UNIVERSITY OF SOUTHAMPTON

# The genetic population structure of the lagoon specialists *Nematostella vectensis*, *Cerastoderma glaucum* and *Gammarus insensibilis* from populations along the southern and eastern coasts of the United Kingdom

by

Catherine Pearson

Ph.D. Thesis

School of Ocean and Earth Science

September 2003

# THE GENETIC POPULATION STRUCTURE OF THE LAGOON SPECIALISTS *NEMATOSTELLA VECTENSIS, CERASTODERMA GLAUCUM* AND *GAMMARUS INSENSIBILIS* FROM POPULATIONS ALONG THE SOUTHERN AND EASTERN COASTS OF THE UNITED KINGDOM

Doctor of Philosophy

Catherine Pearson

September 2003

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF SCIENCE
OCEAN AND EARTH SCIENCE

Doctor of Philosophy

# THE GENETIC POPULATION STRUCTURE OF THE LAGOON SPECIALISTS NEMATOSTELLA VECTENSIS, CERASTODERMA GLAUCUM, AND GAMMARUS INSENSIBILIS FROM POPULATIONS ALONG THE SOUTHERN AND EASTERN COASTS OF THE UNITED KINGDOM

by Catherine Victoria Melvaig Pearson

Molecular markers were used to analyse genetic variation within and between populations of three lagoon specialists; Nematostella vectensis, the lagoon sea anemone, Cerastoderma glaucum, the lagoon cockle and Gammarus insensibilis, the lagoon sand shrimp from lagoons along the southern and eastern coasts of the United Kingdom. Despite their very different life history strategies, all three species demonstrated genetic patterns associated with a high degree of isolation, combined with limited dispersal ability. The random amplified polymorphic DNA (RAPD) data, revealed the mainly asexual sea anemone, Nematostella vectensis to be predominantly clonal throughout its known U.K. range, with monoclonal or near monoclonal populations and rare genotypes confined to single sites. The microsatellite markers employed in this study indicated a high degree of genetic differentiation, severe inbreeding and the absence of Hardy-Weinberg equilibrium conditions in populations of both Cerastoderma glaucum and Gammarus insensibilis.

RAPD data from populations of N. vectensis revealed that 61% of individuals had an identical genotype, the frequency of which varied from 0.01 to 1.00. These data may provide evidence for a single 'general-purpose' genotype. The presence of significant linkage disequilibria ($p<0.05$) in populations of C. glaucum and in populations of G. insensibilis, moreover, provides further support for a 'general-purpose' genotype associated with the lagoon habitat. Additional evidence for such a genotype comes from the extreme excess of homozygosity (two loci showed significantly less observed than expected heterozygosity for all populations; $p<0.01$) in C. glaucum, and extreme excess of heterozygosity (fixed heterozygosity for one locus in all populations; $p<0.01$ and fixed heterozygosity for two loci in one other population; $p<0.01$) in G. insensibilis.

Colonisation to lagoons probably occurs across open water. Dispersal and migration between disjunct populations can have a great effect on the genetic diversity of some populations. While it is often assumed that reduced gene flow is responsible for population differentiation, the possibility that differentiation is occurring as a result of adaptation to a specific habitat niche is considered. Finally, the relevance of these data to the conservation management of lagoon specialists is discussed.

To my son Thomas
(6.3.03 - 22.03.03)

*' What would the world be once bereft of wet and of wildness, let them be left, let them be left, the wildness and the wet, long live the weeds and the wilderness yet. '*

Gerard Manley Hopkins (1881).

# Acknowledgements

I would like to thank my supervisor Dr Alex Rogers for his tireless enthusiasm, his contagious excitement for my work, his gentle but constructive criticism and his never-ending encouragement and of course, for his generosity in the pub. Alex, I would not have finished this work without your support, thank you.

I would like to thank my supervisor Dr Martin Sheader for assistance and advice throughout this study.

I would like to acknowledge NERC for the provision of studentship no. GT04/98/276/MS and to acknowledge the financial support of the Nature Conservancy Council for England for their contribution towards the investigation of the genetic population structure of *Gammarus insensibilis* (project number MAR01/15/02). All specimens of *Gammarus insensibilis* and *Nematostella vectensis* were collected under an English Nature licence held by Martin Sheader.

I would like to thank the following for assistance with sample collection; Dr John Davies, Dr Roger Bamber, Grant Lahore (National Trust, Aldeburgh), Martine Davis, Jo Preston, Dr Kerry Howell, and Andrew Whitehouse. Thanks are also due to Dr Elaine Robson and Dr Cadet Hand for provision of specimens of *Nematostella vectensis* from the U.S.A.

For technical assistance with Sequencing and Gene Scanning I wish to thank Dr Mark Dixon, who is also due thanks for the use of his lab in Boldrewood. I would like to acknowledge Valerie Renault for assistance with the development of microsatellite markers for *Cerastoderma glaucum*.

To the members of MERG, you know who you are. You know what we were. Thank you. Maria, my Portuguese angel, thank you. Rodrigo, how I wish I had some Portuguese so you could really know what I want to say to you; you are an inspirational scientist and a true friend, thank you.

Jo, you always came with me. Whether to the lab or into the field when it was dark, cold and wet or to the pub when it was warm and dry. Thank you.

Kim, Rick, Chris and Mart. You never asked, you always listened. Thank you.

Mum and Dad, well here we are. You know how I appreciate everything you do and have done for me. You care about me as a scientist, you even read what I write! You instilled in me my thirst for knowledge and also the humility to know my limitations. Thank you.

John, from its beginnings in Liverpool to its conclusion here in Sussex. You have held me up on my PhD journey. I cannot hope to be the scientist you are, I admire the understanding and depth you possess even in those subjects beyond your training. You gave me the carrot when I needed it and never resorted to the stick, even when I deserved it. Your patience is incredible. You gave me back my love of science when I thought it was irretrievable, that is the best gift I have ever been given. Thank you.

# Contents

# 1.0 General Introduction

## 1.1 What is a lagoon ?



## 1.2 Islands and lagoons as island-like habitats



## 1.3 Lagoons as marginal habitats



## 1.4 Genetics and conservation

# 1.1 What is a Lagoon?

## *1.1.1 Introduction*

A coastal lagoon is an area of salt or brackish water wholly or partially separated from the adjacent sea by a sand or shingle barrier (Barnes 1980). Wave action, land subsidence or storms may form lagoons naturally (Barnes 1980). They can also be created artificially, and will support characteristic lagoonal flora and fauna (Downie 1996), including some of the rarest plants and animals found in the UK (Bamber *et al.* 2001). The distinction, however, between artificial and natural lagoons is difficult to define, as human intervention has been increasingly necessary to conserve this delicate habitat (Bamber *et al.* 2001).

There has been much discussion over what constitutes a lagoon, in particular their comparison with estuaries (Barnes 1980; Downie 1996; Bamber *et al.* 2001) and several factors may be involved in their classification. These include the way in which the lagoon was formed, the maintenance of its water chemistry and the exchange between the source water and the lagoon. Bamber *et al.* (2001) distinguish lagoons from estuaries and littoral pools by their tideless or tidally restricted state. Lagoons may be categorised into two types, the classical definitions are, *'those formed by sedimentary barriers'* (Barnes 1989) that is the movement of sandbanks or shingle and *'those formed by rocky habitat'* (Covey 1999). The latter is less common, where rocks or other hard substrata separate the lagoon from the surrounding sea. This type of lagoon habitat is predominantly restricted to the fjordic system in Scotland (see Covey 1999).

Two other categories, perhaps as significant, are 'saline lagoons' and 'lagoon like ponds or habitats'. Conservation issues may be as important for these because of the specialist biota found in such habitats (Sheader & Sheader 1989). In order that this habitat may be recognised and protected, classifications may be necessary. At the present time, it is the barrier that separates the lagoon from the sea that is used as a defining feature for statutory sites (Bamber *et al.* 2001). Hereafter all lagoons and lagoon-like habitats will be referred to simply as lagoons.

*Lagoon classification*

A 'typical' lagoon may be tideless or experience very low tidal movement, will contain brackish water and be able to exchange water with the sea/estuary via a channel or by percolation and where a channel is present it is usually smaller than both the lagoon and the separating barrier (Barnes 1980). Based on this physiography lagoons can be classified into five types, (i) isolated saline lagoons, (ii) percolation saline lagoons, (iii) silled saline lagoons, (iv) sluiced saline lagoons and (v) saline lagoon inlets (Downie 1996) (see also Table 1.1a). Bamber *et al.* (2001) report Barnes (1980) as grouping these types by emphasising the degree of connection between the lagoon and sea-water; (i) *open lagoons* - include silled, sluiced and lagoonal inlets, where sea-water exchange, although restricted is open, (ii) *percolation lagoons*, where sea-water exchange occurs through shingle or sand and (iii) *no-flow lagoons*, isolated lagoons, where saline input is via groundwater or over-topping of a non-permeable barrier (distinct from percolation lagoons as there is no exchange of sea-water). The connection between the surrounding sea and the lagoon has implications for the recruitment of species and for the maintenance of the lagoon habitat (Bamber *et al.* 2001). If the link between the sea-water and the lagoon is lost then the lagoon will evolve into a freshwater ecosystem and most if not all associated flora and fauna will be lost (Bamber *et al.* 1992). Conversely, during high tide, lagoons have the potential to become fully saline or even in some conditions hyper-saline (Bamber *et al.* 2001).

**Table 1.1a. The five main sub-types of lagoons (based on Sheader & Sheader 1989) used to identify statutory sites (after Bamber *et al.* 2001)**

| Type of Lagoon | Description |
|---|---|
| Isolated | Completely separated from the sea/estuary by a barrier of rock or sediment. Sea water enters by over-topping or limited ground-water seepage. Predominant inflow is fresh water but salinity may remain high because of evaporation. Tidal range is zero, longevity is likely to be brief |
| Percolation | Normally separated from the sea by a shingle bank. Sea water enters by percolation or occasional over-topping. Tidal range normally significantly reduced or may be unnoticeable. Longevity may be extended but rapid or immediate extinction can occur when surrounding sediment moves. |
| Silled | Lagoons with an open connection to the sea, water is retained at all states of the tide by a horizontal barrier of rock (sill), sea water exchange is regular and frequent. Tidal range may be marked (but reduced), although out of phase with the adjacent sea. Potentially long lived lagoons, extinction is normally an extended and gradual process. |
| Sluiced | Lagoons with an open connection to the sea, water exchange is modified by anthropogenic mechanical structures, valved sluices or sea-doors. Such lagoons restricted by culverts may fall into this or the previous category, although the latter is usually thought to refer to natural lagoons. Longevity as for silled lagoons, but tidal range will depend on the efficiency of the sluice and may be very low. |
| Lagoonal inlets | Open connection to the sea but the connection is narrow. Tidal range may be marked (but reduced), commonly in phase with the adjacent sea; longevity as for silled lagoons. |

### *1.1.2 Formation of Lagoons*

Wave action has been responsible for the creation of many offshore barriers, and thus lagoons, and has been aided by a rise in sea level of 100m over the last 15 000 years (Bird 1970). Shingle is not as common as sand in offshore barriers, owing to its weight, but can sometimes be responsible for barrier formation, especially in circumstances where storm waves have been active (Barnes 1980). Longshore barriers formed through the creation of spits which may seal off bays, have been responsible for the creation of some lagoons e.g. those in the Black Sea (Barnes 1980). Storms can also be responsible for lagoon formation; if storm waves over-top existing barriers or if masses of shingle

from the ends of spits are moved. Additionally land subsidence, sediment redeposition, and river channel changes can all contribute to lagoon formation (Barnes 1980). As well as acting to create lagoons natural processes can act to reduce the number of lagoons. For example, some lagoons evolve into freshwater systems as percolation of sea water is prevented by a build up of debris on the bottom of the lagoon, and over-topping may be reduced by storm raised shingle (Bamber *et al.* 2001). Owing to the many ways in which they may form, lagoons can vary considerably in size. The mean size of lagoons in the UK is 10 ha, but some are as small as 1 ha (Bamber *et al.* 1992). All lagoons, but the smaller sites in particular, have considerable variations in temperature and salinity and these factors, especially when combined with a precarious sea-water link, make lagoons an extreme and marginal habitat (Downie 1996).

### *1.1.3 Distribution of Lagoons*

Owing to the macrotidal conditions of the North Atlantic the lagoon habitat in the UK is not widespread (Barnes 1991), with lagoon sites being restricted primarily to England and Scotland (Bamber *et al.* 2001) (Table 1.1b). The total number of lagoons in England is 177, covering an area of 1200 ha, and these lagoons are clustered together on the south and east coasts (Smith & Laffoley 1992). In these areas there is a predominance of low lying land bordering estuaries and it is particularly here that lagoons are located. Davidson *et al.* (1991) have quantified this phenomenon and report that over 83% of lagoons are closely associated with estuaries (Bamber *et al.* 2001).

In addition to the uneven scattering of lagoons around the UK there is a skewed distribution in the type of lagoon in each of these areas (Table 1.1b). For example, silled lagoons, are predominantly found in the Outer Hebrides (Bamber *et al.* 2001), whereas in East Anglia, glacial offshore deposits have left a mass of shingle, resulting in the formation of many percolation lagoons (Bamber *et al.* 2001). These are examples of how specific local geomorphological processes have shaped the lagoon habitat around Britain.

**Table 1.1b. Frequency of different lagoon types around UK* ((data from Smith & Laffoley 1992; Bamber *et al.* 1999; Covey 1999) after Bamber *et al.* 2001).**

| | Scotland | Wales | England |
|---|---|---|---|
| Total number | 139 | 12 | 177 |
| Sluiced Lagoons | 27% | 59% | 49% |
| Silled Lagoons | 39% | 0% | 0% |
| Lagoonal Inlets | 22% | 25% | 9% |
| Percolation Lagoons | 8% | 8% | 18% |
| Isolated Lagoons | 4% | 8% | 24% |

* There are also 30 lagoons comprising 42 ha in Northern Ireland

### *1.1.4 Biota of lagoons*

The correlation between environmental surroundings, that is the physical habitat created by factors such as the salinity and biological entities, that is the community of species present in that environment is married together by the idea of a 'biotope'. Where a specific habitat type has been defined (see Bamber *et al.* 2001 after Connor *et al.* 1997) along with a known and documented group of species, it is regarded as a biotope (Bamber *et al.* 2001). Bamber (1997) has identified several biotopes found in saline lagoons for England and Wales, the most common of which are characterised by muddy sand to mud-sedimentary substrata inhabited by associated infauna and submerged plants with their associated fauna (Bamber *et al.* 2001). Larger lagoons may often present several biotopes e.g. The Fleet in Dorset, and lagoons which have been long isolated may demonstrate unexpected communities. It is, however, usually possible to predict species composition in a particular lagoon from the physical conditions present.

The species present in lagoons usually comprise a mix of freshwater, marine and estuarine, with the majority being marine or related to marine species (Bamber *et al.* 2001). Barnes (1991) suggests the occurrence of each of these groups will be highly dependent on the salinity in each lagoon. Bamber *et al.* (1992) have gone further and identified six suites of lagoon species and propose that these reflect the degree of connection the lagoon has with the sea (Table 1.1c). For example, sluiced lagoons have reduced water exchange with the surrounding sea and are usually dominated by lagoon specialists. Limited colonisation opportunities for other marine species and the

marginal environment created by the limited water exchange are thought to be responsible (Bamber *et al.* 1992). This hypothesis is partially supported by Covey (1999) who found that the mean species number present in a lagoon increased as the connection with the sea increased. Additionally Sheader & Sheader (1989) noted that the number of species increased as the area of the lagoon increased. This relationship, however, is probably complex as the effect was only seen in silled lagoons (Sheader & Sheader 1989).

**Table 1.1c. The 'six suites' described by Bamber *et al.* (1992) of species found in British lagoons**

| Suite | Description of species type |
|---|---|
| 1 | freshwater/low-salinity species |
| 2 | lagoonal species that are under-recorded |
| 3 | euryhaline specialist species that are tolerant of estuarine conditions |
| 4 | stenohaline marine specialist lagoonal species |
| 5 | estuarine species pre-adapted to lagoonal conditions |
| 6 | estuarine species incidental in lagoons |

In addition to natural lagoons, artificially created lagoons can offer a first hand view of the succession of species in a new lagoon. Some of these developments even attract lagoon specialists, for example six years after recolonisation of redeveloped docks in Merseyside, three species of lagoon specialists were found (Allen *et al.* 1995). It is most likely that their route of colonisation was from local populations in the surrounding estuary area, as other lagoon-like habitats were in the proximity (Allen *et al.* 1995). Barnes (1988) reports the importance of geographical proximity to other lagoons for the colonisation of new lagoons. This places obvious restrictions on the potential for some lagoon communities to develop and in these situations anthropogenic intervention may be necessary. One of the advantages offered by artificial lagoons is as an intermediary between existing lagoons aiding the dispersal of species around the UK (Allen *et al.* 1995).

*Lagoon Specialists*

In Britain there are 9 species that are known only in lagoons and 38 species that are more typical to the lagoon environment than to other aquatic habitats (Barnes 1989) and Bamber *et al.* (2000) offer a list of those species more common to the lagoon habitat than any other, but, report the need for a more comprehensive review in order to produce a more up to date list. A phenomenon characteristic of lagoons is the presence of a high number of individuals from just a few representative species (Bamber *et al.* 1991). Species that are found in the lagoon environment are probably better able to tolerate the wide ranges in temperature and salinity than their marine/estuarine counterparts and as a result of reduced competition are comparatively successful (Bamber *et al.* 1992). In spite of the prevalence of such species in lagoons it is not unusual to find quite common species present e.g. the lug-worm (*Arenicola marina*). Species such as these are able to inhabit a number of wide ranging habitats (Bamber *et al.* 2001). Species diversity in lagoons seems to be better preserved where a high circumference to area ratio is maintained (Bamber *et al.* 1992). The communities supported in lagoons comprise an essential part of the habitat in terms of prey species, potential predators and as sites of attachment (Sheader *et al.* 1997).

*Preference for lagoon specialists*

Many natural lagoons are characterised by occasional anoxia in the benthic environment, caused by high macrophyte production and this is likely to be an important factor in favouring lagoonal species (Allen *et al.* 1995). Lagoon specialists may also out-compete typically marine species in those lagoons where there is regular variation over a wide range in both pH and salinity (Bamber *et al.* 1992; Allen *et al.* 1995). The current flow of a lagoon may have an influence on the species that inhabit them. For example *Nematostella vectensis* is found in abundance in those lagoons showing insignificant or no near-bottom flow rates (Sheader *et al.* 1997). Factors contributing to currents in lagoons include; tidal exchange, riverine input and wind (Sheader *et al.* 1997). Furthermore, gradients in grain size, often seen in lagoons as a result of internal water flow and the influence of shingle barriers (Barnes 1980), create heterogeneity in the sediment and encourage diversity in the benthic biota.

As well as the current and sediment heterogeneity, the substratum and fringing habitat are considered important factors in giving lagoonal biota an advantage over marine species (Allen *et al.* 1995). For example most lagoon specialists are sublittoral and sediment dwelling, (Bamber *et al.* 1992) and therefore require little cryptic fringing habitat or hard substratum to thrive. The variable environment afforded by the conditions in a lagoon is often too stressful for many estuarine and marine species, with the result that true lagoonal communities are often species poor. This unpredictable environment gives an advantage to the better-adapted lagoon species that are not out-competed by their marine/estuarine counterparts. It is interesting to note that the absence of many lagoon specialists from artificially created lagoons is often related to the stability of the habitat (Allen *et al.* 1995).

### *1.1.5 Conservation of lagoons*

The environment and its ecosystems provide the atmosphere, the energy and material sources that are required to sustain life on Earth (Spellerberg 1996). Biodiversity plays an immensely important role. The term biodiversity encompasses species, populations, habitats and ecosystems. Wilson (1992) defines biodiversity as "The variety of organisms considered at all levels, from genetic variants belonging to the same species through arrays of species to arrays of genera, families and still higher taxonomic levels; including the variety of ecosystems, which comprise both communities of organisms within particular habitats and the physical conditions under which they live". There is a need to conserve organisms that provide food, materials, drugs and those which play a vital role in maintaining the balance of the Earth's atmosphere and structure of its soil (Wilson 1992). In addition, it is likely there are many organisms that may prove useful but are as yet, undiscovered (Cox & Moore 2000).

There is a proven link between biodiversity and ecosystem function in terms of nutrient cycling efficiency, resilience to disturbances, and removal of low levels of toxins (Spellerberg 1996). The implication being that where there is greater biodiversity (and thus increased ecosystem service) there will be greater productivity, nutrient recycling and stability. The conservation of biodiversity, is therefore, about more than the

preservation of wildlife, it is about ensuring the continuation of natural ecosystems which act to buffer us from storms and floods and which provide our life support systems. The importance of conservation at the level of ecosystem has been seen in controlling the rate at which natural processes occur. For example deforestation and drainage of swamps have affected the way water moves through and over the land (Spellerberg 1996). The loss of this type of biodiversity and its resulting destruction to biological systems is something that can be understood at large, however, the conservation of a few biological communities is sometimes publicly regarded as unnecessary. There is a need to understand that the biodiversity of communities is immense, and that there are not just a few remaining examples but many. For example, in Europe, there are many varied types of woodland and each community has its own unique assemblage of species (Spellerberg 1996). The biological community is dependent on the variety of species to create the ecosystem. Furthermore, Collins (1995) suggests that the diversity of species in communities may affect the way they and consequently their ecosystems function. There is also a strong aesthetic argument for conserving biodiversity, and an argument that all organisms have a right to exist on the Earth (Wilson 1992). Humankind has a high impact on the ecosystems of the Earth and thus a responsibility to minimise loss and maximise conservation and survival (Anon 1995).

Species richness is often used to describe biodiversity and includes the idea of interactions between species, and because different groups of species can be found together in different sites it is important that any species list includes the names of the species (Spellerberg 1996). On a finer scale, one way to monitor and measure the biodiversity within a species could be to analyse populations of a species and the individuals within each population. Of central importance to all levels of biological diversity is genetic diversity. Genetic diversity has been and is the basis of evolutionary process and is the prerequisite for all other kinds of biological diversity (Wilson 1992). To this end, measuring genetic diversity may provide possible ways to quantify biological diversity within a species. Measuring genetic diversity also provides a means to estimate migration in populations. This is a crucial factor in the regional maintenance/survival of species which conform to a metapopulation structure.

For the purposes of this study it is hoped that an understanding of the genetic diversity and hence the biodiversity of certain key species present in lagoons may go some way to aiding the conservation of the lagoon ecosystem. In the UK there are few studies of the isolated animal populations associated with the lagoon habitat although the distribution and classification of saline lagoons are well-documented and limited data are available on the habitat, community and population structure of these lagoons (Sheader & Sheader 1985a; Barnes 1988; Sheader & Sheader 1989; Bamber *et al.* 1992; Bamber 1997). Bamber (1997) has indicated that an understanding of the distribution of lagoon habitats and the ecology of the associated flora and fauna is essential in managing the fragile lagoon resources in this country. Work done to improve our knowledge of the biodiversity present in lagoons may provide an example of conservation at a specific level of biodiversity

Humans have a high impact on biodiversity sometimes by deliberate extermination e.g. hunting or inadvertently through tourism, recreation or commercial exploitation of species, additionally the effects of pollution on species and ecosystems and the effects of changes in land use have had severe consequences on species, biological communities and ecosystems (Spellerberg 1996). Additionally global climate change can affect ecosystems no matter how far they may be from the source of the change. For example, rising sea levels resulting from an increase in global temperature causing the polar ice caps to melt, may pose a threat to the lagoon habitat because of an increase in storms and the potential for coastal erosion. Furthermore, there are many natural factors which can act to reduce the number of lagoons in the UK, e.g. natural evolution into freshwater systems. Conversely there are many lagoons which have been formed or preserved by human action and/or intervention (Bamber *et al.* 2001). Lagoons are a nationally rare habitat and this is reflected in their status as the only marine priority habitat in the UK (Bamber 1997). It is not just the paucity of lagoons which make them an important habitat. They act as a food source for marine, estuarine and lagoon invertebrate species, and in Europe they provide more habitat for breeding and migrating birds than any other coastal habitat (Bamber *et al.* 2001). They are typically characterised by some of the rarest flora and fauna found in the UK and provide habitat for many rare or scarce insects which require a particular environment to complete life

history stages (Bamber *et al.* 2001). It is hoped that with increased knowledge of the biodiversity, ecology and ecosystems of lagoons we can go some way to increasing the lagoon habitat and its valuable biota.

## 1.2 Islands and Lagoons as an island–like habitat

### *1.2.1 Introduction*

Since Darwin first visited the Galápagos, biologists have used islands to study both ecology and evolutionary concepts. Islands when compared to either continental landmasses or the oceans of the world provide comparatively simple ecosystems (MacArthur & Wilson 1967). Continental landmasses and oceans are under continuous flux of immigrants and emigrants, whereas islands are isolates with reduced immigration and emigration. Moreover, islands can be regarded as special 'traps' for species which disperse to them and successfully colonise them. Islands may be engines of speciation and are often characterised by endemism. To this end island species are often regarded as a 'special case', of communities, driven by the interaction between colonisation and extinction (Wilson 1992). Furthermore, the study and analysis of the biota of islands and the variety found between different islands may be used to predict how organisms may react to changing conditions in other circumstances, and may be an essential component of any conservation management strategy (Cox & Moore 2000).

*Islands as natural laboratories*

Traditional studies classify islands according to geological origins (Table 1.2a) and the biotas associated with these three island types are very different, reflecting their origins. Islands formed from continents would have been populated by organisms native to that continent, although conditions on the island may have changed and evolution and extinction would have acted reflecting those changes (Krebs 1994). On arcs and hotspot islands the flora and fauna would have dispersed to the islands by crossing sea or ocean and as the island matured progressive ecological change would have been observed and eventually evolution and extinction may have acted to change the original colonisers (Cox & Moore 2000). Additionally, islands formed from hotspots are often part of a chain (Table 1.2a) and this may allow for inter-island dispersal which may result in more complex evolution and speciation (Cox & Moore 2000) .

The potential for islands to demonstrate measurable ecological succession has made them the subject of research (Hesse *et al.* 1951; Simberloff & Wilson 1970; Wood &

del Moral 1987; 1988). When established islands have been made sterile, either through extreme environmental disturbance (Hesse *et al.* 1951; Wood & del Moral 1987; 1988) or by purposeful human causes (Simberloff & Wilson 1970) a unique opportunity arises to observe dispersal onto the island and colonisation of the island by new organisms. As well as demonstrating the succession of species, islands have been used as natural laboratories to investigate predator-prey interactions (Schoener & Toft 1983; Spiller & Schoener 1990). Experiments involving progressive ecology have allowed workers to put to test the hypothesis regarding colonisation and extinction of species on islands proposed by MacArthur and Wilson (1967). It is, however, worth mentioning, that historical effects can have an important influence on the biota of an island and may mean the theory of island biogeography is not a paradigm. For example, birds that inhabit the isolated mountain ranges in the Great Basin in the western United States, show no effect of isolation by distance as predicted by the MacArthur-Wilson model and the rate of bird faunal turnover is very low (Krebs 1994).

**Table 1.2a. Islands according to their geological origins (Cox & Moore 2000)**

| Island type | Origin | Example |
|---|---|---|
| One - Continental* | Separated by rising sea levels | Britain |
| | Separated by tectonic processes | New Zealand |
| Two - Volcanic arc** | Arcs form along trenches where ocean crust disappears into the earth. These collide with land masses to form: | |
| | (i) Volcanic islands | Krakatoa |
| | (ii) Mountain ranges | $N^{thm}$ New Guinea Mountain range |
| Three - Hotspot islands** | Plumes of hot material rise to form volcanoes on the Earth's surface. In the ocean, may remain submerged forming a seamount, or above the water will form a volcano. This process forms chains of islands as the oceanic plate is always moving. | Hawaiian chain of islands |

* Also known as landbridge islands ** Also known as oceanic islands

Island biogeography and the equilibrium theory of MacArthur & Wilson (1967) may also be applicable to ecological islands. Ecological islands are patches or areas which are subject to similar dispersal and colonisation pressures as terrestrial and oceanic islands. They consist of populations that may have a fragmented or patchy distribution and colonisation will probably occur as a result of jump dispersal and not diffusion. Jump dispersal is the movement of individual organisms across large distances (of usually unsuitable terrain) followed by the successful founding of a population (Pielou 1979). Dispersal by diffusion, the commonest form of dispersal, involves the gradual movement of a population across hospitable terrain (Pielou 1979). Examples of experimental terrestrial and marine ecological islands can be found in the literature. For example Spiller & Schoener (1990) used 'island' enclosures to monitor lizard and spider competition and reproductive success and Schoener (1974) used plastic 'sponges' in barren parts of a lagoon to observe colonisation. Seamounts are an example of oceanic islands only inhabited by marine species (Table 1.2a) and these appear to be subject to a number of ecological and genetic pressures similar to terrestrial islands (Stockley 2001). Hydrothermal vents, individual patches within coral reefs, rock pools under tidal influence and coastal lagoons are all examples of marine ecological islands. Studies on the ecology of these 'islands' do exist (Johnson & Black 1998; Forrester 1990; Tunnicliffe 1991; Bamber *et al.* 1992), and it is hoped that the present study may go some way to furthering our knowledge of the genetic pressures that 'island' lagoons are subject to. An understanding of the communities of such ecological islands may help us to understand the dispersal of species to them and how successful colonisation arises.

In comparison to terrestrial and oceanic islands, it is not always clear what is preventing species from dispersing to ecological islands and from successfully colonising them. Although ecological islands usually consist of patchy and/or fragmented habitat they do not necessarily have clear boundaries preventing dispersal. In the marine environment, ocean current patterns, sea floor topography and other geographic features may provide isolating mechanisms, but only for a small number of species (Waples 1998). Often, however, it may not be clear why species 'choose' one island over another when they appear to have access to both, a phenomenon seen in some species of coral reef fishes (Krebs 1994). While this aspect of ecological islands is not clearly understood, the fact

that some species inhabit such islands means that their populations will be fragmented and patchy. When a population is disturbed in this way, the potential for gene flow between individuals of these sub-populations alters, and this may have implications for the divided populations and the species as a whole (Cox & Moore 2000). How do species, when limited to small populations, cope with genetic isolation from other populations? Do they evolve into new species? Are they more prone to extinction? It is also important to understand their ability to disperse to other populations or to consider whether their isolation will strand them. Studying all kinds of island populations, including ecological islands such as lagoons may help to give us a clearer understanding of the effects of such isolation. One crucial aim of studying island populations is an understanding of the population dynamics of a species. Species which form metapopulations, made up of several sub-populations, depend on migration and/or dispersal. This movement of individuals maintains the sub-populations, which may be subject to periodic extinctions, and ensures the survival of the species, whereas a species which exists as a large permanent population with several peripheral island populations will not necessarily depend on migration and/or dispersal for its survival. If lagoon species form metapopulations then dispersal is critical as too little will lead to extinction of the metapopulation or even species. It is clear that a knowledge of population dynamics is fundamental to any conservation management plans.

### *1.2.2 Lagoons as islands - the effects of dispersal*

Lagoons are separated from each other by land and expanses of water and thus new colonisation probably has to occur across open waters. It would, therefore, appear that chance is probably responsible for the allocation of some species to lagoons (Barnes 1988). Many marine species have evolved extended pelagic larval stages which are able to disperse widely and some species are able, as adults, to migrate extensive distances (Waples 1987; Waples 1998) (where dispersal refers to the non-directional movement of individuals and migration to the movement of individuals with direction). Many lagoon specialists, however, do not demonstrate a life history stage compatible with long distance dispersal. Moreover, some species appear to have developed life history stages congruent with remaining in the natal lagoon. Where specialist lagoon fauna do

exhibit planktonic larvae this is very often short lived, usually lasting no longer than a week (Barnes 1980). An example of one such species is *Cerastoderma glaucum*, the lagoon cockle. The marine counterpart of this cockle is *Cerastoderma edule*, the common commercial cockle. These two cockles differ in their ecology in several ways (Russell 1972; Brock 1980a) including a larval stage of approximately 1 week and 5 weeks respectively (Boyden 1971; Barnes 1980) and a difference in spawning times (Boyden 1971 & Brock 1980a). It is suggested that different spawning times and the length of survival of the pelagic stage may play a part in the reproductive isolation of these two species (Boyden 1971). This, however, has implications for the dispersal ability of *C. glaucum* and this in turn may have implications for the survival of the species.

Dispersal may be favoured for several reasons; it may have evolved (i) as a risk-spreading strategy, as in metapopulation structure, (ii) as an altruistic trait to reduce kin competition, (iii) it may be favoured to escape fitness depression caused by mating among relatives (Hamilton & May 1977; May 1979; Venable & Brown 1988). Dispersal allows for gene flow, and there is evidence to suggest gene flow opposes the effects of local selection and thus limits adaptation (Lenormand 2002). Conversely dispersal may also recolonise or restock the local population and therefore the local genetic variation, both of which are required for natural selection (Lenormand 2002). So where lagoon species have to disperse long distances to reach the suitable habitat of another lagoon their survival may be compromised. In the UK the lagoon habitat is naturally disjunct as a result of specific local geomorphological processes involved in lagoon formation but has become more fragmented owing to anthropogenic neglect and subsequent loss of the habitat (Bamber *et al.* 2001). Where lagoons form part of a metapopulation this fragmentation is likely to affect the dispersal and migratory potential for species inhabiting them. This may be so great that dispersal and migration are no longer possible with the end result that regional extinctions occur as stochastic processes eliminate each sub-population. It is hoped that examination of the gene flow of three lagoon species may indicate the dispersal success of those species and these data may be a useful indicator for areas needing conservation priority.

### *1.2.3 Population genetics of island habitats*

Frankham (1997) suggests that island populations are susceptible to genetic loss both at foundation, owing to low numbers of individuals founding the populations and following foundation, because of typically low population sizes (c.f. mainland populations). This low level of genetic variation present in island populations may mean short-term evolutionary change is not possible as favourable mutations may appear in small populations infrequently (Ayala 1965). The genetic losses associated with island populations are countered by the gains brought by secondary immigration and new mutations, and may combine to provide increased genetic variation (Jaenike 1973; Lande 1993). Natural selection will, moreover, act to increase the rate of loss, by selection for favourable alleles, but may act to slow it where a heterozygote advantage exists (Frankham 1997).

Where small populations exist, there is an increased likelihood of inbreeding and there is some evidence to suggest that inbreeding leads to extinction, owing to inevitable mating between closely related individuals (Frankham 1998). That inbreeding reduces reproductive success has been shown in both captive and wild populations (Frankham 1995a). Other workers, however, suggest that demographic and environmental factors play a greater role in a population's decline towards extinction (Lande 1988). Work by Saccheri *et al.* (1998) on the Glanville fritillary butterfly (*Melitaea cinxia*) found that where less genetic variation existed in the population there was an increased chance of extinction. Even after accounting for all known demographic and environmental factors, genetic diversity could predict extinction risk (Saccheri *et al.* 1998). Evidence exists which suggest that these results may be applicable to other species. Theoretical models suggest that genetics makes a significant contribution to extinction, relative to other factors (Frankham 1997). Island populations may show high inbreeding levels which, when seen in captive-bred counterparts, have led to an increased risk of extinction (Frankham & Ralls 1998). In field studies of wild plants, the extinction rate was found to be higher in those with low genetic variation (Newman & Pilson 1997). Perhaps the most compelling evidence is the low genetic diversity of endangered species when compared with non-endangered species, which would not be

expected if demographic and environmental factors were the driving force behind extinction (Frankham & Ralls 1998). Evidence such as this provides a persuasive argument for the place of genetics in any management plan for threatened or declining species. It is important to note, however, that the most likely cause of low diversity in populations is often a decrease in population size and an increase in fragmentation.

It was earlier noted that of the species present in lagoons, many are marine, and will not be restricted solely to the lagoon environment. Species which are able to exploit both the marine and lagoon environment will have a greater opportunity for dispersal. Genetic data are available which give insight into the importance of dispersal among marine species and these data suggest that, in general, levels of gene flow and migration for such species are relatively high (Waples 1998). For many lagoon specialists, however, gene flow will be restricted because of the barriers associated with this habitat and because of a reduced or total lack of dispersal ability.

## 1.3 Lagoons as a fragmented and marginal habitat

### *1.3.1 Introduction*

Species living at the edge of their range, as is the case for many of those present in lagoons, often persist in small isolated populations (Nielson *et al.* 1999). Additionally, animals with highly specialised habitat requirements naturally occur in small populations (Stacey & Taper 1992). Low population size may have detrimental effects e.g. inbreeding may increase, there may be a loss of heterozygosity, and of allelic diversity (Hedrick & Miller 1992; Hedrick *et al.* 1995). It is not just isolated habitats, however, that can lower effective population size and limit gene flow. This phenomenon has also been associated with austere habitats, e.g. the lagoon habitat, and this is probably an equal factor which contributes over time to a change in allelic frequencies (Nielson *et al.* 1999).

The ecological isolation of small populations can very quickly lead to changes in genetic diversity in just a few generations (Gavrilets & Hastings 1996). A very real problem associated with changes in genetic diversity between founding populations and refugium populations is the prevention of any back adaptation to founder habitat, should a source population be eradicated (Hedrick & Miller 1992). This may occur if the long term isolation of a population leads to important genetic differences either through drift or mutation.

### *1.3.2 Habitat fragmentation*

Fragmentation is very different for terrestrial and aquatic habitats. Terrestrial fragmentation is often associated with alteration in large tracts of land (Neraas & Spruell 2001). Aquatic fragmentation can, however, take many forms, from isolation in pools along the coast (Johnson & Black 1998) to fragmentation caused by dams and irrigation systems in rivers (Neraas & Spruell 2001). Habitat loss and fragmentation threaten species throughout the world. The adverse effects of fragmented habitats include; reduction in local population size, reduced migration, increased population size

fluctuations and inbreeding depression (Tallmon *et al.* 2002). Fragmentation may limit many species to a small number of habitat patches with often unsuitable environment in-between and this kind of fragmentation can push species to the very edge of their normal range (Tallmon *et al.* 2002). This is especially true for those species with limited dispersal capabilities who will be more susceptible to the negative effects of fragmented habitats.

Dispersal and migration between disjunct populations can have a great effect on the genetic diversity of some populations. For example where a marine species lacks a planktonic larval stage dispersal may be restricted and patterns of isolation by distance may be observed (Johnson & Black 1998). Water gaps of only 500m have been shown to double the genetic subdivision for the intertidal snail *Bembicium viltatum* which relies on adult migration for dispersal (Johnson & Black 1998). Another consequence of fragmentation on the migration and dispersal of individuals has been demonstrated in some fish species. Research has shown, where species have a variable life history, movement between habitats is essential to express the complete life cycle. There is, moreover, some evidence that variable life histories may buffer a population against extinction. If habitat fragmentation prevents expression of all life histories, for example some migratory fish spawn in small streams but require large bodies of water for other life history stages, there may be an increased chance of extinction (Neraas & Spruell 2001).

The distribution of genetic variation is an excellent method for detecting dispersal. In light of the fact that the distribution of genetic variation between populations can be a result of historic *or* current gene flow, however, genetic analysis can sometimes be inadequate for estimating migration among populations (Milligan *et al.* 1994). It is also important to recognise that the dispersal potential inferred from a species larval life history may be considerably different to that observed in the form of gene flow (O'Mullan *et al.* 2001).

Where small new populations are established it would be expected that the founders will contribute only a proportion of their alleles to progeny and that genetic drift will

eliminate some alleles, resulting in an overall loss of heterozygosity (Neeras & Spruell 2001). The same would be true of recently isolated small populations which may demonstrate reduced fitness and lower genetic variation. As a cautionary note, it must be recognised that sample size artefacts can result in fragmented populations demonstrating lowered heterozygosities and a reduced numbers of alleles per locus (Tallmon *et al.* 2002).

Most molecular studies of genetic variation, where little or no loss of genetic variation within populations is observed and where there is also little divergence between the populations, will be interpreted as large effective population sizes or ongoing gene flow between populations (Tallmon *et al.* 2002). Consequently where small fragmented populations are observed to be unaffected by genetic drift, it may be that large $N_e s$ (where $N_e$ is the effective population size and $s$ is the selection coefficient) are maintaining pre-isolation levels of genetic variation. Alternatively, in such populations the gene flow from other populations is sufficiently high to counteract the effect of drift (Tallmon *et al.* 2002).

### *1.3.3 Marginal habitats*

Dispersal to a new habitat usually means adaptation to a marginal habitat will be required by the colonising species. This poses a problem, as natural selection favours individuals that choose a habitat where more of their offspring survive. Individuals dispersing to marginal habitats are colonising poorer habitats and are less likely to raise progeny and so will be selected against (Krebs 1994). In this way marginal populations are regarded as demographic sinks where immigration (at least initially) is necessary to maintain the population (Dias 1996). Conversely the immigrants arriving, carry genes that are adapted to the core (not the new marginal) environment for that species and this gene flow is likely to swamp any local adaptation (Dias 1996). Considering these factors we need to understand when adaptation to a marginal sink will be most likely and what rates of dispersal between the core and marginal habitats will bring the most success (Kawecki 2000).

Some workers have addressed this problem by observing the fate of rare alleles that improve the chances for the population in the marginal habitat, while reducing fitness in the core habitat. There have been two kinds of approaches to understanding this problem; the classic population genetic multiple-niche model (Maynard Smith 1970; Hedrick *et al.* 1976) and those which include explicit source-sink population dynamics (Holt & Gaines 1992; Kawecki 1995). The first assumes dispersal between the habitats affects the genetic composition of the local habitats, but not their size and in this model complete isolation of the new population, that is no dispersal, is the most favourable (Maynard Smith 1970). This model, however, ignores the possibility that the local population may become extinct when cut off from the core population and that dispersal affects local population size (Kawecki 2000). The second approach compares the sensitivity of overall fitness, averaged over habitats, to lifetime reproductive success in different habitats, this approach chooses to sacrifice genetics altogether (Kawecki 2000). In contrast to the first model, where there is less dispersal, there will be more stringent conditions for the invasion of a beneficial rare allele into the marginal habitat (Kawecki 1995). Recent work by Kawecki (2000) reconciles these two models. In this work he reported that the fitness-sensitivity models fit where an allele has an infinitesimally small effect on the fitness of the core habitat. Where the two habitats are, however, of equal quality and exchanging the same number of migrants, a rare allele will be most favourable where complete isolation occurs, which fits the assumptions made by the classic multiple-niche genetic models (Kawecki 2000).

Adaptations to a new natural environment are likely to be polygenic, based on the knowledge that most ecological characters are affected by many loci with small effects (Holt & Gaines 1992). Kawecki (2000) suggests that dispersal and gene flow into a marginal habitat will have qualitatively different effects on the survival of the new population, dependent on whether adaptation to the new environment involves one locus with a large effect, or several loci with small effects. Another important aspect of dispersal and migration into marginal habitats is assortative mating. Research has shown that species adapted to the marginal habitat will preferentially mate with other marginals and not with new arrivals, furthermore, the progeny of marginal crosses are more likely to survive than progeny produced through outcrosses (Rawson & Burton

2002). This behavioural adaptation will act to preserve 'marginal habitat genotypes' and to reduce the effects of immigration. Works such as these highlight how the success of populations in marginal habitats is dependent on ecological, behavioural and genetic aspects of dispersal. Furthermore, studies on the evolution of populations in marginal habitats need to incorporate the effects of all of these.

*Lagoons as a marginal habitat*

Lagoons are considered an extreme marginal environment being both eurythermal and euryhaline. They are, moreover, relatively ephemeral, lasting no longer than 1000 yrs (Downie 1996). Some of the species found in lagoons, which are specialists of the habitat demonstrate great eurytolerance e.g. the lagoon anemone *Nematostella vectensis*. This is a characteristic associated with other colonising species of sea anemone (Shick & Lamb 1977). Other species present in lagoons, that are more common to the marine environment, may be able to succeed in the hostile environment characteristic of lagoons by a switch in life history. This is also a phenomenon demonstrated by *Nematostella vectensis* which is able to reproduce both sexually and asexually and appears to favour the latter life history in the harsh environment of UK lagoons (Pearson *et al.* 2002). Lagoons present a habitat right at the edge of the known species range for many marine and estuarine species. Further investigation is required into the life history strategies, rates of dispersal and the gene flow of these species to evaluate any differences that may be observed as a result of living in this marginal habitat.

## 1.4 Genetics and conservation

### *1.4.1 Introduction*

The best and most cost effective way of preserving biodiversity is to protect and maintain natural habitats (Lande 1988). Understanding the ecological and evolutionary dynamics of small populations is becoming increasingly important as habitats become smaller and more fragmented (Lande 1988). It is important to appreciate how a genetic approach can complement previous ecological studies by giving information about previously unknown parameters. For example, molecular genetics can provide information relating to gene flow, migration, reproductive success and information about the history of a population with regards to its abundance and success (Avise 1994). Such information is of direct relevance to conservation of species present in lagoons.

We can ascertain genetic variation of populations so that those with unique genotypes, if they exist, can be identified (Pearson *et al.* 2002). Data on migration rates can be used to assess likely colonisation rates of new lagoons or those in which populations have been reduced. Where species conform to metapopulation dynamic models, estimates of gene flow between sub-populations can provide 'early warning' of the long term prospects for regional survival. Genetic studies can reveal novel and exciting data about the reproductive biology of target species (Pearson *et al.* 2002). It is becoming increasingly common that molecular genetics data form an integral part of conservation management strategies of threatened species (Mace *et al.* 1996).

### *1.4.2 Population genetic analysis*

Random genetic drift is often responsible for reducing genetic variation in small populations. A reduction in genetic variation may lead to homozygosity and the loss of the ability to adapt to environmental change (Lande 1988). These changes may then result in the depletion or extinction of a population. Should this occur, questions arise such as, (i) how is genetic variability maintained within a finite population and (ii) how

is a depleted population replenished by recruitment from elsewhere, and if so at what rate?

Wright's concept of effective population size (Wright 1943) is a useful way to understand this first question. It considers an ideal population with $N$ individuals, with discrete generations reproducing by random union of gametes (Lande 1988). $N_e$ is the effective size of a population, and refers to the number of individuals in an ideal population that would give the same rate of random genetic drift as in the actual population (Wright 1943). Where factors such as mutation, immigration and selection for heterozygotes are absent, the maintenance of genetic variation will be reduced and the expected rate of loss of heterozygosity is

$1/(2N_e)$ per generation (Lande 1988).

In a single generation of randomly sampled gametes, only rare alleles are likely to be lost, and as these only make a small contribution to heterozygosity in one generation, very little genetic variation will be lost (Lande 1988). This will be additive though, and small populations lasting several generations will be susceptible to a great deal of loss of genetic variability. Additionally an increase in homozygosity of deleterious mutations, kept rare in large populations by selection, may become fixed (Falconer 1981). One important aspect regarding inbreeding and reduction in population size is the speed at which it occurs. When these processes occur gradually, inbreeding depression has little permanent effect as selection acts to rid homozygous deleterious recessive alleles from the population (Falconer 1981).

One approach to answering the question of replenishment would be to estimate the rate of exchange among populations based on genetic data, usually employing the population parameter $F_{ST}$ (Waples 1998), the results of which can be used to estimate the migration parameter $mN_e$, as

$F_{ST} \approx 1(1 + 4\,mN_e)$ (Wright 1943).

Where the term $mN_e$ is the number of genetically effective migrants per generation received by each population .

### *(i) Random sampling error*

A major concern regarding population genetic analysis is random sampling error. There are two elements to this; intra-locus error and inter-locus error. Intra-locus error is a function of the number of individuals sampled and inter-locus error a function of the number of loci examined (Waples 1998). In particular intra-locus error can cause a upward bias in estimated $F_{ST}$ values. This problem can, however, be alleviated by increasing the sample size, where doubling the sample size can half the sampling error (Waples 1998). Chakraborty & Leimar (1987) report little difference between standard and unbiased estimates where a sample size exceeds 50 individuals. When estimating $F_{ST}$ using a sample size of 50 or more, there will be an upward bias of only 0.01 allowing intra-locus sampling error to be safely ignored (Chakraborty & Leimar 1987).

The problems associated with inter-locus error arise because unlike Wright's Island model (1943) the number of sub-populations are finite. In a finite population each gene locus arises independently, as a result of a balance between migration and genetic drift. Consequently the $F_{ST}$ variation between loci may be substantial, even when other island assumptions hold true (Nei *et al.* 1977). This problem can be overcome by increasing the number of loci, as the likelihood of a better estimation of $F_{ST}$ is increased, because using a large number of loci reduces the standard error of the overall estimate of $F_{ST}$.

### *(ii) Non random sampling*

In order that the samples collected for genetic analysis are able to fulfil the assumptions of the models used to analyse them, it is essential that any sampling performed is as random as possible (Waples 1998). To achieve this aim it is essential that knowledge of a species life history, behaviour and ecology is as complete as possible. A need for such knowledge is demonstrated by the factors which may affect the attainment of random samples; sex-biased dispersal patterns, susceptibility to capture and

demographic parameters (Jorde & Ryman 1995). Difficulties with non random sampling occur when interpreting the data. For example any behavioural trait that may have affected capture that is associated with a particular genetic character of interest, even if these traits are not strongly associated with any genetic factor, may still create bias (Waples 1998). This is because they may lead to only a certain subset of the population being sampled for example, if size and age are strongly correlated (as is the case for many marine species), any size selective sampling will bias the sample over-representing some individuals at the cost of others (Waples 1998). Replicates in space and or time can help to negate sampling bias caused by non-random sampling .

### *1.4.3 Methods for estimating gene flow in natural populations*

*F*-statistics have been the recommended method for estimating gene flow in natural populations compared with the rare allele method, owing to difficulties associated with the collection of data for rare alleles (Slatkin & Barton 1989). Statistical tests and significance levels offer a powerful way to gain insight into population structure and, when combined with robust and rich data allow for further understanding of the species under investigation (Waples 1998). It must be remembered, however, that even minor departures from random sampling assumptions could cause misleading results.
One way to deal with problems associated with random sampling assumptions is to change the approach used. For example, evaluation of the magnitude of differences between populations and the significance of these differences can provide useful insight into gene flow between populations, not just differences between the populations (Waples 1998). Importantly there has been a great deal of effort invested in recent years into improving tests for population genetic analysis (e.g. Raymond & Rousset 1995) and into evaluating tests currently available for specific markers (see Balloux & Lugon-Moulin 2002 for full review).

Ideally, when estimating gene flow between populations it is preferable to know whether the use of an isolation or a migration model is more appropriate and unfortunately an estimation of $F_{ST}$ alone does not answer this (Waples 1998). The use of several other *F*-statistics will, however, help to resolve such a question. When used

in conjunction with one another *F*-statistics are a powerful tool as $F_{ST}$ looks at the level of inbreeding in sub-populations with relation to the total (global) population, and $F_{IS}$ and $F_{IT}$ look at individual inbreeding in relation to sub-populations and total population respectively. In addition to gene flow models, $F_{ST}$, when used in conjunction with geographic distances in isolation by distance models, can reveal exciting data with respect to historic processes (Slatkin 1993). For example, Slatkin (1993) proposed that recent colonisation was evident for species with relatively low $F_{ST}$ values showing no evidence of isolation by distance. Conversely, he suggests that no gene flow is present where a lack of isolation by distance is found with high $F_{ST}$ values. Other statistical estimators for measuring the recent historical event of a species are the estimators $h$ and $\pi$, which can be used to measure genetic diversity of mitochondrial and nuclear genomes (Benzie *et al.* 2002; Arnaud-Haond *et al.* 2003).

### *1.4.4 Demography and population genetics for biological conservation*

A fundamental issue that concerns conservationists are those factors causing or contributing to extinction of species. The demographic factors that affect population dynamics include social structure, life history variation caused by environmental fluctuation, dispersal and local extinction and colonisation (Lande 1988). As previously discussed, inbreeding through reduction of population size, can be seen to contribute to a loss of genetic variability leading to a reduction in average fitness. Lande (1988) suggests that there is a 'practical need in biological conservation for understanding the interaction between demographic and genetic factors'.

The Allee effect suggests that there may be a lower threshold for the number of individuals from which a population can recover (Stephens & Sutherland 1999). The population decline may be caused by a number of factors none of which may be genetic. Two random factors which may effect population extinction are demographic and environmental stochasticity (Nunney & Campbell 1993). Demographic stochasticity arises because of sampling variances in the 'vital rates' of individuals, that is the rate of survival and reproduction (Lande 1988). Environmental stochasticity may result in the birth and death rates of the individuals in a population being correlated at any one time,

with some periods of relatively good survival and others relatively poor (Lande 1988), with the consequence that even small populations will be less likely to become extinct when subject to low levels of environmental stochasticity (Nunney & Campbell 1993).

It is important that future conservation plans consider both demographic and genetic factors when assessing factors such as minimum viable population (MVP), for which it has been shown that demographic and genetic models are in good agreement (Nunney & Campbell 1993). Ideally, each case should be considered separately as there may be many factors influencing decisions e.g. biology of the species in question and location and size of habitat. Most workers agree that both genetic and demographic factors are important, however, there may be circumstances where demographic factors do pose more immediate threat to the survival of a species (Lande 1988).

### *1.4.5 The genetic population structure of lagoon invertebrates*

In the present study molecular markers have been employed to determine the genetic variation of UK populations of three lagoon invertebrates; *Nematostella vectensis*, *Cerastoderma glaucum* and *Gammarus insensibilis*. Two of these species are threatened as a result of habitat loss in the UK, so this study aimed to provide important data for conservation management. Fragmentation of habitats can result in population bottlenecks, a reduction in effective population size, and a reduction in dispersal/migration between populations. These factors may lead to a decline in the genetic diversity of a species and an increased chance of sub-population and regional extinction. It appears that there is an increasing role for genetics in understanding the effects of habitat fragmentation on population persistence and evolution (Sih *et al.* 2000) and our ability to investigate the genetic population structure of a species has been made easier by the introduction of new molecular techniques such as those employed in this study.

Chapters 2 and 3 describe in detail the molecular methods used to test several hypotheses. These include the hypotheses (i) because lagoons are relatively small marginal habitats (few niches?), characterised by a high degree of temporal variation in

environmental parameters, they are likely to give rise to 'general-purpose genotypes' and (ii) dispersal between lagoons along the coasts of the United Kingdom is low and probably stochastic (passive). In addition to these hypotheses, molecular markers known as RAPDs (Randomly Amplified Polymorphic DNA) were employed to test specifically for clonality in the lagoon invertebrate *Nematostella vectensis* in Chapter 4. For the invertebrates *Cerastoderma glaucum* and *Gammarus insensibilis* (Chapters 5 & 6 respectively) microsatellite markers were developed and used to test the general hypotheses.

Understanding the patterns of gene flow between lagoons for invertebrates such as those under investigation in this study, may go some way to helping conserve them and other lagoon organisms. The need for such understanding has occurred directly as a result of a desire to preserve the rare and rapidly diminishing lagoon habitat and its associated species. In summary, the present study seeks to determine, by use of molecular markers, the genetic variability and population structure of disjunct populations of the lagoon invertebrates *Nematostella vectensis*, *Cerastoderma glaucum* and *Gammarus insensibilis* from lagoons around the UK.

# 2.0 General Methods

# 2.0 General Methods

## 2.1 Extractions

Extraction of DNA is usually the first procedure for any molecular DNA technique. There are several methods available and these range from traditional and relatively laborious methods, to modern and convenient methods involving commercially purchased kits. The choice of method depends on a number of factors; how 'clean' the required DNA needs to be; the kind of tissue the DNA is to be extracted from and the number of samples to be extracted. The extraction method involves four stages, common to almost all extractions:

1. Lysis of the cell and organelle membranes (and the cell wall in the case of plants) to 'free' DNA from cell nuclei and mitochondria.
2. Removal of organic compounds from the reaction.
3. Precipitation of the DNA.
4. Pelleting, drying and re-suspension of the DNA.

One very notable exception from this extraction protocol is the 'Chelex' ® (Biorad, Hercules, California) method which involves 'suspending' the DNA, after the cell lysis stage, on a bead column (Walsh *et al.* 1991). The DNA is used straight from this solution. This method, although extremely quick, is only suitable for very small or dry samples such as dried blood. DNA samples prepared in this way are less likely to contain PCR inhibitors than protein kinase and phenol/chloroform prepared tissues. The methods used for all extractions in this study were the phenol/chloroform-isoamyl alcohol (PCI) extraction (Sambrook *et al.* 1989), the hexadecyltrimethylammonium bromide (CTAB) extraction (Doyle & Dickson 1987) and in addition some extractions were performed using DNeasy kits (Qiagen, Crawley, West Sussex).

### *2.1.1 Phenol/chloroform-isoamyl alcohol*

A $1mm^3$ piece of tissue was obtained for use in the extraction. Each sample was prepared using a new blade to prevent cross contamination between samples. If the tissue had been stored in ethanol, it was washed in sterile 10 mM Tris-HCl, pH 8.0. Samples were then placed in 420 µl of extraction buffer (350 µl $dH_20$, 40 µl 1M Tris-

HCl, pH 8.0, 20 µl 20% sodium dodecyl sulphate [SDS] ) and 10 µl of 20 mg/ml proteinase K (Roche Diagnostics, Basel, Switzerland). Tissue samples were then incubated at 55°C for up to 3 hrs or until the tissue had digested. During incubation the samples were constantly mixed by being placed in a revolving rack. After 30 min of incubation an additional 10 µl of proteinase K was added.

Samples were then extracted in 300 µl of phenol and 300 µl chloroform-isoamyl alcohol (24:1), and then again in 600 µl of chloroform-isoamyl alcohol. Both of these solutions are organic solvents and so extract proteins from the sample. Precipitation of the DNA was then carried out using 35 µl 3M sodium acetate and 350 µl of 100% ice-cold ethanol. DNA is precipitated by 100% ethanol because of the concentration of monovalent cations in the solution. The addition of the salt increases the concentration of cations and thus the level of precipitation.

To pellet the DNA, the solution was centrifuged at 11 000 rpm at 4°C for 15 min. Finally the pellets were washed in 70% ice cold ethanol, to remove traces of chloroform and salt, dried and then re-suspended in 500 µl "buffer AE" (from DNeasy kit, Qiagen). The buffer provides a stable environment for the DNA and thereby prolongs its shelf life.

### *2.1.2 CTAB*

This extraction method is useful for tissue that contains a high concentration of mucopolysaccharides, for example the foot muscle of molluscs. CTAB is able to bind these compounds and allow the release of DNA into the solution. For the purpose of this thesis the majority of extractions involving the mollusc *Cerastoderma glaucum* were carried out using this method.

A 1mm$^3$ piece of tissue was obtained for use and each sample was prepared using a new blade and washed in 10 mM Tris-HCl, pH 8.0. To each sample 500 µl of CTAB extraction buffer (2% hexadecyltrimethylammonium bromide, 100 mM Tris-HCl (pH 8.0), 1.4 M NaCl, 20 mM EDTA), 1µl of β-mercaptoethanol, which acts as a detergent,

and 5 μl of 20 mg/ml proteinase K were added. Samples were incubated at 60°C in a rotating rack. After 30 min an additional 10 μl of proteinase K was introduced to the sample and incubation continued for a further 30-60 min. Following incubation, extraction in 500 μl of chloroform-isoamyl alcohol was performed three times. The final steps of this extraction procedure; precipitation, pelleting and re-suspension were as per the PCI extraction.

### *2.1.3 DNeasy Kits*

The Qiagen Dneasy kits were used to extract DNA from tissue where the previous 2 methods had failed. The kits are extremely versatile, being able to extract total DNA from many tissue sources, stored in a variety of ways. Following the lysis stage, the method differs dramatically from those already mentioned, as it uses a membrane combined with centrifugation, to isolate the DNA, and not organic extraction followed by ethanol precipitation.

For each sample to be extracted a $1mm^3$ piece of tissue was placed into a 1.5 ml microcentrifuge tube and 180 μl "buffer ATL" was added. 20 μl of proteinase K was added and the tube vortexed. Samples were then incubated in a rotating rack at 55 °C for 1-3 hr. Following incubation the samples were vortexed for 15 sec and 200 μl of "buffer AL" were added. The tubes were vortexed again and incubated at 70 °C for 10 min. 200μl of ethanol was added to each sample which was then vortexed to mix.

Each sample was then gently pipetted into a DNeasy mini column sitting in a 2 ml collection tube, these were then centrifuged at 8 000 rpm for 1 min. The flow through was discarded and the column placed in a new 2 ml collection tube. 500 μl of "buffer AW1" was added and the tubes centrifuged again, with the flow through and collection tubes again being discarded. After the samples were placed in new collection tubes 500 μL of "buffer AW2" was added, and the tubes centrifuged at 13 000 rpm for 3 min, to dry the membrane. The samples were placed in clean 1.5 ml microcentrifuge tubes and 200 μl of "buffer AE" was pipetted directly onto the membrane. Tubes were then

incubated at $37^0$C for 10 min. Finally the tubes were centrifuged at 8 000 rpm for 1 min in order to elute the DNA.

## 2.2 Polymerase Chain Reaction (PCR)

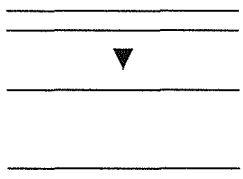
The polymerase chain reaction is an *in vitro* method for amplification of specific DNA sequences (Mullis & Faloona 1987). There are three steps involved in the technique (Saiki *et al.* 1988) which generally, but not always, (see RAPD-PCR) uses two oligonucleotide primers (short DNA molecules) to flank the region of target DNA (Fig. 2.2a).

**Figure 2.2a. The stages involved in the polymerase chain reaction**

### Step 1. Initial Denaturation

This stage involves separating the two strands of helically bound genomic DNA

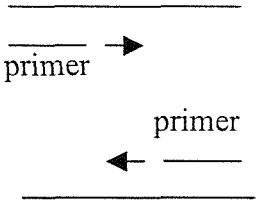
94 – 97°C for 1- 5 min



### Step 2. Annealing

The chosen oligonucleotides or 'Primers' anneal to the single stranded DNA template where they are complementary

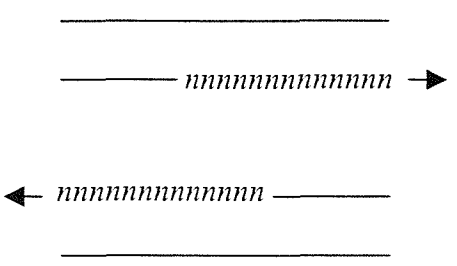
35 – 70°C for 30s – 5 min.




### Step 3. Extension

Now the target DNA is amplified by the addition of complementary nucleotides (the basic units of DNA) Adenine (A), Guanine (G) (the purines) and Thymine (T), and Cytosine (C) (the pyrimidines) (represented here by *n*) by the enzyme *Taq* polymerase.

70 – 75°C for 30s – 2 min




These steps are then repeated for between 25 – 40 cycles. In addition there is a final extension time of between 7 and 30 min to ensure that all resultant amplified DNA strands are double stranded and the reaction is completed.

### *2.2.1 Denaturation*

The DNA needs to be denatured or 'melted' in order to provide the single stranded template necessary for PCR. It also serves to eliminate any secondary structure that may be present in the molecule (Delidow *et al.* 1993). Once the genomic DNA is single stranded it is possible for primers to anneal to known complementary regions of the original template. Primers are short sequences of DNA (usually between 18 – 25 bases long) that are designed to be complementary for a specific, highly conserved region of the genome. They usually come as a pair, with one primer complementary to the single strand of DNA upstream of the target region and one to the single strand downstream of the target region. There are exceptions to this, for example the use of a single primer for the random amplified polymorphic DNA (RAPD) technique (Chp 4).

The enzyme involved in the extension step of the reaction, *Taq* polymerase, is a thermostable DNA polymerase, isolated originally from the thermophilic bacterium *Thermus aquaticus* (hence *Taq*). The use of this enzyme, first demonstrated by Saiki *et al.* (1988), allows repeated synthesis of DNA strands, leading to exponential amplification. Despite the thermostable nature of this enzyme, it is important to ensure that initial melting of high molecular weight DNA is achieved and the activity of *Taq* polymerase is maintained. Consequently the initial cycle is relatively long (e.g. 5 min)

and at a relatively high temperature (e.g. 95°C). The duration and temperature is subsequently reduced for the remainder of the cycles (Delidow *et al.* 1993).

### *2.2.2 Annealing of Primers*

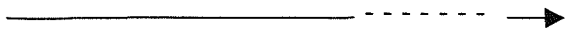
The annealing step involves lowering the temperature to allow the primers to come together with the original, now single stranded, template DNA. The temperature of this step is determined by the strand-melting temperature of the primers (Tm) and by the specificity required (Rychlik 1993). It is vital that the primer/template ratio is correctly determined for each PCR, as the specificity and efficiency of the reaction are strongly influenced by this ratio. The temperature needs to be low enough to allow the strands of DNA to anneal in complementary regions, but not so low that the primers will form bonds in random regions along the template DNA (Palumbi *et al.* 1991).

In order to maintain specificity, it is important to keep the temperature quite high (usually above 40°C), however setting the annealing temperature too high prevents the occurrence of priming. The ideal annealing temperature for new primers can be determined very quickly and efficiently by using a 'gradient block' PCR machine. This method allows the annealing temperature to be set to vary across the sample block thus permitting a range of temperatures to be tested on the same sample simultaneously.

### *2.2.3 Extension of DNA strand*

The ideal temperature for best performance of *Taq* polymerase is between 72-75°C (Palumbi *et al.* 1991). This is why the temperature in the PCR cycle is raised after annealing has been completed. A phosphodiester bond joins the 5' carbon of new nucleotides to the 3' carbon of the last nucleotide added to the growing DNA sequence (Fig. 2.2b). The amplification duration is determined by the length of the sequence amplified and is based on the extension rate of *Taq* which, under optimal conditions, is 2-4 Kb per min (Delidow *et al.* 1993).

**Figure 2.2b. Annealing of primer to a template strand and the extension of nascent DNA**



5' TACAGCCGGGGTTAT 3' Primer: nascent strand extends from here

| | | | | | | | | | | | | | | |

3' ATGTCGGCCCCAATA CCGTAAATAGTCGATCCGTTAG 5' Template

Here the role of the primers becomes obvious. They provide the initiation for the action of the *Taq* polymerase. This is true also of cellular DNA replication, where the addition of nucleotides by DNA polymerase III, cannot begin without the presence of a primer on the template strand (Brown 1992a). The template DNA also acts as the reference for which nucleotides are added to the nascent strand. When DNA is bound in the double helix the correct pairing of the bases, adenine (A) with thymine (T) and cytosine (C) with guanine (G), allows the formation of hydrogen bonds. No other combination allows this type of bond formation, which is fundamental for holding the double helix structure together (Brown 1992b) . In addition, the two strands must run antiparallel to one another in order to stabilise the helix (Brown 1992b), hence one strand runs in the 5' > 3' direction and the other runs 3' > 5'.

There are many potential variations to the PCR cycle, and no rigid rules. For example, primers are not always designed for well known areas of the genome and so 'perfect' matches to the template may not be possible. Alternatively a universal primer for a highly conserved area of the genome can be used with a new previously unrecognised species. In these cases it will be necessary to determine empirically the optimal PCR cycle.

Other than those mentioned thus far (annealing temperature) additional variables to the PCR cycle include, variations in $MgCl_2$ concentration, altering the template concentration, the primer concentration, the time of extension and the number of cycles in the programme. One final variable which is particularly relevant when amplifying unknown regions of the genome, is the adjustment or redesign of the primer sequences.

### *2.2.4 A generalised PCR protocol*

The standard reaction volume for PCRs conducted in this study was 10 μl, except RAPD-PCR reactions (Chp 4) which were conducted using 20 μl. All reactions contained between 5-10 ng template DNA, 10-30 pM primer, 1-2 μl 10x buffer (Perkin Elmer, Warrington, Cheshire or Qiagen), 4 mM dNTP mix (Perkin Elmer) and 1 unit of *Taq* polymerase (Perkin Elmer or Qiagen) or 1 unit of 'HotStar' *Taq* (Qiagen). Some reactions also contained the product 'Q' solution (Qiagen). The solution modifies the melting behaviour of the template DNA, particularly those templates which have a high degree of secondary structure or which are G/C rich (Qiagen). All reactions were brought to volume with sterile $dH_2O$. A 'master mix' of all reagents, excepting template DNA, was prepared for each PCR run and the required amount aliquoted into separate tubes, or wells, if a PCR plate was used. The required amount of DNA was added prior to the PCR run.

PCR buffer, or 10x buffer, usually contains a mixture of Tris-HCl, gelatin, KCl and water. In addition, the buffer can be prepared or purchased to contain $MgCl_2$. The solution helps provide the correct environment for the PCR reaction. Two of the reagents in the buffer act as stabilisers; Tris-HCl stabilises the DNA and gelatin stabilises the enzyme. KCl is present in the reaction mixture to facilitate primer annealing. $MgCl_2$ affects many things in the PCR reaction; primer annealing, strand disassociation, temperature of both template and PCR product, product specificity, formation of primer-dimer artefacts, and enzyme fidelity (Palumbi *et al.* 1991). The concentration of $MgCl_2$ is one of the most influential factors in a PCR reaction; too low and reduced or no enzyme activity will occur leading to a failed PCR; too high and non-specific or smeared products may result (Palumbi *et al.* 1991). For the purposes of the PCR reactions in this study, buffer without $MgCl_2$ was used throughout. This enabled titration of the $MgCl_2$ for each different primer/template combination.

The PCR machines used in this study were a Perkin Elmer (Model '480') and a Hybaid (Model 'Express', with gradient block). Standard cycling parameters were:

95 °C for 5 min } 1 cycle

- 94 °C for 1 min
- determined annealing temp. (dat) for 1 min } 3 cycles
- 72 °C for 1 min

- 94 °C for 30 sec
- dat for 30 sec } 25 cycles
- 72 °C for 45 sec

- 72 °C for 20 min } 1 cycle

An exception to this cycling protocol was used for RAPD-PCR (Chp 4). In addition, where 'HotStar' *Taq* polymerase was used, the 1st cycle was held for 15 min. This *Taq* is provided in an immobilised state and is activated by a 15 min 95 °C incubation step. This step allows complete denaturation of template DNA and helps to prevent non-specific amplification products and so increases specificity (Qiagen).

When performing a PCR reaction, it is essential to run a negative control and helpful to run a positive control. The positive should be DNA template that has previously amplified well with the primers being used, and the negative is all components of the PCR set-up except the template DNA. The use of controls can help eliminate artifacts should problems (e.g. contamination) occur in the PCR reaction.

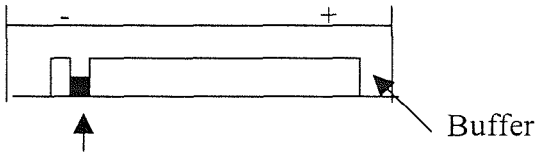
Once the PCR reaction has been completed analysis of the amplified product is usually by electrophoresis on an agarose gel.

## 2.3 Electrophoresis

The Oxford English Dictionary defines electrophoresis as "the movement of charged molecules in an electric field". DNA is similar to other biological compounds, such as proteins, as it carries an electric charge. The charge of DNA is negative, because of the phosphate component of the backbone of the molecule, so when placed in an electric field it migrates towards the positive pole (Figs. 2.3a & 2.3b). There are several electrophoresis methods available to enable the visualisation of both extraction and PCR products. These involve the use of acrylamide and agarose gels, the former giving greater resolution, consequently, acrylamide gels are used for visualising directly labelled PCR products. Agarose gel electrophoresis, however, offers cost effective, rapid and easy visualisation and, if needed, extraction of DNA. Furthermore agarose is non-toxic compared to un-polymerised acrylamide which is a known neurotoxin. Agarose gel electrophoresis, with ethidium bromide staining, is the most commonly used method for checking the size and purity of most PCR products (Delidow *et al.* 1993).

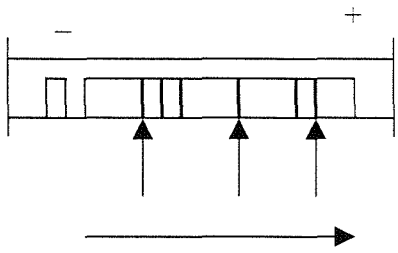
Gels are cast by boiling a buffered agarose solution in a microwave and allowing it to cool in a mould, into which a toothed comb has been set, to produce wells, which divides the gel into a number of lanes. This enables different DNA samples to be analysed simultaneously. Gel electrophoresis is able to separate out DNA molecules based on their size as the dimensions of DNA molecules affects the rate of migration through a gel. To reach the positive electrode, molecules must travel through a network of pores in the gel and it is the relatively small molecules that migrate more rapidly.

**Figure 2.3a. Position of DNA in a gel prior to electrophoresis**




DNA loaded into a well cut out of the gel

**Figure 2.3b. Following electrophoresis DNA separates into bands of different-sized fragments**




Smallest

DNA separating out according to size

Visualisation of the DNA is achieved by removing the gel from the buffer tank and placing it on a source of UV light. The gel is stained with ethidium bromide, either prior to casting, or following electrophoresis. Ethidium bromide binds to the DNA and because it fluoresces under UV light, it enables visualisation.

The size of the DNA fragments run in the gel can be estimated using a size standard, known as a ladder. The most commonly used ladder is the 1 Kilobase (Kb) ladder (Gibco BRL, Life Technologies, Carlsbad, California) which contains 1-12 repeats of a 1018 base pair (bp) DNA fragment. In addition the ladder contains plasmid vector DNA fragments ranging in size from 75–1636 bp. This enables the size of linear double stranded DNA fragments from 500 bp to 12 Kb to be estimated. The ladder is run in a separate lane, simultaneously with the DNA samples to be sized.

Plasmids, along with bacteriophages, are extrachromosomal genetic elements that replicate autonomously within bacterial cells. They are used extensively as cloning vectors e.g. pBR322 (plasmid) and M13 (phage) .

The ladder can also be used to measure approximately the concentration of the DNA fragments present in the gel. This can be achieved roughly by eye, or more accurately using a transilluminator linked to image analysis software.

### *2.3.1 Casting and running a gel*

The majority of gels used for this work were 1% agarose mini gels. A volume of 25 ml is required to cast these gels using Gibco BRL minigel rigs, and for general running, the buffer used was TAE (Tris, Acetic Acid, EDTA). 25 ml of 1x buffer TAE was added to 0.25 g of agarose in a heat resistant conical flask. Perforated clingfilm was used to cover the flask. The gel was then heated at full power in a 750 W microwave for 1 min or until the agarose had dissolved. The flask was then allowed to cool until it could be held by hand, and then 1 $\mu$l of 10 mg/ml ethidium bromide was added and the flask swirled. The gel was then poured steadily into the gel mould with the comb in place. Any bubbles formed as a result of the pouring, were dragged to one side using a pipette tip. The gel was then allowed to set.

Sample DNA and the ladder were combined with gel tracker dye on a piece of parafilm. 1x TAE buffer was poured into the gel bath and the comb and dams were removed from the mould. The samples were then loaded into the wells, with a ladder run at either side of the samples. The gel was run at 50 V for approximately 1 hr, dependent on the fragment size.

When a greater number of samples were required to run simultaneously, or when greater resolution was required (e.g. visualising RAPD-PCR products), a large gel tank was used (volume 250 ml) and Metaphor agarose (FMC Bioproducts, Rockland, Maine) employed (Chp 4).

### *2.3.2 DNA quantification*

When the electrophoresis run was complete, the gel was visualised using the Uvidoc system (Uvidoc 008-XD documentation system with Uvisoft Version 98 gel quantification software). The gel was placed into the transilluminator and the image focused on a screen. The gel was photographed and the image saved onto disc. Using the software supplied the gel image was quantified relative to the ladder.

The software indicated the volume of DNA in each band. The 1636 bp band of the ladder is a known concentration of DNA, this means the volume of this band could be compared to the volumes of the unknown bands. Using the following equation the DNA in the unknown samples was quantified:

$$\left[\frac{\text{vol. of DNA in unknown band}}{\text{vol. of 1636 bp band}}\right] \times 25 = \text{concentration of unknown sample (ng/}\mu\text{l)}$$

Where 1 $\mu$l of ladder and 1 $\mu$l of unknown sample were loaded in the gel.

## 2.4 Sequencing

The ability to determine the exact order of nucleotides in a section of DNA has been one of the greatest achievements of molecular biology. World-wide collaboration using this method has led to the complete sequencing of the human genome (International Human Genome Consortium 2001).

The three methods of sequencing available to molecular biologists are; chemical degradation sequencing, also known as the Maxam-Gilbert method (Maxam & Gilbert 1977), dideoxy or Sanger sequencing (Sanger & Coulson 1975) and automated sequencing which is based on Sanger sequencing.

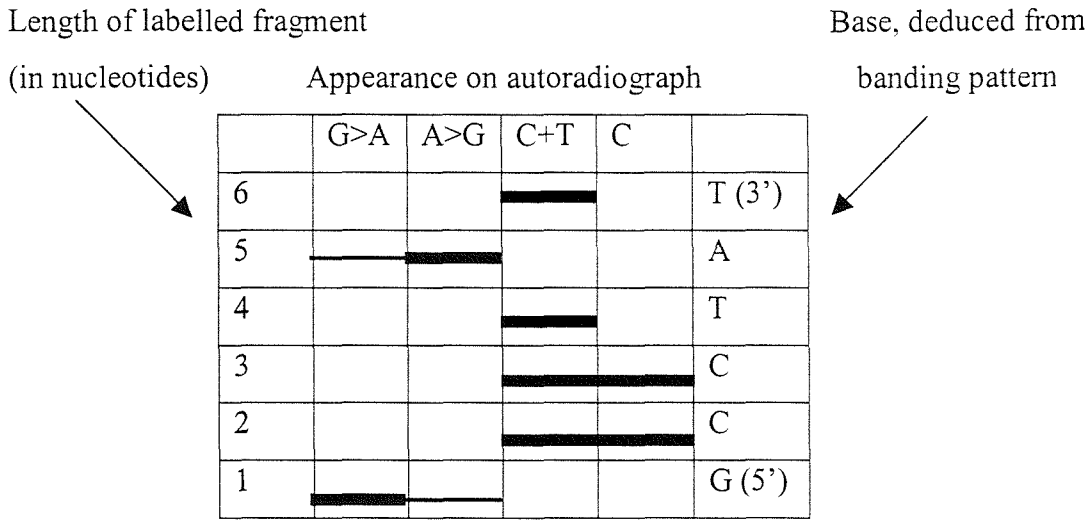
### *2.4.1 Maxam-Gilbert sequencing*

The Maxam-Gilbert method involves, labelling the DNA to be sequenced with $^{32}P$ at either the 3' or 5' end of the sequence and then breaking the DNA sequence at specific nucleotides using specific chemical reactions.

The sample is then divided into four fractions and each treated chemically to modify and remove one of the four bases, breaking the DNA at that point. The first fraction is then chemically treated to cleave at the location of each G in the sequence. In this fraction As are also cleaved, but Gs are cleaved to a greater extent and so is referred to as G>A. The next fraction is treated to cleave As to a greater extent than Gs and so is designated A>G. The third fraction cleaves at C and T equally and the final fraction cleaves at C only.

Polyacrylamide gel electrophoresis is then used to separate the cleaved chains, based on length. This principle is similar to that of agarose electrophoresis, as DNA molecules move through the gel at a rate dictated by their length. Following electrophoresis autoradiography is used to visualise the fragments (Fig. 2.4a).

**Figure 2.4a. A hypothetical example of a Maxam-Gilbert sequencing autoradiograph**




The first lane indicates results for the G>A cleavages, where base 1 is a fragment generated by a cleavage at G, as the band is thicker than that for A. The second lane shows those fragments cleaved at A, thicker bands than for G, so the 5$^{th}$ base is an A. The last two lanes need to be considered together, where a band appears in both lanes a C has been cleaved (bases 2 & 3), and when there is only one band present, a T (bases 4 & 6). The sequence would be read from the band position of the smallest fragment (the one that travelled furthest), so the sequence read for this particular fragment of DNA would be GCCTAG.

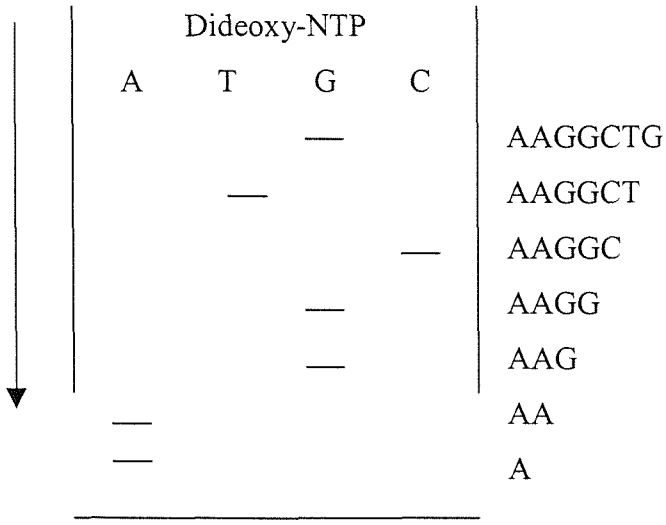
### *2.4.2 Sanger-Coulson sequencing*

This method is also known as the chain-terminating or dideoxy method. It requires the use of single stranded DNA, as chain terminating sequencing involves producing a second strand of DNA based on an existing template. One way by which single stranded DNA may be obtained is to clone it into a M13 vector. M13 phage particles, released by a bacterium are single stranded, so any DNA ligated into a M13 vector will also be single stranded. This is especially useful when sequencing an entirely unknown DNA molecule, such as newly isolated microsatellites (Chp 3).

Once the recombinant molecule has been cloned, the sequencing reaction is initiated by the addition of a primer complementary to a region of the M13 molecule. In addition, dNTPs, *Taq* polymerase and one of 4 dideoxynucleotides are added to the reaction mixture. Dideoxynucleotides can be incorporated into the nascent DNA chain just as efficiently as normal dNTPs, but when added, further synthesis is not possible. This is because the hydroxyl at the 3' carbon on the sugar group is missing and without this group the next nucleotide cannot be attached. Each time a dideoxynucleotide is added, therefore, chain termination occurs. Termination will not always occur at the first position of any given nucleotide, as the mix contains 'normal' as well as chain terminating nucleotides, and these can be incorporated just as readily. Four reactions are carried out in parallel each with a different dideoxynucleotide. Once the reactions are finished, there are four distinct families of DNA, each containing many strands, of differing lengths, each ending in dideoxyGTP, dideoxyCTP, dideoxyTTP or dideoxyATP.

The reactions are then run out on a polyacrylamide gel, separating out the DNA so that the length of each fragment of the target sequence can be determined. The molecules are usually radioactively labelled and the results read using an autoradiograph (Fig. 2.4b):

**Figure 2.4b. An example of the results from Sanger-Coulson sequencing**



The sequence is read by identifying the track each fragment is in, starting with the one which has moved the furthest and then progressing up through the autoradiograph. The final sequence read for this hypothetical example would be AAGGCTG.

### *2.4.3 Automated sequencing*

The dideoxy method for sequencing has been revolutionised by the advent of automated sequencing. Radiolabels in the dideoxy reactions have been replaced with fluorescent labels which can be excited by a laser, detected and data transferred directly to a computer during electrophoresis (Smith *et al.* 1986, Ansorge *et al.* 1987, Prober *et al.* 1987). Individual nucleotides are identified in a single lane (unlike the four reactions required for both Sanger-Coulson and Maxam-Gilbert sequencing) in a temporal manner as opposed to the traditional spatial detection (Prober *et al.* 1987). This is because four different flourescent labels are used that correspond to four different dideoxy terminators.

Direct identification is the greatest advantage of automated sequencing as it removes the time required to dry gels, expose them to film and then analyse the signals on the radiograph (Ansorge *et al.* 1987). A key factor in the popularity and success of this method has, moreover, been the elimination of the use of radioactivity.
Modern day automated sequencing is based on the method first described by Prober *et al.* (1987). This method developed the labelling of chain terminating nucleotides, rather than fluorescently labelled primers (Smith *et al.* 1986 & Ansorge *et al.* 1987). Sequencing methods require a denatured DNA template for the addition of labelled nucleotides, which has previously caused a problem when trying to sequence PCR products. The use of cycle sequencing has, however, overcome this problem. The DNA is denatured and extended several times during a cycle sequencing reaction which guarantees adequate signal for the subsequent sequence detection.

All sequencing reactions undertaken for this study used the automated sequencer ABI Prism 377 with XL Upgrade. All reactions for sequencing were performed with dGTP BigDye terminator reaction mix. This reaction mix is specially modified to use dGTP,

not dITP, in the deoxynucleoside triphosphate mix. dGTP has been shown to produce better sequence data where repeat regions, such as microsatellites, are suspected (PE Applied Biosystems, Warrington, Cheshire 1998).

### *2.4.4 A generalised sequencing reaction*

To obtain optimum results from automated sequencing using dGTP BigDye terminator kits, purification of the preceding PCR reaction products, was necessary. Any traces of unincorporated dNTPs and primers that might have interfered with the cycle sequencing reaction were thus removed.

#### *(i) PCR clean up*

All PCR clean up reactions were performed using Qiagen QIAquick PCR purification kits. The protocol is designed to purify single or double stranded DNA from PCR reactions.

5 volumes of "buffer PB" were added to 1 volume of PCR reaction (e.g.100 $\mu$l of "buffer PB" were added to a 20 $\mu$l PCR reaction). The correct number of QIAquick spin columns corresponding to the number of PCR samples were placed in 2 ml collection tubes and the samples were applied to the columns and centrifuged for 30-60 sec. Flow-through was discarded and the column placed back in the same tube. The DNA now attached to the membrane, was washed by adding 750 $\mu$l of "buffer PE" and the columns were centrifuged for 30-60 sec. The flow-through was again discarded and all samples were spun for a further 1 min to ensure all ethanol residue was removed from the membrane. Columns were then placed in clean 1.5 ml tubes and DNA was eluted by adding 30 $\mu$l of elution buffer to the centre of the membrane, the samples were then incubated for 5 min at 37°C. Finally all samples were centrifuged for 30-60 sec, and the DNA was then ready to use.

Following clean up, the PCR products were prepared for a cycle sequence reaction. The reaction mix contained dNTPs (A, G, C & U), Amplitaq DNA polymerase, labelled

A-Dye terminator, C-Dye terminator, T-Dye terminator and G-Dye terminator, $MgCl_2$ and Tris buffer. Reactions were set up following a standard procedure (Table 2.4a).

**Table 2.4a. Protocol for a sequencing reaction using standard BigDye or dGTP BigDye terminator reaction (PE Applied Biosystems 1998)**

| Step | Action | |
|---|---|---|
| 1 | For each reaction, the following reagents were added to a separate tube | |
| | Reagent | Quantity |
| | Terminator ready reaction mix | 4.0 $\mu$l |
| | Template | * |
| | Primer | 3.2 pmol |
| | $dH_2O$ | Variable ($\mu$l) |
| | Total volume | 10 $\mu$l |
| 2 | Mixed well and spun briefly | |
| 3 | Placed in PCR machine for Cycle sequence reaction | |

*The concentration of DNA used in the sequencing reaction was dependent on the size of the PCR product to be sequenced. Larger PCR products required a greater concentration in the reaction.

Before proceeding to the next stage of the sequencing reaction the extension products were purified, to remove traces of unincorporated dye terminators, which could have interfered with the signal picked up by the automated sequencer.

### *(ii) BigDye Clean-up*

All sequencing reactions were cleaned with Qiagen DyeEx Spin kits. These are designed for easy removal of unincorporated dye terminators. Before beginning the clean up the appropriate speed for the centrifuge (750 xg) was calculated using the following equation:

rpm = 1 000 x $\sqrt{750/1.12}$ r (r = radius of rotor in mm)

The correct number of spin columns corresponding to the number of samples were gently vortexed to resuspend the resin. The caps on all columns were loosened a quarter of one turn (this avoids a vacuum in the column). The bottom closure of all spin

columns were snapped off and each was placed in a 2 ml collection tube. The columns were then centrifuged at the calculated speed for 3 min.

The spin columns were carefully transferred to clean 1.5 μl tubes and the sequencing reaction was applied. Great care was taken to pipette the reaction on to the centre of the gel base, drop by drop so that the reaction mix was absorbed into the bed. The samples were then centrifuged for 3 min. The columns were then discarded as the eluate in the 1.5 μl tubes contained the purified sequencing product. The samples were then dried in a hot block at 70°C.

Once the products were purified they were ready to be sequenced on a polyacrylamide gel. The gel was electrophoresed in the ABI 377, an image of the gel was created and then stored so that it could be analysed at a later date.

### *2.4.5 Casting and running a polyacrylamide gel*

Before the gel was cast, the glass plates were cleaned thoroughly to ensure removal of all debris that may have interfered with pouring the gel. The plates were set into the holder, aligned and then clamped into place. The top and bottom pieces were clamped into place and a comb prepared for setting the gel interface. 40 ml of Genepage Plus (acrylamide) was poured into a small beaker. 28 μl of T-med (N,N,N',N' - Tetramethyethylenediamine) and 200 μl of 10% ammonium persulphate (A.P.S.) were added slowly to the solution which was swirled gently to mix.

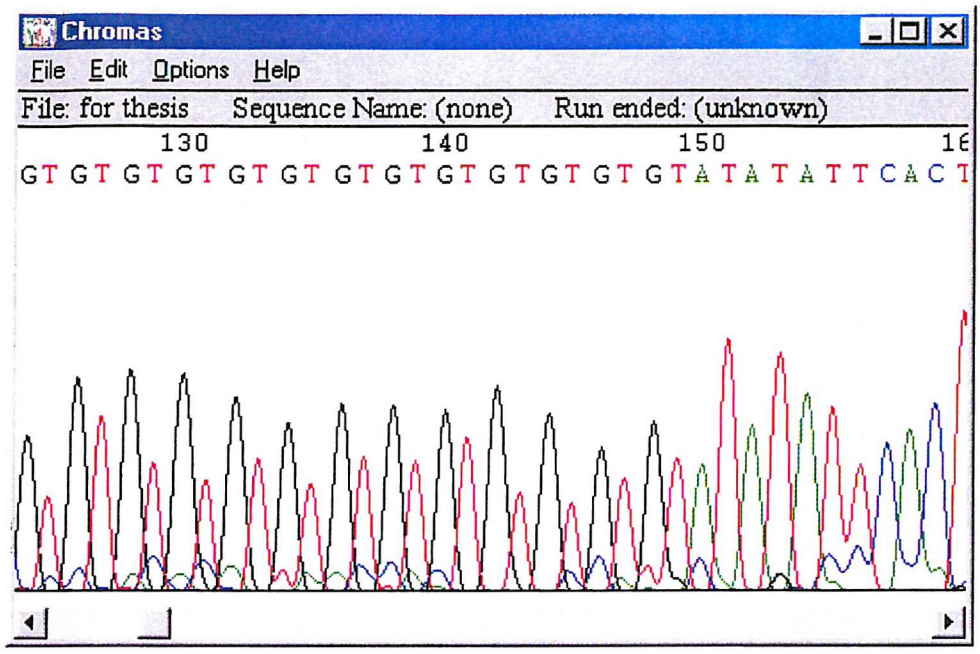
The mixture was drawn immediately into a 50 ml syringe which was introduced into the bottom piece of the casting apparatus and the gel was poured by steadily depressing the syringe allowing the mixture to flow between the plates. To encourage movement of the polyacrylamide through the clamped plates, the glass was gently tapped ahead of the flowing gel. The top piece was removed, a comb set in place, and clamped back into position. This comb created the gel interface, the point at which the samples are introduced.

Samples from the clean-up were resuspended using 4 $\mu$l of loading dye (containing formamide to prevent renaturation of DNA, and blue dextran (5:1)). Once resuspended, they were vortexed and then spun. They were next denatured at 95°C for 2 min prior to loading. Samples were pipetted onto a loading tray and a paper comb was used to 'pick up' the samples. This comb was introduced to the gel and electrophoresis was carried out for 2 minutes to allow the samples to be drawn into the gel. The paper comb was then discarded and the run allowed to continue for at least 13 hours, usually overnight.

### *2.4.6 Tracking and extracting to create an electropherogram image*

Data collected were saved as an image, which was then automatically 'tracked' to ensure the correct samples were associated with the correct lanes. Data were then automatically converted to an electropherogram (Fig. 2.4c). Once sequences could be visualised they were checked for quality. This includes checking that all bases have been called clearly and can be easily read and 'trimming' the beginning and ends of sequences that have been poorly called. In addition sequences are checked for accuracy, for example in Fig. 2.4c a repeat motif corresponding to a GT repeat was required. Sequence information can then be used for a variety of purposes, including primer design or as direct information for genetic analysis. Primer design can be carried out by importing the sequence data into specifically designed software (Chp 3) and for genetic analysis the sequences are imported into the relevant statistical and/or tree drawing software packages.

**Figure 2.4c. An example of an electropherogram showing a sequence read containing a repeat motif.**



## 2.5 Gene scanning & Genotyping

### *2.5.1 Introduction*

Gene scanning allows the detection of DNA fragments labelled with differently coloured fluorescent dyes. PCR products (incorporated with 5' end labelled primers) are run through a polyacrylamide gel and detected using an automated sequencer. The size of each DNA fragment is measured against an internal lane size standard, which is able to precisely call the size of each fragment. Data generated from gene scanning is then exported downstream to software applications for analysis.

### *2.5.2 Preparing samples for gene scan*

DNA samples from individuals were subject to PCR using 5' end labelled fluorescent primers (for primers and labelling used see individual species Chapters 5 & 6). Each PCR reaction was completed with a 30 min hold at 72°C, this procedure forces "A" tailing by *Taq* polymerase and improves allele calling. To improve efficiency, PCR products were multiplexed prior to running on the polyacrylamide gel. No products with the same fluorescent dye were run at the same time. 5$\mu$l of each PCR sample to be multiplexed, were mixed together, and where necessary dilutions were performed using $dH_2O$. Dilutions were calculated empirically for each fluorescent dye, as each had different signal strengths (HEX the lowest and FAM the strongest). Owing to the fluorescent nature of the primers, PCR products were kept in darkness at -20°C to help prevent degeneration prior to gene scanning.

### *2.5.3 Casting and running a polyacrylamide gel for Gene Scan*

A gel for running gene scan products was cast in the same way as a gel for running DNA sequences. Once the gel was set the PCR products were loaded using the following loading mixture for 64 samples:

8μl loading dye
16μl ROX (internal lane size standard)
80μl formamide

1.5μl of loading dye was introduced into a each well in a 64 sample plate. 1μl of each of the mixed PCR products was then added to separate wells. This meant that each well contained 1.5μl of loading dye and 1μl of PCR products that contained a sample that was HEX labelled, a sample that was NED labelled and a sample that was FAM labelled. The sample plate was then put in a hot block at 94°C for 5 minutes to denature the DNA prior to the gene scan run. Once denatured 1μl of each sample was then transferred onto a loading tray and these were "picked" up by a paper comb to be loaded into the gel. This comb was introduced to the gel and electrophoresis was carried out for 2 minutes to allow the samples to be drawn into the gel. The paper comb was then discarded and the run allowed to continue for at least 2.5 hours.

### *2.5.4 Tracking and extracting data*

Once the gene scan run was completed, GeneScan® (Applied Biosystems) analysis software was used to analyse the data. The lanes of the gel were tracked and if any lanes appeared misaligned they were manually tracked. Once this procedure was completed, the data were extracted for further analysis. Samples were analysed using predefined analysis parameters, e.g. analysis range was set; peak amplitude threshold values were set. Additionally, peak assignments for the size standard were defined and selected. Once parameters had been set, the gel was analysed and data were then loaded into Genotyper® (Applied Biosystems) software for further analysis.

### *2.5.5 Genotyping data*

The allele definitions for microsatellite markers are based on the fragment length (size) of the PCR product as estimated by gel electrophoresis. Allele binning using Genotyper® software is a process that allows grouping of allele fragments belonging to a particular size, into a range (bin) centred around the average size with a tolerance

limit. For example an allele definition may look like this; 101.5 ±0.5. This is a necessary procedure as allele sizes tend to vary between electrophoresis runs because of subtle differences in gels or in electrophoresis conditions. For this study alleles were binned using the Histogram window available with the Genotyper® software. Binning was only performed once a data set was complete. Once allele sizes were determined for each microsatellite marker for all samples, data were then exported downstream for further analysis software

## 2.6 Data analysis

Please refer to each individual species chapter under data analysis (Chps 4, 5 & 6).

# 3.0 Microsatellite Development

## 3.1 What are microsatellites?



## 3.2 Microsatellites and conservation



## 3.3 Microsatellite Development

# 3.0 Microsatellite Development

## 3.1 What are microsatellites?

### *3.1.1 Microsatellites*

Advances in molecular techniques have provided biologists with unrivalled access to the genome and this, combined with increasingly powerful statistical analysis, have revolutionised our understanding of genetics. A variety of genetic markers have been employed for this purpose including, allozyme loci, mitochondrial DNA and nuclear DNA (Ward *et. al* 1992; Avise 1994). The latter include the two groups of Variable Number Tandem Repeats (VNTR), minisatellites and microsatellites. VNTR markers comprise a number of tandemly repeated sequences, with the two categories separated based on the length of this repeat sequence (Jarne & Lagoda 1996).

Microsatellites consist of repeated nucleotides, each between 1 and 10 base pairs (bp) in length e.g. $(CA)_n$ (Tautz 1989). Minisatellites have a longer repeat unit, from 15 bp up to 200 bp. Minisatellites were first discovered in 1985 (Jeffreys *et al.* 1985) and their discovery led to the first studies in DNA fingerprinting and so to identification of individuals, parentage and genome mapping (Bruford *et al.* 1996). Minisatellites have been shown to be extremely variable and with a high mutation rate (Bruford *et al.* 1992). Unfortunately, the power of the polymerase chain reaction (PCR) cannot be fully exploited with minisatellites owing to the large size of some of the alleles. The efficiency of the PCR reactions falls as DNA sequences go over 10kb (Bruford *et al.* 1996).

Microsatellites were also identified in the genome over 20 years ago, but were disregarded as sequences of any interest (Jarne & Lagoda 1996). More recently, however, they have become the preferred marker and have found uses in many areas; mapping programmes, investigation into genetic diseases, kinship studies, and population work have all exploited their power (Luikart & England 1999). Microsatellites are now facilitating not only population work, but ecological and behavioural studies, where multi-locus genotyping has been used for pedigree reconstruction, kin based behaviour studies and to reveal the existence of cryptic species

(Jarne & Lagoda 1996). There are three types of microsatellite, pure, compound and interrupted and any combination between these is possible (Table 3.1a). Microsatellites can be highly polymorphic owing to a high mutation rate. This is particularly true of long, (as polymorphism has also been shown to depend on repeat size), uninterrupted microsatellites (Hancock 1998). There is evidence to suggest an interruption within the core sequence stabilises an array of repeats, making interrupted microsatellites the least variable (Jarne & Lagoda 1996). In natural populations, the average expected heterozygosity for microsatellites is well above 50% and can peak at nearly 100%. Levels of polymorphism can, however, decline for interrupted and compound repeats (Jarne & Lagoda 1996).

Microsatellites have been located throughout the eukaryotic genome of all organisms investigated to date (Bruford & Wayne 1993; Hancock 1998). As more organisms are investigated it appears that the prevalence of microsatellite sequence varies between organisms (Beaumont & Bruford 1999). Microsatellites have also shown themselves to be a genetic marker that can be assayed readily by the PCR process when combined with gel electrophoresis (Tautz 1989). A characteristic not generally shared by minisatellites is the cross amplification potential of some microsatellite primers (Bruford *et al.* 1996).

**Table 3.1a. The three families of microsatellite found in a wide variety of eukaryotes and the chloroplastic genome of plants.**

| Microsatellite Name | Ten-repeat dinucleotide array example |
|---|---|
| Pure | CACACACACACACACACACA |
| Compound | CACACACACA**GAGAGAGAGA** |
| Interrupted | CACA**TT**CACACA**TT**CA**TT** CA |

Microsatellite alleles have short lengths, show high levels of variability and they are inherited codominantly making them the ideal marker for many studies (Luikart & England 1999). For example, where multiple alleles are present, the spatial distribution of the alleles can be used to study local gene flow (Queller *et al.* 1993). Codominant,

Mendelian inheritance has been indicated in many studies of pedigree analyses (Jarne & Lagoda 1996).

The most common microsatellite found in all organisms is a poly A/T repeat, but its use is limited in population studies because of its instability during PCR (Hancock 1998). Dinucleotide, trinucleotide and tetranucleotide repeats are the three types of microsatellite commonly used. The most frequently used of these are the dinucleotide repeats, which show a density that peaks around one locus per 5 kb (Jarne & Lagoda 1996). Trinucleotide and tetranucleotide repeats are not used in population studies as much as the dinucleotide repeats, despite the highly polymorphic nature of both.

*Functional significance and neutrality of microsatellites*

It is not known if all microsatellites have any functional significance in the genome (Hancock 1998). Although generally present in non-coding regions of the genome, they have been found in exons (Hancock 1995). In addition, there has been suggestion of a role in gene regulation or in recombination (Stallings *et al.* 1991). More recently microsatellites have been identified ubiquitously in upstream promoter regions of coding sequences, giving more evidence for them having a functional role in the coding process (Kashi & Soller 1998). Furthermore, microsatellites have been shown to bind proteins, a feature typical for upstream activating sequences (Kashi & Soller 1998). Their association, however, with the human genome has mainly been found to be negative, for example they appear to be responsible for a variety of neurological diseases (Moxon & Wills 1999).

What causes the fixation of mutational changes of DNA, such as those seen in microsatellites? In evolutionary biology there are two views; that DNA changes result from evolution by natural selection, and the opposing view, that the majority of mutations have no functional relevance and are spread by chance. This latter view is known as the neutral theory of molecular evolution and was first proposed by Kimura (1968). Kimura suggested the majority of mutations that become fixed have no effect on fitness (meaning they confer no selective cost). These neutral mutations are not subject to natural selection and are usually lost or sometimes fixed by the process of

random genetic drift. Under this hypothesis, natural selection is only involved in removing deleterious mutations and in playing a minor role in the fixation of mutations. Neutralists believe that mutations are either deleterious and selectively removed or neutral, which gives them the chance of becoming fixed. Selectionists believe that substitutions become fixed because they present an advantage, and neutral mutations are rare. Both theories agree that deleterious mutations are in the majority and likely to be removed.

In population studies it is important that the chosen molecular marker is neutral, that is it confers no advantage. Once a genetic marker is under strong selection its use becomes restricted for population genetic studies (Queller *et al.*1993). Neutral markers are excellent for use in population studies for understanding genetic structure, patterns of historical gene flow and demography (Jarne & Lagoda 1996). Other than the few examples given here it is believed that microsatellites are predominantly neutral, making them an ideal choice for population studies (Hancock 1998).

### *3.1.2 Mechanisms of mutation in microsatellites*

In order that statistics may be developed to fully utilise the data generated by microsatellites, understanding the mutation model underlying the evolution of this marker is essential. The evolutionary dynamics of VNTRs have been examined theoretically and experimentally (see Jarne & Lagoda 1996; Armour *et al.* 1999; Estoup & Cornuet 1999 for reviews). It appears that there may be differences in the mutational processes of the two groups, with minisatellites showing a more complex pattern (Estoup & Angers 1998). Although the mechanisms underlying allele frequency changes in populations are not fully understood, the molecular mechanisms involved in repeat number variation are thought to be a result of either intra-allelic or inter-allelic events (Estoup & Angers 1998). Mutation is, therefore, linked with the DNA repair process (intra-allelic) and with recombination (inter-allelic) (Jarne & Lagoda 1996). *In vitro* experiments have provided support for intra-allelic events where mutation occurs by polymerase slippage at DNA replication, resulting in an increase or decrease of the number of repeats by one unit in one generation (Schlötterer & Tautz 1992, Hancock

1998). Other studies (e.g. Strand *et al.* 1993; Weber & Wong 1993) have indicated that larger mutational steps are occurring and that these are generated by processes such as unequal crossing-over between sister chromatids (Estoup *et al.* 1995). It is thought that the majority of mutational events in microsatellites are a result of polymerase slippage. Minisatellites, however, appear to demonstrate an equal occurrence of both intra and inter-allelic mutational events (Estoup & Angers 1998).

***(i) Rates of mutation***

The mutation rate for VNTRs is high enough and back mutations sufficiently common to seriously bias estimates of gene flow if mutations are ignored (Waples 1998). As with many aspects of VNTRs, there is much discussion surrounding the actual rate of mutation for this molecular marker. It is, however, agreed, that the rate is relatively high compared with point mutations ($10^{-9}$ - $10^{-10}$) (Estoup & Angers 1998). Weber & Wong (1993), using human pedigree analysis, put rates for VNTR mutation at around $10^{-3}$ events per locus per generation. Studies of other organisms put rates anywhere between $10^{-4}$ to 6 x $10^{-6}$ (Dallas 1992; Schug *et al.* 1997).

A factor which appears to influence the rate of mutation is the size of the repeat unit (Chakraborty *et al.* 1997). Microsatellites that contain a higher repeat count demonstrate a higher mutation rate. It is thought that this is driven by more opportunity for a stable but misaligned configuration during replication (Estoup & Angers 1998). This has also been shown to be the case when comparing interrupted and non-interrupted microsatellites, with the former being more stable as misalignment chances are reduced (Pèpin *et al.* 1995). In addition, the composition of the repeat unit, particularly those containing A/T nucleotides, appears to influence the polymorphism level of tri- and tetranucleotide microsatellites (Sheffield *et al.* 1995). These factors help to give some understanding to the polymorphic differences that are often observed between loci and/or within a locus. They may also help to explain intraspecific differences among populations (Estoup & Angers 1998).

Some evidence for mutational biases has been found (Estoup & Angers 1998). Gains of repeat units are more frequent than losses and, for microsatellites at least, higher mutation rates are seen during paternal meiosis (Estoup & Angers 1998). Recent research, however, suggests that the magnitude of size alteration when it does occur is greater in females than in males (Primmer *et al.* 1998). Another suggested source of increased mutation has been those heterozygous individuals where the alleles exhibit greatly different lengths (Amos 1999). If this were the case, then heterozygotes would show a greater mutation rate when compared to homozygotes, with the implication that loci in larger populations would evolve at a faster rate than those in smaller populations (Estoup & Angers 1998). This bias remains conjectural and requires further investigation.

***(ii) Size constraints***

The microsatellite loci, examined in natural populations, have a finite size, usually shorter than a few tens of repeat units, implying some kind of size limitation/constraint for a repeat unit (Estoup & Angers 1998). This is in spite of a mutational process which seems to favour gains over losses. No direct evidence, other than for some human diseases, exists for selective constraints on the size of the repeat units. Constraints, if found to be dependent on the size of arrays, will act more strongly on longer repeat units. Several workers have, however, suggested mechanisms which may act to selectively prevent microsatellites from increasing into virtually infinite numbers of repeats (Garza *et al.* 1995; Samedi *et al.* 1998).

***(iii) Size homoplasy***

Size homoplasy occurs where two VNTR PCR products show the same length, but may not share a common ancestral sequence. This means they are identical-in-state (IIS) but not identical by descent (IBD) (Estoup & Angers 1998). Size homoplasy is clearly related to the way that mutations produce new alleles which in turn, links it to the mutation model considered for those loci. For example some models, e.g. the infinite allele model, consider that all new alleles produced in the population are novel, thus

excluding the possibility of size homoplasy (see below). There will be an increased chance of homoplasy, where selection is acting to limit allelic repeats length size, as the number of allelic states allowed is reduced (Nauta & Weissing 1996). Homoplasy is also expected to increase as the time of divergence and the mutation rate increase (Estoup & Angers 1998).

Size homoplasy has been revealed experimentally, and is more commonly detected among populations (Estoup *et al.* 1995; Garza & Freimer 1996) It is thought to affect the resolution of population structure because it can increase the number of significant single locus pairwise tests and can increase the non-stepwise estimators of genetic differentiation (Viard *et al.* 1998). The use, however, of certain models accounts for the problems associated with size homoplasy, reiterating the need to choose loci which demonstrate mutation patterns that are conducive to assessment under such models (see below) (Slatkin 1995; Rousset 1996).

### *3.1.3 Mutation models*

The major limiting step in constructing realistic mutation models for VNTRs is the large variance in the mutation parameters amongst loci (Estoup & Angers 1998). A useful way to overcome this difficulty is to employ loci which demonstrate similar mutation parameters, i.e. follow the same mutation model. This is essential in the accurate estimation of population parameters. Two models of mutation have been provided by population geneticists; the stepwise mutation model (SMM; Kimura & Otha 1978) and the infinite allele model (IAM; Kimura & Crow 1964). These two models are extremes of one another and currently the SMM is believed to reflect most accurately the mechanism of mutation in microsatellites (Balloux & Lugon-Moulin 2002). There are two other models, *K*-allele model (KAM) and the two-phase model (TPM; Valdès *et al.* 1993; Di Rienzo *et al.* 1994). The TPM was developed from the SMM to account for a proportion of larger mutation events , and the KAM is essentially the same as the IAM but unlike the IAM the number of alleles is not infinite (Balloux & Lugon-Moulin 2002). Understanding these models is essential, because the sensitivity to the mutation

model increases with the rate of mutation with the result that the estimation of various population parameters is highly dependent on the model used (Estoup & Cornuet 1999)

***(i) Stepwise mutation model***

In this model, mutational events are assumed to occur as the loss or gain of single repeat units with an equal probability $\mu/2$ in both directions, so alleles can mutate to an allele state already present in the population. This model interprets alleles of very different sizes as being more distantly related than those more equally sized, that is the SMM has a memory of allele size. Studies into microsatellite mutation in humans, however, have revealed that although mutations resulting in gain occur in a single step, mutations resulting in loss can occur in two steps (Bruford *et al.* 1996).

***(ii) Infinite allele model***

The infinite allele model is based on each mutation, involving any number of tandem repeats, creating a novel allele at a given rate, $\mu$. Consequently, the IAM does not allow for homoplasy, when two allelic lineages converge on the same size but have different histories of mutations increasing and decreasing in size, thus size identity is a result of mutation not common ancestry. In contrast to other models, therefore, alleles under IAM are identical by descent.

***(iii) Other mutation models***

The $K$-allele model has exactly $K$ possible alleles. The probability of any mutation event from one allelic state to any other ($K$ - 1) is identical, so a given allele will mutate to any other at a rate $\mu/(K - 1)$. An important component of this model is the allowance for homoplasy so alleles are considered identical-in-state but not identical by descent. Some authors believe that the KAM is a more realistic representation of microsatellite mutation than the IAM, because of the size restraints that appear to act on microsatellites (Estoup & Cornuet 1999). The two-phase model (TPM) considers the loss or gain of $X$ repeats, allowing mutation steps of several repeats. An increase or

decrease in allele size of one unit ($X = 1$) has the probability $p$, (corresponding to the SMM). Gain or loss in allele size of $k$ units has the probability $(1 - p)$, here $X$ follows a geometric distribution (Di Rienzo *et. al.* 1994).

### *3.1.4 Mutation models and natural populations*

Alleles differing by single mutational events and single and multi locus pedigree analysis have shown that the SMM and the TPM best describe the distribution of mutation sizes (Weber & Wong 1993; Primmer *et al.* 1996). Length variation, however, appears to best fit models proposed by the IAM and the KAM (Estoup & Angers 1998). To evaluate how useful the SMM and the TPM are at assessing allelic distributions in natural populations workers have used two approaches (see Shriver *et al.* 1993; Di Rienzo *et al.* 1994; Estoup *et al.* 1995). Results from these tests when applied to real data, have been inconclusive, implying that these models account for only a small amount of the variation observed in VNTR mutation, including the possibility that different mutational processes are affecting different loci (Estoup & Angers 1998).

Whilst the models are useful and can provide interesting and valuable information, recent work on VNTR evolution has revealed factors which are not accounted for by these models (Estoup & Cornet 1999). Sex biased mutation shown by some microsatellites (Weber & Wong 1993) and the constraints which appear to act on allele size (Garza *et al.* 1995) are two such factors.

#### *(i) The estimation of population differentiation*

The models proposed above regarding the mutation of microsatellites have generated varying statistical models for the estimation of genetic differentiation based on data created from microsatellites. Genetic distance is usually investigated by means of Wright's $F_{ST}$ (1951) and or, Slatkin's $R_{ST}$ (1995), which is an analogue of $F_{ST}$ that assumes a SMM. Currently this model is favoured as the most accurate reflection of microsatellite mutation (Balloux & Lugon-Moulin 2002).

There are several definitions of $F_{ST}$. Wright (1921) defined the original fixation index as a way to account for the effect of inbreeding within samples. This was then expanded to account for a population divided into sub-populations, which has led to the traditional hierarchical *F*- statistics, $F_{ST}$, $F_{IS}$, $F_{IT}$ (Wright 1951). Where $I$ = individuals, $S$ = sub-populations and $T$ = total population.

$F_{ST}$, (based on the IAM), fixation index for multiple alleles is calculated as follows:

$F_{ST} = (Ht - Hs) / Ht$ (Nei 1977)

Where *Ht* is the expected heterozygosity of an individual in the total population (assuming all samples are drawn from a single homogenous randomly mating population), and *Hs* is the expected heterozygosity of an individual in a sub-population (calculated separately for each sub-population and then averaged). $F_{ST}$ is therefore a measure of the inbreeding resulting from the correlation among alleles, because they are found in the same sub-population. For two sub-populations, with a two-alleles locus, $F_{ST} = 1$, when the two sub-populations are totally homozygous and fixed for the alternative allele. When the frequencies in the two sub-populations are identical, $F_{ST} = 0$ (Balloux & Lugon-Moulin 2002).

From the above, $F_{ST}$ can be understood to be a measure of heterozygote deficiency owing to population subdivision, this is known as the Wahlund principle (Wahlund 1928). $F_{ST}$ is, therefore, a measure of the heterozygote deficit relative to its expectations under Hardy-Weinberg equilibrium (HWE) (Hartl 2000).

A population where the allele frequencies remain the same in the next generation as they were in the generation before (which means the allele frequencies remain the same, in the proportions $p^2$, $2pq$ and $q^2$ in a two allele locus) is said to be a population that is in HWE (Hartl 2000). This is achieved by Mendelian inheritance keeping the allele frequencies constant, thus preserving genetic variation from generation to generation. There are, however, many assumptions to this principle; that the organism in question is diploid; reproduction is sexual; generations do not overlap; population size is large;

allele frequencies are equal between the sexes; mating is random and there is no mutation, migration or selection (Hartl 2000).

Slatkin (1995) developed a statistical model for microsatellites based on the SMM, $R_{ST}$. $R_{ST}$ is calculated as follows:

$$R_{ST} = (S - Sw) / S$$

Where $S$ is the average squared difference in allele size between all pairs of alleles, and $Sw$ is the average sum of squares of the differences in allele size within each sub-population (Balloux & Lugon-Moulin 2002). These, and hence $R_{ST}$, are derived from the variances in allele sizes, compared to $F_{ST}$ which is calculated from variances in allele frequencies.

$F_{ST}$ (and hence $R_{ST}$) provide the basis for measurement of genetic distance where divergence is caused by drift (Reynolds *et al.* 1983) and are to date the most commonly reported estimators for population structure (Balloux & Lugon-Moulin 2002). In an ideal population where mutation follows an IAM, $F_{ST}$ is a decreasing function of the product of local population size and the sum of migration and mutation ($N(m + \mu)$). However, when mutation is negligible (often not the case for microsatellites) $F_{ST}$ becomes a simple function of the number of migrants (Balloux & Lugon-Moulin 2002).

A major drawback concerning the use of $F_{ST}$ occurs when the mutation model cannot be assumed to be an IAM. This is often the case for microsatellites as the IAM assumes only novel alleles are present in the population, whereas microsatellites are subject to back mutations generating alleles already present in the population. This may lead to an underestimation of true genetic differentiation. Where an IAM cannot safely be assumed, a SMM is employed but using this model the relationship between $F_{ST}$, migrants and mutants is no longer upheld (Rousset 1996), hence the alternative estimation of differentiation, $R_{ST}$, which is independent of the mutation rate under a SMM. There are, however, problems associated with the use of $R_{ST}$. These relate to its high variance (Balloux & Lugon-Moulin 2002). Where sample sizes differ between

populations, smaller populations will contribute less to the total variance and where variance differs greatly between loci, those with low variance will contribute less to the value of $R_{ST}$. Consequently even where microsatellite alleles in a population follow a strict SMM, $F_{ST}$ will be a superior choice to $R_{ST}$ (Gaggiotti *et al.* 1999).

Microsatellites mutate at a high rate and we need an understanding about the manner in which this occurs. To gain a better understanding of migration and divergence, and more accurate ways to measure them statistically, microsatellite mutation must be understood. A precise understanding of the probability of one given allelic state passing to another would allow definition of differentiation statistics on allelic distance (Balloux & Lugon-Moulin 2002).

***(ii) Null alleles***

One of the key problems when assessing microsatellite variability in populations is the presence of "null" alleles. A null allele is a mutation in a region complimentary to a primer, thus PCR fails to amplify one or both alleles in an individual. If null alleles are common in a population then heterozygous individuals may be misinterpreted as homozygotes. Null alleles have been shown to occur in microsatellites, with up to 7 loci out of 27 showing null alleles (Callen *et al.* 1993). The Hardy-Weinberg equilibrium model can be used to detect null alleles where the observed heterozygosity deviates from the expected and where no other cause for the heterozygote deficiency is suspected (Jarne & Lagoda 1996).

## 3.2 Microsatellites and conservation

### *3.2.1 Introduction*

Many studies describing patterns of genetic variation within and between endangered or threatened species, as well as patterns of their evolutionary histories can be found in the current literature (Avise & Hamrick 1996). This work is considered crucial to conservation efforts. Central to the idea of conservation genetics is the retention of genetic diversity within and among populations which provides a buffer essential for the survival of a species (Frankham 1999). Recent research has revealed much about the genetic nature of populations from bottlenecks and random drift to genetic distances (MacHugh *et al.* 1994; Paetkau & Strobeck 1994).

### *3.2.2 Microsatellites as molecular markers for use in conservation*

The use of microsatellites to describe the neutral molecular variation in threatened populations has been increasing in the last decade (Luikart & England 1999). They are especially useful in studies of endangered species, because of their properties of (i) high variability (ii) ease of scoring and (iii) use of only small amounts of DNA (Luikart & England 1999). Another advantage is that costs may be kept to a minimum as microsatellite primers developed for one species may be employed in the study of related taxa (Moore *et al.* 1991).

As well as their use in population work concerning threatened species there has been much interest in the application of microsatellites to hybridisation studies (Bruford *et al.* 1996). Recent research has revealed data such as extensive hybridisation between North American grey wolves and coyotes, and between Ethiopian wolves and domestic dogs (Gottelli *et al.* 1994; Roy *et al.* 1994). Roy *et al.* (1994) also revealed the complete loss of unique alleles in endangered red wolf populations who now appear to share all their alleles with coyotes. Yet another example of the use of microsatellites for conservation research has been demonstrated by Williams *et al.* (1996) who

investigated the genetic effect of the introduction of a founder population of Muntjac deer followed by rapid growth of the population.

It is clear that microsatellites are proving to be an incredibly useful tool for conservation genetics. It is, however, important to highlight the disadvantages associated with this marker. A difficulty facing many workers is locating microsatellites in the genome of the target species. To date there has been some difficulty in locating microsatellites in the DNA of many plant species, several invertebrate groups and some avian groups (Beaumont & Bruford 1999). Other difficulties relate more specifically to the technical aspects of working with this molecular marker. There are also *Taq* generated PCR problems such as slippage when amplifying mono/di nucleotide loci, and the tendency of *Taq* to add extra dATPs causing base shifts (Schlötterer & Tautz 1992; Ginot *et al.* 1996). There has also been some difficulty reported in comparing data between laboratories, in spite of automated procedures and these seem to be exacerbated by changes in the chemicals or hardware used (Beaumont & Bruford 1999).

A particular difficulty relating to conservation genetics and the use of non-invasive samples, such as hair and faeces, are random allelic 'drop out', non-specific amplification, and amplification of incorrect genotypes because of slippage or contamination (Beaumont & Bruford 1999). These problems are thought to occur because of the presence of only one copy of each locus per cell (in comparison to mitochondrial DNA which is present in several copies per cell) and are augmented by altered stringency conditions (Beaumont & Bruford 1999). These problems are not, however, insurmountable (Taberlet *et al.* 1996).

Perhaps the greatest concern regarding the use of microsatellites is the lack of any well established universal evolutionary model. Where conservation questions are being tackled, however, the lack of a universal model is not so much of a problem because genetic divergence is probably a result of bottlenecks, random drift or inbreeding and not mutation (Beaumont & Bruford 1999). A review of current opinion regarding the question of which model applies best, when estimating population differentiation using

microsatellite markers, has recently been addressed in the literature (Balloux & Lugon-Moulin 2002).

In spite of these drawbacks, microsatellites will continue to feature heavily in the study of endangered and threatened species, particularly because of the high resolution and range of information that may be gained from their use. With the use of non-invasive sampling, and the likelihood that more and more degenerate primers will become available, overall cost will reduce and increase their value as a molecular marker. For many marine species, however, genetic data alone cannot resolve key conservation management questions, largely owing to the errors associated with estimating population parameters (Waples 1998). To help eliminate, or at least substantially reduce these errors, it is essential that a thorough knowledge and understanding of the biology and life history of the target species is obtained.

## 3.3 Microsatellite Development

### *3.3.1 Introduction*

The most common method for finding microsatellites involves making and screening a library of short sequences. High molecular weight DNA is extracted from the target species and cut with a restriction enzyme. The DNA is then electrophoresed so that fragments 200-500 bp can be isolated, these fragments (some containing microsatellites) are then ligated into plasmids. *Escherichia coli* is then transformed with these plasmids and plated out to form colonies. A replica plate is then probed with a synthesised labelled repeat sequence, and exposed to film to show those colonies containing the repeat of interest. Positive colonies are located on the live filter and grown up in liquid medium. Finally these are sequenced to verify the presence of the microsatellite and so that the priming regions can be identified.

### *3.1.2 Preparation for enrichment*

#### *(i) Extraction*

To begin microsatellite development the whole genomic DNA from several individuals was extracted using a standard phenol-chloroform method (Chp 2). Samples were then pooled together to increase the total amount of DNA to be screened for microsatellites. This was achieved by conducting a precipitation reaction such as was performed during the final step of the extraction procedure. At the final stage the DNA was diluted in "buffer AE" (Qiagen) to a concentration of $1 \mu g\ \mu l^{-1}$

#### *(ii) Digestion*

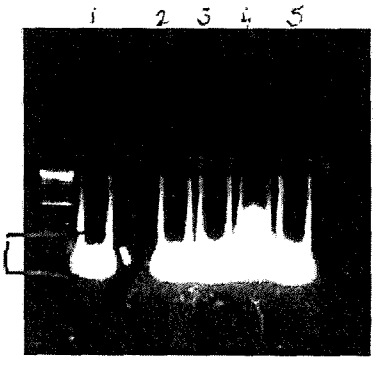
The digestion stage was necessary to produce the right size fragments of DNA (200-500 bp) that correspond to the range of acceptable fragments that could be later inserted into plasmids for cloning. A restriction enzyme was used for this process. This is a type of enzyme known as a endodeoxyribonuclease, which can recognise certain sequences

within DNA molecules and catalyse double strand cleavage of DNA and leave 'sticky ends' (Kandpal *et al.* 1994).

The protocol followed for microsatellite development for these studies used the restriction enzyme *Mbo*1 (Promega – Bio-Tech Ltd.), this is a restriction enzyme that recognises GATC sequences within DNA molecules and cuts the DNA at every such sequence leaving a 'sticky end'. Samples were incubated in a water bath at 37°C for 3 hrs and then the reaction was 'quenched' (this prevents the enzyme from working, usually by denaturation) by placing the reaction into a hot block at 65°C for 15mins. Again the products were pooled together and another precipitation was performed, re-suspending the product in 50µl of TE.

All the sample was then run on 1% LMP (low melting point) agarose in order to separate out the correct size fragments (200-500 bp) (Fig 3.3a). These fragments were excised from the gel and purified using gel extraction kits (Qiagen). After purification the samples were pooled and the DNA precipitated, and the pellet re-suspended in 50µl of TE.

**Figure 3.3a. Gel photograph following digestion of whole genomic DNA with the restriction enzyme Mbo1. Note the increased concentration of DNA in the target region 200-500bp.**




***(iii) Ligation***

This stage was necessary so that adapter molecules complementary to the *Mbo*1 'sticky ends' (the GATC sequence) could be added to the sample fragments of DNA, to allow amplification and later ligation of the unknown fragments of DNA into plasmids (Fig.

3.3b). T4 ligase enzyme was added to the sample along with the adapter molecules and it was left to incubate overnight at 4°C. The ligated fragments (DNA + adapter molecules) were then purified using PCR purification kit (Qiagen).

**Figure 3.3b. pGEM®-T Easy vector circle map and sequence reference points (Promega)**

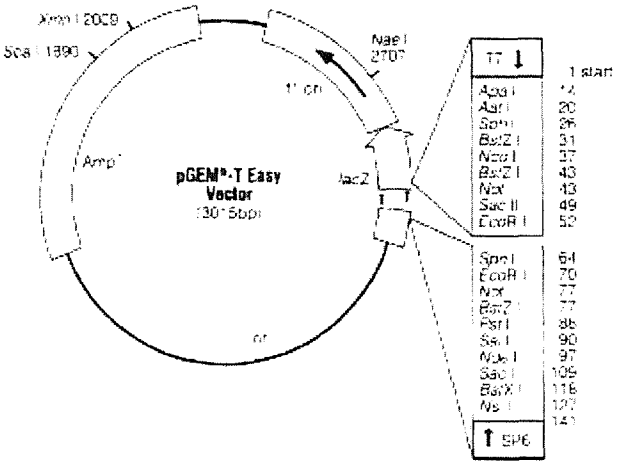



**Figure 3. pGEM®-T Easy Vector circle map and sequence reference points.**

**pGEM®-T Easy Vector Sequence reference points:**

| | |
|---|---|
| T7 RNA Polymerase transcription initiation site | 1 |
| SP6 RNA Polymerase transcription initiation site | 141 |
| T7 RNA Polymerase promoter (−17 to +3) | 2999–3 |
| SP6 RNA Polymerase promoter (−17 to +3) | 139–158 |
| multiple cloning region | 10–128 |
| *lacZ* start codon | 180 |
| *lac* operon sequences | 2836–2996, 166–395 |
| *lac* operator | 200–216 |
| β-lactamase coding region | 1337–2197 |
| phage f1 region | 2380–2835 |
| binding site of pUC/M13 Forward Sequencing Primer | 2956–2972 |
| binding site of pUC/M13 Reverse Sequencing Primer | 176–192 |

### *(iv) Amplification*

At this stage a PCR is performed (Chp 2). Primers complementary to the adapter molecules were added to the reaction mixture along with *Taq* polymerase. Following the PCR the product was purified into 50µl of elution buffer using Qiagen PCR purification kit.

### *3.3.3 Enrichment*

This stage means that fewer clones need to be screened to look for the target microsatellite. This is a result of pre-selecting DNA fragments which contain the microsatellite motif required.

#### *(i) Hybridisation to probe*

A synthetic probe was hybridised to the purified PCR product. These probes were microsatellites labelled with biotin (Genosys). The sort of microsatellite required determined the probe used. For the purpose of this study a dinucleotide repeat $(CA/GT)_{10}$ was used as the probe targeting the microsatellite required. The PCR product was denatured at 100°C for 10 min and then put on ice to prevent re-annealing. Added to the DNA were 2 µl of biotinylated probe, 100µl of buffer (1M Na Phosphate), 10µl of SDS (10%) and 40µl of $dH_2O$. The tube was then centrifuged and incubated at 50°C for 18 hrs.

#### *(ii) Capture of probe*

The biotin of the probe binds strongly to streptavidin. By introducing streptavidin coated magnetic beads (Promega) into the reaction it was possible to collect with a magnet those beads which had bound with the biotin labelled probe. Beads were then washed with increasing stringency and the DNA fragments of interest were eluted. These were the fragments specifically bound to the microsatellite probe.

Before the process began the beads were washed in buffer A (Tris-NaCl). After 3 washes, 600µl of buffer A and 6µl of herring sperm DNA (to block beads with junk DNA and prevent non-target DNA fragments from adsorbing to the beads) were added. The tube was then incubated for 30 min at room temperature. By means of a magnet the supernatant was then removed, and the beads were washed again in buffer A. 200µl of the hybridised PCR product was then added to the 100 µl bead solution, and the tube was then incubated at room temperature for 30 min the supernatant from this

wash formed wash 0. A series of washes, combined with using a magnetised stand to hold the beads were then conducted. Five washes, each of 600 μl, were then conducted at increasing temperature and decreasing salt concentrations, this increased the stringency of the washes (Table 3.3a). All washes were then purified. A PCR was then conducted on each wash to amplify the DNA fragments, primers used were those from earlier amplification reactions.

**Table 3.3a. Conditions used in the series of washes to elute the hybridised DNA fragments bound to the magnetic streptavidin coated beads**

| Solution | Temperature (°C) | Time (min) | Wash |
|---|---|---|---|
| None | Room temperature | 30 | 0 |
| 1 x Buffer A | Room temperature | 10 | 1 |
| 1 x Buffer A | 55 | 10 | 2 |
| 0.1 x Buffer A | 55 | 10 | 3 |
| 0.1 x Buffer A | 65 | 10 | 4 |
| Sterile distilled water | 65 | 10 | 5 |
| 1 x Buffer A | Room temperature | - | 6 |

### *3.3.4 Detection of enriched fragments*

#### *(i) Southern blot*

A 2% TBE gel was run with 4μl of each sample for 1 hour. The gel was then placed on a positively charged nitro-cellulose membrane in a Biorad vacuum blotting apparatus and sealed with 1% agarose solution. 1M NaOH solution was then poured over the set up so that an alkaline transfer could occur. The gel was blotted for 2 hrs at 55 mbar. After blotting the membrane was washed in 2x SSC and the DNA was fixed onto the membrane by cross-linking it under UV light.

### *(ii) Hybridisation of probe to fragments*

The membrane was rolled into a mesh soaked in 2x SSC and placed in a hybridisation flask. 25 µl of pre-hybridisation solution was added, this contained junk DNA, which prevented non-specific hybridisation. The flask was incubated for 1 hr at 42°C. The pre-hybridisation solution was removed and replaced with pre-heated pre-hybridisation solution and 50µl of digoxygenin-labelled oligo probe. This probe specifically hybridises to the target microsatellite. The flask was incubated at 50°C overnight. Hybridisation washes were then carried out to remove any non-specific binding of the probe to the membrane or PCR products. As before washes were carried out by lowering the salt concentration and increasing the temperature, so increasing the stringency of each wash (Table 3.3b). Each wash contained 50 ml of solution in a dish into which the membrane was placed and gently agitated.

**Table 3.3b. Wash conditions used to remove any non-specifically bound probe for $(CA/GT)_{10}$ microsatellites**

| Solution | Time (min) | Temperature (°C) | No. of washes |
|---|---|---|---|
| 2 x SSC, 0.1% SDS | 5 | Room temperature | 2 |
| 2 x SSC, 0.1% SDS | 10 | 55 | 2 |
| 0.5 x SSC, 0.1% SDS | 20 | 55 | 2 |

### *(iii) Detection of fragments*

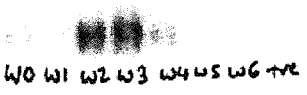
This stage allows identification of those fragments with the required microsatellite sequence. The membrane carrying the hybridised probe and bound antibody conjugate was allowed to react with a chemiluminescent substrate and was then exposed to light sensitive film in order that the signal could be recorded.

The membrane was first washed in 50ml of pre-detection solution (498.5ml maleic acid buffer, 1.5ml Tween (Sigma)) for 5 min. A blocking solution (45ml maleic acid buffer 5ml blocking reagent (Roche)) was then applied to the membrane preventing non-specific attraction of the antibody to the membrane. The membrane was then soaked in the antibody solution for 30 min to allow the anti-digoxygenin antibody alkaline

phosphatase to bind with the hybridised probe. The membrane was then washed in 50 ml of pre-detection solution. It was then placed in a solution of detection buffer (Tris NaCl pH 9.5), then placed on a piece of acetate and covered with CSPD (chemiluminescent alkaline phosphatase substrate) which allowed signal detection. Another piece of acetate was placed on top and then the whole set up was placed in a plastic bag and wrapped in foil. Incubation at 37°C for 20 min was then carried out.

Finally the membrane was exposed to film and the light signal from the site of the hybridised probe was detected (Fig. 3.3c).

**Figure 3.3c. Acetate showing signal from nitro-cellulose membrane. Darker areas relate to stronger signal indicating the presence of many microsatellite motifs.**



### *3.3.5 Cloning*

#### *(i) Insertion into plasmids*

The wash which produced the strongest signal (Fig. 3.3c), the most enriched wash, was used in the next stages. T4 ligase was used to insert the DNA into pGEM®-T plasmids (Fig. 3.1b - see earlier). The ratio of insert to plasmid used was 2:1. A positive control was used to verify the ligation reaction and a background control was used to estimate the proportion of recirculating plasmids i.e. those without the DNA insert. Ligation reactions were incubated overnight at 4°C.

### *(ii) Transformation*

At this stage *E. coli* strain JM 109 were transformed with the pGEM®-T plasmids. 2µl of each of the ligation reactions (DNA, positive and background controls) were placed in separate falcon tubes on ice and 50µl of JM 109 high efficiency competent cells (Promega) were added. After 20 min the tubes were heat shocked for 40 sec at 42°C this allows the transfer of the plasmid through the bacterial membrane. The tubes were then returned to ice for 2 min. 950µl of SOC media was then added to the tubes and they were incubated at 37°C for 1.5 hr, whilst being gently agitated.

100µl of each transformation reaction was then pipetted onto prepared antibiotic (agar made up with 1ml of 100mg $ml^{-1}$ ampicillin (Sigma)) agar plates. The plates had also been prepared with 100 µl IPTG (100 mM) and 20 µl of X-Gal (50 mg $ml^{-1}$). The IPTG and X-Gal are substrates for the enzyme β-galactosidase. The IPTG reacts with the bacteria causing the production of β-galactosidase which in turn reacts with the X-Gal producing a blue colour. If successful cloning of an insert in the vector had occurred the coding sequence of β-galactosidase was interrupted and so the colony remained white.

Inoculated plates were incubated overnight at 37°C. Untransformed JM 109 bacteria were unable to grow as the plasmid gives ampicillin resistance. Identification of the colonies containing insert in the vector is by blue/white selection. White colonies are chosen and plated out onto a gridded dish with LB medium, and incubated overnight at 37°C. Duplicate plates were made, so that one plate could be used for detection and one for later selection and purification.

### *(iii) Blotting*

Colonies were blotted using a nylon membrane and these were treated with alkaline solutions (Table 3.3c) to lyse the cells. The membrane was not wetted too much to prevent loss of the colonies, this was achieved by placing it onto filter paper which had been wetted with the appropriate solution (Table 3.3c). Denatured DNA was then fixed

to the membrane by cross-linking under UV light. 500 μl of a 1:10 dilution of proteinase K (15.6 mg μl) was dribbled over the surface of the membrane, which was then incubated at 37°C for 1 hr to digest any remaining bacterial proteins.

**Table 3.3c. Procedure followed to lyse the blotted colonies and free the target DNA**

| Solution | Time (min) |
|---|---|
| Denaturing | 15 |
| Neutralising | 15 |
| 2 x SSC buffer | 10 |

*Pre-hybridisation, hybridisation and detection*

Protocol for colony blot membranes as previously described for southern blot membrane.

***(iv) Purification of the Plasmids***

Those colonies showing the strongest signal were selected from the duplicate plates. These were grown overnight in a falcon tube containing 6ml of LB medium at 37°C in a shaking incubator. The tubes were then centrifuged for 5 min at 10 000 x g and the supernatant discarded. The Wizard-Plus SV Miniprep DNA purification system (Promega) was used to purify the plasmids before sequencing. 250 μl of cell resuspension solution was added and the tubes vortexed to resuspend the cells. Solution was transferred to 1.5 ml tubes and 250 μl of cell lysis solution was added and the tubes inverted to mix. Tubes were incubated at room temperature until the cell suspension cleared and then 10 μl of alkaline protease solution was added and the tubes mixed by inverting. After incubation at room temperature for 5 min 350 μl of neutralisation solution was added and the tubes mixed by inversion. Following that the samples were centrifuged for 10 min at 14 000 xg.

The supernatant was removed and put in a clean 2 ml tube and a precipitation reaction was conducted (Chp 2), the pellet was resuspended in 30 μl of sterile $H_20$. Following this preparation all samples were subjected to a Qiagen PCR clean-up (Chp 2) and then quantified by electrophoresis on a agarose gel (Chp 2) before being sequenced (Chp 2).

Those sequences containing microsatellites were selected and primers were designed for them.

### *3.3.6 Primer design*

The first stage in the process was to locate the insert sequence by identifying the sequences flanking the insertion site of the plasmid, and the *Mbo* 1 adapter sequences. Once the sequence was identified and the presence of a suitable microsatellite confirmed primer design followed. The freeware programme Primer 3 (Rozen & Skaletsky 1998) was employed to help design and chose primers from the sequence obtained for each sample. There are several criteria to consider when designing primers not least the behaviour of the primers during the PCR process. There are rules, which when applied to the design, help ensure optimum performance of the primers.

Typically a primer pair was designed from the sequence, these two oligonucleotides were between 15 and 25 nucleotides long. For amplification to be successful similar melting temperatures (Tm) for each primer were chosen as well as the selection of primers that would maximise the stability and specificity of binding to the desired template. The Tm is usually about 5°C higher than the annealing temperature and was ideal around 60°C. The calculated Tm from Primer 3 was just a guideline and consideration was given to the concentrations of the ions $Mg^{2+}$ and $K^{+}$ that would be in the solutions as these can affect the Tm.

In order to maximise the stability of the binding the following factors were considered; primer/template composition, primer/template base order and primer and/or template secondary structure. Base composition effects the stability as a result of the G-C content. G-C bonds are more stable and so raise the Tm, so primers with a G-C content of between 40-60% were chosen. This ensured stability of binding but avoided problems with either internal secondary structure or stretches of repeats of any one base. However, G-C content was not as important a consideration as Tm, avoidance of primer/dimers or primer length.

Primer secondary structure was an important consideration as strings of Gs and Cs can form internal non Watson & Crick base pairs which can disrupt stable primer binding. To observe this rule and avoid this problem any areas with greater than 3 Gs were avoided. Problems with self complementarity also needed to be considered, in order to avoid hair pin structures which may have disrupted stable binding.

In order to maximise specificity, primer binding needed to be minimised in areas other than the target region and to other primers. To ensure this, the binding of the primer to template needed to be strongest at the 5' end. One way of avoiding stronger bonding at the 3' end (as this is all the binding required by polymerases to begin extension) was to avoid G-C bound primer/template. Where 5' bonding is strongest this will occur first followed by the 3' end, and then extension begins from the 3' end, which will mean the whole primer is involved in distinguishing the correct region. Minimising binding to other primers was achieved by avoiding complementarity at the 3' end.

All the above criteria were defined when using the Primer 3 programme to chose primers from the plasmid sequences. From the primer pairs provided by the programme a first choice was made for each sequence based on lowest self complementarity and best matches in Tm for the primer pairs. The overall yield of specific DNA target sequences can be affected by many components of the PCR reaction (Chp 2). The design of primers is just the first step to producing the perfect PCR product.

### *3.3.7 Gene Scan*

Once primers had been chosen they were optimised using the parameters described above and once working, one of each set of primers was ordered with a 5' fluor-label. Samples were then subject to PCR and microsatellite alleles were visualised on an ABI 377 automated sequencer using multiplex electrophoresis (Chp 2). Data were then processed (Chp 2) and loaded into downstream software for further analysis (see individual species chapters 5 & 6).

# 4.0 The Genetic Population Structure of *Nematostella vectensis* in the UK

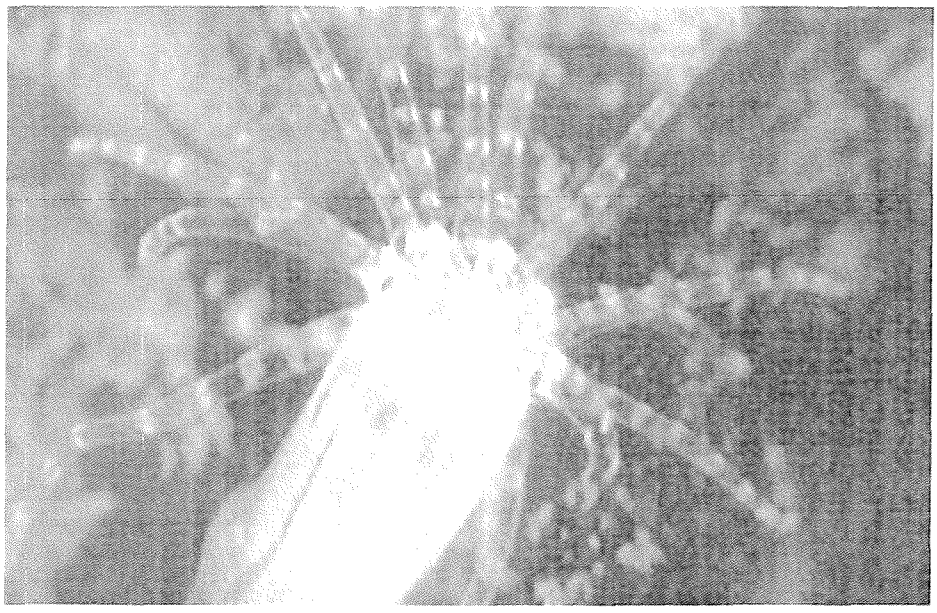
# 4.0 The Genetic Population Structure *of N. vectensis*

# 4.1 Introduction

## *4.1.1* Nematostella vectensis

*Nematostella vectensis* (Stephenson) (Fig. 4.1a), was first described from Bembridge Harbour lagoon on the Isle of Wight . Owing to the current precarious situation facing lagoon habitats in Britain *N. vectensis* has been listed in the IUCN Invertebrate Red Data Book (Williams 1983), where it is considered vulnerable to extinction in the UK. In 1988 it was scheduled for protection in Britain under the Wildlife and Countryside Act 1981. In addition, UK Biodiversity Action Plan objectives have listed this anemone as a key species for which conservation action plans will be implemented (Anon 1995).

**Figure 4.1a. The starlet sea anemone Nematostella vectensis**



### *(i) Ecology and Life History*

### *(a) Distribution*

*Nematostella vectensis* is found in brackish saline ponds or lagoons along the southern and eastern coasts of the UK. It is abundant in North America where it has been recorded in lagoons, salt marsh creeks and also subtidally in certain estuaries of both the Pacific and Atlantic coasts (Sheader *et al.* 1997). To date, in the UK, only female

specimens have been identified (Sheader *et al.* 1997). American populations have been shown to consist of male and female individuals, with single sex groups also being identified at some sites (Hand & Uhlinger 1994). Male specimens from the USA have been mated successfully with British females in the laboratory (M. Sheader pers. comm.).

### *(b) Habitat*

*Nematostella vectensis* is small, usually less than 10 mm, although specimens up to 10 cm have been identified at some sites (C. Pearson pers. obs.). In lagoons it is found burrowed in the muddy sediment leaving only its tentacles spread over the surface. *N. vectensis* makes use of the sediment to protect itself from predators, and if disturbed will withdraw until completely buried below the surface (Harter 1997). It may be found associated with lagoon flora, such as *Chaetomorpha linum* (O.F. Muller). Lagoons in the UK where *N. vectensis* has been identified share several features. These include; a high organic content, the absence of freshwater (excepting rain water), regular or frequent input of sea water and very low flow rates near the bottom of the lagoons (Sheader *et al.* 1997). Principle Component Analysis indicates that salinity range and sediment heterogeneity are the most important factors governing the presence of *N. vectensis* and other species in the lagoon environment (Sheader *et al.* 1997).

Submergence at all states of the tide is another characteristic typically associated with populations of *N. vectensis*. Some populations, however, have recently been identified in lagoons with little or no water coverage during high summer (C. Pearson pers. obs.). Studies using multivariate component analysis group, among others, the lagoon cockle *Cerastoderma glaucum* and the lagoon sand shrimp *Gammarus insensibilis* with *N. vectensis* (Sheader *et al.* 1997).

The lagoons inhabited by *N. vectensis* can be hostile environments with large ranges in salinity and temperature. Salinity data ranging from 10 to 51.54 have been recorded at sites where this anemone has been found and ranges in temperatures from -1 $^{o}$C to 32.5 $^{o}$C (Hand & Uhlinger 1994). Moreover, there is evidence to suggest that this anemone is capable of reproduction under these harsh conditions. Hand & Uhlinger (1992) have

witnessed sexual and asexual reproduction under laboratory conditions in salinities of 12 ‰ to 34 ‰ and 7 ‰ to 42 ‰ respectively.

### *(c) Life history*

Sea anemones may be aclonal, obligately clonal or clonal with a sexual phase (Shick 1989). Those species that are opportunistic and who exploit unstable or marginal habitats, tend to be small and reproduce asexually at a rapid rate when the conditions are right e.g. *Haliplanella luciae* (Hughes 1989). Other, larger, sea anemones e.g. *Anthopleura elegantissima* who reproduce agametically at a slower pace, tend to survive for long periods of time competing for resources with other sexually reproducing species (Sebens 1981). The biological significance of asexual reproduction for these two species of anemone seems to differ (Hughes 1989).

*Nematostella vectensis* is unusual because, in addition to sexual reproduction, it has the ability to reproduce asexually by transverse fission at any time of the year, and this reproductive ability is consistent under both field and laboratory conditions (Sheader *et al.* 1997). It is one of only five species of sea anemone known to possess this vegetative reproductive mode (see Hand & Uhlinger 1995 for full report). Natural populations of *N. vectensis* may be unisex or contain both sexes. Hand & Uhlinger (1994) propose that unisex clonal populations arise because of no or low recruitment of sexually produced individuals.

It is important to note that highly variable asexual division rates have been reported for this species and these may be affected by food availability (Hand & Uhlinger 1992; Hand & Uhlinger 1995). In addition, some individuals never undergo fission and this may impact greatly on the establishment and maintenance of clones in natural populations (Hand & Uhlinger 1995). The hydroid *Protohydra leukarti*, found in brackish water estuaries and lagoons of the Baltic (comparable habitat to *N. vectensis*), shows a similar life history pattern. This species has polyps which may reproduce sexually in the autumn, who survive the winter and then resume asexual reproduction, by transverse fission, in the spring (Hughes 1989). Under laboratory conditions this

species responds to wide fluctuations in temperature and salinity by reproducing sexually. However, fission is resumed immediately after gamete release, if conditions are suitable (Muus 1966).

Sexual reproduction of *N. vectensis* observed in the laboratory, reveals eggs produced in a gelatinous mucoid mass, and free swimming sperm (Sheader *et al.* 1997). It may be that the eggs of *N. vectensis* are produced in this way as an adaptation to remain in the lagoon habitat by preventing or decreasing dispersal. There are no data at present to confirm or refute this hypothesis and so for the time this remains conjecture. In the field the eggs are fertilised by free swimming sperm in the summer and autumn months (Hand & Uhlinger 1992). The resulting planulae are active swimmers and after approximately 7 days display minute tentacles and 10 weeks post fertilisation are themselves sexually active (Sheader *et al.* 1997). Active swimming planulae and an ability to reproduce asexually make this species an excellent candidate for colonising new lagoons.

***(d) Diet***

*Nematostella vectensis* appears to be a non-selective forager. Feeding is merely a neural response to the firing of the stinging organelles (Harter 1997). This anemone has been reported to feed on hydrobiid snails, copepods, midge larvae, egg masses, ostracods, worms, and rotifers (Frank & Bleakney 1978). Insect prey have also been reported in the diet of this anemone (Williams 1983). This is unusual for sea anemones, with only *Actinia tenebrosa*, *Actinia equina* and *Anthopleura elegantissima* cited as having diets which include insects (Ayre 1984a). This difference seen in the diet of *N. vectensis* is thought to be because of the general lack of insects in the marine habitat rather than any particular preference for other prey of this anemone (Hand & Uhlinger 1994).

Another important aspect regarding the diet of this anemone is that it feeds on oyster, mussel and clam larvae (Posey & Hines 1991; Hand & Uhlingher 1992). Although Posey and Hines (1991) did not asses *N. vectensis* as a predator they report an effect on settling larvae, so where present in great numbers at high densities it could have an

impact on those species with a larval stage that occupies its habitat. This is a phenomenon recorded for the sea anemone *Diadumene leucolena* which impacted on the oyster population in Chesapeake Bay, USA (MacKenzie 1977).

Field observations reveal that feeding is halted in *N. vectensis* above salinities of 40 ‰ and under laboratory conditions below 10-12 ‰ supporting the observations of Sheader *et al.* (1997) that in nature, this species is at its greatest densities in lagoons with a salinity range 16 - 36 ‰. Lagoons are hostile environments being both eurythermal and eurysaline, and these factors change from summer to winter, from year to year as well as between the lagoon sites themselves. The influence of these factors in halting activities such as feeding may jointly affect the timing and reproductive effort of *N. vectensis* (Sheader *et al.* 1997). There is, however, laboratory evidence to suggest that while starvation can depress asexual reproduction it does not halt it altogether, and similarly extreme starvation (6 months) does not lead to the death of *N. vectensis* (Hand & Uhlinger 1995). The population dynamics of this sea anemone may therefore be relatively well adapted to the harsh environment proffered by the lagoon habitat.

***(ii) Clonal reproduction and genetic implications***

Agametic reproduction is common in organisms where every cell type can be regenerated, for example: cnidarians, ctenophores, platyhelminths, annelids, nemertines and most echinoderms (Hughes 1989). There are several methods of agametic reproduction; fragmentation, budding and fission, all with the same consequence, the production of clones.

Asexual reproduction permits reproduction in isolation which can lead to one individual founding a whole population (Hand & Uhlinger 1994). In addition, energy can be directed into faster propagation because none of the activities associated with sexual reproduction are necessary (Hughes 1989). For any founding species rapid colonisation of a new environment can occur most efficiently by the cloning of a successful genotype (Shick & Lamb 1977). It is thought that for species living in marginal habitats, or in

rapidly changing environments, selection for tolerance can proceed more rapidly and efficiently among clones (Jackson & Coates 1986).

It has long been assumed that natural selection works on genotypic diversity created as a result of sexual reproduction (Hughes 1989). Lynch (1984a, b) challenges this idea suggesting it only accounts for evolution of "the species" and ignores the ability of clonal lineages to respond to changing environments. There is no argument against the idea that genotypic diversity, created by sexual reproduction, enables a species to adapt and respond to selective pressure, but there is an argument against selection changing enough to favour sex over a clonal mode of reproduction (Hughes 1989).

Vrijenhoek (1985) suggested that one of the most important implications of clonal reproduction was "Heterozygosity Assurance". Cloning protects against loss of heterozygosity and inbreeding depression as the parent genome is preserved by clonal reproduction (Vrijenhoek 1998). This longevity, however, can be affected by environmental changes in the short term and ultimately it may be lost by the accumulation of deleterious mutations (Hughes 1989). Muller (1964) suggested that asexual lineages have a high rate of extinction because deleterious mutations can accumulate in a "ratchet mechanism". Individuals are unable to produce offspring more successful than themselves (except when rare back mutations occur) and once the best clone is lost, the second best will follow and so on until mean fitness is reduced so much that extinction is inevitable (Lynch *et al.* 1993). The effects of Muller's ratchet are more evident in organisms with large numbers of genes existing in finite populations as lucky clones are more easily lost by chance (Vrijenhoek 1998). Theoretical analysis suggests that a population such as the one described above may become extinct in 10 000 - 100 000 generations (Vrijenhoek 1998). For an opportunistic species such as *N. vectensis* which appears to follow a 'boom and bust' life history strategy there would not be time for this to pose a serious problem.

The loss of genotypic variation from populations is prevented by cloning which lowers the chance of genomic extinction (Hughes 1989). Localised populations of anemones such as *Anthopleura elegantissima* demonstrate the true genotypic diversity of a

panmictic population as there appears to be no local selection and individuals of different clones exist together (Smith & Potts 1987). This, however, is not true for all species of anemone. Many appear to be dominated by a few clones e.g. Australian populations of *Actinia tenebrosa*, where interclonal aggression separates polyps into groups of clonemates (Ayre 1984b). For species such as these, with locally structured populations, genetic diversity will be typically reduced, certainly much lower than the entire panmictic population (Ayre 1984b). Successful clones will give temporal stability to genotypic frequencies. Since arrival over 90 years ago *A. tenebrosa* clones in south-western Australia have maintained low levels of genetic diversity, indicating little sexual recruitment, or intense selection following sexual reproduction (Ayre 1985). Either way there is little indication that these clones will not continue to be successful for many more decades.

Genotypic frequencies can be distorted by a sequence of clonal generations, for example heterozygote clones that have the advantage will increase in frequency through cloning (Hughes 1989). Clonal populations that possess selectively neutral genotypes, initially in Hardy-Weinberg equilibrium (HWE), will demonstrate an equal probability of them drifting towards homozygosity or heterozygosity (if sexual reproduction does not occur) (Herbert 1987). However, founder effects and inbreeding can cause a deficiency in heterozygotes in a population, but any mixis that may occur will tend to restore HWE (Herbert 1987). Other effects on the genotype of clonal animals will depend on whether or not asexual reproduction is agametic. Mutations are not as likely to be transferred in agametic reproducers as offspring are not produced from a single cell (Hughes 1989). Through the processes mentioned above it can be seen, therefore, that cloning can have a marked effect on the genetics of populations.

### *4.1.2 Conservation of the species*

#### *(i) Introduction*

*Nematostella vectensis* is considered rare and endangered in the UK (Williams 1983) principally through loss of its lagoonal habitat (Downie 1996; Bamber 1997). This is

recognised by its appearance on the list of globally threatened and declining species (Downie 1996). Molecular techniques are now widely available which allow access to the genetics of endangered populations (Sunnucks 2000). In the present study the RAPD technique (Williams *et al.* 1990) was used to determine the genetic variation among UK populations and assess the mode of reproduction in *N. vectensis*. Other studies report on the genetic population structure of some sea anemones (Black & Johnson 1979; Ayre 1985; Ayre *et al.* 1991), but no such studies on the euryhaline anemone *N. vectensis* have been undertaken.

### *(ii) Use of RAPDs in conservation genetics*

RAPDs have many advantages over other molecular markers for use in conservation genetics. The genome of the target species does not need to be known, a potentially unlimited number of loci can be utilised for genetic analysis, the method is not as labour intensive or expensive as others and perhaps most importantly (for threatened and endangered species), it requires very little DNA (Fritsch & Rieseberg 1996). There is a disadvantage to the use of RAPDs, relating to the question of dominance (the appearance of a locus cannot be distinguished as either homozygote or heterozygote) and this may create difficulties in describing genetic variation from the data (Fritsch & Rieseberg 1996). In spite of the problems associated with RAPDs relating to dominance it has been shown to be a useful method for investigating genetic variation within species, and a robust method for testing hypotheses of clonal structure (Fritsch & Rieseberg 1996).

One important use of RAPDs in conservation biology has been as a diagnostic marker where rare species are morphologically very similar to common relatives. RAPDs have been used successfully as species specific markers to identify closely related species of mosquito (Kambhampati *et al.* 1992). The taxonomic status of an endangered or rare organism may be of great importance in terms of its conservation. A study of the autumn buttercup (*Ranunculus acriformis* var. *aestivalis*) used RAPD markers, where morphological characters were not of use, to confirm its status as more closely related to one species of buttercup than another (Van Buren *et al.* 1994). Templeton (1991)

describes the importance of understanding the reproductive system of endangered species in order that effective and successful conservation plans may be implemented. RAPDs have shown themselves to be a useful marker for distinguishing methods of reproduction, determining paternity and for estimating outcrossing rates in plant species (Fritsch & Rieseberg 1992; Hadrys *et al.* 1993; Pearson *et al.* 2002).

It is becoming increasingly common for molecular genetics data to form an integral part of conservation management strategies for threatened species (Scher 1996). For some declining plant species the importance of understanding clonality for their conservation management has been demonstrated (Stewart & Porter 1995; Sydes & Peakall 1998). Huff *et al.* (1998) suggest that it is essential to gain knowledge of the genotype of any species for which transplantation is intended, as local species may have ecologically different histories. It is conceivable that transplantation may be a useful strategy for conservation of *N. vectensis* within the UK. For the successful implementation of a conservation action plan for *N. vectensis* it is important to understand the genetic population structure both within and between lagoons. The present study seeks to determine the genetic structure of *N. vectensis* populations, by using RAPD markers, and to address the extent to which they are clonal and provide data to assist their conservation management.

## 4.2 Methods

### *4.2.1 RAPD-PCR as a molecular marker*

The RAPD technique has been shown to be a useful tool for investigating clonal population structure among plants (e.g. van de Ven & McNicol 1995; Gabrielsen & Brochmann 1998; Bush & Mulcahy 1999) and to a lesser degree in animal populations (e.g. Tyler-Waters & Hawkins 1995). RAPDs have the potential to reveal large numbers of polymorphisms by screening large numbers of alleged non-coding regions of the genome and there is evidence that RAPD markers have a high mutation rate (Comes & Abbott 2000).

#### *(i) Robustness and repeatability of RAPDs*

Despite the extensive use of RAPDs as a molecular marker, concerns have been raised over the lack of repeatability for some loci (see Hadrys *et al.* 1992; Grosberg *et al.* 1996 for full review). This repeatability is increased greatly when weaker bands are omitted from the scoring protocol, as these vary the most widely on multiple runs (Bowditch *et al.* 1994; Cooper 2000). Consequently, a control series was set up using five individuals from N1 (Gilkicker). These individuals were amplified for each primer a total of 10 times and only the brightest bands were used to set the scoring protocol. Across these data, bands chosen for the scoring protocol showed a repeatability of 100%.

As suggested by several workers (e.g. Welsh & McClelland 1990; Borowsky & Vidthayanon 2001) we performed all reactions with uniform template concentrations for all sample sites. To verify the accuracy of the data from selected loci, all reactions were repeated 1 year later. The total number of bands amplified for the replicate series was 10051 (from a possible total of 10447) and gave a repeatability of RAPD band amplification of 96.2%.

### *(ii) The RAPD technique*

In this procedure a single oligonucleotide primer of arbitrary sequence is used in the PCR reaction (Chp 2) which contains genomic DNA of the organism under investigation. The primers are single stranded and usually 10 nucleotides in length (decamers). If the primer binds to areas of the genome within approximately 3 kilo bases (kb) of each other, the region between will be amplified. More than one DNA fragment may be amplified in such a reaction, as the primer may bind at any number of locations along the genome. In addition, many different primers can be used and these may produce yet more amplified fragments. These fragments are usually visualised on agarose gels after staining with electrophoresis (Chp 2).

In order to enhance the priming potential of RAPDs, the environment in the PCR is quite different to that of a standard reaction (Chp 2). A high $MgCl_2$ concentration is required to augment primer annealing, and to increase enzyme activity. This results in many more fragments being amplified by the arbitrary primers as specificity between primer site and target DNA is reduced. In addition, the annealing temperature in the reaction is typically low, usually 35-36 $^{o}$C, encouraging non-specific binding. These factors combined with an increased number of cycles, around 40, maximise the potential for amplification of loci using this technique.

## *4.2.2 Materials and methods*

### *(i) Sample collection*

Whole animal samples were collected from lagoons between October 1998 and August 1999 (Table 4.2a). Sampling was undertaken at several locations at each lagoon thus reducing the probability of collecting individuals of the same parentage. Between 40 and 60 individuals were collected at each lagoon. Where possible the maximum number of anemones were collected, but at sites where individuals appeared relatively scarce (e.g. Seafield Bay) fewer were sampled. Samples were collected and placed directly into 95% ethanol.

### *(ii) DNA extraction*

The entire anemone was used for DNA extraction, except in a few individuals that were over 1cm in length. In these individuals a $1mm^3$ piece of tissue was obtained for use in the extraction. The PCI extraction method (Chp 2) was used for all samples. Sterilised $dH_20$ was used to dilute the extractions to 5 ng/µl. Samples were subsequently stored at -20°C until required for RAPD-PCR reactions.

**Table 4.2a.** ***Nematostella vectensis*, numbers of individuals collected and sample locations**

| Name of site | Code | UK Ordnance Survey Map Grid Reference | No. of individuals used for RAPD profiling |
|---|---|---|---|
| Fort Gilkicker Lagoon | N1 | SZ 608 978 | 40 |
| Bembridge Harbour Lagoons & Harbour Farm Lagoons | N2 | SZ 633 884 to SZ 642 884 | 40 |
| Yar Bridge Lagoon | N4 | SZ 349 897 | 39 |
| Normandy Farm Lagoon | N5 | SZ 332 947 | 39 |
| Half Moon Lagoon | N6 | TG 049 453 | 40 |
| Reedland Marshes | N7 | TM 483 071 to TM 483 726 | 37 |
| Kings Marsh Lagoons | N8 | TM 445 450 | 39 |
| Seafield Bay | N10 | TM 122 330 | 27 |
| The Naze Lagoon | N11 | TM 217 238 | 36 |

### *(iii) RAPD-PCR*

Initially RAPD amplification was performed on a subset of 5 individuals from Gilkicker lagoon (N1) using Operon primers (Operon Technologies, Alameda, California, USA). Primers OPF-1 to OPF-20, OPU-1 to OPU-20 and OPK-1 to OPK-20 were screened. Of these, six primers (OPF-06, OPF-14, OPU-08, OPU-11, OPK-10 & OPK-17) consistently produced easily visualised and reproducible bands (Table 4.3a). For these and subsequent reactions the following protocol was used. Each 20 µl PCR reaction contained between 5-10 ng template DNA, 30 pM primer, 2 µl 10X buffer (Perkin Elmer), 3mM $MgCl_2$ (Perkin Elmer), 4 mM of dNTP mix (Perkin Elmer) and 1 unit of

*Taq*-polymerase (Perkin Elmer). Amplification was performed using a Perkin Elmer thermocycler (Model 480) programmed for 3 cycles of 94°C for 3 min, 36°C for 30s and 72°C for 90s followed by 42 cycles of 60s at 94°C, 60s at 36°C and 90s at 72°C. For the final cycle the extension step was held for 10 min. Once suitable primers had been chosen 40 individuals from N1 were profiled with all primers. The *N. vectensis* population from N1 was then used as the reference population against which to compare the remaining sites. All successfully extracted individuals from each site were then profiled against the same five individuals from the reference population, which were also considered as the control. For four sites (N1, N2, N4 and N6) all reactions were performed twice. All samples were visualised using the Uvidoc system (Uvidoc 008-XD documentation system with Uvisoft Version 98 gel quantification software) after being run on 2% gels made with Metaphor agarose (FMC Bioproducts, Rockland, Maine) stained with ethidium bromide. Gels were run at 150 V for 3 h in 1x TBE.

***(iv) Analyses***

Individuals from the reference population, used to screen the primers, were utilised to standardise a scoring protocol. The size of each locus produced by each primer was measured using a 1kb ladder standard (Gibco BRL, Paisley, Scotland, UK) and Uvidoc software. Bands were scored for each primer as either present (1) or absent (0). Only those loci that consistently amplified brightly, in replicate samples, were counted. Each individual was then profiled according to the following protocol. The ARLEQUIN package (version 2.000, Schneider *et al.* 2000) was used to carry out an analysis of molecular variance (AMOVA; Excoffier *et al.* 1992) using a Euclidean distance matrix. This method is unaffected by the dominance problem associated with RAPDs and is thus considered a useful method for analysing this type of genetic data (Huff *et al.* 1993). Two different similarity coefficients (Jaccard and Dice) were used to calculate similarities between pairs of individuals using the RAPIDISTANCE programme (J. Armstrong, A. Gibbs, R. Peakall, G. Weiller, Australian National University, Canberra, Australia).

The Jaccard coefficient is:

$$n11/(n-n00)$$

The Dice coefficient is:

$$2*n11/((2*n11)+n01+n10)$$

Where n = The number of band positions

- n00 = The number of positions where sample 1 and sample 2 exhibit band absence
- n01 = The number of positions where sample 1 exhibits no band and sample 2 exhibits the band
- n10 = The number of positions where sample 1 shows the band and sample 2 shows no band
- n11 = The number of positions where sample 1 and sample 2 exhibit band presence

The RAPDISTANCE programme was also used for cluster analysis to produce UPGMA trees (Sneath & Sokal 1973) from these data.

## 4.3 Results

### *4.3.1 RAPD profiles*

The six primers used in this study gave a total of 31 bands (Table 4.3a). All of these markers were found to be 100% polymorphic when considered across all the populations. This situation, however, changed dramatically when considering each primer and each population separately (Table 4.3b). All primers ranged from being monomorphic to 100% polymorphic, with only primers OPF-14 and OPU-08 showing these two extremes throughout all populations. The median percentage polymorphism shown by OPU-08 across all the populations was the lowest at 0%. In one population, N6, all primers for all individuals gave identical RAPD profiles (Fig. 4.3a).

**Table 4.3a.** ***Nematostella vectensis.*** **Sequences of RAPD primers along with numbers of bands scored per primer (with upper and lower quartiles of bands present per individual) for 337 individuals.**

| Primer | 5' to 3' sequence | No. of bands | Median (quartiles) no. of bands present per individual |
|---|---|---|---|
| **OPF-06** | **GGGAATTCGG** | 6 | 6 (6,6) |
| OPF-14 | TGCTGCAGGT | 4 | 4 (4,4) |
| OPU-08 | GGCGAAGGTT | 5 | 5 (5,5) |
| OPU-11 | AGACCCAGAG | 5 | 5 (4,5) |
| OPK-10 | GTGCAACGTG | 5 | 5 (5,5) |
| OPK-17 | CCCAGCTGTG | 6 | 6 (6,6) |

**Figure 4.3a. Metaphor agarose gel showing RAPD profiles using primer OPF-06 for 5 individuals from the reference population, N1 (Gilkicker) and 13 individuals from N6 (Half Moon).**

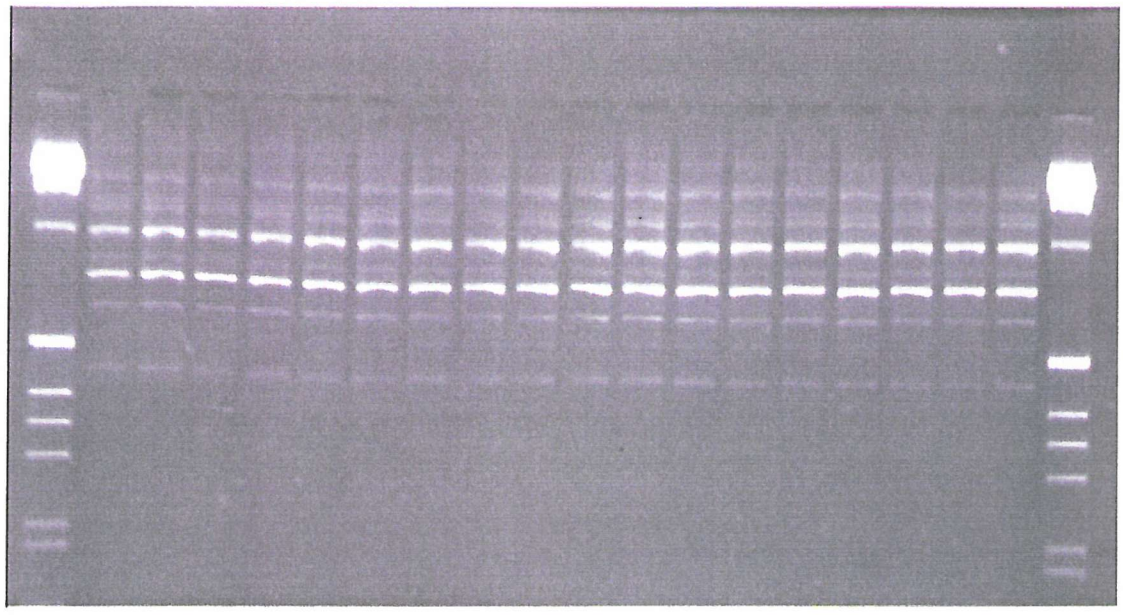
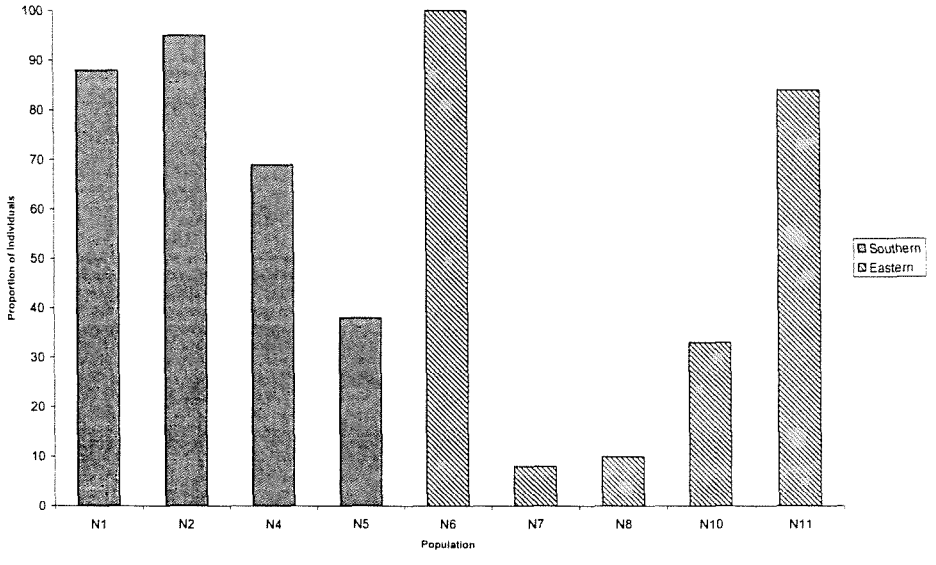


**Table 4.3b. *Nematostella vectensis*. Percentage of polymorphic markers across all populations and within each population.**

| % polymorphic markers | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | In populations | | | | | | | | | |
| Primer | No. of bands | Total | N1 | N2 | N4 | N5 | N6 | N7 | N8 | N10 | N11 | Median (quartiles) |
| OPF-06 | 6 | 100 | 33.3 | 16.7 | 83.3 | 100 | 0 | 100 | 100 | 100 | 66.7 | 83.3 (41.7,100) |
| OPF-14 | 4 | 100 | 0 | 0 | 100 | 100 | 0 | 100 | 100 | 100 | 100 | 100 (25,100) |
| OPU-08 | 5 | 100 | 0 | 0 | 100 | 100 | 0 | 100 | 100 | 0 | 0 | 0 (0,100) |
| OPU-11 | 5 | 100 | 60 | 0 | 100 | 100 | 0 | 100 | 100 | 100 | 0 | 100 (15,100) |
| OPK-10 | 5 | 100 | 60 | 0 | 80 | 100 | 0 | 100 | 100 | 100 | 100 | 100 (65,100) |
| OPK-17 | 6 | 100 | 50 | 0 | 83.3 | 100 | 0 | 100 | 100 | 100 | 100 | 100 (58.3,100) |

One RAPD profile generated by the 31 markers was shared by 61% of all individuals across the entire data set (Fig. 4.3c). Only one other RAPD profile appeared at a level equal to or greater than 5% of individuals in a single population (profile 38 in population N7). In total, 79 RAPD profiles were generated across all populations with the dominant RAPD phenotype being distributed across all sites (Fig. 4.3d). In addition to Profile 1, eight other profiles occurred in more than one population. Profile 18 was the most frequent of these, occurring in four of the nine populations but this still only represented 2.1% of the entire data set (Appendix 1).

**Figure 4.3c.** ***Nematostella vectensis*****. Proportion of individuals across the entire data set displaying the dominant RAPD profile, profile 1.**




### *4.3.2 AMOVA*

The data from the AMOVA calculated from a Euclidean distance matrix, revealed very low levels of geographical variation when the east coast sites were compared to the south coast sites. The "among" group (that is, south coast sites compared to east coast sites) variation accounted for only 2.95% of the total variance (Table 4.3c). The greatest amount of genetic variation, 78.24%, was seen among individuals within populations. In addition, the AMOVA showed highly significant genetic divergence ($P$<0.0001) among the populations within the groups, but no significant divergence between sites on the south coast and east coast.

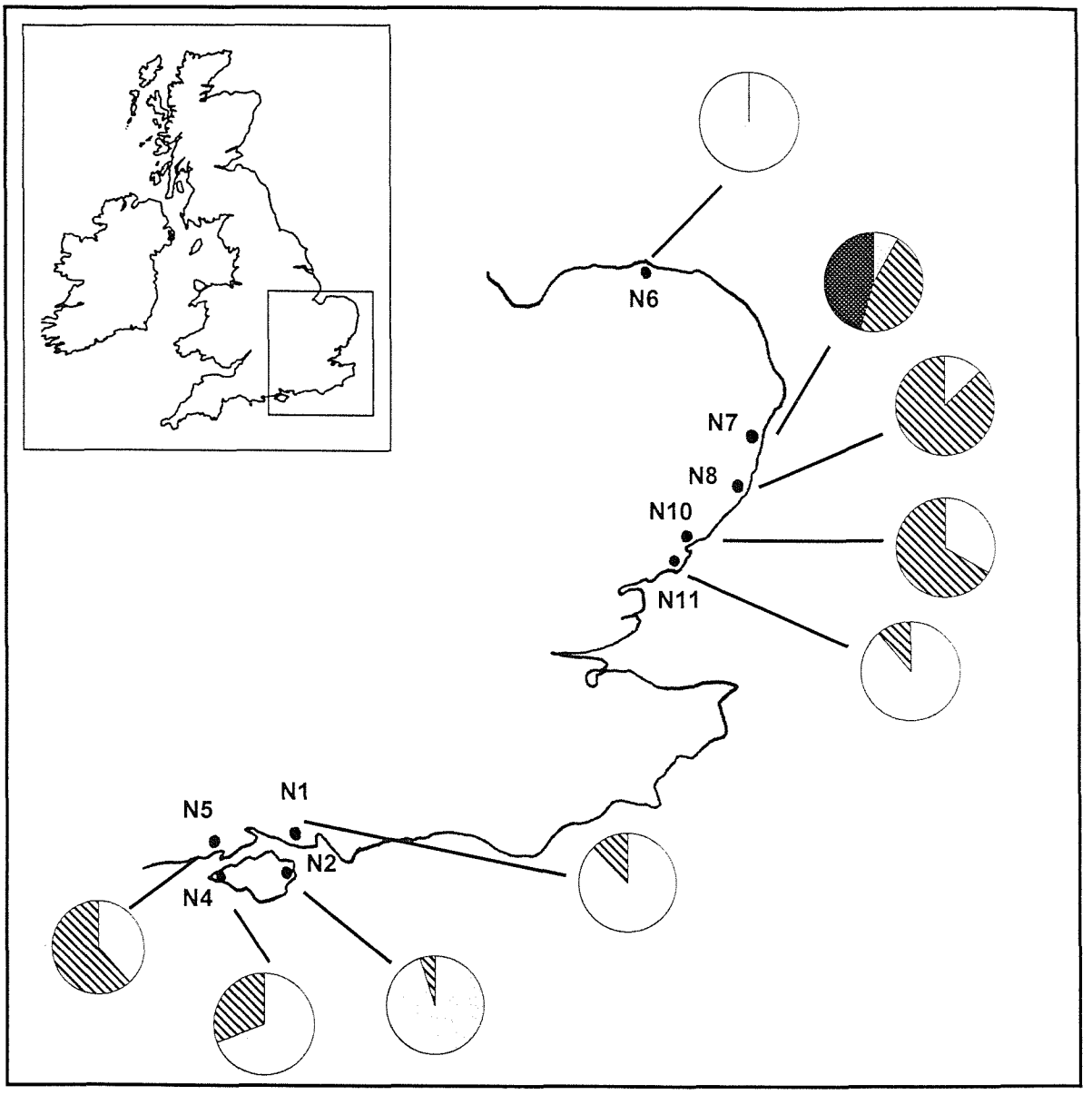
**Table 4.3c. Analysis of molecular variance (AMOVA) among groups (southern group N1, N2, N4, N5, eastern group N6, N7, N8, N10 & N11), among the populations within those groups and within all populations regardless of group. Total % distribution of the variance components and *P*-values for corresponding $\Phi$-tests are given.**

| Source of variation | d.f. | Variance components | % Total | *P*-value |
|---|---|---|---|---|
| Among groups | 1 | 0.08 | 2.95 | ns |
| Among populations within groups | 7 | 0.55 | 18.81 | <0.0001 |
| Within populations | 328 | 2.29 | 78.24 | <0.0001 |

d.f., degrees of freedom

The UPGMA trees produced using the similarity coefficients, Jaccard and Dice, were identical in their clustering. These trees, however, showed no evident geographic pattern of variation and have been omitted.

**Figure 4.3d. RAPD profiles identified in the sea anemone *Nematostella vectensis* from 9 populations in the UK**




Key: The grey areas relates to the dominant RAPD profile, profile 1. The banded areas are all profiles represented by <1% of the total sample, grouped together. The black area is the only other profile representing >5% of the sample, profile 38

## 4.4 Discussion

### *4.4.1 Evidence for clonal reproduction in* Nematostella vectensis

Morphological data indicated populations comprised females only and suggested that the anemone *Nematostella vectensis* was either clonal throughout its known southern range or a sequential hermaphrodite (Sheader *et al.* 1997). RAPD analysis supported these findings, indicating that the species is predominantly clonal throughout its known UK range. Of those individuals studied, 61% revealed identical RAPD banding patterns and some sites were monoclonal or near monoclonal, with nearly all the rarer genotypes confined to single populations (Appendix 1).

The observed population structure of *Nematostella vectensis* within the UK is likely to result from the species' reproductive mode. Clonal reproduction, combined with low levels of dispersal between lagoons, favours the significant variance detected among populations within both the southern and eastern groups of lagoons (Table 4.3c). Occasional long distance dispersal, either artificially, by rafting events or by passive transport followed by asexual proliferation of populations has resulted in the lack of genetic structure at larger spatial scales. There is no direct evidence for a larval phase of *N. vectensis* in populations found in the UK, so it is unlikely that dispersal would occur by this method. One possible method of dispersal is passive transport by wading birds which regularly travel between the lagoons of the Solent and east coast (J. Gill pers. comm.). In the absence of empirical evidence this must remain conjectural.

Passive dispersal is a feature often used to explain patterns of distribution and community composition of non-mobile organisms (Jenkins & Buikema 1998). Waterbirds have long been considered a major disperser of aquatic plant and invertebrate propagules either by endo- or ectozoochory (Figuerola & Green 2002). Evidence for such transport, however, has been at best speculative. Most recently indirect evidence has come from genetic studies involving aquatic organisms where results indicate that geographical distance and genetic distance between the populations are unrelated (Herbert & Finston 1996; Vanoverbeke & DeMeester 1997; Freeland *et al.*

2000). It is important to note that the lack of association between genetic and geographic distances is not direct evidence for dispersal by waterbirds. It is perhaps more likely an implication that for these populations colonisation events may be independent of geographic distance, or it may suggest that the populations are not at equilibrium (Figuerola & Green 2002).

Cloning as a mode of reproduction tends to coincide with situations that favour a rapid increase in population numbers, e.g. seasonal or unstable habitats (Hughes 1989). Evidence to support this is seen not just in agametic reproducers but in cyclical parthenogens, where species are able to exploit 'pulsed' resources by rapid clonal expansion (Hughes 1989). In addition to anthozoans, this life history strategy is shared by some turbellarians, oligochaetes, polychaetes and echinoderms all of whom demonstrate a general capacity for opportunism. As mentioned previously there are clonal reproducers where competition for substratum is more significant than explosive population growth, and for these species cloning is a way to avoid senescence by redistributing biomass (Shick 1989). This strategy enables survival under continuing conditions of limited resources.

In summary, cloning allows for great flexibility in life history by combining extremes in development. Where reproduction is concerned the most opportunistic animals reproduce asexually, e.g. aphids, and in terms of longevity, large clonal anemones have been known to survive longer than most aclonal animals (Hughes 1989). The life history of *N. vectensis* indicates it is an opportunistic species well adapted to life in an unpredictable environment. The ability to reproduce asexually lends itself well to the colonisation of new sites and to recovery following environmental adversity.

### 4.4.2 *Causes of observed genetic variation*

The presence of a single dominant genotype could be attributed to populations being established by a single founding individual and examples of the founder effect have been reported previously for sea anemones where genetically identical individuals have been reported at high densities (e.g. Shick & Lamb 1977). The movement of shellfish

such as oysters (*Crassostrea* spp) may be responsible for the original introduction of a successful member of this species to the UK from North America (Sheader *et al.* 1997). Repeated introductions associated with regular imports could be the cause of the higher levels of genetic variation (78.24%) seen in the lagoons clustered around the east coast where there is a high incidence of oyster farms. Oysters were repeatedly brought to this region from North America until WWII (I. Laing, pers. comm.).

The genetic diversity revealed by the RAPD markers in some of the UK populations of *N. vectensis* may, in part, be caused by non lethal somatic mutations (Hughes 1989). Lagoons are a relatively ephemeral environment, usually lasting no more than 1000 years (Downie 1996), thus deleterious mutations associated with asexual life histories may have had insufficient time to emerge (Muller 1964).

### *4.4.3 Evidence to support a clonal mode of reproduction*

Physiological evidence supports a clonal mode of reproduction for this species of sea anemone. *N. vectensis* is very eurytolerant, a characteristic in common with other colonising species of sea anemone (Shick & Lamb 1977). This eurytolerance may be associated with a single, successful dominant RAPD profile, an occurrence which has been reported using allozyme markers for the anemone *Haliplanella luciae* (Shick & Lamb 1977). It is, however, important to note that allozymes are functional enzymes which may be subject to selection whereas RAPDs, representing unknown regions of the genome, may not be under such selective pressure.

### *4.4.4 Implications for the conservation of* Nematostella vectensis

The discovery of a high incidence of clonality in UK populations of *Nematostella vectensis* has implications for the conservation of this species. The data indicate that abundance does not necessarily equate with diversity. For example, Gilkicker lagoon (N1), has previously recorded extremely high densities of *N. vectensis* (Sheader *et al.* 1997) but this study revealed relatively low genetic diversity compared with other less populated lagoons (such as Seafield Bay). Management strategies must be tailored to

protect those populations with monoclonal or low genetic diversity which may be more susceptible to mass extinction, were tolerance limits reached (Shick & Lamb 1977). It is, moreover, essential that any conservation plan must aim to preserve the already low genetic diversity present in the UK populations of *N. vectensis*.

Despite the usefulness of genetic data in conservation management it is important to appreciate that conserving the genetic diversity of a species alone is insufficient (Mace *et al.* 1996). Long term conservation relies on the preservation of habitats and ecosystems and the diversity within them (Anon 1995). Lagoons are extremely vulnerable habitats and may be lost intermittently or permanently, through natural or human activity (Bamber *et al.* 2001). The dominant clone revealed by this study occurred in all lagoons throughout the known distribution of *N. vectensis* within the UK. Where new lagoons are formed, either naturally or artificially, then introductions of specimens of the dominant genotype would be likely to be successful and may be used to maintain the distribution of this species.

### *4.4.5 The use of RAPD markers in population studies*

The current study supports previous works which show RAPD markers to be a consistent method for identifying clonal structure of marine invertebrate populations (e.g. Tyler-Waters & Hawkins 1995; Sommerfeldt & Bishop 1999). The amplified band repeatability of 96.2%, demonstrated in this study, indicates that reliable and reproducible results can be attained through careful laboratory procedure. In our, and other workers experience one of the most influential factors on repeatability was uniform template concentrations (Borowsky & Vidthayanon 2001). To obtain the most reliable results all DNA templates concentrations in this study were calculated against DNA of known concentration (Borowsky *et al.* 1995). Skroch & Nienhuis (1995) report an increase in reliability of RAPD band reproduction when amplification environments are kept consistent. For the purpose of this study all amplification procedures were performed using the same PCR thermocycler and the same reagents. Weak amplification of bands seen for some RAPD primers is probably related to low specificity between the primer and primer site (Skroch & Nienhuis 1995). We took care

to standardise a scoring protocol using only those bands showing strong amplification, and with these, obtained a 100% band repeatability for the control series. The availability of RAPDs for use with anonymous genomes and their relative lack of expense make them potentially valuable genetic markers for population studies (Hadrys *et al.* 1992).

# 5.0 Genetic Structure of *Cerastoderma glaucum*

# 5.0 Genetic Structure of *Cerastoderma glaucum*

# 5.1 Introduction

## *5.1.1* Cerastoderma glaucum

*Cerastoderma glaucum* Bruguière (Fig 5.1a) is the lagoon cockle (Lamellibranchia: Cardiidae). More genetic and ecological data exist for this species than for the other lagoon specialists being discussed in this study (e.g. Brock 1978; Brock 1979; Hummel *et al.* 1994). This may, in part, be because of interest in the relationship between *C. glaucum* and the marine cockle *Cerastoderma edule* (L.). Until Mars (1951) and Høpner Petersen (1958) divided these species on morphological characters they were considered synonymous (Table 5.1a). Kingston (1974) reports a reduced vigour in crosses of the two species when compared to homozygous fertilisations, however, Barnes (1980) reports successful hybridisation for these species in the laboratory, resulting in the production of fertile offspring. Clearly more work needs to be done to clarify the hybridisation potential of this species. These data, however, do indicate a very close relationship between these putative species. In the field the two species rarely occur in the same habitat (Barnes 1980) (Table 5.1b). Where the two species are found to coexist, it is usually in tidal, estuarine habitats. For example Boyden & Russell (1972) surveyed 43 British and Irish sites where *C. glaucum* was known from and found *C. edule* at only 6 of those sites, all of which were tidal in nature.

**Figure 5.1a. The lagoon cockle *Cerastoderma glaucum***

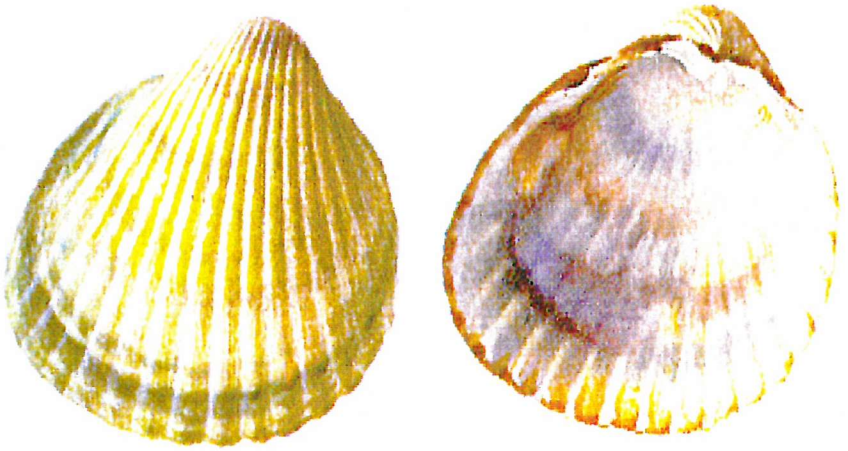


**Table 5.1a. Characters used to separate between *C. glaucum* and *C. edule* after Boyden (1971); extracted from Høpener-Petersen (1958)**

| Character | *C. glaucum* | *C. edule* |
|---|---|---|
| Shell shape | Posteriorly elongated | Almost oval |
| Ligament length | Short | Long |
| Periostracum | Extensive (often covering 75% of shell surface) | Poorly developed |
| Ribs on middle shell | Sharply pointed | Flattened |
| Posterior valve junction | Straight | Crenulate |
| Shell breadth (width) | Broad | Narrow |
| Digestive gland colouration | Always black | Variable - pale brown to dark green |
| Internal shell pigmentation | Pigment spot is absent in some individuals | Brown pigment spot always present at shell posterior |

**Table 5.1b. Habitat preference of *Cerastoderma glaucum* and *Cerastoderma edule* in Britain. After Boyden and Russell (1972)**

| **Species** / **Preference** | *C. edule* | *C. glaucum* |
|---|---|---|
| Habitat | Estuary | Lagoon |
| Temperature (°C) | 3-20 | 0-25 |
| Salinity (°/oo) | 15-35 | 5-40 |
| pH | 7.5-8.5 | 7.7 -10.1 |
| Oxygen (% saturation) | 90-105 | 0-200 |
| Substratum | Mud/sand | Mud/sand/shingle/vegetation |
| Exposure | Sheltered | Extremely sheltered |
| Tidal range (M) | 1.5-10 | 0-3.0 |
| Exposure to air (% time) | 0-50 | 0-5 |
| Habit | Buried | Buried/surface/on vegetation |

Another important distinguishing factor is the ability of *C. edule* to air breathe, which *C. glaucum* appears unable to do (Boyden 1972). This physiological inability clearly limits *C. glaucum* to areas of habitat that do not become uncovered by the water for any length of time. The instability of lagoon temperatures, which may rise above 30°C in the summer months, seems to be a factor which may play a part in excluding *C. edule* from the lagoon environment (Russell 1972). Furthermore, *C. edule* is less tolerant of the low salinities which may be characteristic of lagoons in the winter months.

Russell (1972) concluded that neither $O_2$ tension or pH were causes for habitat choice in either *C. glaucum* or *C. edule*, but that wave action which exerts an influence by crushing settling larvae was a major determining factor. Seemingly, *C. glaucum* favours quieter, non-tidal areas such as are found in lagoons and *C. edule* is better adapted to tidal conditions and the heavy wave action associated with the coastal habitat. This may control their distribution to some extent throughout Europe (Russell 1972; Barnes 1980). It is most likely that a combination of factors affect the distribution of these two species. For example, any physical factor such as $O_2$ tension or pH will be a stress to either species when outside their normal range and to cope with stress a species will normally produce stress proteins or alter gene expression. Table 5.1b shows clearly that the two species have very different abilities to cope with this kind of physiological adversity.

*Cerastoderma glaucum* has been suggested as an indicator species for lagoons (Bamber *et al.* 2000). As the lagoon habitat varies over time, changes in the abundance of this species may be used to interpret the biological impact of these variations. *C. glaucum* can be considered a good candidate as an indicator species as it is widely distributed and relatively long lived (Bamber *et al.* 2000).

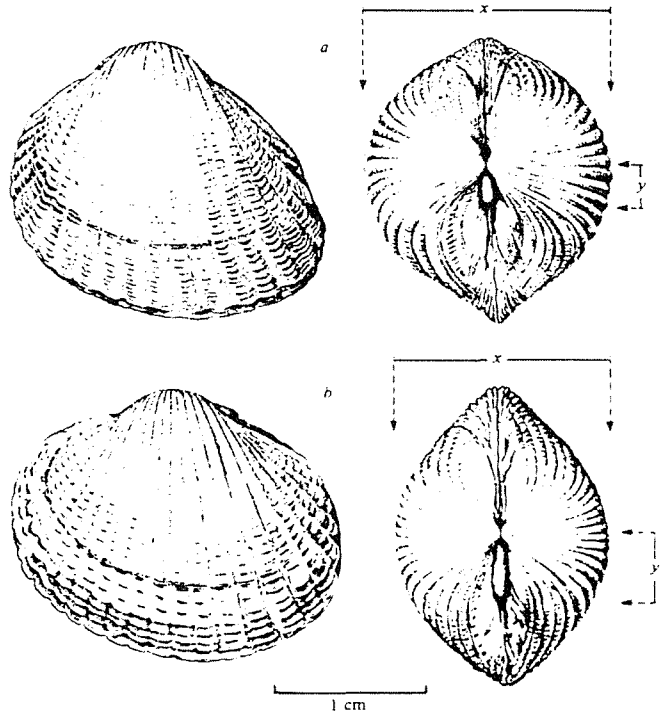
### *(i) Ecology and Life History*

#### *(a) Distribution*

*C. glaucum* is found in the Baltic, Black, Red, Mediterranean and Caspian Seas with its northern limit the Atlantic coast of Norway (Hummel *et al.* 1994). Its distribution in these areas includes tidal sites, adjacent lakes, estuaries, brackish lagoons and lagoon-like habitats (Hummel *et al.* 1994). Where this species has been identified in sympatric populations with *C. edule* confusion regarding identification of each species may arise, as individuals of intermediate morphology are often found (Brock 1979; Barnes 1980). There is some speculation as to whether these intermediate forms are hybrids of the species or a result of convergent evolution (Barnes 1980). Hybridisation is the most likely especially where barriers to reproduction in these incipient species are not fully in

place. The suggestion of convergent evolution is doubtful as it implies a process that has taken place over a long time which would likely produce a third or fourth species. There is no evidence to support this latter hypothesis, so it must remain conjectural. Fig. 5.1b shows typical forms of the two species.

**Figure 5.1b. A comparison between the typical morphologies of the lagoon cockle, *Cerastoderma glaucum* (a) and the marine cockle, *Cerastoderma edule (b)*. The ratio y:x is considered a reliable means of identification. After Barnes (1980)**




### *(b) Habitat*

It appears that *C. glaucum* is able to tolerate a wide range of environmental conditions (Brock 1980b). One factor, however, that may be important in limiting the distribution of *C. glaucum* is the suitability of the substrate (Brock 1979 & Brock 1980b). Brock (1979) reports the lagoon cockle as shunning areas of loose sediment where it may be easily disturbed by tidal currents. This would explain its preference for lagoons around the UK coast, some of which are characterised with muddy/sandy bottoms and little or no tidal movement.

Although typically a benthic species, *Cerastoderma glaucum* is often found associated, usually in low/moderate numbers, with macrophytes in the lagoon habitat (Al-Suwailem 1991). This is a distinct habitat choice, which is rarely seen, in the marine cockle. This climbing behaviour has been interpreted as a way of increasing the access of the cockle to its required food source, plankton, by placing it in a better position in the water column (Brock 1979). It also serves to move it away from anoxic sediment and the benthic boundary layer. The habitat of lagoons is often characterised by dense stands of vegetation, offering reduced feeding potential for a suspension feeder such as *Cerastoderma glaucum* (Barnes 1980).

### *(c) Life History*

The life history and reproductive cycle of *Cerastoderma glaucum* appears to vary depending on its locality. For example northern European populations are characterised by a monocyclic reproductive pattern (Boyden 1971). Mediterranean populations, however, may spawn up to three times in the same year (Ivell 1979a). More recently, however, there has been some evidence for a second spawning period in populations found in northern regions, occurring around October (Al-Suwailem 1991). This pattern of reproduction, influenced by locality, is also seen in other bivalves (Boyden 1971).

Spawning usually occurs between May and July in northern localities (Høpner Petersen 1958; Boyden 1971; Kingston 1974a). The sheltered nature of many lagoons means that spawning usually occurs early on in this period, whereas, those populations of *C. glaucum* found in estuaries may not spawn until later, usually July (Boyden 1971, Barnes 1980). This is because of the need for a relatively high temperature to initiate spawning in *C. glaucum*. Several spawning records for *C. glaucum* show good agreement for the early spawning of lagoonal populations (Kingston 1974a; Ivell 1979b; Al-Suwailem 1991). Adults in these populations are mature by April, ready for spawning in May and by July only spent individuals may be found (Al-Suwailem 1991).

It is thought that gametogenesis in *C. glaucum* begins in September/October, and by the start of the winter the number of immature individuals is still high (Kingston 1974a).

As the season progresses, however, the number of individuals with mature gonads increases and a large percentage of the population possess mature gametes by late March early April (Kingston 1974a; Al-Suwailem 1991). The usual age for onset of spawning is two years. It is not uncommon, however, for the males of *C. glaucum* to reach sexual maturity earlier than the females, with some spawning in their first year (Al-Suwailem 1991). A major factor, which distinguishes the *Cerastoderma* from other lamellibranchs (e.g. *Mercenaria mercenaria*; *Macoma balthica*), is the pattern of gametogenetic development. Other lamellibranchs demonstrate a major period of redevelopment immediately after spawning with the gametes developing through the autumn and maintained over winter, going into a second major developmental phase in spring (Boyden 1971). *Cerastoderma* are more similar to a second group of bivalves e.g. *Mytilus* spp and *Ostrea* spp, where gametogenesis does not occur directly after spawning, but is replaced by a period of laying down fat and glycogen stores (Boyden 1971).

Following spawning planktonic larvae are produced and within one week metamorphosis occurs (Barnes 1980). Maturity is often reached at a young age and thus small size, a phenomenon common in lagoon species relative to their marine counterparts (Barnes 1980). There is much debate regarding the growth rate of *C. glaucum* which seems to relate to the environment from which it is sampled (Al-Suwailem 1991). Kingston (1974b) reports that the larvae of *C. glaucum* are able to withstand and grow in a variety of salinities and temperatures such as are common to the lagoon environment. The normal life span for a lagoon cockle is 3-4 years, however, some individuals have been recorded living for up to six years (Boyden 1971; Ivell 1979b). However, in some very isolated lagoons, e.g. Holkham Salts Hole, Norfolk, this cockle has been observed at ages exceeding this (R. Bamber pers. comm.).

In lagoons *C. glaucum* is most likely found on the submerged macrophytes, *Zostera* and *Ruppia*, both common to the lagoon environment (Barnes 1980). *Zostera marina* is well reported in the literature for its association with both macro- and microfauna in temperate regions (e.g. Marsh 1973; Kikuchi 1980). Furthermore, much has been published concerning the colonisation and larval settlement of fauna associated with

macroalgae, particularly *Macrocystis* spp. and *Cladophora* spp (e.g. Bishop & Bishop 1973; Preston & Moore 1988). Very little has been published regarding the relationship between lagoon fauna and alga, such as *Chaetomorpha linum*. Al-Suwailem (1991), however, suggests an association between this algae and 0 group and newly settled spat of *C. glaucum*. In addition, it is not uncommon to see adults of the species resting in mats of *C. linum* (C. Pearson pers. obs.). This association may result in a possible mode of colonisation into new lagoons for this species if *C. linum* were to disperse between lagoons by passive means. At the present time, however, no evidence for this hypothesis exists.

One final interesting aspect of the life history, is the skewed sex ratio, which has on occasion been recorded for this species. Boyden (1971) assumed equal numbers of males and females for populations of the lagoon cockle, however, Al-Suwailem (1991) reports a ratio of 0.77:1 for *C. glaucum* populations in Gilkicker lagoon, Hampshire. Furthermore, Kingston (1974a) reported significantly more male than female individuals in four populations of *C. glaucum* from the southeast coast of England. This may have important implications for the genetic structuring of such populations as it may cause a reduction in the effective population size and so increase genetic drift (Sukumar 1991).

***(d) Diet***

*C. glaucum*, like many bivalves is a suspension feeder, feeding on small suspended macrophyte detrital particles broken down by primary and secondary feeders. The availability of food for this species of cockle seems to influence the timing and even the synchrony of spawning (Høpner Peterson 1958) even more than temperature and salinity (Kingston 1974). Synchronised spawning and the availability of suitable food may play an important role in the ability of this species to disperse to new sites. It would appear that most *C. glaucum* larvae enter the plankton after the spring, when food availability has fallen, with the result that the larvae are small and live for a short time in the water column compared to other bivalves such as *C. edule* in which larvae live for up to 5 weeks (Boyden 1971). There is evidence to suggest that unlike most

lamellibranchs, including *C. edule*, *C. glaucum* does show some degree of synchrony in spawning times (Boyden 1971). Synchronised spawning occurring at times of low food availability may result in populations of this cockle having a reduced pelagic phase. This may be an adaptation to retain larva in the environment to which they are best adapted, i.e. their natal lagoon. Some marine species which have to rely on a limited food source and/or live in a comparatively rare habitat do show limited pelagic stages for their larvae. For example, under 'comfortable' laboratory conditions the larvae of the nudibranch *Adalaria proxima* survives for longer compared to observations of its survival in the field (Todd *et al.* 1991).

### *5.1.2 Population genetics of* Cerastoderma glaucum

Palumbi (1994) suggests the accumulation of genetic differences between populations of the same species, may be related to biological causes such as life history patterns and habitat choice or may result from historical barriers to gene flow e.g. geographic and hydrographic factors. The population genetics of species inhabiting lagoons in the UK is little understood. Many of these lagoons are completely isolated from the sea with inflow of sea water via percolation only, leaving little or no potential for dispersal of organisms and hence reducing gene flow (Barnes 1988). Understanding the patterns of gene flow between lagoons for a species such as *Cerastoderma glaucum*, may go some way to helping conserve this and other lagoon organisms. The need for such understanding has occurred directly as a result of a desire to preserve the rare and rapidly diminishing lagoon habitat and its associated species.

Our ability to investigate the genetic population structure of a species with a view to conserving it, has been made increasingly easier by the introduction of new molecular techniques such as microsatellite analysis (Bruford *et al.* 1996) (and see Chp 3). The present study seeks to determine the genetic variability and population structure of populations of *Cerastoderma glaucum* from lagoons around the UK, using microsatellite primer pairs developed specifically for this species.

## 5.2 Methods

### *5.2.1 Microsatellite Development*

(For detailed methodology see Chp 3)

An enriched library technique was used to isolate and characterise microsatellites in the genome of *Cerastoderma glaucum* using methods adapted from Kandpal *et al.* (1994). This library construction was performed by a supervised summer student. Primers were then designed from flanking sequences for 9 loci using Primer 3 (Rozen & Skaletsky 1998). Optimisation of primers was performed, with adjustment to variables such as primer annealing temperature and $MgCl_2$ concentration. In spite of optimisation two primers failed to amplify the target region and one produced multiple amplification products, so these were discarded. Of the original 9 primer pairs, therefore, only 6 were used to screen the populations of *C. glaucum*.

### *5.2.2 Materials and Methods*

Genetic diversity within lagoons was analysed for *Cerastoderma glaucum* populations; two isolated populations on the south coast and three isolated populations on the east coast (Fig. 5.2a). Microsatellite alleles were amplified from 6 loci for 310 lagoon cockles.

#### *(i) Sample collection*

Whole animal samples were collected from lagoons between October 2000 and April 2001 (Table 5.2a). Between 60 and 100 individuals were collected at each lagoon. Where possible the maximum number of cockles were collected, but at sites where individuals appeared relatively scarce (e.g. Gilkicker) fewer were sampled. Samples were collected and placed directly into a saturated solution of 7% magnesium chloride (J. Preston pers. comm.). This relaxed the molluscs causing them to open their shells so that their soft bodies were exposed. Once relaxed, the cockles were placed into 95% ethanol.

**Figure 5.2a. Map depicting sample collection localities around the UK. The codes represent; C1 Fort Gilkicker lagoon, C2 Cockle Pond lagoon, C3 Reedland Marshes at Dunwich, C4 Holkham Salts Hole lagoon and C5 Abrahams Bosom lagoon**

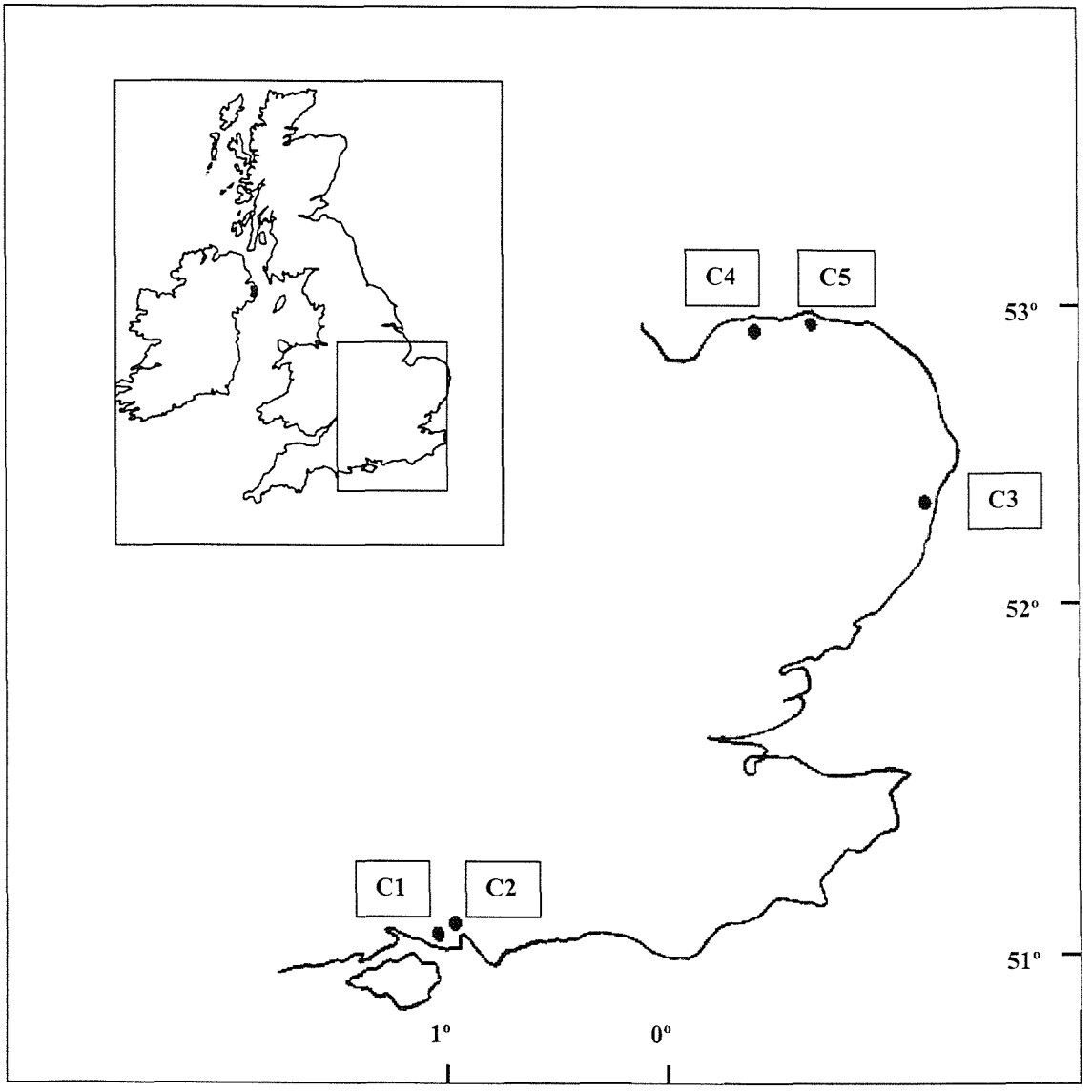


**Table 5.2a. Sampling localities and numbers of individuals of *Cerastoderma glaucum* analysed for 6 microsatellite loci**

| Name of site | Code | UK Ordnance Survey Map Grid Reference | No. of individuals DNA extracted |
|---|---|---|---|
| Fort Gilkicker Lagoon | C1 | SZ 608 978 | 54 |
| Cockle Pond | C2 | SZ 617 998 | 64 |
| Reedland Marshes | C3 | TM 483 071 to TM 483 726 | 64 |
| Holkham Salts Hole | C4 | TF 886 451 | 64 |
| Abrahams Bosom | C5 | TF 912 452 | 64 |

### *(ii) DNA extraction*

The adductor muscles and a piece of foot muscle measuring approximately $10mm^3$ were removed from each cockle and placed into a labelled tube containing 95% ethanol. A $1mm^3$ piece of muscle tissue was taken from each tube for the extraction procedure, using a clean blade for each sample. DNA extractions were performed using the CTAB method (Chp 2) for all samples, from all sites, except Gilkicker. 30 individuals from this site were extracted using the DNAeasy kits (Qiagen) (Chp 2). All extractions were quantified (Chp 2) and diluted to a final concentration of 5ng/$\mu$l using sterilised $dH_2O$.

### *(iii) PCR*

Optimisation reactions were performed in a 10$\mu$l volume containing either 1x Qiagen or Perkin Elmer PCR buffer (Chp 2), varying concentrations of $MgCl_2$ (Table 5.2b), 200$\mu$M each dNTP, 0.5 $\mu$M each primer,1 unit of *Taq* DNA polymerase (Table 5.2b), 5-10ng template DNA and sterilised chemical $H_2O$ (Sigma). The thermocyclers used in the reactions were either a Hybaid PCR-Express or a MWG-Biotech 'Primus 96 Plus'.

Optimised cycling parameters were as follows:

initial denaturing step at 96°C for 5 min;
3 cycles of 1 min at 95°C, 1 min at primer annealing temp, 1 min at 72°C;
25 cycles of 30 s at 95°C, 30 s at primer annealing temp, 45 s at 72°C;
final elongation cycle of 15 min at 72°C (to help ensure d-ATP tailing).

PCR conditions for amplification of DNA used in the genotyping of the *Cerastoderma glaucum* populations were based on these parameters and on varying concentrations of reagents mentioned above (Table 5.2b). In addition one of the primers in the reaction was substituted with a fluor-labelled primer in the same concentration (Table 5.2b). PCR reactions for genotyping were conducted using 0.2 ml 96 well covered microtitre plates.

**Table 5.2b. PCR conditions used to amplify and genotype 6 microsatellite loci in five populations of lagoon cockle *Cerastoderma glaucum***

| Locus (Dye) | | | Annealing temp (°C) | [$MgCl_2$] (mM) | HotStar *Taq** |
|---|---|---|---|---|---|
| 6Fam | Hex | Ned | | | |
| Cg 1 | | | 58 | 1 | yes |
| Cg 2 | | | 64 | 1 | yes |
| | Cg 4 | | 56 | 1.5 | no |
| | Cg 7 | | 55 | 2 | yes |
| | | Cg 9 | 55 | 2 | yes |
| | | Cg 11 | 57 | 1 | yes |

* HotStar Taq (Qiagen) where this was not used Perkin Elmer reagents and *Taq* used

### *(iv) Genotyping*

Microsatellite alleles were visualised on an ABI 377 automated sequencer, using 5' fluor-labelled primers and multiplex electrophoresis. PCR products were diluted between 2-5 times before screening. The fluor-label used for each microsatellite locus can be seen in Table 5.2b. The primer pair sequences used to amplify each locus are published for the first time in these works (Table 5.2c). The number of individuals genotyped from each sample site was 64, except for Fort Gilkicker lagoon where only 54 individuals were sampled. 64 individuals were sampled as this gave a sufficient sample size for estimating population genotype frequencies and screening could be achieved on one electrophoresis gel using a 64 lane paper comb.

**Table 5.2c. Primer pairs used to amplify 6 microsatellite loci in the lagoon cockle *Cerastoderma glaucum***

| Locus | Primer Sequence (5'-3') | Microsatellite |
|---|---|---|
| **Cg1** | F: TTTAATCCCTAATGAAGATTGA<br>R: TTTGTCCACCCACTGACC* | $(GT)_8(CG)_4GTCGGC(GT)_6$ |
| **Cg2** | F: CCATTGTCACGTGGGTCTTC<br>R: AGCGAGGCAGTCACACACTC* | $(GT)_{36}$ |
| **Cg4** | F: GTGTTGGACTCGCCATACC*<br>R: GACACAAGTAAAAACAATGTCT | $(GT)_{13}$ |
| **Cg7** | F: GATCCAGCCGTTCAAGTCC*<br>R: CGAAATAATGCGCGATGC | $(GT)_{13}$ |
| **Cg9** | F: CCATATTACCACTGCCCACAC<br>R: TGACCCCCTCCAGTGATTC* | $(GACA)_5GGAA(GACA)_3$ |
| **Cg11** | F: GGGGCGATTCTGGAGTAGTAG<br>R: GTCAAACCAGGCGCTAAGTC* | $(C)_{11}(CA)_7$ |

* Indicates primer with fluor-label

### *(v) Analyses*

The number of alleles per locus and the frequency of each allele in each population were estimated using GENEPOP version 3.1, (updated from Raymond & Rousset 1995). The programme was also used to perform global tests for linkage disequilibrium between all possible pairs of microsatellite loci. An unbiased estimate of the *P*-value, Fishers Exact test, was performed by means of a Markov chain (slightly different from that described by Guo & Thompson in 1989 in an unpublished report (Technical report No. 187, Department of Statistics, University of Washington, Seattle, USA)), using 100 000 steps.

ARLEQUIN version 2.00, (updated from Schneider *et al.* 2000) was used to estimate observed and expected heterozygosity for each locus according to methods in Guo & Thompson (1992). The test for fit to Hardy-Weinberg expectations is analogous to Fisher's Exact test using a Markov chain with 100 000 steps in the chain. The programme was also used for analysis of molecular variance (AMOVA) (Excoffier *et al.* 1992) to detect significant genetic differences between the groups, between

populations within the groups and for significant genetic differences within populations regardless of grouping.

The software FSTAT version 2.9.3.2, (updated from Goudet 1995) was used to obtain *F*-statistics according to Weir & Cockerham (1984). Significant differences in gene frequencies from those expected under Hardy-Weinberg equilibrium within each population were tested through $F_{IS}$ values. $F_{IS}$ values for each locus were tested for significant departures from zero by permuting alleles among individuals within samples using the statistic *smallf*. Testing over all loci was performed using Fisher's procedure (Sokal & Rohlf 1981) again using the statistic *smallf*. Levels of genetic divergence at microsatellite loci were investigated by computing $F_{ST}$. These values were tested for significant departures from zero using 5000 permutations of genotypes among samples. This test generates contingency tables of alleles by samples and these are classified by the log-likelihood statistic *G* (Goudet *et al.* 1996). Once again Fisher's procedure was used to obtain overall significance for allele frequency changes of the five sampling sites.

Finally FSTAT was used to estimate $F_{ST}$ per pair of samples. For each pair of samples multi-loci genotypes are randomised between the two samples and the overall loci *G*-statistic is used to classify the tables. The table wide level of significance was set at 1%. Where individual population tests or individual locus tests were performed, Bonferroni correction procedures (Rice 1989) were applied; these reduce the likelihood of Type I errors.

## 5.3 Results

### *5.3.1 Microsatellite Loci*

*Polymorphism, genetic variability of microsatellites and linkage disequilibrium*

Six polymorphic microsatellite loci were amplified for all cockle samples (Table 5.3a). Nine new microsatellite loci had been isolated for the lagoon cockle, however optimisation was only possible for 6 loci. Of these Cg1, Cg2, Cg4, Cg7, Cg9 & Cg11 were found to be polymorphic across all populations considered (Table 5.3b, Table 5.3c). The number of alleles per locus ranged from eight to twenty-six (Table 5.3a). Expected heterozygosities for all populations over all loci ranged from 0.35104 to 0.88715 (Table 5.3d), with an average of 0.6491 (± 0.0285). Observed heterozygosity values ranged from 0.03333 to 0.66071. The average value was 0.47087 (± 0.0344). The mean number of alleles detected in each population ranged from 4.7 (± 0.9) to 11.0 (± 1.8) and the proportion of unique alleles in each population varied from 0.01 to 0.15 with the two southern sites showing a higher proportion (Table 5.3c).

The analysis revealed significant linkage disequilibria in all populations ($p< 0.05$). C5 showed significant disequilibria for 4 loci ($p< 0.05$) and this was the most shown by any population. Summarising the results over all populations gave a significant linkage disequilibrium for two loci ($p< 0.05$), these loci were Cg7 and Cg9. The overall significance disappeared when these two loci were omitted. These loci were therefore excluded from further analysis on genetic differentiation. No other loci across all populations were found to show significant linkage disequilibrium and were therefore considered as statistically independent.

**Table 5.3a. Number of alleles and allelic size ranges for the six microsatellite loci isolated for the lagoon cockle *Cerastoderma glaucum***

| Locus | Source | No. of alleles | Allelic size range (bp) |
|---|---|---|---|
| **Cg 1** | V. Renault* pers. comm. | 22 | 136 - 224 |
| **Cg 2** | " | 26 | 161 - 279 |
| **Cg 4** | " | 13 | 150 - 216 |
| **Cg 7** | " | 10 | 102 - 126 |
| **Cg 9** | " | 16 | 176 - 240 |
| **Cg 11** | " | 8 | 83 - 107 |

* Southampton University, Southampton, UK

**Table 5.3b. Observed allele frequencies at 6 microsatellite loci in five lagoon cockle (*Cerastoderma glaucum*) populations**

| Cg4 | | Allele size | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | n | **150** | **156** | **164** | **166** | **170** | **172** | **174** | **176** | **178** | **180** | **182** | **214** | **216** |
| **C1** | 53 | - | 0.009 | 0.009 | 0.623 | 0.019 | - | 0.047 | 0.189 | 0.066 | 0.028 | 0.009 | - | - |
| **C2** | 56 | - | - | - | 0.473 | - | - | 0.071 | 0.286 | 0.098 | 0.045 | - | 0.009 | 0.018 |
| **C3** | 64 | 0.156 | - | - | 0.656 | 0.016 | - | - | - | 0.172 | - | - | - | - |
| **C4** | 59 | - | - | - | 0.669 | - | - | - | - | 0.331 | - | - | - | - |
| **C5** | 61 | - | - | - | 0.779 | - | 0.008 | - | - | 0.213 | - | - | - | - |

| Cg7 | | Allele size | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | n | **102** | **104** | **112** | **114** | **116** | **118** | **120** | **122** | **124** | **126** |
| **C1** | 53 | 0.010 | 0.058 | 0.010 | 0.173 | 0.077 | 0.077 | 0.375 | 0.202 | 0.010 | 0.010 |
| **C2** | 55 | - | 0.018 | 0.018 | - | 0.409 | 0.045 | 0.327 | 0.136 | 0.045 | - |
| **C3** | 48 | - | - | - | 0.042 | 0.271 | 0.052 | - | 0.615 | 0.021 | - |
| **C4** | 61 | - | - | - | 0.475 | 0.090 | - | 0.336 | 0.090 | 0.008 | - |
| **C5** | 56 | - | - | - | 0.402 | 0.089 | - | 0.420 | 0.089 | - | - |

| Cg11 | | Allele size | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | n | **83** | **85** | **97** | **99** | **101** | **103** | **105** | **107** |
| **C1** | 41 | 0.11 | 0.049 | 0.061 | 0.305 | 0.402 | 0.049 | 0.024 | - |
| **C2** | 48 | 0.052 | 0.104 | 0.188 | 0.229 | 0.333 | 0.083 | - | 0.01 |
| **C3** | 55 | 0.055 | 0.482 | 0.091 | 0.055 | 0.318 | - | - | - |
| **C4** | 49 | - | - | - | 0.133 | 0.184 | 0.684 | - | - |
| **C5** | 30 | - | - | - | 0.15 | 0.25 | 0.6 | - | - |

**Table 5.3b. Continued**

| Cg1 | | Allele size | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | n | **136** | **138** | **146** | **148** | **152** | **154** | **156** | **162** | **170** | **178** | **180** | **184** | **186** | **188** | **190** | **192** | **194** | **196** | **198** | **200** | **202** | **224** |
| **C1** | 50 | 0.010 | - | - | 0.100 | 0.030 | 0.010 | 0.010 | 0.010 | - | 0.020 | - | 0.040 | 0.010 | 0.080 | 0.200 | 0.060 | 0.060 | 0.040 | 0.040 | 0.240 | 0.030 | 0.010 |
| **C2** | 30 | 0.050 | - | 0.033 | 0.100 | 0.017 | 0.017 | - | - | 0.017 | - | 0.017 | - | 0.017 | 0.117 | 0.233 | 0.033 | 0.067 | 0.083 | 0.050 | 0.150 | - | - |
| **C3** | 51 | 0.324 | - | - | - | - | 0.010 | - | - | - | - | - | 0.010 | - | - | 0.167 | 0.382 | - | - | - | 0.108 | - | - |
| **C4** | 47 | - | - | - | - | - | - | - | - | - | - | - | - | 0.021 | 0.117 | 0.787 | - | - | 0.011 | - | - | 0.064 | - |
| **C5** | 43 | 0.616 | 0.035 | - | - | - | - | - | - | - | - | - | - | 0.012 | 0.012 | 0.326 | - | - | - | - | - | - | - |

| Cg2 | | Allele size | | | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | n | **161** | **163** | **177** | **181** | **183** | **191** | **193** | **197** | **199** | **201** | **247** | **251** | **253** | **255** | **257** | **259** | **261** | **263** | **265** | **267** | **269** | **271** | **273** | **275** | **277** | **279** |
| **C1** | 53 | - | - | - | 0.009 | 0.009 | - | - | 0.009 | - | - | - | - | 0.009 | - | 0.019 | 0.066 | 0.245 | 0.123 | 0.104 | 0.17 | 0.085 | 0.057 | 0.057 | - | 0.019 | 0.019 |
| **C2** | 55 | 0.009 | 0.018 | 0.009 | - | - | - | - | - | - | - | 0.009 | 0.018 | 0.045 | - | 0.018 | 0.082 | 0.064 | 0.109 | 0.264 | 0.127 | 0.091 | 0.100 | 0.018 | 0.009 | 0.009 | - |
| **C3** | 48 | - | - | - | - | - | 0.021 | 0.021 | 0.052 | 0.302 | 0.031 | - | - | - | - | - | - | 0.021 | 0.177 | 0.292 | 0.052 | 0.031 | - | - | - | - | - |
| **C4** | 59 | - | - | - | - | - | 0.008 | - | - | - | - | - | - | - | 0.008 | - | 0.025 | 0.364 | 0.364 | 0.153 | 0.068 | 0.008 | - | - | - | - | - |
| **C5** | 43 | - | - | - | - | - | - | - | - | - | - | - | - | - | - | - | 0.012 | 0.023 | 0.453 | 0.291 | 0.186 | 0.035 | - | - | - | - | - |

| Cg9 | | Allele size | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | n | **176** | **196** | **200** | **202** | **208** | **220** | **222** | **224** | **226** | **228** | **230** | **232** | **234** | **236** | **238** | **240** |
| **C1** | 47 | - | - | 0.170 | - | - | 0.287 | - | - | 0.021 | 0.106 | - | 0.298 | 0.085 | - | 0.032 | - |
| **C2** | 48 | - | - | 0.031 | - | - | 0.417 | 0.021 | 0.031 | - | 0.281 | 0.010 | 0.115 | 0.010 | 0.063 | - | 0.021 |
| **C3** | 50 | 0.010 | 0.010 | 0.160 | 0.010 | - | 0.670 | - | - | 0.010 | - | - | 0.010 | 0.110 | - | 0.010 | - |
| **C4** | 47 | - | - | - | - | - | - | - | - | - | 0.202 | 0.043 | 0.585 | 0.170 | - | - | - |
| **C5** | 31 | - | - | - | - | 0.016 | - | - | 0.032 | - | 0.452 | 0.032 | 0.419 | 0.048 | - | - | - |

**Table 5.3c. Genetic variation detected by 6 polymorphic microsatellite loci in five UK lagoon populations of the lagoon cockle *Cerastoderma glaucum***

| Population | Mean sample size/locus (± SE) | Mean no. of alleles/locus (± SE) | Proportion of unique alleles |
|---|---|---|---|
| **C1** | 49.5 ±2.0 | 11.0 ±1.8 | 0.09 |
| **C2** | 48.7 ±4.0 | 10.5 ±1.8 | 0.15 |
| **C3** | 52.5 ±2.6 | 5.2 ±0.9 | 0.07 |
| **C4** | 53.7 ±2.7 | 5.7 ±0.9 | 0.01 |
| **C5** | 44.0 ±5.2 | 4.7 ±0.9 | 0.03 |

### *5.3.2 Hardy-Weinberg Equilibrium (HWE)*

*Heterozygosity*

Table 5.3c shows observed and expected heterozygosity for the 6 loci across all populations. Differences in observed versus expected heterozygosity were found for at least 2 loci for all populations considered ($p< 0.05$) and population C3 showed a significant difference for all loci. Population C4 showed significant departure from HWE at only 2 of the six loci considered ($p< 0.01$).

Two of the six loci (Cg2 and Cg11) showed significantly less observed than expected heterozygosity for all populations ($p<0.01$). All other significant deviations from HWE were also observed as a deficit in heterozygotes except for the two loci, Cg4 and Cg9, which showed a significant excess in observed heterozygosity ($p< 0.05$) in populations C3 and C5 respectively.

**Table 5.3d. *Cerastoderma glaucum* observed and expected heterozygosity all populations**

| Pop$^n$ | C1 | | | C2 | | | C3 | | |
|---|---|---|---|---|---|---|---|---|---|
| **locus** | **n** | ***H*exp** | ***H*obs** | **n** | ***H*exp** | ***H*obs** | **n** | ***H*exp** | ***H*obs** |
| **Cg1** | 50 | 0.87657 | 0.74 | 30 | 0.84915 | 0.63333* | 51 | 0.70452 | 0.43137** |
| **Cg2** | 53 | 0.88715 | 0.41509** | 55 | 0.87773 | 0.54545** | 47 | 0.82132 | 0.2766** |
| **Cg4** | 53 | 0.58562 | 0.67925 | 56 | 0.68951 | 0.66071* | 64 | 0.51919 | 0.65625* |
| **Cg7** | 52 | 0.78662 | 0.61538** | 55 | 0.7613 | 0.72727 | 48 | 0.58662 | 0.375** |
| **Cg9** | 47 | 0.75086 | 0.70213 | 48 | 0.73531 | 0.64583* | 50 | 0.51798 | 0.5** |
| **Cg11** | 41 | 0.75429 | 0.41463** | 48 | 0.81491 | 0.5** | 55 | 0.68874 | 0.4** |

*H*obs = observed heterozygosity; *H*exp = expected heterozygosity; * *H*obs & *H*exp = significantly different $p<0.05$; ** *H*obs & *H*exp = significantly different $p<0.01$.

**Table 5.3d. Continued**

| Pop[n] | C4 | | | C5 | | |
|---|---|---|---|---|---|---|
| **locus** | **n** | ***H*exp** | ***H*obs** | **n** | ***H*exp** | ***H*obs** |
| **Cg1** | 47 | 0.38252 | 0.3617 | 43 | 0.61067 | 0.27907** |
| **Cg2** | 59 | 0.65363 | 0.10169** | 43 | 0.72695 | 0.2093** |
| **Cg4** | 59 | 0.44633 | 0.45763 | 61 | 0.35104 | 0.32787 |
| **Cg7** | 61 | 0.55521 | 0.60656 | 56 | 0.51303 | 0.58929 |
| **Cg9** | 47 | 0.38961 | 0.40426 | 31 | 0.55526 | 0.6129* |
| **Cg11** | 49 | 0.48517 | 0.22449** | 30 | 0.59661 | 0.03333** |

*H*obs = observed heterozygosity; *H*exp = expected heterozygosity; * *H*obs & *H*exp = significantly different $p<0.05$; ** *H*obs & *H*exp = significantly different $p<0.01$.

### *5.3.3* F-*statistics*

*Genetic Structuring*

*F*-statistics were estimated according to Weir & Cockerham (1984). Where low genetic exchange occurs among units (such as might be expected where populations are isolated from one another) sub-structuring of populations can occur. $F_{IS}$ values were calculated to provide further information. Values of $F_{IS}$ are negative when more heterozygotes than expected are present in a population and positive when more homozygotes are present (Hartl & Clark 1997). $F_{IS}$ values per population over all loci are shown in Table 5.3e. All $F_{IS}$ values for all populations were positive and ranged from moderately high to high (0.270 - 0.602). All of the values calculated were significantly different from zero (after Bonferroni corrections). The overall within-population heterozygote deficiency $F_{IS}$ ($F_{IS}$ = 0.364) was highly significant ($p< 0.001$). The null hypothesis that the populations under investigation are in HWE was therefore rejected. Levels of genetic divergence ($F_{ST}$) at each of the microsatellite loci were tested by looking for significant departures from zero. The high and significant overall $F_{ST}$ value ($F_{ST}$ = 0.151) for all loci across all populations, suggests a high level of genetic structuring. Exact *G*-tests for each locus and overall loci were highly significant (Table 5.3f).

Analysis of genetic differentiation between populations showed a range of pairwise $F_{ST}$ values between 0.0104 (low differentiation) and 0.2595 (high genetic differentiation). All pairwise comparisons were found to be statistically significant (exact *G*-test; $p<0.005$) (Bonferroni corrected) except between population C1 (Gilkicker lagoon) and population C2 (Cockle Pond lagoon) (Table 5.3g).

**Table 5.3e. $F_{IS}$ values calculated according to Weir & Cockerham (1984) with FSTAT version 2.9.3.2 (Goudet 1995) for *Cerastoderma glaucum* populations from lagoons around the UK**

| Population | $F_{IS}$ | $p$ |
|---|---|---|
| C1 | 0.270 | 0.0001* |
| C2 | 0.304 | 0.0001* |
| C3 | 0.422 | 0.0001* |
| C4 | 0.602 | 0.0001* |
| C5 | 0.395 | 0.0001* |

*Values of statistical significance (Bonferroni corrected)

**Table 5.3f. $F_{ST}$ values calculated according to Weir & Cockerham (1984) with FSTAT version 2.9.3.2 (Goudet 1995) per locus and over all loci**

| Locus | $F_{ST}$ | $p$ |
|---|---|---|
| Cg1 | 0.216 | < 0.001* |
| Cg2 | 0.081 | < 0.001* |
| Cg4 | 0.078 | < 0.001* |
| Cg11 | 0.206 | < 0.001* |
| All loci | 0.151 | < 0.001* |

* *Exact G-test*

**Table 5.3g. $F_{ST}$ pairwise values calculated according to Weir & Cockerham (1984) with FSTAT version 2.9.3.2 (Goudet 1995) between all pairs of populations of *Cerastoderma glaucum* populations investigated in the UK**

| **Pop$^n$** | **C1** | **C2** | **C3** | **C4** | **C5** |
|---|---|---|---|---|---|
| **C1** | 0 | - | - | - | - |
| **C2** | 0.0104 $^{ns}$ | 0 | - | - | - |
| **C3** | 0.1113* | 0.0975* | 0 | - | - |
| **C4** | 0.1736* | 0.1776* | 0.2595* | 0 | - |
| **C5** | 0.1652* | 0.1529* | 0.1758* | 0.1516* | 0 |

* $p$ <0.005

### *5.3.4 AMOVA*

Partitioning of microsatellite allelic variance was performed using analysis of molecular variance (AMOVA) (Excoffier *et al.* 1992) based on *F*-statistics. AMOVA partitions total variance into three covariance components. The 'within sites' component quantifies variation among individuals in the samples. The remaining variance, divided

into 'among sites within groups' and 'among groups' components, can potentially give insight into relationships among sites by reflecting genetic diversity at different levels of a hierarchy. The covariance components are used to calculate fixation indices (*F*-statistics). The significance of the fixation indices is tested using a non-parametric permutation approach (Excoffier *et al.* 1992). Here hierarchical AMOVA was used to test for genetic structuring defined by the geographical regions (south or east coast) where the sampled lagoons occur. Significance was tested using 16 000 permutations. AMOVA revealed that the greatest percentage of the variance (88%) was explained by within-sub-population variation and the least (4%) by different sampling regions (Table 5.3g).

**Table 5.3h. Analysis of molecular variance (AMOVA) among groups (southern group C1, C2, eastern group C4, C5 & C6), among the populations within those groups and within all populations regardless of group. Total % distribution of the variance components and *p*-values for corresponding $\Phi$-tests are given**

| Source of variation | d.f. | Variance components | % of variation | *p* |
|---|---|---|---|---|
| **Among Groups** | 1 | 0.03 | 3.54 | ns |
| **Among populations within groups** | 3 | 0.09 | 8.61 | <0.001 |
| **Within populations** | 593 | 0.02 | 87.85 | <0.001 |

## 5.4 Discussion

In spite of the recent boom in the development of DNA markers for investigating the population structure of marine organisms (Nielson *et al.* 1999; Smith *et al.* 2001), previous work on *Cerastoderma glaucum* using protein markers has concentrated on understanding the species complex existing between *C. glaucum* and *C. edule* (Brock 1978). This study, therefore, represents the first known work on the population structure of the lagoon cockle *Cerastoderma glaucum*, using microsatellite DNA markers

The majority of marine invertebrates have dispersive larvae or propagules with the ability to survive in the open seas for extended periods of time, resulting in populations with a low degree of genetic structure (Waples 1998). Consequently, some authors regard the marine environment as open habitat with barriers to dispersal occurring infrequently, if at all (Palumbi 1994; Waples 1998). Hedgecock (1986), however, noted that habitat fragmentation can and does occur, and where such a division arises it may cause adaptive differentiation. For a species such as *Cerastoderma glaucum* with a clear habitat preference for sheltered, isolated lagoons, the potential for a high degree of genetic structure between the populations is considerable.

### *5.4.1 Evidence for low levels of dispersal and high levels of genetic structuring in populations of* C. glaucum *in the UK*

Ecological and physiological data indicated that populations of *C. glaucum* in lagoons around the UK may demonstrate low levels of dispersal and suggested a high degree of genetic structuring for these populations (Boyden 1972; Russell 1972; Brock 1979; Barnes 1980). Microsatellite analysis supported these findings, indicating significant genetic differentiation between populations and high levels of inbreeding. This differentiation is not translated to the regional level which suggests a pattern of random dispersal (probably by passive means) followed by isolation of founders which may have led to subsequent inbreeding.

### $F_{ST}$ *and pairwise* $F_{ST}$

$F_{ST}$ values were used to investigate genetic differentiation at separate loci and over all loci in each population of *C. glaucum*. The values obtained for each locus and over all loci were statistically significant. The pairwise $F_{ST}$ values obtained in this study have revealed statistically significant population sub-structuring for *C. glaucum* populations in the UK (Table 5.3g). It is important to note that these values may even be an underestimation of the degree of sub-structuring between these populations. When there are large numbers of multiple alleles per locus, multi-allele estimates of $F_{ST}$ can underestimate genetic sub-structuring and any barrier effect, and this can occur whether or not alleles are shared among populations under consideration (Hedrick 1999). Pairwise $F_{ST}$ values for the *C. glaucum* populations in this study ranged between 0.0104 and 0.1776. Although there are no other published works on the genetic differentiation revealed by microsatellite markers of lagoon molluscs, these values are within and exceed those reported for local differentiation of other marine molluscs by microsatellite markers (0.00192 - 0.0542; Rogers unpblished data). Previous works on genetic differentiation of *Cerastoderma* have focused on separating the two species (Jelnes *et al.* 1971; Brock 1978; Gosling 1980 & Brock 1987; Hummel *et al.* 1994). However, values obtained for genetic differentiation between populations within species in these studies (using allozymes) show less differentiation probably because of the low polymorphism associated with these markers. Interestingly a greater degree of genetic differentiation was found between populations of *C. glaucum* compared to *C. edule* (Hummel *et al.* 1994).

Other work using allozymes on marine bivalves demonstrating high gene flow, put average $F_{ST}$ values at between 0.01-0.03 for populations separated by hundreds of kilometres and 0.04-0.06 for those separated by thousands of kilometres (Skibinski *et al.* 1983; Dillon & Manzi 1992; Grant *et al.* 1992; Sarver *et al.* 1992; Saavedra *et al.* 1993). Even taking into account the low polymorphism of the molecular markers used in those studies, data presented here show a greater, and in all but one case, a significant differentiation, indicating low levels of gene flow (dispersal). Moreover, the data in the present study are for populations, which in some cases, are separated by just tens of kilometres.

*Amova*

The examination of genetic differentiation is one way that we are able to draw some conclusions about sub-structuring and gene flow between populations (Gerlach & Musolf 2000). The greater differentiation among nearby populations shown by the results of the AMOVA (Table 5.3h) and supported in the eastern region by the significant pairwise $F_{ST}$ value (Table 5.3g), indicates more differentiation among populations within a region than among regions, reflecting a local dynamic. This dynamic has affected the genetic variation of the local populations resulting in the observed low levels of heterozygosity (Table 5.3d).

### *5.4.2. Causes of genetic structuring*

*Founder effect*

The large $F_{ST}$ values may indicate it is likely that the populations investigated were founded from small numbers of individuals (Beaumont & Nichols 1996) probably as a result of low levels of passive dispersal. Where sexually reproducing marine species demonstrate limited dispersal capabilities they will effectively exist as a series of closed populations (Ayre *et al.* 1997). As a result of the combined effects of genetic drift (founder effect), mutation and site specific natural selection, populations may become highly differentiated (Ayre *et al.* 1997). Evidence that this has been occurring in populations of *C. glaucum* is characterised by the consistent deficits of heterozygotes in all populations studied. The lagoon habitat is an extreme and marginal one where certain types of genotype may be strongly selected. Sexual reproduction, recombination and gene flow (dispersal) all act to dilute the effect of favoured genotypes which may prevent successful adaptation to such habitats. *C. glaucum* shows a strong habitat preference for lagoons over the 'normal' marine environment (Brock 1979; Brock 1980b) and is able to outcompete marine counterparts. The life history of this species appears to show an important adaptive feature that may be acting to retain larvae in the natal lagoon. Spawning occurs when food availability is low leading to a significant reduction in the larval dispersal period (1 week) compared to the marine cockle (5 weeks). This is an adaptation seen in many other lagoon species over their marine counterparts (Barnes 1980).

### *5.4.3. Evidence for inbreeding*

*Low levels of heterozygosity*

The reduction in the dispersing ability of *C. glaucum* larvae has led to high levels of inbreeding in individual lagoons. Evidence for this is revealed by low levels of heterozygosity. With few exceptions, observed levels of heterozygosity for the populations of *C. glaucum* throughout the UK was significantly lower than expected ($p<$ 0.05; Table 5.3d) (not considering the non-independent loci Cg7 & Cg9). Populations C2 (Cockle Pond) and C5 (Abrahams Bosom) demonstrated a lower than expected heterozygosity for all loci considered (after removal of Cg7 & Cg9) (Table 5.3d). Population C5 for locus Cg11 showed a dramatic loss in heterozygosity ($H_{exp}$ = 0.596 and $H_{obs}$ = 0.0333 ; $p<$ 0.01). Although the sample size for this locus was low (n = 30) it does exceed that required to detect polymorphism (<0.05) at a diallelic locus with 95% confidence (Sjögren & Wyöni 1994). The level of variation detected for the same locus in the other four populations was an order of magnitude higher. Detection of so little genetic variation in a natural population at a hypervariable microsatellite locus is vary rare (Eldridge *et al.* 1999).

These low heterozygosity levels have also been demonstrated with allozyme markers in populations of *C. glaucum* in other parts of Europe (Hummel *et al.* 1994). Studies of allozymes (e.g. LAP) in marine bivalves commonly reveal heterozygote deficiencies but the reasons for this are not well understood (Berger 1983; Zouros & Mallet 1989). There is some evidence in the case of allozymes to suggest that different environmental conditions favour different genotypes (Koehn *et al.* 1980). Moreover, recent studies of the lagoon sea anemone *Nematostella vectensis* have used RAPD markers to reveal maintenance of a single genotype by asexual reproduction (Pearson *et al.* 2002). This genotype is thought to be the most successful, and best adapted to the extreme and marginal lagoon environment. Taxa that live in other extreme habitats such as hydrothermal vent species, have also been found to demonstrate significant heterozygote deficiencies (Creasy *et al.* 1996). Microsatellites have revealed heterozygote deficiencies in marine bivalves and in other molluscs, demonstrating a

degree of genetic structure for these species and stressing the importance of local populations (Huang *et al.* 2000; Casu *et al.* 2002)

Where heterozygote deficiency exists for multiple loci this could be because of pooling of several sub-populations (Wahlund effect), inbreeding or the effect of null alleles (Callen *et al.* 1993). The entire genome is affected by inbreeding and the Wahlund effect, whereas null alleles are locus specific (Lehmann *et al.* 1999). Evidence from significant linkage disequilibrium for these data suggest that null alleles are not the cause of the heterozygote deficit. Linkage disequilibrium is expected where inbreeding or the Wahlund effects are acting because members of different sub-populations have different probabilities to carry certain combination of alleles (Lehmann *et al.* 1999). Moreover, the strong linkage disequilibrium detected in at least 2 loci for all populations of *C. glaucum* studied are indicative of inbreeding maintaining combinations of successful genes (genotypes) in populations around the UK. Linkage disequilibrium is often identified in small founding populations where it has had insufficient time to decay through recombination (Hartl 2000) or where inbreeding has been occurring. Populations that have recently been subject to mixing of genotypes (for example where immigrants have entered the population and mated with residents) may also demonstrate significant linkage disequilibrium. Moreover, functionally important genes that are linked have been shown to be of major importance to the persistence of threatened species and microsatellites may be linked to such genes (Frankham *et al.* 2002). Selection may therefore play a major role in maintaining the linkage disequilibrium seen in populations of *C. glaucum*.

*Inbreeding*

It is expected that long term isolation of small populations will lead to inbreeding (Frankham 1998). The effective inbreeding coefficient $F_e$ calculated by Frankham (1998) for a number of island populations was averaged at 0.29 with 29% of populations obtaining inbreeding coefficients above 0.5. These values are calculated according to;

$F_e = 1\text{-}H_{IS} / H_M$

where $H_{IS}$ and $H_M$ are the heterozygosities for island and mainland populations respectively. In domestic species higher inbreeding coefficients have been associated with reduced fitness and higher rates of extinction (Frankel & Soulé 1981; Frankham & Ralls 1998). Demographically, this may be revealed as a reduction in fecundity and/or by a skewed sex ratio (Eldridge *et al.* 1999).

$F_{IS}$ *values*

While no direct comparison can be made with the data shown by Frankham (1998), as $F_e$ cannot be calculated. In this study $F_{IS}$ values were used as an estimation of inbreeding, that is an excess of homozygotes at the sub-population level (Balloux *et al.* 1998). It is interesting to note that all $F_{IS}$ values calculated for all the populations under investigation were positive and statistically different from zero ($p$= 0.0001). These figures indicate a level of inbreeding in all populations of *C. glaucum* in the UK. Work by Kingston (1974a) found significantly more males in four populations of *C. glaucum* on the south east coast of England and Al-Suwailem (1991) reported a ratio of 0.77:1 for *C. glaucum* populations in Gilkicker lagoon (site C1 in this study). This ratio may be further evidence supporting the presence of inbreeding in this particular population of the lagoon cockle. Boyden (1971) argues that where two populations of the same species are separated and subject to different environmental conditions spawning may occur at different times and thus allow reproductive isolation to occur between such populations. The conditions in any two lagoons can be very different at the same time of year and it is known that temperature plays an important role in determining the release of *C. glaucum* gametes (Boyden 1971; Barnes 1980) thus the timing of the release of larvae into the plankton may be a major contributing role to the isolation of these populations. It is clear, however, that more research is needed to understand whether dispersal is costly to *C. glaucum*, whether inbreeding is associated with any kind of fecundity depression and how important the synchrony of larval release is to the isolation of each population.

### *5.4.4. The role of dispersal in the genetic structuring of populations of* C. glaucum

Balloux *et al.* (1999) stress the importance that a lack of dispersal may have among breeding groups, with resulting genetic differentiation creating barriers to the loss of genetic variance at the sub-population level. Charlesworth & Charlesworth (1987) argue, however, that where isolation occurs for long periods of time populations are able to adapt to inbreeding by purging deleterious alleles. However, Hedrick (1994) suggests that some effects of inbreeding depression may result from the presence of weakly deleterious alleles (difficult to eliminate through selection) and not just a few highly deleterious alleles.

For all populations of the lagoon cockle considered in this study, high levels of dispersal are unlikely for two reasons. Firstly, the physical means for larvae to enter the surrounding seas from some of the lagoons can be limited, for example where connection to the sea is via percolation through ground water (C5; Abrahams Bosom and C4; Holkham Salts Hole lagoons ). This physical difficulty combined with the low survival rate of *C. glaucum* larvae reduces the chances of spat from one lagoon ever reaching and settling in a new lagoon. The presence of higher numbers of unique alleles in the southern populations of the lagoons (Table 5.3c) provides some evidence that dispersal into the open seas is limited/absent between the east and south coasts. It may also be suggestive of a different demographic history for the southern populations compared to those from the east coast (Hartl 2000). A higher number of alleles would be expected where populations were larger and more stable (Hartl 2000), and while the lagoon sites investigated on the south coast (C1 & C2) are slightly larger in size than two of those on the east coast (C4 & C5) the relative population densities appeared comparable (C. Pearson pers. obs.). Historically some factor may have acted to reduce the population sizes of those lagoons on the east coast. The potential for re-colonisation from other populations is limited for the east coast lagoons in this study because of the nature of the sea-lagoon connections, none of which involve an open connection to the sea (C. Pearson pers. obs). Lagoon sites C1 & C2, however, both have open connections with the adjacent sea and are relatively close to one another. Both of these lagoons are involved in human leisure activities, C1 is in the middle of a golf course and

C2 is used as a boating pond. There is less than 1 mile between these two lagoons by sea or overland, moreover, both lagoons are situated adjacent to roads and coastal footpaths. There is, therefore, a possibility of anthropogenic transport between the two lagoons. C1 and C2 lagoons had a low and non-significant pairwise $F_{ST}$ value (Table 5.3g) probably as a result of gene flow between them. The same result may arise, however, if they were founded by the same genotype (Pàlsson 2000).

### *5.4.5 Conservation implications for* C. glaucum

These data demonstrate and quantify for the first time the genetic effect that isolation in the lagoons is having on the populations of *Cerastoderma glaucum* around sites in the UK. It appears that populations of *Cerastoderma glaucum* may be under strong selection to maintain certain genotypes that are specifically adapted to the narrow niche offered by their particular lagoon habitat. Inbreeding combined with low dispersal are potential mechanisms by which this is occurring. There may, however, be other systems such as assortative mating or some form of chromosomal mechanism acting to maintain combinations of genes (genotypes). Where assortative mating is occurring there will be a positive correlation between the phenotypes of mating pairs; the implication being individuals choosing to mate with individuals more similar to themselves than with others less similar (Hartl 2000). If assortative mating is acting between individuals better adapted to the habitat, then progeny produced in this way are more likely to succeed in the lagoon environment than the progeny of outcrosses.

These data have implications for the conservation of individual lagoons and the species within them. If lagoon specialists have adaptive mechanisms to limit dispersal resulting in the maintenance of a particular successful genotype then translocation of individuals between lagoons may be a conservation catastrophe. This is because it may result in favourable genotypes being diluted, or lost, as a consequence of outcrossing, leading to possible extinction of the population. Additionally, there is evidence to suggest it would be judicious to tailor management strategies to protect populations with low genetic diversity as these may be susceptible to extinction, should conditions alter and if

tolerance limits were reached (Shick & Lamb 1977), especially as many of the observed phenomena may indicate small population size.

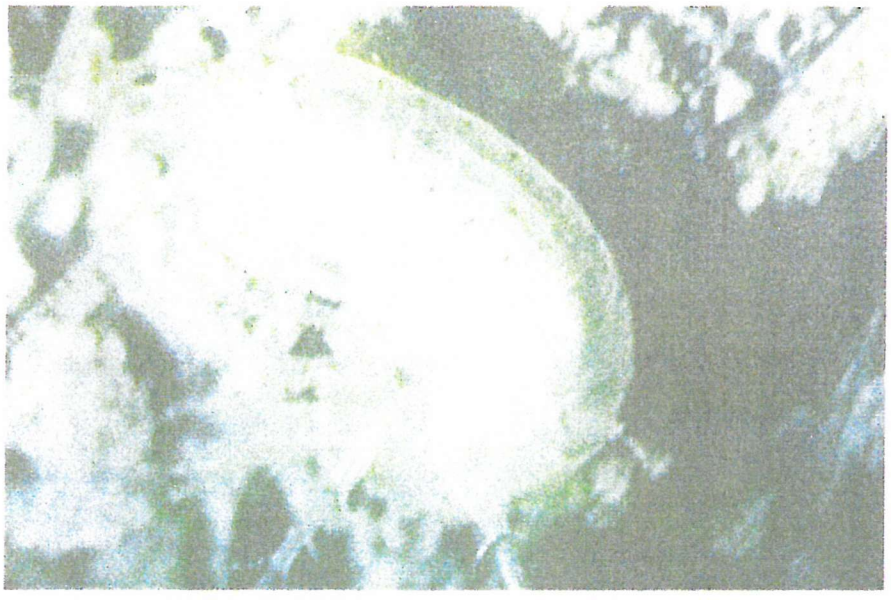
Much work has been conducted both theoretically (Halley & Manasse 1993; Mills & Smouse 1994) and experimentally (Newman & Pilson 1997; Saccheri *et al.* 1998) on the contributing role of genetic factors to extinction. Furthermore, Eldridge *et al.* (1999) believe that it would be imprudent to disregard the effects that genetic factors may have on the contribution to the extinction of small isolated populations. The heterozygosity deficiency and the levels of inbreeding reported here for populations of *C. glaucum* around the UK may indicate that genetic factors are having a negative influence. There is, however, strong supporting evidence that the population structure of *C. glaucum* revealed by microsatellite markers in this study is a result of adaptation of a few successful genotypes to a marginal habitat. This study therefore highlights the importance of considering ecological, physiological, and morphological data before interpreting any genetic data.

# 6.1 Introduction

## *6.1.1* Gammarus insensibilis

*Gammarus insensibilis* Stock (Fig. 6.1a) is considered a globally threatened/declining species (Downie 1996) and in the UK is protected under Schedule 5 of the Wildlife and Countryside Act of 1988. Although a number of studies have examined the ecology of *G. insensibilis* in Britain (Sheader & Sheader 1985b; Sheader 1996), no data exist on the genetic variability both within and between these populations.

**Figure 6.1a. The lagoon sand shrimp Gammarus insensibilis**



### *(i) Ecology and life history*

### *(a) Distribution*

World-wide *G. insensibilis* has been recorded from the North Sea to its southern limits in the Black and Mediterranean Seas, where it is known to occur in fully marine conditions (Sheader 1996). Typically this species occurs in sheltered brackish conditions where it is often misidentified as the fully marine *G. locusta*. There are,

however, several characteristics, which distinguish these species (see Sheader & Sheader 1985b for full review). Moreover, there appear to be reproductive barriers to hybridisation between these species, as fertilised eggs from crosses only pass through the initial stages of cleavage and then development is halted (Sheader & Sheader 1985b). In its southern range *G. insensibilis* is found in sublittoral habitats and in sheltered brackish sites down to 15 m. In the UK lagoons and saline ponds provide an ideal habitat for *G. insensibilis*. The English coast, although towards the northern limit for this species, shows a relatively wide geographic distribution of *G. insensibilis*. The present known range for this amphipod, includes 22 lagoon or lagoon like habitats, ranging from Dorset to Lincolnshire (Sheader pers. comm.).

### *(b) Habitat*

Lagoon sites, where *G. insensibilis* is found, have been shown to be characterised by high salinity, regular tidal input of sea water, small tidal range and low or absent fresh water input (excepting rain water). They also retain water, irrespective of tide or season (Sheader pers. comm.). These features, common to lagoons, encourage the growth of attached and free-floating plants (Sheader 1985). The physical features described, are essential to the success of *G. insensibilis* in lagoons. This has been demonstrated at sites where seawater input has been changed and *G. insensibilis* has disappeared (Sheader pers. comm.). Mats of *Chaetomorpha linum* are abundant in lagoons where this amphipod is found, indicating these two species have an association. In addition to *C. linum* this amphipod has been found associated with both *Ruppia* and *Zostera* (Sheader & Sheader 1985b).

Many sites where *G. insensibilis* is recorded also support characteristic weed associated fauna; isopods e.g. *Idotea chelipes*; amphipods e.g. *Corophium insidiosum* ; chironomids e.g. *Chironomous salinarius* ; bivalves e.g. *Cerastoderma glaucum*; gastropods e.g. *Hydrobia ventrosa* and anthozoans e.g. *Nematostella vectensis* (Sheader & Sheader 1985b). However, the abundance and relative importance of each of these species varies between sites and with the season (Sheader & Sheader 1985b).

Al-Suwailem (1991) reports an abundance of *G. insensibilis* associated with the lagoonal macrophytes, compared to samples taken from the sediment alone. Furthermore, in his study, Al-Suwailem (1991) found crustaceans to be the dominant taxa associated with macrophytes, accounting for greater than 78% of the total number of individuals.

### *(c) Life history*

Much work has been published on the reproductive biology of shallow water and sublittoral gammaridean amphipods (e.g. Nelson 1980; Sainte-Marie 1991; Steele & Steele 1991). The genus shows diverse life history strategies, which may vary with latitude, salinity, and depth (Sainte-Marie 1991). Diversification of life history, e.g. variation in reproductive parameters, has allowed for adaptation into a wide range of environments, including the marginal habitat of lagoons (Sheader 1996). In addition to spatial differences between the species some seasonal variations are also observed, and although these are far greater between species, some variation is observed within species (Sheader 1996).

*G. insensibilis*, in common with other gammarid amphipods, breeds throughout much of the year, showing a marked decline in reproduction throughout the winter months (Sheader & Sheader 1985b). At some sites e.g. Gilkicker lagoon, reproductive activity is reported throughout the year maintaining a peak in the summer months (Sheader 1996). In lagoons, environmental conditions alter not just on a seasonal basis but can also change daily, this is particularly true of small lagoons. Mature *G. insensibilis* females may produce a succession of broods throughout the course of a year and thus within a lagoon environment, they and their broods will be subjected to a wide range of environmental fluctuations and in turn changes in energy budgets (Sheader 1996). It would be of some benefit to organisms living in such habitats if their reproductive strategy were to reflect such potential short term changes

Sheader (1983) describes the life cycle of the brackish/freshwater amphipod *G. duebeni* as consisting of two major overlapping generations, and in addition, suggests that this species is able to alter its breeding strategy in response to seasonal changes. These

adaptations are extremely useful for an organism living in an unpredictable environment and it is reasonable to suppose that *G. insensibilis* in the lagoon environment, would benefit from such strategies. Upon maturity, *G. insensibilis* females produce eggs at each moult, however in the absence of males, females will delay the moult for up to 2 days, furthermore they will reabsorb oocytes if mating does not occur (Sheader 1996). This behaviour provides some evidence of the ability of this species to alter its reproductive strategy according to local conditions. Fertilised eggs, held in their mother's brood pouch, take between 11-15 days (at 20°C) to develop (Sutcliffe 1992).

During the summer months, growth of juveniles is rapid, with some females ready to breed within four or five months (Karakiri & Nicolaidou 1987). This is not uncommon for amphipod species and it is likely that temperature, food supply and predation are all important factors (Sheader 1983). When cultured under laboratory conditions *G. insensibilis* has been shown to grow at a rate of 2.3 mm per month indicating that maturity can be reached 3 months after release from the marsupial pouch (Al-Suwailem 1991). These data, however, should be interpreted with caution as field observations have shown differences in the brooding size of eggs for winter and summer months (Steele & Steele 1975a; Sheader 1983).

This seasonal variation in egg size is another feature distinguishing *G. insensibilis* from *G. locusta*, along with a much shorter precopula period (less than 10 hrs for *G. insensibilis*) (Sheader 1985). The seasonal variation observed in the eggs of *G. insensibilis* is a result of the amount of yolk laid down in the egg, with winter eggs containing a greater quantity of yolk and producing a larger sized juvenile at hatching (Sheader 1996). Eggs are deposited in the marsupium just prior to a moult so any factor delaying or bringing on an early moult will affect the size of the egg. Temperature has been shown to control the intermoult period, in both field and laboratory studies, so at lower temperatures (e.g. winter temperatures), where the intermoult period is extended, Stage I eggs achieve their size early leaving plenty of time for yolk deposition which increases the size of the egg (Sheader 1996). Large juveniles are thought to have a greater chance of survival than small juveniles (Skadsheim 1984; Rabalais & Gore

1985), thus the production of large juveniles in the colder months has implications for the survival of populations as a whole. Populations of short lived gammarids, such as *G. insensibilis*, are at low abundance in the winter and early spring and rely on low mortality of winter eggs/juveniles to boost the population in the spring and summer months (Sheader 1996). Temperature alone does not appear to be the only parameter that controls the intermoult period, as alterations to the length of the intermoult period have been observed in females in order to achieve synchronous release of juveniles from the marsupium (Sheader 1996). The implications of this adaptive behaviour are discussed below.

Young of *Gammarus insensibilis* develop directly, thus this amphipod relies on adult migration to colonise new sites. The amphipod *G. duebeni* has an observed summer rhythm of peak mating and egg production at the time of spring tides and it is thought this may facilitate the dispersal of juveniles released prior to this time (Sheader 1983). This rhythm of egg production is known for a range of littoral and shallow water species of amphipoda (Fish 1975; Sheader 1978;) and Sheader (1996) observed *G. insensibilis* females synchronising egg production in the summer months. Moreover, there is some evidence to suggest that the timing of the release of juveniles from the marsupium is dependent on local conditions (Steele & Steele 1975b). Population synchrony could be an important factor in determining the ability of some populations to disperse into the open seas and thus colonise new lagoons. There is some evidence to suggest an association between the juveniles of *G. locusta* and floating driftweed as a means for dispersal (Tully & Céidigh 1986). Although no data exist to support this as a method of dispersal for *G. insensibilis*, there are abundant data suggesting a relationship between lagoon/marine flora and this amphipod (Sheader & Sheader 1985b). It would therefore, not be unreasonable to suggest this as a means of dispersal between lagoons.

Where peak mating and peak egg production occur at the time of spring tides certain advantages may be conferred such as a reduced stress environment for newly laid eggs (following the release of the juveniles), owing to the recent arrival of the spring tide and the facilitation of adult movement through the habitat, aiding mate location (Sheader 1983). At temperatures greater than 19.5°C *G. duebeni* females are capable of

producing a brood for each spring tide, as the intermoult period at such temperatures is extremely short (Sheader 1983). It is reasonable to assume *G. insensibilis* females could be capable of this as they are able to alter the intermoult period, and if this were the case dispersal of high numbers of juveniles from some sites could be quite possible.

*G. duebeni* shows a different life cycle dependent on local conditions (Hynes 1954; Steele & Steele 1969; Sheader 1983), particularly with respect to the number of generations produced per year. Sheader (1983) believes that for *G. duebeni* there is probably a link between generations produced and food availability and temperature; as in the Test Estuary at the southern limit of the species, 2 generations per year were observed as opposed to one recorded for other more northern populations. All the characters observed in *G. duebeni* which allowed for more than one generation per year have been observed in *G. insensibilis*, for example females maturing rapidly and at a small body size (Sheader 1996). Evidence for such adaptations mean it is entirely reasonable to suppose that *G. insensibilis*, in those populations where breeding occurs throughout much of the year, is capable of producing more than one generation per year.

*G. insensibilis* has been described as an annual species (Karakiri & Nicolaidou 1987). There are, however, behavioural tendencies in this species which may cause the death of many individuals before completion of their annual cycle . During the summer months, on hot days, both adults and juveniles of this species have been observed to swarm in very large numbers to the shore where death can occur rapidly by desiccation and anoxia (C. Pearson pers. obs.). This is a behaviour which is likely to exert strong selection on this species.

***(d) Diet***

*Chaetomorpha linum* forms a major component of the diet of this amphipod (Sheader & Sheader 1985, Bamber *et al.* 1992). Furthermore, *G. insensibilis* is able to assimilate this food resource at a much more efficient rate than its marine counterpart, *G. locusta* (Sheader & Sheader 1985). This amphipod grazes directly on macrophytes and along with *Idotea chelipes* has a relatively significant impact on lagoon macrophytes (Al-

Suwailem 1991). *Gammarus insensibilis* is, however, regarded as omnivorous and grazes on macrophytes according to the availability of other, preferable food sources (including epiphytic microflora and macrophyte detritus) and also according to the age of the animal (Sainte-Marie 1991). This author describes *G. insensibilis* has having a broad feeding niche which is advantageous in a habitat subject to a high degree of temporal and spatial variability.

## 6.2 Methods

### *6.2.1 Microsatellite Development*

(For detailed methodology see Chp 3)

An enriched library technique was used to isolate and characterise microsatellites in the genome of *Gammarus insensibilis*. A genomic library enriched for (CA/GT) microsatellites was then constructed using methods adapted from Kandpal *et al.* (1994). Two rounds of enrichment were performed; from the first 95 recombinant clones were screened and 8 resulting positive clones were cycle sequenced, from the second 102 recombinant clones were screened and 15 resulting positive clones were cycle sequenced. Primers were designed from the flanking sequences for 1 locus from the first round and for 9 loci from the second round using Primer 3 (Rozen & Skaletsky 1998).

Optimisation of primers was then performed, with adjustment of variables such as primer annealing temperature and $MgCl_2$ concentration. In spite of optimisation some primers failed to amplify the target region and gave non-specific amplification, so were discarded. Of the original 10 primer pairs only 8 were used to screen the populations of *G. insensibilis*.

### *6.2.2 Materials and Methods*

#### *(i) Sample collection*

Whole animal samples were collected from lagoons between April 2001 and May 2001 (Table 6.2a; Fig. 6.2a). Between 43 and 70 individuals were collected at each lagoon. Where possible the maximum number of *Gammarus insensibilis* were collected, but at sites where individuals appeared relatively scarce (e.g. Salthouse Broad) fewer were sampled. Samples were collected and placed directly into 95% ethanol.

**Figure 6.2a. Map depicting sample collection localities around the UK. The codes represent; G1 Fort Gilkicker lagoon, G2 The Salterns lagoon, G3 Salthouse Broad lagoon and G4 Humberston Fitties lagoon.**

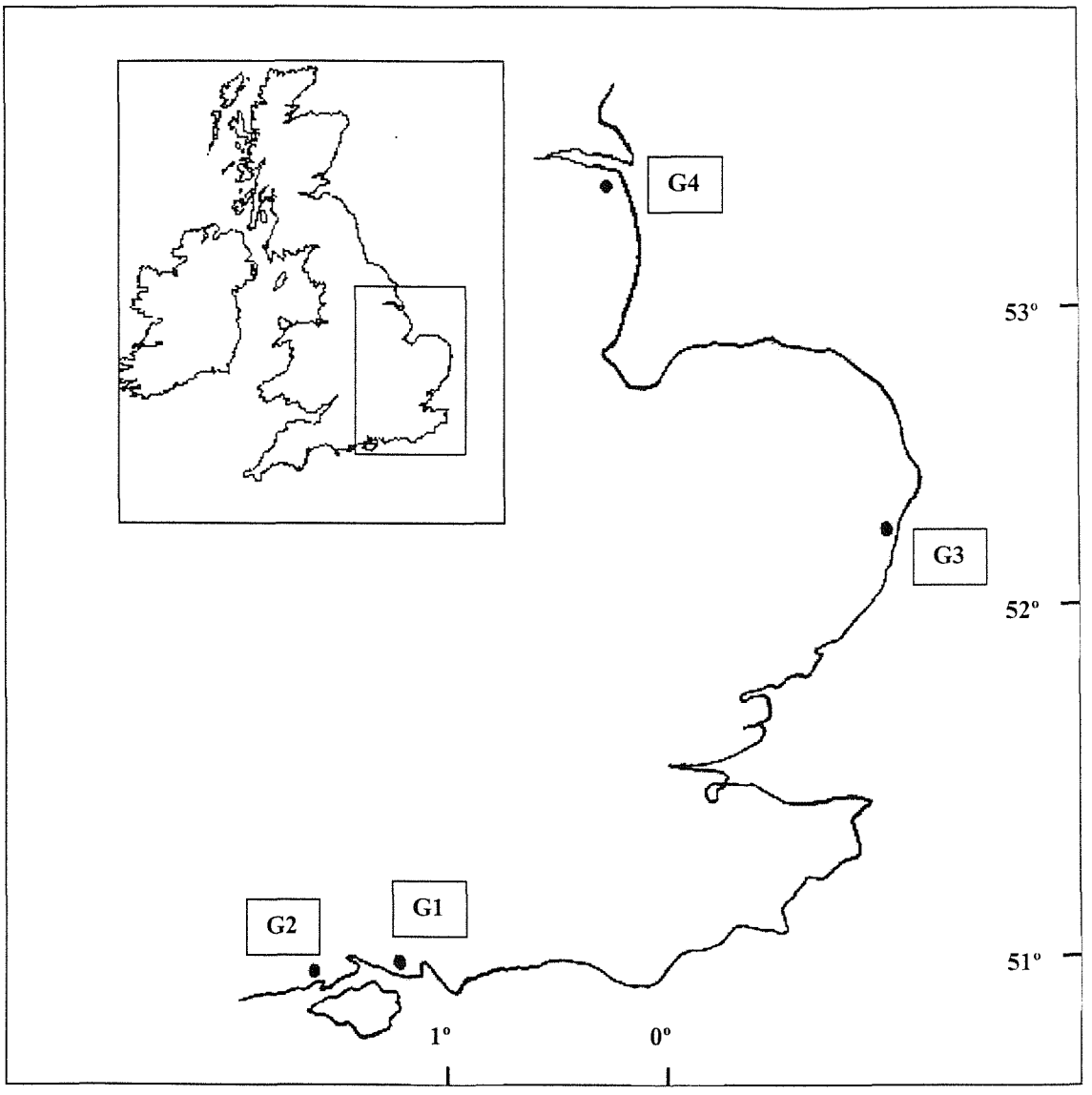


**Table 6.2a. Sampling localities and numbers of individuals of *Gammarus insensibilis* analysed for 8 microsatellite loci**

| Name of site | Code | UK Ordnance Survey Map Grid Reference | No. of individuals DNA extracted |
|---|---|---|---|
| Fort Gilkicker Lagoon | G1 | SZ 608 978 | 60 |
| The Salterns | G2 | SZ 328 935 | 56 |
| Salthouse Broad | G3 | TG 068 446 | 39 |
| Humberston Fitties | G4 | TA 336 048 | 60 |

### *(ii) DNA extraction*

To avoid contamination from either gut contents or from possible brooded young a leg was removed from each animal for the purpose of DNA extraction. The extractions for this study were performed using the PCI extraction protocol (Chp 2). Following extraction all samples were diluted to a final concentration of 5 ng/$\mu$l using sterile $dH_2O$.

### *(iii) PCR*

Optimisation reactions were performed in a 10 $\mu$l volume containing either 1x Qiagen or Perkin Elmer PCR buffer (Chp 2), 1 x Qiagen 'Q' solution (Table 6.2b), varying concentrations of $MgCl_2$ (Table 6.2b), 200$\mu$M each dNTP, 0.5 $\mu$M each primer,1 unit of *Taq* DNA polymerase (Table 6.2b), 5-10 ng template DNA and sterilised chemical $H_2O$ (Sigma). The thermocyclers used in the reactions were either Hybaid PCR-Express or MWG-Biotech 'Primus 96 Plus'.

Optimised cycling parameters were as follows:

initial denaturing step at 96°C for 5 min;
3 cycles of 1 min at 95°C, 1 min at primer annealing temp, 1 min at 72°C;
25 cycles of 30 s at 95°C, 30 s at primer annealing temp, 45 s at 72°C;
final elongation cycle of 15 min at 72°C (to help ensure d-ATP tailing).

PCR conditions for amplification of DNA used in the genotyping of the *Gammarus insensibilis* populations were based on these parameters and on varying concentrations of reagents mentioned above (Table 6.2b). In addition one of the primers in the reaction was substituted with a fluor-labelled primer in the same concentration (Table 6.2b). PCR reactions for genotyping were conducted using 0.2 ml 96 well microtitre plates.

**Table 6.2b. PCR conditions used to amplify and genotype 6 microsatellite loci in five populations of lagoon sand shrimp *Gammarus insensibilis*.**

| Locus (Dye) | | | Annealing temp (°C) | [$MgCl_2$] (mM) | 'Q' Solu$^n$ | Hot star *Taq** |
|---|---|---|---|---|---|---|
| 6Fam | Hex | Ned | | | | |
| Gi 1 | | | 55 | 2 | 2x | no |
| | Gi 2 | | 55 | 2 | 2x | no |
| | | Gi 3 | 60 | 2 | - | no |
| Gi 4 | | | 60 | 2 | - | no |
| | Gi 5 | | 55 | 2 | 2x | no |
| | | Gi 6 | 55 | 2 | 2x | no |
| Gi 8 | | | 55 | 2 | - | yes |
| | Gi 9 | | 55 | 2 | - | yes |

### *(iv) Genotyping*

Microsatellite alleles were visualised on an ABI 377 automated sequencer, using 5' fluor-labelled primers and multiplex electrophoresis. PCR products were diluted between 2-5 times before screening. The fluor-label used for each microsatellite locus can be seen in Table 6.2b. The primer pair sequence used to amplify each locus are published for the first time in these works (Table 6.2c). The number of individuals genotyped from each sample site was 60, except for The Salterns lagoon where only 56 individuals were sampled and Salthouse Broad lagoon where just 39 individuals were genotyped.

**Table 6.2c. Primer pairs used to amplify 8 microsatellite loci in the lagoon sand shrimp *Gammarus insensibilis*.**

| Locus | Primer Sequence (5'-3') | Microsatellite |
|---|---|---|
| **Gi1** | F: CTTCGGGGCAAAACAAAG*<br>R: TGAGCGTCCTGAGCCTATC | $(CA)_5CG(CA)_4$ |
| **Gi2** | F: GGATCGTTGTGCATGGAAC*<br>R: CAAAGTCCCAGTTGCCTTTC | $(GA)_5(CA)_3$ |
| **Gi3** | F: TTATCTCCCCCGTTCTTCTC*<br>R: AGCAGTCCTTGTGTCCTTGC | $(GT)_7$ |
| **Gi4** | F: GGCTACTATGCACCGCTCTC*<br>R: AGGTCAAGGATATGCAACGG | $TCA(TC)_2ATCCTCGTCC(TC)_3$ |
| **Gi5** | F: TCATCCGAAGCCATCTACG*<br>R: CCCATTAAGGAGACGACGAC | $(GCT)_4(ACT)_4GCTGTCTCC(GCT)_4$ |
| **Gi6** | F: GGACAGGTCAGGGAACAAAG*<br>R: GTGGTGCATGGAACTCAAAG | $(TC)_5$ |
| **Gi8** | F:CATAACGGTGGTAAAACCCC*<br>R:AACAAACACGGACTCAATGC | $(TAT)_2T(CAA)_3$ |
| **Gi9** | F:CACTTTTGTTGAGGGGAACC<br>R:TTCGGTCAGTCACTCCGTC* | $(CA)_8CTGCT(GT)_5$ |

* Indicates primer with fluor-label

### *(v) Analyses*

The number of alleles per locus and the frequency of each allele in each population were estimated using GENEPOP version 3.1, (updated from Raymond & Rousset 1995). The programme was also used to perform global tests for linkage disequilibrium between all possible pairs of microsatellite loci. An unbiased estimate of the *p*-value, Fishers Exact test, was performed by means of a Markov chain (slightly different from that described by Guo & Thompson in 1989 in an unpublished report (Technical report No. 187, Department of Statistics, University of Washington, Seattle, USA)), using 100 000 steps.

ARLEQUIN version 2.00, (updated from Schneider *et al.* 2000) was used to estimate observed and expected heterozygosity for each locus according to methods in Guo & Thompson (1992), the test for fit to Hardy-Weinberg expectations is analogous to

Fisher's Exact test using a Markov chain with 100 000 steps in the chain. The programme was also used for analysis of molecular variance (AMOVA) (Excoffier *et al.* 1992) to detect significant genetic differences between the groups, between populations within the groups and for significant genetic differences within populations regardless of grouping.

The software FSTAT version 2.9.3.2, (updated from Goudet 1995) was used to obtain *F*-statistics according to Weir & Cockerham (1984). These were performed as for *Cerastoderma glaucum* (Chp 5).

## 6.3 Results

### *6.3.1 Microsatellite Loci*

*Polymorphism, genetic variability of microsatellites and linkage disequilibrium*

Ten microsatellite loci were isolated from the enriched library, however, of these it was possible to optimise only eight. These eight microsatellite loci were amplified for all *Gammarus insensibilis* samples (Table 6.3a). All these microsatellites were found to be polymorphic across all populations considered (Table 6.3b). The number of alleles per locus ranged from eight to eighteen (Table 6.3a). Expected heterozygosities for all populations over all loci ranged from 0.3027 to 0.81495 (Table 6.3d), with an average of 0.58371 (± 0.0173). Observed heterozygosity values ranged from 0.46552 to 1.0. The average value was 0.80985 (± 0.033). The mean number of alleles detected in each population ranged from 3.8 (± 0.7) to 7.9 (±2.0) and the proportion of unique alleles in each population ranged from 0.03 to 0.27 (Table 6.3c).

The analysis revealed significant linkage disequilibria in all populations, ($p< 0.05$). G2 showed significant disequilibria for four loci ($p< 0.05$) and this was the most shown by any population. Summarising the results over all populations gave a significant linkage disequilibrium for two loci, these loci were Gi4 and Gi5. The overall significance disappeared when these two loci were omitted. These loci were therefore excluded from analysis of genetic structure of populations. No other loci across all populations were found to show significant linkage disequilibrium and were therefore considered as statistically independent.

**Table 6.3a. Number of alleles and allelic size ranges for the eight microsatellite loci isolated for the lagoon sand shrimp *Gammarus insensibilis*.**

| Locus | No. of alleles | Allelic size range (bp) |
|---|---|---|
| **Gi 1** | 9 | 140 - 158 |
| **Gi 2** | 18 | 94 - 174 |
| **Gi 3** | 8 | 152 -198 |
| **Gi 4** | 6 | 91 - 157 |
| **Gi 5** | 18 | 266 - 336 |
| **Gi 6** | 11 | 129 - 175 |
| **Gi 8** | 15 | 140 - 210 |
| **Gi 9** | 12 | 98 - 138 |

**Figure 6.3b. Observed allele frequencies at 8 microsatellite loci in four lagoon sand shrimp *Gammarus insensibilis* populations**

| Gi1 | | Allele size | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | n | **140** | **142** | **144** | **146** | **148** | **150** | **152** | **154** | **158** |
| **G1** | 54 | - | - | - | - | 0.46 | 0.04 | 0.50 | - | - |
| **G2** | 51 | - | - | - | - | 0.45 | 0.04 | 0.51 | - | - |
| **G3** | 39 | 0.03 | 0.03 | 0.03 | 0.04 | 0.38 | 0.01 | 0.49 | - | - |
| **G4** | 58 | 0.01 | - | 0.02 | 0.02 | 0.43 | 0.01 | 0.50 | 0.01 | 0.01 |

| Gi 2 | | Allele size | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | n | **94** | **100** | **110** | **116** | **120** | **122** | **126** | **128** | **130** | **134** | **138** | **140** | **152** | **156** | **160** | **162** | **164** | **174** |
| **G1** | 54 | - | - | - | - | 0.03 | 0.03 | 0.02 | 0.47 | 0.03 | - | - | - | - | - | 0.02 | 0.35 | 0.06 | - |
| **G2** | 51 | - | - | - | - | - | - | - | 0.08 | 0.58 | - | - | - | - | - | - | 0.31 | 0.03 | - |
| **G3** | 39 | - | - | - | - | - | - | - | - | 0.5 | - | - | - | - | - | - | 0.41 | 0.09 | - |
| **G4** | 35 | 0.01 | 0.01 | 0.01 | 0.01 | - | - | - | 0.07 | 0.36 | 0.01 | 0.03 | 0.01 | 0.01 | 0.01 | - | 0.31 | 0.09 | 0.03 |

| Gi 3 | | Allele size | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | n | **152** | **154** | **156** | **158** | **192** | **194** | **196** | **198** |
| **G1** | 36 | - | 0.08 | 0.15 | 0.03 | 0.08 | 0.33 | 0.17 | 0.15 |
| **G2** | 43 | - | - | - | - | 0.02 | 0.43 | 0.07 | 0.48 |
| **G3** | 38 | 0.11 | - | 0.37 | - | - | 0.53 | - | - |
| **G4** | 58 | - | - | 0.23 | - | - | 0.77 | - | - |

| Gi 4 | | Allele size | | | | | |
|---|---|---|---|---|---|---|---|
| | n | **91** | **93** | **97** | **99** | **101** | **157** |
| **G1** | 49 | - | - | 0.03 | 0.29 | 0.68 | - |
| **G2** | 51 | - | - | - | - | 0.56 | 0.44 |
| **G3** | 38 | 0.46 | 0.03 | - | 0.20 | 0.32 | - |
| **G4** | 49 | 0.18 | 0.05 | - | 0.07 | 0.69 | - |

**Figure 6.3b. Continued**

| Gi 5 | | Allele size | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | n | **266** | **272** | **274** | **276** | **278** | **280** | **282** | **284** | **288** | **294** | **296** | **298** | **302** | **304** | **308** | **312** | **328** | **336** |
| **G1** | 53 | - | 0.50 | - | - | 0.50 | - | - | - | - | - | - | - | - | - | - | - | - | - |
| **G2** | 50 | - | 0.50 | - | - | 0.50 | - | - | - | - | - | - | - | - | - | - | - | - | - |
| **G3** | 37 | - | - | 0.03 | 0.46 | - | 0.01 | 0.47 | - | - | 0.01 | - | - | 0.01 | - | - | - | - | - |
| **G4** | 59 | 0.01 | 0.03 | 0.02 | 0.42 | 0.03 | 0.01 | 0.42 | 0.01 | 0.01 | - | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 |

| Gi 6 | | Allele size | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | n | **129** | **131** | **133** | **135** | **137** | **139** | **141** | **169** | **171** | **173** | **175** |
| **G1** | 54 | 0.03 | 0.05 | 0.01 | - | - | - | 0.62 | - | - | 0.30 | - |
| **G2** | 50 | 0.02 | - | 0.02 | 0.03 | - | 0.54 | - | - | 0.02 | 0.37 | - |
| **G3** | 39 | - | - | - | 0.06 | 0.46 | 0.01 | 0.01 | 0.31 | 0.03 | 0.05 | 0.06 |
| **G4** | 60 | - | - | - | - | - | 0.49 | 0.01 | - | - | 0.50 | - |

| Gi 8 | | Allele size | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | n | **140** | **142** | **144** | **148** | **150** | **154** | **156** | **160** | **164** | **168** | **176** | **178** | **184** | **200** | **210** |
| **G1** | 32 | - | - | - | - | - | - | 0.31 | - | 0.66 | - | - | - | - | 0.02 | 0.02 |
| **G2** | 40 | - | - | - | - | - | - | 0.24 | - | 0.76 | - | - | - | - | - | - |
| **G3** | 37 | 0.05 | 0.01 | - | 0.01 | - | 0.01 | 0.39 | 0.01 | 0.45 | 0.03 | 0.01 | 0.01 | - | - | - |
| **G4** | 58 | 0.03 | 0.03 | 0.01 | 0.01 | 0.01 | 0.02 | 0.36 | - | 0.50 | 0.01 | - | 0.01 | 0.01 | - | - |

| Gi 9 | | Allele size | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | n | **98** | **106** | **120** | **122** | **124** | **126** | **128** | **130** | **132** | **134** | **136** | **138** |
| **G1** | 41 | - | 0.01 | 0.02 | 0.01 | 0.37 | 0.12 | 0.41 | - | 0.04 | 0.01 | - | - |
| **G2** | 36 | 0.01 | - | - | - | 0.60 | 0.14 | 0.10 | - | - | 0.01 | 0.03 | 0.11 |
| **G3** | 38 | - | - | - | - | - | 0.17 | 0.33 | - | - | - | - | 0.50 |
| **G4** | 59 | - | - | - | - | - | 0.02 | 0.50 | 0.02 | - | - | - | 0.47 |

**Table 6.3c. Genetic variation detected by 8 polymorphic microsatellite loci in four UK lagoon populations of the lagoon sand shrimp *Gammarus insensibilis*.**

| Population | Mean sample size/locus (± SE) | Mean no. of alleles/locus (± SE) | Proportion of unique alleles |
|---|---|---|---|
| **G1** | 46.6 ±3.2 | 5.0 ±0.9 | 0.14 |
| **G2** | 46.5 ±2.1 | 3.8 ±0.7 | 0.03 |
| **G3** | 37.9 ±0.4 | 5.4 ±0.9 | 0.08 |
| **G4** | 53.1 ±3.6 | 7.9 ±2.0 | 0.27 |

### *6.3.2 Hardy-Weinberg Equilibrium (HWE)*

Table 6.3d shows observed and expected heterozygosity for the 8 loci across all populations. Differences in observed versus expected heterozygosity were found for at least 7 loci for all populations considered ($p< 0.05$) and populations G3 & G4 showed a significant difference for all loci considered. Locus Gi3 showed significantly less

observed heterozygosity than expected in two of the four populations and Gi9 in 1 of the four populations. However, all other loci in all populations showed a greater observed than expected heterozygosity and this was statistically significant ($p< 0.05$) for all but one locus (Gi4 in population G1).

Locus Gi5 showed a statistically significant observed heterozygosity of 1 ($p< 0.01$) in all populations. The only other loci to demonstrate such an excess in heterozygosity were Gi2 & Gi9 in population G3 and Gi6 in population G4.

**Table 6.3d. *Gammarus insensibilis* observed and expected heterozygosity in all populations**

| Pop[n] | G1 | | | G2 | | |
|---|---|---|---|---|---|---|
| locus | n | *H*exp | *H*obs | n | *H*exp | *H*obs |
| **Gi 1** | 54 | 0.53981 | 0.94444** | 51 | 0.54106 | 0.90196** |
| **Gi 2** | 54 | 0.65386 | 0.85185** | 51 | 0.58649 | 0.66667** |
| **Gi 3** | 36 | 0.81495 | 0.52778** | 43 | 0.60027 | 0.53488** |
| **Gi 4** | 49 | 0.49632 | 0.63265 | 51 | 0.49796 | 0.88235** |
| **Gi 5** | 53 | 0.50476 | 1** | 50 | 0.50505 | 1** |
| **Gi 6** | 54 | 0.52925 | 0.75926** | 50 | 0.56505 | 0.78** |
| **Gi 8** | 32 | 0.47867 | 0.65625** | 40 | 0.36677 | 0.475** |
| **Gi 9** | 41 | 0.69798 | 0.58537** | 36 | 0.65649 | 0.58333 |

*H*obs = observed heterozygosity; *H*exp = expected heterozygosity; ** *H*obs & *H*exp = significantly different $P<0.01$.

**Table 6.3d. Continued**

| Pop[n] | G3 | | | G4 | | |
|---|---|---|---|---|---|---|
| locus | n | *H*exp | *H*obs | n | *H*exp | *H*obs |
| **Gi 1** | 39 | 0.61905 | 0.89744** | 58 | 0.56822 | 0.93103** |
| **Gi 2** | 39 | 0.58908 | 1** | 35 | 0.77805 | 0.97143** |
| **Gi 3** | 38 | 0.58386 | 0.94737** | 58 | 0.36027 | 0.46552* |
| **Gi 4** | 38 | 0.60211 | 0.97368** | 49 | 0.6659 | 0.46939** |
| **Gi 5** | 37 | 0.57164 | 1** | 59 | 0.65826 | 1** |
| **Gi 6** | 39 | 0.68765 | 0.89744** | 60 | 0.51247 | 1** |
| **Gi 8** | 37 | 0.65272 | 0.83784** | 58 | 0.63013 | 0.81034** |
| **Gi 9** | 38 | 0.62772 | 1** | 59 | 0.53687 | 0.9322** |

*H*obs = observed heterozygosity; *H*exp = expected heterozygosity; * *H*obs & *H*exp = significantly different $P<0.05$; ** *H*obs & *H*exp = significantly different $P<0.01$.

### *6.3.3* F-*statistics*

*Genetic Structuring*

*F*-statistics were estimated according to Weir & Cockerham (1984). First $F_{IS}$ values were calculated to provide information about possible inbreeding within the populations under investigation. $F_{IS}$ values per population over all loci are shown in Table 6.3e. All

$F_{IS}$ values for all populations were negative (this occurs when more heterozygotes than expected are present in a population) and ranged from moderately low to extremely low (-0.171 to -0.529). All calculated values were significantly different from zero (after Bonferroni corrections). The overall within-population inbreeding coefficient $F_{IS}$ ($F_{IS}$ = -0.373) was highly significant ($p$<0.001). Leading to a rejection of the null hypothesis that the populations are in HWE for the loci under investigation. Levels of genetic divergence ($F_{ST}$) at each of the microsatellite loci were tested by looking for significant departures from zero. The high and significant overall $F_{ST}$ value ($F_{ST}$ = 0.174) for all loci across all populations, suggests a high level of genetic structuring. Exact *G*-tests for each locus, except Gi8, and overall loci were significant (Table 6.3f).

Analysis of genetic differentiation between populations showed a range of pairwise $F_{ST}$ values between 0.1173 (moderate differentiation) and 0.2073 (high genetic differentiation) (Table 6.3g). All pairwise comparisons were found to be statistically significant (exact G-test; p<0.005) (Bonferroni corrected).

**Table 6.3e. $F_{IS}$ values calculated according to Weir & Cockerham (1984) with FSTAT version 2.9.3.2 (Goudet 1995) for *Gammarus insensibilis* populations from lagoons around the UK**

| Population | $F_{IS}$ | $p$ |
|---|---|---|
| G1 | -0.171 | 0.0004* |
| G2 | -0.212 | 0.0004* |
| G3 | -0.504 | 0.0004* |
| G4 | -0.529 | 0.0004* |

*Values of statistical significance (Bonferroni corrected)

**Table 6.3f. $F_{ST}$ values calculated according to Weir & Cockerham (1984) with FSTAT version 2.9.3.2 (Goudet 1995) per locus and over all loci**

| Locus | $F_{ST}$ | $p$ |
|---|---|---|
| Gi1 | 0.000 | <0.05* |
| Gi2 | 0.159 | <0.05* |
| Gi3 | 0.180 | <0.05* |
| Gi6 | 0.339 | <0.05* |
| Gi8 | 0.037 | ns |
| Gi9 | 0.227 | <0.05* |

*Values of statistical significance (Exact *G* -test) $p$ <0.05

**Table 6.3g.** $F_{ST}$ **pairwise values calculated according to Weir and Cockerham (1984) with FSTAT version 2.9.3.2 (Goudet) between all pairs of populations of *Gammarus insensibilis* populations investigated in the UK**

| Pop[n] | G1 | G2 | G3 | G4 |
|---|---|---|---|---|
| G1 | 0 | - | - | - |
| G2 | 0.1732* | 0 | - | - |
| G3 | 0.1912* | 0.2073* | 0 | - |
| G4 | 0.1822* | 0.1541* | 0.1173* | 0 |

*$p$ <0.005

### *6.3.4 AMOVA*

Partitioning of microsatellite allelic variance was performed using analysis of molecular variance (AMOVA) (Excoffier *et al.* 1992) based on *F*-statistics. AMOVA partitions total variance into three covariance components. The 'within sites' component quantifies variation among individuals in the samples. The remaining variance, divided into 'among sites within groups' and 'among groups' components, can potentially give insight into relationships among sites by reflecting genetic diversity at different levels of a hierarchy. The covariance components are used to calculate fixation indices (*F*-statistics). The significance of the fixation indices are tested using a non-parametric permutation approach (Excoffier *et al.* 1992). Here hierarchical AMOVA was used to test for genetic structuring defined by the three geographical regions (south, east or north-east coast) where the sampled lagoons occur. Significance was tested using 16 000 permutations. AMOVA revealed that the greatest percentage of the variance (85%) was explained by within-sub-population variation and the least (-2%) by different sampling regions (Table 6.3h).

**Table 6.3h. Analysis of molecular variance (AMOVA) among groups (southern group G1, G2 eastern group G3 & north-eastern group G4), among the populations within those groups and within all populations regardless of group. Total % distribution of the variance components and *p*-values for corresponding $\Phi$-tests are given.**

| Source of variation | d.f. | Variance components | % of variation | *p* |
|---|---|---|---|---|
| **Among Groups** | 2 | -0.04 | -2.42 | ns |
| **Among populations within groups** | 1 | 0.29 | 17.87 | < 0.001 |
| **Within populations** | 414 | 1.36 | 84.55 | < 0.001 |

## 6.4 Discussion

### *6.4.1 Genetic variation*

#### *(i) Heterozygosity*

Once linked loci had been disregarded (Gi4 & Gi5) it was observed that in population G4 (Humberstone Fittes) locus Gi6 was nearly a single heterozygotic genotype (the raw data show only one individual with a genotype differing by a single dinucleotide repeat, (n= 60); Table 6.3b). Furthermore, population G3 (Salthouse Broad) had two loci (locus Gi2 & Gi9) where just a few individuals showed different genotypes (different only by a single dinucleotide repeat; Table 6.3b). These data gave highly significant observed heterozygosities of 1.0 ($p< 0.01$). In some invertebrate populations this has been seen as evidence for obligate asexuality (Pàlsson 2000).

There is no evidence to date of asexuality in any species of amphipod, however, the biology of amphipods is poorly understood. The reproduction of *G. duebeni*, presently better studied than that of *G. insensibilis*, shows no evidence for parthenogenesis. In the former species, unfertilised eggs will pass through initial cleavage stages but development then ceases, in addition, laboratory studies of this species reveal that egg laying does not occur in those individuals where successful mating has not taken place (Sheader 1983). Moreover, within the phylum Crustacea, only the Branchipoda contain species which reproduce by parthenogenesis (although both Cephalocaridia and Cirripedia contain hermaphrodites) (Barnes *et al.* 1988). It would appear, on the basis of this evidence, unlikely that asexuality is a plausible explanation for the significant heterozygote excess revealed by this study. It is possible, however, that this species is exhibiting sperm-dependent parthenogenesis. Taxa which demonstrate this life history strategy often consist of sexual and sperm-dependent forms. The latter produce fertile sperm and mate, but received sperm is only used to induce parthenogenetic embryo development (Weinzierl *et al.* 1999). The freshwater planarian *Schmidtea polychroa*, is one such species and there does appear to be a weak niche differentiation between the

forms. The reason for this differentiation, however, remains unclear (Weinzierl *et al.* 1999).

Obligatory parthenogenetic populations of *Daphnia pulex* have been shown to consist of only a few multilocus genotypes with excess or fixed heterozygosity (Innes *et al.* 1986). In extreme situations, however, similar patterns have been observed, created by associated over-dominance from linkage and deleterious mutations in small sexual populations (Pàlsson & Pamilo 1999). It is plausible that both these factors could be affecting the genotypic diversity of populations of *G. insensibilis* in the UK.

Pàlsson (2000) suggests that populations existing with single heterozygotic genotypes, which contain alleles shared with neighbouring populations, are probably experiencing a bottleneck event rather than demonstrating any long term asexual strategy. Moreover, bottleneck events may be compounded by selection for heterozygous genotypes. Where the single heterozygotic genotypes were observed in this study, the alleles concerned were also present in other populations. Interestingly, the heterozygotic genotype (Gi6) found in population G4 (Humberstone Fittes) contains alleles common only to the southern populations of *G. insensibilis* (Table 6.3b).

For 4 of the 6 loci, observed heterozygosities were greater than expected for all populations considered. Such a pattern would be observed by admixture of different allele sizes resulting from long distance migration, or may arise from an expansion in population sizes after a bottleneck event (Boileau *et al.* 1992). A bottleneck event, that is a contraction in the population, will severely affect the number of alleles, resulting in more heterozygotes present in the population than would be expected (Cornuet & Luikart 1996). Where populations are unstable, as may be the case for a species living in a marginal habitat, such as that afforded by the lagoon environment, changes in population size seem a plausible explanation. It is important to note, however, that not all loci demonstrate the same pattern. Locus Gi3 showed significantly less observed heterozygosity ($p< 0.01$) in 2 of the 4 populations and locus Gi9 in 1 ( $p< 0.01$). Where assumptions associated with particular mutational models are being violated such locus specific patterns may be observed (Pàlsson 2000).

Significant heterozygote excess in local groups has been associated with migration of one sex from the population (Balloux *et al.* 1998). Should populations of *G. insensibilis* exist with uneven sex ratios, this may go some way to explaining the significant heterozygote excess seen in these populations. Sheader (1983) found a skewed sex ratio in a population of *G. duebeni* studied over a period of one year in the Test Estuary. During October - February mature females outnumbered mature males by a minimum of 2:1 and in early August by a minimum of 3:1. It is likely, however, that genetic factors and photoperiod, which are thought to control sex determination in *G. duebeni* and other amphipods, are altering the sex ratio, rather than an exodus of males from the populations. Thomas *et al.* 1996 report the operational sex ratio of *G. insensibilis* to be severely biased in favour of males in populations that they have studied. Where a skewed sex ratio exists it may act to lower the effective population size and therefore increase the likelihood of inbreeding. Whether such a skewed sex ratio is affecting the heterozygosity in populations of *G. insensibilis* remains an open, and relevant empirical question.

### *(ii) Effective population size*

Some workers (Gerlach & Musolf 2000) have tested to see if significant numbers of loci within sub-populations exhibit heterozygote excess, as evidence for a reduction in effective population size. Such data may be evidence for a severe, range-wide bottleneck, which could be the cause of any genetic differentiation observed. Within a population it is unusual for each individual to have an equal chance of contributing to the next generation, meaning that the total size of a population is unlikely to reflect the size of the breeding population. This may be a result of unequal sex ratios, differences in individual fertility, overlapping generations, fluctuations in the number of animals of breeding age etc. (Hartl 2000). The effects of the population size on inbreeding, expressed as the co-efficient $F$, can be shown to be dependent on the effective population size $N_e$, and on changes over time:

$\Delta F = 1 / (2 N_e)$ per generation (Wright 1931)

this means that populations with a smaller effective size will become inbred more rapidly. Other consequences of a small effective population size are random fluctuations and loss of alleles because of genetic drift. The chances of alleles being lost or fixed will be increased as the effective population size decreases, with alleles at a low frequency having a greater chance of being lost from a population of a small effective size (Hartl 2000). Genetic drift increases the differentiation between sub-populations, as chance alleles may be lost from small isolated populations (Goropashnaya *et al.* 2001). Such a pattern has been observed for all sub-populations of *G. insensibilis* in this study, with significant pairwise $F_{ST}$ values between all sites (Table 6.3g).

Populations of *G. insensibilis* may have been subjected to temporary reductions in population size as a result of high predation, disease or by the longer term effects of isolation. Alternatively re-colonisation of the sites under investigation may be occurring which would give the appearance of the 'bottleneck' because of a founder event by just a few individuals. The nature of the lagoon habitat is one of 'boom and bust', common to many marginal habitats where migrant flow is increased from densely populated areas into sparsely populated areas (Lenormand 2002). Bamber *et al.* (1992) refer to the 'ephemerality' of the lagoon environment which may provide opportunity for extinction and re-colonisation of populations. Unfortunately surveys of lagoons do not occur on a frequent temporal basis and ecological data which may provide evidence for either founding events, or events which may be causing the disappearance of populations of *G. insensibilis* are scarce. However, given the genetic evidence combined with available ecological data (Sheader & Sheader 1985a; Sheader & Sheader 1985b; Bamber *et al.* 2000; Bamber *et al.* 2001), it would seem plausible that the populations of the lagoon sand shrimp are experiencing a range-wide population bottleneck; either as a result of a decline in effective population size or perhaps more optimistically prior to a population expansion following recent founding events.

### *(iii) Selection & genetic distance*

*Genetic distance*

Where homozygosity is associated with deleterious mutations, selection will favour a heterozygotic genotype, this may lead to an over-dominance of heterozygotes at neutral markers. This could lead to distinct lineages of microsatellites of a certain size in a population, resulting in observed heterozygosities exceeding that expected (Pàlsson 2000). For small populations, for populations experiencing bottlenecks and for those reproducing asexually such an effect would be magnified (Pàlsson 1999). Furthermore, there is evidence to suggest that genetic distance can be increased by population bottlenecks (Chakraborty & Nei 1976) and the pairwise $F_{ST}$ values for all *G. insensibilis* populations showed significant moderate to high genetic differentiation ($p< 0.005$; Table 6.3g). Moreover the relatively high proportion of rare alleles across all populations, particularly G1 & G4 (Table 6.3c), provide additional evidence for recent bottleneck events (Goropashnaya *et al.* 2001).

*Selection*

For small, sufficiently isolated populations such as those found in lagoons, mildly deleterious mutations, as well as severe ones, can be a strong force for extinction. Individually such mutations are effectively 'invisible' to selection but their combined affect of reducing population viability can be as catastrophic as a disastrous environmental event (Higgins & Lynch 2001). Theoretical models have shown that where mutation is absent increased carrying capacity in the populations cause time to extinction to be massively increased (Higgins & Lynch 2001). It would appear, however, that for the populations of *G. insensibilis* in this study, carrying capacity has been reduced resulting in a reduced effective population size which may be a real threat to the survival of the populations. Alternatively, since populations of *G. insensibilis* are still evident, it appears that selection may be acting to maintain genotypes better adapted to the lagoon environment. The strong linkage disequilibria detected in at least 2 loci for all populations of *G. insensibilis*, and in 4 loci for one population studied are indicative of fixation of successful genes (genotypes) in populations around the UK.

## *6.4.2 AMOVA & Population structuring*

The results from the AMOVA indicate that there is a lack of macrogeographic patterns as there is more difference among populations within a region than among regions (Table 6.3h). Studies on the genetic variation of *Daphnia pulex* living in temporary freshwater pond habitats have revealed similar patterns; that is, large differences in populations in close proximity but a lack of overall macrogeographical differentiation (Crease *et al.* 1990; Innes 1991). The dynamic revealed in the present study appears to have affected the heterozygosities of the populations resulting in the levels of divergence seen. The large divergence in neighbouring populations may have resulted from founding events, which have led to bottlenecks in the populations (Boileau *et al.* 1992). Additional evidence for recent founding events comes from the presence of linked loci in the populations of *G. insensibilis*. Linkage disequilibrium may be identified in founding populations where it has had insufficient time to decay through recombination (Hartl 2000) or where inbreeding has been occurring. As mentioned previously, linkage disequilibrium identified for loci in populations of *G. insensibilis* provide further evidence for recent founding events. McPhail & Carveth (1993) document biological disjunction between sub-populations of several riverine/coastal fish species. Explanations used to explain this differentiation include a combination of founding events and selective pressures (Smith *et al.* 2001). Both these factors could be invoked to account for the genetic pattern that has emerged regarding the UK population of *G. insensibilis*.

The $F_{ST}$ and AMOVA data revealed in this study may be a result of the introduction of individuals from neighbouring populations or rare long distance migrants (Rousset 1997). Although migration is unlikely given the life history strategy of *G. insensibilis*, it is possible that dispersal via mats of lagoon flora may occur between lagoons allowing the introduction of new individuals. Dispersal via floating mats of driftweed for juveniles of the marine amphipod *G. locusta* is known (Tully & Céidigh 1986) and there is a well established relationship between *G. insensibilis* and some species of marine/lagoon flora (Sheader & Sheader 1985b). Without empirical data and given the high levels of separation seen between populations of *G. insensibilis*, it seems improbable that this may be a cause of the observed patterns. Even moderate levels of

gene flow, such as may result from a few migrants, would be expected to reduce genetic differentiation, between even distant populations (Carmichael *et al.* 2001).

*Fixed heterozygosity*

Given that dispersal (gene flow) is likely to be low in this species, founder events leading to high levels of inbreeding are expected. There is evidence of inbreeding from the high degree of linkage disequilibria seen in all populations of *G. insensibilis* studied. Where extreme inbreeding occurs with selection against homozygotes, fixed heterozygosity is a possibility (Bierne *et al.* 1998). This may explain the fixed heterozygosity seen for some loci, in some of the populations (Table 6.3d) (although where $N$ is small, as is the case for some populations, homozygotes may be lost by chance). Alternatively, hybridisation combined with sperm-dependent parthenogenesis may result in fixed heterozygosities such as those seen in populations *of Gammarus insensibilis*. Hybridisation of a species followed by a sex-related cytogenetic factor has been observed in species belonging to the *Daphnia pulex* complex. Here a hybridisation event was followed by either the ejection or silencing of the maternal nuclear genome, resulting in offspring (clones) demonstrating near complete homozygosity (Dufresne & Hebert 1994). There are species, however, that make transitions to polyploidy without interspecific hybridisation. Allozymes and mtDNA have been used to demonstrate internal genome-fusion events in the freshwater ostracod *Cyprinotus incongruens* (Chaplin & Hebert 1997). Clearly more work into the life history strategies of *G. insensibilis* and the chromosomal mechanisms associated with the reproduction of this species is needed. Finally consideration must be given to the idea that this species may, as data for other lagoon specialists (*Nematostella vectensis* and *Cerastoderma glaucum*) indicate, be adapted to a narrow niche particular to a specific lagoon. If this were the case maintenance of an optimal genotype may be occurring through translocation or other means, e.g. assortative mating or sperm-dependent parthenogenesis.

# 7.0 General Discussion

## 7.1 Island biodiversity - crossing the Rubicon

## 7.2 The Population genetics of the lagoon specialists; *Nematostella vectensis*, *Cerastoderma glaucum* & *Gammarus insensibilis*



## 7.3 Conclusions

## 7.1 Island biodiversity - crossing the Rubicon

The Earth's biodiversity is being reduced at an unprecedented rate because of the destruction to and the resulting loss of natural habitats (Eldridge 1999). Habitat fragmentation may lead to fewer areas supporting ever decreasing populations (Gerlach & Musolf 2000). Factors associated with fragmentation and the reduction in population size, such as reduced gene flow, increased inbreeding coefficients and reduced fecundity are all contributing to an increase in extinction (Hartl *et al.* 1992). "Gene flow" is understood as the change in gene frequency owing to the movement of gametes, individuals or groups of individuals from one place to another, whereas "genetic drift" is the unpredictable change in gene frequency owing to finite population size (Slatkin 1987). Demographic, environmental and genetic stochasticity are all greatly affecting the increasing number of small isolated populations that are proving unable to adapt (Caughley 1994). The role played by gene flow in structuring the genetics of a particular species, in a particular population, will depend on both the geographic distribution of that species and on other important evolutionary forces (Slatkin 1987). There is much discussion in the literature regarding the influence of each of the above factors (Lande 1988; Caughley 1994; Frankham 1995; Hedrick *et al.* 1996; Brookes 1997; Lacy 1997). Which factor is most influential in causing extinction is still widely debated.

Whatever the causes, however, it is an accepted fact that we need a greater understanding of the process of population extinction, if we are to achieve the aim of protecting the planet's biodiversity. Island populations are highly vulnerable to extinction because of their small size and their limited dispersal ability (Frankham 1997, 1998). Species with high dispersal (high gene flow) will demonstrate little geographic (inter-population) differentiation, and the same alleles will occur with similar frequencies over large areas, whereas species with limited or no dispersal (low gene flow) will show distinct inter-population variability and different allele frequencies will be observed (Hummel *et al.* 1994). Higher measures of variability, such as may be found in larger lagoons, where the populations are more stable, may imply those populations are less effected by isolation (Johnson & Black 1998). Understanding the

population dynamics of species that live in small isolated groups, often in marginal habitats, may increase our understanding of population extinction in fragmented habitats and go some way to helping preserve biodiversity.

In the present study RAPDs and microsatellite markers have been used to examine the population genetics of three lagoon specialists, including any existing relationships or divergences between spatially disjunct populations of these invertebrates. The availability of RAPDs for use with anonymous genomes and their relative lack of expense make them potentially valuable genetic markers for population studies (Hadrys *et al.* 1992). Microsatellites may offer an advantage over protein studies for many reasons, for example; the ease of preservation of DNA compared to protein, high mutation rates and high degrees of polymorphism. The latter two reasons may also confer an advantage over other nuclear and mitochondrial DNA studies where resolution or fine scale genetic differentiation may be lacking (Hughes & Queller 1993; Forbes *et al.* 1995). Furthermore, the high mutation rates, which give microsatellites their high polymorphism, make them a useful tool for studying geographically proximate populations (Jarne & Logoda 1996).

## 7.2 The Population genetics of the lagoon specialists; *Nematostella vectensis*, *Cerastoderma glaucum* & *Gammarus insensibilis*

### *7.2.1 Summary of main findings*

Genetic data suggest that the fragmented nature of the lagoon habitat has greatly influenced the genetic population structure of the three lagoon specialists under investigation. In spite of three very different life history strategies, all three species demonstrate genetic patterns associated with a high degree of isolation, combined with limited dispersal ability. The reproductive modes of the three organisms are very disparate, the anemone reproducing mainly asexually by transverse fission, the sand shrimp brooding its developing young in a specialised marsupium and the cockle adopting a method common to many marine species, broadcast spawning. The RAPD molecular markers used, reveal the mainly asexual sea anemone, *Nematostella vectensis* to be predominantly clonal throughout its known UK range, with monoclonal or near monoclonal populations and rare genotypes confined to single sites. The microsatellite markers employed in this study indicate a high degree of genetic differentiation, severe inbreeding and the absence of Hardy-Weinberg equilibrium conditions in populations of the lagoon sand shrimp, *Gammarus insensibilis* and the lagoon cockle *Cerastoderma glaucum*.

The population structure of *Nematostella vectensis*, revealed by RAPD analysis, is thought to result from its reproductive mode combined with low levels of dispersal (Pearson *et al.* 2002). Cloning as a mode of reproduction tends to coincide with situations that favour a rapid increase in population numbers, e.g. seasonal or unstable habitats (Hughes 1989). The ability of this anemone to reproduce asexually lends itself well to the colonisation of new sites and to recovery following environmental adversity.

In the isolated populations of *Cerastoderma glaucum* the proportion of heterozygotes was significantly lower than expected, probably as a result of the fixation or near fixation of alleles (Madsen *et al.* 1996). It is likely that a reduction in effective population size has occurred in this species indicated by skewed sex ratios and

significant $F_{IS}$ values that revealed severe inbreeding. Where inbreeding continues in a population, selection will act to reduce the effects by reducing the frequency of deleterious recessive alleles (Frankel & Soulé 1981; Van Noordwijk & Scharloo 1981; Jarvinen & Varvio 1985). Further work investigating the prevalence of reproductive abnormalities, e.g. reduced fecundity, inviable offspring etc. in populations of *C. glaucum* may indicate if there has been sufficient time for selection to be effective against deleterious mutations.

Populations of *Gammarus insensibilis* were characterised by significant excesses in heterozygosity. This is a phenomenon commonly associated with bottleneck events, that is, contractions in population sizes (Boileau *et al.* 1992). These data combined with evidence of skewed sex ratios and significant inbreeding in all populations investigated, lead to the conclusion that the effective population size of this species is extremely low. The genetic data provide evidence to indicate that a wide range bottleneck has occurred, and theoretical models suggest this sometimes precludes an extinction event (Frankham 1998). The high degree of genetic differentiation between the populations points to little or no dispersal which may reduce chances for recovery from this situation. For isolated animal populations, however, population genetic theory predicts that inbreeding will expose deleterious alleles to natural selection, thereby purging the genetic load resulting in further inbreeding causing little or no reduction in fitness (Waller 1993). This assumption will be violated if sufficient migrants re-introduce purged deleterious alleles (Keller & Waller 2002).

The population structure of all three species appears to have been affected by their limited dispersal capabilities. In spite of the ability demonstrated by the anemone, *N. vectensis*, for rapid increase in population size, the genetic diversity, as for the other two species under investigation, remains low. The incidence of rare genotypes for the anemone *N. vectensis* in single populations and the presence of private alleles in populations of both *G. insensibilis* and *C. glaucum* confirm that levels of dispersal are low for lagoon organisms regardless of life history. It is thought that approximately 10% of the lagoon habitat will be lost in the next 20 years (Downie 1996). Should this occur, the further fragmentation of this habitat will exacerbate the situation and it will

become even harder for organisms living in these island-like habitats to disperse and maintain gene flow.

### *(i) Habitat fragmentation*

Connection between the surrounding sea and any one lagoon may be an important factor in aiding dispersal of lagoon species. Furthermore, the specific nature of that connection may be more influential in dispersal/recolonisation from one lagoon to another. For example an occasional storm surge over a narrow low barrier may be an important dispersal aid in one lagoon whereas tidal range may be more important to another. In other marine habitats the dispersal of the intertidal snail *Bembicium viltatum* is dependent on the type of connection between its 'pond' habitat and the sea (Johnson & Black 1998).

Widewater lagoon (Sussex, UK) has been landlocked for approximately the last 300 years. The only means for seawater to enter is by percolation, which effectively prevents colonisation by any macrofauna from the surrounding sea. A survey of this site in 1978/9 recorded a small population of *G. insensibilis*, a subsequent survey in 1985 found the population absent (Sheader & Sheader 1985b). Moreover, they noted that compared to lagoons in the surrounding area the general level of species diversity in Widewater lagoon was low, reflecting the inability for recolonisation of such an isolated lagoon following extinction events.

The example of Widewater lagoon probably reflects a general trend around the UK. Firstly, because of limited connection to the surrounding seas, even species such as *C. glaucum* with planktonic larvae, find dispersal difficult. Local heterogeneity is known for some species of marine invertebrates (see Hedgecock 1986 for review) and such heterogeneity may result from physical barriers or from selective processes. Secondly, the ever increasing distance from one lagoon site to another, makes potential gene flow (dispersal) an unlikely event. In the case of *C. glaucum* it is probably the difficult access to lagoon sites that is preventing colonisation by immigrants. This species is able to survive and is found in other marine habitats; such as estuaries and sheltered

shore lines (Boyden & Russell 1972; Russell 1972; Brock 1980b). It therefore has 'stepping stone' habitats that could provide potential refuge between its preferred lagoon habitat. The high degree of genetic differentiation between populations of *C. glaucum* indicates that this is not happening, except perhaps between the populations of C1 and C2, where no significant genetic differentiation was observed. These lagoons are distinct from the others in the study in their connection to the surrounding sea, which in both cases is an open channel (Sheader & Sheader 1985a) and they are also less than 1 mile apart by land or sea. It would seem most likely that it is the connection between lagoon and sea that plays a major role in influencing the genetic population structure of *C. glaucum*. Where populations exist as part of a metapopulation failure to migrate/disperse between sub-populations may increase the likelihood of extinction.

A population of *N. vectensis* has been identified in the estuarine environment in the UK (C. Pearson pers. obs.) and its recorded absence from this habitat is more likely a result of being overlooked, than an inability to survive there. It is an abundant species in the estuarine environment in the U.S.A (Harter 1997). This anemone, like the lagoon cockle, is able to make use of a variety of marine habitats, implying that inaccessibility to individual lagoons must be a stronger influence on gene flow than the distance between them. *Gammarus insensibilis*, however, is a true lagoon specialist, that is, one that is better suited and will survive to out-compete marine counterparts in such an environment (Barnes 1980). It is possible that for this species, both dispersal and survival in the open seas combined with difficulty in entering new lagoon sites are major influences on its population genetic structure.

The presence of other species may distinguish between possible founder events associated with colonisation of a lagoon, subsequent bottlenecks and the effectiveness of barriers in preventing occasional gene flow. Johnson & Black (1998) note that where fish are present in a lagoon there must be some degree of connection between the sea and the habitat and sufficient volume and persistence of water to sustain the fish. They suppose that the presence of fish may reflect a reduced chance of bottlenecks for other species in the lagoon, but do note that this is a very complex relationship with many factors to consider. Although not a factor considered in this investigation, this might be

an initial way to establish which cause (reduced dispersal or physical obstacle) is a potential barrier to gene flow in lagoon species.

### *(ii) Founder events & bottlenecks*

The ephemeral nature of many lagoon sites means that they may be subjected to marked variation in population size, leading to population bottlenecks and even extinction. As a result of this, genetic drift will probably have considerable influence in shaping the genetic structure of lagoon species (Städler & Jarne 1997). It is important to note that although genetic drift will affect all loci in a population in the same way, the same is not true for selection. Natural selection for locally important adaptations can cause substantial differences at a few loci, with loci that are neutral or weakly selected being uniform throughout a species' geographical range (Slatkin 1987). Crow & Kimura (1970) uphold that in the absence of strong selection very low levels of gene flow should be sufficient to increase genetic variation. Slatkin (1985) suggests, that exchange of a single individual every other generation would be enough to prevent genetic differentiation caused by random genetic drift but Mills & Allendorf (1996) report between 1 and 10 migrants are required every generation to reduce a loss of heterozygosity and polymorphism within sub-populations. The genetic data obtained in this study would indicate that no such influx of migrants is occurring.

Where local variation in population size exists affecting the effective population size, a loss of variability and a temporary increase in genetic distance among populations may be generated (Nei *et al.* 1975). This is clearly the case for all three lagoon organisms investigated in this study. The most extreme form of this, a founder event, which may be caused by the colonisation of new habitats is evident for the lagoon sand shrimp *G. insensibilis* and the anemone, *N. vectensis*. Bottlenecks are often observed in recently founded populations by the presence of a reduced number of genotypes, such as have been recorded for both *G. insensibilis* and *N. vectensis*. For clonal species, population bottlenecks may be particularly severe as a single individual is able start a new population (Samadi *et al.* 1999). Hoffmann (1987), using electrophoretic techniques, has shown that for the sea anemone *Metridium senile*, asexual reproduction and the

founder effect are important in stabilising and establishing populations of this species when the recruitment of sexually reproduced larvae is rare. Recent colonisation by *N. vectensis* and *G. insensibilis* seems unlikely for many of the populations, because of physical barriers preventing immigration into many of the lagoons. Chakraborty & Nei (1977) note that following a bottleneck event the genetic effects on the population will depend on *N*, on the size of the bottleneck and on the rate of recovery from the bottleneck. This may have important implications for species which have a low rate of increase. The cause of the reduced variation and bottlenecks observed is most probably inbreeding and genetic drift combined with insufficient time, since the reduction in effective population size (that is the initial founding event) for either mutation and/or selection to act.

Slatkin (1985; 1987) argues that extinction and recolonisation may be an important source of gene flow, lowering genetic differentiation between subdivided populations. He proposes that where the time to extinction (in generations) of a local population is less than or equal to the effective number of locally breeding adults, extinction and recolonisation can reduce or eliminate the effects of genetic drift. Wade & McCauley (1988) propose a contrary argument, in agreement with Wright 's (1940) proposal that the effect of extinction and recolonisation events is to create local genetic differentiation. Specifically Wade & McCauley (1988) note that the effects of extinction and colonisation on the genetic differentiation of local populations depends on how the colonising groups are formed, and on the relationship between colonisation and migration. Data here indicate that both these factors may be important to the genetic population structure of populations of *G. insensibilis* throughout the UK.

### *(iii) General purpose genotype*

Bamber & Henderson (1989) proposed a model of speciation for silverside fishes, predicting that isolated and variable environments, such as found in coastal lagoons, act to select for general purpose genotypes, where a generalist genotype enables an organism to adjust their morphology, physiology and behaviour to a wide range of conditions (Beheregaray & Sunnucks 2001). Such general purpose genotypes may be a

wide ranging feature of species living in marginal habitats; allowing them to invade, radiate and colonise such habitats, by out competing both marine and freshwater counterparts. Ward (1989) states that while it is often assumed gene flow (or lack of it) is responsible for population differentiation, the possibility that differentiation is occurring as a result of selection specific to the individual habitat must be considered. *N. vectensis* is very eurytolerant and this may be associated with a single, successful dominant RAPD profile, an occurrence which has been reported using allozyme markers for the anemone *Haliplanella luciae* (Shick & Lamb 1977). It is, however, important to note that allozymes are functional enzymes that may be subject to selection whereas RAPDs representing unknown regions of the genome may not be under such selective pressure. Should such a genotype exist for this species when combined with rapid reproduction short term success in the marginal habitat may be guaranteed.

The presence of linkage disequilibrium in populations of both *G. insensibilis* and *C. glaucum* may be evidence for 'general purpose' genotypes associated with the lagoon habitat. Where linkage disequilibrium persists in a population it could be that natural selection is favouring certain genotypes with a force strong enough to overcome the natural tendency for linkage disequilibrium to be purged from that population (Pàlsson 1999). Further evidence for 'general purpose' genotypes lie in the presence of fixed heterozygosity seen in populations of *G. insensibilis* and in the presence of heterozygote deficiency seen in populations of *C. glaucum*. This hypothesis may be tested empirically by observing the presence of linked loci and through measurements of HWE in different populations of *C. glaucum*, those from the lagoon habitat and those from the less marginal estuarine habitat. Additionally cytogenetic studies may reveal chromosomal mechanisms, such as translocation, acting to link favourable genes/alleles.

## 7.3 Conclusions

### *7.3.1 Implications for conservation*

Inbreeding depression may occur in populations that show high levels of inbreeding, particularly where other factors such as skewed sex ratios are identified. Inbreeding depression is increasingly being identified in small isolated wild populations (Frankham 1995; Madsen *et al.* 1996; Lacy 1997). Forbes & Hogg (1999) suggest populations may be a cause for management concern when they exhibit both increased genetic differentiation and reduced genetic diversity. Some of the populations of lagoon invertebrate under investigation in this study show high levels of inbreeding and significant genetic differentiation e.g. the population of the lagoon cockle *Cerastoderma glaucum* in Abrahams Bosom lagoon on the north Norfolk coast. These low levels of genetic variability have major implications for conservation and management. Where a loss of heterozygosity reduces a population's ability to respond to environmental change, the probability of extinction increases (Caro & Laurenson 1994).

Lagoons in the UK are an important reservoir of unique biodiversity. If low levels of genetic diversity are found to be typical among lagoon specialists, the implication is that a lagoon population will preserve little of the genetic diversity of the original species, limiting its usefulness as a source to establish new populations (Eldridge *et al.*1999). However, where populations of a species exist at several isolated locations a more genetically diverse population could be produced by interbreeding, furthermore this will have a chance of increasing the fitness of populations suffering from inbreeding depression (Frankham 1995). Although increased dispersal to isolated sites would allow outcrossing and an increase in the supply rate of new mutations it will also act to decrease the probability of fixation in isolated populations (Krebs 1994), the implication being, interbreeding with translocated individuals may dilute a locally adapted genotype, and this may actually increase a population's chance of extinction. Furthermore, immigration into an isolated population may result in the re-introduction of previously purged deleterious alleles, and this may act to *increase* the risk of extinction of that population (Keller & Waller 2002).

Higgins & Lynch (2001) suggest that a very important factor in reducing extinction in small isolated populations is thought to be the carrying capacity of the population. They hypothesise that an increased carrying capacity may provide additional assurance against extinction by increasing the efficiency of selection against deleterious mutations. Traditionally the demography of a population has been used in determining minimum viable population sizes, more recently, however, population genetics has revealed that the rate of accumulation of deleterious mutations may be a greater threat to small isolated populations. This is especially true when considering time scales relevant to conservation of small isolated populations (Higgins & Lynch 2001).

Lynch (1997) reports on recent studies in conservation genetics needing to assess the risks of inbreeding relative to the risk of introducing non-endemic individuals, with particular reference to the possibility of 'obliterating important adaptations to specific geographic locations and ecosystems'. Lande (1988) suggests that conservation and recovery plans should include demography and population genetics data in order that survival of populations may be properly assessed. It would be vital to the success of any management plan involving lagoon organisms that consideration be given to local adaptation and the risk of diluting local genotypes. The strong linkage disequilibrium in populations of *C. glaucum* and *G. insensibilis*, the fixed heterozygosity in populations of *G. insensibilis*, the reduced levels of heterozygosity in populations of *C. glaucum* and the presence of a single dominant genotype in populations of *N. vectensis* may all be evidence of adaptation to a specific habitat niche.

### *7.3.2 Use of* F *statistics for estimating population structure*

There has been much discussion in the literature about the use of various methods, Slatkin's $R_{ST}$ (1995), unbiased $R_{ST}$ ($UR_{ST}$; Goodman 1997) and Wright's $F_{ST}$ (1965) for estimating genetic structuring (Valsecchi *et al.* 1997; Gaggiotti *et al.* 1999; Balloux & Lugon-Moulin 2002). Valsecchi *et al.* (1997) concluded unbiased $R_{ST}$ as the most reliable for microsatellite data as it takes into account unequal sample sizes. It must be emphasised, however, that $F_{ST}$ estimators are not generally sample-size dependent (Lugon-Moulin *et al.* 1999) and both Gaggiotti *et al.* (1997) and Balloux & Lugon-

Moulin (2002) suggest the most conservative approach is to use this estimator. Carmichael *et al.* (2001) suggest that $F_{ST}$ performs better as an estimator (than those based in SMM) where imperfect microsatellites are included in the analysis; this was the case for both *C. glaucum* and *G. insensibilis*.

The $F_{ST}$ estimator used for the analysis in this study was the unbiased estimator devised by Weir & Cockerham (1984), and therefore allowed for unequal sample sizes. Furthermore, evidence suggests that sample size is not as much of an influence on the accuracy of estimates of $F_{ST}$ as the number of independent loci tested (Beaumont & Nichols 1996). Unfortunately, in this study although several microsatellite loci were isolated for both *C. glaucum and G. insensibilis*, some proved impossible to optimise and others were found not to be statistically independent. Further work may benefit from time spent attempting to isolate more independent loci, and perhaps trying to cross amplify with primers designed for microsatellite loci from closely related species. This was attempted in this study with a few primers for both *C. glaucum* from *C. edule* and for *G. insensibilis* from *G. pulex* without success. Failure to amplify loci of *C. edule* using primers developed for *C. glaucum* populations provides further evidence that these two are separate species. Additional support is given to this hypothesis by the low levels of gene flow between populations *of C. glaucum*.

The mode of mutation for any given microsatellite locus is generally unknown, and as a result using $R_{ST}$ as an estimator may distort the level of differentiation between populations (Estoup & Cornuet 1999). Using $F_{ST}$ eliminates this difficulty, furthermore Gaggiotti *et al.* (1999) recommend the use of $F_{ST}$ where population sizes and the number of loci used are low, both of which are the case for my data.

### *7.3.3 Considerations & recommendations for further study*

In addition to the statistical testing for population differences and the estimations performed to measure levels of genetic variation, several other approaches may be useful to look at population structure of lagoon organisms. Gene diversity analysis, which partitions genetic variation into a range of spatial and temporal components,

could provide insight into population subdivision and temporal stability. This may give an indication as to the suitability of individual sites for the survival of different populations over time. Furthermore, performing a Mantel test correlation between microsatellite data and geographical distance may provide empirical evidence for isolation by distance.

The fragmentation and loss of the lagoon habitat, combined with low dispersal rates possibly resulting from physical difficulties associated with immigration into new lagoon sites, are contributing to the reduction in genetic variation and gene flow of the three lagoon invertebrates investigated in this study. There is evidence in the literature to suggest that should this situation be allowed to continue, the levels of inbreeding and reduction in effective population sizes may lead to the extinction of these species from some sites (Frankham 1995b; Keller & Waller 2002). While the lagoon cockle, *Cerastoderma glaucum*, is not considered endangered and is abundant in the lagoon and other marine habitats, the genetic circumstances facing this species may well be indicative of the situation facing other lagoon organisms. The present study, however, indicates that although low genetic diversity is present in populations of lagoon invertebrates (probably a result of low levels of gene flow and founder effects or bottlenecks) it is likely that populations are well adapted to their individual niches. Evidence for this is seen in the continued survival of populations living in the lagoon habitat. Conservation strategy must therefore provide a balance between maintaining sufficient diversity *and* preserving successful genotypes. The most likely way to achieve this would be through preservation of existing lagoon habitats, ensuring that environmental parameters remain the same and that no further lagoon sites are lost.

Data provided by highly variable molecular DNA markers such as microsatellites and RAPDs are impacting on conservation biology in an unprecedented way because of their ability to help resolve many previously unanswered questions. The power of these tools is not to be underestimated. However, prudence must be used when unravelling data generated by these markers. Hedrick (1999) urges caution particularly when interpreting data from microsatellite loci, because of the statistical power associated with their analysis, the connection between any biological significance and statistical

significance may be weak. Future management plans will need to consider many factors, genetic, demographic and environmental if endangered species are to survive. This study has gone some way to demonstrate the need for detailed genetic work in giving insight into the 'health' of a species. For example, ecological surveys have shown the lagoon cockle to be abundant in many lagoon sites, giving a sense this species is surviving well in this habitat (Sheader & Sheader 1985a). Such surveys, however, do not reveal information regarding the *effective* population size or data concerning interactions occurring between sub-populations. It is only through a combination of all these factors that a true picture of a species can be built.

Opening of the more isolated lagoons to the surrounding seas has been suggested by Bamber *et al.* (1992). They note the importance of migration between lagoons on population densities and propose that isolation combined with a hypothesis of genetic plasticity could result in differentiation between lagoon populations. Data from these work indicate that this is indeed the case. However, providing access to lagoons that presently receive low numbers of migrants may jeopardise the survival of present populations because of 'flooding' of successful genotypes. Further work on understanding the mechanisms maintaining such genotypes is clearly needed. There is some evidence, however, that the development of 'artificial' lagoon sites may prove to be a useful tool in maintaining the biodiversity of some lagoon specialists (Allen *et al.*1995). The success of such sites considered with data from this thesis highlight the need to understand the biological and genetic requirements of each individual organism if the continued diversity of the lagoon environment is to be assured.

In spite of the usefulness of data provided by studies such as these, it must be remembered that the environmental consequences of human activity are by far the greatest threat to wild populations. These are occurring at a far greater rate, e.g. through loss of habitat, than any genetic challenges facing populations, such as inbreeding (Caro & Laurenson 1994).

# Chapter 8.0 References

Al-Suwailem AM (1991) The Ecology of a Saline Lagoon in Southern England. PhD thesis, University of Southampton, UK.

Allen JR, Wilkinson SB, Hawkins SJ (1995) Redeveloped docks as artificial lagoons: The development of brackish-water communities and potential for conservation of lagoonal species. *Aquatic Conservation: Marine and Freshwater Ecosystems*, **5**, 299-309.

Amos W (1999) A comparative approach to the study of microsatellite evolution. In: *Microsatellites Evolution and Applications*. Eds Goldstein DB, Schlötterer C, OUP, Oxford, UK.

Anon (1995) *Global Biodiversity Assessment*. United Nations Environment Programme, CUP, Cambridge, UK.

Ansorge W, Sproat B, Stegemann J, Schwager C, Zenke M. (1987) Automated DNA sequencing: ultrasensitive detection of fluorescent bands during electrophoresis. *Nucleic Acids Research*, **15**, 4593-4602.

Armour JAL, Alonso Alegre S, Miles S, Williams LJ, Badge RM (1999) Minisatellites and mutation processes in tandemly repetitive DNA. In: *Microsatellites Evolution and Applications*. Eds Goldstein DB, Schlötterer C, OUP, Oxford, UK.

Arnaud-Haond S, Bonhomme F, Blanc F (2003) Large discrepancies in differentiation of allozymes, nuclear and mitochondrial DNA loci in recently founded Pacific populations of the pearl oyster *Pinctada margaritifera*. Journal of Evolutionary Biology, **16**, 388-398.

Avise JC (1994) *Molecular markers, natural history and evolution*. Chapman & Hall, London, UK.

Avise JC, Hamrick JL (1996) *Conservation genetics*. Champman & Hall, New York, USA.

Ayala FJ (1965) Evolution of fitness in experimental populations of *Drosophila serrata*. *Science*, **150**, 903-905.

Ayre DJ (1984a) The sea anemone *Actinia tenebrosa*: An opportunistic insectivore. *Ophelia*, **23**, 149-153.

Ayre DJ (1984b) The effects of sexual and asexual reproduction on geographic variation in the sea anemone *Actinia tenebrosa*. *Oecologica (Berlin)*, **62**, 222-229.

Ayre DJ (1985) Localized adaptation of clones of the sea anemone *Actinia tenebrosa*. *Evolution*, **39**, 1250-1260.

Ayre DJ, Read J, Wishart J (1991) Genetic subdivision within the eastern Australian population of the sea anemone *Actinia tenebrosa*. *Marine Biology*, **109**, 379-390.

Ayre DJ, Hughes TP, Standish RJ (1997) Genetic differentiation, reproductive mode, and gene flow in the brooding coral *Pocillopora damicornis* along the Great Barrier Reef, Australia. *Marine Ecology Progress Series*, **159**, 175-187.

Balloux F, Goudet J, Perrin N (1998) Breeding system and genetic variance in the monogamous, semi-social shrew, *Crocidura russula*. *Evolution*, **52**, 1230-1235.

Balloux F, Lugon-Moulin N (2002) The estimation of population differentiation with microsatellite markers. *Molecular Ecology*, **11**, 155-165.

Bamber RN (1997) Assessment of saline lagoons within Special Areas of Conservation. *Peterborough, English Nature Research Reports*, No. **235**.

Bamber RN, Battern SD, Bridgewater ND (1991) The brackish ponds at Killingholme, Humberside, UK. *Aquatic Conservation Marine and Freshwater Ecosystems*, **1**, 173-181.

Bamber RN, Battern SD, Sheader M, Bridgewater ND (1992) On the ecology of brackish water lagoons in the United Kingdom. *Aquatic Conservation: Marine and Freshwater Ecosystems*, **2**, 65-94.

Bamber RN, Evans NJ, Whittall A (2000) Survey of potential coastal saline lagoons and pools in Wales, December 1998. *Countryside Council for Wales Report*, No. **377**.

Bamber RN, Gilliland PM, Shardlow MEA (2001) Saline lagoons: a guide to their management and creation (interim version). Peterborough, English Nature.

Bamber RN, Henderson PA (1989) Preadaptive plasticity in atherinids and the estuarine §eat of teleost evolution. *Journal of Fish Biology*, **33**, Supplement A, 17-23.

Barnes RSK (1980) *Coastal Lagoons*, CUP, Cambridge, UK.

Barnes RSK (1988) The faunas of land locked lagoons: chance differences and the problems of dispersal. *Estuarine and Coastal Shelf Science*, **26**, 309-318.

Barnes RSK (1989) The coastal lagoons of Britain: an overview and conservation appraisal. *Biological Conservation*, **49**, 295-313.

Barnes RSK (1991) Dilemmas in the theory and practice of biological conservation as exemplified by British coastal lagoons. *Biological Conservation*, **55**, 315-328.

Barnes RSK, Calow P, Olive PJW (1988) *The Invertebrates: a new synthesis*. Blackwell Scientific Publications, Oxford, UK.

Beaumont MA, Bruford MW (1999) Microsatellites in conservation genetics. In: *Microsatellites Evolution and Applications*. Eds Goldstein DB, Schlötterer C, OUP, Oxford, UK.

Beaumont MA, Nichols RA (1996) Evaluating loci for the use in the genetic analysis of population structure. *Proceedings of the Royal Society of London Series B*, **263**, 1619-1626.

Beheregaray LB, Sunnucks P (2001) Fine-scale genetic structure, estuarine colonization and incipient speciation in the marine silverside fish *Odontesthes argentinensis*. *Molecular Ecology*, **10**, 2849-2866.

Benzie JAH, Ballment E, Forbes AT, Demetriades NT, Sugama K, Haryanti MS (2002) Mitochondrial DNA variation in Indo-Pacific populations of the giant tiger prawn, *Penaeus monodon*. *Molecular Ecology*, **11**, 2553-2569.

Berger EM (1983) Population genetics of marine gastropods and bivalves. In: *The Mollusca, Volume 6, Ecology*. Ed Russell-Hunter Bird WD, Academic Press, Orlando, USA.

Bierne N, Launey S, Naciri-Graven Y, Bonhomme F (1998) Early effects of inbreeding as revealed by microsatellite analyses on *Ostrea edulis* larvae. *Genetics*, **148**, 1893-1906.

Bird ECF (1970) *Coasts*. MIT Press, Cambridge, Massachusetts, USA.

Bishop MJ, Bishop SJ (1973) Associations of molluscs and marine plants at San Diego, California. *Journal of Conchology*, **28**, 43-54.

Black R, Johnson MS (1979) Asexual viviparity and population genetics of *Actinia tenebrosa*. *Marine Biology*, **53**, 27-31.

Boileau MG, Herbert PDN, Schwartz SS (1992) Non-equilibrium gene frequency divergence: persistent founder effects in natural populations. *Journal of Evolutionary Biology*, **5**, 25-39.

Borowsky RL, McClelland M, Cheng R, Welsh J (1995) Arbitrarily primed DNA fingerprinting for phylogenetic reconstruction in vertebrates. *Molecular Biology and Evolution*, **12**, 1022-1032

Borowsky RL, Vidthayanon C (2001) Nucleotide diversity in populations of balitorid cave fishes from Thailand. *Molecular Ecology*, **10**, 2799-2805.

Bowditch BM, Albright DG, Williams JKG, Braun, MJ (1994) Use of Randomly Amplified Polymorphic DNA markers in comparative genetic studies. *Methods in Enzymology*, **224**, 294-309.

Boyden CR (1971) A comparative study of the reproductive cycle of the cockles *Cerastoderma edule* and *C. glaucum*. *Journal of the Marine Biological Association of the UK*, **51**, 605-622.

Boyden CR (1972) The behaviour, survival and respiration of the cockles *Cerastoderma edule* and *C. glaucum* in air. *Journal of the Marine Biological Association of the UK*, **52**, 661-680.

Boyden CR, Russell PJC (1972) The distribution and habitat range of the brackish water cockle (*Cardium (Cerastoderma) glaucum*) in the British Isles. Journal *of Animal Ecology*, **41**, 719-34.

Brock V (1978) Morphological and Biochemical criteria for the separation of *Cardium glaucum* (Bruguière) from *Cardium edule* (L.). *Ophelia,* **17**, 207-214.

Brock V (1979) Habitat selection of two congeneric bivalves, *Cardium edule* and *C. glaucum* in sympatric and allopatric populations. *Marine Biology,* **54**, 149-156.

Brock V (1980a) Evidence for niche differences in sympatric populations of *Cerastoderma edule* and *C. lamarcki*. *Marine Ecological Progress Ser*ies, **2**, 75-80.

Brock V (1980b) The geographical distribution of *Cerastoderma [Cardium] edule* (L.) and *C. lamarcki* (Reeve) in the Baltic and adjacent seas related to salinity and salinity fluctuations. *Ophelia,* **19**, 207-214.

Brock V (1987) Genetic relations between the bivalves *Cardium* (*Cerastoderma*) *edule*, *Cardium lamarcki* and *Cardium glaucum*, studied by means of crossed immunoelectrophoresis. *Marine Biology*, **93**, 493-498.

Brooks M (1997) A clean break. Conservation and genetics make bad bed fellows. *New Scientist*, **156**, 64.

Brown TA (1992a) Replication of DNA molecules. In: *Genetics a Molecular Approach*. Chapman & Hall, London, UK.

Brown TA (1992b) The Structure of DNA. In: *Genetics a Molecular Approach*. Chapman & Hall, London, UK.

Bruford MW, Cheesman DJ, Coote T, Green HAA, Haines SA, O'Ryan C, Williams TR (1996) *Microsatellites and Their Application to Conservation Genetics*. In: Molecular Genetic Approaches in Conservation. Eds Smith TB, Wayne RK, OUP, Oxford, UK.

Bruford MW, Hanotte O, Brookfield JFY, Burke T (1992) *Multi- and single-locus DNA fingerprinting*. In: Molecular analysis of populations: A practical approach. IRL Press, Oxford, UK.

Bruford MW, Wayne RK (1993) Microsatellites and their application to population genetics. *Current Opinions in Genetic Development,* **3**, 939-943

Bush SP, Mulcahy DL (1999) The effects of regeneration by fragmentation upon clonal diversity in the tropical forest shrub *Poikilacanthus macranthus*: random amplified polymorphic DNA (RAPD) results. *Molecular Ecology*, **8**, 865-870.

Caddy JF (1967) Maturation of gametes and spawning in *Macoma balthica* (L.) *Canadian Journal of Zoology*, **45**, 955-65.

Callen DF, Thimpson AD, Shen Y, Phillips HA, Richards RI, Mulley JC, Sutherland GR (1993) Incidence and origin of "null" alleles in the (AC)n microsatellite markers. *American Journal of Human Genetics*, **52**, 922-927.

Carmichael LE, Nagy JA, Larter NC, Strobeck C (2001) Prey specialization may influence patters of gene flow in wolves in the Canadian Northwest. *Molecular Ecology*, **10**, 2787-2798.

Caro TM, Laurenson MK (1994) Ecological and genetic factors in conservation: A cautionary tale. *Science*, **263**, 485-486.

Casu M, Maltagliati F, Meloni M, Casu D, Cossu P, Binelli G, Curini-Galletti M, Castelli A (2002) Genetic structure of *Octupus vulgaris* (Mollusca: Cephalopoda) from the Mediterranean Sea as revealed by a microsatellite locus. *Italian Journal of Zoology*, **69**, 295-300.

Caughley G (1994) Directions in conservation biology. *Journal of Animal Ecology*, **63**, 215-244.

Chakraborty R, Kimmel M, Stivers DN, Davidson J, Deka R (1997) Relative mutation rates at di- tri- and tetrasnucleotide microsatellite loci. *Proceedings of the National Academy of Sciences USA*, **94**, 1041-1046.

Chakraborty R, Leimar O (1987) Genetic variation within a subdivided population. In: *Population genetics and fishery management*. Eds Ryman N, Utter F, University of Washington Press, Seattle, USA.

Chakraborty R, Nei M (1977) Bottleneck effects on average heterozygosity and genetic distance with the stepwise mutation model. *Evolution*, **31**, 347-356.

Chaplin JA, Hebert PDN (1997) *Cyprinotus incongruens* (Ostracoda): An ancient asexual? *Molecular Ecology*, **6**, 155-168.

Charlesworth D, Charlesworth B (1987) Inbreeding depression and its evolutionary consequences. *Annual Review of Ecology and Systematics*, **18**, 237-268.

Collins SL (1995) The measurement of stability in grassland. *Trends in Ecology and Evolution*, **10**, 95-6.

Comes HP, Abbott RJ (2000) Random amplified polymorphic DNA (RAPD) and quantitative trait analyses across a major phylogeographical break in the Mediterranean ragwort *Senecio gallicus* Vill. (Asteraceae). *Molecular Ecology*, **9**, 61-76.

Connor DW, Braizier DP, Hill TO, Northen K (1997) Marine biotope classification for Britain and Ireland: Volume 1. Littoral biotopes Version 97.06. *JNCC Report* No. **229**. Peterborough, Joint Nature Conservation Committee.

Cornuet JM, Luikart G (1996) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics*, **144**, 2001-2014.

Covey R (1999) *The saline lagoon survey of Scotland.* In: Scotland's living coastline. Eds Baxter JM, Duncan K, Atkins S, Lees G, HMSO, London, UK.

Cox CB, Moore PD (2000) *Biogeography: An ecological and evolutionary approach.* Blackwell Science, Oxford, UK.

Crease TJ, Lynch M, Spitze K (1990) Hierarchical analysis of population genetic variation in mitochondrial and nuclear genes of *Daphnia pulex*. *Molecular Biology and Evolution*, **7**, 444-459.

Crow JF, Kimura M (1970) *An Introduction to Population Genetic Theory*. Harper and Rowe, New York, USA.

Dallas JF (1992) Estimation of microsatellite mutation rates in recombinant inbred strains of mouse. *Mammalian Genome*, **3**, 452-456.

Delidow BC, Lynch JP, Peluso JJ, White BA (1993) Polymerase Chain Reaction, Basic Protocols. In: *PCR Protocols Current Methods and Applications*. Ed White BA, Humana Press, New Jersey, USA.

Dias PC (1996) Sources and sinks in population biology. *Trends in Ecology and Evolution*, **11**, 326-330.

Dillon RT, Manzi JJ (1992) Population genetics of the hard clam, *Mercenaria mercenaria*, at the northern limit of its range. *Canadian Journal of Fisheries and Aquatic Science*, **49**, 2574-2578.

DiRienzo A, Peterson AC, Garza JC, Valdes AM, Slatkin M, Freimer NB (1994) Mutational processes of simple-sequence repeat loci in human populations. *Proceedings of the National Academy of Sciences of the USA*, **91**, 3166-3170,

Downie AJ (1996) Saline lagoons and lagoon-like saline ponds in England. *Peterborough, English Nature Science Series*, No. **29**.

Doyle JJ, Dickson EE (1987) Preservation of plant samples for DNA restriction endonuclease analysis. *Taxon*, **36**, 715-722.

Dufresne F, Hebert PDN (1994) Hybridization and the origins of polyploidy. *Proceedings of the Royal Society of London Series B- Biological Sciences*, **258**, 141-146.

Eldridge MDB, King JM, Loupis AK, Spencer PBS, Taylor AC, Pope LC, Hall GP (1999) Unprecedented low levels of genetic variation and inbreeding depression in an island population of the black-footed rock wallaby. *Conservation Biology*, **13**, 531-541.

Estoup A, Angers B (1998) Microsatellites and minisatellites for molecular ecology: Theoretical and empirical considerations. In: *Advances in Molecular Ecology.* Ed Carvalho GR, IOS Press, Oxford, UK.

Estoup A, Cornuet J-M (1999) Microsatellite evolution: inferences from population data. In: *Microsatellites Evolution and Applications*. Eds Goldstein DB, Schlötterer C, OUP, Oxford, UK.

Estoup A, Tailliez C, Cornuet J-M, Solignac M (1995) Size homoplasy and mutational processes of interrupted microsatellites in two bee species, *Apis mellifera and Bombus terrestris* (Apidae). *Molecular Biology and Evolution*, **12**, 1074-1084,

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*, **86**, 991-1000.

Falconer DS (1981) *Introduction to Quantitative Genetics*. Longman, London, UK.

Figuerola J, Green AJ (2002) Dispersal of aquatic organisms by waterbirds: a review of past research and priorities for future studies. *Freshwater Biology*, **47**, 483-494.

Fish JD (1975) Development, hatching and brood size in *Bathyporeia pilosa* and *B. pelagica* (Crustacea: Amphipoda). *Journal of the Marine Biological Association of the United Kingdom*, **55**, 357-368.

Forbes SH, Hogg JT (1999) Assessing population structure at high levels of differentiation: microsatellite comparisons of bighorn sheep and large carnivores. *Animal Conservation*, **2**, 223-233.

Forbes SH, Hogg JT, Buchanan FC, Crawford AM, Allendorf FW (1995) Microsatellite evolution in congeneric mammals - Domestic and bighorn sheep. *Molecular Ecology*, **12**, 1106-1113.

Forrester GE (1990) Factors influencing the juvenile demography of a coral reef fish. *Ecology*, **71**, 1666-1681.

Frank PG, Bleakney JS (1978) Asexual reproduction, diet and anomalies of the anemone *Nematostella vectensis* in Nova Scotia. *Canadian Field Naturalist*, **92**, 259-263.

Frankel OH & Soulé ME (1981) *Conservation and evolution*. CUP, Cambridge, UK.

Frankham R (1995a) Conservation genetics. *Annual Review of Genetics*, **29**, 305-327.

Frankham R (1995b) Inbreeding and Extinction: a threshold effect. *Conservation Biology*, **9**, 792-799.

Frankham R (1997) Do Island populations have less genetic variation than mainland populations? *Heredity*, **78**, 311-327.

Frankham R (1998) Inbreeding and extinction: island populations. *Conservation Biology* **12**, 665-675.

Frankham R, Ralls K (1998) Inbreeding leads to extinction. *Nature*, **392**, 441-442.

Frankham R (1999) Quantitative genetics in conservation biology. *Genetic Research*, **74**, 237–244.

Frankham R, Balloux JD, Briscoe DA (2002) *Introduction to Conservation Genetics*. CUP, Cambridge, UK.

Freeland JR, Noble LR, Okamura B (2000) Genetic diversity of North American populations of *Cristatella mucedo*, inferred from microsatellite and mitochondrial DNA. *Molecular Ecology*, **9**, 1375-1389.

Fritsch P, Rieseberg LH (1992) Outcrossing rates are high in androdiocious populations of the flowering plant *Datisca glomerata*. *Nature*, **359**, 633-636.

Fritsch P, Rieseberg LH (1996) The use of Random Amplified Polymorphic DNA (RAPD) in Conservation Genetics. In: *Molecular Genetic Approaches in Conservation*. Eds Smith TB, Wayne RK, OUP, Oxford, UK.

Gabrielsen TM, Brochmann C (1998) Sex after all: High levels of diversity detected in the arctic clonal plant *Saxifraga cernua* using RAPD markers. *Molecular Ecology*, **7**, 1701-1708.

Gaggiotti OE, Lange O, Rassman K, Gliddon C (1999) A comparison of two indirect methods for estimating average levels of gene flow using microsatellite data. *Molecular Ecology*, **8**, 1513-1520.

Garza JC, Slatkin M, Freimer NB (1995) Microsatellite allele frequencies in humans and chimpanzees, with implications for constraints on allele size. *Molecular Biology and Evolution*, **12**, 594-603.

Garza JC, Freimer NB (1996) Homoplasy for size at microsatellite loci in humans and chimpanzees. *Genome Research*, **6**, 211-217.

Gravrilets S, Hastings A (1996) Founder effect speciation: A theoretical reassessment. *American Naturalist*, **147**, 466-491.

Gerlach G, Musolf K (2000) Fragmentation of landscape as a cause for genetic subdivision in bank voles. *Conservation Biology*, **14**, 1066-1074.

Ginot F, Bordelais I, Nguyen S, Gyapay G (1996) Correction of some genotyping errors in automated fluorescent microsatellite analysis by enzymatic removal of one base overhangs. *Nucleic Acids Research*, **24**, 540-541.

Goodman SJ (1997) Rstcalc, a collection of computer programmes for calculating estimates of genetic differentiation from microsatellite data and determining their significance. *Molecular Ecology*, **6**, 881-885.

Goropashnaya AV, Seppa P, Pamilo P (2001) Social and genetic characteristics of geographically isolated populations in the ant *Formica cinerea*. *Molecular Ecology*, **10**, 2807-2818.

Gosling E (1980) Gene frequency changes and adaptations in marine cockles. *Nature*, **286**, 601-602.

Gottelli D, Sillero-Zubiri C, Applebaum GD, Roy MS, Girman DJ, Garcia-Moreno J, Ostrander EA, Wayne RK (1994). Molecular genetics of the most endangered canid: The Ethiopian Wolf, *Canis simensis*. *Molecular Ecology*, **3**, 301-312.

Goudet J (1995) FSTAT (Version 1.2): a computer programme to calculate *F*-statistics. *Journal of Heredity*, **86**, 485-486.

Goudet J, Raymond M, DeMeeüs T, Rousset F (1996) Testing differentiation in diploid populations. *Genetics*, **144**, 1933-1940.

Grant WS, Schneider AC, Leslie RW, Cherry MI (1992) Population genetics of the brown mussel *Perna perna* in southern Africa.. *Journal of Experimental Marine Biology and Ecology*, **165**, 45-58.

Grosberg RK, Leviton, DR, Cameron BB (1996) Characterization of genetic structure and genealogies using RAPD-PCR markers: A random primer for the novice and nervous. In: *Molecular Zoology, Advances, strategies and protocols*. Eds Ferraris JD, Palumbi SR, Wiley-Liss, New York, USA.

Guo S, Thompson E (1992) Performing the exact test of Hardy-Weinberg proportion for multiple alleles. *Biometrics,* **48**, 361-372.

Hadrys H, Balick M, Schierwater B (1992) Applications of random amplified polymorphic DNA (RAPD) in molecular ecology. *Molecular Ecology*, **1**, 55-63.

Hadrys H, Schierwater B, Dellaporta SL, Desalle R, Buss LW (1993) Determination of paternity in dragonflies by random amplified polymorphic DNA fingerprinting. *Molecular Ecology*, **2**, 79-87.

Halley JM & Manasse RS (1993) A population dynamics model for annual plants subject to inbreeding depression. *Evolutionary Ecology*, **7**, 15-24.

Hamilton WD, May RM (1977) Dispersal in stable habitats. *Nature*, **269**, 578-581.

Hancock JM (1995) The contribution of slippage like processes to genome evolution. *Journal of Molecular Evolution*, **41**, 1038-47.

Hancock JM (1998) Microsatellites and other simple sequences: genomic context and mutational mechanisms. In: *Microsatellites Evolution and Applications*. Eds Goldstein DB, Schlötterer C, OUP, Oxford, UK.

Hand C, Uhlinger KR (1992) The culture, sexual and asexual reproduction, and growth of the sea anemone *Nematostella vectensis*. *Biological Bulletin*, **182**, 169-176.

Hand C, Uhlinger KR (1994) The unique widely distributed, estuarine sea anemone, *Nematostella vectensis* Stephenson: A review, new facts, and questions. *Estuaries*, **17**, 501-508.

Hand C, Uhlinger KR (1995) Asexual reproduction by transverse fission and some anomalies in the sea anemone *Nematostella vectensis*. *Invertebrate Biology*, **114**, 9-18.

Harter VL (1997) *The use of* Nematostella vectensis *Stephenson in aquatic toxicity bioassays: Setting the ground work*. MSc thesis, University of Western Washington, USA.

Hartl DL (2000) *A primer of population genetics*. Sinauer Associates, Sunderland, Massachusetts, USA.

Hartl DL, Clark AG (1997) *Principles of population genetics*. Sinauer Associates, Sunderland, Massachusetts, USA.

Hartl GB, Marowski J, Swiatecki A, Janiszewski T, Willing R (1992) Genetic diversity in the polish brown hare *Lepus europaeus* Pallas 1778: implications for conservation and management. *Acta Theriologica*, **37**, 15-25.

Hedgecock D (1986) Is gene flow from pelagic larval dispersal important in the adaptation and evolution of marine invertebrates? *Bulletin of Marine Science*, **39**, 550-564.

Hedrick PW (1994) Purging inbreeding depression and the probability of extinction: full-sib mating. *Heredity*, **73**, 363-372.

Hedrick PW (1999) Perspective: highly variable loci and their interpretation in evolution and conservation. *Evolution*, **53**, 313-318.

Hedrick PW, Ginevan, ME, Ewing EP (1976) Genetic polymorphism in heterogeneous environments. *Annual Review of Ecology and Systematics*, **7**, 1-32.

Hedrick PW, Hedgecock D, Hamelberg S (1995) Effective population size in winter run chinook salmon. *Conservation Biology*, **9**, 615-624.

Hedrick PW, Lacy RC, Allendorf FW, Soulé ME (1996) Directions in conservation biology: comments on Caughley. *Conservation Biology*, **10**, 1312-1320.

Hedrick PW, Miller PS (1992) Conservation genetics - techniques and fundamentals. *Ecological Applications*, **2**, 30-46.

Herbert PDN (1987) Genotypic characteristics of cyclical parthenogens and their obligatory asexual derivatives. In: *The Evolution of Sex and its Consequences*. Ed Stearns SC, Birkhauser Verlag, Basel, Switzerland.

Herbert PDN, Finston TL (1996) Genetic differentiation in *Daphnia obtusa*: a continental perspective. *Freshwater Biology*, **35**, 311-321.

Higgins K, Lynch M (2001) Metapopulation extinction caused by mutation accumulation. *Proceedings of the National Academy of Science, USA*, **98**, 2928-2933.

Hoffman RJ (1987) Short-term stability of genetic structure in populations of the sea anemone *Metridium senile*. *Marine Biology*, **93**, 499-507.

Holt RD, Gaines MS (1992) Analysis of adaptation in heterogeneous landscapes: implications for the evolution of fundamental niches. *Evolutionary Ecology*, **6**, 433-447.

Høpner-Petersen G (1958) Notes on the growth and biology of the different *Cardium* species in Danish brackish water areas. *Meddr Danm Fisk-og Havenders*, **2**, 1-31.

Huang BX, Peakall R, Hanna PJ (2000) Analysis of genetic structure of blacklip abalone (*Haliotis rubra*) populations using RAPD, minisaatellite and microsatellite markers. *Marine Biology*, **136**, 207-216.

Huff DR, Peakall R, Smouse PE (1993) RAPD variation within and among natural populations of outcrossing buffalograss (*Buchlöe dactyloides* (Nutt.) Engelm.). *Theoretical and Applied Genetics*, **86**, 927-934.

Huff DR, Quinn JA, Higgins B, Palazzo AJ (1998) Random Amplified Polymorphic DNA (RAPD) variation among native bluestem [*Schizachyrium scoparium* (Michx.) Nash] populations from sites of high and low fertility in forest and grassland biomes. *Molecular Ecology*, **7**, 1591-1597.

Hughes RN (1989). *A functional biology of clonal animals*. Chapman & Hall, London, UK.

Hughes CR, Queller DC (1993) Detection of highly polymorphic microsatellite loci in a species with little allozyme polymorphism. *Molecular Ecology*, **2**, 131-137.

Hummel H, Wolowicz M, Bogaards RH (1994) Genetic variability and relationships for populations of *Cerastoderma edule* and of the *C. glaucum* complex. *Netherlands Journal of Sea Research*, **33**, 81-89

Hynes HBN (1954) The ecology of *Gammarus dubeni* Lilljeborg and its occurrence in freshwater in western Britain. *Journal of Animal Ecology*, **23**, 38-84.

Innes D (1991) Geographic patterns of genetic differentiation among sexual populations of *Daphnia pulex*. *Canadian Journal of Zoology*, **69**, 995-1003.

Innes D, Schwartz SS, Herbert PDN (1986) Genotypic diversity and variation in mode of reproduction among populations in the *Daphnia pulex* group. *Heredity*, **77**, 282-291.

International Human Genome Consortium (2001) Initial sequencing and analysis of the human genome. *Nature* **409**, 860-921.

Ivell R (1979a) The biology and ecology of a brackish lagoon bivalve, *Cerastoderma glaucum*, Bruguèire, in Lago Lungo, Italy. *Journal of Molluscan Studies*, **45**, 364-382.

Ivell R (1979b) The biology and ecology of a brackish lagoon bivalve, *Cerastoderma glaucum*, Bruguèire, in an English lagoon, the Widewater, Sussex. *Journal of Molluscan Studies*, **45**, 383-400.

Jackson JBC, Coates AG (1986) Life cycles and evolution of clonal (modular) animals. *Philosophical Transactions of the Royal Society of London, series B*, **313**, 7-22.

Jaenike JR (1973) A steady state model of genetic polymorphism on islands. *American Naturalist*, **107**, 793-795.

Jarne P, Lagoda PJL (1996) Microsatellites, from molecules to populations and back. *Trends in Ecology and Evolution*, **11**, 424- 429.

Jarvinen O, Varvio S (1985) Proness to extinction of small populations of seals: demographic and genetic stochasticity vs environmental stress. *Finnish Game Research*, **44**, 6-18.

Jeffreys AJ, Wilson V, Thein L (1985) Hypervariable "Minisatellite" regions in human DNA. *Nature,* **314**, 67-73

Jelnes JE, Petersen GH, Russell PJC (1971) Isoenzyme taxonomy applied on four species of *Cardium* from Danish and British waters with a short description of the distribution of the species (Bivalvia). *Ophelia*, **9**, 15-19.

Jenkins DG, Buikema ALJ (1998) Do similar communities develop in similar sites? A test with zooplankton structure and function. *Ecolocgical Monographs*, **68**, 421-443.

Johnson MS, Black R (1998) Increased genetic divergence and reduced genetic variation in populations of the snail *Bembicium vittatum* in isolated tidal ponds. *Heredity*, **80**, 163-172.

Jorde PE, Ryman N (1995) Temporal allele frequency change and estimation of effective size in populations with overlapping generations. *Genetics*, **139**, 1077-1090.

Kambhampati S, Black WC, Rai IV, Rai KS (1992) Random amplified polymorphic DNA of mosquito species and populations (Diptera: Culicidae): Techniques, statistical analysis and applications. *Journal of Medical Entomology*, **29**, 939-945.

Kandpal RP, Kandpal G, Weissman SM (1994) Construction of libraries enriched for sequence repeats and jumping clones, and hybridization selection for region-specific markers. *Proceedings of the National Academy of Sciences of the USA*, **91**, 88-92.

Karakiri M, Nicolaidou A (1987) Populations studies on the Amphipoda of Mazoma lagoon (Greece). *Helgolander Meeresunters*, **41**, 453-464.

Kashi Y, Soller M (1998) Functional roles of microsatellites and minisatellites. In: *Microsatellites Evolution and Applications*. Eds Goldstein DB, Schlötterer C, OUP, Oxford, UK.

Kawecki TJ (1995) Demography of source-sink populations and the evolution of ecological niches. *Evolutionary Ecology*, **9**, 38-44.

Kawecki TJ (2000) Adaptation to marginal habitats: contrasting influence of the dispersal rate on the fate of alleles with small and large effects. *Proceedings of the Royal Society of London Series B*, **267**, 1315- 1320.

Keller LF, Waller DM (2002) Inbreeding effects in wild populations. *Trends in Ecology and Evolution*, **17**, 230-241.

Kikuchi T (1980) Faunal relationships in temperate seagrass beds. In: *Handbook of Seagrass Biology An Ecosystem Perspective*. Eds Phillips RC, McRoy CP, Garland STPM, London, UK.

Kimura M (1968) Evolutionary rate at the molecular level. *Nature*, **217**, 624-626.

Kimura M, Crow JF (1964) The number of alleles that can be maintained in a finite population. *Genetics*, **49**, 725-738.

Kimura M, Otha T (1978) Stepwise mutation model and distribution of allelic frequencies in a finite population. *Proceedings of the National Academy of Sciences of the USA*, **75**, 2868-2872.

Kingston PF (1974a) Studies on the reproductive cycle of *Cardium edule* and *C. glaucum*. *Marine Biology*, **28**, 317-323.

Kingston PF (1974b) Some observations on the effects of temperature and salinity upon the growth of *Cardium edule* and *Cardium glaucum* larvae in the laboratory. *Journal of the Marine Biological Association of the UK*, **54**, 309-317.

Koehn RK, Bayne BL, Moore MN, Siebenaller JF (1980) Salinity related physiological and genetic differences between populations of *Mytilus edulis*. *Biological Journal of the Linnean Society*, **14**, 319-334.

Krebs CJ (1994) *Ecology*. Harper Collins, England, UK.

Lacy RC (1997) Importance of genetic variation to the viability of mammalian populations. *Journal of Mammalogy*, **78**, 320-335.

Lande R (1988) Genetics and demography in biological conservation. *Science*, **241**, 1455-1460.

Lande R (1993) Risks of population extinction from demographic and environmental stochasticity and random catastrophes. *American Naturalist*, **142**, 911-927.

Lenormand T (2002) Gene flow and the limits to natural selection. *Trends in Ecology & Evolution*, **17**, 183-189.

Lehmann T, Hawley WA, Grebert H, Danga M, Atielli F & Collins FH (1999) The rift valley complex as a barrier to gene flow for *Anopheles gambiae* in Kenya. *The Journal of Heredity*, **90**, 613-621.

Lugon-Moulin N, Brünner H, Wyttenbach A, Hausser J, Goudet J (1999) Hierarchical analysis of genetic differentiation in a hybrid zone of *Sorex araneus* (Insectivora, Soricidae). *Molecular Ecology*, **8**, 419-431.

Luikart G, England PR (1999) Statistical analysis of microsatellite DNA data. *Trends in Ecology and Evolution*, **14**, 253-256.

Lynch M (1984a) Destabilizing hybridization, general-purpose genotypes and geographic parthenogenesis. *Quarterly Review of Biology*, **59**, 257-290.

Lynch M (1984b) The limits to life history evolution in *Daphnia*. *Evolution*, **38**, 465-82.

Lynch M (1997) Mutation accumulation in nuclear, organelle and prokaryotic transfer RNA genes. *Molecular Biology and Evolution*, **14**, 914-925.

Lynch M, Bürger R, Butcher D, Gabriel W (1993) The mutational meltdown in asexual populations. *Journal of Heredity*, **84**, 257-290.

MacArthur RH, Wilson EO (1967) *The theory of Island Biogeography*. Princeton University Press, Princeton, USA.

Mace GM, Smith TB, Bruford MW, Wayne RK (1996) An Overview of the Issues. In: *Molecular Genetic Approaches in Conservation*. Eds Smith TB, Wayne RK, OUP, Oxford, UK.

MacHugh DE, Loftus RT, Bradley DG, Sharp PM Cunningham P (1994) Microsatellite DNA variation within and among European cattle breeds. *Proceedings of the Royal. Society of London, series B*, **256**, 25-31.

Madsen T, Stille B, Shine R (1996) Inbreeding depression in an isolated population of adders *Vipera berus*. *Biological Conservation*, **75**, 113-118.

Mars P (1951) Essai d'interprétation des formes généralment groupeés sous le nom de *Cardium edule* Linné. *Bulletin du Muséum d'Histoire naturelle, Marseille,* **11**, 2-31.

Marsh GA (1973) The *Zostera* epifaunal community in the York River, Virginia, Chesapeake. *Science*, **14**, 87-97.

Maxam AM, Gilbert W (1977) A new method for sequencing DNA. *Proceedings of the National Academy of Sciences USA,* **74**, 560-564.

May RM (1979) When to be incestuous. *Nature*, **279**, 192-194.

Maynard Smith J (1970) Genetic polymorphism in a varied environment. *American Naturalist*, **104**, 487-490.

McPhial JD, Carveth R (1993) *A foundation for conservation: the nature and origin of the freshwater fish fauna of British Columbia*. BC Environment, Victoria, Canada.

Milligan BG, Leebensmack J, Strand AE (1994) Conservation genetics - beyond the maintenance of marker diversity. *Molecular Ecology*, **3**, 423-435.

Mills LS, Allendorf FW (1996) The one migrant per generation rule in conservation and management. *Conservation Biology*, **10**, 1509-1518.

Mills LS & Smouse PE (1994) Demographic consequences of inbreeding in remnant populations. *American Naturalist*, **144**, 412-431.

Moore SS, Sargeant LL, King TJ, Mattick JS, Georges M, Hetzel DJS (1991) The conservation of dinucleotide microsatellites among mammalian genomes allows the use of heterologous PCR primer pairs in closely related species. *Genomics*, **10**, 654-660.
Moxon ER, Wills C (1999) DNA Microsatellites: Agents of evolution? *Scientific American*, **280**, 94-99.

Muller HJ (1964) The relevance of mutation to mutational advance. *Mutation Research*, **1**, 2-9

Mullis KB, Faloona FA (1987) Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. *Methods in Enzymology*, **155**, 335-350.

Muus K (1966) Notes on the biology of *Protohydra leuckarti* Greef (Hydroidea: Protohydridae). *Ophelia*, **3**, 141-50.

Nauta MJ, Weissing FJ (1996) Constraints on allele size at microsatellite loci: implications for genetic differentiation. *Genetics*, **143**, 1021-1032.

Nei M (1977) *F*-statistics and analysis of gene diversity in sub-divided populations. *Annals of Human Genetics*, **41**, 255-233.

Nei M, Chakravarti A, Tateno Y (1977) Mean and variance of $F_{ST}$ in a finite number of incompletely isolated populations. *Theoretical Population Biology*, **11**, 291-306.

Nei M, Maruyama T, Chakraborty R (1975) The bottleneck effect and genetic variability in populations. *Evolution*, **29**, 1-10.

Nelson WG (1980) Reproductive patterns of gammaridean amphipods. *Sarsia*, **65**, 61-71.

Neraas LP, Spruell P (2001) Fragmentation of riverine systems: the genetic effects of dams on bull trout (*Salvelinus confluentus*) in the Clark Fork River system. *Molecular Ecology*, **10**, 1153-1164

Newman D, Pilson D (1997) Increased probability of extinction due to decreased genetic effective population size: experimental populations of *Clarkia pulchella*. *Evolution*, **51**, 354-362.

Nielson JL, Crow KD, Fountain MC (1999) Microsatellite diversity and conservation of a relict trout population: McCloud River redband trout. *Molecular Ecology*, **8**, 129-142.

Nunney L, Campbell KA (1993) Assessing minimum viable population size: Demography meets population genetics. *Trends in Evolution and Ecology*, **8**, 234-239.

O'Mullan GD, Maas PAY, Lutz RA, Vrijenhoek RC (2001) A hybrid zone between hydrothermal vent mussels (Bivalvia: Mytilidae) from the Mid Atlantic Ridge. *Molecular Ecology*, **10**, 2819-2831.

Paetkau D & Strobeck C (1994) Microsatellite analysis of genetic variation in black bear populations. *Molecular Ecology* **3**, 489-495.

Paetkau D, Strobeck C (1995) The molecular basis and evolutionary history of a microsatellite null allele in bears. *Molecular Ecology*, **4**, 519-20.

Pàlsson S (1999) *On the evolutionary effects of linkage and deleterious mutations in small populations*. PhD thesis. Uppsala University, Sweden.

Pàlsson S (2000) Microsatellite variation in *Daphnia pulex* from both sides of the Baltic Sea. *Molecular Ecology*, **9**, 1075-1088.

Pàlsson S, Palmilo P (1999) The effect of deleterious mutations on linked neutral variation in small populations. *Genetics*, **153**, 475-483.

Palumbi S (1994) Genetic divergence, reproductive isolation and marine speciation. *Annual Review of Ecology and Systematics*, **25**, 547-572.

Palumbi S, Martin A, Romano S, McMillan WO, Stice L, Grabowski, G (1991) *The Simple Fool's Guide To PCR*. Kewalo Marine Laboratory, University of Hawaii.

PE Applied Biosystems (1998) ABI Prism® dGTP BigDye™ Terminator Ready Reaction Kit Protocol. Perkin Elmer Corporation, USA.

Pearson CVM, Rogers AD, Sheader M (2002) The genetic structure of the rare lagoonal sea anemone, *Nematostella vectensis* Stephenson (Cnidaria; Anthozoa) in the UK based on RAPD analysis. *Molecular Ecology*, **11**, 2285-2293.

Pèpin L, Amigues Y, Lèpingle A, Berthier JL, Bernsaid A, Vaiman D (1995) Sequences conservation of microsatellites between cattle (*Bos taurus*), goat (*Capra hircus*) and related species. Examples of use in parentage testing and phylogeny analysis. *Heredity*, **74**, 53-61.

Pielou EC (1979) *Biogeography*. Wiley, New York, USA.

Porter HJ (1964) Seasonal gonadal changes in the adult clams *Mercenaria mercenaria* (L) in North Carolina. *Proceedings of the National Shellfish Association*, **55**, 35-52.

Preston A, Moore PG (1988) The flora and fauna associated with *Cladophora albida* (Huds.) Kütz from rock pools on Great Cumbrae Island, Scotland. *Ophelia*, **29**, 169-186.

Primmer CR, Ellegren H, Saino N, Moller AP (1996) Directional evolution in germline microsatellite mutations. *Nature Genetics*, **13**, 391-393.

Primmer CR, Saino N, Moller AP, Ellegren H (1998) Unravelling the processes of microsatellite evolution through analysis of germline mutations in barn swallows, *Hirundo rustica*. *Molecular Biology and Evolution*, **15**, 1047-1054.

Prober JM, Trainor GL, Dam RJ, Hobbs FW, Robertson CW, Zagursky RJ, Cocuzza AJ, Jensen MA, Baumeister K (1987) A system for rapid DNA sequencing with fluorescent chain-terminating dideoxynucleotides. *Science*, **238**, 336-341.

Queller DC, Strassmann JE, Hughes CR (1993) Microsatellites and Kinship. *Trends in Ecology and Evolution*, **8**, 285-288.

Rabalais NN, Gore RH (1985) Abbreviated development in decapods. In: *Larval growth*. Ed Wenner A, AA Balkema, Rotterdam, Netherlands.

Rawson PD, Burton RS (2002) Functional coadaption between cytochrome *c* and cytochrome *c* oxidase within allopatric populations of a marine copepod. *PNAS*, **99**, 12955-12958.

Raymond M, Rousset F (1995) An exact test for population differentiation. *Evolution*, **49**, 1280-1283.

Reynolds J, Weir BS, Cockerham CC (1983) Estimation of the coancestory coefficient: basis for a short-term genetic distance. *Genetics*, **105**, 767-779.

Rice WR (1989) Analyzing tables of statistical tests. *Evolution*, **43**, 223-225.

Rousset F (1996) Equilibrium values of measures of population subdivision for stepwise mutation processes. *Genetics*, **142**, 1357-1362.

Rousset F (1997) Genetic differentiation and estimation of gene flow from *F*-statistics under isolation by distance. *Genetics*, **145**, 1219-1228.

Roy MS, Geffen E, Smith D, Ostrander O, Wayne RK (1994) Patterns of differentiation and hybridization in North American wolflike canids, revealed by analysis of microsatellite loci. *Molecular Biology and Evolution*, **11**, 553-570.

Rozen S, Skaletsky HJ (1998) *Primer 3*. Code available at http://www-genome.wi.mit.edu/genome-software/other/primer3.html

Russell PJC (1972) Biological studies on *Cardium glaucum*, based on some Baltic and Mediterranean populations. *Marine Biology*, **16**, 290-296.

Rychlik, R (1993) Selection of Primers for Polymerase Chain Reaction. In: *PCR Protocols Current Methods and Applications*. Ed White BA, Humana Press, New Jersey, USA.

Saavedra C, Zapata C, Guerra A, Alvarez G (1993) Allozyme variation in European populations of the oyster *Ostrea edulis*. *Marine Biology*, **115**, 85-95.

Saccheri I, Brakefield PM, Nichols RA (1998) Inbreeding and extinction in a butterfly metapopulation. *Nature*, **392**, 491-494.

Saiki RK, Gelfland DH, Stoffel S, Scharf SJ, Higuchi R, Horn GT, Mullis KB, Erlich HA (1988). Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science*, **239**, 487-491.

Sainte-Marie B (1991) A review of the reproductive bionomics of aquatic gammaridean amphipods: variation of life history traits with latitude, depth, salinity and superfamily. *Hydrobiologia*, **223**, 189-227.

Samadi S, Erard F, Estoup A, Jarne P (1998) The influence of mutation, selection and reproductive systems on microsatellite variability: a simulation approach. *Genetical Research Cambridge*, **71**, 213-222.

Samadi S, Mavárez J, Pointier P, Delay B, Jarne P (1999) Microsatellite and morphological analysis of population structure in the parthenogenetic freshwater snail *Melanoides tuberculata*: insights into the creation of clonal variability. *Molecular Ecology*, **8**, 1141-1153.

Sambrook J, Fritsch E, Maniatis T (1989) *Molecular Cloning: a Laboratory Manual.* Cold Spring Harbor Laboratory Press, New York, USA.

Sanger F, Coulson AR (1975) A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *Journal of Molecular Biology,* **94**, 441-448.

Sarver SK, Landrum MC, Foltz DW (1992) Genetics and taxonomy of ribbed mussels (*Geukensia* spp.). *Marine Biology*, **113**, 385-390.

Scher S (1996) Genetic structure of natural *Taxus* populations in western North America. In: *Molecular Genetic Approaches in Conservation.* Eds Smith TB, Wayne RK, OUP, Oxford, UK.

Schlötterer C, Tautz D (1992) Slippage synthesis of microsatellites. *Nucleic Acids Research,* **20**, 211-215

Schneider S, Kueffer JM, Roessli D, Excoffier L (2000) *ARLEQUIN, version 2.000: A Software for Population Genetic Data Analysis.* Genetics and Biometry Laboratory, University of Geneva, Geneva.

Schoener A (1974) Experimental zoogeography: Colonization of marine mini-islands. *American Naturalist*, **108**, 715-738.

Schoener TW, Toft CA (1983) Spider populations: Extraordinarily high densities on islands without top predators. *Science*, **219**, 1353-1355.

Schug MD, Mackay TFC, Aquadro CF (1997) Low mutation rates of microsatellite loci in Drosophila melanogaster. *Nature Genetics*, **15**, 99-102.

Sebens KP (1981) The allometry of feeding, energetics and body size in three species of sea anemone. *Biological Bulletin*, **161**, 152-171.

Sheader M (1978) Distribution and reproductive biology of *Corophium insidiosum* (Amphipoda) on the north-east coast of England. *Journal of the Marine Biological Association of the UK*, **58**, 585-596.

Sheader M (1983) The reproductive biology and ecology of *Gammarus duebeni* (Crustacea : Amphipoda) in southern England. *Journal of the Marine Biological Association of the UK*, **63**, 517-540.

Sheader M, Sheader A (1985a) Survey of brackish coastal lagoons. Sussex to Dorset, 1984-85. *Peterborough, Nature Conservancy Council CSD Report*, No. **739**.

Sheader M, Sheader A (1985b) New distribution records for *Gammarus insensibilis* Stock, 1966, in Britain. *Crustaceana*, **49**, 101-105

Sheader M, Sheader A (1989) Coastal saline ponds of England and Wales an overview. *Peterborough, Nature Conservancy Council CSD Report*, No. **1009**.

Sheader M (1996) Factors influencing egg size in the gammarid amphipod *Gammarus insensibilis*. *Marine Biology*, **124**, 519-526.

Sheader M, Suwailem AM, Rowe GA (1997) The anemone, *Nematostella vectensis*, in Britain: Considerations for conservation management. *Aquatic Conservation*, **7**, 13-25.

Sheffield VC, Weber JL, Buetox KH, Murray JC, Even DA, Wiles K, Gastier JM, Pulido JC, Yandava C, Sunden SL, Mattes G, Businga T, McClain A, Beck J, Scherpier T, Gilliam J, Zhong J, Duyk GM (1995) A collection of tri- and tetranucleotide repeat

markers used to generate high quality, high resolution genome-wide linkage maps. *Human Molecular Genetics*, **4**, 1837-1844.

Shick JM, Lamb AN (1977) Asexual reproduction and genetic population structure in the colonizing sea anemone *Haliplanella luciae*. *Biological Bulletin*, **153**, 604-617.

Shick JM (1989) *A functional biology of sea anemones*. Croom Helm, London, UK.

Shriver MD, Jin L, Chakraborty R, Boerwinkle E (1993) VNTR allele frequency distributions under the stepwise mutation model: A computer simulation approach. *Genetics*, **13,** 983-993.

Sih A, Gunnar Jonsson B, Luikart G (2000) Habitat loss: ecological, evolutionary and genetic consequences. *Trends in Ecology and Evolution*, **15**, 132-134.

Sjögren P, Wyöni PI (1994) Conservation genetics and the detection of rare alleles in finite populations. *Conservation Biology*, **8**, 267-270.

Skadsheim A (1984) Coexistence and reproductive adaptations of amphipods: the role of environmental heterogeneity. *Oikos*, **43**, 94-103.

Skibinsky DOF, Beardmore JA, Cross TF (1983) Aspects of the population genetics of *Mytilus* (Mytilidae: Mollusca) in the British Isles. *Biological Journal of the Linnaen Society*, **19**, 137-183.

Skroch P, Nienhuis J (1995) Impact of scoring error and reproducibility of RAPD data on RAPD based estimates of genetic distance. *Theoretic and Applied Genetics*, **91**, 1086-1091.

Slatkin M (1985) Gene flow in natural populations. *Annual Review of Ecology and Systematics*, **16**, 393-430.

Slatkin M (1987) Gene flow and the geographic structure of natural populations. *Science*, **236**, 787-792.

Slatkin M (1993) Isolation by distance in equilibrium and non-equilibrium populations. *Evolution*, **47**, 264-279.

Slatkin, M (1995) A measure of population subdivision based on microsatellite allele frequencies. *Genetics*, **139**, 457-462.

Slatkin M, Barton NH (1989) A comparison of three indirect methods for estimating average levels of gene flow. *Evolution*, **43**, 1349-1368.

Smith BL, Potts DC (1987) Clonal and solitary anemones (*Anthopleura*) of western North America: population genetics and systematics. *Marine Biology* **94**, 537-546.

Smith CT, Nelson RJ, Wood CC, Koop BF (2001) Glacial biogeography of North American coho salmon (*Oncorhynchus kisutch*). *Molecular Ecology*, **10**, 2775-2786.

Smith LM, Sanders JZ, Kaiser RJ, Hughes P, Dodd C, Connell CR, Heiner C, Kent SBH, Hood LE (1986) Fluorescence detection in automated DNA sequence analysis. *Nature*, **321**, 674-679.

Sneath PHA, Sokal RE (1973) *Numerical Taxonomy*. WH Freeman and Co., San Francisco, USA.

Sokal RE, Rohlf FJ (1981) *Biometry*. WH Freeman and Co., New York, USA.

Sommerfeldt AD, Bishop JDD (1999) Random amplified polymorphic DNA (RAPD) analysis reveals extensive natural chimerism in a marine protochordate. *Molecular Ecology*, **8**, 885-890.

Spellerberg IF (1996) Changes in biological diversity. In: *Conservation Biology*. Ed Spellerberg IF, Longman, England, UK.

Spiller DA, Schoener TW (1990) Lizards reduce food consumption by spiders: Mechanisms and consequences. *Oecologia*, **83**, 150-161.

Stacey PB, Taper M (1992) Environmental variation and persistence of small populations. *Ecological Applications*, **2**, 18-29.

Städler T, Jarne P (1997) Population biology, genetic structure and mating system parameters in freshwater snails. In: *Ecology and Evolution of Freshwater Organisms*. Eds Streit B, Städler T, Lively CM, Birkhäuser, Basel, Switzerland.

Stallings RL, Ford AF, Nelson D, Torney C, Hildebrand CE, Moyzis RK (1991) Evolution and distribution of (GT)n repetitive sequences in mammalian genomes. *Genomics*, **10**, 807-815.

Steele DH, Steele, VJ (1969) The biology of *Gammarus* (Crustacea, Amphipoda) in the northwestern Atlantic. I. *Gammarus dubeni*. Lillj. *Canandian Journal of Zoology*, **47**, 235-244.

Steele DH, Steele, VJ (1975a) Egg size and duration of embryonic development in Crustacea. *Internationale Revue der gesamten Hydrobiologie*, **60**, 711-715.

Steele DH, Steele, VJ (1975b) The biology of *Gammarus* (Crustacea, Amphipoda) in the northwestern Atlantic. XI. Comparison and discussion. *Canadian Journal of Zoology*, **52**, 1116-1126.

Steele DH, Steele VJ (1991) Morphological and environmental restraints on egg production in amphipods. In: *Crustacean egg production*. Eds Wenner A, Kuris A, AA Balkema, Rotterdam, Netherlands.

Stephens PA, Sutherland WJ (1999) Consequences of the Allee effect for behaviour, ecology and conservation. *Trends in Evolution and Ecology*, **14**, 401-405.

Stewart CN, Porter DM (1995) RAPD profiling in biological conservation – an application to estimating clonal variation in rare and endangered Iliamna in Virginia. *Biological Conservation*, **74**, 135-142.

Stockley BM (2001) *Spatial and temporal population genetics of two commercially exploited deepsea fish from the Azores Archipelago*. PhD Thesis, University of Southampton, UK.

Strand M, Prolla TA, Liskay RM, Petes TD (1993) Destabilization of tracts of simple repetitive DNA in yeast by mutations affecting DNA mismatch repair. *Nature*, **365**, 274-276.

Sukumar R (1991) The management of large mammals in relation to male strategies and conflict with people. *Biological Conservation*, **55**, 93-102.

Sunnucks P (2000) Efficient genetic markers for population biology. *Trends in Ecology & Evolution*, **15**, 199-203.

Sutcliffe DW (1992) Reproduction in *Gammarus* (Crustacea: Amphipoda): basic reproductive processes. *Freshwater Forum*, **2**, 102-128.

Sydes MA, Peakall R (1998) Extensive clonality in the endangered shrub *Haloragodendron lucasii* (Haloragaceae) revealed by allozymes and RAPDs. *Molecular Ecology*, **7**, 87-93.

Taberlet P, Griffen S, Goossens B, Questiau S, Manceau V, Escaravage N, Waits LP, Bouvet J (1996) Reliable genotyping of samples with very low DNA quantities using PCR. *Nucleic Acids Research*, **24**, 3189-3194.

Tallmon DA, Draheim HM, Scott Mills L, Allendorf FW (2002) Insights into recently fragmented vole populations from combined genetic and demographic data. *Molecular Ecology*, **11**, 699-709.

Tautz D (1989) Hypervariability of simple sequences as a general source for polymorphic DNA markers. *Nucleic Acids Research*, **17**, 6463-6471.

Templeton AR (1991) Off-site breeding of animals and implications for plant conservation strategies. In: *Genetics and conservation of rare plants*. Eds Falk DA, Holsinger KE, Oxford University Press, New York, USA.

Thomas F, Renaud F, Cezilly F (1996) Assortative pairing by parasitic prevalence in *Gammarus insensibilis* (Amphipoda): patterns and processes. *Animal Behaviour*, **52**, 683-690.

Todd CD, Bentley MG, Havenhand JN (1991) Larval metamorphosis of the opisthobranch mollusk *Adalaria proxima* (Gastropoda: Nudibranchia). The effects of choline and elevated potassium ion concentration. *Journal of the Marine Biological Association of the UK*, **71**, 53-72.

Tully O & Céidigh PÓ (1986) The ecology of *Idotea* species (Isopoda) and *Gammarus locusta* (Amphipoda) on surface driftweed in Galway Bay (west of Ireland). *Journal of the Marine Biological Association of the UK*, **66**, 931-942.

Tyler-Waters H, Hawkins AR (1995) The application of RAPD markers to the study of the bivalve mollusc *Lasaea rubra*. *Journal of the marine Biological Association of the UK*, **75**, 563-569.

Valdes AM, Slatkin M, Freimer NB (1993) Allele frequencies at microsatellite loci: the stepwise mutation model revised. *Genetics*, **133**, 737-749.

Valsecchi E, Palsboll P, Hale P, GlocknerFerrari D, Ferrari M, Clapham P, Larsen F, Mattila D, Sears R, Sigurjonsson J, Brown M, Corkeron P, Amos B (1997) Microsatellite genetic distances between oceanic populations of the humpback whale (*Megaptera novaeangliae). Molecular Biology and Evolution*, **14**, 355-362.

Van Buren R, Harper KT, Anderson WR, Stanton DJ, Seyoum S, England JL (1994) Evaluating the relationship of autumn buttercup (*Ranunculus acriformis* var. *aestivalis*) to some close congeners using random amplified polymorphic DNA. *American Journal of Botony*, **81**, 514-519.

Van Noordwijk AJ, Scharloo W (1981) Inbreeding in an island population of the great tit. *Evolution*, **35**, 674-687.

Vanoverbeke J, DeMeester L (1997) Among-populational genetic differentiation in the cyclical parthenogen *Daphnia magna* (Crustacea, Anomopoda) and its relation to geographic distance and clonal diversity. *Hydrobiologia*, **360**, 135-142.

van de Ven WTG, McNicol RJ (1995) The use of RAPD markers for the identification of Sitka spruce (*Picea sitchensis*) clones. *Heredity,* **75**, 126-132.

Venable DL, Brown JS (1988) The selective interactions of dispersal, dormancy and seed size as adaptations for reducing risk in variable environment. *American Naturalist*, **131**, 360-384.

Viard F, Franck P, Dubois M-P, Estoup A, Jarne P (1998) Variation of microsatellite size homoplasy across electromorphs, loci and populations in three invertebrate species. *Journal of Molecular Evolution*, **47**, 42-51.

Vrijenhoek RC (1985) Animal population genetics and disturbance: the effects of local extinctions and recolonizations on heterozygosity and fitness. In: *The Ecology of Natural Disturbances and Patch Dynamics*. Eds Pickett STA, White P, Academic Press, New York, USA.

Vrijenhoek RC (1998) *Clonal Organisms and the Benefits of Sex*. In: Advances in Molecular Ecology. Ed Carvalho GR, IOS Press, Oxford, UK.

Wade MJ & McCauley DE (1988) Extinction and reclonization: Their effects on the genetic differentiation of local populations. *Evolution*, **42**, 995-1005.

Wahlund (1928) Zustammensetzung von Population und Korrelationserscheinungen von Standpunkt der Vererbungslehre aus betrachtet. *Hereditas*, **11**, 65-106.

Waller DM (1993) The statics and dynamics of mating system evolution. In: *The Natural History of Inbreeding and Outbreeding*. Ed Thornhill N, University of Chicago Press, USA.

Walsh PS, Metzger DA, Higuchi R (1991) Chelex-100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. *Botechniques,* **10**, 506-513.

Waples RS (1987) A multispecies approach to the analysis of gene flow in marine shore fishes. *Evolution*, **41**, 385-400.

Waples RS (1998) Separating the wheat from the chaff: Patterns of genetic differentiation in high gene flow species. *The American Genetic Association*, **89**, 438-459.

Ward RD (1989) Molecular population genetics of marine organisms. In: *Reproduction, Genetics and Distributions of Marine Organisms*. Eds Ryland JS, Tyler PA., 23rd European Marine Biology Symposium, Olsen & Olsen, Fredensborg, Denmark.

Ward RD, Skibinski DOF, Woodwark M (1992) Protein heterozygosity, protein structure and taxonomic differentiation. *Evolutionary Biology*, **26**, 73-159.

Weber JL, Wong C (1993) Mutation of human short tandem repeats. *Human Molecular Genetics*, **2**, 1123-1128.

Weir BS, Cockerham CC (1984) Estimating *F*-statistics for the analysis of population structure. *Evolution*, **38**, 1358-1370.

Welsh J, McClelland M (1990) Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acids Research*, **18**, 7213-7218.

Weinzierl RP, Beukeboom LW, Gerace L, Michiels, NK (1999) Spatial and ecological overlap between coexisting sexual and parthenogenetic *Schmidtea polychroa* (Tricladida; Platyhelminthes). *Hydrobiologia*, **392**, 179-185.

Williams RB (1983) The Starlet sea anemone: *Nematostella vectensis*. In: *The IUCN Invertebrate Red Data Book*. Eds Wells SM, Pyle RM, Collins NM, International Union for Conservation of Nature and Natural Resources, Gland, Switzerland.

Williams TR, Bruford MW, Chapman N, Wayne RJ, Harris S (1996) Molecular genetics of an invading species: Reeves' muntjac in southern England. Proceedings of the 3rd International Congress of Biology - Deer.

Williams JGK, Kubelik AR, Livak KJ, Rafalski JA, Tingey SV (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Research*, **18**, 6531-6535.

Wilson EO (1992) *The diversity of Life*. Penguin, London.

Wright S (1921) Systems of mating, I-V. *Genetics*, **6**, 111-178.

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

Wright S (1940) Breeding structure of populations in relation to speciation. *American Naturalist*, **74**, 232-248.

Wright S (1943) Isolation by distance. *Genetics*, **28**, 114-138.

Wright S (1951) The genetical structure of populations. *Annals of Eugenics*, **15**, 323-354.

Wright S (1965) The interpretation of population structure by *F*-statistics with special regard to systems of mating. *Evolution*, **19**, 395-420.

Zouros E, Mallet AL (1989) Genetic explanations of the growth/heterozygosity correlation in marine mollusks. In: *Reproduction, genetics and distribution of marine organisms*. Eds Ryland JS, Tyler PA., 23rd European Marine Biology Symposium, Olsen & Olsen, Fredensborg, Denmark.

# Appendix 1

*Nematostella vectensis* RAPD profile distribution throughout the populations. Populations shown with numbers of individuals RAPD profiled in parenthesis. Percentage total refers to the percent value of that profile throughout the total data set.

| Profile | N1 (40) | N2 (40) | N4 (39) | N5 (39) | N6 (40) | N7 (37) | N8 (39) | N10 (27) | N11 (36) | % Total |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 35 | 38 | 27 | 15 | 40 | 3 | 5 | 9 | 32 | 61.0 |
| 2 | 1 | | | | | | | | | 0.3 |
| 3 | 1 | | | | | | | | | 0.3 |
| 4 | 1 | | | | | | | | | 0.3 |
| 5 | 1 | | | | | | | | | 0.3 |
| 6 | 1 | | | | | | | | | 0.3 |
| 7 | | 2 | | | | | | | | 0.6 |
| 8 | | | 1 | | | | | | | 0.3 |
| 9 | | | 3 | | | 2 | | | | 1.5 |
| 10 | | | 1 | | | | | | | 0.3 |
| 11 | | | 1 | | | | | | | 0.3 |
| 12 | | | 2 | | | | 1 | 2 | | 1.5 |
| 13 | | | 1 | | | | | | | 0.3 |
| 14 | | | 1 | | | | | | | 0.3 |
| 15 | | | 1 | | | | | | | 0.3 |
| 16 | | | 1 | | | | | | | 0.3 |
| 17 | | | | 1 | | | 1 | | | 0.6 |
| 18 | | | | 1 | | | 2 | 3 | 1 | 2.1 |
| 19 | | | | 1 | | | | | | 0.3 |
| 20 | | | | 1 | | | | | | 0.3 |
| 21 | | | | 1 | | | | | | 0.3 |
| 22 | | | | 1 | | | | | | 0.3 |
| 23 | | | | 1 | | | | | | 0.3 |
| 24 | | | | 2 | | | | | | 0.6 |
| 25 | | | | 2 | | | | | | 0.6 |
| 26 | | | | 1 | | 1 | | | | 0.6 |
| 27 | | | | 3 | | | | | | 0.9 |
| 28 | | | | 1 | | | | | | 0.3 |
| 29 | | | | 1 | | | | | | 0.3 |
| 30 | | | | 2 | | | | | | 0.6 |
| 31 | | | | 1 | | | | | | 0.3 |
| 32 | | | | | | 1 | | | | 0.3 |
| 33 | | | | | | 1 | | | | 0.3 |
| 34 | | | | | | 1 | | | | 0.3 |
| 35 | | | | | | 3 | | 1 | | 1.2 |
| 36 | | | | | | 2 | 6 | | | 2.4 |
| 37 | | | | | | 1 | 2 | | | 0.9 |
| 38 | | | | | | 17 | | | | 5.0 |
| 39 | | | | | | 1 | | | | 0.3 |
| 40 | | | | | | 1 | | | | 0.3 |
| 41 | | | | | | 1 | | | | 0.3 |
| 42 | | | | | | 1 | | | | 0.3 |

| Profile | N1 (40) | N2 (40) | N4 (39) | N5 (39) | N6 (40) | N7 (37) | N8 (39) | N10 (27) | N11 (36) | % Total |
|---|---|---|---|---|---|---|---|---|---|---|
| 43 | | | | | | 1 | | | | 0.3 |
| 44 | | | | | | | 3 | | | 0.9 |
| 45 | | | | | | | 1 | | | 0.3 |
| 46 | | | | | | | 1 | | | 0.3 |
| 47 | | | | | | | 1 | | | 0.3 |
| 48 | | | | | | | 2 | | | 0.6 |
| 49 | | | | | | | 1 | | | 0.3 |
| 50 | | | | | | | 1 | | | 0.3 |
| 51 | | | | | | | 1 | | | 0.3 |
| 52 | | | | | | | 1 | | | 0.3 |
| 53 | | | | | | | 1 | | | 0.3 |
| 54 | | | | | | | 2 | | | 0.6 |
| 55 | | | | | | | 1 | | | 0.3 |
| 56 | | | | | | | 1 | | | 0.3 |
| 57 | | | | | | | 1 | | | 0.3 |
| 58 | | | | | | | 1 | | | 0.3 |
| 58 | | | | | | | | 1 | | 0.3 |
| 60 | | | | | | | | 2 | | 0.6 |
| 61 | | | | | | | | 1 | | 0.3 |
| 62 | | | | | | | | 1 | | 0.3 |
| 63 | | | | | | | | 1 | | 0.3 |
| 64 | | | | | | | | 1 | | 0.3 |
| 65 | | | | | | | | 1 | | 0.3 |
| 66 | | | | | | | | 1 | | 0.3 |
| 67 | | | | | | | | 1 | | 0.3 |
| 68 | | | | | | | | 1 | | 0.3 |
| 69 | | | | | | | | 1 | | 0.3 |
| 70 | | | | | | | | | 1 | 0.3 |
| 71 | | | | | | | | | 1 | 0.3 |
| 72 | | | | | | | | | 1 | 0.3 |
| 73 | | | | 1 | | | | | | 0.3 |
| 74 | | | | 1 | | | | | | 0.3 |
| 75 | | | | 1 | | | | | | 0.3 |
| 76 | | | | | | | 1 | | | 0.3 |
| 77 | | | | | | | 1 | | | 0.3 |
| 78 | | | | | | | 1 | | | 0.3 |
| 79 | | | | 1 | | | | | | 0.3 |