gene (COI) and the adjacent tRNA^ala, tRNA^leu, tRNA^asn, tRNA^gln and tRNA^pro genes (Waters et al., 2004), in conjunction with morphological data (O’Loughlin and Waters 2004). The resulting maximum parsimony analyses of the sequence data can be seen in Fig. 1.1.

Statistical resampling supported the monophyly of three Australian Meridiastra species, but indicated that both Asterina and Patiriella are paraphyletic assemblages (Waters et al., 2004). Six well supported clades were reported, with a clear link between phylogenetic and biogeographic relationships. Interestingly, four of the five Japanese species were grouped together in Clade IV, much of Australia’s diversity was placed in Clade III and Clade V consisted of southern temperate Patiriella species. A. gibbosa and A. phylactica (Clade VI) were, furthermore, highly distinct from all other asterinid species. As a result Waters et al. (2004) hypothesised that evolutionary radiations within the Asterinidae may be geographically localised phenomena.
Figure 1.1. Phylogenetic relationships of asterinids based on mtDNA sequences (taken from Waters et al., 2004). Statistical support for particular clades is indicated by bootstrap values, and by decay indices (in italics). MtDNA clades I – IV were identified on the basis of high bootstrap support and new generic assignments suggested by Waters et al. (2004) appear on the right.
1.3. The Study Species: *Asterina gibbosa* and *Asterina phylactica*

1.3.1. Morphology

*A. gibbosa* is an abundant sea star ranging in size from 0.75mm (newly metamorphosed animals) to 50mm, with records existing of specimens of up to 70mm in the Mediterranean (Ludwig, 1879). *A. phylactica* is smaller than *A. gibbosa*; although newly metamorphosed animals are the same size, the largest *A. phylactica* found is approximately 15mm (Emson and Crump, 1979).

The colour of *A. gibbosa* varies from a pale brown colour through orange and green-brown to dark olive; mottled individuals are present in some populations. The predominant colour pattern varies geographically, for example, Exmouth, Devon, *A. gibbosa* are predominantly grey/green; those from pools near Start Point, Devon, mottled in colour; and, near Angle, Pembrokeshire, pool populations are dark olive (Crump and Emson, 1983). The pattern also varies with habitat, animals from low shore situations being on the whole paler, invariably pale orange or pale grey (Emson and Crump, 1979). Some small *A. gibbosa* may show central dark pigmentation but without the green background colour.

In contrast to *A. gibbosa*, *A. phylactica* has a very consistent and dark colour pattern. The general background colour of the animal is a dark green with a distinctive chocolate brown substar on the aboral (dorsal) surface (Crump and Emson, 1983). While the intensity of the background pigmentation may vary from population to population, it is always green and the overlying star pattern is consistently a darker green with the spine clumps red and orange (Emson and Crump, 1979). When seen on a pale background, *A. phylactica* appears contrastingly coloured but against a dark or varied background, and particularly against algal holdfasts, the colour pattern is highly cryptic (Emson and Crump, 1979). The colour pattern only begins to develop as individuals approach one year of age and it is only at this stage that distinction between these two species becomes possible (Emson and Crump, 1979).
There are no morphological characteristics that consistently distinguish *A. gibbosa* from *A. phylactica* apart from colour pattern and size (Emson and Crump, 1979). This can be seen in Fig 1.2. When a comparison is made between the morphology of *A. phylactica* and *A. gibbosa* of the same size, two differences are usually apparent. First, there are rarely spines on the ventral side of the mouth plates of *A. phylactica*, whereas there are two in *A. gibbosa* (Emson and Crump, 1979). Secondly, because *A. phylactica* of more than 6mm in diameter are mature adults, the gonopores are readily distinguishable, whereas they are not obvious in *A. gibbosa* within the size range of *A. phylactica*, i.e. less than 15mm (Emson and Crump, 1979). However, *A. phylactica* may be readily distinguished from *A. gibbosa* in the field by the distinctive colour pattern and small size (Emson and Crump, 1979). The differences can be seen in Table 1.1.
<table>
<thead>
<tr>
<th>Feature of comparison</th>
<th>A. phylactica</th>
<th>A. gibbosa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Colour pattern</td>
<td>Consistent dark pattern</td>
<td>Variable</td>
</tr>
<tr>
<td>Size</td>
<td>Up to 15 mm in diameter</td>
<td>Up to 50 mm in diameter</td>
</tr>
<tr>
<td>Reproductive method</td>
<td>Large, directly developing</td>
<td>Large, directly developing</td>
</tr>
<tr>
<td></td>
<td>eggs brooded to mobile juvenile state</td>
<td>eggs</td>
</tr>
<tr>
<td>Age at first egg production</td>
<td>Two years (in Britain)</td>
<td>Four years (in Britain)</td>
</tr>
<tr>
<td>Aggregative behaviour</td>
<td>Lengthy aggregation (three weeks) for brooding</td>
<td>Brief aggregation for egg laying</td>
</tr>
<tr>
<td>Reproductive pattern</td>
<td>Simultaneous hermaphrodite</td>
<td>Protandrous hermaphrodite</td>
</tr>
<tr>
<td>Number of broods</td>
<td>One to three</td>
<td>Three to seven</td>
</tr>
<tr>
<td>Life span</td>
<td>Up to four years</td>
<td>Up to seven years or more</td>
</tr>
<tr>
<td>Shore type in respect to exposure</td>
<td>Very exposed to exposed</td>
<td>Exposed to sheltered</td>
</tr>
<tr>
<td>Vertical distribution on the littoral environment</td>
<td>Restricted to tide pools concentrated near MHWN</td>
<td>On boulder scree mostly around ELWS or in rock pools near MHWN</td>
</tr>
<tr>
<td>Distribution on boulders in sympatric populations</td>
<td>Principally tops and sides</td>
<td>Principally sides and undersides</td>
</tr>
<tr>
<td>Association with Corallina</td>
<td>Adults often very abundant living in Corallina</td>
<td>Adults very rarely found in Corallina</td>
</tr>
<tr>
<td>Reaction to light (sensitivity)</td>
<td>Low sensitivity (greater than 10,000 lux in summer)</td>
<td>High sensitivity (5,000 lux in summer)</td>
</tr>
<tr>
<td>Movement / activity</td>
<td>Diurnally active</td>
<td>Principally nocturnal</td>
</tr>
</tbody>
</table>

Table 1.1. Summary of differences between A. gibbosa and A. phylactica.
1.3.2. Ecology

A. gibbosa is most commonly found closely adpressed to the undersides of large boulders on the low shore at low tide. A. gibbosa can inhabit a wide range of shores from sheltered to semi-exposed, indeed, wherever boulder scree dominates the lower part of the shore, around the extreme low-water spring (ELWS). A. gibbosa is susceptible to desiccation and is rarely found in the open except on low shore gully walls, overhangs and in caves. A. gibbosa can also be found in intertidal rock pools, below mean high-water neap (MHWN) on exposed, semi-exposed and sheltered shores. In the pool habitat, A. gibbosa are usually found underneath boulders and in crevices on the sides, although, at night individuals often move onto the top of boulders and onto the algae of the pool bottom and sides, where the animals can use a different food supply (Crump and Emson, 1983). Pools with stable boulders are more likely to contain A. gibbosa than those without such large boulders and large pools are, unsurprisingly, more likely to contain large populations (Emson and Crump, 1979), with some tide pools containing an abundance of A. gibbosa (up to 10 per square metre). Such pools lie near MHWN on moderately-exposed shores with broken contours protected from heavy wave action but receiving considerable water exchange daily (Emson and Crump, 1979). They are usually found in areas where
the maximum exposure to air does not exceed three hours on low spring tides (Emson and Crump, 1979). *A. gibbosa* can be found on bedrock and boulders in the kelp zone (*Laminaria hyperborea*), in areas of moderate exposure down to 10m, and even as deep as c. 125m (Mortensen, 1927).

In British waters *A. phylactica* is found in the intertidal zone, mainly in large, deep high level rocks pools on relatively exposed shores which bear an abundant flora of perennial algae characteristically including *Cystoseira* and *Corallina* spp. (Emson and Crump, 1979). Unlike *A. gibbosa*, *A. phylactica* show an apparent disregard for light, with no detectable difference in distribution between day and night. However, the distribution changes during the breeding season, when they show temporary rugophilia and are sensitive to high light levels. A high proportion of adults will be found clustered in crevices in the rock and where these are not available, in and adjacent to the holdfasts of the algae (Crump and Emson, 1983). In the Mediterranean and Adriatic, *A. phylactica* can be common on sheltered and semi-exposed boulder shores, on the undersides of large and small boulders. An example of a typical intertidal site of *A. phylactica* can be seen in Fig 1.3.

*A. phylactica* is also known to occur at sublittoral sites. Specimens have been found at 18m depth near Ramsey Island, South Wales; in an area of strong current flow on the tops of boulders covered with bushy Polyzoa; other specimens have been collected at 1m depth in Lough Hyne, County Cork, Eire, where the animal appears to be abundant (R. Emson, *pers comm.*).

![Image of Chapel Point, Cornwall](image_url)

**Fig. 1.3.** Chapel Point, Cornwall, this is a typical location for *A. phylactica* to inhabit as there is an exposed rocky spur, on which there are high- and mid-shore rockpools containing an abundance of *Corallina* spp. (*pers. obs.*).
A. gibbosa and A. phylactica are opportunistic, omnivorous scavengers, which feed primarily on surface films of diatoms, detritus and bacteria and to some extent on decaying plant or animal material (Crump and Emson, 1983). During observations A. gibbosa has been seen to feed on dead Carcinus maenas, Littorina littorea, Gibbula umbilicalis, A. phylactica (alive), unidentified hydroids, Ulva lactuca, decaying Laminaria fronds and other unidentifiable drift algae, but its normal diet is predominantly microphagous (Crump and Emson, 1983). A. phylactica typically feeds on small epiphytic animals and microorganisms living on perennial algae. At sites where both species co-exist, and occupy the same niche, no difference in food or feeding has been observed (Crump and Emson, 1983).

1.3.3. Distribution

The centre of distribution of A. gibbosa is to the south of the British Isles, with its northernmost limit in Scotland (Crump and Emson, 1983). It is most abundant on the southern coast of the South West of England, with its eastern limit in the Channel being Poole, Dorset. It is also abundant throughout Wales, Ireland, and the Isle of Mann, however, there are only isolated records of the species further north, in West Scotland, and South East Scotland / North East England. In mainland Europe, it is found on the Atlantic coast of France and Spain, and throughout the Mediterranean and Adriatic Seas, as far south as Israel.

A. phylactica is scarce in comparison to A. gibbosa, with a patchy distribution. As for A. gibbosa, A. phylactica is most abundant on the coast of the south west of the UK, with a range which spans from the Isle of Man at its most northernmost limit, with records of it being found sporadically on the Mediterranean coast of France and in the Adriatic Sea at Croatia. In British waters A. phylactica is known from rock pools at several sites in South Wales, Devon and Cornwall. The type population is at West Angle Bay, Pembrokeshire.

Although there are very few published observations of either A. gibbosa or A. phylactica being found in the sub-littoral it is likely that there are some populations in both the Atlantic and Mediterranean. Contact was made with a large number of marine research institutes and marine biology departments around Europe, only one
respondent reported observing *A. gibbosa* in the sub-littoral, where it has been found in trawls of the shallow substratum in the Strait of Sicily (M. Gristina *pers comm*).

**Reproduction**

*A. gibbosa* is a protandrous hermaphrodite which, in British waters, becomes mature as a male at 2 years old and enters the female phase when 4 years old, when 18 – 22mm in diameter (Emson and Crump, 1979). It is not uncommon to find large *A. gibbosa* animals with both ripe spermatozoa and ova present in the gonads and self-fertilisation is possible in principle. External fertilisation, however, is normally the rule (Crump and Emson, 1983).

In late May, females are found laying egg masses of up to 1000 large (0.5mm diameter), bright orange, directly developing eggs on the underside of rocks and in crevices. Individuals show a tendency to aggregate, and to adopt a humped posture while laying the eggs, but once the process is complete the eggs are abandoned. If not consumed by predators such as prawns (*Palaemon elegans, Palaemon serratus*) (Emson and Crump, 1979) these eggs develop to metamorphosis into mobile juveniles in 16 – 21 days (Marthy, 1980).

The juveniles (0.75mm diameter) disperse and may be found on the underside of rocks on the lower shore and in high shore rock pools in early June and grow quite rapidly, relative to growth rates later in the life cycle, attaining a size of 2-3mm diameter by October of the year they hatch (Crump and Emson, 1983). Growth during the winter is minimal but a further burst in April and May enables the animals to achieve an average diameter of 5mm in the first year (Crump and Emson, 1983).

*A. phylactica* is a simultaneous hermaphrodite, mature male gonads being present concurrently with female ova (Emson and Crump, 1979). *A. phylactica* is, at maturity, amongst the smallest starfish known (Crump and Emson, 1983). Compared with *A. gibbosa*, *A. phylactica* matures at a much smaller size, laying viable eggs when a diameter of 5mm is reached (Emson and Crump, 1979). Fecundity increases with size; the smallest mature individuals producing up to 22 eggs and the largest egg mass found having 128 eggs (Emson and Crump, 1979).
late May, *A. phylactica* individuals form aggregations of typically two to five animals two weeks prior to the laying of eggs. Animals remain in these aggregations until breeding is complete (Crump and Emson, 1983). Within these aggregations individuals lay large, directly developing eggs, which appear identical to those laid by *A. gibbosa* in terms of size and rate of development (Emson and Crump, 1979). The most distinctive feature of the reproductive biology of *A. phylactica* is the protection of the developing ova by the adults which adopt a humped posture over the egg mass and remain so until the juveniles have metamorphosed (Crump and Emson, 1983). This can be seen in Fig 1.4. The species name, taken from the Greek *phylacticos*, 'guarding,' refers to this habit. Metamorphosis from fertilised egg to competent juvenile takes up to three weeks depending on temperature. However, not all *A. phylactica* that release eggs brood them to metamorphosis. Crump and Emson (1979), observed some animals moving away from the brooding site during the brooding period.

Fig 1.4. *A. phylactica* aggregating on the underside of a boulder to brood their eggs.

1.3.4. Dispersal

*A. gibbosa* and *A. phylactica* both have an entirely benthic life history, except when juveniles float to the surface when rafting, which may be a dispersal mechanism. Despite this, both species occur on the intertidal throughout the Mediterranean Sea and the North East Atlantic. This geographical size of distribution range is not unusual for species with limited or sporadic dispersal; indeed species with low potential dispersal ability have been identified with far larger ranges, such as the
cheilostome bryozoan, *Callpora weslawski*, which has a brooding life history and a bipolar geographical distribution (Kuklinski and Barnes, 2010), that is, it is found at both poles, but not in-between.

Within the distribution range of *A. gibbosa* and *A. phylactica* there are three well documented major biogeographical boundaries: the English Channel; the Strait of Gibraltar/Almeria-Oran Front; and the Strait of Sicily, as well as numerous oceanic current systems. Species with limited or sporadic dispersal are generally hypothesized to have high inter-population genetic variation (Riginos and Victor, 2001; Sponer and Roy, 2002).

It has been hypothesized that the sea star *Parvulastra* (formerly *Patiriella*) *exigua* can passively disperse for many months or even more than a year in the open ocean (Waters and Roy, 2004), therefore, adult and juvenile rafting on wood or macro algae is the main mechanism proposed for dispersal (Mortensen, 1933; Fell, 1962; Clark and Downey, 1992; Hart et al., 1997; Waters and Roy, 2004).

1.4. The Study Location: the North East Atlantic Ocean and the Mediterranean Sea

1.4.1. The Mediterranean Sea

The Mediterranean is a fully enclosed sea, with the exception of the 12.9km wide, 286m deep Strait of Gibraltar on the western edge and the man-made Suez Canal (opened in 1869) in the South East. The opening of the Suez Canal has led to 343 non-indigenous species being introduced into the Mediterranean Sea from the Indo-Pacific region (Streftaris et al., 2005).

The Mediterranean consists of two basins separated by the 350m sill at the Strait of Sicily. The western basin has an approximate depth of 3400m and the depth of the eastern basin is approximately 4200m. A 40-70m sill (the Bosporous-Dardanelles sill) separates the Black Sea from the Mediterranean Sea and the 160m Pelagosa sill separates the Adriatic from the Eastern Mediterranean Sea (Patarnello et al.,...
These shallow sills have enabled the basins to become isolated from each other over the Quaternary, which may have facilitated the conditions required for allopatric speciation events to occur.

Circulation within the Mediterranean Sea is forced by topography (water exchange through the various Straits), wind stress and buoyancy flux at the surface due to fresh water and heat fluxes (Robinson et al., 2001). At the western edge of the Mediterranean is the Almeria-Oran Front (AOF). The AOF is an oceanographic front between Almeria in South East Spain and Oran in Algeria, which has a pronounced step in temperature (1.4°C) and salinity (2 ppt) over a 2km distance. The AOF has an average water current speed of 4cm/s flowing South East from the coast of Spain towards North Africa (Tintore et al., 1988), it then flows along the North African coast (Algerian current) (Alhammoud et al., 2005). Surface circulation in the Eastern Mediterranean is counter clockwise (Hamad et al., 2005). The location of the AOF and major Mediterranean currents can be seen in Figure 1.6.

Today, there are more than 8500 species of macroscopic organisms inhabiting the Mediterranean Sea (Longhurst, 1988; Bianchi and Morri, 2000). This represents between 4% and 18% of the World’s marine biodiversity, despite the Mediterranean Sea having only 0.82% of the surface area and 0.32% of the World’s oceans. Of these 8500 species, more than 25% are endemic to the Mediterranean Sea. This high level of biodiversity maybe the result of a rich geological history.

During the Cenozoic era, between 5.96 and 5.33 million years ago, the Messinian Salinity Crisis (MSC) occurred. The MSC is the name given to describe the desiccation of the Mediterranean Sea during this period. The MSC is believed to have been triggered by the isolation of the Mediterranean Sea from the World’s oceans by a combination of global sea level lowering and tectonic uplift closing the Mediterranean-Atlantic connections. Prior to the MSC, the Mediterranean Sea was connected to the Atlantic Ocean via several Iberian and North African gateways (Esteban et al., 1996), the deepest of which, the Rifean Corridor in Morocco, was estimated to be 600-800m in depth. The Iberian gateways became reduced in number by the middle to late Miocene, with the closure of the last connection
between the two basins, the Guadalhorce River valley occurring during the early Messinian (Patarnello et al., 2007).

The isolation of the Mediterranean Sea from the Atlantic Ocean resulted in its desiccation into a series of hypersaline and freshwater lakes (occurring in the proximity of river outflows) (Krijgsman et al., 1999). Few coastal marine taxa survived this event (Penzo et al., 1998; Hrbek and Meyer, 2003; Huyse et al., 2004), but some did survive, perhaps in refuge areas, through the Neogene (Stanley, 1990; Bellan-Santini et al., 1992; Myers, 1996). The opening of the Strait of Gibraltar allowed contact between the Atlantic Ocean and the Mediterranean Sea, which initiated the end of the MSC (Krijgsman et al., 1999).

There have been a number of extended glacial and shorter, warmer interglacial events, during the Quaternary, which have primarily shaped the present day distribution of European flora and fauna. The interglacial periods allowed the northward expansion in the range of species that subsequently persisted throughout the glacial maxima in one or more refugia (for reviews see Hewitt, 1999, 2000).

During the Last Glacial Maximum (LGM), which occurred 25-18 kya (thousands of years ago), oscillations of sea level occurred, for example, 12 kya the sea level in the Mediterranean Sea was 65m lower than today, whereas 21 kya the sea level was approximately 110-150m lower than today (Maggs et al., 2008). These fluctuations in sea level led to periods of reduced connectivity between the Eastern and Western basins, which stabilised at about 11 kya (Collina-Girard, 2001).

The area comprising the Strait of Gibraltar and the AOF represents a biogeographical boundary separating the Mediterranean region (to the East), the Lusitanian region (to the North West) and the Mauritanian region (to the South West) (Briggs, 1974). Two factors have been proposed to account for the biogeographical barrier in the area: the one-way surface current of water from the Atlantic Ocean flowing through the Strait of Gibraltar into the Mediterranean Sea (Parilla and Kinder, 1992; Bryden et al., 1994); and the presence of gyres forming a well defined hydrogeographical boundary of surface waters, the AOF (Tintore et al., 1988).
Recent molecular studies have suggested moderately strong to strong genetic
differentiation between each side of the Strait of Gibraltar in swordfish (Kotoulas et
al., 1995), a number of sparid fishes (Bargelloni et al., 2003), sea urchins (Duran et
al., 2004a), sponges (Duran et al., 2004b), Krill (Zane et al., 2000), the Atlantic
bonito (Viñas et al., 2004), prawns (Reuschel et al., 2010), and even A. gibbosa
(Baus et al., 2005). Hypotheses for the observed patterns of differentiation include
isolation-by-distance with limited gene flow or secondary contact of previously
isolated and divergent populations (Perez-Losada et al., 2002). However, no genetic
differentiation was observed in other species including flat oysters (Launey et al.,
2002), sparid fishes (Bargelloni et al., 2003), lobster (Stamatis et al., 2004), mackerel
(Zardoya et al., 2004) and chthamalids (Shemesh et al., 2009).

1.4.2. The Northeast Atlantic Ocean

Passive dispersal, as expected with rafting juveniles, or adults, in the North Eastern
Atlantic Ocean is influenced by the extensive oceanic and coastal currents (Gysels
et al., 2004). The most important of these are the North Atlantic Current (NAC) and
the Shelf Edge Current (SEC), flowing northward along the western coasts of the
British Isles towards the Norwegian Trench. Water also travels eastward from the
Atlantic Ocean along the English Channel towards the North Sea. The major current
systems can be seen in Figure 1.6.

During the LGM, as recent as the Younger Dryas cold phase (12.8-11.5 kya) an ice
sheet covered Scandinavia and some of the British Isles (Lowe and Walker, 1997)
with sea ice possibly extending as far south as the Bay of Biscay (Frenzal et al.,
1992; Svendsen et al., 2004). Marine populations either became extinct or were
forced to retreat southward into one or more southern glacial refugia including the
Spanish-Portuguese-North African coast and the Mediterranean Sea (Garcia Manin
et al., 1999; Koljonen et al., 1999; Verspoor et al., 1999) and/or smaller periglacial
refugia (isolated northern ice-free areas) such as the western English Channel
(Provan et al., 2005; Hoarau et al., 2007) and Southwest Ireland (Provan et al.,
2005; Hoarau et al., 2007). The re-establishment of the NAC occurred 10 kya BP,
allowing marine organisms to recolonize northward from their southern refugia and
the eastern English Channel was gradually flooded, with a connection to the North Sea being established ca. 9 kya (Behre, 2007).

It is expected that marine organisms which inhabited the shallow subtidal and intertidal zones and which have short distance dispersal would have experienced direct loss of habitat due to ice scouring, local extinction and eventually, strong genetic differentiation between isolated refugia (Chevolot et al., 2006).

Figure 1.6. The main current patterns of Western Europe, compiled after Turrell (1992), Millot (1999) and Hansen and Østerhaus (2000). NAC, North Atlantic Current; SEC, Shelf Edge Current (from Gysels et al., 2004).
The Iroise Sea and the north western part of the Brittany coast (France), is a biogeographic transition zone between temperate and cold-temperate/boreal marine faunal assemblages (Cox and Moore, 2000; Dinter, 2001). There is an inflow of warmer, higher salinity water derived from the NAC which enters the western English Channel from the NAC along the north coast of Brittany. Most of the water then circulates northwards to return to the Celtic Sea along the south coast of Cornwall (Jolly et al., 2005), with the remainder flowing north-eastwards on the English side of the Channel (see Beaumont, 1982). The eastern and western province of the English Channel and the English and French coasts in the eastern province are separated by a strong current that flows east to west. In the summer, a frontal system known as the Ushant Front develops between North Brittany and the tip of Cornwall, at the confluence of mixed and stratified waters (Pingree et al., 1975). Deep sequence divergence of the tubeworm Pectinaria koreni has been observed between populations in Brittany and populations in the English Channel using both mtDNA and allozyme markers (Jolly et al., 2005). A genetic break at the English Channel has also been observed for the common goby Pomatoschistus microps (Gysels et al., 2004) and the prawn Palaemon elegans (Reuschel et al., 2010).

The historical separation of the Atlantic and Mediterranean basins and present day gene flow has made species with an Atlantic-Mediterranean distribution of special interest.

1.5. Genetic studies

1.5.1. Choice of molecular marker

The genetic composition of an organism is influenced by the biology and circumstances of the individuals through which it passes, including reproductive success, migration, population size, natural selection and historical events (Sunnucks, 2000). Genetic markers are simply heritable characters with multiple states at each character. All genetic markers reflect differences in DNA sequences, usually with a trade-off between precision and convenience (Sunnucks, 2000).
Population genetics has historically used single locus techniques to investigate specific genes or non-coding regions in either the mitochondrial or the nuclear genome, or investigate allele frequencies at highly polymorphic loci (e.g. microsatellites). The single locus approach is a relatively cheap and easy approach that provides data that are comparable across species and requires a low quantity of DNA (Sunnucks, 2000). Sequence data can detect both silent (i.e. synonymous changes that do not alter the amino acid) and non-silent changes (i.e. non-synonymous changes that result in a change in the amino acid) (Singh, 2003). Markers giving both allele and haplotype frequency and sequence data can be informative over a range of timescales. Problems associated with analyzing only a few loci can include data that provide an incomplete or biased view of the genome and population history (Luikart et al., 2003).

1.5.2. Mitochondrial Genetics

Animal mitochondrial DNA is a small, closed, circular, double stranded DNA molecule approximately 15-17 kb, however this number varies among taxa (Ballard and Whitlock, 2004; Burger et al., 2003). MtDNA encodes 37 genes in most species, 24 of which encode the translation machinery of the mtDNA itself (Ballard and Whitlock, 2004; Ballard and Rand, 2005; Burger et al., 2003). No introns, intergenic sequences or interrupted genes are present in the conserved mtDNA gene content (Moritz et al., 1987). Many phyla have large length variations resulting from insertions in the control region, duplication or deletion of sequences or replication slippage (Moritz et al., 1987), in addition to this gene arrangement varies in some taxa.

MtDNA has a number of attributes that make it a useful tool in evolutionary biology (Wilson et al., 1985; Avise et al., 1987; Harrison, 1989). The mitochondrial protein genes, including the Cytochrome Oxidase I gene (one of the genes under investigation in this study) are extensively used in phylogenetic study as they can reveal divergence at the species or recently diverged species level (Rokas et al., 2003). MtDNA protein genes possess a triplet code for the assembly of proteins that places strong constraints on nucleotide changes at first and second codon positions (Li, 1997, and references therein). This variability of the evolutionary rate of different
nucleotide sites is a general property of DNA sequences (Pesole and Saccone, 2001). However, because of the degenerate nature of the amino acid code, many third and to a lesser extent first codon positions are less constrained and have been observed to evolve at a higher rate (Li, 1997, and references therein). These changes are called synonymous substitutions and they do not cause amino acid changes (Li, 1997).

MtDNA occurs in multiple copies in every cell of an organism, lying outside the cell’s nucleus. This makes mtDNA abundant and easy to amplify, which has enabled it to be used in thousands of population and evolutionary studies (reviewed by Ballard and Rand, 2005) since its inception by Avise et al. (1987). MtDNA evolves rapidly relative to nuclear DNA, with the rate of synonymous substitutions around ten times greater for mtDNA relative to protein coding nuclear genes (Brown et al., 1982).

MtDNA has been used for genetic barcoding. This is an approach which has been utilised to classify individuals as a member of a species using a universal primer pair to amplify an approximately 650-bp region from the 5’ end of the COI gene (Herbert et al., 2003). This technique has been used to classify echinoderms such as holothurian species (bèche-de-mer) (Uthicke et al., 2010) and Echinoidea, Asteroidea, Holothuroidea, Crinoidea and Ophiuroidea (Ward et al., 2008). The perceived advantage of using DNA barcodes is using them is much cheaper, much quicker relative to more traditional approaches and would require no taxonomic experts (Rubinoff et al., 2006). Limitations of DNA barcoding include using only a single gene as a barcode, which largely ignores the issues with mtDNA outlined below. In addition to this, studies tend to use one of the most basic phylogenetic methods available – simple pairwise distances interpreted through phenetic clustering to produce tree-like representatives of species clusters (Neighbour-Joining phenograms).

Within an individual it is the norm to have a single copy of the mitochondrial genome in all cells, homoplasmy, however heteroplasmy (where there is more than one haplotype within a single cell) has been reported in many species of animal and plant (reviewed by Kmiec et al., 2006). Paternal leakage, recombination and small-scale mutations have been identified as causes of heteroplasmy (Kmiec et al., 2006).
Heteroplasmy can lead to unreliable population or evolutionary inference as recombination could produce a novel haplotype (Ladoukakis and Zouros, 2001).

MtDNA tends to be inherited only through the female line as the paternal mitochondria are usually eliminated during fertilization of an egg (Kvist et al., 2003). The consequence of strict maternal inheritance is a simple pattern of inheritance with no mixing between lineages, either through reassortment or through recombination. Any mutation in a female’s mtDNA can be passed onto her progeny and hers alone; the behaviour of males is irrelevant (Birky et al., 1989). However, paternal transmission has been reported in mussels (Ladoukakis and Zouros, 2001), some other invertebrate species (Rokas et al., 2003) and even humans (Kraytsberg et al., 2004). If paternal leakage were to occur then a state of mitochondrial heteroplasmy can be expected, with a mitochondrial genome derived from both parents.

Nuclear mitochondrial pseudogenes (NUMTS) are non-functional copies of mtDNA transferred and incorporated into the nuclear genome, which can retain close homology to the original mitochondrial genes (Bensasson et al., 2001; Tourmen et al., 2002; Woischnik and Moraes, 2002). NUMTS may get co-amplified with the mitochondrial sequence which can lead to inaccurate estimation of heteroplasmy when heteroplasmic sites are incorrectly identified (Hirano et al., 1997). Song et al. (2008) reported that NUMTS if not checked would have overestimated the number of unique species of grasshopper and crayfish using genetic barcoding. In grasshoppers 17 species were identified then reduced to 6 after checking the data for NUMTS, in crayfish the number was reduced from 25 to 7 after checking for NUMTS (Song et al., 2008). Sequences can be assessed for NUMTS by examining the amino acid composition. NUMTS accumulate in-frame stop codons because they become non-functional after nuclear integration and are not under pressure to conserve an open reading frame (Song et al., 2008). Although it is worth noting that even if sequences do not contain in-frame stop codons it does not exclude those sequences from being NUMTS.

MtDNA was considered to be selectively neutral however, first and second position codons evolve less rapidly than third position codons (Stewart et al., 2005). These non-synonymous base changes are often slightly deleterious mutations that are less
favoured by selection, leaving the less harmful third position codon mutations which do not change the amino acid (Nachman, 1998). In addition to this, there are large differences in the rate of mutation of different mtDNA sections, for example, the non-coding control region are far more variable than other mtDNA regions as they are not constrained by function (Ballard and Whitlock, 2004).

MtDNA potentially provides a “molecular clock”. If a DNA sequence accumulates changes at a constant rate, then there exists the possibility of using the degree of genetic divergence to estimate the time of coalescence (Avise, 2000). MtDNA offers arguably the best combination of characteristics for this: it is thought to be predominantly selectively neutral, it is constrained in size such that most mutations are point mutations, and its evolutionary rate is high enough to be informative over relatively short evolutionary periods (Ingman et al., 2000). Since its inception, the concept of the molecular clock has formed the focus of heated debate (Smith and Peterson, 2002). Calibrations of molecular clocks are subject to large errors and there are also large errors associated with their application to entire phylogenies (Heads, 2005). There are no fossil records for Asteroidea species, therefore many studies use a rate of 3.1-3.5% / million years which was calibrated by comparing the level of genetic differentiation between species of the pantropical sea urchin genus Eucidaris on each side of the Isthmus of Panama (Lessios et al., 1999; McCartney et al., 2003; Waters and Roy, 2004). The Isthmus of Panama finally closed 3.1 mya (Lessios et al., 1999) separating the Atlantic and Pacific oceans. This event has been used to calibrate the molecular clocks of many species. However, this technique has been called into doubt as the rates of molecular evolution vary widely within and between lineages (Gillooly, 2005) and molecular estimates of divergence time often disagree with the fossil record (Alroy, 1999; Smith and Peterson, 2002). Variations in rates of nucleotide substitution have been correlated with body size, metabolic rate (Martin and Palumbi, 1993), generation time (Laird et al., 1969) and environmental temperature (Bleiweiss, 1998; Wright et al., 2003). New methods of estimating branch lengths do not assume a strict clock, however this has limitations as the number of models that could be selected from is effectively infinite, so this method is only as accurate as its model (Near and Sanderson, 2004). A further limitation of the molecular clock is that researchers cannot be sure that Atlantic/Pacific species pairs diverged at the time of the rise of the Isthmus or before (Heads,
2005). Therefore, a rate of 3.1-3.5% per million years may be overestimating the rate of mutations for echinoderms.

1.5.3 Asteroidea mtDNA

The phylum Echinodermata comprises five classes: the Crinoidea (sea lilies and feather stars), Ophiuroidea (basket stars and brittle stars), Asteroidea (sea stars), Echinoidea (sea urchins), and Holothuroidea (sea cucumbers) (e.g., Littlewood et al., 1997). Within Asteroidea the mtDNA genomes from seven species belonging to four families are known (Perseke et al., 2010). Phylogenetic analyses using entire mtDNA sequences has discriminated three distinct lineages within Echinodermata: the Crinoidea, the Ophiuroidea, and a group containing Echinoidea, Holothuroidea and Asteroidea based on both amino acid sequence comparisons and gene order evolution (Perseke et al., 2010). Within the group containing Asteroidea there has been a major inversion in the mtDNA of Asteroidea compared to Echinoidea, of a 4.6 kb portion of the tRNA cluster, the 16S rRNA, and the ND1 and ND2 genes. It is hypothesised that these two orders have been separated for at least 225 million years (Smith et al., 1993).

The entire mtDNA genome of A. gibbosa or A. phylactica has not been sequenced, however it has been for the closely related asteroid, Asterina pectinifera (Asakawa et al., 1995). The mtDNA genome in A. pectinifera is 16,260 base pairs long, and contains the genes for cytochrome oxidase subunits I, II and III (COI, II and III), cytochrome b (Cyt b), NADH dehydrogenase subunits 1 – 6 and 4L (ND1/6 and 4L), ATPase subunits 6 and 8 (ATPase 6 and 8), two rRNAs and 22 tRNAs (Asakawa et al., 1995). This can be seen diagrammatically in figure 1.7.
1.5.4 Amplified Fragment Length Polymorphisms

Amplified Fragment Length Polymorphism (AFLP) analysis (Vos et al., 1995) is an effective PCR-based multilocus DNA fingerprinting technique for revealing genetic diversity and differentiation within and among populations (Liu and Cordes, 2004; Bensch and Akesson, 2005). AFLP analysis is highly reproducible, it gives a large number of gene fragments that are treated as bi-allelic gene loci, and it requires minimal development time (Sunnucks, 2000). AFLP has two major disadvantages – fragments must be treated as dominant gene loci and nothing is known about the loci. Despite these disadvantages, the AFLP technique has been used extensively to elucidate genetic diversity and differentiation among populations in plants (e.g. Haldimann et al., 2003; Kjølner et al., 2004) and animals (e.g. Seki et al., 1999; Ajmone-Marsan et al., 1997, 2001; Kai et al., 2002; Ogden and Thorpe, 2002; Takami et al., 2004).

Fig 1.7. Gene organisation of the mitochondrial DNA genome of A. pectinifera (taken from Asakawa et al., 1995).
The per-locus type of genetic information obtained by AFLP is relatively poor, the presence or absence of a DNA fragment can be detected at a given locus but in most studies it is impossible to differentiate between dominate homozygous (1/1) and heterozygous (1/0) genotypes (Bensch and Akesson, 2005).

The AFLP method produces bands that are separated on differences in length using polyacrylamide gel electrophoresis or sequencing robots. A band of a certain length represents a presence allele (scored 1) at such an AFLP locus, if a band of a certain length is not present it is scored as an absence allele (scored 0) (Bensch and Akesson, 2005).

Most population genetic analyses based on AFLP data assume that an absent allele really is absent from the data. This will happen if there is a base substitution in the sequence corresponding to the restriction site for the restriction enzyme or in the sequence corresponding to the additional bases in the pre-amplification (1 bp + 1 bp) and selective (2 bp + 2 bp) amplification. Other types of mutations, such as indel variations or substitutions, may result in a DNA fragment of a different length resulting in bands in a different position on the gel. Two alleles representing the same locus may mistakenly be scored as presence alleles at two different loci, these alternative presence alleles are not independent and thus violate the assumption in analysis of population structure and estimate of population genetic diversity (Bensch and Akesson, 2005). This assumption can also be violated if size homoplasmy occurs i.e. that bands of the same length are not homologous and thus representing two or more AFLP loci, but appear to be one band on the gel. This is of a particular concern in studies of genetic diversity and phylogenetic reconstructions (O’Hanlon and Peakall, 2000; Vekemans et al., 2002).

The protocol starts with the digestion of genomic DNA with two different restriction enzymes, a rare cutter and a frequent cutter (Vos et al., 1995). The frequent cutter generates small DNA fragments that will amplify well and are in the optimal size range for separation on denaturing gels, whereas the rare cutter reduces the number of fragments to be amplified as only the rare cutter/ frequent cutter fragments are amplified (Vos et al., 1995). The two restriction enzymes produce different sticky ends which serve as binding sites for the PCR amplification. The DNA digestion is
followed by two PCR stages, firstly a pre-amplification PCR then by a selective amplification of sets of restriction enzymes is performed. The final stage is the analysis of the amplified fragments on a denaturing gel (Vos et al., 1995).

AFLP’s have proven to be a powerful tool for assessing genetic diversity and population structure in a wide range of plants (Tremetsberger et al., 2003; He et al., 2004; Juan et al., 2004; Muller et al., 2004), fungi (Laitung et al., 2004) and animals (Ajmone-Marsan et al., 2002; Wang et al., 2003; Takami et al., 2004; Baus et al., 2005).

Several studies have used AFLP analyses for the study of marine invertebrate species. Weetman et al. (2007) used AFLP markers to show Atlantic population differentiation among locations from western Britain, the eastern English Channel and the Baltic Sea, in the coastal shrimp *Crangon crangon*. Brazeau et al. (2005) used AFLP markers to show significant genetic differences among the four populations (in the Bahamas, at Key Largo, Florida and in the Gulf of Mexico) of the coral *Agaricia agartcites*. Douek et al., (2002) studied the sea anemone *Actinia equina* from four sampling sites along a 25km strip of the coast of the Istria Peninsula, Croatia, revealing significant genetic differences between two clades, strongly supporting the assumption of cryptic speciation in Adriatic populations of *A. equina*.

### 1.6. Phylogeography

Phylogeography is concerned with the principles and processes governing the spatial arrangements of genealogical lineages, especially those of closely related species (Avise, 2000; 2009). It deals with the historical, phylogenetic components of spatial distribution attempting to characterise the phylogenetic deployment of genealogical lineages across the geographical landscape. Population structure through space and time is shaped by mechanisms that might favour local differentiation, including historical vicariance, habitat discontinuity, larval behaviour, marine currents and local adaptation (Baus et al., 2005). It is these ancestral
interactions, in combination with present day environmental patterns that are the focus of marine phylogeography investigations.

Marine invertebrate species without planktonic larvae are often characterised by low levels of gene flow among populations separated by large distances. However, some organisms with no larval dispersal stage appear widespread and with no significant genetic structure (Oosthuizen et al., 2004; Waters and Roy, 2004; Teske et al., 2007).

The genetic structure of marine organisms has been shaped by several factors, such as: behaviour and reproductive mode of the species; oceanic currents as a means for either dispersal or retention (Lessios et al., 1999); and, past climatic and vicariance events (Wilson and Bernatchez, 1998). It is difficult to disentangle the relative importance of each of these factors for determining population structure in the marine environment. This requires the integration of data from different fields such as paleoclimatology, geology, (pale)oceanography and the biology of the species (Gysels et al., 2004).

1.7. An Overview of Recent Sea Star Genetic Studies

Baus et al. (2005) used AFLP to investigate genetic diversity and gene flow in eight populations from across the range of *A. gibbosa* to: reveal the level of genetic differentiation between populations; determine the current patterns of gene flow between populations; and, evaluate potential natural barriers, such as the Strait of Gibraltar/AOF.

*A. gibbosa* was characterised by high levels of genetic diversity, with the percentage of polymorphic bands across the populations ranging from 48.4% to 78.7%. High levels of genetic differentiation were reported between the Mediterranean and the Atlantic populations, however, there was as much variation attributed to within populations as to between the two geographical groups. This was supported by the loosely grouped structure of the PCA analysis.
Gene flow was evident between the Atlantic populations, with a significant number of individuals being reallocated to a different but closely located population. There was evidence of a strong correlation between genetic differentiation and geographical distance and an isolation-by-distance differentiation pattern. The relatively isolated population at Lough Hyne, Ireland, a landlocked seawater loch, appeared to have maintained some level of gene flow with populations on the South West coast of England.

In the Mediterranean, gene flow appeared to be much more restricted, with all individuals except one being re-allocated to their original population during the assignment test. High levels of genetic differentiation were observed between the Naples Bay, Italy population and the French Mediterranean populations. A barrier at the Strait of Gibraltar or Almeria-Oran Front was given as a potential reason for the genetic differentiation between the groups, however within the Mediterranean there was strong structuring. Therefore, the difference between the two groups is as likely to be caused by a lack of gene flow within the Mediterranean as a result of the various geographical processes occurring there. An isolation-by-distance model is suggested to account for the genetic structuring seen within the Atlantic populations, it is possible that the Mediterranean populations are an extension of that, with little or no gene flow occurring as a result of the geographical distances involved between the two groups.

The sampling was inadequate, with only seven sites sampled (including an aquarium), three on the coast of South West England, one on the Atlantic coast of France, in the Pyrénées-Atlantiques region, two on the French Mediterranean coast and one at Naples, Italy. The sample locations were sporadic, with large distances with no sites, due to the difficulty in finding locations where A. gibbosa could be found. If more populations were found and sampled it may provide a more accurate picture for the large amount of differentiation found between the Atlantic and Mediterranean populations.

Crandall et al., (2008) studied two co-distributed species of sea star using a segment of the COI gene to elucidate their phylogeography. Linckia laevigata has a negative geotaxic planktotrophic life history, with a larval stage of 22-28 days and can be
found from South Africa to the Cook Islands on coral reefs to a depth of 30m. *Protoreaster nodosus* is expected to have a more limited dispersal potential, having a positive geotaxic planktotrophic life history, with a shorter larval stage of 10-14 days and has a smaller range from Sri Lanka to New Caledonia, in lagoons or seagrass meadows to a depth of 5m. Both species showed signs of a recent demographic expansion dating to the Pleistocene, consistent with a range expansion following periods of lowered sea level (Crandall *et al.*, 2008). Indian and Pacific populations of *L. laevigata* were separated at some point in the Pliocene or Pleistocene (Williams and Benzie, 1998; Williams *et al.*, 2002) but with only a moderate genetic structure, whereas *P. nodosus* exhibited strong genetic structure presumably because of the short pelagic duration of its larvae (Crandall *et al.*, 2008). The authors concluded that although it is likely that patterns of genetic diversity and structure have been shaped by sea-level fluctuations, species-specific responses have resulted in different phylogeographic patterns in this shared environment (Crandall *et al.*, 2008).

The brooding sea star *Astrotoma agassizii*, which is widely distributed throughout Antarctica and southern South America, was studied using two mtDNA gene fragments – COI and 16S rDNA (16S). *A. agassizii* was found to be genetically discontinuous across its range with three separate cryptic species that lack morphological distinction (Hunter and Halanych, 2008). Within each of the clades there were high levels of genetic continuity across large geographical distances (>500km), suggesting that having a life history with a low potential for dispersal is not a barrier to dispersal (Hunter and Halanych, 2008).

The spatial genetic variation of *Patiria miniata*, a broadcast-spawning sea star with high dispersal potential was analysed by Keever *et al.* (2009), using microsatellites, mtDNA (the five tRNA genes and part of the COI gene) and nuclear DNA introns (the alpha subunit of the ATP synthetase gene (ATPS) and the glucose-6-phosphate isomerase gene (GPI)). *P. miniata* has a geographic range in the intertidal zone of the northeast Pacific Ocean from Alaska to California. A genetic break was found at Queen Charlotte Sound, British Columbia, which corresponds to genetic breaks in other species such as the kelp *Eisenia arborea* and the turban snail *Astraea gibberosa* (Keever *et al.*, 2009). The authors were surprised that there was no
genetic differentiation across the majority of the range of *P. miniata*, from Vancouver Island to Southern California, which includes a distributional gap in Washington, Oregon and northern California. The authors hypothesized that this distribution was the result of high gene flow across the range disjunction, a recent colonization of Vancouver Island from California migrants, or, a recent extirpation of populations in Washington, Oregon and northern California that has resulted in the fragmentation of a formerly continuous range.

Boissin et al. (2008) used microsatellites and mtDNA (16S) and seven nuclear DNA introns to study the brooding brittle star *Amphipholis squamata*, on the French Mediterranean coast. They found that there were at least four cryptic species within the *A. squamata* complex. Genetic differentiation and an isolation-by-distance pattern were displayed in all of the lineages (Boissin et al., 2008). Within all of the lineages there were haplotypes shared between individuals sampled from remote locations and all major haplotypes were widespread (Boissin et al., 2008). It should be noted that the microsatellite data displayed a total lack of heterozygotes, suggesting a very high selfing rate, which is likely to have favoured the formation of the species complex observed (Boissin et al., 2008).

### 1.8. Hypothesis and aims

The main aim of this project is to examine the comparative phylogeography of the congener species of cushion star *A. gibbosa* and *A. phylactica*. Specific hypotheses to be tested include:

- *A. gibbosa* and *A. phylactica* are two genetically distinct species (Chapter Four);
- there is genetic structuring between populations, concordant with the differences in life history between these two species (Chapter Four);
- there is genetic differentiation between populations in the Atlantic Ocean and Mediterranean Sea as a result of barriers to dispersal and low levels of gene flow (Chapters Two, Three and Four);
• the northern Atlantic Ocean populations are the result of a recent postglacial range expansion, in contrast to the stable Mediterranean populations (Chapters Two and Three).
• and, that due to range expansion at the end of the LGM both species will exhibit an excess of recent mutations, with some populations showing high levels of haplotype diversity associated with periglacial refugia (Chapters Two and Three).

Chapters two, three and four are written as stand-alone entities and will be condensed and reformatted with the intention of submission for publication.

1.9. References


around the Iberian Peninsula indicated by microsatellite DNA variation. *Heredity* 89, 417–424.


Chapter 2

A phylogeographic study of the sea star *Asterina gibbosa* using Mitochondrial DNA
2.1 Abstract

The cushion star *Asterina gibbosa* is a common species inhabiting the intertidal environment around the coast of the North East Atlantic Ocean and Mediterranean Sea. *A. gibbosa* lays eggs which directly develop in lecithotrophic larvae, lacking a pelagic dispersal phase. We studied sequence variation in the mitochondrial DNA (mtDNA), specifically, the five transfer RNA genes and the 5’ end of the cytochrome c oxidase subunit I (COI) gene for 157 individuals from three Mediterranean and 18 Atlantic populations across most of its presently known range of distribution, and compared it to AFLP data collected by Baus *et al.* (2005). Ten haplotypes were identified, with one common haplotype occurring in every population, as well as haplotypes private to both basins. Two of these haplotypes are shared with the closely related sea star *Asterina phyliactica*. Two divergent haplogroups were found, both of which contained individuals from both basins. The haplotypic distribution found in *A. gibbosa*, with the presence of one abundant haplotype and a large number of closely related haplotypes, is typical of species experiencing reduction in variability and subsequent expansions, with fragmented populations coming into secondary contact. The AFLP data showed that there was clear structuring between the two basins, with some gene flow occurring between Atlantic populations but very little or no gene flow occurring within the Mediterranean; these signals were probably lost from the mtDNA data as a result of the range expansion. Climatic fluctuations related to glacial cycles during the Pleistocene could explain the present level of variability and nucleotide diversity found.

2.2 Introduction

Starting in the Pliocene (~2.5 million years ago (mya)) there was a global cooling which resulted in permanent ice sheets being formed at high northern latitudes. There have been approximately 40 glacial and interglacial phases, as well as smaller cycles of a few tens of thousands of years and shorter duration (Lambeck *et al.*, 2002). As a consequence, both water temperature and sea level underwent considerable changes over the ongoing Quaternary with major sea-level cycles occurring at intervals of ~100,000 years (100 kya) over the past ~800 kya, with
maximum amplitudes of 120–140m. Throughout this period the connection of the Mediterranean to the Atlantic was not broken (Flores et al., 1997). Different regions of the North East Atlantic Ocean and the Mediterranean Sea are thought to have been differentially affected by reduction in sea surface temperature associated with the Pleistocene glaciations (Lambeck et al., 2002). The Last Glacial Maximum (LGM) at 20 kya greatly influenced present-day distributions of marine species in the North Atlantic (Olsen et al., 2010). The result of the glacial periods was local extinctions and latitudinal shifts of many taxa, followed by recolonisation from less affected regions that may have acted as refugia during the glacial periods (Almada et al., 2001; Domingues et al., 2006), for example there were areas of the Atlantic coast which acted as marine refugia for species including brown algae (Hoarau et al., 2007) and green crab *Carcinus maenas* (Roman and Palumbi, 2004).

Given this historical backdrop, it may be expected that northern populations will show the genetic signature of a short demographic history, consistent with a population expansion from southern refugia after the LGM (Hewitt 2001, 2004). However, this signal can be lost in certain circumstances, including during admixture, when formerly isolated populations come into secondary contact with each other producing an even greater level of diversity than either parent population (Petit et al., 2003). The response to the LGM is not consistent across taxa inhabiting the rocky intertidal shore, with many species showing patterns of differentiation and divergence consistent with long-term regional persistence (e.g. Hickerson and Ross, 2001; Wares and Cunningham, 2001; Marko, 2004; Hickerson and Cunningham, 2005; Hoarau et al., 2007; Marko et al., 2010). It has been hypothesized that because of the large variability in temperature experienced by rocky shore species, simple climate-driven changes to a species geographical range might therefore not occur (Marko et al., 2010; Helmuth et al., 2002). However, within the North East Atlantic the transfer of water to land-based ice sheets caused sea-level to drop rapidly, by 30-40m within 1000-2000 years; during such times the coastline retreated seaward by up to several hundred kilometres (Maggs et al., 2008).

Population connectivity is mainly determined by the potential dispersal of a species, therefore a species with a restricted dispersive ability should present genetically structured populations (Palumbi, 2003). For marine invertebrates with a direct
developing life history, dispersal and gene flow is likely to be the result of drifting or rafting on natural and anthropogenic floating material; this has been an inferred dispersal mechanism for littoral marine invertebrates (Ingólfsson, 1995; Ó Foighil et al., 1999, 2001; Castilla & Guíñez, 2000). Some species possess traits that make them much more suitable for rafting than others (Thiel, 2003). If rafting is deemed to be the most likely method for dispersal then the extensive oceanic and coastal currents need to be considered as a factor influencing the distribution and genetic variability of a species (Domingues et al., 2007). The most important currents around the UK and Atlantic coasts of France, Spain and Portugal are the North Atlantic Current (NAC) and the Shelf Edge Current (SEC), flowing northward to the western coasts of the British Isles then towards the Norwegian trench (Gysels et al., 2004). Atlantic water also flows northward through the English Chanel towards the North Sea (Gysels et al., 2004). On the Atlantic coast of Northwest Europe there are large tidal movements and strong coastal currents are common. Within the Mediterranean many small eddies and other local currents exist (Send et al., 1999) with a much smaller tidal range and with a lower amplitude of the associated tidal currents (Baus et al., 2005).

The Atlantic and the Mediterranean are linked by the Gibraltar Strait, which is characterized by a two-layer flow regime. Atlantic waters inflow into the Mediterranean in the upper layer and Mediterranean waters outflow into the Atlantic in the lower layer (Malanotte-Rizzoli and Bergamasco, 1989; Özgökmen et al., 2001). Close to the Strait of Gibraltar is the Almería-Oran Front, which is a quasi-permanent anticyclonic gyre (Millot, 1999). However, these barriers are species dependent, with some species such as the prawn Palaemon elegans (Reuschal et al., 2010), the shrimp Crangon crangon (Luttikhuizen et al., 2008), seabream Diplodus puntazzo (Bargelloni et al., 2005), sea bass Dicentrarchus labrax (Lemaire et al., 2005), cuttlefish Sepia officinalis (Pérez-Losada et al., 2002) and the sponge Crambe crambe (Duran et al., 2004) as well as the subject of this study, A. gibbosa (Baus et al., 2005), showing high levels of genetic differentiation between Atlantic and Mediterranean populations, while others such as damselfish Chromis chromis (Domingues et al., 2005), sea bream Diplodus sargus (Bargelloni et al., 2005; Domingues et al., 2007a), wrasse Thalassoma pavo (Costagliola et al., 2004) and Norway lobster Nephrops norvegicus (Stamatis et al., 2004) show no genetic
partition between the Atlantic and the Mediterranean. Both vicariance and dispersal are likely to have played an important role in the evolutionary history of the marine fauna of the North East Atlantic Ocean and the Mediterranean Sea (Domingues et al., 2007).

*A. gibbosa* can be found in rock pools and under boulders in the littoral zone, as well as in the sublittoral (Crump and Emson, 1983). The species distribution is considered to be Lusitanian (Bullimore and Crump, 1982) stretching from the British Isles into the Mediterranean Sea. *A. gibbosa* is a protandrous hermaphrodite where self-fertilisation is possible (Crump and Emson, 1983). The life history of *A. gibbosa* consists of the mature individuals laying up to 1000 sticky, bright orange, yolky eggs which are laid in masses attached to the substratum, which develop directly into lecithotrophic brachiolariae which, from hatching until metamorphosis adhere to the substratum (Ludwig, 1882; MacBride, 1896; Marthy, 1980), therefore lacking a planktonic larval stage. *A. gibbosa* is an opportunist, omnivorous scavenger, which feeds primarily on surface films of diatoms, detritus and bacteria and to some extent on decaying plant or animal material (Crump and Emson, 1983). It is this life history and behaviour that is expected to influence migration in *A. gibbosa* among different populations.

The present study analyses the Mitochondrial DNA (mtDNA), specifically five transfer RNA genes, a non-coding intron and the 5' end of the COI gene of *A. gibbosa* to establish a phylogeographic comparison with the AFLP population genetic data generated by Baus et al. (2005). MtDNA sequence data have been used in a large number of studies for inferring the evolutionary and demographic past of both populations and species (Avise et al., 1987; Avise, 2000; Palumbi et al., 1997), while also giving a picture of gene flow between populations (Grant and Waples, 2000). Mitochondrial genetic variability is most appropriate for phylogeny, phylogeography and population genetic approaches as this molecule, due to its rapid evolutionary rate, contains the signatures of historical demographic events that can play a major role in establishing contemporary biogeographic patterns and population structure (Couceiro et al., 2007). However, a matrilineal phylogeny inferred from mitochondrial data gives us only a fraction of the genealogical information (Avise, 1998). Different genetic markers can provide different ecological and demographic information about
a species (Avise, 2004) therefore it is advantageous to use both a molecular and mtDNA marker to obtain a more comprehensive picture of phylogeographical patterns (Ballard and Whitlock, 2004).

This study aims to examine the following hypotheses - that mitochondrial DNA analysis will reveal strong genetic differentiation between populations in the Mediterranean Sea and the Atlantic Ocean, as was found previously using AFLP by Baus et al. (2005); that due to its limited dispersal ability, regional genetic structure can be detected throughout the range of *A. gibbosa*; and, that due to the biogeographic history of the region in which it is found, there is a postglacial pattern of range expansion from South to North resulting from the LGM.

### 2.3 Methods

#### 2.3.1 Sampling

*A. gibbosa* samples were collected from under boulders and stones of rocky sea shores and tide pools from seventeen locations selected across the species' range: twelve sites in the British Isles were chosen, as was one site on the Atlantic Coast of Ireland, one site on the Atlantic coast of France and two sites on the French Mediterranean coast and one site on the west coast of Italy (See Table 2.1 and Fig 2.1 for details). All specimens were stored for up to five days in absolute ethanol at ambient temperature and then for three days up to a three months at -80°C.
<table>
<thead>
<tr>
<th>Basin</th>
<th>Country</th>
<th>Population</th>
<th>Latitude</th>
<th>Longitude</th>
<th>Abbreviation</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Atlantic</td>
<td>Outer Hope, S. Devon</td>
<td>50°14'N</td>
<td>3°51'W</td>
<td>APP</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>Atlantic</td>
<td>Ayrmer Cove, S. Devon</td>
<td>50°17'N</td>
<td>3°54'W</td>
<td>ACG</td>
<td>13</td>
</tr>
<tr>
<td>3</td>
<td>Atlantic</td>
<td>Bucks Mill, N. Devon</td>
<td>50°59'N</td>
<td>4°20'W</td>
<td>BMAG</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>Atlantic</td>
<td>Chapel Point, S. Cornwall</td>
<td>50°15'N</td>
<td>4°46'W</td>
<td>CPAG</td>
<td>5</td>
</tr>
<tr>
<td>5</td>
<td>Atlantic</td>
<td>Gara Point, S. Devon</td>
<td>50°18'N</td>
<td>4°04'W</td>
<td>GPAG</td>
<td>10</td>
</tr>
<tr>
<td>6</td>
<td>Atlantic</td>
<td>Guéthary, Pyrénées-Atlantiques</td>
<td>43°25'N</td>
<td>1°37'W</td>
<td>FGAG</td>
<td>17</td>
</tr>
<tr>
<td>7</td>
<td>Atlantic</td>
<td>Ladram Bay, S. Devon</td>
<td>50°39'N</td>
<td>3°16'W</td>
<td>LBG</td>
<td>14</td>
</tr>
<tr>
<td>8</td>
<td>Atlantic</td>
<td>Littleham Cove, S. Devon</td>
<td>50°36'N</td>
<td>3°21'W</td>
<td>LCG</td>
<td>5</td>
</tr>
<tr>
<td>9</td>
<td>Atlantic</td>
<td>Lough Hyne, Co. Cork</td>
<td>50°30'N</td>
<td>9°18'W</td>
<td>LHAG</td>
<td>10</td>
</tr>
<tr>
<td>10</td>
<td>Atlantic</td>
<td>Prawle Point, S. Devon</td>
<td>50°13'N</td>
<td>3°40'W</td>
<td>PAG</td>
<td>18</td>
</tr>
<tr>
<td>11</td>
<td>Atlantic</td>
<td>Rockham Bay, N. Devon</td>
<td>51°10'N</td>
<td>4°12'W</td>
<td>RBG</td>
<td>10</td>
</tr>
<tr>
<td>12</td>
<td>Atlantic</td>
<td>Marazion, S. Cornwall</td>
<td>50°07'N</td>
<td>5°28'W</td>
<td>SMMAG</td>
<td>6</td>
</tr>
<tr>
<td>13</td>
<td>Atlantic</td>
<td>Welcombe Mouth, N. Devon</td>
<td>50°55'N</td>
<td>4°32'W</td>
<td>WMG</td>
<td>8</td>
</tr>
<tr>
<td>14</td>
<td>Atlantic</td>
<td>West Angle Bay, Pembrokeshire</td>
<td>51°41'N</td>
<td>5°06'W</td>
<td>WAB</td>
<td>8</td>
</tr>
<tr>
<td>15</td>
<td>Mediterranean France</td>
<td>Banyuls-sur-Mer, Pyrénées-Orientales</td>
<td>42°29'N</td>
<td>3°07'E</td>
<td>FBAG</td>
<td>12</td>
</tr>
<tr>
<td>16</td>
<td>Mediterranean Italy</td>
<td>Naples Bay, Naples</td>
<td>40°49'N</td>
<td>14°14'E</td>
<td>INAG</td>
<td>13</td>
</tr>
<tr>
<td>17</td>
<td>Mediterranean France</td>
<td>Les Embiez, Provence-Alpes-Côte d'Azur</td>
<td>43°04'N</td>
<td>5°47'E</td>
<td>FEAG</td>
<td>3</td>
</tr>
</tbody>
</table>

**Table 2.1.** Sampling information for *A. gibbosa.*
Fig. 2.1. A map of the sampling locations of *A. gibbosa*. The numbers relate to the numbers assigned to populations in Table 2.1.

### 2.3.2 DNA extraction

DNA extractions were performed on 2–4 mm$^3$ of tissue (arm tip) using a DNeasy extraction kit (QIAGEN, catalogue # 69506) or by phenol-chloroform and CTAB purification following the method of Arndt *et al.* (1996). DNA quality and quantity was assessed by electrophoresing samples on 1% agarose gels alongside dilutions of phage lambda DNA (Promega, catalogue # D1501). The samples were subsequently stored at -20°C.

### 2.3.3 mtDNA amplification and sequencing

The amplification of the mitochondrial genome that contains five transfer RNA genes and the 5' end of the COI gene (Hart *et al.*, 1997) was accomplished using primers initially designed by M.W. Bruford, which were synthesised by GIBCO.
Primer 1 (F121-146) is 25 bases long with a 5' to 3' sequence consisting of:
CAG TAA CCA CTT TGC TAC CCC AGT C

Primer 2 (500-476 BACK) is 25 bases long with a 5' to 3' sequence consisting of:
GTA TGA TAA CGC TCA TTG CAG TTC C

All PCRs were performed in a PE 9700 thermal cycler. MtDNA amplification was performed using the following program: 94°C for 5 minutes; followed by 94°C for 30 seconds, 50°C for 30 seconds, 72°C for 1 minute for 35 cycles with a final extension of 72°C for 10 minutes. The reaction conditions used were 1x PCR buffer, 1.5 mM MgCl₂; 0.2 mM dNTPs; 1 μM of each primer; 0.1 U Taq (Invitrogen); and 1 μl DNA template (diluted between 1/10 and 1/1000) in a final volume of 15 μl.

PCR products were purified using a BIO 101 Geneclean Turbo PCR Kit (Q-BIOgene) or ExoSap (Amersham Biosciences) according to the manufacturer's instructions. Sequencing PCR reactions were performed in 7.5 μl reactions. The PCR program involved an initial denaturation at 96°C for 1 minute 30 seconds for 1 cycle; denaturation at 96°C for 10 seconds; annealing at 50°C for 5 seconds; extension at 60°C for 4 minutes for 25 cycles. Both forward primer and reverse primer reactions were performed. The reaction conditions were as follows: PCR H₂O and template combined 3 μl; Better Buffer (Web Scientific Ltd) 2.5 μl; Big Dye 0.5 μl (ABI PRISM® Big Dye TM Terminator dye vs. 3.1); either forward or reverse primer (1.6 μM) 1.5 μl.

An ABI Prism 3100 semi-automated genetic analyser (Applied Biosystems) was used for the sequencing according to the manufacturer's instructions.

### 2.3.4 Authentic sequence verification
The mtDNA forward and reverse sequences were aligned using Geneious vs. 4.0.2 (Biomatters Ltd.) and then verified by eye. The amino acid reading frame was identified by aligning the full sequence (Accession no. U50058, Hart et al. 1997) with the haplotypes identified and checking the echinoderm mitochondrial genetic code.

### 2.3.5 Analysis of Genetic diversity and differentiation
Estimates of nucleotide diversity and haplotype diversity were obtained using the program Arlequin vs. 3.11 (Excoffier et al., 2005). Comparison of haplotype and
nucleotide diversity was used to reveal information about patterns of historical demography. High haplotype diversity in conjunction with low nucleotide diversity can suggest recent population growth while high haplotype diversity with high nucleotide diversity is indicative of a stable population (Mila et al., 2000).

To visualize the relationship between the haplotypes and their frequencies, a median joining network (MJN) (Bandelt et al., 1999) was constructed in NETWORK vs. 4.1.1.2. (www.fluxus-engineering.com). Phylogenetic estimations were also employed using Neighbour Joining (NJ; Saitou and Nei, 1987) implemented in Geneious Pro (ver. 5.0.4) using the Tamura Nei model of evolution. The Tamura Nei model (Tamura and Nei, 1993) corrects for multiple hits, taking into account the differences in substitution rate between nucleotides and the inequality of nucleotide frequencies. It distinguishes between transitional substitution rates within purines or pyrimidines and transversional substitution rates between purines and pyrimidines. In addition to this it assumes equality of substitution rates among sites. Node support was tested using 1000 bootstrap replicates. A bootstrap consensus tree was inferred from 1000 replicates (Felsenstein, 1985). Branches corresponding to partitions reproduced in less than 50% of bootstrap replicates were collapsed and the percentage of replicate trees in which associated haplotypes clustered together is shown.

Pairwise divergences between populations (sampling locations) and divergence within and between the Atlantic Ocean and the Mediterranean Sea and between haplogroups were calculated using MEGA 3 (Kumar et al., 2004) using the Tamura Nei model of evolution (Tamura and Nei, 1993).

To estimate population structure we conducted an analysis of molecular variance (AMOVA) using Arlequin vs. 3.11. AMOVA was used to investigate genetic differentiation between Atlantic and Mediterranean populations and to confirm the subdivided groups identified by the MJN and NJ Tree. AMOVA uses the frequencies of haplotypes and the number of mutations between them to test the significance of the variance components associated with hierarchical levels of genetic structure (within populations, among populations within groups, and among groups) by means of non-parametric permutation (Excoffier et al. 1992).
2.3.6 Analysis of Population Demography

To infer the demographic history of populations, we compared mismatch distributions of pairwise nucleotide differences among mtDNA haplotypes with expectations of a sudden-expansion model (Rogers, 1995) using Arlequin (vs. 3.11) (Excoffier et al., 2005). If a population has undergone rapid expansion, a unimodal mismatch distribution approximating a Poisson curve is expected, the shape of which is expected to be not significantly different from the model of sudden expansion (i.e. P>0.05) (Rogers and Harpending, 1992), whereas populations approaching mutation drift equilibrium are expected to produce a multimodal or ‘ragged’ mismatch distribution and a significantly different P value. The raggedness statistic quantifies the smoothness of the observed pairwise differences distribution.

Tests based on pairwise comparisons, such as the mismatch distribution, lose much of the historical information available in the data (Felsenstein, 1992). Neutrality tests, especially Fu’s (1997) FS, have been shown to be much more powerful in detecting signals of population growth (Ramos-Onsins & Rozas, 2002). Therefore, Tajima’s D (Tajima, 1989), Fu’s Fs (Fu, 1997) and Fu and Li’s F* and D* (Fu and Li, 1997), were conducted. Tajima’s D and Fu’s Fs (10000 simulated samples) were calculated in ARLEQUIN and Fu and Li’s F* and D* statistics were calculated in DNASP vs. 4.0 (Rozas et al., 2003).

Fu (1997) demonstrated that comparisons of neutrality tests can distinguish between genetic hitch-hiking and population expansion, two processes that can leave similar imprints on the data. If Fu & Li’s (1993) F* and D* are significant but Fu’s (1997) FS is not, then background selection is indicated. If the reverse is true (FS is significant, but F* and D* are not), then population expansion is indicated (Russell et al., 2005). A negative Tajima’s D statistic and significance value indicates either population expansion or background selection. P is the probability of obtaining the observed D value under the neutral mutation hypothesis (P< 0.05 = the data differs significantly from neutral mutation, P> 0.05 the data does not differ significantly from zero).

An approximate estimation of time since expansion was calculated using \( \tau = 2ut \) (Rogers and Harpending, 1992), where \( \tau \) (tau) is the mode of the mismatch distribution, i.e. the estimation of the age of expansion (obtained from ARLEQUIN vs.
3.11), \( u \) = the mutation rate per sequence per generation and \( t \) = time in generations since expansion. The value of \( t \) was calculated using the equation \( t = \frac{\tau}{2u} \), and the value of \( u \) was calculated using the equation \( u = 2\mu k \), where \( 2\mu = \) the nucleotide divergence rate (twice the mutation rate per nucleotide) per million years (3.1 - 3.5%, Lessios et al., 1999; McCartney et al., 2003; Waters and Roy, 2004) and \( k \) = sequence length. Time in generations since expansion (\( t \)) was then multiplied by the generation time - four years (Emson and Crump, 1978, cited in Emson and Crump, 1979) which gave an estimated time in years since expansion.
2.4 Results

2.4.1 Sequences and Haplotype Analysis

Table 2.2. mtDNA haplotypes of *A. gibbosa*. Data comprises: total number of haplotypes identified in each sampling location and the number of individuals found with each haplotype in each sampling location. Populations above the line are found in the Atlantic Ocean whereas populations found below the line are found in the Mediterranean. The blue haplotypes are grouped into haplogroup 1 whereas the red haplotypes are grouped into haplogroup 2.

A 332-bp sequence from 157 individuals from 18 populations, 15 in the North East Atlantic Ocean and 3 situated in the Mediterranean Sea, revealed 18 polymorphic sites, giving 10 haplotypes (Table 2.2). Of these polymorphisms, five occurred at first
codon positions, four occurred at second codon positions and nine occurred at third codon positions. There was a two base pair insertion in the non-coding region prior to the COI gene.

Fig. 2.2. Geographical positions of sampling localities (see Table 2.1) and distribution of *A. gibbosa* haplotypes. The size of the circle corresponds to the number of individuals sampled from a population, with each coloured slice corresponding to a haplotype.

Of the ten haplotypes, six were present in the Mediterranean, two of which were exclusive to that basin and eight were present in the Atlantic, four of which exclusively (Fig 2.2). The haplotype with the greatest frequency in the Mediterranean (AG4) was only found in two individuals with the Atlantic populations sampled, in single individuals on the Atlantic coast of South West France and Lough Hyne, Ireland. The two private Mediterranean haplotypes have a low frequency, they are found in single individuals only, likewise, the four private Atlantic haplotypes are also found in a similar low frequency.
Fig. 2.3. MJN for *A. gibbosa* mtDNA haplotypes, indicating the haplotype relationships. Circles represent haplotypes, the circle size is proportional to the haplotype frequency. Each cross-line is indicative of a site which has undergone a substitution. Blue = Individuals from Atlantic Populations, Red = Individuals from Mediterranean Populations, Yellow = median vectors of un-sampled or extinct ancestral sequences.
Fig. 2.4. Neighbour joining phenogram tree based on Tamura Nei genetic distances for *A. gibbosa* mtDNA haplotypes, with bootstrap values on the branches. The tree is rooted by *Patiriella exigua* (Genbank No. U50053 (Hart et al., 1997)).

The MJN (Fig. 2.3) shows that there are two divergent clades (haplogroups), with the most common haplotype within each haplogroup being separated by nine nucleotide substitutions. The two haplogroups overlap in their distribution, neither is confined to a specific geographical area. Phylogenetic reconstruction of the ten haplotypes (plus an outgroup of *P. exigua*) revealed two main groups with NJ bootstrap support of 100% (Fig. 2.4), this supported the data seen in the MJN.
Table 2.3. Comparative summary statistics for mtDNA diversity within *A. gibbosa* within both the Atlantic and Mediterranean basins.

<table>
<thead>
<tr>
<th></th>
<th>Atlantic</th>
<th>Mediterranean</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sample Size</strong></td>
<td>129</td>
<td>28</td>
</tr>
<tr>
<td><strong>No. of Populations</strong></td>
<td>14</td>
<td>3</td>
</tr>
<tr>
<td><strong>No. of Haplotypes</strong></td>
<td>8</td>
<td>6</td>
</tr>
<tr>
<td><strong>Nucleotide Diversity</strong></td>
<td>0.008 +/- 0.005</td>
<td>0.023 +/- 0.013</td>
</tr>
<tr>
<td><strong>Haplotype (gene) Diversity</strong></td>
<td>0.278 +/- 0.050</td>
<td>0.720 +/- 0.052</td>
</tr>
<tr>
<td><strong>Mean no. pairwise differences</strong></td>
<td>2.735 +/- 1.460</td>
<td>5.669 +/- 2.802</td>
</tr>
<tr>
<td><strong>Fu's Fs Statistic</strong></td>
<td>2.283 (P &gt; 0.10)</td>
<td>4.537 (P &gt; 0.10)</td>
</tr>
<tr>
<td><strong>Fu and Li's F</strong></td>
<td>1.089 (P &gt; 0.10)</td>
<td>-0.046 (P &gt; 0.10)</td>
</tr>
<tr>
<td><strong>Fu and Li's D</strong></td>
<td>1.173 (P &gt; 0.10)</td>
<td>0.285 (P &gt; 0.10)</td>
</tr>
<tr>
<td><strong>Tajimas D</strong></td>
<td>-0.025 (P &gt; 0.10)</td>
<td>1.950 (P &gt; 0.10)</td>
</tr>
</tbody>
</table>

The number of haplotypes found in each basin was eight in the Atlantic and six in the Mediterranean. The number of haplotypes found in the Mediterranean is surprisingly high considering there were only 28 individuals sequenced, compared to 129 individuals sequenced in the Atlantic. Differences between the Atlantic and Mediterranean populations can also be seen in the nucleotide and haplotype diversities, both were far greater in the Mediterranean than in the Atlantic (Table 2.3). The mean number of pairwise differences was twice as great for individuals in the Mediterranean compared to individuals in the Atlantic.

Within both basins there is high haplotype diversity and low nucleotide diversity, which suggests that a population expansion event or a selective sweep may have occurred, however, all of the neutrality tests gave non-significant values (Table 2.3).
2.4.2 Genetic Diversity and Population Structure

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>d.f.</th>
<th>Sum of Squares</th>
<th>Sum of Variance</th>
<th>Percentage of Variation</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Components</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No Regional Groups</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Among populations</td>
<td>17</td>
<td>70.052</td>
<td>0.236</td>
<td>10.16</td>
<td>&lt;0.05</td>
</tr>
<tr>
<td>Within populations</td>
<td>136</td>
<td>290.566</td>
<td>2.090</td>
<td>89.84</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Total</td>
<td>156</td>
<td>360.618</td>
<td>2.327</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No Regional Groups – Atlantic Ocean Only</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Among populations</td>
<td>14</td>
<td>34.302</td>
<td>0.144</td>
<td>10.46</td>
<td>&lt;0.05</td>
</tr>
<tr>
<td>Within populations</td>
<td>114</td>
<td>140.722</td>
<td>1.234</td>
<td>89.54</td>
<td>&lt;0.05</td>
</tr>
<tr>
<td>Total</td>
<td>128</td>
<td>175.023</td>
<td>1.379</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No Regional Groups – Mediterranean Sea Only</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Among populations</td>
<td>2</td>
<td>5.536</td>
<td>0.163</td>
<td>10.32</td>
<td>&lt;0.1</td>
</tr>
<tr>
<td>Within populations</td>
<td>25</td>
<td>35.5</td>
<td>1.42</td>
<td>89.68</td>
<td>&lt;0.1</td>
</tr>
<tr>
<td>Total</td>
<td>27</td>
<td>41.036</td>
<td>1.583</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Two Groups: Atlantic/Mediterranean</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Among Groups</td>
<td>1</td>
<td>4.303</td>
<td>-0.011</td>
<td>-0.49</td>
<td>&gt;0.05</td>
</tr>
<tr>
<td>Among populations within groups</td>
<td>16</td>
<td>65.749</td>
<td>0.240</td>
<td>10.35</td>
<td>&lt;0.05</td>
</tr>
<tr>
<td>Within Populations</td>
<td>139</td>
<td>290.566</td>
<td>2.090</td>
<td>90.14</td>
<td>&lt;0.05</td>
</tr>
<tr>
<td>Total</td>
<td>156</td>
<td>360.618</td>
<td>2.319</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2.4. ΦST AMOVA for *A. gibbosa* samples indicating the groups tested, and showing the percentage of genetic variation accounted for when the data is divided up into different groups.

An AMOVA was performed (Table 2.4) to test for population structure between the Mediterranean and Atlantic, as well as to test for population structure within each basin. When no regional groups were tested the large majority of the variation (89.84% (P<0.01)) occurs within populations. When the Atlantic populations were analysed alone, within the Atlantic 90% (P<0.05) was attributed to diversity within populations, with just 10% (P<0.05) being attributed to differentiation between populations. When the Mediterranean populations were analysed alone, within the Mediterranean 90% (P<0.10) was attributed to diversity within populations, with just 10% (P<0.10) being attributed to differentiation between populations. Furthermore,
when the range was split into the Atlantic Ocean and Mediterranean Sea most variation was still partitioned within populations (90.14% (P<0.05)).

The Tamura Nei pairwise divergence revealed very little population structure, with values ranging from 0 (between 14 pairs of populations) to 0.036 (between FBAG (Mediterranean) and PPAG (Atlantic)) (Table 2.5). There was an average Tamura Nei pairwise genetic distance of 0.032% between all populations and 0.016% between groups of populations found in the Mediterranean and Atlantic basins.
<table>
<thead>
<tr>
<th></th>
<th>ACG</th>
<th>LCG</th>
<th>INAG</th>
<th>GPAG</th>
<th>PPAG</th>
<th>FBAG</th>
<th>LHAG</th>
<th>SPAG</th>
<th>BMAG</th>
<th>LBG</th>
<th>FEAG</th>
<th>WAB</th>
<th>RBAG</th>
<th>WMG</th>
<th>SMMAG</th>
<th>FGAG</th>
<th>CPAG</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACG</td>
<td>72</td>
<td>3782</td>
<td>10</td>
<td>19</td>
<td>3143</td>
<td>436</td>
<td>82</td>
<td>291</td>
<td>82</td>
<td>3285</td>
<td>325</td>
<td>307</td>
<td>291</td>
<td>126</td>
<td>853</td>
<td>60</td>
<td></td>
</tr>
<tr>
<td>LCG</td>
<td>0.031</td>
<td>3835</td>
<td>82</td>
<td>55</td>
<td>3196</td>
<td>502</td>
<td>10</td>
<td>357</td>
<td>10</td>
<td>3338</td>
<td>394</td>
<td>372</td>
<td>335</td>
<td>190</td>
<td>904</td>
<td>129</td>
<td></td>
</tr>
<tr>
<td>INAG</td>
<td>0.013</td>
<td>0.033</td>
<td>3777</td>
<td>3764</td>
<td>959</td>
<td>3808</td>
<td>3841</td>
<td>3853</td>
<td>3841</td>
<td>785</td>
<td>3887</td>
<td>3854</td>
<td>3832</td>
<td>3714</td>
<td>3518</td>
<td>3735</td>
<td></td>
</tr>
<tr>
<td>GPAG</td>
<td>0.006</td>
<td>0.031</td>
<td>0.017</td>
<td>27</td>
<td>3138</td>
<td>428</td>
<td>90</td>
<td>285</td>
<td>90</td>
<td>3281</td>
<td>317</td>
<td>298</td>
<td>283</td>
<td>122</td>
<td>851</td>
<td>50</td>
<td></td>
</tr>
<tr>
<td>PPAG</td>
<td>0.026</td>
<td>0.031</td>
<td>0.03</td>
<td>0.027</td>
<td>3092</td>
<td>454</td>
<td>55</td>
<td>314</td>
<td>55</td>
<td>3243</td>
<td>344</td>
<td>320</td>
<td>288</td>
<td>143</td>
<td>832</td>
<td>74</td>
<td></td>
</tr>
<tr>
<td>FBAG</td>
<td>0.013</td>
<td>0.04</td>
<td>0.023</td>
<td>0.019</td>
<td>0.036</td>
<td>3170</td>
<td>3203</td>
<td>3215</td>
<td>3203</td>
<td>225</td>
<td>3248</td>
<td>3182</td>
<td>3193</td>
<td>3075</td>
<td>2847</td>
<td>3061</td>
<td></td>
</tr>
<tr>
<td>LHAG</td>
<td>0.007</td>
<td>0.032</td>
<td>0.018</td>
<td>0.012</td>
<td>0.028</td>
<td>0.018</td>
<td>510</td>
<td>351</td>
<td>510</td>
<td>3312</td>
<td>291</td>
<td>356</td>
<td>338</td>
<td>323</td>
<td>1067</td>
<td>385</td>
<td></td>
</tr>
<tr>
<td>SPAG</td>
<td>0.015</td>
<td>0.031</td>
<td>0.023</td>
<td>0.019</td>
<td>0.029</td>
<td>0.027</td>
<td>0.02</td>
<td>364</td>
<td>0</td>
<td>3344</td>
<td>398</td>
<td>380</td>
<td>341</td>
<td>200</td>
<td>912</td>
<td>137</td>
<td></td>
</tr>
<tr>
<td>BMAG</td>
<td>0.000</td>
<td>0.031</td>
<td>0.013</td>
<td>0.006</td>
<td>0.026</td>
<td>0.013</td>
<td>0.007</td>
<td>0.015</td>
<td>0.000</td>
<td>344</td>
<td>398</td>
<td>380</td>
<td>341</td>
<td>200</td>
<td>912</td>
<td>137</td>
<td></td>
</tr>
<tr>
<td>LBG</td>
<td>0.000</td>
<td>0.031</td>
<td>0.013</td>
<td>0.006</td>
<td>0.026</td>
<td>0.013</td>
<td>0.007</td>
<td>0.015</td>
<td>0.000</td>
<td>344</td>
<td>398</td>
<td>380</td>
<td>341</td>
<td>200</td>
<td>912</td>
<td>137</td>
<td></td>
</tr>
<tr>
<td>FEAG</td>
<td>0.002</td>
<td>0.034</td>
<td>0.015</td>
<td>0.009</td>
<td>0.029</td>
<td>0.012</td>
<td>0.009</td>
<td>0.018</td>
<td>0.002</td>
<td>0.002</td>
<td>3389</td>
<td>3333</td>
<td>3336</td>
<td>3217</td>
<td>3003</td>
<td>3225</td>
<td></td>
</tr>
<tr>
<td>WAB</td>
<td>0.015</td>
<td>0.031</td>
<td>0.023</td>
<td>0.019</td>
<td>0.029</td>
<td>0.027</td>
<td>0.02</td>
<td>0.023</td>
<td>0.015</td>
<td>0.015</td>
<td>0.018</td>
<td>85</td>
<td>100</td>
<td>209</td>
<td>996</td>
<td>274</td>
<td></td>
</tr>
<tr>
<td>RBAG</td>
<td>0.000</td>
<td>0.031</td>
<td>0.013</td>
<td>0.006</td>
<td>0.026</td>
<td>0.013</td>
<td>0.007</td>
<td>0.015</td>
<td>0.000</td>
<td>0.000</td>
<td>0.002</td>
<td>0.015</td>
<td>39</td>
<td>192</td>
<td>975</td>
<td>257</td>
<td></td>
</tr>
<tr>
<td>WMG</td>
<td>0.008</td>
<td>0.031</td>
<td>0.018</td>
<td>0.012</td>
<td>0.027</td>
<td>0.02</td>
<td>0.013</td>
<td>0.019</td>
<td>0.008</td>
<td>0.008</td>
<td>0.01</td>
<td>0.019</td>
<td>0.008</td>
<td>177</td>
<td>940</td>
<td>241</td>
<td></td>
</tr>
<tr>
<td>SMMAG</td>
<td>0.000</td>
<td>0.031</td>
<td>0.013</td>
<td>0.006</td>
<td>0.026</td>
<td>0.013</td>
<td>0.007</td>
<td>0.015</td>
<td>0.000</td>
<td>0.000</td>
<td>0.002</td>
<td>0.015</td>
<td>0.008</td>
<td>0.018</td>
<td>813</td>
<td>74</td>
<td></td>
</tr>
<tr>
<td>FGAG</td>
<td>0.001</td>
<td>0.032</td>
<td>0.013</td>
<td>0.007</td>
<td>0.027</td>
<td>0.013</td>
<td>0.008</td>
<td>0.016</td>
<td>0.001</td>
<td>0.001</td>
<td>0.003</td>
<td>0.016</td>
<td>0.001</td>
<td>0.008</td>
<td>0.001</td>
<td>819</td>
<td></td>
</tr>
<tr>
<td>CPAG</td>
<td>0.000</td>
<td>0.031</td>
<td>0.013</td>
<td>0.006</td>
<td>0.026</td>
<td>0.013</td>
<td>0.007</td>
<td>0.015</td>
<td>0.000</td>
<td>0.000</td>
<td>0.002</td>
<td>0.015</td>
<td>0.000</td>
<td>0.008</td>
<td>0.000</td>
<td>0.001</td>
<td></td>
</tr>
</tbody>
</table>

**Table 2.5.** Below diagonal, average Tamura Nei pairwise divergences (%) between sampling locations for *A. gibbosa* mtDNA sequences (calculated in Mega 3). Above diagonal, approximate coastal distances (kilometres) between sampling locations.
The AMOVA result, in addition to the median joining network and average Tamura Nei pairwise divergence estimate, surprisingly indicates that there was no significant pattern of separation between individual populations or even between populations in the Atlantic and Mediterranean basins.

2.4.3 Haplogroup Analysis

Fig. 2.5. Geographical positions of sampling localities (see Table 2.1) and distribution of *A. gibbosa* haplogroups. The size of the circle corresponds to the number of individuals sampled from a population, with each slice corresponding to a haplogroup. The blue slices correspond to haplogroup 1 and the red slices correspond to haplogroup 2.

The MJN and NJ data shows that there is some phylogenetic structure, with two distinct haplogroups.
Haplogroup 1 contains individuals from all 17 populations sampled, across the entire geographical range, whereas haplogroup 2 contains individuals from ten populations, on both the north and south coasts of the South West England peninsula, Southern Ireland and the Mediterranean, spanning 3832 kilometres. Of the 157 individuals sampled, 131 were assigned to haplogroup 1 whereas just 20 were assigned to haplogroup 2. An equal number of haplotypes were assigned to both haplogroups despite only 21 individuals being placed in haplogroup 2 compared to 143 individuals in haplogroup 1.

Differences between the two haplogroups can also be seen in the nucleotide and haplotype diversities, both were two to three times greater in Haplogroup 2 compared to Haplogroup 1 (Table 2.6). The mean number of pairwise differences was three times as great for Haplogroup 2 compared to Haplogroup 1.

The distribution of individuals within each haplogroup can be seen in Fig. 2.5. Although both haplogroups are present in both the Atlantic and Mediterranean, their frequencies vary. Haplogroup 1 is the dominant haplogroup in both the Atlantic and Mediterranean.
2.4.4 Population Demography

The mismatch distributions (Fig 2.6) were consistent with a demographic expansion at some time in the past for haplogroup 1, showing a negative binomial curve, reflecting that most individuals in those respective groups possessed the same abundant haplotype (AG1 for haplogroup 1 and AG2 for haplogroup 2). Haplogroup 2 also showed a small secondary peak created by the presence of a small number of relatively divergent haplotypes.

![Mismatch distribution graphs for AG Haplogroups 1 and 2](image)

**Fig. 2.6.** Mismatch distribution showing frequencies of the two haplogroups found in *A. phylactica*. The red curve is the expected distribution under the sudden expansion model.
The neutrality tests (Table 2.3) of the populations grouped by geographical location are all non-significant. When the individuals are grouped according to haplogroup some of the neutrality tests are statistically significant (Table 2.6). For haplogroup 1, the Tajima’s D result is very slightly negative (-1.668) but statistically significant, the Fu’s F result was positive but was not statistically significant at the 95% level, but it was significant at the 90% level (P = 0.07), this result could be indicative of range expansion. In haplogroup 2, only the Fu’s F neutrality test was statistically significant at the 95% level all of the other tests were non-significant. This may be the result of a recent population expansion; alternatively, the lack of significant results may be caused by the small sample size. The time since expansion for haplogroup 1 was estimated at approximately 670–722 kya (τ = 3, SD = 0.004), there was no evidence of a demographic expansion for haplogroup 2 (τau = 0), although this might be the result of the small sample size.

Comparison with A. phylactica

It is worth noting that both A. gibbosa and A. phylactica share two haplotypes – AG1 and AG2 are both present in both species. This will be explored in more detail in Chapter 4.

2.5 Discussion

The results derived from the study of mtDNA in A. gibbosa show a contrasting pattern to that found with the AFLP markers by Baus et al. (2005). The largest difference observed between the two studies corresponds to the level of genetic structure seen both between populations which are geographically close, but also to the hydro-geographical barrier at the Strait of Gibraltar.

The geographical history of the study area has been extremely volatile over the past 6 million years, for example, in the area of greatest concentration of populations – the South coast of Devon and Cornwall there have been a number of climatic fluctuations. The ‘Younger Dryas’ cold phase (12.9–11.5 kya) allowed the English Channel to be partly open to the sea in the west (Lambeck, 1997) and where
continuous permafrost was located above 55°N (Renssen and Vandenberghe, 2003), the North Sea, and English Channel were also isolated from each other, this isolation ended after the catastrophic opening of the Dover Straights in the late Pleistocene (some 10 kya) when a great volume of water spilled out of the North Sea across the emerged eastern and central portions of the Channel (Smith, 1989). Major events such as these can lead to shifts in a species range, extinctions, separation of populations, range expansion and secondary contact, all of which will be imprinted into the DNA of extant species. It is possible that events like this within the species range provided *A. gibbosa* with temporary refugia during the LGM. In these refuges and the Mediterranean local diversification would have allowed the evolution of the two haplogroups from which secondary contact, intermingling and dispersal could have occurred.

The pairwise \( \theta^B \) values, AMOVA, PCA analysis and re-allocation test carried out in the study performed by Baus *et al.* (2005) all concurred and showed high levels of genetic differentiation between the Atlantic and the Mediterranean populations, with private alleles being found for the Atlantic (7 loci) and the Mediterranean (5 loci) populations. These results strongly suggest that the Strait of Gibraltar represents a major barrier to dispersal for *A. gibbosa*. However, the mtDNA AMOVA result indicated that there was little genetic structure of *A. gibbosa* between the Mediterranean and the Atlantic basins, even though there are both private haplotypes and different frequencies of common haplotypes in both the regions. This could be the result of the frequency and widespread occurrence of the most common haplotype (AG1) (Fordyce and Nice, 2003) and/or the small sample within the Mediterranean where average sample size was 9.3 individuals per population (Wang *et al*., 2009). Therefore, the original hypothesis that mtDNA would provide support for the AFLP data is rejected.

The AFLP data showed that *Asterina gibbosa* showed high levels of genetic diversity, for example, each of the 159 individuals tested except one pair displayed a unique AFLP band pattern (Baus *et al.*, 2005). Assignment tests suggested that there is gene flow occurring among populations within the Atlantic, with individuals being reallocated to different, but geographically close populations. Baus *et al.* (2005) suggested that there was a strong correlation between genetic differentiation and
geographical distance and an isolation-by-distance differentiation pattern within the
Atlantic is plausible. The lack of intra- and inter-population differentiation observed
within the mtDNA marker could be the consequence of historical events. If there
were extinctions in the Atlantic the observed lack of differentiation might be the result
of a recolonisation of the Atlantic from the Mediterranean. The AFLP analysis is likely
to be data-rich and therefore sensitive enough to detect this signature of population
structure, whereas the mtDNA analysis is not, i.e. there has not been enough time
for the populations to diverge again.

The Mediterranean populations exhibited less evidence for gene flow than the
Atlantic counterparts according to the AFLP analysis. There were particularly high
levels of differentiation between the Italian and French Mediterranean populations of
A. gibbosa. This was not the case for the mtDNA where 90% of the total genetic
variation was attributed to diversity within populations rather than differentiation
between populations. This statistic is a little surprising as there are three haplotypes
found in the Naples population not found with either of the other Mediterranean
populations.

Of the six mitochondrial haplotypes found in the Mediterranean, two correspond to
the most common haplotypes within the range, one is shared with a single individual
found at Lough Hyne, one is shared with a single individual from Ladram Bay, south
Devon and two haplotypes are private to the Mediterranean. Of the eight haplotypes
found in the Atlantic, four are private to the Atlantic and four are shared as discussed
above. Of the four private Atlantic haplotypes three are found in single individuals
and one is found in single individuals in two populations, one at Guéthary the other
at Gara Point, 851 kilometres apart.

The median joining network shows two clear haplogroups, however these
haplogroups are not geographically isolated as both occur throughout the range of A.
gibbosa. This suggests a pattern of secondary contact and large amounts of historic
gene flow between populations or a population expansion. The AFLP data suggests
that although there is some gene flow between populations it does not appear to be
common, therefore the gene flow and population expansion is believed to be
historical.
With the mtDNA sequence data, nucleotide diversity was low whereas haplotype diversity was high. Other studies of echinoderms have obtained similar results for sequences of the COI region (e.g., Zulliger et al., 2009, McCartney et al., 2000; Uthicke and Benzie, 2003; Duran et al., 2004a). It has been documented that marine invertebrates with large population sizes can retain numerous haplotypes during periods of population growth or expansion (Watterson, 1984). High haplotype diversity and low nucleotide diversity could therefore be the result of a rapid population expansion as new mutations are retained (Avise et al., 1984; Watterson, 1984).

The Atlantic populations exhibit lower haplotype diversity than the Mediterranean populations. A general trend of declining genetic diversity with latitude has been observed in northern temperate species as a result of founder events during postglacial expansions (Hewitt, 2000; Adams et al., 2006; Muhlin and Brawley, 2009). Older populations have more time to accumulate patterns of regional migration–drift equilibrium resulting in greater genetic diversity than the younger northern populations.

Although *A. gibbosa* does not have a highly dispersive larval stage long range dispersal may be possible through either adult or juvenile passively drifting or rafting on natural and anthropogenic floating material. Drifting or rafting has been inferred as an important dispersal mechanism for many littoral marine invertebrates (Ingólfsson, 1995; Ó Foighil et al., 1999, 2001; Castilla and Guíñez, 2000). Organisms have been found rafting on a wide variety of different substrata of natural (wood, seagrasses, macroalgae, volcanic pumice, corals) and anthropogenic (plastics, tar balls, manufactured wood) origin.

It may be plausible to infer from the mtDNA data that the two haplogroups seen in the MJN are actually two different taxa, which may have originated from a common ancestor in allopatry followed by secondary contact and interbreeding. The time since expansion for the two haplogroups are very different, with haplogroup 1 expanding approximately 56,000 – 63,000 years ago and haplogroup 2 not showing a signature of a population expansion. However, the AFLP data do not support the
idea of two distinct species spanning the Atlantic and Mediterranean as no genetic signals were observed which could lead to this conclusion. These results provide yet another example of why multi-locus studies should be used to make inferences about the phylogeny and phylogeography of species, as individual markers can give misleading results.

2.6 References


Chapter 3

Population genetics of an enigmatic sea star *Asterina phylactica*
3.1 Abstract

In this study we analysed mitochondrial DNA (mtDNA) and Amplified Fragment Length Polymorphism (AFLP) variation in the brooding sea star *Asterina phylactica* collected from populations in the North East Atlantic Ocean and Mediterranean Sea, in order to verify if any population structuring exists over both small and large geographical scales, across well known phylogeographic boundaries. The AFLP analysis showed that the Mediterranean and Atlantic populations were deeply divergent, however limited sampling in the Mediterranean means that the reason for this cannot be ascertained, therefore it could be the result of one or more of the following: the barrier at the Strait of Sicily or the Strait of Gibraltar; the low tidal range and current system of the Mediterranean; or, isolation-by-distance. MtDNA revealed a single common haplotype found in both basins, with all of the other haplotypes found in only a single basin. Two of the haplotypes are shared with the congeneric species *Asterina gibbosa*. Two mtDNA haplogroups were identified, both of which exhibited a pattern of a recent range expansion. In addition to the divergence between the Atlantic and Mediterranean, the Hartland Quay population on the coast of North Devon shows signs of divergence and high phylogenetic diversity – four haplotypes are found here, two of which are private, including the haplotype found in the greatest frequency indicating that this population may have survived the last glacial maxim in a northern refugium, or may be a point of contact between divergent populations.

3.2 Introduction

In the marine realm there are no obvious physical barriers to prevent gene flow between populations – it is expected that species with pelagic eggs or larvae, or species with highly mobile adults would show little sign of population structuring, whereas less mobile species or species which have sedentary larvae or juveniles should show signs of substantial population structure with distinct populations on even very small geographical scales (e.g. Lee and Boulding, 2009; Baus *et al.*, 2005; Doherty *et al.*, 1995; Planes, 1998).
The main mechanism proposed for dispersal for a species with an entirely benthic life history is adult or juvenile rafting on wood or macro algae (Mortensen, 1933; Clark, 1992; Hart et al., 1997; Waters and Roy, 2004). It has been hypothesized that organisms can passively disperse for many months in the open ocean (Waters and Roy, 2004), therefore it might be expected that there is a pattern of sporadic unidirectional gene flow following the major oceanic currents, showing an isolation-by-distance model, with highly genetically structured populations. Within the range of A. phylactica, a scarce and patchily distributed sea star, strong oceanic currents are capable of transporting passive individuals uni-directionally north from the Bay of Biscay, northward along the western coasts of the British Isles towards the Norwegian Trench. Water also travels eastward from the Atlantic Ocean along the English Channel towards the North Sea (Gysels et al., 2004).

Modern day population structure is not only the result of the life history of a species or past geographical events, but the result of many biotic and abiotic processes in the marine environment, including anthropogenic processes which can cause significant changes to populations (Palstra and Ruzzante, 2008). One such example which had a devastating effect on A. phylactica was the Sea Empress oil spill which almost eliminated the ‘type’ population at West Angle Bay, Pembrokeshire, in 1996 (Sea Empress Evaluation Committee Final Report, 1998).

Population genetics, along with its allied discipline, phylogeography (Avise, 2000) is concerned with the principles and processes governing the genetic structure of populations and their genealogical lineages. Population structure through space and time is shaped by mechanisms that might favour local differentiation including historical vicariance, habitat discontinuity, larval behaviour, marine currents and local adaptation (Baus et al., 2005). It is these intrinsic and extrinsic processes, in combination with present day environmental patterns that are the focus of genetic studies of marine populations. However, using only a unipaternally inherited single locus marker, such as mtDNA, can give a limited, incomplete or misleading view of population history (Pamilo and Nei, 1988; Palumbi and Baker, 1994; Hare and Avise, 1998; Alves et al., 2006), therefore, in the present study a multi-locus approach using both mtDNA and AFLP’s was taken to make reliable inferences about the
genetic signals underlying the disjunct population structure and phylogeography of *A. phylactica*.

*A. phylactica* (Crump and Emson, 1983) is a small sea star, which can be found in rock pools and in the sublittoral environment on the coast of the North East Atlantic Ocean and Mediterranean Sea. *A. phylactica* is a congeneric species to the cushion star *A. gibbosa*, and was only identified as being a separate species in 1979 (Emson and Crump, 1979) whereas previously it was misidentified as a juvenile form of *A. gibbosa*. *A. phylactica* may be readily distinguished from *A. gibbosa* in the field by its distinctive colour pattern (Emson and Crump, 1979) and reproductive biology. In contrast to *A. gibbosa*, *A. phylactica* protects its developing offspring with adults adopting a humped posture over the egg mass and remains in contact with its offspring until the juveniles have metamorphosed (Crump and Emson, 1983). *A. phylactica* individuals have a strong tendency to form aggregations up to 2 weeks before egg laying and to remain in the aggregations until breeding is complete (Crump and Emson, 1983). *A. phylactica* spawns directly developing benthic lecithotrophic larvae, and they therefore have an entirely benthic life history. It occupies a disjunct distribution throughout the Mediterranean Sea and the North East Atlantic.

Species of marine invertebrates which brood their young are expected to have a high potential for successful range expansion as colonisation of a new population can be achieved with a clutch of brooded young (Higgs *et al.*, 2009). Brooding marine invertebrates tend to be smaller than closely related species which do not brood their young which is advantageous for dispersal, as small animals are more likely to live on potential rafting substrata prior to it becoming detached (Cheetham, 1960; Highsmith, 1985).

An understanding of mechanisms that have lead to the formation of contemporary populations of marine invertebrates, especially scarce species such as *A. phylactica*, is of vital importance for the planning of conservation strategies which enable the creation of marine protected areas and parks. As the potential for dispersal according to life history strategy often does not necessarily correlate with the amount of dispersal and gene flow that actually occurs, an understanding of the genetic
structure of a species such as *A. phylactica* is important, as it can assist in determining the distinctiveness of individual populations and therefore aid in setting priorities for management and conservation programmes (Moritz, 1994).

This study therefore sets out to test the hypotheses that due to its life history strategy, strong genetic differentiation exists between populations in the Mediterranean Sea and the Atlantic Ocean; that due to its low putative dispersal capability, strong localised population structure occurs throughout the range of *A. phylactica*; and that it would be possible to detect lower genetic diversity in populations in the Atlantic compared with the Mediterranean, where it is thought to have survived during the last glacial maximum. These hypotheses will allow both a contextualisation of the populations where the species is most under threat and conservation prioritisation of those littoral populations that contain the most unique and diverse gene-pools.

### 3.3 Materials and methods

#### 3.3.1 Sampling

*A. phylactica* samples were collected under boulders and stones of rocky sea shores and tide pools, and from within *Corallina* algae at eight locations across its range: seven sites in the British Isles and one site in continental Europe: Rovinj (Croatia, Adriatic coast). Unfortunately it was not possible to locate and access samples from additional sites within the Mediterranean. See Table 3.1 and Fig. 3.1 for details. All specimens were stored for up to five days in absolute ethanol at ambient temperature and then for three days up to a three months at -80°C.
DNA Extraction

DNA extractions were performed on 2–4 mm³ of tissue (arm tip) using a DNeasy extraction kit (QIAGEN, catalogue # 69506) or by phenol-chloroform and CTAB purification following the method of Arndt et al., (1996). DNA quality and quantity was assessed by electrophoresing samples on 1% agarose gels alongside dilutions of phage lambda DNA (Promega, catalogue # D1501). The samples were subsequently stored at -20°C.

3.3.2 mtDNA Laboratory Procedures

The amplification, sequencing and alignment of the mitochondrial genome that contains five transfer RNA genes, a short non-coding intron sequence and the 5' end of the COI gene (Hart et al., 1997) was performed as in Chapter 2.
<table>
<thead>
<tr>
<th>Basin</th>
<th>Population</th>
<th>Latitude</th>
<th>Longitude</th>
<th>Abbreviation</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Atlantic</td>
<td>Chapel Point, S. Cornwall, UK</td>
<td>50°15'N</td>
<td>4°46'W</td>
<td>CPAP</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>23</td>
</tr>
<tr>
<td>2 Atlantic</td>
<td>Gara Point, S. Devon, UK</td>
<td>50°18'N</td>
<td>4°04'W</td>
<td>GPAP</td>
<td>21</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>3 Atlantic</td>
<td>Hartland Quay, N. Devon, UK</td>
<td>50°59'N</td>
<td>4°32'W</td>
<td>HQAP</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>34</td>
</tr>
<tr>
<td>4 Atlantic</td>
<td>Lizard Point, S. Cornwall, UK</td>
<td>49°57'N</td>
<td>5°12'W</td>
<td>LPAP</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>5 Atlantic</td>
<td>Port Erin, Isle of Man</td>
<td>54°05'N</td>
<td>4°45'W</td>
<td>IOMP</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>6 Atlantic</td>
<td>Prawle Point, S. Devon, UK</td>
<td>50°13'N</td>
<td>3°40'W</td>
<td>PPAP</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>28</td>
</tr>
<tr>
<td>7 Atlantic</td>
<td>Rosemullion Head, S. Cornwall, UK</td>
<td>50°06'N</td>
<td>5°05'W</td>
<td>RMAP</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>8 Mediterranean</td>
<td>Rovinj, Istrian Peninsula, Croatia</td>
<td>45°03'N</td>
<td>13°39'E</td>
<td>RAP</td>
<td>6</td>
</tr>
</tbody>
</table>

Table 3.1. Sampling information for *A. phylactica*. Characteristics of the samples collected from eight populations of *A. phylactica*. For each sample site, the number of individuals analysed (*n*) with both mtDNA and AFLP markers.
3.3.3 AFLP Laboratory Procedures

The AFLP procedure was performed following the methods published by Baus et al. (2005), which was based upon the approach published by Ajmone-Marsan et al. (1997) that uses one pre-selective primer combination and five selective primer combinations.

Firstly, 200 ng of DNA was restricted to ensure complete digestion, instead of 400 ng (Ajmone-Marsan et al., 1997). Secondly, three independent pre-selective PCR reactions were performed on the same digested DNA. The products from one of these reactions were visualised on agarose gel to check that the digestion had proceeded efficiently and produced a consistent smear with no banding present. If bands were present in the smear, the sample was discarded. To maximise the probability of amplifying all restriction fragments produced by the digestion reaction, the PCR products of the three pre-selective PCRs were combined, diluted and used in the selective PCR reactions. To check and maintain consistent results, positive and negative control samples were incorporated into every procedure from the selective PCR stage onwards and comparisons between plates were made to assess repeatability.

Digestion of genomic DNA and ligation of adaptors

The digestion reaction contained 2.5 μl of One-Phor-All Buffer PLUS; 1.25 μl of DTT (100 mM); 3.1 μl of BSA (0.4 mg/ml); 0.5 μl of Taq (10 U/μl); 200 ng of DNA and adjusted to 25 μl final volume with ddH2O. This was incubated for 1 h at 65°C before the following solution was added: 10.52 μl ddH2O; 1.5 μl of One-Phor-All Buffer PLUS; 0.75 μl of DTT (100 mM); 1.88 μl of BSA (0.4 mg/ml); 0.34 μl of EcoR1 at 15 U/μl. This 40 μl reaction volume was incubated for a further 1 h at 37°C immediately prior to ligation of adaptors. The ligation reaction contained: 4.15 μl of ddH2O; 1 μl of EcoRI adaptors stock (5 pmol/μl; 2.5 pmol/μl of each EcoRI adaptor); 1 μl of Taq adaptors stock (50 pmol/μl) 25 pmol/μl of each TaqI adaptor; 0.1 μl of ATP (100 mM); 1 μl of One-Phor-All Buffer PLUS; 0.5 μl of DTT (100 mM); 1.25 μl of BSA (0.4 mg/μl) and 1 μl of DNA ligase (1 /μl) total volume 10 μl. This mix was added to the 40 μl of digested DNA and incubated at 37°C for 3 h. This mix was then diluted 1:10

103
with a low TE buffer (1 ml Tris-HCl 1M pH 7.5, 20 ml 0.5 EDTA pH 8.0, up to 100 ml with ddH₂O) 1:100.

**Pre-selective PCR**

The reactions contained: 19.3 µl of ddH₂O; 5 µl of Invitrogen Taq polymerase buffer; 1.5 µl MgCl₂ (50 mM); 4 µl dNTPs (10 mM); 7.5 µl of E01 (10 ng/µl) (Pre-selective EcoRI primer E01 5'...GAC TGC GTA CCA ATT CA...3'); 7.5 µl of T01 or T02 (10 ng/µl) (Pre-selective Taq primer T02 5'...GAT GAG TCC TGA CCG AC…3'); 0.2 µl of Invitrogen Taq polymerase 5 u/µl) and 5 µl of diluted template DNA in a total volume of 50 µl (Ajmone-Marsan et al., 1997). PCRs were performed in a Perkin Elmer thermal cycler at 72°C for 1 - 2 min for 1 cycle; 94°C for 30 s, 56°C for 1 min and 72°C for 1 min for 30 cycles; 72°C for 10 min for 1 cycle. The PCR products were then checked on a 1.25% agarose gel to see a smear between 100 and 1000 bp. The pre-amplified template was then diluted 20 fold with a low TE buffer.

**Selective PCR and Primer combinations**

Reactions contained: 7.3 µl of ddH₂O; 2 µl of Invitrogen Taq polymerase Buffer; 0.6 µl MgCl₂ (50 mM); 1.6 µl dNTPs (10 mM); 0.5 µl of EcoRI primer (labelled with 6 FAM), (10 ng/µl); 3 µl of unlabelled TaqI primer (10 ng/µl); 0.08 µl of Invitrogen Taq 5 U/µl); 5 µl of the diluted pre-amplified template in a total volume of 20 µl (Ajmone-Marsan et al., 1997). The PCR conditions were as follows: Initial denaturation at 94°C for 2 min for 1 Cycle. The PCR cycle was repeated 36 times using the following conditions: denaturation at 94°C for 30 s; annealing at 65°C for 30 s (this annealing temperature was then reduced by 0.7°C each cycle to 56°C (13 cycles) and thereafter kept constant until the completion of the PCR run (a remaining 23 cycles)); extension at 72°C for 1 min. After the cycle completion, the final extension was at 72°C for 10 min for 1 cycle. Five primer combinations were used for the selective PCRs (Table 3.2, (Ajmone-Marsan et al., 1997)).
Detection of AFLP bands

1 μl of the labelled amplification product was mixed with 0.5 μl of ROX 500 size standard (Applied Biosystems) and 10 μl of formamide (Applied Biosystems). The products were denatured at 94°C for 2 min and then the fragments were separated using an ABI prism 3100 genetic analyzer. The samples were analyzed using GENESCAN vs. 2.0 (Applied Biosystems) and the size standard peaks were checked for accuracy. The fragments were visualised with GENOTyper vs. 3.6 (Applied Biosystems). The fragments for each sample were scored automatically by using a function in GENOTyper which scores all fragments between specified size and intensity ranges. The size and intensity parameters were set according to the maximum accuracy and repeatability of the control samples, and were adjusted for each primer pair. The accuracy of the AFLP method and the repeatability was assessed using the positive and negative controls on each plate.

3.3.4 Data analysis

3.3.4.1 mtDNA

The mtDNA was analysed as per Chapter 2. A brief description of the tests performed is described below.

Analysis of Genetic diversity and differentiation

Estimates of nucleotide diversity and haplotype diversity were obtained using the program Arlequin vs. 3.11 (Excoffier et al., 2005). A median joining network (MJN) (Bandelt et al., 1999) and Neighbour Joining Tree using the Tamura Nei model of
evolution were constructed to visualise the relationship between haplotypes. The
Tamura Nei (1993) evolutionary model corrects for multiple hits taking into account
the differences in substitution rate between nucleotides, distinguishing between
transition and transversion substitution rates and the inequality of nucleotide
frequencies, but assuming equality of substitution rates among sites. Pairwise
divergences were calculated using the Tamura Nei model of evolution and an
analysis of molecular variance (AMOVA) was performed to investigate the levels of
differentiation between populations and groups of populations.

Analysis of Population Demography
To infer the demographic history of populations, mismatch distributions of pairwise
nucleotide differences among mtDNA haplotypes were compared with expectations
of a sudden-expansion model (Rogers, 1995) using Arlequin (vs. 3.11) (Excoffier et
al., 2005). Tajima’s D (Tajima, 1989), Fu’s Fs (Fu, 1997) and Fu and Li’s F and D
(Fu and Li, 1997), were conducted to investigate if there was a divergence from
neutrality. In the event of a demographic expansion, an approximate estimation of
time since expansion was calculated using $t = 2ut$ (Rogers and Harpending, 1992).

3.3.4.2 AFLP
Genetic diversity was evaluated as the percentage of polymorphic loci (calculated
using Aflp-surv version 1.0; Vekemans, 2002) and the number of shared multilocus
AFLP patterns (evaluated with Arlequin version 3.11; Excoffier et al., 2005).

Genetic differentiation among populations was evaluated using a hierarchical
Bayesian approach developed by Holsinger et al. (2002) that does not assume any
prior knowledge of the degree of within-population inbreeding and is therefore not
subject to the problems of traditional methods of analysis using dominant markers.
The software HICKORY, version 1.0.3 (Holsinger et al., 2002) was used to estimate
$\theta^B$, a Bayesian analogue of $F_{ST}$, across all populations and for each pairwise
combination of populations. The data were run two to three times with the default
parameters (burn-in = 50 000, number of samples = 250 000, thinning factor = 50)
using four models: a full model, a model that assumes no inbreeding within
populations ($F_{IS} = 0$ model), a model that assumes no differentiation among
populations ($\theta^B = 0$ model) and a model that does not attempt to estimate $F_{IS}$ (f-free
model). Because estimates of $F_{IS}$ derived from dominant marker data may be unreliable (Holsinger and Wallace, 2004), we used the $f$-free analysis as our preferred method to calculate estimates of $\theta^B$. The deviance information criterion (DIC) values for the $F_{IS} = 0$, $\theta^B = 0$ and full models were used to estimate how well each model fitted the data (a smaller DIC value indicates a better fit) and which model should be preferred. A difference of more than six units is required to indicate strong support for one model over another (Holsinger et al., 2002). The $F_{IS}$ value can be unreliable for estimating within-population inbreeding (Holsinger et al., 2002), therefore only the results on overall population differentiation are reported. $F_{ST}$ values across all populations were calculated using Aflp-surv version 1.0 as a comparison to the $\theta^B$ data.

A Mantel test was performed using the program TFPGA (Miller, 1997) to evaluate the geographical structure among sample sites. The matrix used both pairwise $\theta^B$ and $F_{ST}$ values, and geographical distances calculated as the shortest distance between two populations that did not involve crossing land.

To estimate population structure we conducted an analysis of molecular variance (AMOVA) using Arlequin vs. 3.11. AMOVA was used to investigate genetic differentiation between Atlantic and Mediterranean populations and to confirm the subdivided groups identified by the MJN and NJ Tree. AMOVA uses the frequencies of haplotypes and the number of mutations between them to test the significance of the variance components associated with hierarchical levels of genetic structure (within populations, among populations within groups, and among groups) by means of non-parametric permutation (Excoffier et al., 1992).

Principal coordinate analysis (PCA) based on $F_{ST}$ and $\theta^B$ distances between AFLP multilocus phenotypes was performed using Genalex version 5.1 (Peakall and Smouse, 2006). The PCA via covariance matrix with data standardization method was chosen. Finally, assignment tests were carried out using the re-allocation procedure of aflpop version 1.1 (Duchesne and Bernatchez, 2002) with the default settings (fixed correction value for zero frequencies = 0.001, minimal log-likelihood
difference to allocate specimens = 0, number of artificial genotypes to compute $P$ values = 500).

3.4 Results

3.4.1 Mitochondrial DNA
Sequences and Haplotype Analysis
A 332-bp sequence from 108 individuals from eight populations, seven in the North East Atlantic Ocean and one situated in the Mediterranean Sea, revealed 16 polymorphic sites, giving seven haplotypes (Table 3.3). The haplotypes can be seen in Appendix 1. Of these polymorphisms, four occurred at first codon positions, two occurred at second codon positions (including one in the non-coding intron section) and ten occurred at third codon positions. Of the seven haplotypes, only one haplotype (AP1) was shared across both basins, all of the other six haplotypes were exclusive either to the Mediterranean Sea or the Atlantic Ocean (Fig. 2.1). Five of these haplotypes were private to individual populations, the most notable of which was AG6. AG6 was found at a high frequency at the Hartland Quay population (16 of the 24 individuals), the other private haplotypes were sampled in smaller numbers than AG6.
Table 3.3. mtDNA haplotypes of *A. phylactica*. Data comprises: total number of haplotypes identified in each sampling location and the number of individuals found with each haplotype in each sampling location. Populations above the line are found in the Atlantic Ocean whereas populations found below the line are found in the Mediterranean. The blue haplotypes are grouped into haplogroup 1 whereas the red haplotypes are grouped into haplogroup 2.

<table>
<thead>
<tr>
<th></th>
<th>AP1</th>
<th>AP2</th>
<th>AP3</th>
<th>AP5</th>
<th>AP6</th>
<th>AP7</th>
<th>AP8</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPAP</td>
<td>13</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>14</td>
</tr>
<tr>
<td>GPAP</td>
<td>16</td>
<td>4</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>21</td>
</tr>
<tr>
<td>HQAP</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>16</td>
<td></td>
<td></td>
<td></td>
<td>24</td>
</tr>
<tr>
<td>LPAP</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>10</td>
</tr>
<tr>
<td>IOMP</td>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>8</td>
</tr>
<tr>
<td>PAP</td>
<td>11</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>16</td>
</tr>
<tr>
<td>RMP</td>
<td>8</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>9</td>
</tr>
<tr>
<td>RAP</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td>3</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>69</td>
<td>16</td>
<td>1</td>
<td>1</td>
<td>16</td>
<td>2</td>
<td>3</td>
<td></td>
</tr>
</tbody>
</table>
Fig. 3.2. Geographical positions of sampling localities (see Table 3.1) and distribution of A. phylactica haplotypes. The size of the circle corresponds to the number of individuals sampled from a population. Each colour corresponds to a different haplotype.
Fig. 3.3. Median Joining Network (MJN) for *A. phylactica* mtDNA haplotypes, indicating the haplotype relationships. Circles represent haplotypes, circle area represents number of individuals within that haplotype. Each cross-line is indicative of a site which has undergone a substitution. Blue = Individuals from Atlantic Populations, Red = Individuals from Mediterranean Populations.
Fig. 3.4. Neighbour joining phenogram tree based on Tamura Nei genetic distances for *A. phylactica* mtDNA haplotypes, with bootstrap values on the branches. The tree is rooted by *Patiriella exigua* (Genbank No. U50053 (Hart *et al.*, 1997)).
The Median Joining Network (Fig. 3.3) shows that there are two divergent clades (haplogroups) separated by ten nucleotide substitutions. Haplogroup 1 (the most frequently observed haplogroup) contains individuals from both the Atlantic and Mediterranean, haplogroup 2 has individuals from the Atlantic only. Phylogenetic reconstruction of the 7 haplotypes (plus an outgroup of *P. exigua*) revealed two main groups with NJ bootstrap support of 100% (Fig. 3.4), this supported the data seen in the MJN. The Tamura Nei pairwise distances between haplotypes can be seen in Table 3.4. Two haplogroups can be clearly seen with these values.

<table>
<thead>
<tr>
<th></th>
<th>AP01</th>
<th>AP02</th>
<th>AP03</th>
<th>AP05</th>
<th>AP06</th>
<th>AP07</th>
<th>AP08</th>
</tr>
</thead>
<tbody>
<tr>
<td>AP01</td>
<td>0.048</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AP02</td>
<td>0.048</td>
<td>0.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AP03</td>
<td>0.048</td>
<td>0.001</td>
<td>0.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AP05</td>
<td>0.048</td>
<td>0.001</td>
<td>0.001</td>
<td>0.048</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AP06</td>
<td>0.001</td>
<td>0.048</td>
<td>0.048</td>
<td>0.048</td>
<td>0.048</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AP07</td>
<td>0.005</td>
<td>0.043</td>
<td>0.043</td>
<td>0.043</td>
<td>0.043</td>
<td>0.005</td>
<td></td>
</tr>
<tr>
<td>AP08</td>
<td>0.009</td>
<td>0.038</td>
<td>0.038</td>
<td>0.038</td>
<td>0.038</td>
<td>0.009</td>
<td>0.005</td>
</tr>
</tbody>
</table>

**Table 3.4.** Tamura Nei pairwise divergences (%) between mtDNA haplotypes for *A. gibbosa* and *A. phylactica*.

<table>
<thead>
<tr>
<th>Sample Size</th>
<th>Atlantic</th>
<th>Mediterranean</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>102</td>
</tr>
<tr>
<td></td>
<td>No. of Populations</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>No. of Haplotypes</td>
<td>5</td>
</tr>
<tr>
<td>Nucleotide Diversity</td>
<td>0.012 +/- 0.007</td>
<td>0.002 +/- 0.003</td>
</tr>
<tr>
<td>Haplotype (gene) Diversity</td>
<td>0.511 +/- 0.049</td>
<td>0.933 +/- 0.122</td>
</tr>
<tr>
<td>Mean no. pairwise differences</td>
<td>4.090 +/- 2.055</td>
<td>0.333 +/- 0.380</td>
</tr>
<tr>
<td>Fu’s Fs Statistic</td>
<td>8.302 (P &gt;0.10)</td>
<td>-7.143* (P &lt;0.05)</td>
</tr>
<tr>
<td>Fu and Li’s F</td>
<td>1.011 (P &gt;0.10)</td>
<td>-0.965 (P &gt;0.10)</td>
</tr>
<tr>
<td>Fu and Li’s D</td>
<td>0.825 (P &gt;0.10)</td>
<td>-0.950 (P &gt;0.10)</td>
</tr>
<tr>
<td>Tajima's D</td>
<td>0.917 (P &gt;0.10)</td>
<td>-0.933 (P &gt;0.10)</td>
</tr>
</tbody>
</table>

**Table 3.5.** Comparative summary statistics for mtDNA diversity within *A. phylactica* within both the Atlantic and Mediterranean basins.
There were five haplotypes found in the Atlantic and three haplotypes found in the Mediterranean. As with *A. gibbosa*, the number of haplotypes observed in the Mediterranean is high considering the small sample size, in this case just six individuals. Nucleotide diversity is marginally lower in the Mediterranean than in the Atlantic however the haplotype diversity is almost double (Table 3.5). In contrast to this, the mean number of pairwise differences was far greater in the Atlantic than the Mediterranean.

Within both basins there was high haplotype diversity and low nucleotide diversity, which indicates that a population expansion event or a selective sweep may have occurred. The Atlantic population's neutrality tests gave non-significant results indicating that there is a stable population structure in this basin (Table 3.5). In contrast to this, the Mediterranean population had a large negative and statistically significant *Fu’s F* value, which indicates population expansion, although as with the *A. gibbosa* results in Chapter 2, all of the other neutrality tests were not significant.

An AMOVA was performed (Table 3.6) to test for population structure between the Mediterranean and Atlantic, as well as to test for population structure within the Atlantic Ocean. When no regional groups were tested the large majority of the variation (92.00% (*P*<0.05)) was partitioned within populations. When the Atlantic populations were analysed alone, within the Atlantic 93% (*P*<0.05) was attributed to diversity within populations, with just 7% (*P*<0.05) being attributed to differentiation between populations. Finally, when the range was split into the Atlantic Ocean and Mediterranean Sea an extremely large proportion of the variation was partitioned within populations (93.03% (*P*<0.05)).
Table 3.6. ΦST AMOVA for *A. phylactica* samples indicating the groups tested, and showing the percentage of genetic variation accounted for when the data is divided up into different groups.

Haplogroup Analysis

There were two distinct haplogroups observed, which when analysed give different results (Table 3.7) to those seen when analysing the data as populations or groups of populations (Table 3.5). As with *A. gibbosa* the majority of individuals are assigned to the largest haplotype (AP1), which is placed into haplogroup 1.

Table 3.7. Summary statistics for *A. phylactica* when the mtDNA haplotypes are classified into haplogroups.
Of the 108 individuals sampled, 90 were assigned to haplogroup 1 and 18 were assigned to haplogroup 2. Haplogroup 1 contains the largest haplotype (AP1), which contains individuals from all of the populations sampled, the high frequency private haplotype found at Hartland Quay, north Devon and all of the Mediterranean individuals. Haplogroup 2 is localised to the South West peninsular of England with all of the individuals sampled from populations on the south coast of Devon and Cornwall and Hartland Quay (Fig. 3.5).

The haplotype diversity was much greater than the nucleotide diversity (Table 3.7) within both haplogroups.
Fig. 3.5. Geographical positions of sampling localities (see Table 3.1) and distribution of *A. phylactica* haplogroups. The size of the circle corresponds to the number of individuals sampled from a population, with each slice corresponding to a haplogroup. The blue slices correspond to haplogroup 1 and the red slices correspond to haplogroup 2.

**Population Demography**

The mismatch distributions show that both haplogroup 1 and haplogroup 2 have a negative binomial curve, consistent with a demographic expansion at some time in the past (Fig. 3.6). As with *A. gibbosa* in Chapter 2, both haplogroups display a large peak at the origin, reflecting that most individuals in those respective groups
possessed the same abundant haplotypes (AP1 for haplogroup 1 and AP2 for haplogroup 2).

![Mismatch distribution showing frequencies of the two haplogroups found in A. phylactica. The red curve is the expected distribution under the sudden expansion model.](image)

The neutrality tests (Table 3.5) of the populations grouped by geographical location are non-significant with the exception of the Fu’s $F$ value which in the Mediterranean had a large negative statistically significant value. When the individuals were
grouped according to haplogroup (Table 3.7) both haplogroups gave a similar result: both the Fu's Fs and Tajima's D tests gave statistically significant negative values, this coupled with the non-significant Fu and Li's values suggest that a population expansion has occurred. The time since expansion for haplogroup 1 was estimated at approximately 640,000 – 722,000 years ago (Tau mean = 3, SD = 0.005), this date is congruent with haplogroup 2 (Tau mean = 3, SD = 0.002).

*A. gibbosa*

It is worth noting that both *A. gibbosa* and *A. phylactica* share the two most frequently detected haplotypes. This will be explored in more detail in Chapter 4.

### 3.4.2 AFLP

One hundred and three specimens of *A. phylactica* sampled from four populations were selected across the species range. These specimens were analysed using five selective primer pair combinations and generated a total of 292 bands. The percentage of polymorphic bands across the 4 natural populations ranged between 24.3% (HQAP, South Devon) and 39.7% (RAP, Croatia), as detailed in Table 3.8.

<table>
<thead>
<tr>
<th>Basin</th>
<th>Country</th>
<th>Population</th>
<th>Abbreviation</th>
<th>Latitude</th>
<th>Longitude</th>
<th>n</th>
<th>%P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atlantic</td>
<td>UK</td>
<td>Chapel Point, S. Cornwall</td>
<td>CPAP</td>
<td>50°15'N</td>
<td>4°46'W</td>
<td>23</td>
<td>25.3</td>
</tr>
<tr>
<td>Atlantic</td>
<td>UK</td>
<td>Hartland Quay, N. Devon</td>
<td>HQAP</td>
<td>50°59'N</td>
<td>4°32'W</td>
<td>34</td>
<td>24.3</td>
</tr>
<tr>
<td>Atlantic</td>
<td>UK</td>
<td>Prawle Point, S. Devon</td>
<td>PPAP</td>
<td>50°13'N</td>
<td>3°40'W</td>
<td>28</td>
<td>30.1</td>
</tr>
<tr>
<td>Mediterranean</td>
<td>Croatia</td>
<td>Rovinj, Iberian Peninsula</td>
<td>RAP</td>
<td>45°03'N</td>
<td>13°39'E</td>
<td>18</td>
<td>39.7</td>
</tr>
</tbody>
</table>

**Table 3.8.** Characteristics of the samples collected from four populations of *A. phylactica*. For each sample site, the number of individuals analysed (*n*) and the percentage of polymorphic AFLP loci (%P) are presented.

Mean genetic differentiation between populations estimated using the Bayesian hierarchical method of Holsinger ($\Theta^B$) was 0.319. As a comparison, the mean estimate of $F_{ST}$ obtained using Arlequin was 0.388. A matrix of both $\Theta^B$ and $F_{ST}$ values between populations is shown in Table 3.9. It shows relatively high levels of
genetic differentiation between Atlantic and Mediterranean populations in comparison to the genetic differentiation between the Atlantic populations.

The deviance information criteria (DIC) statistic provided by the Bayesian factor analysis were used as a model choice criterion between three models: the full model, the $F_{IS} = 0$ model and the $\theta^B = 0$ model. Here, the DIC values were, respectively, 2143 using the full model, 2168 for the $F_{IS} = 0$ model and 8440.52 for the $\theta^B = 0$ model. The full model was thus clearly preferred to the $\theta^B = 0$ model, supporting the existence of a significant level of differentiation among populations. However, there was only weak evidence that the full model should be preferred to the $F_{IS} = 0$ model since the difference in DIC (25 units) between the two models arises as a result of differences in model dimensions ($pD = 451$ for the $F_{IS} = 0$ model and 419 for the full model).

<table>
<thead>
<tr>
<th></th>
<th>CPAP</th>
<th>HQAP</th>
<th>PPAP</th>
<th>RAP</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPAP</td>
<td>0</td>
<td>224</td>
<td>74</td>
<td>5058</td>
</tr>
<tr>
<td>HQAP</td>
<td>0.0767 (0.069)</td>
<td>0</td>
<td>286</td>
<td>5172</td>
</tr>
<tr>
<td>PPAP</td>
<td>0.0381 (0.074)</td>
<td>0.0571 (0.016)</td>
<td>0</td>
<td>5082</td>
</tr>
<tr>
<td>RAP</td>
<td>0.576 (0.733)</td>
<td>0.5921 (0.7295)</td>
<td>0.5736 (0.7424)</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 3.9. Pairwise genetic differentiations estimated $\theta^B$ values, with $F_{ST}$ values in brackets, of four populations of *A. phylactica* (bottom left triangle). Geographical distances between populations in kilometres (top right triangle).

As shown in Table 3.10, AMOVA revealed that 54% of the total genetic variation was attributed to differences between individuals within populations ($P<0.05$), while 46% was attributed to differences among populations ($P<0.05$). When the Atlantic populations were analysed alone, within the Atlantic 93% ($P<0.05$) was attributed to diversity within populations, with just 7% ($P<0.05$) being attributed to differentiation between populations. Finally, when populations were grouped into Atlantic and Mediterranean regions, 73% of the variance was attributed to differences among these regions ($P>0.10$) and 25.5% to differences between populations within geographical regions ($P<0.05$), suggesting a high level of differentiation between these two basins at the nuclear genomic level.
<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>d.f.</th>
<th>Sum of Squares</th>
<th>Variance Components</th>
<th>Percentage of Variation</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>No Regional Groups</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Among populations</td>
<td>3</td>
<td>1041.178</td>
<td>13.275</td>
<td>54.08</td>
<td>&lt;0.05</td>
</tr>
<tr>
<td>Within populations</td>
<td>99</td>
<td>1116.123</td>
<td>11.274</td>
<td>45.92</td>
<td>&lt;0.05</td>
</tr>
<tr>
<td>Total</td>
<td>102</td>
<td>2157.301</td>
<td>24.549</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>No Regional Groups – Atlantic Ocean Only</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Among populations</td>
<td>2</td>
<td>58.034</td>
<td>0.697</td>
<td>6.83</td>
<td>&lt;0.05</td>
</tr>
<tr>
<td>Within populations</td>
<td>82</td>
<td>780.178</td>
<td>9.514</td>
<td>93.17</td>
<td>&lt;0.05</td>
</tr>
<tr>
<td>Total</td>
<td>84</td>
<td>838.212</td>
<td>10.211</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Two Groups: Atlantic/Mediterranean</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Among Groups</td>
<td>1</td>
<td>983.145</td>
<td>32.288</td>
<td>73.06</td>
<td>&gt;0.10</td>
</tr>
<tr>
<td>Among populations within groups</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>&lt;0.05</td>
</tr>
<tr>
<td>Within Populations</td>
<td>99</td>
<td>1116.123</td>
<td>11.274</td>
<td>25.51</td>
<td>&lt;0.05</td>
</tr>
<tr>
<td>Total</td>
<td>102</td>
<td>2157.301</td>
<td>44.196</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 3.10.** AMOVA, based on ΦST values between AFLP multilocus phenotypes, for 103 *A. phylactica* individuals sampled from four populations, with and without geographical structuring.

The result of the PCA analysis is shown in Fig. 3.7. The plot of the first and second principal coordinates, which accounted for 74.5% and 7.5% of the total variation, respectively, provided a visual representation of genetic similarity between populations in the Atlantic, and provides strong evidence of a split between the Atlantic and Mediterranean Sea populations in PC1 whereas PC2 is accounted for by a single individual located in the Rovinj population. This supported the results given by AMOVA. The Atlantic and Mediterranean samples clustered separately from each other. As stated above, there is a single outlier from the Mediterranean population that is situated away from both clusters. $F_{ST}$ values and not $\theta^B$ values are displayed below to show a direct comparison to the result obtained by Baus *et al.* (2005). The $\theta^B$ data gave a similar pattern to the $F_{ST}$ data displayed in Fig 3.7.
Fig. 3.7. Principal coordinate analysis (PCA), based on Pairwise genetic differentiations (estimated as $F_{ST}$ values using AFLP-SURV) between AFLP multilocus phenotypes, for 103 *A. phylactica* sampled from 4 populations.

<table>
<thead>
<tr>
<th>Allocated to:</th>
<th>PPAP</th>
<th>CPAP</th>
<th>RAP</th>
<th>HQAP</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPAP</td>
<td>18</td>
<td>7</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>CPAP</td>
<td>7</td>
<td>14</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>CRAP</td>
<td>0</td>
<td>0</td>
<td>18</td>
<td>0</td>
</tr>
<tr>
<td>HQAP</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>29</td>
</tr>
</tbody>
</table>

Table 3.11. Results of the assignment test performed using the reallocation procedure of afpop 1.1 (Duchesne and Bernatchez, 2002) for the 103 *A. phylactica* individuals sampled from four populations.
This general pattern of genetic structure was also illustrated by the assignment test carried out on all individuals (Table 3.11). In the Mediterranean basin, all individual multilocus genotypes were re-allocated to their population of origin. In the Atlantic basin, however, a high proportion of individuals were misassigned. The misassigned individuals were generally re-allocated to populations that were geographically close to their original location, for example Prawle Point and Chapel Point, both of which were located on the south coast of Devon and Cornwall, respectively. There are also approximately similar amounts of gene flow between Hartland Quay and the two south coast populations, but the level of gene flow is inferred as being relatively low with five individuals from a total of 34 from Hartland Quay being allocated to other populations and five individuals out of a total of 53 individuals from the south coast populations being re-allocated to Hartland Quay.

A Mantel test was performed in order to evaluate the correlation between genetic differentiation ($\theta^B$ value) and geographical distance (measured as the shortest distance without crossing land, in km) between populations within each basin. A correlation between genetic differentiation and geographical distance was observed within the Atlantic group ($r = 0.9995$). However, the level of significance of 5% was not reached ($P = 0.089$) due to the small number of populations analysed combined with two populations separated by 224km showed a greater degree of genetic differentiation ($\theta^B = 0.077$) than two populations separated by 286km ($\theta^B = 0.057$). A similar result was seen with the $F_{ST}$ data, with a correlation between genetic differentiation and geographic distance ($r = 0.999$) but again it was not statistically significant ($P = 0.084$).

3.5 Discussion

As different genetic markers can provide different ecological and demographic information about a species (Avise, 2004), both mtDNA sequences (tRNA's and COI) and AFLPs were used as molecular markers to determine the diversity of and connectivity among A. phylactica populations along the coast of the North East Atlantic Ocean and Mediterranean Sea.
Contemporary population structures cannot be explained by modern gene flow in isolation, past geological and climatological events must also be taken into consideration (Avise, 2004), although due to the open character of the sea with a lack of obvious permanent barriers to gene flow, the influence of historical events can be reduced by gene flow between populations (Palumbi, 1994). It can be difficult to disentangle historical signals from present day patterns of gene flow, as the signature of historical processes might obscure the current patterns of intraspecific processes, particularly with mitochondrial genes (Patarnello et al., 2007).

Little or no population differentiation was detected with the mtDNA sequence data – for AMOVA the within population variation contributed to in excess of 90% of the variation when analysed with and without geographical groupings. The observed signal is a little unexpected, as there is only a single haplotype shared between the Atlantic and Mediterranean populations, but no variation was attributed to differences between the groups. However, the differences between haplotypes is shallow, usually just a single substitution. A similar pattern was observed for the direct developers *Haustrum vinosa* and *Parvulastra exigua* (an asterinid sea star), where population differentiation was not affected by a phylogeographic barrier (Southeast Australian biogeographic barrier), with AMOVA revealing no significant differentiation between regions (Ayre et al., 2009).

Although the mtDNA AMOVA result showed no differentiation between the Atlantic and Mediterranean populations the MJN shows that there are two haplogroups present, one consists entirely of individuals from Atlantic populations, the other is made up from individuals from both basins. These divergent haplogroups exert a greater influence on the genetic analysis of the mtDNA than the effect of geography. However, the AFLP data provides a different perspective. The pairwise $\theta^B$ values (Table 3.9), AMOVA (Table 3.10), PCA (Fig. 3.7) and the re-allocation test (Table 3.11) all concur and show high levels of genetic differentiation between the Atlantic and the Mediterranean populations. These results suggest that there may be a barrier to dispersal for *A. phylactica* somewhere between the South West of England and Rovinj, on the Istrian peninsula of Croatia, the precise location of this is unknown.
Within this range there are a number of well documented barriers to gene flow, although the importance or severity of marine barriers differs from species to species. These barriers can be physical imposed by the complex shorelines but also ephemeral barriers including local upwelling, current patterns and isotherms that prevent gene flow and promote the isolation of populations. Probably the most documented barrier in this range is the Strait of Gibraltar, or more likely, the Oran-Almeria Front which acts as a barrier between the Atlantic Ocean and the Mediterranean Sea, this prevents gene flow in many species, including *A. gibbosa* (Baus et al., 2005), for a review see Patarnello et al. (2007). The English Channel has also been proposed as a phylogeographic barrier (Jolly et al., 2005) for the polychaete *Pectinaria koreni*. The Italian peninsula acts as a barrier to gene flow between the western and eastern Mediterranean basins with the Strait of Sicily and the Strait of Messina acting as barriers (e.g. *S. solea*, Kotoulas et al., 1995; *Trisopterus minutus*, Mattiangeli et al., 2003; *Psetta maxima*, Suzuki et al., 2004). In the Southern Adriatic there is a temperature barrier for some species with water temperatures getting as high as 24°C (Gysels et al., 2004), although this is unlikely to be a barrier to dispersal for organisms who’s main habitat is in rock pools, on the littoral environment.

There is a pattern of low nucleotide diversity and high haplotype diversity, which is often related to expansions after a period of small effective population size (Avise et al., 1984). The values observed for the haplotype and nucleotide diversities are similar to those of other species of seastar including the brooding brittle star *Astrotoma agassizii* (Hunter and Halanych, 2008) and the two co-distributed seastars *Linckia laevigata* and *Protoreaster nodosus* (Crandall et al., 2008).

The neutrality tests were significantly negative and the mismatch distribution displayed a negative binomial curve, all of which indicates that a recent demographic expansion has occurred. The time since expansion was estimated at 267,000 – 301,000 years ago for both haplogroups. This result indicates that there has been a shared pattern of population expansion, probably linked to a glacial/ interglacial event at around this time. During the Pleistocene, the Mediterranean and Atlantic populations would have been isolated during glacial periods. During these periods of
isolation genetic divergence can occur, followed by secondary contact during more favourable environmental conditions. The impact of glacial periods on demographic fluctuations has been remarked in different species (Stamatis et al., 2004; Bargelloni et al., 2005; Domingues et al., 2005).

For most extant species, the glaciation cycles the Pleistocene (18,000–11,500 years ago) were arguably the most important climatological events during their evolutionary life span (Hewitt, 2000). Indeed, the late Pleistocene has been reported as the epoch of population expansion for many marine organisms from disparate geographic regions (Benzie et al., 2002; Gopurenko and Hughes, 2002; Uthicke and Benzie, 2003). It is likely that the genetic footprint detected is heavily influenced by these glaciation cycles, with contemporary gene flow occurring via drifting or rafting of either juvenile or adult individuals.

The Atlantic populations exhibit lower haplotype diversity than the Mediterranean populations. A general trend of declining genetic diversity with latitude has been observed in northern temperate species as a result of founder events during expansions (Hewitt, 2000; Adams et al., 2006; Muhlin and Brawley, 2009). The older populations have more time to accumulate patterns of regional migration–drift equilibrium resulting in greater genetic diversity than the younger northern populations. There was only one haplotype individual from the Mediterranean population with a common haplotype with the Atlantic populations, this could be explained by one or more of the following: the time since divergence is ancient; there is limited gene flow between the two basins (in this case an individual would have rafted into the Mediterranean, which seems unlikely due to the hydrogeographic barriers); and/ or, there are un-sampled populations between the sets of populations (this is likely).

_A. phylactica_ showed high level of genetic diversity, with AFLP genetic diversity ranging between 24.3% at Hartland Quay, north Devon and 39.7%, at Rovinj, Croatia. The assignment test (Table 3.11) suggested that gene flow occurs among populations within the Atlantic. Indeed, a significant amount of Atlantic individuals were reallocated to a different but closely located population from the same basin, with apparent gene flow occurring between the north and south coasts of Devon and
Cornwall, the South West peninsula of England does not appear to act as a boundary to gene flow. All of the Mediterranean individuals were re-allocated to their original population in the assignment test.

In order for a pattern to population connectivity to occur as seen by both the mtDNA and AFLP markers there must be gene flow between populations. It is likely that the populations are connected by stochastic rafting events, which may allow for the rare dispersal of individuals (or groups) over both short and long geographical distances, allowing gene flow to occur between any pair of populations (Ayre et al., 2009). *A. phylactica* displays considerable geographic differentiation but only weak relationships between genetic and geographic distance, this observation is consistent with other species that are dependent upon rafting for dispersal (Waters and Roy, 2004; Thiel and Haye, 2006).

The Hartland Quay population on the north coast of Devon, provides an interesting data set. Both the mtDNA and the AFLP datasets show that this population is distinct from all of the other populations. The AFLP data suggests that there is gene flow between this and other populations on the southwest coast of England, but the amount is small. The mtDNA sequences show that the majority of individuals sampled in this population are assigned to a haplotype which is private to this population. It is possible that this population has survived the last glacial maxima in a different refugium to *A. phylactica* from different populations.

It appears that some of the results obtained from the mtDNA and AFLP analysis give discordant results. Previous studies using non-nuclear DNA and AFLP’s have also found this to be the case. Després et al. (2003) found that chloroplast DNA was unable to resolve the phylogeny of *Trollius* spp. due to the lack of intragenus nucleotide variability. The authors believed that this lack of diversity was a result of a rapid and recent radiation of the genus and/or gene flow removing previous genetic differentiation (Després et al., 2003). The AFLP data enabled the construction of a phylogenetically informative tree with high bootstrap values (Després et al., 2003). Riberon et al. (2004) found that AFLP markers gave good intra-population diversity when mtDNA data showed very little or no differentiation in *S. a. aurorae*, the low
diversity seen with the mtDNA marker was a population bottleneck or the persistence of low population sizes.

3.6 References


Mortensen T (1933) Echinoderms of southern Africa (Asteroidea and Ophiuroidea). 


Chapter 4

A comparative genetic analysis of the cryptic sea stars *Asterina gibbosa* & *Asterina phylactica* using mitochondrial and AFLP markers
4.1 Abstract

The coast of the North East Atlantic Ocean and Mediterranean Sea includes a number of complex biogeographic barriers, which may limit the dispersal of marine invertebrates and maintain genetically structured biogeographic regions. This study compares genetic diversity using mitochondrial DNA (mtDNA) sequences and Amplified Fragment Length Polymorphism's (AFLP) between two co-distributed congener sea stars that have similar life history strategies, but which differ in some components of these strategies, that are expected to impact on their population structure. Both mtDNA and AFLP data supported the evolutionary distinctiveness of A. gibbosa and A. phylactica, with significant differences observed with the AFLP data, with no evidence for breeding between the two species with the exception of a single individual which was assigned across the species barrier in a re-assignment test. This could result from introgression or the individual could have been incorrectly classified as A. phylactica at the time of sampling. There was no other evidence of introgression with the AFLP marker. There is incomplete mtDNA lineage sorting between these taxa, with the two most common haplotypes being shared between the two species. There were eight haplotypes private to A. gibbosa and five haplotypes private to A. phylactica. Within the Atlantic, the AFLP data showed that gene flow was detected between proximate populations of both species, but much more restricted gene-flow was inferred for Mediterranean A. gibbosa. The mtDNA showed there was little genetic structuring with >90% of the genetic variation situated within populations as opposed to between populations, within both species situated in both basins. Two haplogroups were common in both species and present in both basins, with the presence of a common haplotype in all populations throughout the entire 5400 km range. Both haplogroups/s species showed congruent departures from neutrality that are consistent with major range expansions, that date to within the Pleistocene epoch. There are a number of haplotype rich populations present around the coast of Devon, hinting that either populations persisted in one of more northern refugia during the last glacial maxima or that this region represents an unusually rich mixture of population genotype, possibly as the result of prevailing currents. In addition to this, the genetic data from A. phylactica individuals found at Rovinj,
Croatia implies that this population is extremely distinct, perhaps warranting separate taxonomic status.

### 4.2 Introduction

In marine environments, species with either a short or no planktonic stage are in general expected to have a small geographic range, whereas a long planktonic stage promotes a wider distribution range (Mileikovsky, 1971; Scheltena, 1978, 1986; Crisp, 1978; Jablonski, 1986). However, some of the least mobile marine invertebrates exhibit some of the largest geographic ranges (Johannesson, 1988; Sponer and Roy, 2002; Kuklinski and Barnes, 2010) presumably as a result of a variety of biotic and abiotic factors including biological, ecological, physiological, physical and geological factors. It is widely accepted that there is a lack of real understanding of the interactions of all of these factors in determining genetic structure in marine populations. There is evidence that vicariant effects of earlier (Plio-Pleistocene) patterns of oceanic circulation, physical barriers (Taylor and Hellberg, 2006; Waters, 2008; Ayre et al., 2009) and contemporary restrictions to dispersal including ephemeral hydrological and ecological barriers such as currents, temperature, salinity and behavioural responses (Ayre et al., 2009) can all affect the distribution of populations and their genetic structure. For example, historical barriers and geological features can permanently or temporarily prevent an influx of individuals and can contribute to the structuring of populations by allowing mechanisms that may favour local differentiation to occur (e.g. Palumbi, 1994; Riginos and Nachman, 2001; Bierne et al., 2003).

In addition to discovering geographically mediated genetic structure, molecular studies have led to the discovery of substantial cryptic diversity within the marine environment (e.g. Knowlton, 1993, 2000; Thorpe et al., 2000), with many morphologically similar marine organisms now believed to comprise cryptic species (Knowlton, 1993; Bickford et al., 2007). Cryptic diversity has been found across a number of groups of marine taxa, including algae (Wright et al., 2000), crustaceans (Kitaura et al., 2002), cnidarians (Dawson and Jacobs, 2001), molluscs (Collin,
Gene flow within and among populations and taxa that are not genetically isolated can arise as a result of behavioural, anthropogenic and stochastic factors. Organisms, including several echinoderm species, have been found drifting or rafting on a wide variety of different substrates including natural materials (wood, seagrasses, macroalgae, volcanic pumice, corals) and materials of anthropogenic (plastics, tar balls, manufactured wood) origin, as well as fouling in ballast (Carlton, 1985; Sponer and Roy, 2002; Waters and Roy, 2003, 2004; Thiel, 2003). Rafting has been proposed as an effective dispersal mode in marine species inhabiting isolated locations such as oceanic islands (Parker and Tunnicliffe, 1994; Morton and Britton, 2000). Littoral and shallow water species may therefore have ample opportunity for long distance dispersal as rafting agents, such as algae, flotsam and jetsam, are common in these habitats.

Within the North East Atlantic and Mediterranean Sea there are a number of potential barriers to gene flow, including the Strait of Gibraltar/ Almeria Oran Front (AOF) (dividing the Mediterranean and the Atlantic), the Strait of Sicily (separating the Eastern and Western Mediterranean) and the English Channel. The most well documented barrier is the Strait of Gibraltar which has been found to act as a barrier to gene-flow in some organisms, including one of the species under investigation in this study, *Asterina gibbosa* (Baus et al, 2005; see Patarnello et al., 2007 for review).

The geological history of a region can therefore have a profound effect on the modern day genetic structure of a species. The North East Atlantic Ocean and Mediterranean Sea have experienced a turbulent past, with a rich geological history. The Last Glacial Maximum (LGM), which occurred ca. 20 kya, within the last glacial period (110 to 20 kya) is the most recent of these and has greatly influenced present-day distributions of marine species in the North Atlantic. This event caused a shift in the range for many organisms, with many northern species displaced southward into periglacial refugia. When the glaciers retreated, ranges began to expand northwards once more (Hewitt, 1996, 2004). This pattern of range contraction and expansion occurred repeatedly as a result of the glacial episodes of
the Quaternary (2.6 mya to present) and many of these episodes are detectable within contemporary intertidal populations of invertebrate and seaweed species (Wares and Cunningham, 2001; Addison and Hart, 2005; Provan et al., 2005; Hoarau et al., 2007; Provan and Bennett, 2008).

In addition to the Quaternary glacial episodes, the Messinian Salinity Crisis (MSC) also exerted a major effect on biotic distribution in the Mediterranean Sea. The desiccation of the Mediterranean Sea 5.96-5.33 mya resulted in the isolation of the Mediterranean Sea from the Atlantic Ocean caused by the closure of the Rifean and Baetic gateways as a result of a global lowering of the sea level and tectonic uplift (Maldonado, 1985; Duggen et al., 2003). During the MSC, the Mediterranean Sea was reduced to a series of hypersaline lakes with thick evaporate deposition. The MSC ended with the opening of the Strait of Gibraltar, allowing contact between the Mediterranean and the Atlantic at the end of the Miocene. These isolation events cause once continuous populations to become fragmented, allopatrically divergent and may be followed by secondary contact, having a great impact on the genetic diversity and structure of species and can promote reproductive incompatibility with full or partial speciation (Avise, 2000; Provan and Bennett, 2008; Makino, 2009).

Patterns of range expansion during favourable periods are likely to be species-specific (Chevolet et al., 2006; Larmuseau et al., 2009). It would be expected that many marine organisms with limited dispersal potential whose range has suffered a contraction followed by an expansion in response to climatic fluctuations would demonstrate a pattern of leptokurtic dispersal. In this scenario, a recently colonised area should show lower genetic diversity, possibly dominated by few genotypes and a high frequency of alleles identical to or recently descended from the founding populations (Maggs et al., 2008; Ibrahim et al., 1996). However, this model can lead to false assumptions, as periglacial refugia might also experience genetic bottlenecks resulting in low genetic diversity (Brochmann et al., 2003), or the high genetic diversity of presumed refugial populations may in fact be the result of secondary contact (Maggs et al., 2008). Southern marine refugia are known to have existed around the Iberian Peninsula and Mediterranean, however, small glacial refugia may have existed further north, around the British Isles and South West Ireland (Jolly et al., 2006).
In this study, we used a population genetic and phylogeographic approach to study the post-glacial distribution of genetic variation in the two congeneric and phenotypically cryptic cushion stars Asterina gibbosa (Echinodermata, Asteroidea) (Pennant, 1897) and A. phylactica (Emson and Crump, 1979). These species are congeners that have a Lusitanian distribution, sharing many characteristics and known littoral population locations. The range of these species is in fact almost identical, spanning from the Adriatic Sea to the west coast of Scotland, although A. gibbosa is far more common and found at many more locations than A. phylactica. Both species can be found in the littoral zone (Emson and Crump, 1979; Crump and Emson, 1983) but in addition are likely to feature substantial sublittoral populations. Both species possess an entirely benthic, lecithotrophic development, keeping an almost uninterrupted association with the substratum on which they live (Haesaerts et al., 2006; Crump and Emson, 1983). However, there are some important behavioural and life history differences between these presumed taxa: for example A. phylactica has a tendency to form a humped posture over egg masses and brood the eggs until metamorphosis, whereas A. gibbosa abandon the egg masses deposition (Crump and Emson, 1983), A. gibbosa grow to a far larger size than A. phylactica, growing up to 50mm compared to 15mm for A. phylactica and A. gibbosa is principally found on the sides and underside of boulders, whereas A. phylactica is usually found on the sides and top of the boulders.

However, both A. gibbosa and A. phylactica have been observed moving to the surface of the water floating upside down, a few days after metamorphosis (Haesaerts et al., 2006; Marthy, 1980; Soliman and Nojima, 1984; Byrne, 1995; Chen and Chen, 1992). Studies have reported that this juvenile rafting is likely to be a mechanism for gene-flow and dispersal (Byrne, 1995; Waters and Roy, 2004).

The aim of this study was to test the hypotheses that A. gibbosa and A. phylactica are two genetically distinct taxa that do not exchange genes where they occur in sympathy and that genetic structure among conspecific populations is concordant with the differences in life history in these two species – specifically that A. phylactica exhibits lower levels of population genetic diversity and higher population genetic structure than A. gibbosa.
4.3 Methods

The raw data analysed in this chapter is the same as that used for Chapters Two and Three. The raw data has been re-analysed to investigate the hypotheses outlined above. The details below provide a complete overview of the sampling locations and a brief description on the laboratory procedures and data analysis performed.

4.3.1 Sampling

*A. gibbosa* and *A. phylactica* samples were collected from under boulders and stones of rocky sea shores and tide pools at locations selected across the species' range (see Table 4.1 and Fig 4.1 for details). All specimens were stored for up to five days in absolute ethanol at ambient temperature and then for three days up to a three months at -80°C.
<table>
<thead>
<tr>
<th>Basin</th>
<th>Species</th>
<th>Population</th>
<th>Latitude</th>
<th>Longitude</th>
<th>Population Code</th>
<th>N (mtDNA)</th>
<th>N (AFLP)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atl AG</td>
<td>Outer Hope, S. Devon, UK</td>
<td>50°14'N  3°51'W</td>
<td>APP</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atl AG</td>
<td>Ayrmer Cove, S. Devon, UK</td>
<td>50°17'N  3°54'W</td>
<td>ACG</td>
<td>13</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atl AG</td>
<td>Bucks Mill, N. Devon, UK</td>
<td>50°59'N  4°20'W</td>
<td>BMAG</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atl AG</td>
<td>Chapel Point, S. Cornwall, UK</td>
<td>50°15'N  4°46'W</td>
<td>CPAG</td>
<td>5</td>
<td>31</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atl AG</td>
<td>Gara Point, S. Devon, UK</td>
<td>50°18'N  4°04'W</td>
<td>GPAG</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atl AG</td>
<td>Guéthary, Pyrénées-Atlantiques, Fr</td>
<td>43°25'N  1°37'W</td>
<td>FGAG</td>
<td>17</td>
<td>16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atl AG</td>
<td>Ladram Bay, S. Devon, UK</td>
<td>50°39'N  3°16'W</td>
<td>LBG</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atl AG</td>
<td>Littleham Cove, S. Devon, UK</td>
<td>50°36'N  3°21'W</td>
<td>LCG</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atl AG</td>
<td>Lough Hyne, Co. Cork, Ireland</td>
<td>50°30'N  9°18'W</td>
<td>LHAG</td>
<td>10</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atl AG</td>
<td>Prawle Point, S. Devon, UK</td>
<td>50°13'N  3°40'W</td>
<td>PPAG</td>
<td>18</td>
<td>30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atl AG</td>
<td>Rockham Bay, N. Devon, UK</td>
<td>51°10'N  4°12'W</td>
<td>RBG</td>
<td>10</td>
<td>33</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atl AG</td>
<td>Marazion, S. Cornwall, UK</td>
<td>50°07'N  5°28'W</td>
<td>SMAG</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atl AG</td>
<td>Welcombe Mouth, N. Devon, UK</td>
<td>50°55'N  4°32'W</td>
<td>WMG</td>
<td>8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atl AG</td>
<td>West Angle Bay, Pembrokeshire, UK</td>
<td>51°41'N  5°06'W</td>
<td>WAB</td>
<td>8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Med AG</td>
<td>Banyuls-sur-Mer, Pyrénées-Orientales, Fr</td>
<td>42°29'N  3°07'E</td>
<td>FBAG</td>
<td>12</td>
<td>16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Med AG</td>
<td>Naples Bay, Naples, Italy</td>
<td>40°49'N  14°14'E</td>
<td>INAG</td>
<td>13</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Med AG</td>
<td>Les Embiez, Provence-Alpes-Côte d'Azur, Fr</td>
<td>43°04'N  5°47'E</td>
<td>FEAG</td>
<td>3</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atl AP</td>
<td>Chapel Point, S. Cornwall, UK</td>
<td>50°15'N  4°46'W</td>
<td>CPAP</td>
<td>14</td>
<td>23</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atl AP</td>
<td>Gara Point, S. Devon, UK</td>
<td>50°18'N  4°04'W</td>
<td>GPAP</td>
<td>21</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atl AP</td>
<td>Hartland Quay, N. Devon, UK</td>
<td>50°59'N  4°32'W</td>
<td>HQAP</td>
<td>24</td>
<td>34</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atl AP</td>
<td>Lizard Point, S. Cornwall, UK</td>
<td>49°57'N  5°12'W</td>
<td>LPAP</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atl AP</td>
<td>Port Erin, Isle of Mann</td>
<td>54°05'N  4°45'W</td>
<td>IOAP</td>
<td>8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atl AP</td>
<td>Prawle Point, S. Devon, UK</td>
<td>50°13'N  3°40'W</td>
<td>PPAP</td>
<td>16</td>
<td>28</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atl AP</td>
<td>Rosemullion Head, S. Cornwall, UK</td>
<td>50°06'N  5°05'W</td>
<td>RMAP</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Med AP</td>
<td>Rovinj, Istrian Peninsula, Croatia</td>
<td>45°03'N  13°39'E</td>
<td>RAP</td>
<td>6</td>
<td>18</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.1. Sampling information for A. gibbosa (AG) and A. phylactica (AP). Med = Mediterranean Sea; Atl = Atlantic Ocean; Fr = France.
Fig. 4.1. A map of the sampling locations of *A. gibbosa* (blue) and *A. phylactica* (red). The numbers relate to the numbers assigned to populations in Table 4.1.

4.3.2 Laboratory Procedures
The amplification, sequencing and alignment of the mitochondrial genome that contains five transfer RNA genes and the 5' end of the COI gene (Hart *et al.*, 1997) was performed in Chapter 2 and the AFLP laboratory procedures were performed in Chapter 3 for *A. phylactica* and by Erika Baus (Baus *et al.*, 2005) for *A. gibbosa*. The approach taken for both species were the same.

4.3.3 Data Analysis
The mtDNA was analysed as per Chapter 2 and the AFLP data was analysed as per Chapter 3. A brief description of the tests performed can be found below.
4.3.3.1 mtDNA

Analysis of Genetic diversity and differentiation

Estimates of nucleotide diversity and haplotype diversity were obtained using the program Arlequin vs. 3.11 (Excoffier et al., 2005). A median joining network (MJN) (Bandelt et al., 1999) and Neighbour Joining Tree using the Tamura Nei model of evolution were constructed to visualise the relationship between haplotypes. Pairwise divergences between haplotypes were calculated using the Tamura Nei model of evolution. This evolutionary model (Tamura and Nei, 1993) corrects for multiple hits, taking into account the differences in substitution rate between nucleotides, distinguishing between transition and transversion substitution rates and the inequality of nucleotide frequencies, but assuming equality of substitution rates among sites. An analysis of molecular variance (AMOVA) was performed to investigate the levels of differentiation between populations and groups of populations.

Analysis of Population Demography

To infer the demographic history of populations, mismatch distributions of pairwise nucleotide differences among COI haplotypes were compared with expectations of a sudden-expansion model (Rogers, 1995) using Arlequin (vs. 3.11) (Excoffier et al., 2005). Tajima's D (Tajima, 1989), Fu's Fs (Fu, 1997) and Fu and Li's F* and D* (Fu and Li, 1997), were conducted to investigate if there was a divergence from neutrality. In the event of a demographic expansion, an approximate estimation of time since expansion was calculated using \( t = 2ut \) (Rogers and Harpending, 1992).

4.3.3.2 AFLP

Genetic diversity was evaluated as the percentage of polymorphic loci (calculated using Aflp-surv version 1.0; Vekemans, 2002) and the number of shared multilocus AFLP patterns (evaluated with Arlequin version 3.11; Excoffier et al., 2005). Genetic differentiation among populations was evaluated using a hierarchical Bayesian approach developed by Holsinger et al. (2002) using the software HICKORY version 1.0.3 (Holsinger et al., 2002) to estimate \( \Theta_B \), a Bayesian analogue of \( F_{ST} \), across all populations and for each pairwise combination of populations. This Bayesian approach does not assume any prior knowledge of the degree of within-population
inbreeding and is therefore not subject to the problems of traditional methods of analysis using dominant markers.

Four models were assessed: a full model, a model that assumes no inbreeding within populations \( (F_{IS} = 0 \text{ model}) \), a model that assumes no differentiation among populations \( (\theta^B = 0 \text{ model}) \) and a model that does not attempt to estimate \( F_{IS} \) (f-free model). Because estimates of \( F_{IS} \) derived from dominant marker data may be unreliable (Holsinger and Wallace, 2004), we used the f-free analysis as our preferred method to calculate estimates of \( \theta^B \). The deviance information criterion (DIC) values for the \( F_{IS} = 0 \), \( \theta^B = 0 \) and full models were used to estimate how well each model fitted the data (a smaller DIC value indicates a better fit) and which model should be preferred. Pairwise FST values across all populations were calculated using AMOVA vs. 3.11 as a comparison to the \( \theta^B \) data. \( F_{ST} \) values across all populations were calculated using Aflp-surv version 1.0 as a comparison to the \( \theta^B \) data.

A Mantel test was performed using the program TFPGA (Miller, 1997) to evaluate the geographical structure among sample sites. The matrix used pairwise \( \theta^B \) and geographical distances calculated as the shortest distance between two populations that did not involve crossing land.

To estimate population structure and differences between \( A. \ gibbosa \) and \( A. \ phylactica \) we conducted a series of AMOVA tests using Arlequin version 3.11. AMOVA was used to investigate genetic differentiation between \( A. \ gibbosa \) and \( A. \ phylactica \) and to assess any geographic structuring. Paired t-tests using arcsin transformed fragment frequencies (which were normally distributed) calculated in AFLP-Surv version 1.0 (Vekemans, 2002) were performed using Minitab version 15 to investigate differences between and within \( A. \ gibbosa \) and \( A. \ phylactica \). Principal coordinate analysis (PCA) based on FST distances between AFLP multilocus phenotypes (calculated with Arlequin 2.000) was performed using Genalex version 5.1 (Peakall and Smouse 2001). The PCA via covariance matrix with data standardization method was chosen. Assignment tests were carried out using the re-allocation procedure of aflpop version 1.1 (Duchesne and Bernatchez, 2002).
4.4 Results

4.4.1 Mitochondrial DNA

Sequences and Haplotype Analysis

A 332-bp mtDNA fragment incorporating five transfer RNA genes, a non-coding intron and the 5' end of the COI gene, from 272 individuals from 26 populations, (18 *A. gibbosa* and eight *A. phylactica*) was sequenced. The aligned sequences and inferred protein translations are presented in Appendix 1. Of these populations, three sites had both species present (Prawle Point, S. Devon; Gara Point, S. Devon; Chapel Point, S. Cornwall). The sequences sampled revealed 20 polymorphic sites, producing 15 unique haplotypes, with two shared between the two species (Tables 4.2a and 4.2b). Of these polymorphisms, five occurred at first codon positions, four occurred at second codon positions and 11 occurred at third codon positions. Of the 17 haplotypes, two were shared across both species (all of the other haplotypes were private either to *A. gibbosa* or *A. phylactica*). However, the shared haplotypes are those with the greatest frequency in both species (AG1 and AP2; AG2 and AP1). There is a two base pair insertion in the intron prior to the 5' end of the COI gene in haplotype AG9, this insertion is also present in *Patiriella exigua*, which was used as an outgroup (Hart *et al.*, 1997). In addition, there are two stop codons present in the tRNA-Gln gene present in all of the sequences. These stop codons are also present in the original sequence published by Hart *et al.* (1997) and it is likely that in these echinoderms, these sequences are not acting as stop codon.
<table>
<thead>
<tr>
<th>Population Code</th>
<th>AG1</th>
<th>AG2</th>
<th>AG3</th>
<th>AG4</th>
<th>AG5</th>
<th>AG6</th>
<th>AG7</th>
<th>AG8</th>
<th>AG9</th>
<th>AG10</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>APP</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>ACG</td>
<td>13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>13</td>
</tr>
<tr>
<td>BMAG</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>CPAG</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5</td>
</tr>
<tr>
<td>GPAG</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>10</td>
</tr>
<tr>
<td>FGAG</td>
<td>16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>17</td>
</tr>
<tr>
<td>LBG</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6</td>
</tr>
<tr>
<td>LCG</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td>5</td>
</tr>
<tr>
<td>LHAG</td>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td>PAG</td>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td>18</td>
</tr>
<tr>
<td>RBG</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>10</td>
</tr>
<tr>
<td>SPG</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td>1</td>
<td>8</td>
</tr>
<tr>
<td>SMMAG</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6</td>
</tr>
<tr>
<td>WMG</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>8</td>
</tr>
<tr>
<td>WAB</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td></td>
<td>8</td>
</tr>
<tr>
<td>FBAG</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>10</td>
<td></td>
<td></td>
<td>12</td>
</tr>
<tr>
<td>INAG</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4</td>
<td>1</td>
<td></td>
<td>13</td>
</tr>
<tr>
<td>FEAG</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>Total</td>
<td>117</td>
<td>16</td>
<td>1</td>
<td>13</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>157</td>
</tr>
</tbody>
</table>

**Table 4.2(a).**

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>AP1</th>
<th>AP2</th>
<th>AP3</th>
<th>AP5</th>
<th>AP6</th>
<th>AP7</th>
<th>AP8</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPAP</td>
<td>13</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>14</td>
</tr>
<tr>
<td>GPAP</td>
<td>16</td>
<td>4</td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>21</td>
</tr>
<tr>
<td>HQAP</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>16</td>
<td></td>
<td></td>
<td></td>
<td>24</td>
</tr>
<tr>
<td>LPAP</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>10</td>
</tr>
<tr>
<td>IOMP</td>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>8</td>
</tr>
<tr>
<td>PAP</td>
<td>11</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>16</td>
</tr>
<tr>
<td>RMP</td>
<td>8</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>9</td>
</tr>
<tr>
<td>RAP</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Total</td>
<td>69</td>
<td>16</td>
<td>1</td>
<td>1</td>
<td>16</td>
<td>2</td>
<td>3</td>
<td>108</td>
</tr>
</tbody>
</table>

**Table 4.2(b).** COI haplotypes of *A. gibbosa* (4.2 (a)) and *A phylactica* (4.2 (b)). Data comprises: total number of haplotypes identified in each sampling location and the number of individuals found with each haplotype in each sampling location. Populations above the line are found in the Atlantic Ocean whereas populations found below the line are found in the Mediterranean. The blue haplotypes are grouped into haplogroup 1 whereas the red haplotypes are grouped into haplogroup 2.
Fig. 4.2. Geographical positions of sampling localities (see Table 4.1) and distribution of *A. gibbosa* and *A. phylactica* haplotypes. The size of the circle corresponds to the number of individuals sampled from each population. Each colour corresponds to a different haplotype.

The median joining network (MJN) revealed a pattern reminiscent of admixture by secondary contact between two groups of widely distributed and ancestral haplotypes (Hap AG1/ AP2 and Hap AG2/ AP1) (Fig. 4.3). The two haplogroups are separated by 10 nucleotide substitutions, with individuals from both basins and both species present in both haplogroups. The frequencies of these haplogroups are very different between the two species, with 88% of *A. gibbosa* individuals being assigned to haplogroup 1 compared to 16% of *A. phylactica* individuals and 12% of *A. gibbosa* individuals are assigned to haplogroup 2 compared to 84% of *A. phylactica* individuals. Overall 59% of individuals are assigned to haplogroup 1 and 41% of
individuals are assigned to haplogroup 2. Within the Atlantic 57% of individuals are assigned to haplogroup 1 (87% *A. gibbosa* and 17% *A. phylactica*) and 43% of individuals are assigned to haplogroup 2 (13% *A. gibbosa* and 83% *A. phylactica*), whereas in the Mediterranean 74% of individuals are assigned to haplogroup 1 (89% *A. gibbosa* and 0% *A. phylactica*) and 26% of individuals are assigned to haplogroup 2 (11% *A. gibbosa* and 100% *A. phylactica*). The two divergent haplogroups are evident not only from the MJN but also the neighbour joining tree, based on Tamura Nei genetic distances (Fig. 4.4), which shows 100% bootstrap support for the two clades. The haplotype pairwise Tamura Nei values can be seen in Table 4.3.

**Fig. 4.3.** Median joining network for *A. gibbosa* and *A. phylactica* mtDNA haplotypes, indicating the haplotype relationships. Circles represent haplotypes, the circle size is proportional to the haplotype frequency. Each cross-line is indicative of a site which has undergone a substitution. Blue = *A. gibbosa* individuals from the Atlantic; Red = *A. gibbosa* individuals from the Mediterranean; Light blue = *A. phylactica* from the Atlantic; Pink = *A. phylactica* individuals from the Mediterranean; Yellow = median vectors of un-sampled or extinct ancestral sequences.
<table>
<thead>
<tr>
<th></th>
<th>AG1/AG2</th>
<th>AG2/AG1</th>
<th>AG3</th>
<th>AG4</th>
<th>AG5</th>
<th>AG6</th>
<th>AG7</th>
<th>AG8</th>
<th>AG9</th>
<th>AG10</th>
<th>AP3</th>
<th>AP5</th>
<th>AP6</th>
<th>AP7</th>
<th>AP8</th>
</tr>
</thead>
<tbody>
<tr>
<td>AG1</td>
<td>0.054</td>
<td>0.027</td>
<td>0.060</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AG2</td>
<td>0.027</td>
<td>0.054</td>
<td>0.054</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AG3</td>
<td>0.005</td>
<td>0.060</td>
<td>0.032</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AG4</td>
<td>0.048</td>
<td>0.005</td>
<td>0.054</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AG5</td>
<td>0.048</td>
<td>0.016</td>
<td>0.043</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AG6</td>
<td>0.005</td>
<td>0.054</td>
<td>0.032</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AG7</td>
<td>0.005</td>
<td>0.060</td>
<td>0.032</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AG8</td>
<td>0.048</td>
<td>0.005</td>
<td>0.054</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AG9</td>
<td>0.048</td>
<td>0.005</td>
<td>0.054</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AG10</td>
<td>0.001</td>
<td>0.054</td>
<td>0.027</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AP3</td>
<td>0.054</td>
<td>0.000</td>
<td>0.060</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AP5</td>
<td>0.049</td>
<td>0.005</td>
<td>0.055</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AP6</td>
<td>0.043</td>
<td>0.010</td>
<td>0.049</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.3. Tamura-Nei pairwise divergences (%) between mtDNA haplotypes for *A. gibbosa* and *A. phylactica*.
Fig. 4.4. Neighbour joining phenogram tree based on Tamura Nei genetic distances for *A. gibbosa* and *A. phylactica* mtDNA haplotypes, with bootstrap values on the branches. The tree is rooted by *Patiriella exigua* (Genbank No. U50053 (Hart *et al.*, 1997)).
<table>
<thead>
<tr>
<th>Sample Size</th>
<th>A. gibbosa</th>
<th>A. phylactica</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of Populations</td>
<td>18</td>
<td>8</td>
</tr>
<tr>
<td>No. of Haplotypes</td>
<td>10</td>
<td>7</td>
</tr>
<tr>
<td>Nucleotide Diversity</td>
<td>0.009 +/- 0.006</td>
<td>0.012 +/- 0.007</td>
</tr>
<tr>
<td>Haplotype (gene) Diversity</td>
<td>0.423 +/- 0.046</td>
<td>0.552 +/- 0.048</td>
</tr>
<tr>
<td>Mean no. pairwise differences</td>
<td>2.457 +/- 1.336</td>
<td>3.704 +/- 1.886</td>
</tr>
<tr>
<td>Fu's Fs Statistic</td>
<td>0.743 (P &gt; 0.10)</td>
<td>4.786 (P &gt; 0.10)</td>
</tr>
<tr>
<td>Fu and Li's F</td>
<td>-0.224 (P &gt; 0.10)</td>
<td>0.970 (P &gt; 0.10)</td>
</tr>
<tr>
<td>Fu and Li's D</td>
<td>-0.319 (P &gt; 0.10)</td>
<td>0.820 (P &gt; 0.10)</td>
</tr>
<tr>
<td>Tajimas D</td>
<td>-0.183 (P &gt; 0.10)</td>
<td>0.815 (P &gt; 0.10)</td>
</tr>
</tbody>
</table>

Table 4.4. Comparative summary statistics for mtDNA diversity within *A. gibbosa* and *A. phylactica*.

Table 4.4 shows that the 157 *A. gibbosa* individuals belong to ten haplotypes, whereas the 108 *A. phylactica* individuals belong to seven haplotypes (although since only eight populations could be sampled for *A. phylactica*, we did not test whether this was statistically significantly different). For both species nucleotide diversity is low and haplotype diversity is high, with relatively similar values. The values are also similar for the mean number of pairwise differences. Furthermore, none of the neutrality tests are statistically significant. These results are expected as the two divergent haplogroups are present within both species.
When individuals were grouped according to basin there were a total of 231 individuals from 22 populations sampled in the Atlantic and 34 individuals from four populations in the Mediterranean, which yielded 11 haplotypes present in the Atlantic and nine haplotypes in the Mediterranean. Seven haplotypes were private to the Atlantic and four were private to the Mediterranean, only four haplotypes were present in both basins. Similar results were observed for both the Mediterranean and Atlantic groups, with populations in both basins showing low nucleotide diversity and high haplotype diversity (Table 4.5). In the Mediterranean none of the neutrality tests are statistically significant; however, in the Atlantic the Fu and Li’s F and D statistics are both statistically significant whereas the Fu’s Fs and Tajima’s D statistics are both non-significant, this is indicative of background selection occurring.

**Genetic Diversity and Population Structure**

An Analysis of Molecular Variance (AMOVA) was then performed (Table 4.6) to test for population structure for *A. gibbosa* and *A. phylactica* populations across their geographical range, to compare population structure within and among each species. When no species groups were specified, most of the variation (55.91% (P < 0.01)) was found to occur within populations (although both within and among population variance components were significant). When *A. gibbosa* and *A. phylactica* populations were grouped together by basin, the percentage of variation explained by basin (38.39%, P < 0.01) was significantly lower than for *A. phylactica* populations, which showed a significant variation (61.81%, P < 0.01). This suggests that *A. phylactica* populations have a more distinct population structure than *A. gibbosa* populations, which may be due to the differences in their habitats and life histories.

Table 4.5. Summary statistics for mtDNA diversity within *A. gibbosa* and *A. phylactica* when analysing them together, divided into the basin they were sampled from.

<table>
<thead>
<tr>
<th></th>
<th>Atlantic</th>
<th>Mediterranean</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sample Size</strong></td>
<td>231</td>
<td>34</td>
</tr>
<tr>
<td><strong>No. of Populations</strong></td>
<td>22</td>
<td>4</td>
</tr>
<tr>
<td><strong>No. of Haplotypes</strong></td>
<td>11</td>
<td>9</td>
</tr>
<tr>
<td><strong>Nucleotide Diversity</strong></td>
<td>0.016 +/- 0.008</td>
<td>0.018 +/- 0.010</td>
</tr>
<tr>
<td><strong>Haplotype (gene) Diversity</strong></td>
<td>0.522 +/- 0.045</td>
<td>0.530 +/- 0.057</td>
</tr>
<tr>
<td><strong>Mean no. pairwise differences</strong></td>
<td>5.118 +/- 2.494</td>
<td>4.103 +/- 2.061</td>
</tr>
<tr>
<td><strong>Fu’s Fs Statistic</strong></td>
<td>3.029 (P &gt;0.10)</td>
<td>3.350 (P &gt;0.10)</td>
</tr>
<tr>
<td><strong>Fu and Li’s F</strong></td>
<td>2.548* (P &lt;0.02)</td>
<td>0.561 (P &gt;0.10)</td>
</tr>
<tr>
<td><strong>Fu and Li’s D</strong></td>
<td>1.472* (P &lt;0.10)</td>
<td>0.320 (P &gt;0.10)</td>
</tr>
<tr>
<td><strong>Tajimas D</strong></td>
<td>1.874 (P &gt;0.10)</td>
<td>1.421 (P &gt;0.10)</td>
</tr>
</tbody>
</table>
Phylactica were specified in the test, the results changed, with the majority of the variation being partitioned among groups – i.e. putative species (55.05% (P <0.01)), with 40.77% occurring within populations (P <0.01), recapitulating the extreme frequency differences found in the major haplogroups in both species.

Furthermore, when A. gibbosa and A. phylactica individuals from the Atlantic were analysed the greatest proportion of variation was again attributed to differences between putative A. gibbosa and A. phylactica (55% (P <0.01)), with 42% (P <0.01) of the difference being attributed to within population variation and small but significant component of the variation attributed to among population variation within species. The results were similar for individuals within the Mediterranean with 67% of the variation being attributed to differences between A. gibbosa and A. phylactica (P >0.10) and 29% of the variation difference being attributed to within population variation (P <0.01). The non-significance of the first result is most likely to be the result of the small sample size.
<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>d.f.</th>
<th>Sum of Squares</th>
<th>Variance Components</th>
<th>Percentage of Variation</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Groups</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Among populations</td>
<td>25</td>
<td>402.538</td>
<td>1.420</td>
<td>44.09</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Within populations</td>
<td>239</td>
<td>430.288</td>
<td>1.800</td>
<td>55.91</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Total</td>
<td>264</td>
<td>832.826</td>
<td>3.220</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Two Groups: A. gibbosa/A.phylactica</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Among Groups</td>
<td>1</td>
<td>315.449</td>
<td>2.431</td>
<td>55.05</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Among populations within groups</td>
<td>24</td>
<td>87.089</td>
<td>0.185</td>
<td>4.18</td>
<td>&lt;0.05</td>
</tr>
<tr>
<td>Within Populations</td>
<td>239</td>
<td>430.288</td>
<td>1.800</td>
<td>40.77</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Total</td>
<td>264</td>
<td>832.826</td>
<td>4.416</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Two Groups: Atlantic A. gibbosa/ A. phylactica</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Among Groups</td>
<td>1</td>
<td>363.447</td>
<td>3.138</td>
<td>54.08</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Among populations within groups</td>
<td>20</td>
<td>99.669</td>
<td>0.253</td>
<td>4.36</td>
<td>&lt;0.05</td>
</tr>
<tr>
<td>Within Populations</td>
<td>209</td>
<td>503.884</td>
<td>2.4113</td>
<td>41.56</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Total</td>
<td>230</td>
<td>967.000</td>
<td>5.802</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Two Groups: Mediterranean A. gibbosa/ A. phylactica</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Among Groups</td>
<td>1</td>
<td>30.925</td>
<td>2.873</td>
<td>66.78</td>
<td>&gt;0.10</td>
</tr>
<tr>
<td>Among populations within groups</td>
<td>2</td>
<td>5.536</td>
<td>0.185</td>
<td>4.29</td>
<td>&lt;0.10</td>
</tr>
<tr>
<td>Within Populations</td>
<td>30</td>
<td>37.333</td>
<td>1.244</td>
<td>28.93</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Total</td>
<td>33</td>
<td>73.794</td>
<td>4.302</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Two Groups: Atlantic A. gibbosa &amp; A. phylactica/ Mediterranean A. gibbosa &amp; A. phylactica</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Among Groups</td>
<td>1</td>
<td>13.898</td>
<td>-0.060</td>
<td>-1.89</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Among populations within groups</td>
<td>24</td>
<td>388.640</td>
<td>1.434</td>
<td>45.17</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Within Populations</td>
<td>239</td>
<td>430.288</td>
<td>1.800</td>
<td>56.72</td>
<td>&gt;0.10</td>
</tr>
<tr>
<td>Total</td>
<td>264</td>
<td>832.826</td>
<td>3.174</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Four Groups: Atlantic A. gibbosa / Atlantic A. phylactica/ Mediterranean A. gibbosa/ Mediterranean A. phylactica</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Among Groups</td>
<td>3</td>
<td>324.129</td>
<td>1.953</td>
<td>49.64</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Among populations within groups</td>
<td>22</td>
<td>83.659</td>
<td>0.203</td>
<td>5.16</td>
<td>&lt;0.02</td>
</tr>
<tr>
<td>Within Populations</td>
<td>239</td>
<td>425.038</td>
<td>1.778</td>
<td>45.20</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Total</td>
<td>264</td>
<td>832.826</td>
<td>3.935</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.6. ΦST AMOVA for *A. gibbosa* and *A. phylactica* samples indicating the groups tested, and showing the percentage of genetic variation accounted for when the data is divided up into different groups.
When the genotypes from both species are pooled as if they belonged to a single species and geographic differences assessed using AMOVA, it can be seen that the majority of the variation was attributed to within populations (57% (P>0.10)), 45% (P<0.01) of the variation is attributed to variation among populations within groups and there was no variation attributed to differences between the Atlantic and Mediterranean. This data suggests that the difference between \textit{A. gibbosa} and \textit{A. phylactica} greatly exceeds the difference resulting from the geographic isolation of populations in the Atlantic and Mediterranean basins. When the populations were grouped into both species and geographical location most of the variation was attributed to between groups (49.64% (P <0.01), with almost as large a portion of the variation attributed to within populations (45.20% (P <0.01).

Finally, when analysing \textit{A. gibbosa} and \textit{A. phylactica} as putative species it was found that when the range was split into the Atlantic Ocean and Mediterranean Sea most variation was partitioned within populations – within the Atlantic it was 90% (P<0.05) for \textit{A. gibbosa} and 93% (P<0.05) for \textit{A. phylactica}, in the Mediterranean it was also 90% (P<0.10) for \textit{A. gibbosa} (data not shown, see Chapters 2 and 3).
**Fig. 4.5.** Geographical positions of sampling localities (see Table 4.1) and distribution of *A. gibbosa* and *A. phylactica* haplogroups. The size of the circle corresponds to the number of individuals sampled from a population, with each slice corresponding to a haplogroup. The blue slices correspond to *A. gibbosa* haplogroup 1, the red slices correspond to *A. gibbosa* haplogroup 2, the orange slices correspond to *A. phylactica* haplogroup 1 and the green slices correspond to *A. phylactica* haplogroup 2. *A. gibbosa* individuals are a darker shade than the *A. phylactica* individuals.

**Haplogroup Analysis**

The MJN and NJ data shows that there is some genetic structure, with two distinct haplogroups. The data obtained by grouping individuals into haplogroups (Table 4.7) gives different results to the data seen when comparing the Atlantic and Mediterranean populations (Table 4.5). This is not surprising as the majority of
individuals present in both basins are assigned to the largest haplotype (AG1), which is placed into haplogroup one.

<table>
<thead>
<tr>
<th></th>
<th>Haplogroup 1</th>
<th>Haplogroup 2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sample Size</strong></td>
<td>161</td>
<td>111</td>
</tr>
<tr>
<td><strong>No. of Populations</strong></td>
<td>21</td>
<td>18</td>
</tr>
<tr>
<td><strong>No. of Haplotypes</strong></td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td><strong>Nucleotide Diversity</strong></td>
<td>0.001 +/- 0.001</td>
<td>0.001 +/- 0.001</td>
</tr>
<tr>
<td><strong>Haplotype (gene) Diversity</strong></td>
<td>0.401 +/- 0.046</td>
<td>0.576 +/- 0.047</td>
</tr>
<tr>
<td><strong>Mean no. pairwise differences</strong></td>
<td>0.307 +/- 0.319</td>
<td>0.312 +/- 0.323</td>
</tr>
<tr>
<td><strong>Fu’s Fs Statistic</strong></td>
<td>-7.079* (P &lt;0.01)</td>
<td>-9.416* (P &lt;0.01)</td>
</tr>
<tr>
<td><strong>Fu and Li’s F</strong></td>
<td>-3.771* (P &lt;0.02)</td>
<td>-0.715 (P &gt;0.10)</td>
</tr>
<tr>
<td><strong>Fu and Li’s D</strong></td>
<td>-3.912* (P &lt;0.02)</td>
<td>-0.288 (P &gt;0.10)</td>
</tr>
<tr>
<td><strong>Tajimas D</strong></td>
<td>-1.885* (P &lt;0.01)</td>
<td>-1.601* (P &lt;0.05)</td>
</tr>
</tbody>
</table>

Table 4.7. Summary statistics for *A. gibbosa* and *A. phylactica* when the mtDNA haplotypes are classified into haplogroups.

Haplogroup 2 contained individuals from all eight *A. phylactica* populations sampled, across the entire geographical range, as well as ten of the *A. gibbosa* populations. Haplogroup 1 contained individuals from all of the *A. gibbosa* populations, and five of the eight *A. phylactica* populations. The three populations it is absent from are the Isle of Mann, Croatia and Lizard Point populations. Of the 265 individuals sampled, 155 (137 *A. gibbosa* and 18 *A. phylactica*) were assigned to haplogroup 1 and 110 were assigned to haplogroup 2 (20 *A. gibbosa* and 90 *A. phylactica*). Haplogroup 2 included one more haplotype than haplogroup 1. Neither haplogroup is confined to a specific geographical region, with representatives of both species in both haplogroups being present in both the Atlantic and Mediterranean (Fig. 4.5).

Population Demography

A negative binomial curve is observed in the mismatch distribution for both haplogroups. This indicates that a demographic expansion has occurred at some time in the past (Fig. 4.6). As with the analysis of *A. gibbosa* and *A. phylactica* individually in Chapters Two and Three respectively both haplogroups showed a large peak at the origin, reflecting that most individuals in those respective groups possessed the same abundant haplotypes.
Fig. 4.6. Mismatch distribution showing frequencies of the two haplogroups found in *A. gibbosa* and *A. phylactica*. The red curve is the expected distribution under the sudden expansion model.

When assessed as individual species (Table 4.3) none of the neutrality tests gave statistically significant results. When the data was analysed as haplogroups (Table 4.7) the results were different - the negative Tajima's D test statistic for haplogroup 1 was highly significant for deviations from the model of neutrality, indicating either range expansion or background selection. Fu's Fs for haplogroup 1 was highly negative and highly significant and Fu and Li's F* and D* were both negative and
significant, often indicating background selection. This suggests that haplogroup 1 is either expanding or has undergone a mitochondrial selective sweep (Rogers 1995). Haplogroup 2 showed strong significant negative values for both the Tajima’s D and Fu’s F tests, indicating a population expansion has occurred. This result in conjunction with the mismatch distributions and the networks suggests that a recent demographic expansion has occurred.

The number of generations since expansion ranged from 160,000 – 181,000 ($\tau = 3.00$, $SD = 0.003$) for haplogroup 1 and 160,000 – 181,000 for haplogroup 2 ($\tau = 3.00$, $SD = 0.013$). As both species have different generation times (four years for *A. gibbosa* and two years for *A. phylactica*) the time since expansion is in the range of 341,000 – 681,000 for haplogroup 1 (with *A. gibbosa* being the dominant species) or 640,000 – 722,000 for haplogroup 2 (with *A. phylactica* being the dominant species).

4.4.2 AFLP Data

One hundred and fifty-nine specimens of *A. gibbosa* sampled from eight populations (seven natural populations and one population from a local aquarium) and one hundred and three specimens of *A. phylactica* sampled from four populations (all natural) were sampled across the species range (as shown in Fig. 4.7). These specimens were analysed using five selective primer pair combinations and generated a total of 292 bands. All individuals except two displayed unique AFLP band patterns, suggesting a high level of genetic variability. As shown in Table 4.8, the percentage of polymorphic bands across the 11 natural populations ranged between 19.2% and 39.7% (Table 4.8). The population sampled in the aquarium of the ‘Institut Océanographique Paul Ricard’ in Les Embiez showed an even lower level of polymorphism (7.5%).
Fig. 4.7. A map to show the sampling locations for the AFLP study. Red arrows = *A. gibbosa* sites, green arrows = *A. phylactica* sites and blue arrows = *A. gibbosa* and *A. phylactica* sites.
Table 4.8. Characteristics of the samples collected from twelve populations of *A. gibbosa* and *A. phylactica*. For each sample site, the number of individuals analysed (n) and the percentage of polymorphic AFLP loci (%P) are presented.

A matrix of pairwise θ^B and FST values between populations is shown in Table 4.9. It shows high levels of genetic differentiation between populations consisting of *A. gibbosa* and *A. phylactica* individuals with an average θ^B value of 0.702, ranging from 0.528 (Rovinj, Croatia – Naples, Italy) to 0.778 (Chapel Point, S. Devon - Les Embiez, Atlantic France). Within both species the average values were 0.319 for *A. phylactica* with a range of 0.038 (Chapel Point, S. Devon – Prawle Point, S. Devon) to 0.592 (Rovinj, Croatia – Hartland Quay, N. Devon) and an average of 0.300 for *A. gibbosa*, with a range of 0.0274 (Chapel Point, S. Devon – Prawle Point, S. Devon) to 0.5093 (Lough Hyne, Ireland - Les Embiez, Atlantic France). The paired t-test showed that there was a significant difference between *A. gibbosa* and *A. phylactica* within the Atlantic populations (t = 4.1355, p <0.001). The average pairwise θ^B value within pairs of *A. gibbosa* population is 0.300 (FST 0.302), which is close to the figure observed for *A. phylactica* 0.319 (FST 0.307), the figure is far greater between the two species with an average θ^B value of 0.702 (FST 0.717).
The deviance information criteria (DIC) statistic provided by the Bayesian factor analysis was used as a model choice criterion between three models: the full model, the $F_{IS} = 0$ model and the $\theta^B = 0$ model. Here, the DIC values were, respectively, 5400 using the full model, 5486 for the $F_{IS} = 0$ model and 56069 for the $\theta^B = 0$ model. The full model was thus clearly preferred to the $\theta^B = 0$ model, supporting the existence of a significant level of differentiation among populations. However, as for *A. phylactica* in Chapter 3 and there was only weak evidence that the full model should be preferred to the $F_{IS} = 0$ model since the difference in DIC (86 units)

<table>
<thead>
<tr>
<th>CPAP</th>
<th>HQAP</th>
<th>PPAP</th>
<th>RAP</th>
<th>CPAG</th>
<th>PPAG</th>
<th>RBAG</th>
<th>LHAG</th>
<th>FGAG</th>
<th>FBAG</th>
<th>FEAG</th>
<th>INAG</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>224</td>
<td>74</td>
<td>5058</td>
<td>74</td>
<td>257</td>
<td>385</td>
<td>819</td>
<td>3061</td>
<td>3225</td>
<td>3735</td>
<td></td>
</tr>
<tr>
<td>0.0767</td>
<td>0.039</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.0381</td>
<td>0.018</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.5760</td>
<td>0.565</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.7443</td>
<td>0.755</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.7423</td>
<td>0.759</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.7238</td>
<td>0.716</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.7288</td>
<td>0.745</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.7377</td>
<td>0.757</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.7510</td>
<td>0.776</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.7778</td>
<td>0.828</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.6939</td>
<td>0.706</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 4.9.** Pairwise genetic differentiations estimated $\theta^B$ values, with FST values in brackets, between eight populations of *A. gibbosa* and 4 populations of *A. phylactica* (bottom left triangle). Geographical distances between populations in kilometres (top right triangle).
between the two models arises as a result of differences in model dimensions (pD = 1109 for the $F_{IS} = 0$ model and 1013 for the full model).

A Mantel test was performed in order to evaluate the correlation between genetic differentiation (estimated as $\theta^B$ values) and geographical distance (measured as the shortest distance without crossing land in kilometres) between all populations of both species. There was not a statistically significant correlation between genetic differentiation and geographical distance ($r = 0.2303, p = 0.072$).

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>d.f.</th>
<th>Sum of Squares</th>
<th>Variance Components</th>
<th>Percentage of Variation</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>No Regional Groups</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Among populations</td>
<td>11</td>
<td>9150.685</td>
<td>38.209</td>
<td>78.14</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Within populations</td>
<td>250</td>
<td>2671.803</td>
<td>10.687</td>
<td>21.86</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Total</td>
<td>261</td>
<td>11822.489</td>
<td>48.896</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Two Groups: Atlantic/ Mediterranean</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Among Groups</td>
<td>1</td>
<td>877.024</td>
<td>1.652</td>
<td>3.31</td>
<td>&gt;0.10</td>
</tr>
<tr>
<td>Among populations within groups</td>
<td>10</td>
<td>8273.661</td>
<td>37.570</td>
<td>75.28</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Within Populations</td>
<td>250</td>
<td>2671.803</td>
<td>10.687</td>
<td>21.41</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Total</td>
<td>261</td>
<td>11822.489</td>
<td>49.910</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Two Groups: A. gibbosa/ A. phylactica</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Among Groups</td>
<td>1</td>
<td>6895.56</td>
<td>52.93</td>
<td>71.69</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Among populations within groups</td>
<td>10</td>
<td>2255.12</td>
<td>10.21</td>
<td>13.84</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Within Populations</td>
<td>250</td>
<td>2671.80</td>
<td>10.68</td>
<td>14.47</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Total</td>
<td>261</td>
<td>11822.48</td>
<td>73.84</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Four Groups: Atlantic A. gibbosa / Atlantic A. phylactica/ Mediterranean A. gibbosa/ Mediterranean A. phylactica</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Among Groups</td>
<td>3</td>
<td>8557.965</td>
<td>48.002</td>
<td>77.87</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Among populations within groups</td>
<td>8</td>
<td>592.720</td>
<td>2.956</td>
<td>4.79</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Within Populations</td>
<td>250</td>
<td>2671.803</td>
<td>10.687</td>
<td>17.34</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Total</td>
<td>261</td>
<td>11822.489</td>
<td>61.645</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.10. AMOVA, based on FST values between AFLP multilocus phenotypes, for 159 A. gibbosa and 103 A. phylactica individuals sampled from twelve populations, with and without structuring.

AMOVA revealed (Table 4.10) that 22% of the total genetic variation was attributed to differences between individuals within populations ($P <0.01$), while 78% was
attributed to differences among populations (P <0.01). When populations were grouped into Atlantic and Mediterranean regions, 75% (P <0.01) of the variance was attributed to differences among populations within groups, 21% (P <0.01) to differences between within populations and just 3% (P >0.1) to differences among groups. This data suggests that the difference between the two species is far greater than the differences attributed to geographical location. 72% of the total genetic variation was attributed to differences between groups (A. gibbosa and A. phylactica) (P <0.01), with just 14% attributed to differences among populations within groups and within populations (P <0.01). When populations were grouped according to both species and geographical location, 78% (P <0.01) of the variance was attributed to differences among populations within groups, 17% (P <0.01) to differences between within populations and just 3% (P <0.01) to differences among groups, which reinforces the large genetic distances between both species and basins. When analysing A. gibbosa and A. phylactica as putative species it was found that when the range was split into the Atlantic Ocean and Mediterranean Sea most variation was partitioned within populations – within the Atlantic it was 78% (P<0.05) for A. gibbosa and 93% (P<0.05) for A. phylactica, in the Mediterranean it was 51% (P<0.05) for A. gibbosa (data not shown, see chapters 2 and 3).
Fig. 4.8. Principal coordinate analysis (PCA), based on Pairwise genetic differentiations (estimated as $F_{ST}$ values using AFLP-SURV) between AFLP multilocus phenotypes, for 141 *A. gibbosa* sampled from 4 populations and 103 *A. phylactica* sampled from 8 populations.

The PCA analysis provides a visual representation of genetic similarity between both species. The plot of the first and second principal coordinates, which accounted for 74.5% and 7.5% of the total variation can be seen in Fig. 4.8. There is clear differentiation between *A. gibbosa* and *A. phylactica*, however, there is a single *A. phylactica* outlier, which is closer to the *A. gibbosa* cluster than the *A. phylactica* clusters. The Croatian population (RAP) is separate from the main *A. phylactica* cluster, towards the Mediterranean end of the *A. gibbosa* cluster, suggesting that this population is highly distinct from the other populations. The Atlantic *A. phylactica* cluster appears to be more homogenous than the *A. gibbosa* cluster, which is quite elongated with Atlantic individuals prevalent at one end of the cluster and Mediterranean individuals prevalent at the other. The $\theta^B$ values gave a similar result to the FST data (data not shown).
In the Mediterranean basin, all individual multilocus genotypes except for one were re-allocated to their population of origin (Table 4.11). The exception was one of the *A. phylactica* individuals collected from Croatia, which was re-allocated to the Naples *A. gibbosa* population. There were no other instances of cross species re-allocation.

A high proportion of individuals were misassigned in the Atlantic basin, but none were assigned to the other species. The misassigned individuals were generally re-allocated to populations that were geographically close to their original location. Although both *A. gibbosa* and *A. phylactica* were collected from the same geographic location in two instances there appears to be no evidence for hybridisation, as the t-tests showed that the two species are significantly different at the shared locations and no individuals were misassigned (Chapel Point, $t = 3.8655$, $p < 0.001$; Prawle Point, $t = 3.7074$, $p < 0.003$).

### Discussion

Comparative studies using multiple lineages and genetic markers, such as this, can be used to estimate the phylogenetic correlation between life history and population
genetic structure, and the residual contributions of other (geological or ecological) processes in shaping the evolution of population genetic structure.

The most striking results seen in this study are (i) the genetic differentiation between A. gibbosa and A. phylactica; (ii) the genetic isolation of the Croatian A. phylactica population; (iii) the significant differentiation between populations in the Mediterranean Sea and those of the Atlantic Ocean, seen with the AFLP marker; (iv) the high level of endemism of the Mediterranean alleles; (v) the presence of a common haplotype observed throughout the ranges of both A. gibbosa and A. phylactica; (vi) the presence of northern refugia during the LGM, and; (vii) the discordant results observed between the two molecular markers.

The analysis of the mtDNA sequences suggests that there is incomplete lineage sorting between A. gibbosa and A. phylactica, with two haplotypes shared across both species. There are a number of haplotypes private to both species (eight and five, in A. gibbosa and A. phylactica respectively), with the mtDNA AMOVA data attributing 55% (P <0.01) of the variation observed to the difference between A. gibbosa and A. phylactica, which is supported by the AFLP data where the AMOVA data (72% of variation observed is between A. gibbosa and A. phylactica), the PCA analysis and paired t-tests all indicate that the two species are reproductively isolated, or inter-species hybridisation is rare or non-existent. Indeed, the θB mean pairwise value is doubled when analysing A. gibbosa and A. phylactica as a single species rather than within either of the putative species. These results indicate that a lineage split has occurred, with the two species being isolated from each other followed by secondary contact. Genetic differences have not been completely manifested within the mtDNA sequences but are evident with the AFLP analysis with no confirmed AFLP introgression. Some mitochondrial introgression is common between closely related species, but it is rare for introgression to occur throughout the entire distribution range of reproductively distinct species (Kemppainen et al., 2009), such as A. gibbosa and A. phylactica. A lack of mitochondrial divergence has mostly been observed in local populations, along hybrid zones (e.g. Ruedi et al., 1997; Wilson and Bernatchez, 1998; Ruber et al., 2001; Melo-Ferreira et al., 2005; Roca et al., 2005; Berthier et al., 2006), where contemporary hybridization is known
to occur (e.g. Ruedi et al., 1997; Bachtrog et al., 2006; Carson and Dowling, 2006) or in endemic species with limited geographic distributions (Bachtrog et al., 2006; Carson and Dowling, 2006). Incomplete lineage sorting of mtDNA between two non-hybridizing species may be the result of a recent divergence and/or the effective population size is too large for ancestral polymorphism to have sorted to reciprocal monophyly (Funk and Omland, 2003). Incomplete mtDNA lineage sorting between closely related but reproductively isolated species has also been reported in the sea stars Parvulastra dyscrita and Parvulastra exigua (Dunbar, 2006) and the periwinkles L. fabalis and L. obtusata (Kemppainen et al., 2009), in both cases this was attributed to a recent divergence between the two lineages.

The results show that within the mtDNA there are two haplogroups, both of which contain individuals from both A. gibbosa and A. phylactica. Within both haplogroups, there is a pattern of low nucleotide diversity and high haplotype diversity, strongly negative Fu’s F and Tajima’s D values, this together with star shape haplotype networks indicates that both haplogroups have undergone a range expansion, which was estimated to be a far older date than the LGM, at between 320 kya to 361 kya for haplogroup 1 and maybe as long ago as 640 kya to 722 kya for haplogroup 2, placing both of the expansions within the Pleistocene epoch. These dates are comparable to dates for range expansions of the sea stars Linckia laevigata and Protoreaster nodosus in the Coral Triangle (Crandall et al., 2008), during a time of large fluctuations in sea level.

Both A. gibbosa and A. phylactica show a higher level of haplotype diversity in the Mediterranean Sea than in the Atlantic Ocean. Genetic diversity patterns are frequently employed to delineate potential routes for a species range expansion (Couceiro et al., 2007). There should be less genetic diversity in younger populations than that found in older, more established populations, as older populations have a greater length of time to accumulate mutations. In areas affected by the last LGM many intertidal species exhibit more genetic differentiation in the southern part of the range when compared with populations further north (Patarnello et al., 2007), which is usually attributed to founder events during range expansion following deglaciation (Hewitt, 2000). However, other studies have shown the opposite to be true (Coyer et al., 2003; Olsen et al., 2004; Provan et al., 2005). Alternatively, the Mediterranean is
characterized by local upwelling, eddies, other current patterns and isotherms that reinforce isolations among populations, in addition to the physical barriers imposed by complex shorelines (Magoulas et al., 2006) as well as numerous small islands. The amplitude of tidal currents is low in the Mediterranean, this could reduce the amount of exchange between the littoral and sublittoral environments, resulting in a lower number of instances of juveniles or adults drifting or rafting to other populations, allowing drift to occur.

The results show that within population diversity is low for both A. gibbosa and A. phylactica, with single haplotypes recovered from individual populations in many cases. The lack of mtDNA differentiation detected between populations possessing only the most dominant haplotypes, particularly in the Atlantic, is likely to reflect the colonisation of these haplotypes, rather than high levels of gene flow. This observation was also made for the intertidal snail Zeacumantus subcarinatus within the Otago region of New Zealand (Keeney et al., 2009). Despite this lack of mtDNA differentiation, most of the genetic diversity occurred within and not between populations for both species, in both basins for A. gibbosa and just in the Atlantic for A. phylactica (we were unable to assess the Mediterranean due to only one population being sampled).

The A. gibbosa AFLP data contrasts with the mtDNA structure, which showed that haplogroups are not restricted to any particular geographical area. MtDNA is useful when inferring past processes, but the genomic approach using AFLP reflects present day gene flow and as such can highlight very different demographic processes. The AFLP analysis suggests that gene flow is more restricted for A. phylactica than A. gibbosa and, for A. gibbosa, it is more restricted in the Mediterranean than the Atlantic. The current study provides evidence of dispersal distances of A. gibbosa and A. phylactica up to a few hundred kilometres around the coasts of the North East Atlantic, with sea stars remaining close to their natal populations and those adjacent to them. This dispersal capability is reflected in the isolation by distance genetic structure evident but does not indicate that long distance dispersal (above a few hundred kilometres) regularly occurs.
Genetic homogeneity spanning large geographic distances has been reported for the brooding sea star *A. squamata* which displays several haplotypes spread over populations separated by 1000 km in New Zealand (Sponer and Roy, 2002), another brooding sea star *Astrotoma agassizii* was found to exhibit genetic homogeneity distances greater than 500 km within the Antarctic continental shelf (Hunter and Halanych, 2008), two brooding amphipods *Leucothoe kensleyi* and *Leucothoe ashleyae* showed a lack of genetic structure across 355 km along the coast of Florida (Richards *et al.*, 2007) and *Z. subcarinatus* which has several haplotypes shared among distant regions (Keeney *et al.*, 2009).

A lack of regional genetic differentiation does not necessarily mean that there is frequent dispersal across the barrier but rather reflects the lack of any incremental effect of the barrier upon species that exhibit high levels of variation within regions (Ayre *et al.*, 2009). The AFLP data does suggest that modern day gene flow is not very pronounced, with genetic structure evident and with most individuals being reassigned to their parent population in the re-allocation test, although there is likely to be some gene flow occurring. This is the expected population structure for a marine invertebrate with a life history without a high-dispersal larval stage. Even with the mtDNA sequences, there are a number private haplotypes within some populations, which provides a small amount genetic structuring.

There is a marked difference within the AFLP data between the Atlantic and Mediterranean, although the sampling intensity is unable to pinpoint where the barrier to gene flow is. There is a highly documented barrier at the Strait of Gibraltar/AOF (see Paternello *et al.*, 2007 for review). There are many mesoscale processes caused by the incoming Atlantic waters which could be responsible for further genetic differentiation both within the Mediterranean and between the Atlantic and Mediterranean, as the $\delta^{18}$ differences between populations within the Mediterranean are as great as they are between populations found in the Atlantic and Mediterranean. Large differences between populations in the Mediterranean are not confined to these species, the brooding brittle star *Amphipholis squamata* also displayed strong differentiation and a positive correlation between geographical and genetic distances within lineages in the Mediterranean (Boissin *et al.*, 2008). Within
the range of *A. phylactica* there is another potential phylogeographic break at the Strait of Sicily, where there is strong local currents caused by the bathymetry of the region (Mejri *et al.*, 2009). A break has been observed here between flounder (Borsa *et al.*, 1997), sea bass (Bahri-Sfar *et al.*, 2000), sandy goby (Stefanni and Thorley, 2003), mackerel (Zardoya *et al.*, 2004) and Tortonese’s goby (Mejri *et al.*, 2009). The differences seen with the AFLP data is less pronounced with the mtDNA sequences, with only a small percentage of the variation being attributed to geographical location - there is less than 5% of the variation attributed to differences between the Atlantic and Mediterranean. However, this may be caused by the frequency of the most common haplotype within each species, which present in all of the populations. There are four and six haplotypes exclusive to the Mediterranean and Atlantic respectively with only six shared haplotypes indicating that there is genetic differentiation between the two basins for the mtDNA sequences.

Some of the AMOVA results gave non-significant results, despite showing apparent high levels of genetic structure. Nested (hierarchical) randomisation tests, such as AMOVA will not reject the null hypothesis of no group structure if there are insufficient numbers of populations per group (Fitzpatrick, 2009). Overall, it appears that the only major barriers occur between the Western Mediterranean and the Adriatic and the Atlantic Ocean and Mediterranean Sea, although contemporary gene-flow may be restricted as a result of the English Channel and the entrance to Lough Hyne. Haplotypes are shared between specimens from remote locations, which may be unexpected for species such as these that do not have a life history with a high potential for long distance dispersal.

Due to the developmental mode of both species it is likely that the mechanism for dispersal is drifting or rafting on different substrata of natural (wood, seagrasses, macroalgae, volcanic pumice, corals) and anthropogenic (plastics, tar balls, manufactured wood) origin (Thiel, 2003). Macroalgae are generally considered to be poor dispersers across ocean basins (Hoek, 1987), successful dispersal on drifting macroalgal thalli tends to be an uncommon and opportunistic method of dispersal (Hoek, 1987), which is often unsuccessful as they are seldom transported to suitable habitat (Phillips, 2001), it can be inferred that this is also the case for other drifting substrata. Rafting on macroalgae or other substrata as a mechanism for long-
distance dispersal has been proposed for several other marine invertebrates e.g., the oyster *Ostrea chilensis* (O’Foighil et al., 1999), sea star *Patiriella exigua* (Waters and Roy, 2004), gastropods *Diloma* spp. (Donald et al., 2005) and mussels *Perna* spp. (Wood et al., 2007).

Haplotypes shared among distant populations may represent either very low levels of contemporary gene flow or historic long distance dispersal coupled with incomplete lineage sorting. This means that dispersal via drifting or rafting must have occurred over large geographical ranges, unless there are a number of unsampled intermediate populations or there has been an extirpation of intermediate populations (Keeney et al., 2009). The patterns observed are likely to indicate the stochastic nature of rafting and its capacity to allow rare dispersal over short and long distances and between any pair of populations. Indeed, species that are dependent upon rafting typically display considerable geographic differentiation but only weak relationships between genetic and geographic distance (Waters et al., 2004; Thiel and Haye, 2006).

Strong population structure has been found in asterinid species with benthic development of brooded larvae (Matsuoka and Asano, 2003; Waters et al., 2004; Baus et al., 2005; Colgan et al., 2005; Sherman et al., 2008) as well as asterinid species with planktotrophic development (Waters and Roy, 2004). A study across a phylogenetic break in Australia found that life history could not predict connectivity between populations on both sides of the break. Clear phylogeographic breaks were found for only four of the six species that span the barrier and yet have planktonically dispersed larvae, however, the direct developing asterinid *Parvulastra exigua* and *Haustrum vinosa*, did not show any effect of the barrier in either the study with a COI marker (Ayre et al., 2009) or with allozyme surveys (Sherman et al., 2008). These studies suggest that modern day barriers to gene flow as well as historical processes might have a large influence on population genetic variation in asterinids and other marine invertebrates (Keever et al., 2009).

The intertidal communities around the British Isles consist of species that have persisted in northern glacial refugia during Pleistocene glacial maxima and species who’s range contracted and expanded in response to the glacial and interglacial
periods (Jolly et al., 2006). Within the range of A. gibbosa and A. phylactica there are four reported areas of refugia from the LGM, these are located in the Iberian Peninsula (Hoarau et al., 2007), the Mediterranean Sea (Patarnello et al., 2007), the Western English Channel (Provan et al., 2005; Hoarau et al., 2007) and Southwest Ireland (Provan et al., 2005; Hoarau et al., 2007). The evolutionary rate of mtDNA implies that most polymorphisms involving more than one base difference are likely to have pre-dated the LGM (Anderson et al., 2006) therefore distributions of haplotypes between isolated geographical areas are expected to represent either longer-term isolation or the effects of lineage sorting by genetic drift, with a very limited contribution from haplotypes derived by new mutations (Maggs et al., 2008). Sympatric mtDNA clades are often interpreted as evidence for vicariance followed by reinvasion (Avise, 2000). With the recent geological history of the range of both species it is likely that more than one refugium were used during the LGM, followed by secondary contact.

The genetic structure of the Hartland Quay population of A. phylactica is quite distinct from the other populations, with both the mtDNA and AFLP markers. The mtDNA data shows that there are two private haplotypes present within this population. The populations along the English Channel also show high levels of genetic diversity relative the populations further west and north (with the exception of the Hartland Quay population).

The Lough Hyne population of Southwest Ireland contains a haplotype common with the three French populations and is a fairly diverse population relative to most, containing three haplotypes. Within the Western English Channel, there is a haplotype present which is found otherwise only in the Naples population and there are private haplotypes, suggesting that some of these populations maybe older than the 10 kya expected if there was a range expansion after the LGM.

The Croatian population of A. phylactica is very different to all of the other populations sampled. A cluster of closely related private haplotypes suggests an extended period of genetic isolation and a recovery after a low effective population size, with the accumulation of new mutations that are one or two mutational steps away from a frequent ancestor. It is possible that this population was founded by a
very small number of individuals, or even a single organism or egg mass, which migrated northwards from the Mediterranean and became trapped in the cul-de-sac of the Adriatic Sea by the extensive gyres and high surface temperatures in the Southern Adriatic (although it is likely that these temperatures should not be an issue for an organism which inhabits intertidal rock pools which are subject to a wide range of temperature fluctuations). Alternatively, the north Adriatic Sea may have provided a refugia with the population becoming genetically isolated due to the barriers to gene flow found in the Adriatic and Mediterranean Sea’s. The AFLP data also shows the Rovinj population to be distinct from both the *A. phylactica* and *A. gibbosa* populations. There is a single *A. phylactica* individual (from Croatia) grouped with *A. gibbosa* in the PCA analysis, with the same individual being the only individual to be re-allocated across the species barrier. This could be the result of introgression or, more likely, the individual could have been incorrectly classified as *A. phylactica* at the time of sampling.

**Discordant Molecular Markers?**

It appears that discordant results have been obtained between the two molecular markers employed in this study. This has been observed in other studies (e.g. Ujvari *et al.*, 2005; Kawakami *et al.*, 2007; Tzika *et al.*, 2008; Ujvari *et al.*, 2008). In plants, selective sweeps and/or rapid demographic expansion on chloroplast DNA have been suggested as important mechanisms for causing lower population subdivision of maternal inherited markers compared to paternal or biparental inherited genes (Muir and Filatov, 2007). The results produced in this study suggest that the mtDNA had been subjected to a selective sweep or demographic expansion, so these scenarios may explain the lower population subdivision seen with mtDNA compared to the AFLPs.

In conclusion, the present study shows that present oceanographic processes and paleoecological history (e.g., glaciations) have a role in shaping the genetic variability and population structure of *A. gibbosa* and *A. phylactica*. Despite having a developmental mode with a low potential for dispersal, both species have haplotypes which are found throughout the species range, in both the Atlantic and Mediterranean basins, across a geographic range of >5000 km indicating that dispersal occurs probably through drifting or rafting on various substrates, as
reported for other marine invertebrates, although gene flow is likely to be restricted. The AFLP analysis suggests that gene flow is more restricted for *A. phylactica* than *A. gibbosa* and, for *A. gibbosa*, it is more restricted in the Mediterranean than the Atlantic. Dispersal within the North East Atlantic does occur up to a few hundred kilometres, however sea stars usually remain close to their natal populations and those adjacent to them. The data presented here confirms that *A. gibbosa* and *A. phylactica* are distinct species, although there is incomplete mtDNA lineage sorting between the two species as has been found in the sea stars *P. dyscrita* and *P. exigua* (Dunbar, 2006) and the periwinkles *L. fabalis* and *L. obtusata* (Kemppainen et al., 2009). Within both the species there are deeply divergent haplogroups, which are shared across both species, throughout the geographical range. The lack of evidence for a recent range expansion around the time of the LGM in either species coupled with the observed patterns of haplotype distribution suggests that *A. gibbosa* and *A. phylactica* survived the LGM in one or more northern refugia. In addition to this, cryptic diversity has been found within *A. phylactica* found at Rovinj in Croatia.

### 4.6 References


Riginos C, Nachman MW (2001) Population subdivision in marine environments: the contributions of biogeography, geographical distance and discontinuous habitat to


Chapter 5

Discussion
As the first genetic analysis of North East Atlantic and Mediterranean populations of the genus *Asterina*, this study has both resolved a number of key problems and identified new questions about the phylogeny and phylogeography of this ecologically relatively well studied sea star taxon. For example, this is the first study to address the genetic distinctiveness and comparative phylogeography of *Asterina gibbosa* and *Asterina phylactica*. Furthermore, a localised, divergent, and potentially important population for conservation within *A. phylactica* is identified. The following sections aim to synthesise the data from Chapters 2 - 4 and highlight gaps in our knowledge and key questions that remain to be answered.

Comparative studies such as those carried out in this thesis, potentially allow researchers to evaluate the effects of shared historical biogeographical processes in driving the evolution and the regional distribution of biodiversity (Bermingham and Moritz, 1998; Avise, 2000). In general, *A. phylactica* and *A. gibbosa* have been shown to exhibit concordant patterns of divergence across the Atlantic Ocean and Mediterranean Sea, with a strong signal of a shared demographic history.

### 5.1. *A. gibbosa* versus *A. phylactica*

Both molecular approaches employed in this study suggest that *A. gibbosa* and *A. phylactica* comprise two evolutionarily divergent lineages. The mtDNA data showed that the two most common and probably ancestral mitochondrial haplotypes are shared between the species, indicating that mtDNA lineage sorting is not yet complete, and only a relatively modest 55% of the molecular variance observed with the AMOVA test was attributable to differences between the two species. Incomplete mtDNA lineage sorting between closely related putative littoral invertebrate species has recently been reported in the sea stars *Parvulastra dyscrita* and *Parvulastra exigua* (Dunbar, 2006) and the periwinkles *L. fabalis* and *L. obtusata* (Kemppainen et al., 2009), and is most likely a result of recent divergence between the two lineages. In support of this inference, the AFLP data showed strong nuclear DNA evidence that these are distinct reproductively isolated species. AMOVA showed that 72% of the variation seen was partitioned between the two species and there was no molecular evidence of hybridisation, even in the locations where both species were
present, often in the same rock pools. It can therefore be concluded that *A. gibbosa* and *A. phylactica* are likely to be reproductively isolated and therefore, under the traditional Biological Species Concept, seem to conform with the expectations of distinct species.

5.2. **Putative *A. phylactica* population at Rovinj, Croatia.**

The two different genetic approaches (mtDNA and AFLP) both showed signs of divergence between the Rovinj population in Croatia and both the *A. gibbosa* and *A. phylactica* lineages. This was the only population sampled in the Adriatic Sea and is at least 1,900 kilometres from the nearest *Asterina* population sampled. The results show that this population is distinct from populations of both *A. gibbosa* and *A. phylactica*. The three mtDNA haplotypes observed at this population were found within the haplogroup dominated (albeit not exclusively) by *A. phylactica*, indicating that if secondary contact has occurred within *A. phylactica* between the two haplogroups, it has probably not occurred at Rovinj. The two of the three haplotypes are private to this population, with one haplotype separated by the most frequent haplotype found in *A. phylactica* by two and the second being separated by four substitutions. The AFLP data shows that the Rovinj population has by far the greatest level of polymorphisms of all of the *A. phylactica* populations, with 40%, however a proportion of this diversity may be attributed to differences between the majority of the individuals found at the site and an “outlier” found at that population.

It is unlikely that this population has arisen as the result of a recent founder event due to the high levels of genetic heterogeneity found at this population. Both datasets indicate that this is not a recently established population. Although the mtDNA haplotypes are within the haplogroup dominated by *A. phylactica* the AFLP data shows the population is significantly different to the other *A. phylactica* populations, although for it to represent a common ancestral sequence would require a major shift of haplotype frequencies in the descendant population and concurrent elimination of the two other haplotypes present. The current genetic data provides strong evidence for cryptic diversity within this species. Since both haplogroups are found in both species, genetically individuals from this population could be inferred
as being as closely related to *A. gibbosa* as they are to *A. phylactica*, although morphologically they resemble *A. phylactica*.

Considering the isolated location of this population and the physical boundaries caused by currents, eddys and gyres within the Mediterranean and Adriatic seas, it is likely that the genetic composition of the population has evolved in recent allopatry to other populations. For full allopatric speciation to occur there are at least three stages – (1) an isolated population must form - that is, become separated from the parental population; (2) it must persist long enough to allow genetic differentiation into a new species; and (3) it must actually undergo that differentiation (Allmon, 1992, 1994; McKinney and Allmon, 1995; Allmon et al., 1998). Currently, it is not known whether this population is reproductively isolated from either *A. gibbosa* or *A. phylactica* and this question deserves further investigation.

It might be expected that other isolated pockets of cryptic diversity may exist within the Adriatic or eastern Mediterranean but these populations are either absent (unlikely) or remain undetected (highly probable). Further surveying and sampling of populations should occur to analyse the geographic distribution of such populations. It was noted that during the time of the field visit to Croatia that many market stall holders were selling dried sea shore fauna, such as sea stars, as souvenirs and this may have contributed to the small number of individuals observed at Rovinj.

### 5.1. *A. phylactica* Population at Hartland Quay

The Hartland Quay population of *A. phylactica* was found to be distinct from the other populations due to the presence of a private mtDNA haplotype, which is the dominant haplotype in the population. In addition to this frequent, dominant haplotype there is also greater than normal haplotype diversity, with four haplotypes being identified in this population. However, the AFLP data show that there is gene flow between this and other populations on the southwest coast of England.

The differences observed between this population and many of the other populations could be the result of the possibility that this population survived the LGM in a
different refugium to the other populations. The remnants of populations which have glacial refugia may be inferred from the levels of divergence and population subdivision. Refugia for northern species suggested are the Iberian Peninsula (Gysels et al., 2004a; Hoarau et al., 2007), the Mediterranean Sea (Olsen et al., 2004; Sa-Pinto et al., 2005), the Bay of Biscay (Nesbø et al., 2000), the south-western coast of Ireland (Jolly et al., 2006; Hoarau et al., 2007) and Hurd Deep in the English Channel (Coyer et al., 2003; Provan et al., 2005). An alternative explanation maybe a mixture of chance related to the sampling or due to local current patterns. The observed pattern seen could be the result of currents bringing together an unusual combination of haplotypes which have converged at this location, the same currents then act as a barrier to individuals successfully dispersing from this population.

5.2. Dispersal

Some haplotypes were shared between specimens from locations separated by distances in excess of 5000 km. These findings are perhaps surprising considering the hypothesised restricted ability to disperse of directly developing species. It was expected that there is strong genetic structure associated with geographical location because of the restricted dispersal ability and therefore the lower potential for long distance gene flow to occur. However, recently high levels of gene flow resulting in mtDNA homogeneity was demonstrated across a 500 km range throughout the Antarctic Peninsula for the brooding sea star Astrotoma agassizii, also an unexpected result given the brooding nature of this species according to the authors (Hunter and Halanych, 2008).

Within the Atlantic, many Asterina populations were found to contain low levels of haplotype diversity, typically comprising of just one or two haplotypes. Recent founder effects, due to range expansion following the LGM may partially explain the lack of diversity within and among these sites, although the dominance of the most common haplotype within the region also increases the likelihood of colonizing individuals possessing it. The most likely method for dispersal for organisms such as
A. gibbosa and A. phylactica which do not have a life history strategy which promotes dispersal, is rafting on natural and anthropogenic floating materials. Sponer and Roy (2002) found that within the brooding brittle star Amphipholis squamata, several haplotypes were spread over populations separated by up to 1000 km, with this range only possible through sporadic long-distance dispersal through passive transport by rafting or drifting on macroalgae.

The observed spatial genetic patterns among A. gibbosa and A. phylactica localities are also likely to be influenced by oceanographic currents and fronts. Oceanic currents are responsible for both dispersal and gene flow on a large scale, but they can also act to promote genetic differentiation between populations by causing physical barriers (Quinterio et al., 2007; Palumbi, 1994). Oceanic fronts, sharp discontinuities of physical and biochemical variables, are generated by various physical processes, occur in all oceans, and are likely to represent barriers to faunal exchange (Millot, 2005). Within the North East Atlantic and Mediterranean Sea there are complicated oceanic current patterns that can potentially isolate populations, increasing the tendency to self-recruit and acting as a barrier to gene-flow.

Strong tidal currents, especially in the English Channel, may facilitate gene flow within species such as A. gibbosa and A. phylactica. If present-day gene flow between the Atlantic populations is maintained by passive rift or rafting via currents, limited genetic differentiation on a small scale, along with a cline of isolation-by-distance on a larger geographical scale throughout the north-eastern Atlantic basin would be expected.

The AFLP data confirms this to be the case in the Atlantic. In both A. gibbosa and A. phylactica there is strong evidence for gene flow between local populations with less gene flow as the distances increase. The population assignment tests revealed that, despite its relative isolation from oceanic waters, the population inhabiting Loch Hyne in Ireland appears to have maintained some level of gene flow with Cornish populations (Baus et al., 2005). The one surprising result is at Rockham Bay population located on the coast of north Devon. This population possesses as many individuals re-allocated to the Prawle Point population (situated on the south coast) as it does to itself and this observation coupled with the low amount of pairwise
genetic differentiation between the two populations, the mitochondrial DNA data which shows a single, dominant haplotype (AG1 the most common haplotype) and the oceanic current patterns, it is likely that this population has experienced or periodically experiences an influx of individuals from the Prawle Point population. Thus it does not appear that the North Atlantic Current (NAC) which flows eastward along the English Channel from Plymouth prevents individuals rafting from South Devon along the coast past a potential barrier at the Lizard peninsula.

In the Mediterranean the AFLP data shows gene flow between the populations to be almost absent, even in populations that are geographically close, for example, the Banyuls-sur-Mer and Les Embiez populations are separated by just 225 km but the pairwise differentiation is 0.344 (the Banyuls-sur-Mer and Rockham Bay population result is only slightly higher with 0.386 despite being separated by >3000 km) and no individuals are misallocated between populations. Retention of individuals within these populations appears to be high, probably as a result of the differing hydrographic properties of the two basins. The Mediterranean consists of a number of semi-enclosed circulation systems and shelf areas (Agostini and Bakun, 2002) and has small tidal currents, whereas Atlantic continental shelf is characterized by the NAC, and moderate to strong tidal currents. Atlantic populations may therefore be more prone to effective migration than those in the Mediterranean.

In contrast, mtDNA haplotypes are shared across the entire geographic ranges of both species it is more difficult to discern any patterns. There some notable exceptions, for example the Rovinj and Hartland Quay populations as discussed previously. There is an observable pattern when looking at the two haplogroups, the prevalent haplogroup of A. gibbosa across the range is haplogroup 1, in A. phylactica the dominant haplogroup throughout the range is haplogroup 2. The presence of sympatric mtDNA clades within species has been interpreted as evidence for vicariance followed by reinvasion and secondary contact (Avise, 2000). This has been inferred for Atlantic mackerel (Nesbo et al., 2000; Zardoya et al., 2004), swordfish (Bremer et al., 1995, 2005; Buonnacorsi et al., 2001; Graves and McDowell, 2003), blue marlin (Buonnacorsi et al., 2001), sailfish (Graves and McDowell, 2003), Atlantic bonito (Viñas et al., 2004), Atlantic big eye (Martínez et al., 2006) and Scabbardfish (Stefanni and Knutsen, 2007). The incomplete lineage
sorting observed between *A. gibbosa* and *A. phylactica* is likely to be the result of the two species being isolated from each other followed by secondary contact. Mitochondrial introgression is common between closely related species, including the sea stars *P. dyscrita* and *P. exigua* (Dunbar, 2006).

Genetic differences have not been completely manifested within the mtDNA sequences but are evident with the AFLP analysis with no confirmed AFLP introgression. Some mitochondrial introgression is common between closely related species, but it is rare for introgression to occur throughout the entire distribution range of reproductively distinct species (Kemppainen *et al.*, 2009).

Generally, there is very little genetic differentiation between populations within each of the two basins, with the vast majority of genetic variance occurring within populations as opposed to between populations. This observation is in agreement with many other genetic studies of marine species with a low dispersal potential, including the brooding sea star *Astrotoma agassizii* (Hunter and Halanych, 2008) and the brooding brittle star *Amphipholis squamata* (Sponer and Roy, 2002).

However, there do appear to be some phylogeographic breaks in the data, possibly associated with physical barriers throughout the range of these species. For example, in the Atlantic, 84% of *A. gibbosa* individuals contain the mtDNA haplotype AG1, whereas the proportion in the Mediterranean containing this haplotype is just 29%. In the Atlantic, 67% of *A. phylactica* individuals contain the haplotype AP1, whereas the proportion in the Mediterranean containing this haplotype is 17%. The Strait of Gibraltar is a break for many species (see Patarnello *et al.*, 2007) including *A. gibbosa* (Baus *et al.*, 2005) but it is unclear if this acts as a break for *A. phylactica* due to the locations of the sampled populations. There is a clear barrier between the Rovinj population and the populations on the coast of the UK but the exact location is unknown due to a lack of sampling sites in the western Mediterranean. In the Mediterranean, restricted gene flow between populations across the Siculo-Tunisian barrier in the caramote prawn *Penaeus (Melicertus) kerathurus* (Zitari-Chatti *et al.*, 2009) has been observed. Populations in the Adriatic may be isolated as a result of gyral circulations (Artegiani *et al.*, 1993), the planktonic chaetognath *Sagitta setosa* was found to have a genetic break between populations in the Mediterranean and
the Adriatic Sea (Peijnenburg et al., 2004). Likewise, the English Channel was found to act as a barrier to gene flow in the prawn *Palaemon elegans* separating Great Britain from northern France (Reuschel et al., 2010) and in the common goby *Pomatoschistus microps* with distinct haplotypes dominating at either side of the English Channel (Gysels et al., 2004a).

### 5.3. The effect of the Pleistocene Glaciations

The earth's climate in the recent past appears to have warmed and cooled in cycles of approximately 100,000 years, with the last ice sheet retreating approximately 10,000 years ago (Webb and Bartlein, 1992). Cycles of glacial advance and retreat from the late Pliocene to the Pleistocene epoch (which began 1.8 Mya) are arguably the most important climatological events during the evolutionary life span of most extant species (Hewitt, 2000), with major changes in the coastal marine environment, including variation in sea surface and air temperatures and changes in sea level and coastal hydrography (Foltz, 2008), causing cycles of range contraction and expansion coupled in some instances with range fragmentation and vicariance into disjunct refugia (Larmuseau, 2009).

The current demography of Atlantic populations results mainly from the LGM as earlier demographic signals were probably blurred or even eradicated by this event. However, demographic expansions of marine species along the north-eastern Atlantic coast generally pre-date the LGM events, with expansions dated as between 1.7 and 0.11 Ma; for example algae (Provan et al., 2005; Hoarau et al., 2007; Calderón et al., 2008), polychaetes (Jolly et al., 2006), bivalves (Luttikhuizen et al., 2003), urchins (Calderón et al., 2008), crustaceans (Stamatis et al., 2004), rays (Chevolot et al., 2006) and teleosts (Gysels et al., 2004a; Aboim et al., 2005; Bremer et al., 2005; Charrier et al., 2006). The time since expansion for the two haplogroups for *A. gibbosa* and *A. phylactica* is within this range with a range expansion occurring at 341,000 – 681,000 and 640,000 – 722,000 for haplogroups 1 and 2 respectively.

As a result of the slower evolutionary rates, most polymorphisms within mtDNA will likely pre-date the LGM (Anderson et al., 2006). Therefore the lineage split seen is
much older than the LGM. The differential distributions of haplotypes seen within the ranges of both *A. gibbosa* and *A. phylactica*, and indeed between both species, represent longer-term isolation. This lends some weight to the hypotheses that the Hartland Quay population descends from one that may have existed in separate refugia to the rest of the *A. phylactica* populations sampled. The Pleistocene glaciation events that occurred in the North East Atlantic may have allowed species with a once continuous distribution to become fragmented in pockets of warmer currents which could allow allopatric speciation events to occur. Secondary contact could have ensued causing geographic introgression but not genetic introgression.

The strongly negative neutrality test values, together with multiple star-polytomies and mismatch distributions indicate a recent range expansion has occurred probably in response to the Pleistocene glacial episodes.

### 5.4. Brooding and Colonisation

It has been hypothesised that brooding species may, counterintuitively, have a high potential for successful range expansion as new areas could be colonised by just a single female with a large clutch of brooded young as a “seed population” (Higgs *et al.*, 2009). Brooders tend to be smaller than closely related species with a non-brooding life history strategy. Small size is an advantage for shallow water brooding species with respect to dispersal by floating or rafting and small animals are more likely to live on potential rafting substrata before it becomes detached (Cheetham, 1960; Highsmith, 1985). However, here there is little evidence to suggest that the brooding behaviour of *A. phylactica* has conferred an advantage or given rise to a different phylogeographic pattern to the strategy employed by *A. gibbosa* as there are similar levels of genetic structuring within both species.

### 5.5. Mitochondrial DNA Data versus Nuclear AFLP Data

The results obtained show that the mtDNA and AFLP data are not completely congruent. Mitochondrial DNA is regarded as a useful tool for examining
phylogeographic structure however, when using it to compare geographic locations or species it relies on intraspecific genetic variation being much less than interspecific genetic variation (Dasmahapatra et al., 2010). Mitochondrial DNA reflects evolution only of a single, non-recombining, maternally inherited genome, which can be affected by factors such as coalescent stochasticity, interspecific hybridization and selection, whereas AFLP markers give a genome wide perspective of evolution and on average are less likely to be affected by selection signatures. For example, in the butterfly genus *Mechanitis*, Dasmahapatra et al. (2010) compared AFLP data to mtDNA sequences and found that there was a lack of agreement between the population structures detected by the two sets of markers. It was found that the mtDNA was unable to reveal any correlation between wing patterning and mtDNA haplogroup, whereas the AFLP genetic clusters were correlated with wing phenotype (Dasmahapatra et al., 2010).

5.6. Conservation Zones

One of the key aims of this study was to find any populations, areas or regions that are of particular importance which may help in the planning of conservation strategies which enable the creation of marine protected areas and parks. Although the conservation of all populations is important for a species such as *A. phylactica*, an uncommon intertidal species, it is recognised that only a small number are ever likely to receive the protection they require. The two key populations that I would recommend to receive protection (along with one or more of the other populations containing individuals from both mtDNA haplogroups) are the populations at Hartland Quay, north Devon and Rovinj on the Istrian Peninsula of Croatia because of their uniqueness at a genetic level. For the same reasons I recommend that the *A. gibbosa* populations found at Naples, Italy; Banyuls-sur-Mer, France; Prawle Point and Chapel Point, both south Devon should receive conservation status. Perhaps the biggest puzzle, especially for *A. phylactica*, is the disjunct nature of its distribution. So perhaps the most powerful recommendation for this species is that more surveys should be carried out along the coast of the north east Atlantic and the whole of the Mediterranean, which could be potentially carried out by the various marine stations found in this region. In this way a more realistic and accurate
knowledge of this enigmatic sea star can be obtained and its conservation status can be assessed more rationally.

5.7. Further Work

As mentioned above, future work should focus on surveys and “filling in the gaps” within the sampling range, with sampling along the coasts of the Iberian peninsula, along the Mediterranean French coast, the Italian coast and the Adriatic Sea. In addition to this, large scale genome-wide molecular markers such as SNPs (Single Nucleotide Polymorphisms) should be utilised to establish the extent and time of divergence between the species tested and to confirm the lack of population differentiation found on the coast of the Atlantic Ocean.

This body of work has focussed on intertidal populations of *A. gibbosa* and *A. phylactica*. The extent of gene flow that occurs between littoral and sub-littoral populations is unknown and the influence of sub-littoral populations upon the phylogeography of the littoral populations for these species. A study should be conducted to examine these issues.

The data collected in this study suggests that there is no hybridization occurring between *A. gibbosa* and *A. phylactica* even when individuals co-inhabit rock pools. A laboratory based study should be conducted to examine if hybridization is possible between the two putative species or to confirm that they are reproductively isolated.

The cryptic diversity discovered at Rovinj should be studied to investigate the morphology and the life history of individuals from this population identified, as this population was sampled in September it is unknown if this species broods its eggs like *A. phylactica* or leaves them unattended like *A. gibbosa*.
5.8. References


211


**APPENDIX 1**

**Figure A1. Alignment of the sequence for *Asterina gibbosa* for all haplotypes, with missing data marked as a dash.**

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>10</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>50</th>
<th>60</th>
</tr>
</thead>
<tbody>
<tr>
<td>AG01</td>
<td>GTGTCTCTCTGTTTTACCGCCATTTGCTTTTCATGGCTACACCGCAATAGAGTAG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AG02</td>
<td>GTTCTCTCTCTCCTTGTTTTACCGCCATTTGCTTTTCATGGCTACACCGCAATAGAGTAG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AG03</td>
<td>-----CTCCTCTCTCTCCTTTACCGCCATTTGCTTTTCATGGCTACACCGCAATAGAGTAG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AG04</td>
<td>-----TCCTCTCTCTCCTTTACCGCCATTTGCTTTTCATGGCTACACCGCAATAGAGTAG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AG05</td>
<td>-----CTCCTCTCTCTCCTTTACCGCCATTTGCTTTTCATGGCTACACCGCAATAGAGTAG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AG06</td>
<td>-----TCCTCTCTCTCCTTTACCGCCATTTGCTTTTCATGGCTACACCGCAATAGAGTAG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AG07</td>
<td>-----TCCTCTCTCTCCTTTACCGCCATTTGCTTTTCATGGCTACACCGCAATAGAGTAG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AG08</td>
<td>-----TCCTCTCTCTCCTTTACCGCCATTTGCTTTTCATGGCTACACCGCAATAGAGTAG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AG09</td>
<td>-----TCCTCTCTCTCCTTTACCGCCATTTGCTTTTCATGGCTACACCGCAATAGAGTAG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AG10</td>
<td>-----TCCTCTCTCTCCTTTACCGCCATTTGCTTTTCATGGCTACACCGCAATAGAGTAG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure A1 (continued) and Table:**

<table>
<thead>
<tr>
<th></th>
<th>61</th>
<th>70</th>
<th>80</th>
<th>90</th>
<th>100</th>
<th>110</th>
<th>120</th>
</tr>
</thead>
<tbody>
<tr>
<td>AG01</td>
<td>TATAAGGTAAACGAAGAATTTTGATTTCTAAAATATAAGTTCAACTCTTATCTTTCTA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AG02</td>
<td>TATAAGGTAAACGAAGAATTTTGATTTCTAAAATATAAGTTCAACTCTTATCTTTCTA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AG03</td>
<td>TATAAGGTAAACGAAGAATTTTGATTTCTAAAATATAAGTTCAACTCTTATCTTTCTA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AG04</td>
<td>TATAAGGTAAACGAAGAATTTTGATTTCTAAAATATAAGTTCAACTCTTATCTTTCTA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AG05</td>
<td>TATAAGGTAAACGAAGAATTTTGATTTCTAAAATATAAGTTCAACTCTTATCTTTCTA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AG06</td>
<td>TATAAGGTAAACGAAGAATTTTGATTTCTAAAATATAAGTTCAACTCTTATCTTTCTA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AG07</td>
<td>TATAAGGTAAACGAAGAATTTTGATTTCTAAAATATAAGTTCAACTCTTATCTTTCTA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AG08</td>
<td>TATAAGGTAAACGAAGAATTTTGATTTCTAAAATATAAGTTCAACTCTTATCTTTCTA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AG09</td>
<td>TATAAGGTAAACGAAGAATTTTGATTTCTAAAATATAAGTTCAACTCTTATCTTTCTA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AG10</td>
<td>TATAAGGTAAACGAAGAATTTTGATTTCTAAAATATAAGTTCAACTCTTATCTTTCTA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure A1 (continued) and Table:**

<table>
<thead>
<tr>
<th></th>
<th>121</th>
<th>130</th>
<th>140</th>
<th>150</th>
<th>160</th>
<th>170</th>
<th>180</th>
</tr>
</thead>
<tbody>
<tr>
<td>AG01</td>
<td>ATCAGAAAAACAGAAGATTTACACTCACACCCAAACCCCAAGGCTACATTTCTCATA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AG02</td>
<td>ATCAGAAAAACAGAAGATTTACACTCACACCCAAACCCCAAGGCTACATTTCTCATA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AG03</td>
<td>ATCAGAAAAACAGAAGATTTACACTCACACCCAAACCCCAAGGCTACATTTCTCATA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AG04</td>
<td>ATCAGAAAAACAGAAGATTTACACTCACACCCAAACCCCAAGGCTACATTTCTCATA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AG05</td>
<td>ATCAGAAAAACAGAAGATTTACACTCACACCCAAACCCCAAGGCTACATTTCTCATA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AG06</td>
<td>ATCAGAAAAACAGAAGATTTACACTCACACCCAAACCCCAAGGCTACATTTCTCATA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AG07</td>
<td>ATCAGAAAAACAGAAGATTTACACTCACACCCAAACCCCAAGGCTACATTTCTCATA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AG08</td>
<td>ATCAGAAAAACAGAAGATTTACACTCACACCCAAACCCCAAGGCTACATTTCTCATA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AG09</td>
<td>ATCAGAAAAACAGAAGATTTACACTCACACCCAAACCCCAAGGCTACATTTCTCATA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AG10</td>
<td>ATCAGAAAAACAGAAGATTTACACTCACACCCAAACCCCAAGGCTACATTTCTCATA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure A1 (continued) and Table:**

<table>
<thead>
<tr>
<th></th>
<th>181</th>
<th>190</th>
<th>200</th>
<th>210</th>
<th>220</th>
<th>230</th>
<th>240</th>
</tr>
</thead>
<tbody>
<tr>
<td>AG01</td>
<td>AACTTTTTCTCTGATTATATAT---TATATATAAAATAATATAACCCATGACCTAAACAGATG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AG02</td>
<td>AACTTTTTCTCTGATTATATAT---TATATATAAAATAATATAACCCATGACCTAAACAGATG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AG03</td>
<td>AACTTTTTCTCTGATTATATAT---TATATATAAAATAATATAACCCATGACCTAAACAGATG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AG04</td>
<td>AACTTTTTCTCTGATTATATAT---TATATATAAAATAATATAACCCATGACCTAAACAGATG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AG05</td>
<td>AACTTTTTCTCTGATTATATAT---TATATATAAAATAATATAACCCATGACCTAAACAGATG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AG06</td>
<td>AACTTTTTCTCTGATTATATAT---TATATATAAAATAATATAACCCATGACCTAAACAGATG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AG07</td>
<td>AACTTTTTCTCTGATTATATAT---TATATATAAAATAATATAACCCATGACCTAAACAGATG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AG08</td>
<td>AACTTTTTCTCTGATTATATAT---TATATATAAAATAATATAACCCATGACCTAAACAGATG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AG09</td>
<td>AACTTTTTCTCTGATTATATAT---TATATATAAAATAATATAACCCATGACCTAAACAGATG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AG10</td>
<td>AACTTTTTCTCTGATTATATAT---TATATATAAAATAATATAACCCATGACCTAAACAGATG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note: The sequence alignment for *Asterina gibbosa* shows the alignment of the sequence for all haplotypes, with missing data marked as a dash. The table provides the alignment details for specific positions across the sequence.*
<table>
<thead>
<tr>
<th></th>
<th>241</th>
<th>250</th>
<th>260</th>
<th>270</th>
<th>280</th>
<th>290</th>
<th>300</th>
</tr>
</thead>
<tbody>
<tr>
<td>AG01</td>
<td>ATTTTTTTTCTACTAAACACAAGGACATAGGTACTCTATATCTTATATTGGAGGCTGAGGC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AG02</td>
<td>ATTTTTTTTCTACTAAACACAAGGACATAGGTACTCTATATCTTATATTGGAGGCTGAGGC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AG03</td>
<td>ATTTTTTTTCTACTAAACACAAGGACATAGGTACTCTATATCTTATATTGGAGGCTGAGGC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AG04</td>
<td>ATTTTTTTTCTACTAAACACAAGGACATAGGTACTCTATATCTTATATTGGAGGCTGAGGC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AG05</td>
<td>ATTTTTTTTCTACTAAACACAAGGACATAGGTACTCTATATCTTATATTGGAGGCTGAGGC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AG06</td>
<td>ATTTTTTTTCTACTAAACACAAGGACATAGGTACTCTATATCTTATATTGGAGGCTGAGGC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AG07</td>
<td>ATTTTTTTTCTACTAAACACAAGGACATAGGTACTCTATATCTTATATTGGAGGCTGAGGC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AG08</td>
<td>ATTTTTTTTCTACTAAACACAAGGACATAGGTACTCTATATCTTATATTGGAGGCTGAGGC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AG09</td>
<td>ATTTTTTTTCTACTAAACACAAGGACATAGGTACTCTATATCTTATATTGGAGGCTGAGGC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AG10</td>
<td>ATTTTTTTTCTACTAAACACAAGGACATAGGTACTCTATATCTTATATTGGAGGCTGAGGC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>301</th>
<th>310</th>
<th>320</th>
<th>330</th>
</tr>
</thead>
<tbody>
<tr>
<td>AG01</td>
<td>CGGAATGGCGGAACACTGCAATGAGCTTGTTATCATAG</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AG02</td>
<td>CGGAATGGCGGAACACTGCAATGAGCTTGTTATCATAG</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AG03</td>
<td>CGGAATGGCGGAACACTGCAATGAGCTTGTTATCATAG</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AG04</td>
<td>CGGAATGGCGGAACACTGCAATGAGCTTGTTATCATAG</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AG05</td>
<td>CGGAATGGCGGAACACTGCAATGAGCTTGTTATCATAG</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AG06</td>
<td>CGGAATGGCGGAACACTGCAATGAGCTTGTTATCATAG</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AG07</td>
<td>CGGAATGGCGGAACACTGCAATGAGCTTGTTATCATAG</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AG08</td>
<td>CGGAATGGCGGAACACTGCAATGAGCTTGTTATCATAG</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AG09</td>
<td>CGGAATGGCGGAACACTGCAATGAGCTTGTTATCATAG</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AG10</td>
<td>CGGAATGGCGGAACACTGCAATGAGCTTGTTATCATAG</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure A2. tRNA-Indel Translation of *A. gibbosa* for all haplotypes, with missing data marked as a dash. Stop codons are marked with a *

AG01
VVSWLTANCLISLQPNSK*YGKTKNFD*NMSSTLIFLISKTSLSHTYTSNSQSYYSNLYFLIMYMMNYK

AG02
VISWLTANCLISLQPNSK*YNNTKNSFD*NISSTLIFLISNTLHYTHTTNSQSYYYSNLYFLILIYIINYN

AG03
---------NCLISLQPNSK*YNNTKNSFD*NISSTLIFLISNTLSLHACTNSQSYYSNLYFLILIYIINYN

AG04
--SWLTANCLISLQPNSK*YNNTKNSFD*NISSTLIFLISNTLHYTHTTNSQSYYYSNLYFLILIYIINYN

AG05
---------NCLISLQPNSK*YNNTKNSFD*NISSTLIFLISNTLSIYHTHTNSQSYYSNLYFLILIYIINYN

AG06
--SWLTANCLISLQPNSK*YNNTKNSFD*NISSTLIFLISNTLSLHYTHTTNSQSYYYSNLYFLILIYIINYN

AG07
---------NCLISLQPNSK*YNNTKNSFD*NISSTLIFLISNTLSIYHTHTNSQSYYSNLYFLILIYIINYN

AG08
--SWLTANCLISLQPNSK*YNNTKNSFD*NISSTLIFLISNTLSLHYTHTTNSQSYYYSNLYFLILIYIINYN

AG09
---------------------*YNNTKNSFD*NISSTLIFLISNTLSLHYTHTTNSQSYYYSNLYFLILIYI**II

AG10
------LTANCLISLQPNSK*YNNTKNSFD*NISSTLIFLISNTLSIYHTHTNSQSYYSNLYFLILIYIINYN

Figure A3. Cytochrome Oxidase I (COI) translation of *A. gibbosa* for all haplotypes, with missing data marked as a dash.

AG01
MQLRWFFSTNHKDIGTLYLIFGAWAGMAGTAMSIII

AG02
MQLFRWFFSTNHKDIGTLYLIFGAWAGMAGTAMSIII

AG03
MLRNWFFSTNHKDI------------------------

AG04
MLRNWFFSTNHKDIGT------------------------

AG05
MLTRWFFSTNHKDIGTLYLIFGAWAGMAGTAMSIII

AG06
MLTRWFFSTNHKDIGTLYLIFGAWAGMAGTAVSVII

AG07
MLRNWFFSTNHKDIGT------------------------

AG08
MLSRWFFSTNHKDIGT------------------------

AG09
MLRNWFFSTNHKDI------------------------

AG10
MLTRWFFSTNHKDIGTLYLIFGAWAGMAGTAMSIII
**Figure A4.** Alignment of the sequence for *Asterina phylactica* for all haplotypes, with missing data marked as a dash.

<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>AP01</td>
<td>GTTATCTCTCGTTAACGCGCAAATGGCATTCTCATTAGAGAAATAGAAAGTAG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AP02</td>
<td>GTTATCTCTCGTTAACGCGCAAATGGCATTCTCATTAGAGAAATAGAAAGTAG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AP03</td>
<td>ATGCTCGTGCGTTAACGCGCAAATGGCATTCTCATTAGAGAAATAGAAAGTAG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AP04</td>
<td>ATGCTCGTGCGTTAACGCGCAAATGGCATTCTCATTAGAGAAATAGAAAGTAG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AP05</td>
<td>ATGCTCGTGCGTTAACGCGCAAATGGCATTCTCATTAGAGAAATAGAAAGTAG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AP06</td>
<td>GTTATCGTTAACGCGCAAATGGCATTCTCATTAGAGAAATAGAAAGTAG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AP07</td>
<td>ATGCTCGTGCGTTAACGCGCAAATGGCATTCTCATTAGAGAAATAGAAAGTAG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AP08</td>
<td>ATGCTCGTGCGTTAACGCGCAAATGGCATTCTCATTAGAGAAATAGAAAGTAG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

61 70 80 90 100 110 120

| AP01 | TATAAGGTTAAAACAAAGAATTTTGATTTCTAAAATATAAGTTCAACTCTTATCTTTCTA |   |   |   |   |   |
| AP02 | TATAAGGTTAAAACAAAGAATTTTGATTTCTAAAATATAAGTTCAACTCTTATCTTTCTA |   |   |   |   |   |
| AP03 | TATAAGGTTAAAACAAAGAATTTTGATTTCTAAAATATAAGTTCAACTCTTATCTTTCTA |   |   |   |   |   |
| AP04 | TATAAGGTTAAAACAAAGAATTTTGATTTCTAAAATATAAGTTCAACTCTTATCTTTCTA |   |   |   |   |   |
| AP05 | TATAAGGTTAAAACAAAGAATTTTGATTTCTAAAATATAAGTTCAACTCTTATCTTTCTA |   |   |   |   |   |
| AP06 | TATAAGGTTAAAACAAAGAATTTTGATTTCTAAAATATAAGTTCAACTCTTATCTTTCTA |   |   |   |   |   |
| AP07 | TATAAGGTTAAAACAAAGAATTTTGATTTCTAAAATATAAGTTCAACTCTTATCTTTCTA |   |   |   |   |   |
| AP08 | TATAAGGTTAAAACAAAGAATTTTGATTTCTAAAATATAAGTTCAACTCTTATCTTTCTA |   |   |   |   |   |

121 130 140 150 160 170 180

| AP01 | ACTGAAAACAAACAGAATTTTACACTCACACCAATAACTCCCAAAGTTACTATTCTAACTTA |   |   |   |   |   |
| AP02 | ACTGAAAACAAACAGAATTTTACACTCACACCAATAACTCCCAAAGTTACTATTCTAACTTA |   |   |   |   |   |
| AP03 | ACTGAAAACAAACAGAATTTTACACTCACACCAATAACTCCCAAAGTTACTATTCTAACTTA |   |   |   |   |   |
| AP04 | ACTGAAAACAAACAGAATTTTACACTCACACCAATAACTCCCAAAGTTACTATTCTAACTTA |   |   |   |   |   |
| AP05 | ACTGAAAACAAACAGAATTTTACACTCACACCAATAACTCCCAAAGTTACTATTCTAACTTA |   |   |   |   |   |
| AP06 | ACTGAAAACAAACAGAATTTTACACTCACACCAATAACTCCCAAAGTTACTATTCTAACTTA |   |   |   |   |   |
| AP07 | ACTGAAAACAAACAGAATTTTACACTCACACCAATAACTCCCAAAGTTACTATTCTAACTTA |   |   |   |   |   |
| AP08 | ACTGAAAACAAACAGAATTTTACACTCACACCAATAACTCCCAAAGTTACTATTCTAACTTA |   |   |   |   |   |

181 190 200 210 220 230 240

| AP01 | GAATGGCCGGAACTGCAATGAGCGTTATCATA |   |   |   |   |   |
| AP02 | GAATGGCCGGAACTGCAATGAGCGTTATCATA |   |   |   |   |   |
| AP03 | GAATGGCCGGAACTGCAATGAGCGTTATCATA |   |   |   |   |   |
| AP04 | GAATGGCCGGAACTGCAATGAGCGTTATCATA |   |   |   |   |   |
| AP05 | GAATGGCCGGAACTGCAATGAGCGTTATCATA |   |   |   |   |   |
| AP06 | GAATGGCCGGAACTGCAATGAGCGTTATCATA |   |   |   |   |   |
| AP07 | GAATGGCCGGAACTGCAATGAGCGTTATCATA |   |   |   |   |   |
| AP08 | GAATGGCCGGAACTGCAATGAGCGTTATCATA |   |   |   |   |   |

241 250 260 270 280 290 300

| AP01 | TTTTTTCCACTAAACACAAGGACATAGTTACTCTCCTGACTCTTTATATTGAGCCTGAGCCG |   |   |   |   |   |
| AP02 | TTTTTTCCACTAAACACAAGGACATAGTTACTCTCCTGACTCTTTATATTGAGCCTGAGCCG |   |   |   |   |   |
| AP03 | TTTTTTCCACTAAACACAAGGACATAGTTACTCTCCTGACTCTTTATATTGAGCCTGAGCCG |   |   |   |   |   |
| AP04 | TTTTTTCCACTAAACACAAGGACATAGTTACTCTCCTGACTCTTTATATTGAGCCTGAGCCG |   |   |   |   |   |
| AP05 | TTTTTTCCACTAAACACAAGGACATAGTTACTCTCCTGACTCTTTATATTGAGCCTGAGCCG |   |   |   |   |   |
| AP06 | TTTTTTCCACTAAACACAAGGACATAGTTACTCTCCTGACTCTTTATATTGAGCCTGAGCCG |   |   |   |   |   |
| AP07 | TTTTTTCCACTAAACACAAGGACATAGTTACTCTCCTGACTCTTTATATTGAGCCTGAGCCG |   |   |   |   |   |
| AP08 | TTTTTTCCACTAAACACAAGGACATAGTTACTCTCCTGACTCTTTATATTGAGCCTGAGCCG |   |   |   |   |   |

| AP01 | GAATTGCGGAACTGCAATGAGCGTTATCATA |   |   |   |   |   |
| AP02 | GAATTGCGGAACTGCAATGAGCGTTATCATA |   |   |   |   |   |
| AP03 | GAATTGCGGAACTGCAATGAGCGTTATCATA |   |   |   |   |   |
| AP04 | GAATTGCGGAACTGCAATGAGCGTTATCATA |   |   |   |   |   |
| AP05 | GAATTGCGGAACTGCAATGAGCGTTATCATA |   |   |   |   |   |
| AP06 | GAATTGCGGAACTGCAATGAGCGTTATCATA |   |   |   |   |   |
| AP07 | GAATTGCGGAACTGCAATGAGCGTTATCATA |   |   |   |   |   |
| AP08 | GAATTGCGGAACTGCAATGAGCGTTATCATA |   |   |   |   |   |

301 310 320 330

| AP01 | GAATTGCGGAACTGCAATGAGCGTTATCATA |   |   |   |   |   |
| AP02 | GAATTGCGGAACTGCAATGAGCGTTATCATA |   |   |   |   |   |
| AP03 | GAATTGCGGAACTGCAATGAGCGTTATCATA |   |   |   |   |   |
| AP04 | GAATTGCGGAACTGCAATGAGCGTTATCATA |   |   |   |   |   |
| AP05 | GAATTGCGGAACTGCAATGAGCGTTATCATA |   |   |   |   |   |
| AP06 | GAATTGCGGAACTGCAATGAGCGTTATCATA |   |   |   |   |   |
| AP07 | GAATTGCGGAACTGCAATGAGCGTTATCATA |   |   |   |   |   |
| AP08 | GAATTGCGGAACTGCAATGAGCGTTATCATA |   |   |   |   |   |
**Figure A5.** tRNA-Indel Translation of *A. phylactica* for all haplotypes, with missing data marked as a dash. Stop codons are marked with an *

AP1

VISWLTANCLSISLQPNSK*YNGNTKND*F*NISSTLIFISNTSLHYTNNSQSYSYLNYFLIIYIINYN

AP2

VVISWLTANCLSISLQPNSK*YNGNTKND*F*NISSTLIFISNTSLHYTNNSQSYSYLNYFLIIYIINYN

AP3

------LTANCLSISLQPNSK*YNGNTKND*F*NISSTLIFISNTSLHYTNNSQSYSYLNYFLIIYIINYN

AP5

------LTANCLSISLQPNSK*YNGNTKND*F*NISSTLIFISNTSLHYTNNSQYSYLNYFLIIYIINYN

AP6

VISOVLTANCLSISLQPNSK*YKGKTKND*F*NISSTLIFISNTSLHYTNNSQSYYSYLNYFLIIYIINYN

AP7

---------ISLQPNSK*YNGNTKND*F*NISSTLIFISNTSLHYTNNSQSYYSYLNYFLIIYIINYN

AP8

---------ISLQPNSK*YNGNTKND*F*NISSTLIFISNTSLHYTNNSQSYYSYLNYFLIIYIINYN

**Figure A6.** COI translation of *A. gibbosa* for all haplotypes, with missing data marked as a dash.

AP1

MQLTRWFFSTNHKDIGTLYLFIFGAWAGMGTAMS

AP2

MQLNRWFFSTNHKDIGTLYLFIFGAWAGMGTAMS

AP3

MQLNRWFFSTNHKDIGTLYLFIFGAWAGMGTAMS

AP5

MQLNRWFFSTNHKDVGTLYLFIFGAWAGMGTAMS
AP6  MQLTRWFFSTNHKD<sup>1</sup>G<sup>2</sup>T<sup>3</sup>LY<sup>4</sup>LIFGAW<sup>5</sup>AGMA<sup>6</sup>TAMS<sup>7</sup>VII

AP7  MQLTRWFFST--------------------------

AP8  MQLTRWFFSTNHKD--------------------------
Appendix 2

Gene-flow patterns in Atlantic and Mediterranean populations of the Lusitanian sea star *Asterina gibbosa*.

Baus E, Darrock DJ, Bruford MW

Gene-flow patterns in Atlantic and Mediterranean populations of the Lusitanian sea star *Asterina gibbosa*

E. Baus, D. J. Darrock and M. W. Bruford
Biodiversity and Ecological Processes Group, Cardiff University, Main Building, Park Place, PO Box 915, Cardiff CF10 3TL, UK

**Abstract**

In this study, the population structure of the Lusitanian sea star *Asterina gibbosa* was assessed using amplified fragment length polymorphism (AFLP). One hundred and twenty-two AFLP loci were analysed in 159 individuals from eight populations from across the species’ range and revealed high levels of genetic diversity, with all individuals but two harbouring a unique banding pattern. As reported for other marine invertebrates, we found high levels of genetic differentiation between the Atlantic and Mediterranean basins, suggesting that the Strait of Gibraltar represents a major barrier to dispersal for this sea star. Our assignment studies suggest that, in the Atlantic, a measurable degree of gene flow occurs between populations, which could result in the isolation-by-distance pattern of differentiation observed in this basin. In contrast, no evidence of contemporary gene flow was found in the Mediterranean, suggesting contrasting patterns of dispersal of *Asterina gibbosa* in the Atlantic and Mediterranean basins.

**Keywords:** AFLP, *Asterina gibbosa*, echinoderm, gene flow, Lusitanian distribution, population genetics

Received 21 February 2005; revision accepted 22 June 2005

**Introduction**

Marine species are usually expected to show low levels of geographical differentiation due to the lack of obvious barriers to gene flow in the marine environment (Ward et al. 1994). However, increasing evidence indicates that marine species show more population differentiation than what would be expected on the basis of their dispersal capabilities alone (e.g. Palumbi 1994; Riginos & Nachman 2001). The factors that determine gene-flow patterns among marine populations are still poorly understood. Historical vicariance, habitat discontinuity, larval behaviour, marine currents and local adaptation (to temperature and salinity conditions, for example) are only a few examples of mechanisms that might favour local differentiation (see Palumbi et al. 1994; Riginos & Nachman 2001 and references therein; Bierne et al. 2003).

The Strait of Gibraltar has been shown to play a major role in shaping diversity in marine species with an Atlantic–Mediterranean distribution. Indeed, a clear reduction of gene flow between the Atlantic and the Mediterranean basins has been identified in numerous marine species including invertebrates (Borsa et al. 1997; Pérez-Losada et al. 1999; Zane et al. 2000; Duran et al. 2004; Lo Brutto et al. 2004; Roman & Palumbi 2004). However, a complete Atlantic–Mediterranean divide is not observed at Gibraltar and some species show no differentiation at all between these two basins (Stamatis et al. 2004). A recent study using strictly identical markers and sampling schemes to compare the sensitivity of five closely related teleost fish species to the Atlantic–Mediterranean boundary also led to contrasting results (Bargelloni et al. 2003). Although the five species under investigation shared similar biological features, evidence for a sharp phylogeographical break at Gibraltar was found for only two of them. To date, the historical and/or ecological reasons for these discrepancies have not been fully elucidated.

*Asterina gibbosa* (Pennant, 1897) is a small cushion star (measuring up to 5 cm in diameter) characterized by a subpentagonal body shape with five bluntly rounded arms and a variable colour pattern ranging from light green through muddy brown to bright orange. It is found under boulders and stones in the intertidal zone of rocky sea shores and notably in tide pools. This species occurs along
the western coast of Europe, ranging from the northwest coast of Scotland to throughout the Mediterranean Sea (a 'Lusitanian' distribution) and is particularly abundant in the south of the British Isles (Crump & Emson 1983). Individuals are protandrous hermaphrodites and, in the spring, females lay masses of up to 1000 bright orange, sticky eggs that develop directly into juveniles (Crump & Emson 1983).

Benthic marine species lacking a larval stage have limited intrinsic dispersal capabilities and are therefore expected to display higher population genetic structure than species with a planktonic larval stage (Hunt 1993; Arndt & Smith 1998; Collin 2001). Directly developing species almost exclusively rely on drifting or rafting on macro-algae (or, more recently, on human activities such as shipping) for eggs or juveniles to achieve sporadic long-distance dispersal as has been suggested for several echinoderm species (Spomer & Roy 2002; Waters & Roy 2003; Waters & Roy 2004b) and other marine invertebrates such as sea anemones (Edmands & Potts 1997). They are therefore expected to be sensitive to natural barriers, such as those imposed by marine currents, during dispersal (Spomer & Roy 2002; Waters & Roy 2004b). In this study, we assessed the impact of potential natural barriers, such as the Strait of Gibraltar and the English Channel, on the population structure and dispersal pattern of the direct developing littoral echinoderm species Asterina gibbosa.

Because microsatellite markers proved difficult to isolate in A. gibbosa, amplified fragment length polymorphism (AFLP) was used to conduct this study. This technique has proved to be a powerful tool for assessing genetic diversity and population structure in a wide range of plants (Tremetsberger et al. 2003; He et al. 2004; Juan et al. 2004; Muller et al. 2004), fungi (Laitung et al. 1992; Isabel et al. 2001; Douek et al. 2003; Takami et al. 2004) including marine invertebrates (Triantaphyllidis et al. 1997; Fetznier & Crandall 1999; Bariki et al. 2000; Wilding et al. 2001; Douek et al. 2002). The strength of this technique resides essentially in the speed and ease with which large numbers of high-resolution markers distributed across the genome can be generated without any prior knowledge of the DNA sequence. Dealing with large numbers of markers can be a major advantage since it has been demonstrated that the number of loci can be more critical in generating information content in population genetics than increasing allelic diversity per locus (Bernatchez & Duschesne 2000). Until recently, the dominant nature of AFLP markers was regarded as a potential drawback for population studies. Indeed, approaches to partitioning genetic diversity required either (i) the assumption that populations were in Hardy-Weinberg equilibrium and that the inbreeding coefficient in populations was known (Lynch & Milligan 1994; Zhivotovsky 1999) or (ii) the treatment of the multilocus AFLP phenotypes as haplotypes with the use of similarity indices (Nei & Li 1979) or Euclidian distances (Schneider et al. 2000) to describe distances among haplotypes in an analysis of molecular variance (Excoffier et al. 1992; Isabel et al. 1999). However, Bayesian approaches are now available (Holsinger et al. 2002) that allow the incorporation of uncertainty about the magnitude of within-population inbreeding coefficients into estimates of $F_{ST}$. This method therefore allows fuller advantage to be taken of the information contained in the large amount of markers generated by the AFLP technique and has been used successfully in several recent studies (Tremetsberger et al. 2003; Holsinger & Wallace 2004; Juan et al. 2004; Muller et al. 2004).

In this study, we investigated genetic diversity and gene flow in populations of Asterina gibbosa selected across the species' range, using AFLP, with the following objectives: (i) to reveal the level of genetic differentiation between populations; (ii) to determine the current patterns of gene flow between populations; and (iii) to evaluate the impact of potential natural barriers, such as the Strait of Gibraltar and the English Channel, on the dispersal pattern of this littoral echinoderm.

**Materials and methods**

**Sampling**

*Asterina gibbosa* samples were collected under boulders and stones of rocky sea shores and tide pools from seven locations selected across the species' range: four sites in the British Isles: Loch Hyne ('LH'; Ireland, southern coast), Rockham Bay ('RB'; UK, northern coast of Cornwall), Chapel Point and Prawle Point (respectively 'CP' and 'PP'; UK, southern coast of Cornwall) and three sites in continental Europe: Guéthary ('FG'; France, Atlantic coast), Banyuls-sur-Mer ('FB'; France, Mediterranean coast) and Naples ('IN'; Italy, western coast). Specimens were also collected from aquaria in the 'Institut Océanographique Paul Ricard' in Les Embiez ('FE'; France, Mediterranean coast) which represent individuals taken from an adjacent natural population living at a depth of 4–5 m, offshore from the marine station. All specimens were stored for a few days in absolute ethanol at ambient temperature and then for a few days up to a few weeks at −80 °C.

**DNA extraction**

DNA extractions were performed on 2–4 mm$^3$ of tissue (arm tip) using a DNeasy extraction kit (QIAGEN, catalogue # 69506). DNA quality and quantity was assessed by running samples on 1% agarose gels alongside dilutions of lambda DNA (Promega, catalogue # D1501).
AFLP procedure

AFLP was performed according to Ajmone-Marsan et al. (1997) using the following five selective primer combinations: EcoRI-AAC + TaqI-CCA; EcoRI-AAG + TaqI-CCA; EcoRI-ACT + TaqI-CCA; EcoRI-ACT + TaqI-CAC; EcoRI-AAC + TaqI-ACT. Some adaptations intended to maximize the reproducibility of the procedure were introduced. First, only 200 ng of high quality DNA was used to ensure a full digestion of the starting material. Second, three independent preselective polymerase chain reaction (PCR) products were combined and used as a template for the selective PCR in order to maximize the probability of amplifying all restriction fragments produced by the digestion reaction. Third, all samples were distributed randomly on two 96-well plates and all samples from each plate were processed at the same time, using the same reaction mix. To allow for comparisons between samples processed on different plates, controls were included and used to identify and exclude bands that were not reproducible between experiments. As opposed to the protocol by Ajmone-Marsan et al. (1997), we used a fluorescence-based band detection procedure: EcoRI fluorescent-labelled primers were used during the selective PCR and the detection was performed on an ABI 377 automatic sequencer with the ABI GENESCAN 3.2.1 analysis software (PE Applied Biosystems Inc.). Bands scoring was carried out using GENOGRAPHER software (Benham et al. 1999) and then confirmed visually (blind test). Only bands between 100 and 500 bp with sufficient intensity and which were reliably amplified in both experiments were scored.

Data analysis

Genetic diversity was evaluated as the percentage of polymorphic loci (calculated using AFLP-SURV version 1.0; Vekemans 2002) and the number of shared multilocus AFLP patterns (evaluated with ARLEQUIN version 2.000; Schneider et al. 2000).

Genetic differentiation among populations was evaluated using a hierarchical Bayesian approach developed by Holsinger et al. (2002) that does not assume any prior knowledge of the degree of within-population inbreeding and is therefore not subject to the problems of traditional methods of analysis using dominant markers. We used the software HICKORY version 1.0.3 (Holsinger et al. 2002) to estimate $\theta_0$, a Bayesian analogue of $F_{ST}$, across all populations and for each pairwise combination of populations. The data were run two to three times with the default parameters (burn-in = 50,000, number of samples = 250,000, thinning factor = 50) using four models: a full model, a model that assumes no inbreeding within populations ($F_{IS} = 0$ model), a model that assumes no differentiation among populations ($\theta_0 = 0$ model) and a model that does not attempt to estimate $F_{IS}$ ($f$-free model). Because estimates of $F_{IS}$ derived from dominant marker data may be unreliable (Holsinger & Wallace 2004), we used the $f$-free analysis as our preferred method to calculate estimates of $\theta_0$. The deviance information criterion (DIC) values for the $F_{IS} = 0$, $\theta_0 = 0$ and full models were used to estimate how well each model fitted the data (a smaller DIC values indicates a better fit) and which model should be preferred. A Mantel test was performed using the program TPFA (Miller 1997) to evaluate the geographical structure among sample sites. The matrix used pairwise $\theta_0$ and geographical distances calculated as the shortest distance between two populations that did not involve crossing land.

Analysis of molecular variance (AMOVA), based on Euclidian distances between AFLP multilocus phenotypes, was conducted with ARLEQUIN 2.000 (Schneider et al. 2000). Different regional groupings were tested. Principal coordinate analysis (PCA) based on Euclidian distances between AFLP multilocus phenotypes (calculated with ARLEQUIN 2.000) was performed using GENALEX version 5.1 (Peakall & Smouse 2001). The PCA via covariance matrix with data standardization method was chosen. Finally, assignment tests were carried out using the re-assignment procedure of AFLPOP version 1.1 (Duchesne & Bernatchez 2002) with the default settings (fixed correction value for zero frequencies = 0.001, minimal log-likelihood difference to allocate specimens = 0, number of artificial genotypes to compute $P$ values = 500).

Results

One hundred and fifty-nine specimens of Asterina gibbosa sampled from eight populations (seven natural populations and one population from local aquaria) selected across the species range (as shown in Fig. 1) were analysed using five selective primer pair combinations and generated a total of 183 bands. All bands were polymorphic at the 5% level in at least one population. However, only 122 bands could be unambiguously and reproducibly scored in blind tests and these were retained for further analysis. All individuals tested except two (one pair) displayed unique AFLP band patterns, suggesting a high level of genetic variability. As shown in Table 1, the percentage of polymorphic bands across the seven natural populations ranged between 48.4% and 78.7%. In contrast, the population sampled in the aquaria of the ‘Institut Océanographique Paul Ricard’ in Les Embiez showed a low level of polymorphism (18.0%). A small number of markers were restricted either to the Atlantic (7 loci) or Mediterranean populations (5 loci). Private alleles for the French Atlantic (1 locus) and for each of the Mediterranean populations (PB: 4 loci, FE: 1 locus, IN: 4 loci) were also found.

Mean genetic differentiation between populations estimated using the Bayesian hierarchical method of Holsinger...
Table 1 Characteristics of the samples collected from eight populations of *Asterina gibbosa*. For each sample site, the number of individuals analysed (n) and the percentage of polymorphic AFLP loci (%P) are presented.

<table>
<thead>
<tr>
<th>Basin</th>
<th>Country</th>
<th>Population</th>
<th>Abbreviation</th>
<th>n</th>
<th>%P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atlantic</td>
<td>Ireland</td>
<td>Loch Hyne</td>
<td>LH</td>
<td>8</td>
<td>76.2</td>
</tr>
<tr>
<td>Atlantic</td>
<td>UK</td>
<td>Rockham Bay</td>
<td>RB</td>
<td>33</td>
<td>78.7</td>
</tr>
<tr>
<td>Atlantic</td>
<td>UK</td>
<td>Chapel Point</td>
<td>CP</td>
<td>31</td>
<td>65.6</td>
</tr>
<tr>
<td>Atlantic</td>
<td>UK</td>
<td>Prawle Point</td>
<td>PP</td>
<td>30</td>
<td>68.0</td>
</tr>
<tr>
<td>Atlantic</td>
<td>France</td>
<td>Guéthary</td>
<td>FG</td>
<td>16</td>
<td>77.9</td>
</tr>
<tr>
<td>Mediterranean</td>
<td>France</td>
<td>Banyuls-sur-Mer</td>
<td>FB</td>
<td>21</td>
<td>48.4</td>
</tr>
<tr>
<td>Mediterranean</td>
<td>France</td>
<td>Les Embiez</td>
<td>FE</td>
<td>10</td>
<td>18.0</td>
</tr>
<tr>
<td>Mediterranean</td>
<td>Italy</td>
<td>Naples</td>
<td>IN</td>
<td>10</td>
<td>76.2</td>
</tr>
</tbody>
</table>

Table 2 Pairwise genetic differentiations (estimated as $\theta^B$ values) between eight populations of *Asterina gibbosa*.

<table>
<thead>
<tr>
<th>LH</th>
<th>RB</th>
<th>CP</th>
<th>PP</th>
<th>FG</th>
<th>FB</th>
<th>FE</th>
<th>IN</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.0733</td>
<td>0.0625</td>
<td>0.0645</td>
<td>0.1175</td>
<td>0.4053</td>
<td>0.5093</td>
<td>0.3279</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0.0625</td>
<td>0.0645</td>
<td>0.1175</td>
<td>0.4053</td>
<td>0.5093</td>
<td>0.3279</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0.0741</td>
<td>0.0274</td>
<td>0</td>
<td>0.4315</td>
<td>0.5086</td>
<td>0.3689</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0.1530</td>
<td>0.1653</td>
<td>0.1494</td>
<td>0.4354</td>
<td>0.5082</td>
<td>0.3666</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0.3948</td>
<td>0.4315</td>
<td>0.4354</td>
<td>0.4336</td>
<td>0.4316</td>
<td>0.3412</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0.4564</td>
<td>0.5088</td>
<td>0.5086</td>
<td>0.5082</td>
<td>0.4316</td>
<td>0.3793</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0.3623</td>
<td>0.3689</td>
<td>0.3666</td>
<td>0.3412</td>
<td>0.3793</td>
<td>0</td>
</tr>
</tbody>
</table>

Fig. 1 Map showing the eight *Asterina gibbosa* sampling sites (indicated by arrow heads) selected for this study across the species range. Abbreviations correspond to population names as indicated in Table 1.

As shown in Table 3, AMOVA revealed that 60.49% of the total genetic variation was attributed to differences
Table 3 AMOVA, based on Euclidian distances between AFLP multilocus phenotypes, for 159 Asterina gibbosa individuals sampled from eight populations, with and without regional structuring

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>d.f.</th>
<th>Sum of squares</th>
<th>Variance components</th>
<th>Percentage of variation</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>No regional groups</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Among populations</td>
<td>7</td>
<td>976.577</td>
<td>6.72948</td>
<td>39.51</td>
<td>0</td>
</tr>
<tr>
<td>Within populations</td>
<td>151</td>
<td>1555.681</td>
<td>10.30252</td>
<td>60.49</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>158</td>
<td>2532.258</td>
<td>17.03200</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Two groups: Atlantic/Mediterranean</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Among groups</td>
<td>1</td>
<td>589.823</td>
<td>8.65733</td>
<td>39.77</td>
<td>0.02346</td>
</tr>
<tr>
<td>Among populations</td>
<td>6</td>
<td>386.754</td>
<td>2.80958</td>
<td>12.91</td>
<td>0</td>
</tr>
<tr>
<td>Within populations</td>
<td>151</td>
<td>1555.681</td>
<td>10.30252</td>
<td>47.33</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>158</td>
<td>2532.258</td>
<td>21.76943</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Three groups: Atlantic/Mediterranean France/Mediterranean Italy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Among Groups</td>
<td>2</td>
<td>712.140</td>
<td>9.71138</td>
<td>44.01</td>
<td>0.00399</td>
</tr>
<tr>
<td>Among populations</td>
<td>5</td>
<td>264.437</td>
<td>2.05057</td>
<td>9.29</td>
<td>0</td>
</tr>
<tr>
<td>Within populations</td>
<td>151</td>
<td>1555.681</td>
<td>10.30252</td>
<td>46.69</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>158</td>
<td>2532.258</td>
<td>22.06447</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The plot of the first and second principal coordinates, which accounted for 27.5% and 14.8% of the total variation, respectively, provided a visual representation of genetic similarity between populations (Fig. 2) and supported the results given by AMOVA. The Atlantic and Mediterranean samples clustered separately from each other. Atlantic samples were spread over a relatively large proportion of the two-dimensional space, suggesting high levels of genetic variability in this group. Within the Mediterranean group, the Italian samples were well differentiated from the French samples and were highly variable. It is worth noting that the population from the French aquaria showed very low levels of genetic variability compared to natural populations.

This general pattern of genetic structure was also illustrated by the assignment test (Table 4). In the Mediterranean basin, all individual multilocus genotypes except
Moreover, the percentage of polymorphic bands across the seven natural populations of *A. gibbosa* ranged between 48.4% and 78.7%. These results are consistent with AFLP data reported for other European invertebrates such as the soft coral *Parerythrophora chilensis* (Bakker et al. 2000), the sea anemone *Actinia equina* (Douce et al. 2002) and the dung beetle of the genus *Trypocerus* (Carisio et al. 2004), where all individuals genotyped presented different profiles and the percentage of polymorphic bands across the investigated populations ranged from 36.1% to 47.2% in the sea anemone and 44.2% to 79.7% in the dung beetle, respectively.

The relatively low levels of genetic diversity found in the population from the aquaria in Les Embiez might be consistent with a founder effect. Indeed, this population was probably founded by a few animals pumped at an early stage of development (eggs or juveniles) from a few metres offshore into the water system and thus introduced into the aquaria with the water supply. These individuals are now established in the aquaria and reproduce in probable isolation from the natural population from which they originate. Although *A. gibbosa* is a protandrous hermaphrodite and external fertilization is normally the rule, it is not uncommon to find animals with both ripe spermatozoa and ova present in the gonads. Self-fertilization is therefore considered as a distinct possibility for this species (Cognet & Delavault 1962). This phenomenon could be accentuated in a confined environment and could also partly account for the reduced diversity observed in the population inhabiting the aquaria. The DIC statistics provided by the Bayesian factor analysis indicated that, when all the populations were taken into account, there was little evidence that the $F_{IS} = 0$ model should be rejected.

All together, these results suggest that inbreeding and self-fertilization, if they occur in some circumstances, are probably not a major factor in shaping *A. gibbosa* population genetic structure in natural habitats. This result is in agreement with studies performed in other marine species capable of self-fertilization showing that selfing is rare in the field (Cohen 1990; Hunter & Hughes 1993).

$F_{ST}$ estimated using the Bayesian hierarchical approach was $\Theta^B = 0.364$. It is interesting to note that this value is between those obtained by the traditional approaches, which required either the assumption that populations are in Hardy–Weinberg equilibrium (0.273) or the use of AMOVA based on Euclidian distances between the multilocus AFLP haplotypes (0.395). This is, however, not the rule, since other authors have compared $F_{ST}$ estimates obtained with different approaches and found $\Theta^B$ to be either higher or lower than the other estimates (Holsinger et al. 2002; Tero et al. 2003; Wang et al. 2003; Juan et al. 2004).

The pairwise $\Theta^B$ values (Table 2), AMOVA (Table 3), PCA (Fig. 2) and re-allocation test (Table 4) all concur and show high levels of genetic differentiation between the Atlantic and the Mediterranean populations. Accordingly, private

![Fig. 3 Relationship between genetic differentiation and geographical distance among four Atlantic populations of Asterina gibbosa. Genetic differentiations were estimated as $\Theta^B$ values; geographical distances were measured as the shortest distances without crossing any land and were expressed in km.](image-url)
alleles were found for the Atlantic (7 loci) and the Mediterranean (5 loci) populations. These results strongly suggest that, as described for numerous marine species, the Strait of Gibraltar represents a major barrier to dispersal for *A. gibbosa*. This Atlantic–Mediterranean divide could be explained by different factors acting either singly or in combination. Gene flow could be restricted by marine currents, in particular the strong currents occurring at the narrow and shallow passage of the Strait of Gibraltar, in the Alboran Sea and along the Almeria-Oran line. Investigations conducted around the world indicated that several echinoderm species are sensitive, to various degree, to natural barriers such as marine currents. A phylogeographical analysis of the brooding brittle star *Amphipholis squamata* along the coast of New Zealand revealed that sporadic long-distance dispersal events are consistent with the regime of oceanic circulation (Sponer & Roy 2002). In particular, currents such as the Southland Current were found to have significant effects on the genetic structure of *A. squamata* populations around the east coast of New Zealand. Other phylogeographical investigations have shown that species with a planktonic larval development are also sensitive to natural barriers. A recent study by Waters & Roy (2004a) revealed a significant heterogeneity between populations of the endemic sea star *Patiriella regularis* from northern and southern New Zealand consistent with the hypothesis that coastal upwelling and/or the Cook Strait could disrupt the gene flow between these regions. A study of the gene flow and genetic diversity in the sea urchin genus *Echinometra* across the tropical Pacific indicated that differentiation among populations varied greatly between species and implicated factors such as island position, oceanic currents and random long-distance dispersal events (Palumbi et al. 1997). McCartney et al. (2000) compared the impact of potential natural dispersal barriers on the genetic structure of three Neotropical species of *Echinometra* and found contrasting patterns of gene flow within and between the Caribbean and the rest of the Atlantic, illustrating that echinoderm species with similar biology may be affected by the same potential barrier (i.e. the Caribbean current and the Gulf Stream) in a dissimilar manner. These authors have attributed these discrepancies to the highly stochastic nature of events in which larvae successfully cross large distances. Furthermore, in accordance with our results, a study of the Atlantic–Mediterranean sea urchin *Paracentrotus lividus* revealed significant genetic differentiation between these two basins consistent with restricted gene flow across the geographical boundary imposed by the Strait of Gibraltar (Duran et al. 2004).

In addition to restricted gene flow due to marine currents occurring in the Gibraltar area, other factors might explain the Atlantic–Mediterranean divide in *A. gibbosa*, such as historical events. Mediterranean and Atlantic populations were isolated during glacial period(s), with subsequent genetic divergence and secondary contact. Such a mechanism has been inferred for other marine invertebrates such as the mussel *Mytilus galloprovincialis* (Quesada et al. 1995), the cuttlefish *Sepia officinalis* (Pérez-Losada et al. 1999) and the scallops *Pecten jacobaeus* and *Pecten maximus* (Rios et al. 2002). Selective forces related to physical, chemical and/or ecological conditions present in each basin could also account for the phylogeographical break at Gibraltar. Notably, differences in water temperatures and salinity between the Atlantic and the Mediterranean have been suggested to play a role in maintaining the genetic differentiation between populations of the European hake from the two basins (Cimmaruta et al. 2005).

Our assignment test (Table 4) suggested that gene flow occurs among populations in the Atlantic. Indeed, a significant amount of Atlantic individuals were reallocated to a different but closely located population from the same basin. The Mantel test showed evidence of a strong correlation between genetic differentiation and geographical distance and an isolation-by-distance differentiation pattern within the Atlantic is plausible (Fig. 3). Of note, the Mantel test was not significant at the 5% level but this is likely to be due to the low number of populations (hence degrees of freedom) in the test.

Interestingly, the re-allocation test indicated that a high proportion of samples from two Cornish sites were re-allocated to a population of a third Cornish site, Prawle Point. According to Salomon & Breton's (1993) tidal current model (average tidal amplitude, no wind), the current originating from the Gulf Stream flows towards Plymouth and then separates into two routes, one eastbound towards Normandy and the other, westbound, along the Cornwall coast towards the Scilly Islands. The latter could explain why directional exchanges occur from populations located on the southern Cornish coast (Prawle Point) to populations on the western (Chapel Point) and northern Cornish coast (Rockham Bay). The re-allocation test also revealed that, despite its relative isolation from oceanic waters, the population inhabiting Loch Hyne in Ireland appears to have maintained some level of gene flow with the Cornish populations.

In contrast to the Atlantic, gene flow appears to be much more restricted in the Mediterranean. Indeed, in the Mediterranean, all individuals but one were re-allocated to their original population in the assignment test (Table 4). In agreement with contrasting levels of gene flow between Atlantic and Mediterranean populations, private alleles were found for all of the Mediterranean populations (FB: 4 loci, FE: 1 locus, IN: 4 loci) but only for one (the more distant) of the Atlantic populations (FG: 1 locus) and pairwise θ values among Mediterranean populations were high (Table 2). In particular, high levels of differentiation were found between the Italian and French Mediterranean populations of *A. gibbosa*, as illustrated by the AMOVA (Table 3).
and the PCA (Fig. 2). These results are corroborated by a study of a directly developing marine mollusc that also reported restricted gene flow between populations from the Tyrrhenian Sea and from the northwestern coast of the Mediterranean Sea (Pérez-Losada et al. 1999). The complexity of the northern Mediterranean coastline, as well as the presence of numerous islands, creates many small eddies and other local currents (Send et al. 1999; Encyclopaedia Britannica 2005) that might explain the relative genetic isolation of closely located populations in this area of the basin. Moreover, the low amplitude of tidal currents in the Mediterranean Sea could reduce exchanges between the littoral and sublittoral environment and could partly account for the contrasting dispersal patterns observed in the Atlantic and Mediterranean basins.

In conclusion, this study provides new insights into the dispersal pattern in the Lusitanian cushion star A. gibbosa and might have important implications for the conservation of this species which, like many other littoral organisms, regularly suffers the consequences of human-mediated pollution (such as oil spills).

Acknowledgements

We would like to thank Prof. Paolo Ajmone-Marsan and Dr Elsabete Pires for their help while developing the AFLP technique. We are grateful to Drs Roland Emson and Robin Crump for their inspiration in this project in addition to assistance with sampling and site location in the UK, to Dr Lisa Kirkendale and Dr Trinidad Pérez for their help and assistance during the fieldwork in France and to the ‘Institut Océanographique Paul Ricard’ in Les Embiez for providing us with samples from the aquaria. This research was supported through a European Community Marie Curie Fellowship (HPMF-CT-2002-01780) under the framework of the programme ‘Improving the Human Research Potential and the Socio-Economic Knowledge Base’. The authors are solely responsible for the information published. This work does not represent the views of the European Community.

References


Fetzner JW Jr, Crandall KA (1999) Genetic variability within and among populations of the golden crayfish (Oronectes luteus): a comparison using amplified fragment length polymorphism


Erika Baus was a Marie Curie Research Fellow in the Biodiversity and Ecological Processes Research Group, Cardiff University between 2002 and 2004, where David Darrock is carrying out his PhD studies on the phylogeography of *Asterina* species under the supervision of Mike Bruford and Roland Emson.